

www.engclubs.net

A site for all Engineers



آمار و احتمال

(رشته آمار)

دکتر نرگس عباسی دکتر علی شادرخ دکتر محمدقاسم وحدی اصل

گروه آمار

(آ/۴۴)

دانشگاه پیام نور

۱۵۶۱

سربنشانه: عباسی، نرگس، ۱۳۴۲ -

عنوان و نام پدیدآور: آمار و احتمال / تالیف نرگس عباسی، علی شادرخ، محمدقاسم

وحیدی اصل؛ ویراستار عین‌اله پاشا.

مشخصات نشر: تهران: دانشگاه پیام نور، ۱۳۸۸.

مشخصات ظاهری: ۳۵۷ ص. جدول، نمودار.

فروخت: دانشگاه پیام نور؛ ۱۵۶۱ گروه آمار؛ آ/۴۴

شابک: ۹۷۸-۹۶۴-۳۸۷-۶۰۶-۷

وضعیت فهرست نویسی: فیبا

یادداشت: واژه نامه

موضوع: آمار -- آموزش برنامه‌ای

موضوع: آمار

موضوع: آمار -- آزمون‌ها و تمرین‌ها (عالی)

شناسه افزوده: شادرخ، علی، ۱۳۴۶ -

شناسه افزوده: وحیدی اصل، محمدقاسم، ۱۳۲۶ -

شناسه افزوده: پاشا، عین‌الله، ۱۳۲۸ - ویراستار

شناسه افزوده: دانشگاه پیام نور

رده بندی کنگره: QA۲۷۶/۲۲۸ ۱۳۸۸:

رده بندی دیوبی: ۵۱۹/۵:

شماره کتابشناسی ملی: ۱۸۰۰۷۸۱:



آمار و احتمال

دکتر نرگس عباسی دکتر علی شادرخ دکتر محمدقاسم وحیدی اصل

ویراستار علمی: دکتر عین‌اله پاشا

حروفچینی و نمونه‌خوانی: مدیریت تولید مواد و تجهیزات آموزشی

طراح جلد: سپیده‌سادات مدنی

لیتوگرافی، چاپ و صحافی: انتشارات دانشگاه پیام نور

تعداد: * * *

چاپ..... ۱۳۸۸

قیمت: * * *

کلیه حقوق برای دانشگاه پیام نور محفوظ است.

بسم الله الرحمن الرحيم

پیشگفتار ناشر

کتاب‌های دانشگاه پیام نور حسب مورد و با توجه به شرایط مختلف یک درس در یک یا چند رشته دانشگاهی، به صورت کتاب درسی، متن آزمایشگاهی، فرادرسی، و کمکدرسی چاپ می‌شوند.

کتاب درسی ثمرة کوشش‌های علمی صاحب اثر است که براساس نیازهای درسی دانشجویان و سرفصل‌های مصوب تهیه و پس از داوری علمی، طراحی آموزشی، و ویرایش علمی در گروههای علمی و آموزشی، به چاپ می‌رسد. پس از چاپ ویرایش اول اثر، با نظرخواهی‌ها و داوری علمی مجدد و با دریافت نظرهای اصلاحی و مناسب با پیشرفت علوم و فناوری، صاحب اثر در کتاب تجدیدنظر می‌کند و ویرایش جدید کتاب با اعمال ویرایش زبانی و صوری جدید چاپ می‌شود. متن آزمایشگاهی (م) راهنمایی است که دانشجویان با استفاده از آن و کمک استاد، کارهای عملی و آزمایشگاهی را انجام می‌دهند.

کتاب‌های فرادرسی (ف) و کمکدرسی (ک) به منظور غنی‌تر کردن منابع درسی دانشگاهی تهیه و بر روی لوح فشرده تکثیر می‌شوند و یا در وبگاه دانشگاه قرار می‌گیرند.

مدیریت تولید مواد و تجهیزات آموزشی

فهرست

| نام | پیشگفتار |
|-----|---|
| ۱ | فصل اول آمار توصیفی |
| ۱ | ۱- مقدمه (مروری تاریخی) |
| ۲ | ۱-۱- اعداد |
| ۴ | ۲- آمار توصیفی و آمار استنباطی |
| ۴ | ۳- برخی مفاهیم اولیه |
| ۶ | ۴- مفهوم مقیاس‌سازی |
| ۹ | ۵- منابع داده‌ها |
| ۱۰ | ۲-۱- سازمان‌دهی داده‌ها |
| ۱۱ | ۱-۲-۱- مرتب‌کردن داده‌های عددی |
| ۱۴ | تمرین |
| ۱۴ | ۲-۲-۱- توزیع‌های فراوانی |
| ۲۱ | ۳-۲-۱- توصیف داده‌ها به کمک نمودارها |
| ۲۶ | تمرین |
| ۲۸ | ۳-۱- توصیف عددی داده‌ها |
| ۲۸ | ۱-۳-۱- شاخص‌های گرایش مرکزی |
| ۲۹ | ۱-۱-۳-۱- میانگین |
| ۳۳ | ۲-۱-۳-۱- میانه |
| ۳۷ | ۳-۱-۳-۱- مدل |
| ۳۹ | ۴-۱-۳-۱- دیگر شاخص‌های گرایش مرکزی |
| ۴۲ | ۲-۳-۱- شاخص‌های پراکندگی |
| ۴۲ | ۱-۲-۳-۱- دامنه |
| ۴۳ | ۲-۲-۳-۱- انحراف معیار |
| ۵۳ | ۳-۲-۳-۱- چندک‌ها و دیگر شاخص‌های پراکندگی |
| ۵۷ | ۴-۲-۳-۱- معیارهای ویژه نمودارهای داده‌ها |
| ۶۱ | تمرین |
| ۶۳ | ۴-۱- جدول‌های دو متغیره و معیارهای توصیفی |

| | |
|-----|---------------------------------|
| ۶۴ | ۱-۴-۱ توزیع فراوانی دو متغیره |
| ۶۸ | ۱-۴-۲ معیارهای توصیفی دو متغیره |
| ۷۰ | ۱-۴-۳ ضریب همبستگی خطی |
| ۷۱ | تمرین |
| ۷۳ | فصل دوم قوانین شمارش |
| ۷۳ | ۱-۲ مقدمه |
| ۷۳ | ۲-۲ اصول شمارش |
| ۷۴ | تمرین |
| ۷۶ | ۳-۲ جایگشت |
| ۷۷ | تمرین |
| ۷۸ | ۴-۲ ترتیب |
| ۷۹ | تمرین |
| ۸۰ | ۵-۲ جایگشت دوری |
| ۸۰ | تمرین |
| ۸۱ | ۶-۲ ترکیب |
| ۸۲ | تمرین |
| ۸۲ | ۷-۲ تعمیم ترکیب |
| ۸۳ | تمرین |
| ۸۳ | ۸-۲ ضرایب دو جمله‌ای |
| ۸۷ | تمرین |
| ۸۹ | فصل سوم احتمال |
| ۸۹ | ۱-۳ مقدمه |
| ۸۹ | ۲-۳ آزمایش تصادفی |
| ۹۰ | ۳-۳ فضای نمونه‌ای و برآمد |
| ۹۱ | تمرین |
| ۹۲ | ۴-۳ پیشامد |
| ۹۵ | تمرین |
| ۹۶ | ۵-۳ احتمال |
| ۱۰۳ | ۶-۳ برخی قواعد احتمال |
| ۱۰۸ | تمرین |
| ۱۱۱ | ۷-۳ احتمال شرطی |
| ۱۱۷ | تمرین |
| ۱۱۹ | ۸-۳ پیشامدهای مستقل |
| ۱۲۳ | تمرین |
| ۱۲۴ | ۹-۳ قضیه‌ی بیز |

تمرین

۱۲۹

فصل چهارم متغیرهای تصادفی، توزیع‌ها، و چگالی‌های احتمال

- ۱۳۱ ۱-۴ متغیر تصادفی
۱۳۱ ۲-۴ توزیع‌های احتمال
۱۳۵ تمرین
۱۴۲ ۳-۴ متغیرهای تصادفی پیوسته
۱۴۶ ۴-۴تابع‌های چگالی احتمال
۱۴۸ تمرین
۱۵۳ ۵-۴ توزیع‌های چندمتغیره
۱۵۶ تمرین
۱۶۱ ۶-۴ توزیع‌های حاشیه‌ای
۱۶۲ تمرین
۱۶۵ ۷-۴ متغیرهای تصادفی مستقل
۱۶۵ تمرین
۱۶۷ ۸-۴ توزیع‌های شرطی
۱۶۸ تمرین
۱۶۹ تمرین

فصل پنجم امید ریاضی

- ۱۷۱ ۱-۵ مقدمه
۱۷۱ ۲-۵ مقدار امید ریاضی یک متغیر تصادفی
۱۷۲ تمرین
۱۷۹ ۳-۵ گشتاورها
۱۸۱ تمرین
۱۸۵ ۴-۵ قضیه‌ی چیشیف
۱۸۶ تمرین
۱۸۹ ۵-۵ توابع مولد گشتاورها
۱۹۰ ۶-۵ گشتاورهای ترکیب‌های خطی متغیرهای تصادفی
۱۹۷ تمرین
۲۰۰ تمرین

فصل ششم توزیع‌ها و چگالی‌های احتمال خاص

- ۲۰۳ ۱-۶ مقدمه
۲۰۳ ۲-۶ توزیع یکنواخت گسسته
۲۰۴ ۳-۶ توزیع برنولی
۲۰۵ ۴-۶ توزیع دوجمله‌ای
۲۱۲ تمرین
۲۱۵ ۵-۶ توزیع فوق‌هندسی

| | |
|-----|---------------------------------------|
| ۲۱۸ | تمرین |
| ۲۲۰ | ۶-۶ توزیع پواسون |
| ۲۲۶ | تمرین |
| ۲۲۸ | ۷-۶ چگالی یکنواخت |
| ۲۲۸ | تمرین |
| ۲۲۹ | ۸-۶ توزیع نرمال |
| ۲۳۵ | تمرین |
| ۲۳۷ | ۹-۶ تقریب نرمال برای توزیع دو جمله‌ای |
| ۲۴۱ | تمرین |
| ۲۴۲ | ۱۰-۶ توزیع نرمال دو متغیره |
| ۲۴۴ | تمرین |
| ۲۴۵ | فصل هفتم محاسبات با نرم افزار |
| ۲۴۵ | ۱-۷ مقدمه |
| ۲۴۵ | ۲-۷ ورود و خروج اطلاعات در مینیتوب |
| ۲۵۶ | تمرین |
| ۲۵۸ | ۳-۷ محاسبات در مینیتوب |
| ۲۶۶ | تمرین |
| ۲۶۷ | ۴-۷ رسم نمودار |
| ۲۶۹ | تمرین |
| ۲۷۱ | حل تمرین‌ها |
| ۲۷۱ | حل تمرین‌های فصل اول |
| ۲۷۶ | حل تمرین‌های فصل دوم |
| ۲۸۱ | حل تمرین‌های فصل سوم |
| ۲۹۰ | حل تمرین‌های فصل چهارم |
| ۳۰۲ | حل تمرین‌های فصل پنجم |
| ۳۱۲ | حل تمرین‌های فصل ششم |
| ۳۲۹ | سوالات تستی |
| ۳۴۷ | جداول |
| ۳۵۸ | مراجع |

بسم الله الرحمن الرحيم

پیشگفتار

هدف عمده از درس "آمار و احتمال" یادگیری مفاهیم آماری و پرداختن به روش‌هایی است که به براساس آنها تحلیل داده‌ها صورت می‌پذیرد. برای وارد شدن به مباحث مربوط به روش‌های آماری، ابتدا می‌بایست مقدمات را فرا گرفت.

این کتاب براساس سرفصل مدون وزارت علوم تحقیقات و فناوری برای درس آمار و احتمال یک دوره‌ی کارشناسی رشته‌ی آمار تدوین شده است. برای مطالعه‌ی این کتاب با معلوماتی از دروس ریاضی عمومی I و II کفايت می‌کند.

با مرور فهرست مطالب کتاب، از عنوان‌های کلی و جزئی درس مطلع می‌شوید.

درک مطالب هر فصل وقتی میسر می‌شود که مطالب فصل‌های قبل از آن را به خوبی فهمیده باشید. پس از دریافت چهارچوب هر فصل، به تدریج به مطالعه‌ی عمیق مطالب اهتمام ورزید و تمرین‌های مربوط را حل کنید. با پاسخ‌گویی به پرسش‌ها و حل تمرین‌ها، می‌توانید از کیفیت پیشرفت خود و درک و فهم مطالب، آگاه شوید. ضمناً به تمرین‌های هر فصل در پایان کتاب، پاسخ داده شده است. همچنین در بخشی دیگر برخی پرسش‌های چهارگزینه‌ای مرتبط با مطالب کتاب جمع‌آوری شده که حل آنها را به خواننده وامی‌گذاریم. برای بخش‌هایی که از نرمافزار مینیتب استفاده شده است توصیه می‌شود، آموزش فصل هفتم را ابتدا در کلاس‌های عملی این درس فرا گرفته، سپس وارد فصل‌های اول تا ششم شوند. لازم به ذکر است که آزمون‌های پایان هر نیم‌سال ارزیابی از این فصل نخواهد شد و ارزیابی آن را به عهده‌ی استادی که مجری کلاس‌های عملی هستند واگذار می‌نماییم.

در اینجا لازم می‌دانیم که مراتب قدردانی خود را از استاد گرامی جناب آقای دکتر عین‌الله پاشا استاد محترم دانشگاه تربیت معلم که زحمت ویراستاری کتاب را بر عهده داشتند اعلام نماییم. همچنین از همکاران محترم مدیریت تدوین منابع درسی

خصوصاً جناب آقای اکبری کمال تشکر و سپاس را داریم. جا دارد از آقای محمد عباسی که زحمت تایپ این کتاب را کشیده‌اند قدردانی به عمل آوریم.
هرگونه پیشنهاد و انتقاد در جهت ارتقای کیفیت مطالب علمی کتاب از طرف خوانندگان محترم، مورد استقبال گرم نویسنده‌گان قرار گرفته و موجب سپاس می‌گردد.

نویسنده‌گان

فصل اوّل

آمار توصیفی

۱-۱ مقدمه (مروجی تاریخی)

کلمه‌ی statistics که در زبان فارسی به آمار ترجمه می‌شود به دو معنی به کار می‌رود: یک معنای رایج آن معرف مجموعه‌ای از داده‌هایی است که دامنه‌ی وسیعی از اشیاء، نظیر اندازه‌ی جامعه‌ها، اطلاعات فعالیت‌های تولیدی، قیمت اجناس، درآمدها، میزان بارندگی و غیره را شامل می‌شود. اطلاعات آماری از این نوع به طور متعارف توسط دولت، بخش‌های خصوصی و قسمت‌هایی از زندگی روزمره تولید می‌شود. دیگر آنکه این کلمه به نظریه و روش‌های جمع‌آوری، توصیف، تحلیل داده‌های عددی مربوط می‌شود.

در زمان‌های قدیم دولتها اطلاعات مربوط به جمعیت، ثروت، تعداد سربازها و ... را برای کنترل اوضاع کشور و جمع‌آوری مالیات، گردآوری می‌کردند. در تاریخ، چینی‌ها، چند هزار سال قبل از تولد مسیح، تعداد نفوس را برآورد کرده‌اند و مصریان نیز علاوه بر سرشماری نفوس جداول تحرک جمعیت را نیز تشکیل داده بودند. در ایران و هند نیز جمع‌آوری آمارهای اداری بیش از دو هزار سال سابقه دارد و چینی‌آورده‌اند که در زمان داریوش پادشاه ایران برای شمارش سربازان بدین‌گونه عمل می‌شد که آنان از جایگاه خاصی بگذرند و پاره‌سنگی در آن بگذارند.

در قرن هفدهم آمار حیاتی توسط جان گرانت انگلیسی پایه‌گذاری شد. او را پدر علم آمار حیاتی می‌دانند زیرا او بود که برای اوئین بار آمارهای تولد و مرگ را بررسی کرد و جداول مرگ‌ومیر را تنظیم و امید به زندگی در سنین مختلف را محاسبه کرد و سرانجام در آخر قرن هفدهم اوئین موسسه‌ی بیمه‌ی عمر در لندن پا به عرصه وجود نهاد.

نظریه‌ی آمار از نیمه‌ی قرن هفدهم با واردشدن نظریه‌ی احتمال، و مسائل بازی‌های شناسی توسعه می‌یابد. امروزه "رونالد فیشر" انگلیسی را به عنوان پدر علم آمار می‌شناسند. وی آمار را در زمینه‌های مختلف مانند ژنتیک، بیومتری، آموزش و پرورش، کشاورزی و مانند آن‌ها وارد کرد و در معرفی استنباط آماری شامل برآوردهای نقطه‌ای، توزیع‌های دقیق نمونه، و طرح آزمایش‌ها پیش‌گام بود.

به کارگیری روش‌های آماری در علوم اجتماعی و رفتاری کمی قبل از جنگ جهانی دوم شروع شد. تعداد آمارگیری‌ها در زمینه‌های مختلف افزایش یافت و ضرورت تفسیر اطلاعات مربوط به روان‌شناسی و تعلیم و تربیت آشنازی با علم آمار را اجتناب‌ناپذیر کرد. امروزه موفقیت در بسیاری از زمینه‌های علمی مانند علوم انسانی، پزشکی و فنی و مهندسی بدون داشتن اطلاعات لازم از علم آمار مشکل و گاهی غیرممکن است.

۱-۱-۱ اعداد

به کاربردن مفهوم عدد تحولی در بیان اندیشه‌های انسانی ایجاد کرده است. همچنان که گذشتگان دور نمی‌توانستند، تعداد دقیق فرزندان خود را به اندیشه و گفتار درآورند، وضع بشر، مثلاً تا پیش از اختراع دماسنچ، برای بیان وضع هوا مانند گذشتگان دور خود بود. اگر از وی وضع هوا را از لحظه گرمی و سردی می‌پرسیدند، مثلاً جواب می‌داد: هوا سرد است. اگر می‌گفتند چقدر سرد است، می‌گفت: «خیلی» سرد است. اما با اختراق دماسنچ و کمی‌شدن یا به کمیت درآمدن میزان دما، بشر قادر است به آسانی مناطق «سردسیر» و «گرمسیر» و مانند آنها را به صورتی مشخص‌تر تعریف کند و حتی رده‌بندی را بسیار دقیق‌تر و گستردگر کنند. بیان اینکه فلان کشور «پرجمعیت» است،

نمی‌تواند در این حد از دقت باشد که بگوییم مثلاً جمعیت آن کشور یک «میلیون نفر» است، یا متوسط تعداد کسانی که در هر هکتار می‌زیند مثلاً ۳۰۰۰ نفر است. پس، اگر نتوانیم بگوییم که مفهوم عدد، زندگی انسان را دگرگون کرده‌است، لاقل می‌توانیم بگوییم که با متحول شدن زندگی انسان، عدد نقشی هرچه بهتر و بیشتر در زندگی او داشته است.

عدد در زندگی انسان نقش‌هایی گوناگون دارد. یونانیان حدود سه هزار سال پیش، عدد را تنها به خاطر خود عدد و برای درک روابط اعداد صحیح مورد مطالعه قراردادند و زیباترین قضیه‌های ریاضی و موضوعی بهنام «نظریه‌ی اعداد» را پدید آورند. اما بشر، در تمام عرصه‌های زندگی خود، چه شخصی و چه اجتماعی، به «حساب کردن» نیز نیاز داشت. این نیازهای زندگی، «چهار عمل اصلی حساب» و لذا «علم حساب» را پدید آورد که آن هم جزوی از ریاضیات است.

اما امروزه عدد نقشی بسیار فراتر از نقش خود در علم حساب دارد. در واقع بسیاری از پدیده‌های طبیعی و اجتماعی و مانند آنها را می‌توان به کمک اعداد بیان کرد و یا به تعبیر و تفسیر این پدیده‌ها به کمک اعداد پرداخت. مثلاً بیان شدت زلزله در واحد ریشت، میزان قوت و ضعف نسبی آن را مشخص می‌کند، اما تعداد تلفات انسانی، تعداد بناهای تخریب شده و ... می‌تواند به بهترین نحو، ابعاد فاجعه را نشان دهد. پس در این مورد، چگونگی زلزله و عوارض و عواقب آن را با استفاده از اعداد توصیف کرده‌ایم. امروزه تلاش می‌شود که برای بیان، تجزیه و تحلیل، و تعبیر و تفسیر هر نوع پدیده‌ای از نیروی جادویی اعداد استفاده شود. البته نباید تصوّر کرد که بیان هر نوع پدیده‌ای و تجزیه و تحلیل آن با همان سهولتی انجام می‌شود که در بالا در مورد زلزله انجام شد. برای آنکه تصوّری از برخی از این پدیده‌ها و پیچیدگی نسبی آنها داشته باشیم، به مثال‌های زیر توجه کنید:

الف) نقش افزایش یا کاهش درآمد کشاورزی در مهاجرت‌های روسیایی،
ب) تأثیر آلدگی هوا در سطح سلامت جوامع،
ج) تأثیر استفاده از وسایل فن‌آوری پیشرفته در تغییر ساختار روابط اجتماعی.
برای آنکه بتوانیم به پرسش‌هایی از نوع بالا پاسخ دهیم، ابتدا باید اطلاعاتی در زمینه‌ی موضوع مورد بحث گردآوری کنیم. چنین اطلاعاتی، برحسب مجموعه‌ای از اعداد،

داده‌های آماری نامیده می‌شوند. نحوه‌ی گردآوری، تنظیم و تلخیص، تجزیه و تحلیل، و تعبیر و تفسیر داده‌ها، و استخراج نتایجی درباره‌ی موضوع مورد بحث، موضوع علم آمار را تشکیل می‌دهد.

۱-۱-۲ آمار توصیفی و آمار استنباطی

دو بخش از موضوع علم آمار، آمار توصیفی و آمار استنباطی، را در زیر شرح می‌دهیم.

آمار توصیفی شامل کلیه‌ی اعمالی است که هدف از آن‌ها خلاصه‌کردن داده‌ها یا توصیف جنبه‌های مهم داده‌هاست، بدون آنکه گامی فراتر از آن برداشته شود؛ یعنی بدون آنکه تلاش شود، نتیجه‌گیری‌ها یا استنباط‌هایی درباره‌ی منابعی که داده‌ها از آن استخراج شده‌اند، به عمل آید.

گرچه آمار توصیفی بخشی مهم از آمار است و هنوز هم استفاده‌ی گسترده‌ای از آن، مخصوصاً برای توصیف ویژگی‌های مهم داده‌ها برای عموم به عمل می‌آید، اما در سال‌های اخیر توجه بیشتر به آمار استنباطی، یعنی روش‌هایی است که به کمک آن‌ها می‌توان اطلاعات موجود در مجموعه‌ای متناهی از داده‌ها را به مجموعه‌ای بزرگتر که داده‌ها از آن به دست آمده‌اند، تعمیم داد.

۱-۱-۳ برخی مفاهیم اولیه

جامعه‌ی آماری عبارت است از عناصر مورد نظر که حداقل دارای یک صفت مشخصه‌ی مشترک باشند. مانند جامعه‌ی دانشجویان دانشگاه پیام‌نور، جامعه‌ی ساکنین شهر تهران، جامعه‌ی بیماران قلبی و....

صفت مشخصه‌ی صفتی است که بین همه‌ی عناصر آماری مشترک و متمایزکننده‌ی یک جامعه‌ی آماری از سایر جوامع است. مثلاً در جامعه‌ی ایرانیان افراد جامعه نسبت به ایرانی بودن هیچ وجه تمایزی ندارند.

جامعه‌ی آماری به دو نوع تقسیم می‌شود: متناهی و نامتناهی. اگر مقادیر جامعه از تعداد محدود ثابتی تشکیل شود و پایان‌پذیر باشد آن را متناهی می‌نامیم مثلاً جمعیت ایران در سرشماری سال ۱۳۸۵ که ۷۰۰۴۹۲۶۲ نفر بوده است، یک جامعه‌ی متناهی

است. در مقابل جامعه‌های آماری نامتناهی را داریم، مانند جامعه‌ای با اعضای ۱، ۲، ۳، ۴، ۵، ... یا جمعیت ایران در سال آینده که یک جامعه‌ی نامتناهی را تشکیل می‌دهد.

متغیر: بررسی‌های آماری روی صفات مشخصه انجام نمی‌شود زیرا افراد جامعه نسبت به این صفات هیچ وجه تمایزی با هم ندارند. ولی اگر به افراد جامعه از زاویه‌های دیگر بنگریم تفاوت‌های زیادی می‌بینیم. آن صفاتی را که عامل این تفاوت‌ها هستند صفات متغیر می‌نامیم. پس صفات متغیر صفاتی هستند که در یک جامعه از یک عضو به عضو دیگر تغییر می‌کنند. مثلاً در جامعه‌ی دانشجویان دانشگاه پیام‌نور، شهرستان محل سکونت و رشته‌ی انتخابی صفات متغیر هستند. به طور کلی صفاتی نظری استعداد، هوش، نوع بیماری، گروه خونی، وضع تأهل، سن، طول عمر لامپ، تعداد دفعات تعویض قطعه‌ای از خودرو، میزان تأثیر دارویی خاص، و ...، متغیر در نظر گرفته شوند.

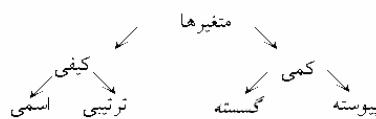
أنواع متغیر: متغیرها به دو دسته تقسیم می‌شوند: متغیرهای کیفی و متغیرهای کمّی.

(الف) متغیرهای کیفی: مربوط به صفاتی هستند که واحد نداشته و قابل اندازه‌گیری نیستند مانند جنس، مرغوبیت، شغل، نوع بیماری، وضع تأهل، رنگ چشم، گروه خونی و نظائر آنها (در بخش بعدی متغیرها متغیرها عمدتاً به گروه متغیرهایی که با مقیاس‌های اسمی و ترتیبی اندازه‌گیری می‌شوند، تعلق دارند).

(ب) متغیرهای کمّی: مربوط به صفاتی هستند که قابل اندازه‌گیری یا شمارش هستند؛ یعنی واحد اندازه‌گیری یا شمارش دارند و قابل مقایسه و سنجش می‌باشند. مانند سن، طول عمر، درجات تحصیلی، وزن، قد، درآمد و غیره که از راه اندازه‌گیری به دست می‌آیند.

متغیرهای کمّی خود به دو دسته تقسیم می‌شوند که عبارتند از متغیرهای کمّی گستته و متغیرهای کمّی پیوسته. **متغیرهای کمّی گستته** آنها بی هستند که قابل شمارش هستند سن، تعداد افراد خانوار، ظرفیت مسافر، تعداد تولید روزانه، جمعیت روستاها و ... **متغیرهای کمّی پیوسته** آنها بی هستند که مقادیر آنها در یک بازه

قرار دارد یا به عبارت دیگر مقادیر آنها بازه‌ای از اعداد حقیقی انتخاب می‌شوند مانند وزن، قد، طول عمر، درآمد، نمره‌ی دروس، هزینه‌ی بیمارستانی برای بیماران و نظایر آن.



۱-۴ مفهوم مقیاس‌سازی

منظور از اندازه‌گیری فرآیندی است که به کمک آن به عضوی از جامعه عدد یا نمره‌ای نسبت داده می‌شود. باید توجه کرد که نمره‌هایی (اعدادی) که به این طریق به دست می‌آیند، معانی متفاوتی دارند. برای توضیح متفاوت بودن این نمره‌ها، سه پاسخ ظاهرآیکسان را که به سه پرسش متفاوت داده شده است، در نظر بگیرید. فرض کنید از کسی که در مسابقات دوی صدمتر کشور شرکت کرده است، به سه سوال زیر تنها با استفاده از عدد پاسخ گوید:

۱. با چه شماره‌ای در مسابقه شرکت کردید؟
۲. نفر چندم شدید؟
۳. در چند ثانیه به خط پایان رسیدید؟

فرض کنید پاسخ وی به هر سه سؤال «۵» باشد. با اینکه در هر سه پاسخ عدد «۵» به کار رفته است، اما این عدد در سه معنی متفاوت به کار برده شده است. او لین ۵ به نشانه‌ی شماره‌ی (شناسایی) شرکت کننده است، دومنی به نشانه‌ی مقام کسب شده در مسابقه و سومی به نشانه‌ی «زمان» پایان مسابقه است. بنابر این هر یک از ۵ ها مجموعه‌ی اطلاعات متفاوتی را شامل می‌شود.

استیونس^۱ (۱۹۴۶ میلادی)، چهار نوع مختلف از مقیاس‌های اندازه‌گیری را مشخص کرده است که هریک از آنها مجموعه‌ی اطلاعات متفاوتی را در بردارند. در هر مقیاس از عدد استفاده می‌شود ولی اطلاعات حاصل از این اعداد به کلی متفاوت است. این چهار مقیاس عبارتند از: اسمی، ترتیبی، فاصله‌ای، و نسبتی. در زیر به شرح هر یک از آنها می‌پردازیم.

۱. مقیاس اسمی

هر مقیاس را که معمولاً یک عدد طبیعی است و تنها برای شناسایی افراد یا اشیا و یا مکان‌ها به کار می‌رود، یک مقیاس اسمی می‌نامند. برای مثال، گروه‌های خونی افراد عبارتند از: A ، B ، AB ، و O ، که آن‌ها را به ترتیب با اعداد $4, 3, 2, 1$ مشخص می‌کنیم. این اعداد صرفاً می‌گویند که هر فردی دارای چه گروه خونی است، فردی که برقسپ ۳ دارد، دارای گروه خونی AB است.

داده‌های اسمی، تنها از نظر اسمی عددی‌اند، زیرا هیچ یک از خصیصت‌های اعدادی را که در حساب معمولی با آنها سروکار داریم، ندارند و اصطلاحاً در مقیاس اسمی، کدگذاری یا اسمگذاری انجام می‌شود. بنابر این توجه به اینکه آیا به کار بردن اعمال حسابی در مورد داده‌ها مجاز است یا خیر، اهمیت زیادی دارد.

۲. مقیاس ترتیبی (رتبه‌ای)

هرگاه مقیاس، یک عدد حقیقی و برتری را بیان کند آن را مقیاس ترتیبی می‌گویند. مواردی پیش می‌آید که صرف‌نظر از تفاوت محتویات یک طبقه یا محتویات طبقه دیگر، یک نوع ارتباط بین آنها برقرار است. روابط موجود بین طبقات با توجه به نوع مقیاس، اکثر مواقع صورت ترجیح دارد "به بالاتر، بیشتر، مشکل‌تر و" بیان می‌شود که چنین روابطی را معمولاً با علائمی نظیر " $>$ " نشان می‌دهند. استفاده از اصطلاحاتی

چون "بالا، وسط، پایین" و "قوی، متوسط، ضعیف" در تحقیقات بیانگر مقیاس ترتیبی است. در این نوع مقیاس مجاز به انجام عملیات جبری نیستیم.

برای مثال، وقتی اعداد $2, 4, 6, 8, 10$ را به ترتیب به پنج فرد A, B, C, D, E اختصاص می‌دهیم به معنی آن است که پنج فرد مذکور از لحاظ صفت مورد نظر به ترتیب A, B, C, D, E قرار می‌گیرند. لذا اگر وضع نسبی افراد یا اشیا را در مقیاس رتبه‌ای بدانیم آن‌ها را به صورت‌های زیر نمایش دهیم: $B > A$; یعنی، B از A بزرگتر است. $C < B$; یعنی، B از C کوچکتر است. $B = B$; یعنی، B با C هم ارزش است.

در مثالی دیگر، در یک مسابقه‌ی دومیدانی دونده‌هایی را که به ترتیب اول، دوّم و سوّم به خط پایان می‌رسند، با اعداد $1, 2$ و 3 مشخص کنیم.

۳. مقیاس فاصله‌ای

وقتی که مقیاسی همه خصوصیات یک مقیاس ترتیبی را دارا باشد و علاوه بر آن فاصله‌ی بین هر دو عدد بر روی مقیاس میزان مشخصی داشته باشد، همچنین نسبت تفاضل با دو فاصله را حفظ کند آن را مقیاس فاصله‌ای می‌نامند. در این نوع اندازه‌گیری نسبت هر دو فاصله مستقل از واحد اندازه‌گیری و مستقل از نقطه‌نظر صفر است. در مقیاس فاصله‌ای مکان نقطه‌ی صفر و واحد اندازه‌گیری اختیاری و قراردادی است.

در اندازه‌گیری دما، با وجود آن‌که دما در دو مقیاس سانتی‌گراد و فارنهایت دارای نقطه‌ی صفر متفاوت است، ولی هر دو برای اندازه‌گیری حرارت به کار می‌روند. برای مثال، پنج کمیت قابل اندازه‌گیری شده x_1, x_2, x_3, x_4, x_5 را در نظر می‌گیریم این کمیت‌ها را به اندازه‌های $11, 14, 17, 20, 23$ مقیاس‌سازی نموده‌ایم (یعنی x_i ‌ها در عددی ثابت ضرب و با عددی ثابت جمع شده‌اند و داده‌های جدید به دست آمده است). ملاحظه می‌شود که اختلاف اندازه‌ی کمیت دوّم از اول ($14 - 11 = 3$) برابر سه برابر اختلاف کمیت اول از چهارم ($20 - 11 = 9$) است.

۴. مقیاس نسبتی

به مقیاسی که علاوه بر همهٔ خصوصیات مقیاس فاصله‌ای دارای نقطهٔ صفر واقعی نیز هست و نسبت را حفظ می‌کند، مقیاس نسبتی می‌گویند. در مقیاس نسبتی، نسبت هر دو قطعهٔ روی مقیاس اندازه‌گیری مستقل از واحد اندازه‌گیری است.

مقیاس‌هایی چون پوند و گرم نقطهٔ صفر واقعی دارند و نسبت هر دو وزن دلخواه از واحد اندازه‌گیری مستقل است. بیشتر اطلاعاتی که با استفاده از وسایل فیزیکی اندازه‌گیری می‌شود از نوع مقیاس نسبتی است. مثلاً، اگر در سطح مقیاس نسبتی، اعداد ۲، ۴، ۸، ۱۶ نمایندهٔ مقدار صفت به خصوصی در اشیای A ، B ، C ، D ، E باشند می‌توان گفت که این صفت در B به اندازهٔ ۲ برابر A و در به اندازهٔ ۳ برابر A و در E ، $\frac{2}{5}$ برابر B و نظایر آن وجود دارد. جدول زیر نشان‌دهندهٔ خلاصهٔ مفاهیم ذکر شده دربارهٔ مقیاس‌های چهارگانه و مقایسهٔ آنها است.

| نسبتی | فاصله‌ای | رتبه‌ای | اسمعی | نوع مقیاس | |
|-------|----------|---------|-------|-------------------|---------------|
| | | | | مرانب | مرانب |
| دارد | دارد | دارد | دارد | ترتیب | |
| دارد | دارد | دارد | دارد | | فواصل |
| دارد | دارد | دارد | دارد | مبدأ صفر قراردادی | |
| دارد | دارد | دارد | دارد | | مبدأ صفر مطلق |

۱-۱-۵ منابع داده‌ها

با توجه به گستردگی مباحث علمی، محققان منابع گوناگونی برای گردآوری داده‌ها در اختیار دارند. بدون واردشدن در جزئیات به انواع مهم داده‌ها بر حسب منابع آنها اشاره می‌کنیم.

داده‌های اوّلیه که اطلاعات خام را که مخصوصاً برای انجام تحلیل گردآوری شده‌اند، در اختیار ما می‌گذارند. این داده‌ها یا از طریق سنجش مستقیم، مانند انجام مصاحبه، مشاهدات میدانی، اندازه‌گیری با ابزارهای ویژه، به دست می‌آیند، یا از طریق

سنجدش غیرمستقیم که در آن سنجدش و ثبت نتایج در مکان‌های مجزا از هم انجام می‌شود و این روش استفاده از مصاحبه‌های تلفنی، شمارنده‌های ترافیک، و ثبات‌های بارندگی خودکار را شامل می‌شود. داده‌های حاصل از روش‌های سنجدش از راه دور نیز در زمرة‌ی داده‌های اویله محسوب می‌شوند.

داده‌های ثانویه که در آن اطلاعات که پیشتر - به هر منظور خاص - گردآوری شده‌اند، منبع اصلی تلقی شده و اطلاعات مربوط به تحلیل مورد نظر از آن استخراج می‌شود. این داده‌ها یا پیامدهای غیرمستقیم فعالیت‌های دستگاه‌های اجرایی دولتی یا خصوصی، نظیر ادارات ثبت احوال و ثبت اسناد، مؤسسات نقشه‌برداری، بانک‌ها، شهرداری‌ها، و غیره‌اند یا داده‌های حاصل از سرشماری‌های رسمی سراسری و یا مقطوعی مؤسساتی نظیر مرکز آمار ایران‌اند. داده‌های مندرج در تصاویر حاصل از دستگاه‌های سنجدش از راه دور نیز از این جمله‌اند.

بديهی است که تفاوت‌های عمدی‌ای بين فرایندهای اندازه‌گیری مربوط به منابع گوناگون، و بين داده‌های اویله و ثانویه موجود است. وقتی از داده‌های اویله استفاده می‌کیم، معمولاً طرح پژوهشی مربوط به تحلیل آسان‌تر است اما فرایند اندازه‌گیری معمولاً دشوار و وقت‌گیر است. با توجه به امکان تحت کنترل داشتن فرایند گردآوری داده‌ها در این حالت، می‌توان متغیرها را به هر شکلی که در تحلیل آماری مورد نیاز باشند، انتخاب و اندازه‌گیری کرد. اما، وقتی از داده‌های ثانویه استفاده می‌کنیم، کنترلی نسبت به متغیرها و فرایند اندازه‌گیری نداریم و این امر ممکن است موجب ایجاد مشکلاتی در موقع تجزیه و تحلیل شود. در این صورت بر عهده‌ی محقق است که ابتدا از مناسب‌بودن داده‌ها اطمینان حاصل کند.

۱-۲ سازماندهی داده‌ها

برای آنکه بتوان نخستین گام‌ها را برای تجزیه و تحلیل داده‌ها برداشت و نتایجی درباره‌ی خصوصیات جامعه‌ای که این اطلاعات از آن استخراج شده‌اند، به دست آورد، لازم است که این داده‌ها سازماندهی (تنظيم) و تلخیص شوند.

۱-۲-۱ مرتب کردن داده‌های عددی

فرض کنید داده‌های زیر میزان بارندگی سالانه ۱۰۰ شهر را بر حسب سانتی‌متر در منطقه‌ای نشان بدهد:

| | | | | | | | | | |
|------|------|------|-------|-------|-------|------|------|------|------|
| ۴۴/۰ | ۳۵/۴ | ۲۸/۴ | ۳۷/۰ | ۴۶/۰ | ۳۵/۴ | ۱۹/۴ | ۲۰/۴ | ۵۶/۴ | ۴۳/۲ |
| ۳۶/۲ | ۳۸/۴ | ۴۹/۲ | ۳۱/۸ | ۸۶/۴ | ۱۲/۶ | ۲۷/۴ | ۱۴/۰ | ۳۹/۴ | ۳۹/۴ |
| ۱۵/۸ | ۲۸/۸ | ۳۸/۰ | ۴۴/۰ | ۳۸/۴ | ۷۶/۰ | ۲۳/۰ | ۱۱/۴ | ۳۹/۸ | ۳۰/۲ |
| ۲۹/۲ | ۴۰/۶ | ۴۹/۶ | ۳۰/۴ | ۱۲/۲ | ۱۲۳/۸ | ۴۲/۰ | ۴۷/۰ | ۳۶/۴ | ۳۹/۲ |
| ۳۵/۲ | ۵۶/۴ | ۳۱/۰ | ۴۵/۰ | ۹۰/۲ | ۱۰۰/۰ | ۳۹/۰ | ۳۷/۰ | ۴۹/۴ | ۲۸/۲ |
| ۱۲/۶ | ۲۷/۰ | ۴۷/۸ | ۵۲/۶ | ۴۱/۰ | ۴۰/۰ | ۲۸/۰ | ۲۳/۶ | ۳۷/۶ | ۳۷/۸ |
| ۳۰/۰ | ۴۸/۵ | ۱۸/۰ | ۴۱/۰ | ۲۲/۶ | ۲۴/۲ | ۸۹/۶ | ۹۰/۴ | ۴۳/۰ | ۲۹/۸ |
| ۵۶/۲ | ۲۴/۸ | ۱۲/۶ | ۵۳/۶ | ۱۲۵/۴ | ۱۶/۲ | ۳۹/۰ | ۴۰/۸ | ۳۳/۶ | ۳۹/۴ |
| ۴۵/۶ | ۳۷/۴ | ۱۸/۰ | ۵۰/۶ | ۱۰۳/۴ | ۵۲/۴ | ۴۵/۲ | ۶۴/۶ | ۲۲/۲ | ۶۰/۰ |
| ۴۲/۲ | ۴۲/۰ | ۱۶/۲ | ۱۰۸/۲ | ۴۶/۰ | ۴۲/۶ | ۳۹/۴ | ۳۷/۶ | ۴۱/۴ | ۴۰/۴ |

شاید در نخستین مرحله، بهتر این باشد که مقادیر فرین، یعنی بیشترین و کمترین مقدار داده‌ها را در این مجموعه معین کنیم. با مقداری وارسی معلوم می‌شود که کمترین مقدار $۱۱/۴$ و بیشترین مقدار $۱۲۵/۴$ است. می‌توان همین شیوه را ادامه داد و دوّمین مقدار، سومین مقدار، و ... را از نظر بزرگی پیدا کرد یا به عبارت دیگر داده‌ها را به ترتیب صعودی مرتب کرد (می‌توان داده‌ها را به ترتیب نزولی نیز مرتب کرد). پس از انجام این کار فهرست زیر را خواهیم داشت:

| | | | | | | | | | |
|------|------|------|------|------|------|------|------|------|-------|
| ۱۱/۴ | ۱۸/۰ | ۲۷/۰ | ۳۰/۴ | ۳۷/۰ | ۳۹/۲ | ۴۱/۰ | ۴۴/۰ | ۴۹/۶ | ۷۴/۰ |
| ۱۲/۲ | ۱۹/۴ | ۲۷/۴ | ۳۱/۰ | ۳۷/۴ | ۳۹/۴ | ۴۱/۰ | ۴۴/۰ | ۵۰/۶ | ۸۶/۴ |
| ۱۲/۶ | ۲۰/۲ | ۴۸/۰ | ۳۱/۸ | ۳۷/۶ | ۳۹/۴ | ۴۱/۴ | ۴۵/۰ | ۵۱/۴ | ۸۹/۶ |
| ۱۲/۶ | ۲۰/۴ | ۴۸/۲ | ۳۲/۴ | ۳۷/۶ | ۳۹/۴ | ۴۲/۰ | ۴۵/۶ | ۵۲/۶ | ۹۰/۲ |
| ۱۲/۶ | ۲۲/۲ | ۲۸/۴ | ۳۳/۶ | ۳۷/۸ | ۳۹/۴ | ۴۲/۰ | ۴۶/۰ | ۵۳/۶ | ۹۰/۴ |
| ۱۴/۰ | ۲۲/۶ | ۲۸/۸ | ۳۵/۲ | ۳۸/۰ | ۳۹/۸ | ۴۲/۲ | ۴۷/۰ | ۵۶/۲ | ۱۰۰/۰ |
| ۱۵/۸ | ۲۳/۰ | ۴۹/۲ | ۳۵/۴ | ۳۸/۴ | ۴۰/۰ | ۴۲/۶ | ۴۷/۸ | ۵۶/۴ | ۱۰۳/۴ |
| ۱۶/۲ | ۲۳/۶ | ۴۹/۸ | ۳۵/۴ | ۳۸/۴ | ۴۰/۴ | ۴۳/۰ | ۴۸/۵ | ۵۶/۴ | ۱۰۸/۲ |
| ۱۶/۲ | ۲۴/۲ | ۳۰/۰ | ۳۶/۲ | ۳۹/۰ | ۴۰/۶ | ۴۳/۲ | ۴۹/۲ | ۵۰/۰ | ۱۲۳/۴ |
| ۱۸/۰ | ۲۴/۸ | ۳۰/۲ | ۳۷/۰ | ۳۹/۰ | ۴۰/۸ | ۴۲/۰ | ۴۹/۴ | ۶۴/۶ | ۱۲۵/۴ |

با اینکه مرتب کردن داده‌ها به ترتیب نزولی یا صعودی، برای مجموعه‌های بزرگ کار چندان آسانی نیست، ملاحظه می‌کنیم که این مجموعه‌ی جدید، چندان تفاوتی با داده‌های قبلی ندارد و نمی‌توان تنها با نظرکردن به آنها از کم و کیف «توزیع» آنها

اطلاعی به دست آورد. بنابر این سودمند خواهد بود که راهی دیگر برای پردازش داده‌ها پیدا کنیم. یکی از روش‌هایی که اخیراً به وجود آمده است، نمودار ساقه و برگ است که تصویری از وضع کلی داده‌ها را در اختیار می‌گذارد. برای سهولت این روش را برای داده‌های مثال زیر توضیح می‌دهیم.

مثال ۱-۱. فرض کنید داده‌های زیر مربوط به تعداد شعب یک بانک در بیست شهر باشند.

| | | | | | | | | | |
|----|----|----|----|----|----|----|----|----|----|
| ۶۹ | ۸۴ | ۵۲ | ۹۳ | ۶۱ | ۷۴ | ۷۹ | ۶۵ | ۸۸ | ۶۳ |
| ۵۷ | ۶۴ | ۶۷ | ۷۲ | ۷۴ | ۵۵ | ۸۲ | ۶۱ | ۶۸ | ۷۷ |

رقم یکان و دهگان هر عدد را جداگانه در نظر بگیرید و اعدادی را که رقم دهگان آنها یکی هستند، یکجا گردآوری کنید. برای این منظور مثلاً عدد ۶۹ را به صورت ۶|۹ در نظر می‌گیریم. رقم دهگان، یعنی ۶ را در ستون اول درجای مناسب خود می‌نویسیم و رقم یکان را در سمت راست علامت | وارد می‌کنیم. برای نوشتمن ۶۴، چون قبلًاً ۶ را در ستون اوّل نوشته‌ایم، کافی است تنها ۴، یعنی رقم یکان را در سمت راست علامت | وارد کنیم. به این ترتیب اعداد به شکل زیر گروه‌بندی می‌شوند:

| | | | |
|---|---|---|---|
| ۵ | ۲ | ۷ | ۵ |
| ۶ | ۹ | ۱ | ۵ |
| ۷ | ۲ | ۹ | ۴ |
| ۸ | ۴ | ۸ | ۲ |
| ۹ | ۳ | | |

اوّین سطر این جدول، یعنی ۵|۷ نشان می‌دهد که این فهرست شامل اعداد ۵۲، ۵۷، ۵۵ است و دومین سطر بیان می‌کند که فهرست شامل ۸ عدد است که رقم دهگان آنها ۶ است.

این جدول، نمایش ساقه و برگ نامیده می‌شود، زیرا هر سطر حالت ساقه‌ای را نشان می‌دهد، و هر رقم واقع در سمت راست خط عمودی | حالت یک برگ را نشان می‌دهد. برای تشکیل دادن این جدول، ابتدا با ساقه‌ها، یعنی

| |
|---|
| ۵ |
| ۶ |
| ۷ |
| ۸ |
| ۹ |

شروع می‌کنیم. سپس با مراجعت به فهرست داده‌ها، به ترتیب هر عدد را در نظر گرفته و دوّمین رقم آن را به عنوان برگ در روی ساقه‌ی مربوط به آن درج می‌کنیم. نمایش ساقه و برگ همان اطلاعاتی را منعکس می‌سازد که در فهرست اصلی موجود است اما بسیار فشرده‌تر از آن است. بخش عمده‌ی کار مرتب کردن داده‌ها، با ایجاد نمایش ساقه و برگ انجام می‌شود. با مرتب کردن برگ‌ها که کار نسبتاً ساده‌تری است، می‌توان کار مرتب کردن داده‌ها را گامی جلوتر برد. در این مثال جدول زیر را به دست می‌آوریم.

| | | | |
|---|---|---|---|
| ۵ | ۱ | ۵ | ۷ |
| ۶ | ۱ | ۱ | ۳ |
| ۷ | ۲ | ۴ | ۴ |
| ۸ | ۲ | ۴ | ۷ |
| ۹ | ۳ | | |

در واقع، نمایش ساقه و برگ و سپس مرتب کردن برگ‌ها شبیه‌کاری است که در فرهنگ‌های لغت انجام می‌شود، واژه‌ها ابتدا بر حسب اولین حرف خود مرتب می‌شوند و سپس بر حسب حروف دوّم و سوم و ... البته در این حالت مانند این است که کلمه دو حرفی است.

توجه می‌کنیم که قرار گرفتن هر فهرستی از داده‌ها در یک نمایش ساقه و برگ به سادگی آنچه در مثال اخیر انجام شد، نیست. گاهی با زیاد شدن ارقام و پراکندگی داده‌ها می‌توان از گردشده‌ی داده‌ها یا حذف برخی رقم‌های اعشاری، استفاده کرد.

در زیر نمودار ساقه و برگ برای مثال ۱-۱، به کمک نرم افزار مینیتوب انجام شده است.

```
MTB > SET C1
DATA> 69 84 52 93 61 74 79 65 88 63
DATA> 57 64 67 72 74 55 82 61 68 77
DATA> END
MTB > Stem-and-Leaf C1;
SUBC> Increment 10.
Character Stem-and-Leaf Display
Stem-and-leaf of C1      N = 20
Leaf Unit = 1.0
 3 5 257
(8) 6 11345789
 9 7 24479
 4 8 248
 1 9 3
```

تمرین

- ۱-۱. برای داده های جدول بارندگی سالانه‌ی ۱۰۰ شهر،
 الف) نمایش ساقه و برگ را به صورت دستی تشکیل دهید.
 ب) باردیگر به کمک نرم افزار با تغییر مقدار در دستور Increment، نمودارهای مختلف ساقه و برگ را رسم کنید.

۲-۲-۱ توزیع های فراوانی

تعداد داده هایی که در یک مطالعه‌ی آماری گردآوری می‌شوند، معمولاً بیشتر از آن است که تنها با نگاه کردن به آنها، بتوان اطلاعی درباره‌ی پدیده‌ی مورد مطالعه به دست آورد و لازم است که این داده‌ها را به نحوی سازمان دهیم. یک راه ساده برای این منظور، گردآوردن آنها در قالب یک نمایش ساقه و برگ است. اما برای مطالعه دقیق‌تر و مفصل‌تر داده‌ها، لازم است آنها را با دقت بیشتری سازماندهی کنیم. یکی از روش‌های بسیار رایج، استفاده از توزیع‌های فراوانی برای رده‌بندی یا گروه‌بندی داده‌هاست. به عنوان مثال فرض کنید که در یک نمونه‌گیری از دانشجویان دانشگاه پیام‌نور در شیراز، ۱۶۱ نفر رشته‌ی مهندسی کامپیوتر، ۴۷ نفر رشته‌ی علوم کامپیوتر، ۲۸ نفر

رشته‌ی ریاضی کاربردی، و ۱۱ نفر رشته‌ی ریاضی محض بوده‌اند. دسته‌بندی این افراد به صورت زیر است.

جدول ۱-۱. دسته‌بندی نمونه‌ای از دانشجویان

دانشگاه پیام‌نور در مرکز شیراز

| | |
|----|-----------------|
| ۴۱ | مهندسی کامپیوتر |
| ۴۷ | علوم کامپیوتر |
| ۲۸ | ریاضی کاربردی |
| ۱۱ | ریاضی محض |

چنین جدولی، یک توزیع رسته‌ای یا توزیع کیفی نامیده می‌شود و برای دسته‌بندی داده‌های غیر عددی در قالب رسته‌ها از آن استفاده می‌شود. عموماً برای سازمان دادن داده‌های غیر عددی یا داده‌هایی که برحسب مقیاس اسمی بیان شده‌اند، از توزیع‌های رسته‌ای استفاده می‌کنیم.

برای داده‌های عددی برحسب مقیاس‌های دیگر، از توزیع‌های کمی یا توزیع‌های عددی استفاده می‌کنیم. در این حالت، به طوری که خواهیم دید، داده‌ها را در بازه‌ها یا فاصله‌هایی از اعداد قرار می‌دهیم. اعم از اینکه برای داده‌ها توزیع‌های رسته‌ای یا عددی تشکیل داده باشیم یا خیر، خواهیم گفت که داده‌ها را در یک جدول فراوانی خلاصه کرده‌ایم. برای آنکه نحوه‌ی تشکیل جدول‌های فراوانی را به تفصیل بررسی کنیم، مثالی می‌آوریم.

مثال ۱-۲. داده‌های جدول ۱-۲ میزان اکسید سولفور منتشر شده در هوا در ۸۰ روز (برحسب تن) را نشان می‌دهند.

| | | | | | | | | | |
|------|------|------|------|------|------|------|------|------|------|
| ۱۵/۸ | ۲۶/۴ | ۱۷/۳ | ۱۱/۲ | ۲۳/۹ | ۲۴/۸ | ۱۸/۷ | ۱۳/۹ | ۹/۰ | ۱۳/۲ |
| ۲۲/۷ | ۹/۸ | ۶/۲ | ۱۴/۷ | ۱۷/۵ | ۲۶/۱ | ۱۲/۸ | ۲۸/۶ | ۱۷/۶ | ۲۳/۷ |
| ۲۶/۸ | ۲۲/۷ | ۱۸/۰ | ۲۰/۵ | ۱۱/۰ | ۲۰/۹ | ۱۵/۵ | ۱۹/۴ | ۱۶/۷ | ۱۰/۷ |
| ۱۹/۱ | ۱۵/۴ | ۲۲/۹ | ۲۶/۶ | ۲۰/۴ | ۲۱/۴ | ۱۹/۲ | ۲۵/۶ | ۱۶/۹ | ۱۹/۰ |
| ۱۸/۵ | ۲۳/۰ | ۲۴/۶ | ۲۰/۱ | ۱۶/۲ | ۱۸/۰ | ۷/۷ | ۱۳/۵ | ۲۳/۵ | ۱۴/۵ |
| ۱۴/۴ | ۲۹/۰ | ۱۹/۴ | ۱۷/۱ | ۲۰/۸ | ۲۴/۳ | ۲۲/۵ | ۲۴/۶ | ۱۸/۴ | ۱۸/۱ |
| ۸/۳ | ۲۱/۹ | ۱۲/۳ | ۲۲/۳ | ۱۳/۳ | ۱۱/۸ | ۱۹/۳ | ۲۰/۰ | ۲۴/۷ | ۳۱/۸ |
| ۲۵/۹ | ۱۰/۵ | ۱۵/۹ | ۲۷/۵ | ۱۸/۱ | ۱۷/۹ | ۹/۴ | ۲۴/۱ | ۲۰/۱ | ۲۸/۵ |

جدول ۲-۱. داده‌های میزان اکسید سولفور منتشره در هوای (برحسب تن) در ۸۰ روز

برای تشکیل جدول فراوانی در نخستین گام باید تعداد رده‌ها یا گروه‌هایی را تعیین کنیم که می‌خواهیم داده‌ها را در آنها جای دهیم. تعیین تعداد رده‌ها تا حدی به سلیقه و نظر محقق بستگی دارد. اما چندین قاعده‌ی کلی برای انتخاب تعداد رده‌ها پیشنهاد شده است که از آن جمله می‌توان استفاده از هریک از فرمول‌های

$$n = 2^k \quad (1-1)$$

$$k = 1 + \lceil \log_2(n) \rceil \quad (2-1)$$

را ذکر کرد که در آن n تعداد داده‌ها و k تعداد رده‌های است. آنچه همواره باید به خاطر داشت این است که:

به ندرت تعداد رده‌ها را کمتر از ۵ و بیشتر از ۱۵ انتخاب می‌کنیم، تعداد دقیق رده‌ها به تعداد اندازه‌گیری‌ها یا مشاهدات بستگی دارد.

وقتی تعداد رده‌ها را بزرگ انتخاب می‌کنیم، با همان مشکلی روبه‌رو هستیم که قبل از تلخیص کردن داده‌ها و کارکردن با داده‌های اصلی در پیش است. وقتی تعداد رده‌ها کم باشد، اطلاعات نسبتاً قابل توجهی درهم ادغام و از دقت نتیجه‌گیری کاسته می‌شود. برای داده‌های مثال حاضر، چون $n = 80$ ، لذا می‌توان تعداد رده‌ها را ۶ یا ۷ گرفت. پس از تعیین تعداد رده‌ها، باید خود رده‌ها یا طبقه‌هایی را که باید داده‌های ما را در برگیرند، مشخص کنیم. در تعیین رده‌ها به قواعد زیر توجه می‌کنیم:

۱. رده‌ها باید طوری انتخاب شوند که همه‌ی داده‌ها را دربرگیرند.

۲. حتماً باید مطمئن شد که هر داده تنها در یک رده قرار می‌گیرد، به عبارت دیگر رده‌ها نباید عضو مشترک داشته باشند.
۳. در صورت امکان طول رده‌ها را برابر انتخاب می‌کنیم، و به علاوه سعی می‌کنیم که کران رده‌ها یا حدود رده‌ها مضری از $5, 10, 100, \dots$ باشد. منظور از حدود هر رده، کمترین و بیشترین مقدار از داده‌های است که می‌توانیم در رده‌ای معین قرار دهیم. در این مثال، چون کمترین مقدار داده‌ها $\frac{1}{2}$ و بیشترین مقدار داده‌ها $\frac{31}{8}$ است، پس برای این داده‌ها مقدار کوچکترین داده - مقدار بزرگترین داده = دامنه تغییرات

$$= \frac{31}{8} - \frac{1}{2} = \frac{25}{6}$$

بنابر این اگر طول رده‌ها را برابر بگیریم طول هر رده بنابر فرمول

$$\text{تعداد رده} / \text{دامنه تغییرات} = \text{طول رده} = c$$

برابر $\frac{4}{14} \equiv \frac{4}{25/6}$ خواهد بود. در این مثال برای سادگی طول مشترک داده‌ها را برابر $\frac{4}{14}$ انتخاب خواهیم کرد. اگر بنابر توصیه‌های ذکر شده، حد پایین رده‌ی اول، یعنی کوچکترین مقداری را که بتوان در رده‌ی اول قرار گیرد برابر $\frac{5}{6}$ انتخاب کنیم، در این صورت می‌توانیم از 7 رده‌ی $\frac{9}{12}, \frac{9}{10}, \dots, \frac{9}{5}$ استفاده کنیم. توجه می‌کنیم که طول هر رده $\frac{4}{14}$ است زیرا مثلاً رده‌ی اول $\frac{9}{6}$ شامل کلیه‌ی مقادیر از $\frac{8}{95}$ تا $\frac{8}{95}$ است به این دلیل که داده‌های ما که در این مثال کمیت‌های پیوسته هستند، به نزدیکترین یک دهم تن گرد شده‌اند. به عنوان مثال داده‌ای که به صورت $\frac{8}{95}$ ثبت شده است در اصل مقداری بین $\frac{8}{95}$ و $\frac{8}{95}$ بوده و به نزدیکترین یک دهم تن گرد شده است. بنابر این حدود واقعی رده‌ی اول مقادیر $\frac{8}{95}, \frac{12}{95}, \dots, \frac{12}{95}$ و حدود واقعی رده‌ی دوم $\frac{8}{95}, \frac{12}{95}, \dots, \frac{12}{95}$ است. حدود واقعی رده‌ها را کرانه‌های رده‌ها می‌نامیم.

اینک برای داده‌های مثال حاضر، جدولی تنظیم می‌کنیم و در اوّلین ستون آن حدود رده‌ها را درج می‌کنیم. ستون دیگری به نام ستون خط‌نشان یا چوب خط در نظر می‌گیریم. اینک داده‌ها را یک‌به‌یک می‌خوانیم و در برابر رده‌ای که داده به آن تعلق دارد، یک خط عمودی به شکل | وارد می‌کنیم. برای سهولت در شمارش آنها، پس از هر چهار علامت خط مورب به صورت / به روی آنها می‌کشیم و هر چنین مجموعه‌ای به نشانه پنج مورد از داده‌هاست. حال به طوری که در جدول دیده می‌شود، تعداد «خط نشان‌ها» را شمرده در ستون سوم همین جدول وارد می‌کنیم و به این ترتیب فراوانی هر رده به دست می‌آید. فراوانی رده‌ی n را با نماد f_i نمایش می‌دهیم، بدینهی است

$$\sum_{i=1}^k f_i = n \quad \text{که}$$

جدول ۱-۳. جدول فراوانی اوّلیه مربوط به میزان اکسید‌سولفور منتشره در هوای

| حدود رده | خط نشان | فراوانی |
|-----------|---------|---------|
| ۵/۰-۸/۹ | | ۳ |
| ۹/۰-۱۲/۹ | | ۱۰ |
| ۱۳/۰-۱۶/۹ | | ۱۴ |
| ۱۷/۰-۲۰/۹ | | ۲۵ |
| ۲۱/۰-۲۴/۹ | | ۱۷ |
| ۲۵/۰-۲۸/۹ | | ۹ |
| ۲۹/۰-۳۲/۹ | | ۲ |
| مجموع | | ۸۰ |

برای آنکه بتوان استفاده‌های گسترده‌تری از جدول فراوانی به عمل آورد، با افزودن ستون‌های دیگری، آن را کامل‌تر می‌کنیم. برای این منظور ستون دیگری به جدول فراوانی اضافه می‌کنیم که ستون نماینده‌ی رده‌ها یا مرکز رده‌ها نامیده می‌شود. برای آنکه نماینده‌ی هر رده را به دست آوریم، حد بالای هر رده را با حد پایین همان رده جمع و حاصل را برابر تقسیم می‌کنیم. نماینده‌ی هر رده را می‌توان با پیدا کردن

مجموع کرانه‌های پایین و بالای هر رده و تقسیم حاصل بر ۲ نیز به دست آورده. برای داده‌های مثال اخیر نماینده‌های رده‌ها به ترتیب مقادیر

$$\frac{۲۹/۰+۳۲/۹}{۲} = ۳۰/۹۵, \frac{۹/۰+۱۲/۹}{۲} = ۱۰/۹۵, \frac{۵/۰+۸/۹}{۲} = ۶/۹۵$$

هستند. نماینده‌ی رده‌ی i نام را با m_i نشان می‌دهیم.

همچنین ستون دیگری به نام ستون فراوانی درصد به جدول فراوانی اضافه می‌کنیم. برای محاسبه‌ی فراوانی درصد هر رده، فراوانی رده را بر تعداد کل داده‌ها تقسیم و حاصل را در 100 ضرب می‌کنیم. برای داده‌های مثال حاضر، فراوانی درصد

$$\text{رده‌ها به ترتیب اعداد } \frac{۳\times ۱۰۰}{۸۰}, \frac{۱۰\times ۱۰۰}{۸۰}, \dots, \frac{۱۲\times ۱۰۰}{۸۰} \text{ خواهند بود.}$$

جدول ۱-۴. جدول فراوانی مربوط به میزان اکسید سولفور متشره در هوا

| حدود رده | کرانه‌های رده | نماینده | فراوانی | فراوانی درصد |
|-------------|---------------|---------|---------|--------------|
| $۵/۰-۸/۹$ | $۴/۹۵-۸/۹۵$ | $۶/۹۵$ | ۳ | $۳/۷۵$ |
| $۹/۰-۱۲/۹$ | $۸/۹۵-۱۲/۹۵$ | $۱۰/۹۵$ | ۱۰ | $۱۲/۵$ |
| $۱۳/۰-۱۶/۹$ | $۱۲/۹۵-۱۶/۹۵$ | $۱۴/۹۵$ | ۱۴ | $۱۷/۵$ |
| $۱۷/۰-۲۰/۹$ | $۱۶/۹۵-۲۰/۹۵$ | $۱۸/۹۵$ | ۲۵ | $۳۱/۲۵$ |
| $۲۱/۰-۲۴/۹$ | $۲۰/۹۵-۲۴/۹۵$ | $۲۲/۹۵$ | ۱۷ | $۲۱/۲۵$ |
| $۲۵/۰-۲۸/۹$ | $۲۴/۹۵-۲۸/۹۵$ | $۲۶/۹۵$ | ۹ | $۱۱/۲۵$ |
| $۲۹/۰-۳۲/۹$ | $۲۸/۹۵-۳۲/۹۵$ | $۳۰/۹۵$ | ۲ | $۲/۵$ |
| مجموع | | | ۸۰ | ۱۰۰ |

داشتن فراوانی‌های درصد برای مقایسه‌ی دو مجموعه از داده‌ها بسیار مفید است.

مثلاً اگر بخواهیم میزان اکسید سولفور متشره از کارخانه‌ای را با میزان اکسید سولفور کارخانه‌ای دیگر مقایسه کنیم، حتماً باید از فراوانی‌های درصد استفاده کنیم. زیرا امکان دارد تعداد داده‌های مربوط به دو کارخانه متفاوت باشد. به علاوه برای این نوع مقایسه‌ها و موارد دیگر می‌توان از فراوانی تجمعی استفاده کرد که تعداد «داده‌های کمتر از» مقداری مشخص را نشان می‌دهد. برای آنکه تعداد داده‌های کمتر از حد پایین هر رده را حساب کنیم، فراوانی‌های تمام رده‌های پیش از آن رده را با هم جمع

می‌کنیم. بدینهی است که فراوانی تجمعی آخرین رده برابر اندازه مشاهدات، n ، است. برای داده‌های مثال حاضر، جدول فراوانی تجمعی زیر را خواهیم داشت:

جدول ۱-۵. جدول فراوانی تجمعی مربوط به میزان اکسید سولفور منتشره در هوا

| فرابانی تجمعی | میزان اکسید سولفور بر حسب تن |
|---------------|------------------------------|
| ۵/۰ کمتر از | ۰ |
| ۹/۰ کمتر از | ۳ |
| ۱۳/۰ کمتر از | ۱۳ |
| ۱۷/۰ کمتر از | ۲۷ |
| ۲۱/۰ کمتر از | ۵۲ |
| ۲۵/۰ کمتر از | ۶۹ |
| ۲۹/۰ کمتر از | ۷۸ |
| ۳۳/۰ کمتر از | ۸۰ |

می‌توان این جدول را در جدول فراوانی قبلی، ادغام کرد و کلیهی این اطلاعات را یکجا در جدولی جامع ارائه داد.

جدول ۱-۶. جدول کامل فراوانی مربوط به میزان اکسید سولفور منتشره در هوا

| حدود رده | حدود رده | کرانهای رده | نماینده | فراوانی | فراوانی درصد | فرابانی تجمعی |
|-----------|-------------|-------------|---------|---------|--------------|---------------|
| ۵/۰-۸/۹ | ۴/۹۵-۸/۹۵ | ۶/۹۵ | ۳ | ۳/۷۵ | ۳ | |
| ۹/۰-۱۲/۹ | ۸/۹۵-۱۲/۹۵ | ۱۰/۹۵ | ۱۰ | ۱۲/۵ | ۱۳ | |
| ۱۳/۰-۱۶/۹ | ۱۲/۹۵-۱۶/۹۵ | ۱۴/۹۵ | ۱۴ | ۱۷/۵ | ۲۷ | |
| ۱۷/۰-۲۰/۹ | ۱۶/۹۵-۲۰/۹۵ | ۱۸/۹۵ | ۲۵ | ۳۱/۲۵ | ۵۲ | |
| ۲۱/۰-۲۴/۹ | ۲۰/۹۵-۲۴/۹۵ | ۲۲/۹۵ | ۱۷ | ۲۱/۲۵ | ۶۹ | |
| ۲۵/۰-۲۸/۹ | ۲۴/۹۵-۲۸/۹۵ | ۲۶/۹۵ | ۹ | ۱۱/۲۵ | ۷۸ | |
| ۲۹/۰-۳۲/۹ | ۲۸/۹۵-۳۲/۹۵ | ۳۰/۹۵ | ۲ | ۲/۵ | ۸۰ | |
| مجموع | | | ۸۰ | ۱۰۰ | | |

ساختن جدول ۱-۶ را به وسیلهی نرمافزار، برای تعیین ستون فراوانی، فراوانی درصد، و فراوانی تجمعی به صورت زیر انجام می‌دهیم.

MTB > SET C1

DATA> 15.8 22.7 26.8 19.1 18.5 14.4 8.3 25.9

DATA> 26.4 9.8 22.7 15.2 23.0 29.0 21.9 10.5

DATA> 17.3 6.2 18.0 22.9 24.6 19.4 12.3 15.9

```

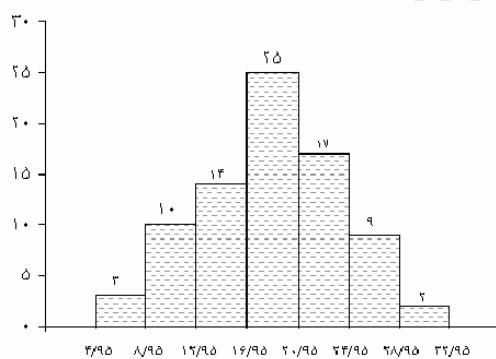
DATA> 11.2 14.7 20.5 26.6 20.1 17.1 22.3 27.5
DATA> 23.9 17.5 11.0 20.4 16.2 20.8 13.3 18.1
DATA> 24.8 26.1 20.9 21.4 18.0 24.3 11.8 17.9
DATA> 18.7 12.8 15.5 19.2 7.7 22.5 19.3 9.4
DATA> 13.9 28.6 19.4 25.6 13.5 24.6 20.2 24.1
DATA> 9.0 17.6 16.7 16.9 23.5 18.4 24.7 20.1
DATA> 13.2 23.7 10.7 19.0 14.5 18.1 31.8 28.5
DATA> END
MTB > NAME C1='DATA' C2='FREQ' C3='PERC' C4='CUMLFREQ'
MTB > LET C2(1)=SUM((C1>4.95)*(C1<8.95))
MTB > LET C2(2)=SUM((C1>8.95)*(C1<12.95))
MTB > LET C2(3)=SUM((C1>12.95)*(C1<16.95))
MTB > LET C2(4)=SUM((C1>16.95)*(C1<20.95))
MTB > LET C2(5)=SUM((C1>20.95)*(C1<24.95))
MTB > LET C2(6)=SUM((C1>24.95)*(C1<28.95))
MTB > LET C2(7)=SUM((C1>28.95)*(C1<32.95))
MTB > LET C3=C2/SUM(C2)*100
MTB > LET C4=PARSUM(C2)
MTB > PRINT C2-C4
Data Display
Row FREQ PERC CUMLFREQ
 1   3  3.75    3
 2  10 12.50   13
 3  14 17.50   27
 4  25 31.25   52
 5  17 21.25   69
 6   9 11.25   78
 7   2  2.50   80

```

۳-۲-۱ توصیف داده‌ها به کمک نمودارها

نمودارها و تصاویر علاوه بر اینکه مشخصه‌های اصلی داده‌ها را به نحو بارزی آشکار می‌کنند، به طوری که بعداً خواهیم دید از لحاظ نظری نیز بسیار مفیدند. رایج‌ترین نمودار آماری، بافت‌نگار نام دارد. برای رسم بافت‌نگار، از یک دستگاه محورهای مختصات متعامد استفاده می‌کنیم. کرانه‌های رده‌ها را روی محور افقی درج می‌کنیم و سپس مستطیل‌هایی بنا می‌کنیم به طوری که طول قاعده مستطیل مربوط به رده‌ی *i*ام، طول آن رده و ارتفاع آن، فراوانی رده باشد.

مثال ۱-۳. داده‌های مربوط به میزان اکسید سولفور منتشره در هوا که در مثال ۲-۱ ارائه شد، یکبار دیگر در نظر می‌گیریم. برای رسم بافت‌نگار مربوط به داده‌ها، به‌طوری که گفته‌یم کرانه‌های رده‌ها را بر روی محور افقی ثبت می‌کنیم و مستطیل‌هایی بر روی این رده‌ها بنا می‌کنیم که ارتفاع آنها برابر با فراوانی رده‌ی مربوط باشد. توجه می‌کنیم که مقیاس روی دو محور را می‌توان متفاوت در نظر گرفت. بافت‌نگار مربوط به این مثال به صورت شکل ۱-۱ است.



شکل ۱-۱. نمودار بافت‌نگار مربوط به میزان اکسید‌سولفور منتشره در هوا

شکل ۱-۱ را با نرم‌افزار با دستورات زیر ترسیم می‌شود.

```
MTB > NAME C5='MIDPOINT'
MTB > Set C5
DATA> 1( 6.95 : 30.95 / 4 )1
DATA> End.
MTB > Histogram 'DATA';
SUBC> MidPoint 'MIDPOINT';
SUBC> Bar.
```

نرم‌افزار به طور خودکار می‌تواند بدون هیچ‌گونه محدودیتی، هیستوگرام را رسم کند اما آنچه که در اینجا ارائه شد، دستوراتی است که خروجی آن با شکل ۱-۱ کاملاً مطابق است.

نمودار دیگری که کمتر مورد استفاده قرار می‌گیرد، چندبر فراوانی نام دارد. در این نمودار، مقادیر مرکز رده‌ها را بر روی محور افقی و فراوانی‌های مربوط را بر روی محور

عمودی درج می‌کنیم و نقطه‌های حاصل را به کمک خط‌های راست بهم وصل می‌کنیم. برای آنکه نمودار حاصل را به محور افقی «بیندیم»، دو رده‌ی فرضی، یکی به مرکز ۲/۹۵ و دیگری به مرکز ۳۴/۹۵، و هر دو با فراوانی صفر، در نظر می‌گیریم، و از اولین نقطه‌ی به دست آمده به نقطه‌ی (۳۰/۹۵, ۳) و از نقطه‌ی (۳۰/۹۵, ۲) به آخرین نقطه، یعنی (۳۴/۹۵, ۰) وصل می‌کنیم تا شکل ۲-۱ حاصل شود.



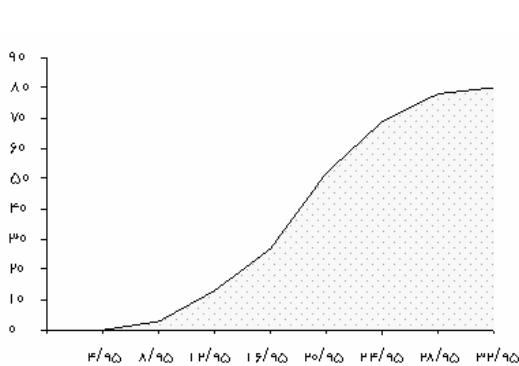
شکل ۲-۱. چندبر فراوانی توزیع داده‌های مربوط به میزان اکسید سولفور

دستورات لازم جهت ترسیم شکل ۲-۱، به شرح زیر است.

```
MTB > NAME C6='CUTPOINT'
MTB > Set C6
DATA> 1( 4.95 : 32.95 / 4 )1
DATA> End.
MTB > Histogram 'DATA';
SUBC> MidPoint 'CUTPOINT';
SUBC> Connect.
```

اگر به همان صورت که در مورد چندبر فراوانی عمل کردیم، نموداری برای فراوانی تجمعی به دست آوریم، نمودار حاصل او جایو نامیده می‌شود. برای رسم او جایو مربوط به فراوانی تجمعی میزان اکسید سولفور منتشره در هوای نقاطی را در دستگاه

مختصات در نظر می‌گیریم که طول هر نقطه، کران بالای هر رده و عرض آن فراوانی تجمعی آن رده باشد. در این صورت شکل ۳-۱ را به دست می‌آوریم.

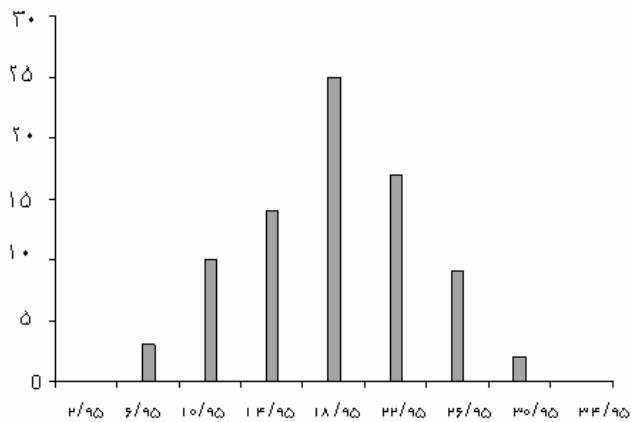


شکل ۳-۱ اوچایو توزیع داده‌های مربوط به میزان اکسید سولفور

شکل ۳-۱ را با دستورات زیر، بار دیگر رسم می‌کنیم.

```
MTB > Histogram 'DATA';
SUBC> Cumulative;
SUBC> Percent;
SUBC> MidPoint 'CUTPOINT';
SUBC> Connect.
```

می‌توان به جای بافت‌نگار، از نمودار میله‌ای نیز استفاده کرد. برای این منظور عیناً مانند زمانی عمل می‌کنیم که می‌خواهیم بافت‌نگار را رسم کنیم، با این تفاوت که به جای استفاده از کرانهای رده‌ها، از نماینده‌ی رده‌ها استفاده می‌کنیم. با اینکه هم برای داده‌های پیوسته و هم گسسته از بافت‌نگار و نمودار میله‌ای هر دو استفاده می‌شود، معمولاً از نمودار میله‌ای برای نمایش داده‌های گسسته و از بافت‌نگار برای داده‌های پیوسته استفاده می‌کنیم. نمودار میله‌ای داده‌های مربوط به میزان اکسید سولفور منتشره در هوا در شکل ۴-۱ رسم شده است.

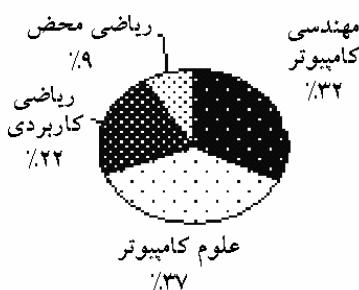


شکل ۱-۴. نمودار میله‌ای داده‌های مربوط به میزان اکسید سولفور

نمودار دیگری که اغلب در روزنامه‌ها و گزارش‌های مختلف مشاهده می‌شود، نمودار کلوچه‌ای یا نمودار دایره‌ای نام دارد و عمدتاً برای ارائه‌ی فراوانی نسبی یا فراوانی درصد به کار می‌رود. برای رسم نمودار دایره‌ای، کل زاویه‌ی مرکزی 360° مربوط به یک دایره را به نسبت فراوانی‌های نسبی یا درصد تقسیم می‌کنیم و سپس قطاع‌هایی در دایره رسم می‌کنیم که زاویه‌ی مرکزی هریک از آنها برابر با اعداد به دست آمده باشد. اگر فراوانی نسبی رده‌ای (نسبت فراوانی هر رده به تعداد کل مشاهده‌ها) f_i باشد، زاویه‌ی مرکزی قطاع مربوط به آن را برابر

$$\alpha_i = 360 \times f_i$$

می‌گیریم.



شکل ۱-۵. نمودار کلوچه‌ای مربوط به نمونه‌ای از دانشگاه پیام‌نور مرکز شیراز

شکل ۱-۵ نمودار کلوچه‌ای، مربوط به نمونه‌ای از دانشجویان دانشگاه پیام‌نور مرکز شیراز است.

با ورود دو ستون، یکی برای تعیین رده‌ها و دیگری تعیین فراوانی، نمودار ۱-۵ به دست می‌آید.

MTB > %Pie C7;
SUBC> Counts C8.

چنانچه که داده‌های اصلی را بدون جدول‌بندی شده وارد نرم‌افزار نموده‌اید لزومی به استفاده از دستور در خط دوم نمی‌پاشد.

تمرین

۱-۲. داده‌های زیر اندازه‌ی قد ۱۰۰ پس ۲۰ ساله در شهر شیراز است.

| | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| ۱۷۲ | ۱۵۸ | ۱۶۲ | ۱۶۳ | ۱۷۲ | ۱۶۴ | ۱۶۳ | ۱۸۴ | ۱۶۶ | ۱۵۷ |
| ۱۸۴ | ۱۶۵ | ۱۵۴ | ۱۷۰ | ۱۶۹ | ۱۶۷ | ۱۷۰ | ۱۷۷ | ۱۶۴ | ۱۶۳ |
| ۱۷۰ | ۱۶۹ | ۱۵۹ | ۱۷۷ | ۱۶۲ | ۱۸۰ | ۱۶۵ | ۱۵۳ | ۱۷۴ | ۱۵۳ |
| ۱۵۰ | ۱۷۲ | ۱۸۲ | ۱۸۴ | ۱۷۲ | ۱۶۹ | ۱۷۲ | ۱۷۰ | ۱۶۹ | ۱۵۸ |
| ۱۶۴ | ۱۶۶ | ۱۷۴ | ۱۷۵ | ۱۷۶ | ۱۵۲ | ۱۸۴ | ۱۶۲ | ۱۷۹ | ۱۵۱ |
| ۱۷۱ | ۱۶۴ | ۱۶۲ | ۱۷۱ | ۱۷۵ | ۱۵۹ | ۱۷۰ | ۱۵۴ | ۱۷۷ | ۱۶۲ |
| ۱۷۸ | ۱۶۹ | ۱۵۱ | ۱۶۴ | ۱۷۴ | ۱۶۱ | ۱۵۰ | ۱۷۳ | ۱۷۹ | ۱۵۹ |
| ۱۷۲ | ۱۵۹ | ۱۶۵ | ۱۵۶ | ۱۵۹ | ۱۶۴ | ۱۶۴ | ۱۷۷ | ۱۷۳ | ۱۶۱ |
| ۱۷۷ | ۱۵۳ | ۱۷۲ | ۱۵۰ | ۱۶۶ | ۱۷۱ | ۱۷۱ | ۱۶۹ | ۱۶۵ | ۱۶۰ |
| ۱۵۱ | ۱۷۰ | ۱۵۶ | ۱۶۵ | ۱۵۴ | ۱۵۶ | ۱۷۰ | ۱۷۲ | ۱۶۵ | ۱۸۳ |

الف) نمودار ساقه و برگ آن رارسم کنید.

ب) با استفاده از روش تعیین تعداد طبقات (یکی از آنها)، جدول فراوانی آن را به دست آورید.

ج. نمودار بافت‌نگار، چندبر آن رارسم کنید.

| | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| ۱۷۴ | ۱۵۸ | ۱۶۲ | ۱۶۳ | ۱۷۴ | ۱۶۲ | ۱۶۳ | ۱۸۴ | ۱۶۶ | ۱۵۷ |
| ۱۸۴ | ۱۶۵ | ۱۵۴ | ۱۷۰ | ۱۶۹ | ۱۶۷ | ۱۷۰ | ۱۷۷ | ۱۶۴ | ۱۶۳ |
| ۱۷۰ | ۱۶۹ | ۱۵۹ | ۱۷۷ | ۱۶۲ | ۱۸۰ | ۱۶۵ | ۱۵۳ | ۱۷۴ | ۱۵۳ |
| ۱۵۰ | ۱۷۲ | ۱۸۲ | ۱۸۴ | ۱۷۴ | ۱۶۹ | ۱۷۲ | ۱۷۰ | ۱۶۹ | ۱۵۸ |
| ۱۵۴ | ۱۶۶ | ۱۷۴ | ۱۷۵ | ۱۷۶ | ۱۵۲ | ۱۸۴ | ۱۶۲ | ۱۷۹ | ۱۵۱ |
| ۱۷۱ | ۱۶۴ | ۱۶۲ | ۱۷۱ | ۱۷۵ | ۱۵۹ | ۱۷۰ | ۱۵۴ | ۱۷۷ | ۱۶۲ |
| ۱۷۸ | ۱۶۹ | ۱۵۱ | ۱۶۴ | ۱۷۴ | ۱۶۱ | ۱۵۰ | ۱۷۳ | ۱۷۹ | ۱۵۹ |
| ۱۷۲ | ۱۵۹ | ۱۶۵ | ۱۵۶ | ۱۶۹ | ۱۶۴ | ۱۶۴ | ۱۷۷ | ۱۷۳ | ۱۶۱ |
| ۱۷۷ | ۱۵۳ | ۱۷۲ | ۱۵۰ | ۱۶۶ | ۱۷۱ | ۱۷۱ | ۱۶۹ | ۱۶۵ | ۱۶۰ |
| ۱۵۱ | ۱۷۰ | ۱۵۶ | ۱۶۵ | ۱۵۴ | ۱۵۶ | ۱۷۰ | ۱۷۲ | ۱۶۵ | ۱۸۳ |

۱-۳. بر اساس آماری از استان سیستان و بلوچستان طی ۲۵ اسفند ماه ۱۳۸۵ تا ۱۵ فروردین ۱۳۸۶، ۳۷۷۷۲ گردشگر از موزه‌ی شهر سوخته، ۱۲۶ گردشگر از موزه‌ی زابل، ۲۴۰۰ گردشگر از موزه‌ی چابهار، ۱۱۲۴ گردشگر از موزه‌ی ایرانشهر و ۱۵۶ گردشگر از موزه‌ی زاهدان دیدن کردند. نمودار مناسب برای توصیف داده‌های را در نظر گرفته آن را رسم کنید.

۱-۴. جدول ناقص زیر مربوط به تعداد روزهای لازم برای اتمام چهارچوب خانه‌های پیش ساخته تنظیم شده است.

| تعداد روزها | تعداد خانه‌ها | نسیم | فرابانی | تجمعی | فرابانی |
|-----------------|---------------|--------|---------|-------|---------|
| کمتر از ۶ - ۵ | ... | ... | ... | ... | ... |
| کمتر از ۷ - ۶ | ۱۹ | ... | ... | ۲۷ | |
| کمتر از ۸ - ۷ | ... | ۰ / ۳۵ | ۰ / ۳۵ | ... | |
| کمتر از ۹ - ۸ | ۱۵ | ۰ / ۱۵ | ۰ / ۱۵ | ... | |
| کمتر از ۱۰ - ۹ | ... | ... | ... | ۹۱ | |
| کمتر از ۱۱ - ۱۰ | ... | ... | ... | ... | |
| | ۱۰۰ | | ۱ / ۱۰۰ | | |

الف) جدول را تکمیل کنید.

ب) نمودار بافت‌نگار را رسیم کنید.

۱-۵. داده‌های زیر مربوط به هزینه‌ی حمل و نقل زمینی سالیانه‌ی شرکتی بر حسب میلیون ریال است.

| | | | | |
|----|----|----|----|----|
| ۲۴ | ۲۱ | ۲۰ | ۳۱ | ۲۵ |
| ۴۱ | ۴۳ | ۳۲ | ۱۷ | ۲۵ |
| ۳۲ | ۳۵ | ۲۶ | ۳۴ | |
| ۳۶ | ۳۹ | ۲۷ | ۲۸ | ۳۱ |
| ۲۱ | ۱۹ | ۲۵ | ۲۶ | ۳۱ |

جدولی بر مبنای پنج رده به طوری که شروع رده‌ی اول آن با پانزده باشد. همچنین نمودار مناسب برای توصیف داده‌ها را رسم کنید.

۱-۳ توصیف عددی داده‌ها

در بخش پیشین نحوه‌ی تلخیص داده‌ها، نحوه‌ی تشکیل جدول فراوانی، و شیوه‌ی توصیف داده‌ها را با رسم نمودارهای مختلف شرح دادیم. طبعاً این‌گونه نمودارها برای بیان چگونگی توزیع داده‌ها به قدر کافی گویا و اطلاع دهنده‌اند، اما جای ویژگی‌هایی از داده‌ها را که به کمک اعداد قابل توصیف است، نمی‌گیرند. منظور از توصیف عددی داده‌ها آن است که کمیت‌هایی عددی را از روی داده‌ها به دست آوریم که به بهترین نحو معرف داده‌ها باشند. این کمیت‌ها، که شاخص یا اندازه نیز نامیده می‌شوند بر دو نوع‌اند: **شاخص‌های گرایش مرکزی** یا اندازه‌های مکانی، و **شاخص‌های پراکندگی**. در زیر به شرح هریک از انواع این اندازه‌ها می‌پردازیم.

۱-۳-۱ شاخص‌های گرایش مرکزی

اغلب لازم است که مجموعه‌ای از داده‌ها را به کمک عدد واحدی بنماییم که به‌گونه‌ای معرف نمایم مجموعه‌ی داده‌ها باشد. اینکه عدد مزبور را چگونه انتخاب کنیم، بستگی به‌ویژگی خاصی دارد که می‌خواهیم به توصیف آن بپردازیم. در مطالعه‌ای معین، ممکن است علاقه‌ی ما به مقادیر فرین (کمترین و بیشترین مقدار) در مجموعه داده‌ها باشد، در مطالعه‌ای دیگر ممکن است توجه ما به عددی باشد که تنها ۱۰ درصد داده‌ها از آن بیشترند، و نظایر آنها. چندین اندازه را در بخش‌های آتی مورد بحث قرار می‌دهیم

که به گونه‌ای مرکز یا وسط مجموعه‌ای از داده‌ها را نشان می‌دهند و به همین دلیل آنها را شاخص‌های گرایش مرکزی می‌نامیم.

۱-۳-۱ میانگین

در بین انواع شاخص‌های گرایش مرکزی، مشهورتر از همه میانگین حسابی یا به اختصار میانگین است. میانگین مجموعه‌ای از مقادیر عبارت از مجموع آنها تقسیم بر تعدادشان است. این همان چیزی است که کم و بیش از بدو ورود به دبستان با آن با نام معدل آشنایی داریم؛ برای آنکه معدل نمرات دانش‌آموزی را پیدا کنیم، مجموع نمرات وی را بر تعداد نمره‌هایش تقسیم می‌کنیم.

می‌توان با استفاده از نمادهای ریاضی، فرمولی کلی برای میانگین ارائه داد: فرض کنید که n عدد داشته باشیم که آنها را حروف x_1, x_2, \dots, x_n نشان می‌دهیم. اگر میانگین این اعداد را با \bar{x} نشان دهیم آنگاه

$$\bar{x} = \frac{x_1 + x_2 + \dots + x_n}{n}$$

می‌توان با استفاده از نماد Σ ، این فرمول را حتی به صورتی ساده‌تر نوشت:

$$\bar{x} = \frac{\sum_{i=1}^n x_i}{n} \quad (۳-۱)$$

مثال ۱-۴. حجم آب حاصل از رسیدن در دریای خزر در پنج سال به قرار زیر بوده است:

$$66/91, 65/23, 81/87, 81/88, 79/76$$

میانگین حجم آب حاصل از رسیدن در این پنج سال چقدر است؟

حل. بنابر فرمول میانگین داریم،

$$\bar{x} = \frac{66/91 + 65/23 + 81/87 + 81/88 + 79/76}{5} = 77/13.$$

محاسبه‌ی میانگین با نرم‌افزار را با دستور زیر می‌توانید انجام دهید.

```

MTB > SET C7
DATA> 66.91 65.23 91.87 81.88 79.76
DATA> END
MTB > MEAN C7 K1
Column Mean
Mean of C7 = 77.130

```

توجه دارید که مقدار میانگین در ثابتی به نام K1 ذخیره شده است.

باید توجه کنیم که فرمول میانگین که در رابطه‌ی (۱-۳) ارائه شد، برای داده‌هایی است که هنوز در جدول فراوانی گروه‌بندی نشده‌اند. اگر داده‌ها را در جدول فراوانی تلخیص کرده باشیم و داده‌های اصلی را در اختیار نداشته باشیم، در این صورت نحوه محاسبه‌ی میانگین متفاوت است. فرض کنید که نماینده یا مرکز رده‌ها را با حروف m_1, m_2, \dots, m_k و فراوانی‌های متناظر را به ترتیب f_1, f_2, \dots, f_k نشان دهیم. در این صورت فرمول زیر برای میانگین داده‌های گروه‌بندی شده به دست می‌آید:

$$\bar{x} = \frac{f_1m_1 + f_2m_2 + \dots + f_km_k}{n},$$

یا به طور خلاصه،

$$\bar{x} = \frac{\sum_{i=1}^k f_i m_i}{n}. \quad (4-1)$$

بدیهی به علت فشرده شدن اطلاعات در جداول آماری و استفاده از نماینده‌های رده‌ها، مقدار میانگین با داده‌های گروه‌بندی نشده متفاوت خواهد شد.

مثال ۱-۵. میانگین داده‌های مربوط به میزان اکسید سولفور منتشره در هوا در ۸۰ روز را که در مثال ۲-۱ ارائه شده است، پیدا کنید.

حل. گرچه برای محاسبه‌ی میانگین داده‌های رده‌بندی شده تنها به نماینده‌های رده‌ها و فراوانی هر رده نیاز داریم، برای استفاده‌های آتی، جدول را به‌طور کامل مجدداً در زیر می‌آوریم.

جدول ۷-۱. جدول فراوانی با یک ستون اضافی برای محاسبه میانگین

| کرانه‌های رده‌ها | m_i | f_i | $f_i m_i$ |
|------------------|-------|-------|-----------|
| ۶/۹۵ - ۸/۹۵ | ۶/۹۵ | ۳ | ۲۰/۸۵ |
| ۸/۹۵ - ۱۰/۹۵ | ۱۰/۹۵ | ۱۰ | ۱۰۹/۵۰ |
| ۱۰/۹۵ - ۱۶/۹۵ | ۱۴/۹۵ | ۱۴ | ۲۰۹/۳۰ |
| ۱۶/۹۵ - ۲۰/۹۵ | ۱۸/۹۵ | ۲۵ | ۴۷۳/۷۵ |
| ۲۰/۹۵ - ۲۴/۹۵ | ۲۲/۹۵ | ۱۷ | ۳۹۰/۱۵ |
| ۲۴/۹۵ - ۲۸/۹۵ | ۲۶/۹۵ | ۹ | ۲۴۲/۵۵ |
| ۲۸/۹۵ - ۳۲/۹۵ | ۳۰/۹۵ | ۲ | ۶۱/۹۰ |
| مجموع | | ۸۰ | ۱۵۰۸ |

به طوری که ملاحظه می‌شود در این جدول علاوه بر اطلاعات موجود در جدول فراوانی پیشین، ستونی را اضافه کردایم که در آن نماینده‌ی هر رده در فراوانی نظیر آن ضرب و حاصل در این ستون درج شده است. مجموع اعداد واقع در این ستون همان

$$\sum_{i=1}^v f_i m_i = 1508 \text{ است.} \quad \sum_{i=1}^v f_i m_i = 1508, \text{ پس}$$

$$\bar{x} = \frac{1508}{80} = 18.85.$$

با توجه به ستون‌های وارد شده به نرم‌افزار در بخش‌های قبلی، مقدار میانگین برای داده‌های مثال ۱-۱ که در ساختار جدولی آماری خلاصه شده است، به شرح زیر است.

MTB > PRINT C5 C2

Data Display

Row MIDPOINT FREQ

| | | |
|---|-------|----|
| 1 | 6.95 | 3 |
| 2 | 10.95 | 10 |
| 3 | 14.95 | 14 |
| 4 | 18.95 | 25 |
| 5 | 22.95 | 17 |
| 6 | 26.95 | 9 |
| 7 | 30.95 | 2 |

MTB > LET K2=SUM(C5*C2)/SUM(C2)

MTB > PRINT K2

Data Display

K2 18.8500

مجدداً تأکید می‌کنیم که این میانگین با میانگین واقعی داده‌ها کمی تفاوت دارد و دلیل آن این است که مثلاً در رده‌ی اوّل به‌جای هر سه قلم از داده‌ها که مقادیری بین ۱۴/۹۵ تا ۸/۹۵ دارند، عدد ۶/۹۵، یعنی نماینده‌ی این رده، را قراردادیم. برای بررسی بیشتر این نکته میانگین داده‌های خام را به دست می‌آوریم

MTB > MEAN C1 K3
 Column Mean
 Mean of DATA = 18.930

خروجی نشان می‌دهد که میانگین داده‌های خام، برابر با ۱۸/۹۳ است.

میانگین، متداول‌ترین اندازه‌ی گرایش مرکزی است. می‌توان دلایل زیر را برای مورد توجه بودن آن برشمود:

۱. می‌توان آن را برای هر مجموعه‌ای از داده‌های عددی محاسبه کرد.
۲. مجموعه‌ای از داده‌های عددی یک و فقط یک میانگین دارد، و به عبارت دیگر منحصر به فرد است.
۳. می‌توان اعمال عددی بیشتری روی آن انجام داد (مثلاً اگر میانگین چند مجموعه از داده‌ها را داشته باشیم، به سادگی می‌توانیم میانگین مجموعه‌ای از داده‌ها را که از تلفیق کردن کلیه‌ی داده‌ها یک‌جا به دست آمده است حساب کنیم).
۴. میانگین نسبتاً قابل اطمینان است به‌این معنی که اگر چند نمونه مختلف را از جامعه‌ای استخراج کنیم، میانگین‌ها تفاوت خیلی زیادی با هم ندارند.
۵. در محاسبه‌ی میانگین همه داده‌ها به حساب می‌آیند.
۶. با اعمال تبدیل زیر بر روی داده‌ها

$$y_i = ax_i + b ; \quad i = 1, 2, \dots, n$$

رابطه‌ی زیر برقرار است:

$$\bar{y} = a\bar{x} + b$$

که در آن a و b اعداد حقیقی ثابتی هستند. خصوصاً با انتخاب $\bar{x} = b = 1$ و $a = 1 - \bar{y}$ ، نتیجه می‌شود.

۲-۱-۳-۱ میانه

در بالا یکی از امتیازات میانگین را در به حساب آوردن کلیه داده‌ها دانستیم. منظور این است که برای محاسبه میانگین باید همه داده‌ها را با هم جمع و بر تعداد کل آنها تقسیم کرد. این امر گرچه ممکن است از جهتی امتیاز به حساب آید، در مواردی نیز ممکن است نقطه‌ی ضعف محضوب شود. به عنوان مثال فرض کنید که اعداد زیر حقوق ماهانه‌ی (برحسب هزار تومان) اعضای شرکتی را نشان دهد که پنج عضو دارد

$$\begin{array}{c} ۲۱۵ \quad ۱۹۵ \quad ۲۴۵ \quad ۱۸۵ \quad ۹۸۵ \\ \bar{x} = \frac{۲۱۵ + ۱۹۵ + ۲۴۵ + ۱۸۵ + ۹۸۵}{۵} = ۱۷۶. \end{array}$$

اگر میانگین حقوق را با \bar{x} نشان دهیم، آنگاه

به طوری که ملاحظه می‌کنیم، رقم ۹۸۵۰۰۰ تومان حقوق یک فرد در مقایسه با حقوق بقیه افراد موجب شده است که میانگین حقوق، ۳۶۵۰۰۰ تومان از کار درآید که نماینده‌ی واقعی حقوق‌های کارمندان این شرکت نیست. دلیل این کار، همان‌طور که گفتیم، به حساب آمدن همه اعضای مجموعه داده‌ها و در نتیجه تأثیر غیرعادی اقلامی از داده‌های است که شاید با بقیه داده‌ها هماهنگ به نظر نرسند. در چنین مواردی می‌توان اندازه گرایش مرکزی دیگری، به نام میانه را محاسبه کرد که آن نیز «وسط» یا «مرکز» داده‌ها را نشان می‌دهد. برای تعیین میانه، داده‌ها را به ترتیب صعودی یا غیرنژولی مرتب می‌کنیم. اگر تعداد داده‌ها فرد باشد، داده‌ی وسطی میانه است. اگر تعداد داده‌ها زوج باشد، میانگین دو داده‌ی وسطی را میانه به حساب می‌آوریم. به عبارت دیگر اگر تعداد داده‌ها را وقتی فرد است با $n = 2k + 1$ نشان دهیم، آنگاه پس از مرتب کردن داده‌ها به ترتیب صعودی یا نژولی، k تا از داده‌ها در سمت چپ داده‌ی $(k+1)$ ام و k تا از داده‌ها در سمت راست آن قرار می‌گیرند و لذا داده‌ی $(k+1)$ امی

میانه است. اگر تعداد داده‌ها زوج باشد و $n=2k$ ، در این صورت پس از مرتب کردن داده‌ها به صورت صعودی یا نزولی $1-k$ تا k تا از داده‌ها در سمت چپ داده‌ی k ام و $-k$ تا از داده‌ها در سمت راست داده‌ی $(k+1)$ ام قرار دارند و لذا میانه برابر با میانگین حسابی داده‌ی k ام و $(k+1)$ ام است. به عنوان مثال اگر حقوق پنج عضو شرکت را که در بالا ذکر شد، به ترتیب صعودی مرتب کنیم، ارقام زیر را خواهیم داشت:

$$185 \quad 195 \quad 215 \quad 245 \quad 985$$

چون تعداد داده‌ها فرد است، لذا میانه‌ی داده‌ها برابر $\frac{215}{2}$ است. معمولاً میانه‌ی داده‌ها را با \tilde{x} (بخوانید ایکس مدل) نشان می‌دهیم.

مثال ۱-۶. تعداد دانشجویان ده مرکز از دانشگاه پیامنور در استانی معین، را به نزدیکترین صد گرد و حاصل را به صد تقسیم کرده‌ایم. اعداد زیر حاصل شده‌اند:

$$7 \quad 120 \quad 4/5 \quad 61 \quad 7/8 \quad 3/2 \quad 46 \quad 81 \quad 3 \quad 2/5$$

میانه‌ی این مجموعه‌ی داده‌ها را به دست آورید.

حل. تعداد این داده‌ها $n=10$ و بنابراین زوج است. اگر این اعداد را به ترتیب صعودی مرتب کنیم، مجموعه‌ی اعداد زیر را خواهیم داشت:

$$2/5 \quad 7 \quad 7/8 \quad 4/5 \quad 3/2 \quad 3 \quad 46 \quad 61 \quad 81 \quad 120$$

در اینجا دو عدد ۷ و $7/8$ در وسط داده‌ها قرار گرفته‌اند ($k-1=4$ عدد کمتر از ۷ و ۴ عدد بزرگتر از $7/8$ است). در نتیجه میانه‌ی داده‌ها به صورت زیر به دست می‌آید

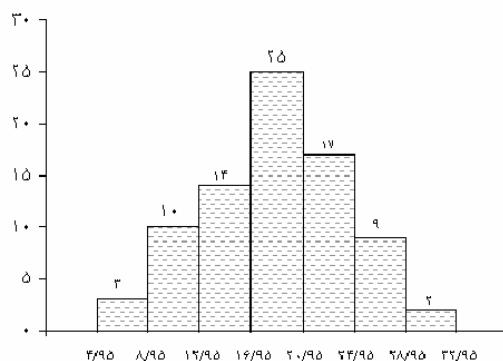
$$\tilde{x} = \frac{7 + 7/8}{4} = 7/4.$$

اگر تعداد داده‌ها زیاد باشد، مرتب کردن داده‌ها به ترتیب صعودی یا نزولی کار چندان آسانی نیست، البته می‌توانیم نمایش ساقه و برگ داده‌ها را تهیه و مرکز داده‌ها را به دست آوریم. اما توجه داریم که معمولاً برای مقاصد دیگر هم، داده‌ها را معمولاً در جدول توزیع فراوانی گروه‌بندی می‌کنیم، بنابر این لازم است روش تعیین میانه‌ی داده‌های جدول‌بندی شده را نیز بدانیم.

وقتی مجموعه‌ای از داده‌ها را در جدول فراوانی خلاصه کرده باشیم، پیدا کردن مقدار دقیق میانه امکان پذیر نیست، زیرا مقادیر واقعی داده‌هایی را که در رده‌ها قرار گرفته‌اند، در اختیار نداریم، اما می‌توان رده‌ای را که میانه در آن قرار گرفته است، پیدا کرد. بنابراین در این حالت میانه را به صورت دیگری تعریف می‌کنیم: میانه‌ی مجموعه‌ای از داده‌های رده‌بندی شده، عددی است که نصف مساحت مستطیل‌های بافت‌نگار در سمت چپ آن و نصف دیگر در سمت راست آن قرار دارد. این تعریف، معادل با این فرض است که مقادیر داده‌هایی که در رده‌ی حاوی میانه قرار دارند، به طور یکنواخت، یعنی با فاصله‌های برابر از هم، توزیع شده‌اند.

برای پیدا کردن خط تقسیم بین دو نیمه‌ی بافت‌نگار (که هر نیمه نماینده‌ی نصف داده‌های رده‌بندی، یعنی $\frac{n}{2}$ است)، باید با شروع از سمت چپ (یا راست) به طرف مرکز بافت‌نگار پیش رویم و به نقطه‌ای برسیم که $\frac{n}{2}$ داده‌ها در چپ (یا راست) آن قرار دارند.

مثال ۱-۷. داده‌های مربوط به میزان اکسیدسولفور مثال ۲-۱ را که بافت‌نگار آن در شکل ۱-۱ رسم شده است، در نظر می‌گیریم. این بافت‌نگار به شکل زیر است:



برای این داده‌ها $n = 80$ و بنابراین $\frac{n}{\mu} = \frac{80}{16/95} = 40$. پس باید مثلاً از سمت چپ (نقطه‌ی شروع) کنیم و فراوانی‌های رده‌ها را با هم جمع کنیم تا به رده‌ای بررسیم که میانه در آن قرار دارد. در این مثال،

$$3+10+14 = 27 < 40 < 3+10+14+25 = 52$$

پس میانه در رده‌ای با کرانه‌های $16/95$ و $20/95$ قراردارد. تعداد $27 = 3+10+14$ تا از داده‌ها در سمت چپ نقطه‌ی $16/95$ قراردارند، طول این رده (مانند همه‌ی رده‌ها در این مثال) ۱۴ است و ۲۵ مورد از داده‌ها در رده‌ی با کرانه‌های $16/95$ و $20/95$ قراردارند. چون طول رده ۱۴ است، پس به هر داده $\frac{14}{25}$ از این طول اختصاص می‌یابد. اما $13 = 40 - 27$ ، پس باید $13 \times \frac{14}{25}$ را به عدد $16/95$ اضافه کنیم تا نقطه‌ای به دست آید که 40 تا از داده‌ها از آن کمترند و 40 مورد دیگر از داده‌ها از آن بیشترند. بنابر این میانه‌ی داده‌های مربوط به این مثال، برابر است با

$$\tilde{x} = 16/95 + 13 \times \frac{14}{25} = 19/03.$$

در حالت کلی، برای محاسبه‌ی مقدار میانه، فراوانی‌های رده‌ها را مثلاً با شروع از طرف مبدأ مختصات با هم جمع می‌کنیم. فرض کنید که

$$f_1 + f_2 + \dots + f_{k-1} < \frac{n}{\mu} \leq f_1 + f_2 + \dots + f_k$$

در این صورت میانه در رده‌ی k قرار دارد. اگر کران پایین این رده را با L_k ، فراوانی آن را با f_k و طول رده را c نشان دهیم و

$$j = \frac{n}{\mu} - (f_1 + f_2 + \dots + f_{k-1})$$

آنگاه

$$\tilde{x} = L_k + \frac{j}{f_k} \times c.$$

مثال ۱-۸-۱ جدول فراوانی زیر، توزیع سنی کارگرانی را که در یک کارخانه‌ی صنعتی به کار اشتغال دارند، نشان می‌دهد. میانه‌ی این توزیع را پیدا کنید.

| سن (بر حسب سال) | فرابانی | تجمعی |
|-----------------|---------|-------|
| ۲۰ - ۲۴ | ۱۲۹ | ۱۲۹ |
| ۲۵ - ۲۹ | ۲۲۱ | ۳۵۰ |
| ۳۰ - ۳۴ | ۳۱۰ | ۶۶۰ |
| ۳۵ - ۳۹ | ۱۶۳ | ۸۲۳ |
| ۴۰ - ۴۴ | ۱۰۵ | ۹۲۸ |
| ۴۵ - ۴۹ | ۶۲ | ۹۹۰ |
| ۵۰ - ۵۴ | ۱۰ | ۱۰۰۰ |
| جمع | ۱۰۰۰ | |

حل. به طوری که ملاحظه می‌کنیم، در این جدول $n=1000$ ، پس $\frac{n}{2} = 500$. با توجه به اینکه

$$129 + 221 = 350 < 500 \leq 350 + 310$$

پس میانه در رده‌ی سوم قراردارد. حد پایین این رده 30 و کران پایین آن $\frac{29}{5} = 5.8$ است (توجه کنید که در فرمول محاسبه‌ی میانه، کران پایین رده را به کار برده‌ایم و نه حد پایین آن را). بنابراین

$$j = 500 - 350 = 150, \quad C = 5, \quad f_{\frac{j}{C}} = 310,$$

$$\tilde{x} = 29/5 + \frac{150}{310} \times 5 = 31/92.$$

۱-۳-۱ مُد

یکی دیگر از شاخص‌های گرایش مرکزی، که «مرکز» مجموعه‌ای از داده‌ها را توصیف می‌کند، مُد است. مُد مجموعه‌ای از داده‌ها، مقداری از داده‌های است که بیشترین فراوانی را داشته باشد. می‌توان مزیت مُد را در این دانست که

۱. نیازی به محاسبه ندارد.

۲. می‌توان را برای داده‌های کیفی نیز تعیین کرد.

مثال ۹-۱. از ۱۰ دانشجو سؤال شده است که به کدام ورزش علاقه دارند و پاسخ‌های زیر دریافت شده است:

فوتبال، فوتبال، والیبال، فوتبال، بسکتبال، شطرنج، فوتبال، والیبال، والیبال، فوتبال.

ملاحظه می‌کنیم که فراوانی فوتبال پنج است و لذا مُدد این مجموعه داده‌ها فوتبال است.

مثال ۱۰-۱. نمره‌ی نهایی ۲۰ دانشجو در درس ریاضی عمومی به شرح زیراست. مدد این داده‌ها را پیدا کنید.

| | | | | | | | | | | | | | | | | | | | | | | |
|------|------|------|------|------|-----|------|------|-----|------|-----|------|------|-----|------|------|-----|-----|------|-----|-----|------|-----|
| ۱۴/۵ | ۱۱/۰ | ۱۵/۰ | ۱۲/۰ | ۱۷/۰ | ۷/۵ | ۱۴/۰ | ۱۲/۰ | ۸/۵ | ۱۷/۰ | ۷/۵ | ۱۴/۰ | ۱۲/۰ | ۸/۰ | ۱۶/۰ | ۱۲/۰ | ۷/۰ | ۵/۵ | ۱۴/۰ | ۹/۰ | ۸/۰ | ۱۲/۰ | ۷/۵ |
|------|------|------|------|------|-----|------|------|-----|------|-----|------|------|-----|------|------|-----|-----|------|-----|-----|------|-----|

با اندک وارسی ملاحظه می‌شود که تعداد دفعات تکرار یا فراوانی ۱۲، عدد ۶ است، لذا مُدد این مجموعه‌ی داده‌ها ۱۲ است.

اگر داده‌ها در جدول فراوانی درج شده باشند، مجدداً باید به تعیین مُدد تقریبی اکتفا کنیم. معمولاً فرمول‌هایی برای محاسبه‌ی مُدد داده‌های رده‌بندی شده ارائه می‌شود. برای سهولت مُدد دسته‌ای از داده‌های رده‌بندی شده را نماینده یا مرکز رده‌ای تعریف می‌کنیم که بیشترین فراوانی را داشته باشد. در نمودارهای بافت‌نگار، مستطیلی که بزرگترین طول را داشته باشد، نماینده‌ی رده‌ی مربوط نمایانگر مُدد خواهد بود.

مثال ۱۱-۱. در مثال ۲-۱، داده‌های مربوط به میزان اکسید سولفور، بیشترین فراوانی ۲۵، مربوط به رده‌ای است که کرانه‌های آن اعداد $16/95$ و $20/95$ است، پس مدد این مجموعه‌ی داده‌ها عدد $x_{m0} = 18/95$ ، یعنی مرکز همین رده است.

۱-۳-۴ دیگر شاخص‌های گرایش مرکزی

میانگین، میانه، و مد هریک به گونه‌ای «مرکز» توزیع را نمایان می‌کنند. با این حال برای برخی داده‌ها به انواع دیگر اندازه‌های گرایش مرکزی نیاز داریم. در زیر دو مورد از این اندازه‌ها را مورد بحث قرار می‌دهیم و مثال‌هایی از کاربرد آنها ارائه می‌کنیم.

میانگین همساز مجموعه‌ای از داده‌ها، معکوس میانگین حسابی معکوس‌های داده‌است. به عبارت دیگر اگر داده‌ها را با x_1, x_2, \dots, x_n و میانگین همساز را با \bar{x}_H نشان دهیم، آنگاه

$$\frac{1}{\bar{x}_H} = \frac{1}{n} \sum_{i=1}^n \frac{1}{x_i},$$

یا

$$\bar{x}_H = \frac{n}{\sum_{i=1}^n \frac{1}{x_i}}.$$

واضح است شرط $x_i \neq 0$ به ازای $i = 1, 2, \dots, n$ ، الزامی است. برای داده‌های ردیبندی شده، فرمول محاسبه‌ی میانگین همساز به صورت زیر در می‌آید:

$$\bar{x}_H = \frac{n}{\sum_{i=1}^n \frac{1}{m_i} f_i}$$

که در آن m_i مرکز ردیه i ام و f_i فراوانی این ردیه است. وقتی مجموعه‌ی داده‌ها دارای واحدهای اندازه‌گیری ترکیبی است مثلاً کیلومتر در ساعت، از میانگین همساز استفاده می‌شود. وقتی مجموعه‌ی داده‌ها مربوط به آهنگ تغییر باشد، میانگین همساز مفیدتر از میانگین حسابی است.

مثال ۱۲-۱. سرعت خودرویی در ساعت اول ۴۸ کیلومتر در ساعت و ساعت دوم ۸۰

کیلومتر در ساعت است، در این صورت میانگین همساز $\frac{48 + 80}{2} = 64$ محاسبه می‌شود.

حال، همین محاسبات را با دستورات نرم‌افزار مرور می‌کنیم.

```

MTB > SET C8
DATA> 48 80
DATA> END
MTB > LET C9=1/C8
MTB > MEAN C9 K4
Column Mean
  Mean of C9 = 0.016667
MTB > LET K5=1/K4
MTB > PRINT K5
Data Display
K5 60.0000

```

مثال ۱۳-۱. سه کارگر به ترتیب دیواری به بلندی یک متر را در $\frac{1}{3}$ ، $\frac{1}{3}$ ، و $\frac{1}{4}$ ساعت می‌سازند. اگر این سه کارگر با هم کارکنند چقدر طول می‌کشد تا یک متر دیوار بنا شود.

حل. پاسخ این سؤال عبارت است $\bar{x}_H = \frac{\frac{1}{3} + \frac{1}{3} + \frac{1}{4}}{3} = \frac{13}{12}$. توجه دارید میانگین نشان دهندهٔ متوسط سرعت بالا بردن یک متر دیوار برای یک نفر است.

میانگین هندسی n عدد مثبت، ریشهٔ n ام حاصل ضرب آنهاست. به عبارت دیگر اگر اعدادهای مثبت را با x_1, x_2, \dots, x_n و میانگین هندسی را با \bar{x}_G نشان دهیم، آنگاه

$$\bar{x}_G = (x_1 x_2 \cdots x_n)^{\frac{1}{n}} = \sqrt[n]{x_1 x_2 \cdots x_n}.$$

با گرفتن لگاریتم از طرفین رابطهٔ اخیر داریم

$$\begin{aligned} \log(\bar{x}_G) &= \frac{1}{n}(\log(x_1) + \log(x_2) + \cdots + \log(x_n)) \\ &= \frac{1}{n} \sum_{i=1}^n \log(x_i). \end{aligned}$$

پس میانگین هندسی، آنتی لگاریتم میانگین حسابی لگاریتم‌های داده‌هاست.

اگر داده‌ها در یک جدول فراوانی تلخیص شده باشند، برای محاسبهٔ میانگین

هندسی، تعریف زیر را به کار می‌بریم:

$$\log(\bar{x}_G) = \frac{1}{n} \sum_{i=1}^k f_i \log(m_i)$$

که در آن m_i مرکز رده‌ی i و f_i فراوانی این رده است.
میانگین هندسی در پیدا کردن متوسط آهنگ‌های تغییر، مثلاً در جمعیت و غیره یا محاسبه‌ی میانگین نسبت تغییرات مانند افزایش قیمت‌ها و غیره، مفید است.

مثال ۱۴-۱. فرض کنید که جمعیت یک روستا ۱۰۰۰ و آهنگ سالانه‌ی رشد جمعیت در طول چهار سال ۸% باشد. جمعیت این روستا در پایان این چهار سال چقدر خواهد بود؟

حل. جمعیت این روستا در شروع و پایان هریک از این چهار سال، به ترتیب عبارت خواهد بود از

$$1000, \quad 1000 \times (1/08)^1, \quad 1000 \times (1/08)^2, \quad 1000 \times (1/08)^3, \quad 1000 \times (1/08)^4$$

میانگین هندسی این پنج عدد، ریشه‌ی پنجم حاصل ضرب این اعداد است، پس

$$\begin{aligned} \bar{x}_G &= \left(1000 \times 1000 \times (1/08)^1 \times 1000 \times (1/08)^2 \times 1000 \times (1/08)^3 \times 1000 \times (1/08)^4 \right)^{\frac{1}{5}} \\ &= \left((1000)^5 \times (1/08)^{1+2+3+4} \right)^{\frac{1}{5}} \\ &= 1000 \times (1/08)^{\frac{5}{5}} = 1166 \end{aligned}$$

محاسبه‌ی میانگین هندسی با مینیتپ نیازی به عملیات فوق ندارد و صرفاً با استفاده از فرمول مربوطه قابل انجام است.

```
MTB > SET C10
DATA> 0:4
DATA> END
MTB > LET C11=1000*((1.08)**C10)
MTB > PARPROD C11 C12
MTB > LET K6=C12(5)**(1/5)
MTB > PRINT K6
Data Display
K6      1166.40
```

نکته. در مشاهدات مثبت، رابطه‌ای بین میانگین‌های حسابی، هندسی، و همساز به شکل زیر برقرار است

$$\bar{x} \geq \bar{x}_G \geq \bar{x}_H.$$

۲-۳-۱ شاخص‌های پراکندگی

یکی از مهم‌ترین مشخصه‌های تقریباً هر مجموعه‌ای از داده‌ها، آن است که مقادیر داده‌ها همانند نیستند. در واقع، میزان ناهمانند بودن، یا تغییرات بین داده‌ها اهمیتی اساسی در آمار دارد. شاخص‌های گرایش مرکزی، جنبه‌ای مهم از مجموعه‌ی داده‌ها را توصیف می‌کنند که همان «معدل» یا «وسط» داده‌هاست اما اطلاعی درباره‌ی این مشخصه اساسی دیگر به ما نمی‌دهند. بنابراین، می‌خواهیم راهی پیدا کنیم که این پراکندگی در داده‌ها، یا تفرق آنها را نشان دهد.

مثال ۱۵-۱. نمرات دو دانشجو رشته علوم کامپیوتر، در جدول زیر درج شده است.

| | آمار و احتمال | اصول کامپیوتر | فیزیک | ریاضی عمومی |
|----------|---------------|---------------|-------|-------------|
| A دانشجو | ۱۴ | ۱۲ | ۱۳ | ۱۳ |
| B دانشجو | ۱۳ | ۸ | ۱۳ | ۱۸ |

مالحظه می‌کنیم که میانگین دو دانشجو در این چهار درس ۱۳ است درحالی که دانشجوی A در همه دروس نمره‌ی قبولی دارد، ولی دانشجوی B در درس ریاضی عمومی قوی و درس اصول کامپیوتر ضعیف است. بنابراین باز هم نیاز به وجود شاخصی که اختلاف موجود بین این دو مجموعه داده‌های متفاوت ولی با میانگین یکسان را نشان دهد، آشکار می‌شود. توجه کنید که علاوه بر میانگین، مدل، و میانه‌ی نمرات هر دانشجو ۱۳ است.

مثال ۱۵-۱. نشان می‌دهد که مفهوم پراکندگی نقشی مهم در تحلیل آماری دارد. در زیر اندازه‌هایی برای پراکندگی ارائه می‌کنیم.

۱-۲-۳-۱ دامنه

ساده‌ترین معیار سنجش پراکندگی، دامنه است که با پیدا کردن تفاضل بزرگ‌ترین مقدار داده‌ها از کوچک‌ترین مقدار داده‌ها به دست می‌آید.

در مثال ۱۵-۱، دامنه نمره‌های دانشجوی A ، $۱۲-۱۶=۲$ و دامنه نمره‌های دانشجوی B ، $۸-۱۸=۱۰$ است. به طوری که ملاحظه می‌کنیم «پراکندگی» نمرات دانشجوی B بیشتر از دانشجوی A ، است. با آنکه دامنه به سادگی به دست می‌آید، عیب عمده‌ی آن این است که تنها داده‌های فرین، یعنی بیشترین و کمترین مقدار داده‌ها را به حساب می‌آورد و کاری به مقادیر بینایینی داده‌ها ندارد. همچنین در برخی مواقع وقتی اندازه‌ی نمونه بزرگ باشد مقدار دامنه افزایش می‌یابد. با این حال، وقتی که تعداد داده‌ها کم است، دامنه معياری مناسب برای پراکندگی است.

برای داده‌های مثال ۱-۱ به کمک نرم‌افزار دامنه را به دست می‌آوریم.

```
MTB > LET K7=MAX(C1)-MIN(C1)
MTB > PRINT K7
Data Display
K7    25.6000
```

۲-۲-۳-۱ انحراف معيار

برای معرفی انحراف معيار که مفیدترین معیار سنجش پراکندگی است، می‌توانیم میزان دوری یا نزدیکی داده‌ها را از میانگین به عنوان ملاکی در نظر گیریم. برای توضیح مطلب، فرض کنید که داده‌های موردنظر را با مقادیر x_1, x_2, \dots, x_n و میانگین آنها را با \bar{x} نشان دهیم. اگر انحراف هر داده را از میانگین، یعنی مقادیر $x_1 - \bar{x}, x_2 - \bar{x}, \dots, x_n - \bar{x}$ را حساب کنیم، به نظر می‌رسد که همین انحراف‌ها از میانگین می‌توانند به عنوان معیاری برای سنجش پراکندگی داده‌ها به کار روند. به عبارت دیگر هر قدر که تفاوت این انحراف‌ها از هم بیشتر باشد، داده‌ها پراکنده‌ترند و هر قدر که این انحراف‌ها تفاوت کمتری باهم داشته باشند، داده‌ها متتمرکز‌ترند. چون مقایسه این n انحراف با هم کار دشواری است، برای سهولت در وله‌ی اوّل سعی می‌کنیم میانگین این n انحراف را به دست آوریم، یعنی

$$\frac{x_1 - \bar{x} + x_2 - \bar{x} + \dots + x_n - \bar{x}}{n}$$

را محاسبه می‌کنیم. به سادگی ملاحظه می‌شود که بنابر تعریف میانگین، صورت این عبارت برابر صفر است و لذا میانگین انحراف‌ها از میانگین صفر است. پس می‌توان نتیجه گرفت: داده‌ها مجموعاً همان‌قدر که در جهت منفی از \bar{x} انحراف دارند مجموعاً همان‌قدر هم در جهت مثبت از \bar{x} انحراف دارند و لذا مجموع انحراف‌ها صفر است. برای رفع این مشکل، باید توجه کرد که علاقه‌ی ما به میزان دوری داده‌ها از میانگین است و نه جهت دوری داده‌ها از آن. به عبارت دیگر می‌توانیم علامت جبری انحراف‌ها از میانگین را نادیده بگیریم و مثلاً

$$\frac{|x_1 - \bar{x}| + |x_2 - \bar{x}| + \dots + |x_n - \bar{x}|}{n}$$

را محاسبه کنیم. این میانگین قدر مطلق انحراف‌ها که به انحراف میانگین موسوم است، کاربردهایی در آمار دارد، اما چون قدر مطلق در این تعریف مشکلاتی را ایجاد می‌کند از در نظر گرفتن آن صرف نظر کرده و سعی می‌کنیم مشکل علامت‌ها را در مقادیر انحراف از میانگین به گونه‌ای دیگر حل کنیم. برای این منظور، مقادیر انحراف از میانگین را به توان دوام می‌رسانیم، یعنی مقادیر

$$(x_1 - \bar{x})^r, (x_2 - \bar{x})^r, \dots, (x_n - \bar{x})^r$$

را محاسبه می‌کنیم. حال برای محاسبه میانگین مجدد انحراف‌ها از میانگین، باید مجموع آنها را محاسبه و بر n ، یعنی تعداد داده‌ها، تقسیم کنیم. بدلاً از که در اینجا نیازی به ذکر نیست، مجموع را به جای تقسیم بر n بر $n-1$ تقسیم می‌کنیم. مقدار حاصل، یعنی

$$\frac{(x_1 - \bar{x})^r + (x_2 - \bar{x})^r + \dots + (x_n - \bar{x})^r}{n-1}$$

را واریانس نمونه‌ای یا به اختصار واریانس می‌نامیم و آن را با s^2 نشان می‌دهیم. پس با استفاده از نماد مجموع یابی Σ داریم

$$s^2 = \frac{\sum_{i=1}^n (x_i - \bar{x})^r}{n-1}.$$

چون واریانس بر حسب مجدد واحدهای که داده‌ها بر حسب آن بیان شده‌اند، محاسبه می‌شود، برای آنکه از واحد اصلی استفاده کنیم، جذر مثبت واریانس را محاسبه و آن را با s نشان می‌دهیم و انحراف معیار نمونه‌ای یا صرفاً انحراف معیار می‌نامیم. پس

$$s = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n-1}}.$$

مثال ۱۶-۱. در مثال ۱۵-۱، ملاحظه کردیم که میانگین نمرات هر دو دانشجو ۱۳ است. مقدار انحراف معیار برای دانشجوی A برابر است با

$$s_1 = \sqrt{\frac{(14-13)^2 + (12-13)^2 + (13-13)^2 + (13-13)^2}{4-1}} = \sqrt{\frac{2}{3}} = 0.816.$$

همچنین انحراف معیار برای دانشجوی B برابر است با:

$$s_2 = \sqrt{\frac{(13-13)^2 + (8-13)^2 + (13-13)^2 + (18-13)^2}{4-1}} = \sqrt{\frac{50}{3}} = 4.082.$$

ملاحظه می‌شود که انحراف معیار نمرات دانشجوی B به مرتب بیشتر از انحراف معیار دانشجوی A است و لذا این معیار پراکنده‌تر بودن نمرات دانشجوی B را در مقایسه با نمرات دانشجوی A به خوبی نشان می‌دهد.

از لحاظ محاسباتی، پیدا کردن انحراف هر داده از میانگین داده‌ها و پیدا کردن مجموع مجددهای انحراف‌ها برای محاسبه انحراف معیار شاید کمی مشکل باشد. به جای آن می‌توان تعریف میانگین را به کار برد و با اندکی اعمال جبری، فرمول انحراف معیار را به صورت معادل زیر در آورد.

$$s = \sqrt{\frac{n \sum_{i=1}^n x_i^2 - (\sum_{i=1}^n x_i)^2}{n(n-1)}}.$$

مثال ۱۷-۱. برای نرم‌افزاری، طول مدت اجرای پنج محاسبه‌ی آماری متفاوت بر حسب یک دهم ثانیه، به شرح زیر است

۰/۹۸ ۰/۹۴ ۱/۹۶ ۱/۴۸ ۲/۰۵

میانگین و انحراف معیار این مجموعه داده‌ها را محاسبه کنید.

حل. بنابر تعریف، داریم

$$\bar{x} = \frac{۰/۹۸ + ۰/۹۴ + ۱/۹۶ + ۱/۴۸ + ۲/۰۵}{۵} = ۱/۴۸۲$$

و چون

$$\sum_{i=1}^5 x_i^2 = (۰/۹۸)^2 + (۰/۹۴)^2 + (۱/۹۶)^2 + (۱/۴۸)^2 + (۲/۰۵)^2 = ۱۲/۰۷۸۵$$

و

$$(\sum_{i=1}^5 x_i)^2 = (۰/۹۸ + ۰/۹۴ + ۱/۹۶ + ۱/۴۸ + ۲/۰۵)^2 = ۵۴/۹۰۸۱$$

پس

$$s = \sqrt{\frac{۵ \times (۱۲/۰۷۸۵) - (۵۴/۹۰۸۱)}{۵ \times ۴}} = ۰/۰۵۲۴$$

با نرم‌افزار مینیتب مقدار انحراف معیار را به دست می‌آوریم.

MTB > SET C13

DATA> 0.98 0.94 1.96 1.48 2.05

DATA> END

MTB > STD C13 K8

Column Standard Deviation

Standard deviation of C13 = 0.52366

اگر داده‌ها در جدول فراوانی رده‌بندی شده باشند، در این صورت با قراردادن

مرکز هر رده به جای داده‌های واقع در آن رده و در نظر گرفتن فراوانی هر رده، فرمول

انحراف‌معیار برای داده‌های رده‌بندی شده به صورت زیر است

$$s = \sqrt{\frac{\sum_{i=1}^k (m_i - \bar{x})^2 f_i}{n-1}}.$$

یا اگر بخواهیم محاسبات را ساده‌تر کنیم، می‌توانیم فرمول معادل زیر را به کار ببریم

$$s = \sqrt{\frac{n \left(\sum_{i=1}^k m_i f_i \right) - \left(\sum_{i=1}^k m_i f_i \right)^2}{n(n-1)}}.$$

مثال ۱۸-۱. می خواهیم واریانس و انحراف معیار داده های مربوط به میزان اکسید سولفور منتشره در هوا را که در مثال ۲-۱ مطرح شده است، محاسبه کنیم. برای این منظور در جدول فراوانی داده های مذکور، ستون جدیدی اضافه می کنیم و در آن مقادیر $m_i^2 f_i$ را درج می کنیم، جدول به صورت زیر خواهد بود.

| کرانه های رددها | m_i | f_i | $f_i m_i$ | $m_i^2 f_i$ |
|-----------------|-------|-------|-----------|-------------|
| ۱۶/۹۵ - ۸/۹۵ | ۶/۹۵ | ۳ | ۲۰/۸۵ | ۱۴۴/۹۰۷۵ |
| ۸/۹۵ - ۱۲/۹۵ | ۱۰/۹۵ | ۱۰ | ۱۰۹/۵۰ | ۱۱۹۹/۰۲۵۰ |
| ۱۲/۹۵ - ۱۶/۹۵ | ۱۴/۹۵ | ۱۴ | ۲۰۹/۳۰ | ۳۱۲۹/۰۳۵۰ |
| ۱۶/۹۵ - ۲۰/۹۵ | ۱۸/۹۵ | ۲۵ | ۴۷۳/۷۵ | ۸۹۷۷/۵۶۲۵ |
| ۲۰/۹۵ - ۲۴/۹۵ | ۲۲/۹۵ | ۱۷ | ۳۹۰/۱۵ | ۸۹۵۳/۹۴۲۵ |
| ۲۴/۹۵ - ۲۸/۹۵ | ۲۶/۹۵ | ۹ | ۲۴۲/۵۵ | ۶۵۳۶/۷۲۲۵ |
| ۲۸/۹۵ - ۳۲/۹۵ | ۳۰/۹۵ | ۲ | ۶۱/۹۰ | ۱۹۱۵/۸۰۵۰ |
| مجموع | | ۸۰ | ۱۵۰۸ | ۳۰۸۵۷/۰۰۰۰ |

جدول ۱۸-۱ جدول فراوانی با دو ستون اضافی برای محاسبه میانگین و واریانس

بنابراین

$$s = \sqrt{\frac{۸۰ \times (۳۰۸۵۷) - (۱۵۰۸)^2}{۸۰ \times ۷۹}} = ۵/۵۵.$$

به طوری که از مقادیر واقع در آخرین ستون جدول فراوانی بالا آشکار می شود، محاسبات مربوط به انحراف معیار، زمانی که نماینده های داده ها ارقام اعشاری و فراوانی ها نسبتاً بزرگ باشند، کار خسته کننده ای است. برای ساده کردن محاسبات، می توان از روش کدگذاری استفاده کرد، به شرط اینکه طول رددها برابر باشند (که در

مثال بالا چنین است). در روش کدگذاری، نماینده‌ی یکی از رده‌ها (ترجیحاً رده‌ی وسطی یا یکی از رده‌های واقع در اواسط جدول) را ۰ اختیار می‌کنیم. سپس نماینده‌های رده‌های بالای این رده را اعداد صحیح منفی، یعنی $-1, -2, \dots$ و نماینده‌های رده‌های واقع در زیر این رده را اعداد صحیح مثبت، یعنی $1, 2, \dots$ اختیار می‌کنیم. اگر این نماینده‌های جدید رده‌ها را با u_1, u_2, \dots, u_n نشان دهیم، فرمول میانگین به صورت

$$\bar{x} = m_{\circ} + \frac{\sum_{i=1}^k u_i f_i}{n} c$$

درمی‌آید. که در آن m_{\circ} نماینده رده‌ای است که در کدگذاری نماینده‌ی جدید ۰ را برای آن در نظر گرفته‌ایم و c طول مشترک رده‌هاست. همچنین فرمول انحراف معیار به صورت زیر است.

$$s = c \sqrt{\frac{n \left(\sum_{i=1}^k u_i^2 f_i \right) - \left(\sum_{i=1}^k u_i f_i \right)^2}{n(n-1)}}$$

نحوه محاسبه میانگین و انحراف معیار را به کمک روش کدگذاری، در مثال زیر تشریح می‌کنیم.

مثال ۱۹-۱. مطابق داده‌های مثال ۱۸-۱ به روش کدگذاری میانگین و انحراف معیار را به دست آورید.

حل. چون تعداد رده‌ها ۷ است، نماینده‌ی جدید رده‌ی چهارم را ۰ می‌گیریم و به ترتیبی که در بالا گفتیم نماینده‌های جدید رده‌های دیگر را به دست می‌آوریم. حال جدول زیر را تشکیل می‌دهیم

| m_i | u_i | f_i | $f_i u_i$ | $u_i^r f_i$ |
|-------|-------|-------|-----------|-------------|
| ۶/۹۵ | -۳ | ۳ | -۹ | ۲۷ |
| ۱۰/۹۵ | -۲ | ۱۰ | -۲۰ | ۴۰ |
| ۱۴/۹۵ | -۱ | ۱۴ | -۱۴ | ۱۴ |
| ۱۸/۹۵ | ۰ | ۲۵ | ۰ | ۰ |
| ۲۲/۹۵ | ۱ | ۱۷ | ۱۷ | ۱۷ |
| ۲۶/۹۵ | ۲ | ۹ | ۱۸ | ۳۶ |
| ۳۰/۹۵ | ۳ | ۲ | ۶ | ۱۸ |
| مجموع | | ۸۰ | -۲ | ۱۵۲ |

جدول ۴-۱ جدول فراوانی داده‌های مثال ۲-۱ برای

محاسبه میانگین و واریانس به روش کدگذاری

با قراردادن $c = ۴$ ، $n = ۸۰$ ، $m_o = ۱۸/۹۵$ در فرمول میانگین به روش کدگذاری داریم:

$$\bar{x} = m_o + \frac{\sum_{i=1}^k u_i f_i}{n} = ۱۸/۹۵ + \frac{-۲}{۸۰} \times ۴ = ۱۸/۸۵.$$

همچنین با قراردادن همین مقادیر همراه با مقدار $\sum_{i=1}^k u_i^r f_i = ۱۵۲$ در فرمول انحراف

معیار به روش کدگذاری خواهیم داشت:

$$s = \sqrt{\frac{۸۰(۱۵۲) - (-۲)^2}{۸۰ \times ۷۹}} = ۵/۵۵.$$

به طوری که ملاحظه می‌کنیم این مقادیر، مطابق انتظار، همان مقادیری هستند که از داده‌های اصلی به دست آمده‌اند.

برخی کاربردهای انحراف معیار را در زیر شرح می‌دهیم.

۱. ضریب تغییر

روش کدگذاری حالت خاصی از تبدیل متغیرها است. اگر مشاهدات در عددی ثابت ضرب و با عددی ثابت جمع کنیم مشاهدات جدیدی به دست می‌آیند.

$$y_i = ax_i + b, \quad i = 1, 2, \dots, n$$

از روی خواص میانگین و واریانس، تأثیر این تبدیل بر روی آنها به شکل زیر می‌شود

$$\bar{y} = a\bar{x} + b, \quad s_y^2 = a^2 s_x^2.$$

میانگین و انحراف معیار دارای واحد سنجش، از نوع واحد سنجش داده‌هاست.

برای مثال ممکن است وزن اشیای مشخصی دارای انحراف معیار $1/0$ اونس یا 2835 میلی گرم باشد که هر دو یکسان هستند، ولی هیچ‌کدام از آنها به راحتی مشخص نمی‌کنند که با یک پراکندگی بزرگ یا یک پراکندگی کوچک سروکار داریم. در این گونه موارد آنچه که مورد نیاز است ارائه‌ی یک مقیاس نسبی پراکندگی است. چنین مقیاسی عبارت است از ضریب تغییر که از تقسیم انحراف معیار بر میانگین حاصل می‌گردد. معمولاً در عدد 100 ضرب شده و به صورت درصد بیان و با حرف V نشان داده می‌شود:

$$V = \frac{s}{\bar{x}} \times 100$$

برای اطلاعات مثال $17-1$ ، مقدار ضریب تغییر برابر است با:

$$V = \frac{5/55}{18/85} \times 100 = 29/4$$

۲. نمره‌ی استاندارد

چنانچه تبدیل خاصی روی x_1, x_2, \dots, x_n اعمال شود:

$$z_i = \frac{x_i - \bar{x}}{s}, \quad i = 1, 2, \dots, n$$

داده‌های z_1, z_2, \dots, z_n را استاندارد شده می‌گوییم و دارای ویژگی میانگین صفر و واریانس یک، $\bar{z} = 0$ و $s^2 = 1$ هستند. هرگاه بخواهیم دو مشاهده از دو نمونه‌ی مختلف را با هم مقایسه کنیم، استاندارد شده آنها کار را آسان می‌کند. مثلاً، میانگین و انحراف معیار نمرات پایان ترم درس زبان خارجه به ترتیب 60 و 16 و میانگین انحراف معیار نمرات پایان ترم درس ریاضی عمومی به ترتیب 58 و 10 است. اگر دانشجویی هر دو درس را امتحان داده باشد و نمره‌ی زبان خارجه 72 و نمره‌ی ریاضی عمومی 68 باشد، وضعیت تحصیلی اش در کدام درس بهتر است؟

$$z_1 = \frac{72 - 60}{16} = 0 / 75$$

$$z_2 = \frac{68 - 58}{10} = 1$$

ملاحظه می‌شود که اگرچه نمره‌ی درس زبان خارجه از نمره‌ی درس ریاضی عمومی او بیشتر است، ولی وضعیت وی در کلاس ریاضی عمومی بهتر از کلاس زبان خارجه است. زیرا در درس ریاضی عمومی به اندازه‌ی یک انحراف معیار بالای میانگین، $58 + 10 = 68$ ، قرار دارد در حالی که در درس زبان خارجه به اندازه‌ی $0 / 75$ انحراف معیار در بالای میانگین، $60 + 0 = 60$ ، قرار دارد.

۳. قضیه‌ی چبیشف

چنانچه پراکندگی دو نمونه از یک جمعیت را با یکدیگر مقایسه کنیم، می‌گوییم نمونه‌ای که دارای انحراف معیار بزرگتری است، پراکندگی‌اش بیشتر است. همچنین عنوان کردیم که هرچه داده‌ها به میانگین نزدیک‌تر باشند و یا پیرامون میانگین تجمع کرده باشند، انحراف معیار به صفر نزدیک است و چنانچه داده‌ها از میانگین دور باشند، آنگاه انحراف معیار بزرگ است. این موضوع به گونه‌ای علمی توسط ریاضی‌دان روسی به نام چبیشف با عنوان قضیه‌ی چبیشف بیان شده است.

قضیه‌ی ۱-۱. برای هر مجموعه از داده‌ها و هر عدد ثابت $k \geq 1$ ، حداقل $(1 - \frac{1}{k}) \times 100$ درصد داده‌ها باید به فاصله‌ی k برابر انحراف معیار در هر دو طرف میانگین قرار گیرند.

قضیه‌ی چبیشف برای حالت‌های خاص k ، چنین بیان می‌شود:

الف) ممکن تعداد خیلی کم داده‌ها به بازه‌ی $(s - s, \bar{x} + s)$ تعلق داشته باشند.

ب) حداقل 75 درصد، $= 100 \times \frac{3}{4}$ داده‌ها به بازه‌ی $(\bar{x} - 2s, \bar{x} + 2s)$ تعلق خواهند داشت.

ج) حداقل $88/9$ درصد، $= 100 \times \frac{8}{9}$ داده‌ها به بازه‌ی $(\bar{x} - 3s, \bar{x} + 3s)$ تعلق خواهند داشت.

مثلاً، در دستگاه اتوماتیک بسته‌بندی زعفران به گونه‌ای تنظیم شده است که درون هر بسته به طور متوسط 16 گرم زعفران و با انحراف معیار 0.5 گرم، زعفران قرار می‌دهد. حداقل درصد بسته‌های تولیدی اش که دارای وزنی بین $15/95$ تا $16/05$ گرم باشند 84 درصد است (توجه دارید که در این مثال $k = 2/5$ در نظر گرفته شده است).

قضیه‌ی چبیشف برای هر نوع داده‌ای قابل کاربرد است، اما چنین بیان می‌شود که حداقل چند درصد داده‌ها به فاصله‌ی معینی تعلق دارند و مقدار دقیق را تعیین نمی‌کند. برای بیشتر داده‌ها، درصد واقعی داده‌ای که به فاصله‌ی معینی تعلق می‌گیرند، بیش از آنچه که قضیه‌ی چبیشف بیان می‌کند، می‌باشد.

برای بیشتر داده‌ها راه کار دقیق‌تری از قضیه‌ی چبیشف با عنوان دستور تجربی بیان شده است. در حقیقت یک نتیجه‌ی نظری است که براساس توزیع نرمال (فصل ششم) بیان می‌شود.

دستور تجربی: برای داده‌ای که منحنی فراوانی آنها زنگی شکل است، داریم:

الف) تقریباً 68 درصد داده‌ها به بازه‌ی $(\bar{x} - s, \bar{x} + s)$ تعلق دارند.

ب) تقریباً 95 درصد داده‌ها به بازه‌ی $(\bar{x} - 2s, \bar{x} + 2s)$ تعلق دارند.

ج) تقریباً $99/7$ درصد (تقریباً تمام داده‌ها) داده‌ها بازه‌ی $(\bar{x} - 3s, \bar{x} + 3s)$ تعلق دارند. با اطلاعات مثال ۱-۱ و شکل ۱-۱ که تقارن تقریبی در نمودار به چشم می‌خورد، می‌توان وضعیت توزیع مشاهدات را در بازه‌های مطرح شده برآورد نمود.

الف) تقریباً 68 درصد داده‌ها به بازه‌ی

$$(18/85 - 5/55, 18/85 + 5/55) = (3/3, 24/4)$$

تعلق دارند.

ب) تقریباً 95 درصد داده‌ها به بازه‌ی

$$(18/85 - 2 \times 5/55, 18/85 + 2 \times 5/55) = (-2/25, 29/95) = (0, 29/95)$$

تعلق دارند.

ج) تقریباً ۷/۹۹ درصد (تقریباً کل) داده‌ها بازه‌ی

$$(18/85 - 3 \times 5/55, 18/85 + 3 \times 5/55) = (0, 35/5)$$

تعلق دارند.

۱-۳-۲-۳ چندک‌ها و دیگر شاخص‌های پراکندگی

به‌خاطر می‌آوریم که میانه‌ی مجموعه‌ای از داده‌ها مقداری بود که نصف داده‌ها در یک طرف و نصف دیگر در طرف دیگر آن قرار داشتند. بهمین قیاس می‌توان چندک‌ها را تعریف کرد که عبارت از مقادیری هستند که بخش معینی از داده‌ها در یک طرف آن و بخش باقی‌مانده در طرف دیگر آن قرار دارند. از جمله‌ی چندک‌های مهم می‌توان از چارک‌ها، دهک‌ها، و صدک‌ها نام برد. این مفاهیم را تنها برای داده‌های رده‌بندی شده تعریف می‌کنیم.

چارک اول نقطه‌ای بر روی محور طول‌هاست که در آن نقطه، مساحت زیر بافت‌نگار به نسبت‌های ۱ و ۳ تقسیم می‌شود، یعنی $\frac{1}{\mu}$ مساحت زیر بافت‌نگار در سمت چپ این نقطه و $\frac{3}{\mu}$ سطح باقی‌مانده در طرف دیگر این نقطه قرار دارد. چارک دوم همان میانه است و چارک سوم نقطه‌ای است که $\frac{3}{\mu}$ مساحت زیر بافت‌نگار در سمت چپ آن و $\frac{1}{\mu}$ در سمت راست آن قرار دارد. چارک اول را با Q_1 و چارک دوم (یا همان میانه) را با Q_2 و چارک سوم را با Q_3 نشان می‌دهیم.

برای محاسبه‌ی Q_1 ، ابتدا $\frac{n}{\mu}$ را محاسبه می‌کنیم و سپس فراوانی‌ها را با شروع از

سمت چپ با هم جمع می‌کنیم. فرض کنید که

$$f_1 + f_2 + \dots + f_{k-1} < \frac{n}{\mu} \leq f_1 + f_2 + \dots + f_k.$$

در این صورت چارک اول در رده‌ی k است. اگر کران پایین این رده را با L'_k نشان

دهیم و

$$m = \frac{n}{\mu} - (f_1 + \dots + f_{k-1})$$

آنگاه

$$Q_1 = L'_k + \frac{m}{f_k} \times c$$

که در آن c طول مشترک رده‌ها و f_k فراوانی رده‌ی k است که چارک اوّل در آن قرار دارد.

عیناً به همین منوال می‌توانیم چارک سوم، Q_3 ، را نیز محاسبه کنیم. در این حالت $\frac{3n}{\mu}$ را محاسبه و با شروع از سمت چپ، فراوانی‌های رده‌ها را با هم جمع می‌کنیم. اگر k رده‌ای باشد به‌طوری که

$$f_1 + f_2 + \dots + f_{k-1} < \frac{3n}{\mu} \leq f_1 + f_2 + \dots + f_k$$

آنگاه Q_3 در رده‌ی k است. اگر کران پایین این رده را با k نشان دهیم و

$$l = \frac{3n}{\mu} - (f_1 + \dots + f_{k-1})$$

آنگاه

$$Q_3 = L''_k + \frac{l}{f_k} \times c$$

یکی دیگر از شاخص‌های پراکندگی، برد میان چارکی است که به صورت $Q_3 - Q_1$ تعریف می‌شود. برد میان چارکی طول بازه‌ای است که پنجاه درصد داده‌ها در آن قرار دارند. برخی پژوهشگران از نیم برد میان چارکی استفاده می‌کنند که به صورت

$$\frac{Q_3 - Q_1}{\mu} \text{ تعریف می‌شود، و انحراف چارکی نیز نامیده می‌شود. یک اندازه‌ی}$$

پراکندگی نسبی، که ضریب تغییر چارکی نام دارد به صورت نسبت نیم برد میان چارکی بر میانگین Q_1 و Q_3 ضرب در 100 ، یعنی به صورت $\frac{Q_3 - Q_1}{Q_3 + Q_1} \times 100$ ، تعریف می‌شود.

به همین ترتیب می‌توان دهک‌های اول، دوم، ...، نهم را تعریف کرد. این دهک‌ها به ترتیب با D_1, D_2, \dots, D_9 نشان داده می‌شوند. مثلاً برای محاسبه‌ی دهک اول، D_1

کافی است $\frac{n}{10}$ را حساب کنیم و k را طوری پیدا کنیم که

$$f_1 + f_2 + \dots + f_{k-1} < \frac{n}{10} \leq f_1 + f_2 + \dots + f_k.$$

در این صورت دهک اول در رده‌ی k است. اگر کران پایین این رده را با L_k نشان دهیم و

$$h = \frac{n}{10} - (f_1 + \dots + f_{k-1})$$

آنگاه

$$D_1 = L_k + \frac{h}{f_k} \times c$$

که مجدداً c طول مشترک رده‌ها و f_k فراوانی رده‌ی k است.

می‌توان سایر دهک‌ها را به همین ترتیب محاسبه و مثلاً $D_2 = D_1 - D_1$ یا $D_3 = D_2 - D_1$... را به عنوان معیارهای دیگری برای سنجش پراکندگی به کار برد.

صدق‌ها عیناً به نحو مشابه تعریف می‌شوند که از ذکر جزئیات مربوط به آنها خودداری می‌کنیم.

مثال ۲۰-۱. مطابق مثال ۱-۲، مربوط به میزان اکسید سولفور، نمودار بافت‌نگار آن داده‌ها را در نظر بگیرید. مقادیر چارک اول، سوم، ضریب تغییر چارکی، دهک اول و دهک نهم را به دست آورید.

حل. چون $n = 80$ ، پس $\frac{n}{10} = 8$. با توجه به اینکه

$$3 + 10 < 20 < 3 + 10 + 14$$

پس چارک اول در رده‌ی سوم قرار دارد که کران پایین آن $L_3 = 12/95$ و فراوانی آن $f_3 = 14$ است. چون $m = 20 - (3 + 10) = 7$ ، بنابر این

$$Q_1 = 12/95 + \frac{7}{14} \times 14 = 14/95.$$

همچنین، چون $\frac{13n}{\mu} = 60$ ، و

$$3 + 10 + 14 + 25 < 60 \leq 3 + 10 + 14 + 25 + 17$$

پس

$$Q_3 = 20/95 + \frac{\lambda}{17} \times 14 = 22/83$$

با توجه به تعریف، برد میان چارکی

$$Q_3 - Q_1 = 22/83 - 14/95 = 7/88$$

است و لذا انحراف چارکی $\frac{Q_3 - Q_1}{\mu} = 3/95$ و ضریب تغییر چارکی

$$\frac{Q_3 - Q_1}{Q_3 + Q_1} \times 100 = 20/86$$

است. چون 8 و $\frac{n}{10} = 1$ پس دهک اول در ردهی دوم با کران پایین

$h = 8 - 3 = 5$ ، پس $h/95 = 8/95$ و فراوانی 10 قرار دارد. چون

$$D_1 = 8/95 + \frac{5}{10} \times 14 = 10/95.$$

به همین قیاس، چون $\frac{9n}{10} = 72$ و

$$3 + 10 + 14 + 25 + 17 < 72 \leq 3 + 10 + 14 + 25 + 17 + 9$$

لذا دهک نهم در ردهی ششم با کران پایین $24/95$ قرار دارد. در نتیجه

$$D_9 = 24/95 + \frac{72 - (3 + 10 + 14 + 25 + 17)}{9} \times 14 = 24/95 + \frac{4}{3} = 26/28$$

همان طور که پیشتر گفتیم، می‌توان مثلاً از $D_9 - D_1$ نیز به عنوان معیاری برای سنجش پراکندگی استفاده کرد. در این مثال $D_9 - D_1 = 15/33$. توجه کنید که 80% داده‌ها

در بازه‌ی از D_1 تا D_9 قرار دارد.

دستوری که در نرم‌افزار مینیتوب را ارائه‌ی گزارشی از کمیت توصیفی، پیش‌بینی شده است. خواننده می‌تواند توضیحات بیشتر درباره‌ی خروجی زیر را در فصل هفتم بیابید.

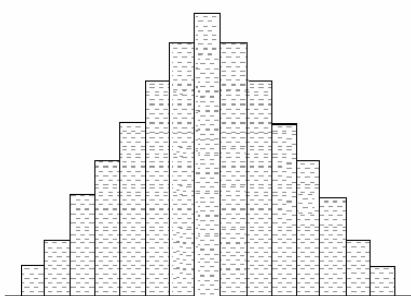
MTB > Describe 'DATA'.

Descriptive Statistics

| Variable | N | Mean | Median | TrMean | StDev | SEMean |
|----------|-------|--------|--------|--------|-------|--------|
| DATA | 80 | 18.930 | 19.050 | 18.963 | 5.670 | 0.634 |
| Variable | Min | Max | Q1 | Q3 | | |
| DATA | 6.200 | 31.800 | 14.825 | 23.375 | | |

۴-۲-۳-۱ معیارهای ویژه‌ی نمودارهای داده‌ها

تاکنون انواع معیارهایی را که بهنحوی «مرکز» داده‌ها و «تغییرپذیری» یا «پراکندگی» داده‌ها را توصیف می‌کردند، شرح دادیم. در واقع نمی‌توان حدی برای تعداد راههای توصیف داده‌ها قائل شد، زیرا آماردانان مرتباً روش‌های جدیدی برای توصیف مشخصه‌های عددی داده‌های مورد مطالعه‌ی خود در مسائل معین به وجود می‌آورند. در این بخش، مسأله‌ی توصیف شکل کلی توزیع داده‌ها را به اختصار مورد بحث قرار می‌دهیم.

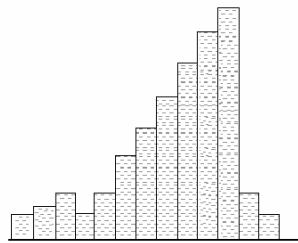


شکل ۶-۱. توزیع زنگ شکل متقارن

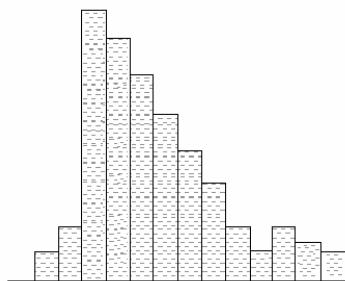
گرچه توزیع‌های فراوانی می‌توانند هر شکل و وضعی داشته باشند، اغلب توزیع‌هایی را که در عمل با آنها برخورد می‌کنیم، می‌توان با استفاده از یکی از چند نوع استاندارد توزیع‌های فراوانی بهنحو نسبتاً رضایت‌بخشی توصیف نمود. در این انواع توزیع‌های

فراآنی، مهم‌تر از همه توزیع زنگ شکل متقارن است که در شکل ۶-۱ نشان داده شده است. در واقع دلایل نظری برای اینکه چرا بسیاری از توزیع‌ها از این الگو تبعیت می‌کنند، وجود دارد.

دو توزیع دیگر در شکل‌های ۷-۱ و ۸-۱ را هنوز هم می‌توان با اندک مسامحه زنگ شکل نامید، اما این توزیع‌ها دیگر متقارن نیستند. توزیع‌های نظیر اینها که «دمچه»‌ای در یک طرف دارند، چوله نامیده می‌شوند، اگر دمچه در سمت چپ باشد، گوییم که توزیع چولگی منفی دارد (شکل ۷-۱) و اگر دمچه در راست باشد، گوییم توزیع چولگی مثبت دارد (شکل ۸-۱). توزیع‌های درآمدها یا دستمزدها اغلب به دلیل وجود تعدادی مقدار نسبتاً بزرگ، چولگی مثبت‌دارند. در این توزیع‌ها، مقادیر کوچک قادر به خنثی کردن اثر مقادیر بسیار بزرگ نیستند.



شکل ۱-۷. توزیع با چولگی منفی



شکل ۱-۸. توزیع با چولگی مثبت

راههای گوناگونی برای اندازه‌گیری میزان چولگی موجود است. یک روش نسبتاً آسان، بر این واقعیت مبنی است که برای یک توزیع زنگ شکل کاملاً متقارن، از نوع توزیع توصیف شده در شکل ۱-۶، مقادیر میانگین و میانه برهم منطبق‌اند. چون وجود تعدادی مقدار نسبتاً بزرگ که مقادیر کوچک نظیر قادر به جبران آنها نیستند، موجب می‌شود که میانگین بزرگتر از میانه باشد (و وجود مقادیر نسبتاً کوچک که مقادیر نسبتاً بزرگ نظیر قادر به جبران آنها نیستند موجب می‌شود که میانگین کوچکتر از میانه باشد)، می‌توانیم از این رابطه‌ی بین میانگین و میانه استفاده و معیار ساده‌ای برای میزان چوله بودن توزیع، تعریف کنیم. این معیار، ضریب چولگی پی‌رسن نامیده می‌شود و مقدار آن چنین است:

$$SK = \frac{3(\bar{x} - \tilde{x})}{s}$$

برای یک توزیع کاملاً متقارن، مقدار SK برابر ۰ است و در حالت کلی این مقدار بین -3 و $+3$ تغییر می‌کند.

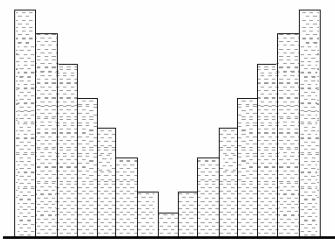
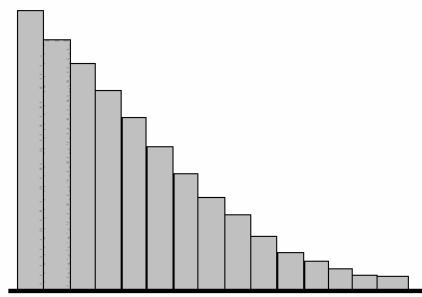
مثال ۲۱-۱. مطابق مثال ۲-۱، مربوط به میزان اکسید سولفور، دیدیم که

$$\bar{x} = 18/85, \quad \tilde{x} = 19/03, \quad s = 5/55$$

بنابراین

$$SK = \frac{3(18/85 - 19/03)}{5/55} = -0.01.$$

این مقدار بسیار کوچک است و لذا می‌توان توزیع را تقریباً متقارن دانست. دو نوع دیگر از توزیع‌هایی که در عمل پیش می‌آیند، توزیع‌های J معکوس شکل و U شکل، که در شکل‌های ۹-۱ و ۱۰-۱ نشان داده شده‌اند. همان‌گونه که در نمودار این توزیع‌ها آشکار است، نام این توزیع‌ها تا حدی شکل توزیع‌ها را معین می‌کنند.

شکل ۹-۱. توزیع U شکلشکل 10-1. توزیع J معکوس شکل

معیار دیگری در نمودارها آماری مطرح است که به میزان کشیدگی یا بر جستگی منحنی توجه دارد. معمولاً این معیار از مقایسه با منحنی نرمال به دست می‌آید.

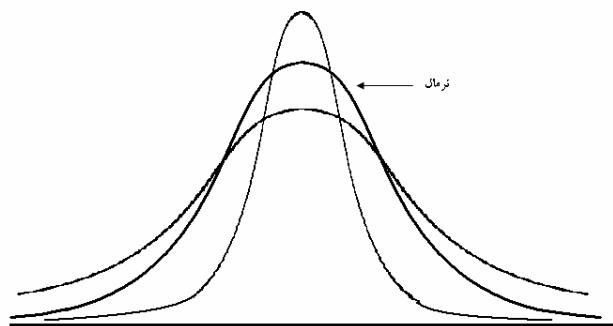
بر جستگی به وسیله‌ی فرمول زیر محاسبه می‌شود:

$$K = \frac{M_4}{s^4} - 3$$

که در آن $M_4 = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^4$ گشتاور مرکزی مرتبه‌ی چهارم، و s انحراف معیار نمونه است. چون مقدار $\frac{M_4}{s^4}$ برای توزیع نرمال به عدد 3 نزدیک است معیار بر جستگی را

از تفاوت دو کمیت محاسبه می‌کنند. بسته به آنکه این کمیت مثبت و یا منفی باشد به

ترتیب بر جستگی بیشتر و یا کمتر از منحنی نرمال خواهد داشت.



شکل ۱۱-۱. وضعیت کشیدگی یا برجستگی منحنی‌ها نسبت به منحنی نرمال

براساس داده‌های مثال ۱-۱، و انجام دستورات زیر، در خروجی مقدار کشیدگی و چولگی به ترتیب $0/645$ و $0/1177$ - نمایش داده می‌شود.

```
MTB > %Describe 'DATA';
SUBC> Confidence 95.0.
```

تمرین

۱-۶. برای مشاهده‌های زیر میانه، نما و چارک اول را محاسبه کنید.

4 0 1 3 2 5 1 0 3 1 2 4 3 1 6 0 7 3 8 2

۷-۱. سه نقطه‌ی A ، B ، و C به یک فاصله از یکدیگر قرار دارند. متحرکی از نقطه‌ی A به نقطه‌ی B ، از نقطه‌ی B به نقطه‌ی C و از نقطه‌ی C به نقطه‌ی A به ترتیب با سرعتی معادل 20 متر در ساعت، 40 متر در ساعت و 80 متر در ساعت حرکت می‌کند. سرعت متوسط این متحرک در مسیر ذکر شده را محاسبه کنید.

۸-۱ رابطه‌ی زیر را به ازای هر a ای ثابت کنید

$$\sum_{i=1}^n (x_i - a)^2 = \sum_{i=1}^n (x_i - \bar{x})^2 + n(\bar{x} - a)^2.$$

۹-۱. آسانسوری به گونه‌ای طراحی شده است که بیشترین بار آن 2000 پوند باشد. اگر یک جابه‌جایی مسافران این آسانسور 8 مسافر زن که میانگین وزن آنها 123 پوند

است و ۵ مرد را که میانگین وزن آنها ۱۷۴ پوند است سوار کند، آیا آسانسور دارای بار بیش از حد مجاز است؟

۱۰-۱. سه نمایندگی فروش یخچال از شرکتی، در سه شهر بزرگ تعداد ۵۷۵، ۱۴۱۰ و ۱۳۱۰۵ یخچال را به ترتیب با متوسط قیمت ۱۲۰۰۰۰ تومان، ۱۱۷۰۰۰ تومان و ۱۲۱۰۰۰ تومان فروخته‌اند. متوسط قیمت یخچال‌های فروخته‌شده چقدر است؟

۱۱-۱. اگر واریانس یک جامعه برابر با ۲ و میانگین آنها برابر با ۳ و مجموع توان‌های دوام آنها برابر با ۵۵ باشد، اندازه‌ی جامعه چقدر است؟

۱۲-۱. اگر میانگین داده‌های x_1, \dots, x_n برابر با \bar{x} باشد، آنگاه میانگین داده‌های $(x_n + 1), (x_1 + 1), \dots, (x_p + 1)$ چقدر است؟

۱۳-۱. اگر میانگین داده‌های x_1, \dots, x_p برابر با ۱۰ باشد، آنگاه میانگین داده‌های $(x_9 + ۹), (x_۸ + ۸), \dots, (x_۱ + ۱)$ را محاسبه کنید.

۱۴-۱. اگر میانگین داده‌های $x_۱, \dots, x_p$ برابر با ۱۵ باشد، آنگاه میانگین داده‌های $(x_p + ۸), (x_{۱۰} + ۸), \dots, (x_۱ + ۱)$ را محاسبه کنید.

۱۵-۱. اگر میانگین داده‌های $x_۱, \dots, x_p$ برابر با ۱۵ باشد، آنگاه میانگین داده‌های $x_۱, \dots, x_p, ۳۵, ۳۶, ۳۷, ۳۸, ۳۹$ را به دست آورید.

۱۶-۱. اگر انحراف متوسط داده‌های $x_۱, \dots, x_p$ برابر با صفر باشد، آنگاه میانگین داده‌های $x_۱, \dots, x_p, ۵۰$ را به دست آورید.

۱۷-۱. اگر به هریک از داده‌ها عدد ۱۵ را بیفزاییم، در مقدار انحراف میانگین چه تغییری رخ می‌دهد؟

۱۸-۱. اگر انحراف معیار داده‌هایی $x_۱, \dots, x_{۱۰}$ برابر با صفر باشد، میانه‌ی داده‌های $(x_۱ + ۱), (x_۲ + ۱), \dots, (x_{۱۰} + ۱)$ را محاسبه کنید.

۱۹-۱. اگر انحراف معیار اعداد ۳۶، ۳۶، ۲۴، ۲۴ و ۸ برابر با $9/8$ باشد، انحراف معیار اعداد ۹، ۶، ۵، ۳ و ۲ چقدر است؟

۲۰-۱. برای ۱۲۰ مشاهده اطلاعات زیر به دست آمده است.

الف) میانگین، میانه، واریانس را به دست آورید.

ب) چارک‌های اوّل و سوم را محاسبه کنید.

ج) مقدار چولگی را به دست آورید.

| ردی | فراوانی |
|--------------|---------|
| ۴۴/۵ - ۵۴/۵ | ۵ |
| ۵۴/۵ - ۶۴/۵ | ۱۵ |
| ۶۴/۵ - ۷۴/۵ | ۱۳ |
| ۷۴/۵ - ۸۴/۵ | ۱۹ |
| ۸۴/۵ - ۹۴/۵ | ۷ |
| ۹۴/۵ - ۱۰۴/۵ | ۱ |
| جمع | ۱۲۰ |

۲۱-۱. طول عمر ۱۰۰ باتری خودرو به ترتیب دارای میانگین، میانه و انحراف استاندارد $\frac{۵}{۴۸}, \frac{۳}{۴۸}$ و $\frac{۱}{۶۵}$ سال است.

الف) ضریب چولگی، و

ب) بر جستگی،
را محاسبه کنید.

۱-۴ جدول‌های دو متغیره و معیارهای توصیفی

تاکنون درباره‌ی یک متغیر بحث کردیم، ولی در اکثر اوقات تحلیل اطلاعات درباره‌ی یک متغیر کافی نیست، زیرا ممکن است چند متغیر مورد مطالعه، به لحاظ داشتن ارتباط با یکدیگر، تغییر کنند. بنابراین مطالعه‌ی جداگانه‌ی آنها، این روابط یا بستگی‌ها را نشان نمی‌دهد و ناچاریم که چند متغیر را یک‌جا در نظر بگیریم.

در این بخش مشاهدات جفتی مربوط به دو صفت از جامعه که به صورت زیر نمونه‌گیری شده‌اند را در جدول‌های دو بعدی مرتب می‌کنیم:

$$(X_1, Y_1), (X_2, Y_2), \dots, (X_n, Y_n).$$

۱-۴-۱ توزیع فراوانی دو متغیره

روشی را که به کار می‌بریم برای متغیرهای کیفی و کمی، گسته و پیوسته یکسان است. فرض کنید که متغیر X مقادیر x_1, x_2, \dots, x_s و متغیر Y مقادیر y_1, y_2, \dots, y_t را اختیار کنند. بنابراین هر عضو جامعه‌ای مقداری از جامعه X مثلاً x_i و اندازه‌ای دیگر از متغیر Y ، مانند y_i را دارد. در نتیجه اعضای جامعه برای دو متغیر X و Y با هم مجموعه‌ای از زوج‌های مرتب (X_i, Y_i) را تشکیل خواهند داد که عبارتند از: هر X_i ‌ای می‌تواند یکی از مقادیر از حوزه تغییرات متغیر X و هر Y_i ‌ای می‌تواند یکی از مقادیر متغیر Y اختیار کنند و حالت‌های ممکنه در جامعه به صورت زیرند:

$$(X_i, Y_i) : \begin{array}{cccc} (x_1, y_1) & (x_1, y_2) & \cdots & (x_1, y_t) \\ (x_2, y_1) & (x_2, y_2) & \cdots & (x_2, y_t) \\ \vdots & \vdots & \ddots & \vdots \\ (x_s, y_1) & (x_s, y_2) & \cdots & (x_s, y_t) \end{array}$$

معمولًا هنگامی که تعداد نمونه یعنی n کوچک باشد برای تعیین هر نوع رابطه، مستقیماً از همین جدول ساده استفاده می‌شود ولی اگر حجم جامعه بزرگ باشد جدول فراوانی متغیرها به صورت جدول دو بعدی نوشته می‌شود و محاسبات با استفاده از این جدول صورت می‌گیرد. امروزه به دست آوردن جداول توزیع فراوانی با استفاده از کامپیوتر و ماشین‌های حساب، بسیار ساده است ولی آنچه را که در اینجا ذکر می‌کنیم در بررسی‌های فردی است که با حجم نسبتاً کوچک صورت می‌گیرد. برای این کار اطلاعات مربوط به هر فرد نمونه یا جامعه را روی یک کارت نوشته سپس کارت‌ها را نسبت به یک متغیر مرتب می‌کنیم. به این ترتیب اگر نسبت به متغیر X مرتب شده باشند یک سری تغییرات از x_1, x_2, \dots, x_s خواهیم داشت، سپس گروه x_i را جدا کرده و نسبت به تغییرات Y مجددًا مرتب می‌کنیم تا t گروه y_1, y_2, \dots, y_t به دنبال هم قرار گیرند و این کار را تا گروه s تکرار می‌کنیم که در پایان به طور معمول $s \times t$ دسته کارت خواهیم داشت که در هر رشته X ‌ها با هم و Y ‌ها نیز با هم برابرند و

تعداد کارت‌های هر رده فراوانی آن رده را بیان می‌کند که برای رده کارتی k است با n_{ij} نشان داده می‌شود.

$$(X = x_i, Y = y_j)$$

نتایج گروه‌بندی را در یک جدول دو بعدی به صورت زیر می‌آوریم.

| X | y_1 | y_p | \cdots | y_t | n_{i_0} |
|----------|----------|----------|----------|----------|-----------|
| x_1 | n_{11} | n_{1p} | \cdots | n_{1t} | n_{1o} |
| x_p | n_{p1} | n_{pp} | \cdots | n_{pt} | n_{po} |
| \vdots | \vdots | \vdots | \ddots | \vdots | \vdots |
| x_s | n_{s1} | n_{sp} | \cdots | n_{st} | n_{so} |
| n_{oj} | n_{o1} | n_{op} | \cdots | n_{ot} | n |

که برای فراوانی‌ها روابط زیر را می‌توان نوشت.

$$\begin{aligned} \sum_{i=1}^s \sum_{j=1}^t n_{ij} &= \sum_{j=1}^t \sum_{i=1}^s n_{ij} = n, \\ \sum_{i=1}^s n_{ij} &= n_{i1} + n_{i2} + \cdots + n_{is} = n_{i\bullet}, \\ \sum_{j=1}^t n_{ij} &= n_{1j} + n_{2j} + \cdots + n_{sj} = n_{\circ j}, \\ \sum_{j=1}^t n_{\circ j} &= n_1 + n_2 + \cdots + n_s = n_{\bullet}. \end{aligned}$$

مثال ۲۱-۱. برای یک نمونه به اندازه‌ی $n=26$ ، از دانشجویان یک کلاس، تعداد فرزندان پسر خانواده، X ، و تعداد فرزندان دختر خانواده، Y ، را سوال کرده‌ایم، نتایج زیر به دست آمده است.

توزيع فراوانی یا جدول دو بعدی به صورت زیر بیان خواهد شد.

| $X \backslash Y$ | ١ | ٢ | ٣ | ٤ | ٥ | ٦ | ٧ | $n_{i,j}$ |
|------------------|---|---|---|---|---|---|---|-----------|
| $n_{\circ,j}$ | ٠ | ٢ | ٣ | ٢ | ٧ | ٥ | ٦ | ٢٦ |
| ٠ | ٠ | ٠ | ٠ | ١ | ٤ | ٠ | ٠ | ٥ |
| ١ | ٠ | ٠ | ١ | ١ | ٣ | ٣ | ٢ | ١٠ |
| ٢ | ٠ | ٠ | ٠ | ٠ | ٠ | ٢ | ٤ | ٦ |
| ٣ | ٢ | ٢ | ١ | ٠ | ٠ | ٠ | ٠ | ٥ |

با نرم افزار مینیتوب

```

MTB > READ C14 C15
DATA> 2 6
DATA> 2 7
DATA> 2 7
DATA> 1 6
DATA> 1 7
DATA> 1 6
DATA> 0 5
DATA> 1 5
DATA> 0 5
DATA> 0 5
DATA> 1 5
DATA> 1 5
DATA> 1 4
DATA> 2 6
DATA> 2 7
DATA> 2 7
DATA> 1 7
DATA> 1 6
DATA> 1 3
DATA> 0 5
DATA> 0 4
DATA> 3 3
DATA> 3 2
DATA> 3 1
DATA> 3 2
DATA> 3 1
DATA> END
      26 rows read.
MTB > Table C14 C15;
SUBC> Counts.
Tabulated Statistics
ROWS: C14    COLUMNS: C15
      1      2      3      4      5      6      7     ALL

```

| | | | | | | | | |
|------------------|---|---|---|---|---|---|---|----|
| 0 | 0 | 0 | 0 | 1 | 4 | 0 | 0 | 5 |
| 1 | 0 | 0 | 1 | 1 | 3 | 3 | 2 | 10 |
| 2 | 0 | 0 | 0 | 0 | 0 | 2 | 4 | 6 |
| 3 | 2 | 2 | 1 | 0 | 0 | 0 | 0 | 5 |
| ALL | 2 | 2 | 2 | 2 | 7 | 5 | 6 | 26 |
| CELL CONTENTS -- | | | | | | | | |
| COUNT | | | | | | | | |

مثال ۱-۲۲. یک نمونه ۲۰ تایی از کارگران یک کارخانه را انتخاب و آنها را بر حسب جنس، X ، و به تفکیک مهارت، Y ، مورد مطالعه قرار داده ایم که داده ها عبارتند از:
 (زن، نیمه ماهر)، (زن، نیمه ماهر)، (مرد، ماهر)، (مرد، ماهر)، (زن، ساده)، (زن، ساده)،
 (مرد، ساده)، (مرد، ساده)، (مرد، نیمه ماهر)، (زن، ساده)، (زن، ساده)، (زن، ماهر)، (مرد، ماهر)،
 (مرد، ماهر)، (مرد، ساده)، (مرد، ساده)، (مرد، نیمه ماهر)، (زن، ساده)
 که جدول توزیع فراوانی به صورت زیر بیان می شود.

| $x \backslash y$ | Maher | Nimaher | Sade |
|------------------|-------|---------|------|
| Zen | ۲ | ۲ | ۵ |
| Merd | ۴ | ۲ | ۵ |

مثال ۱-۲۳. لیست مشاهدات نمرات درس آمار و احتمال، X ، و نمرات درس ریاضی عمومی، Y ، دانشجویان به شرح زیر است:

| | | | | | | | | | | | | | | | |
|-----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|
| x | ۱۰ | ۱۵ | ۱۴ | ۵ | ۰ | ۱۶ | ۱۲ | ۱۰ | ۱۳ | ۱۴ | ۱۴ | ۲۰ | ۲۰ | ۱۵ | ۶ |
| y | ۱۰ | ۱۴ | ۱۴ | ۰ | ۰ | ۱۶ | ۱۰ | ۱۴ | ۱۲ | ۱۳ | ۱۲ | ۱۸ | ۲۰ | ۱۷ | ۳ |
| x | ۶ | ۶ | ۱۲ | ۱۱ | ۱۰ | ۱۷ | ۱۷ | ۱۶ | ۳ | ۳ | ۲ | ۱۴ | ۱۹ | ۵ | ۱۰ |
| y | ۸ | ۸ | ۱۲ | ۱۰ | ۱۴ | ۱۵ | ۱۴ | ۱۶ | ۴ | ۲ | ۲ | ۱۶ | ۲۰ | ۶ | ۷ |

اگر X و Y را به پنج ردۀ با عرض مساوی تقسیم کنیم نتیجه گروه‌بندی به شرح جدول زیر خواهد بود.

| $x \backslash y$ | ۰-۱۴ | ۱۴-۱۸ | ۱۸-۲۲ | ۲۲-۲۶ | ۲۶-۳۰ | n_{ij} |
|------------------|------|-------|-------|-------|-------|----------|
| ۰-۱۴ | ۴ | ۰ | ۰ | ۰ | ۰ | ۴ |
| ۱۴-۱۸ | ۲ | ۱ | ۲ | ۰ | ۰ | ۵ |
| ۱۸-۲۲ | ۰ | ۱ | ۲ | ۲ | ۰ | ۵ |
| ۲۲-۲۶ | ۰ | ۰ | ۱ | ۶ | ۳ | ۱۰ |
| ۲۶-۳۰ | ۰ | ۰ | ۰ | ۱ | ۵ | ۶ |
| $n_{\circ j}$ | ۶ | ۲ | ۵ | ۹ | ۸ | ۳۰ |

۱-۴-۲ معیارهای توصیفی برای دو متغیر

به شرط آنکه دو متغیر مورد مطالعه کمی باشند، با تنظیم جداول دو بعدی مقادیر میانگین و واریانس برای هریک از صفات به دست می آیند.

$$\begin{aligned}\bar{x} &= \frac{1}{n} \sum_{j=1}^t \sum_{i=1}^s n_{ij} x_i = \frac{1}{n} \sum_{i=1}^s n_{i\circ} x_i, \\ \bar{y} &= \frac{1}{n} \sum_{j=1}^t \sum_{i=1}^s n_{ij} y_j = \frac{1}{n} \sum_{j=1}^t n_{\circ j} y_j, \\ s_x^2 &= \frac{1}{n} \sum_{j=1}^t \sum_{i=1}^s n_{ij} (x_i - \bar{x})^2 = \frac{1}{n} \sum_{i=1}^s n_{i\circ} (x_i - \bar{x})^2, \\ s_y^2 &= \frac{1}{n} \sum_{j=1}^t \sum_{i=1}^s n_{ij} (y_j - \bar{y})^2 = \frac{1}{n} \sum_{j=1}^t n_{\circ j} (y_j - \bar{y})^2.\end{aligned}$$

کمیتی ترکیبی، که می تواند تاثیرات دو متغیر بر یکدیگر را اندازه گیری کند را در زیر معرفی می کنیم:

$$s_{xy} = \frac{1}{n} \sum_{j=1}^t \sum_{i=1}^s n_{ij} (y_j - \bar{y})(x_i - \bar{x}).$$

s_{xy} کوواریانس دو متغیر X و Y است. فرمول فوق به صورت زیر ساده می شود:

$$s_{xy} = \frac{1}{n} \sum_{j=1}^t \sum_{i=1}^s n_{ij} x_i y_j - \bar{x} \bar{y}.$$

انتظار داریم اگر دو متغیر هم سو باشند، یعنی با افزایش یکی، دیگری افزایش یابد و بالعکس، مقدار کوواریانس مثبت شود و اگر دو متغیر مورد مطالعه هم سو نباشند، یعنی با افزایش یکی، دیگری کاهش یابد، مقدار کوواریانس منفی شود. کمیت s_{xy} نسبت به تغییر واحد اندازه‌گیری حساسیتی مشابه واریانس دارد.

$$s_{ax+b, cy+d} = ac s_{xy}.$$

مثال ۱-۲۴. با توجه به اطلاعات زیر مقدار کوواریانس بین x و y را به دست آورید.

| x | y | $\sum xy$ | x^2 | y^2 |
|---------------|-----|----------------|------------------|-------------------|
| ۳ | ۱۱ | ۳۳ | ۹ | ۱۲۱ |
| ۵ | ۲۰ | ۱۰۰ | ۲۵ | ۴۰۰ |
| ۴ | ۱۶ | ۶۴ | ۱۶ | ۲۵۶ |
| ۷ | ۲۴ | ۱۶۸ | ۴۹ | ۵۷۶ |
| ۹ | ۲۶ | ۲۳۴ | ۸۱ | ۶۷۶ |
| ۶ | ۱۵ | ۹۰ | ۳۶ | ۲۲۵ |
| ۵ | ۲۱ | ۱۰۵ | ۲۵ | ۴۴۱ |
| ۴ | ۱۸ | ۷۲ | ۱۶ | ۳۲۴ |
| ۸ | ۲۷ | ۲۱۶ | ۶۴ | ۷۲۹ |
| $\sum x = 51$ | | $\sum y = 178$ | $\sum xy = 1082$ | $\sum x^2 = 321$ |
| | | | | $\sum y^2 = 3748$ |

پس،

$$\bar{x} = \frac{51}{9} = 5.667, \quad S_x^2 = \frac{321}{9} - (5.667)^2 = 3.555, \quad s_x = 1.88,$$

$$\bar{y} = \frac{178}{9} = 19.778, \quad S_y^2 = \frac{3748}{9} - (19.778)^2 = 25.27, \quad s_y = 5.02,$$

بنابراین،

$$s_{xy} = \frac{1}{9} (1082) - (5.667) \times (19.778) = 8.141.$$

با مینیتب کمیت کوواریانس را به دست می‌آوریم.

```
MTB > SET C16
DATA > 3 5 4 7 9 6 5 4 8
DATA > END
```

```

MTB > SET C17
DATA> 11 20 16 24 26 15 21 18 27
DATA> END
MTB > LET K7=SUM((C16-MEAN(C16))*(C17-MEAN(C17)))/N(C16)
MTB > PRINT K7
Data Display
K7    8.14815

```

اگر از دستور موجود در نرم افزار استفاده شود خروجی آن یک ماتریس که روی قطر اصلی واریانس هر ستون با احتساب ضریب $1 - \frac{1}{n}$, خواهد شد.

```

MTB > COVARIANCE C16 C17
Covariances
      C16   C17
C16    4.00000
C17   9.16667 28.44444

```

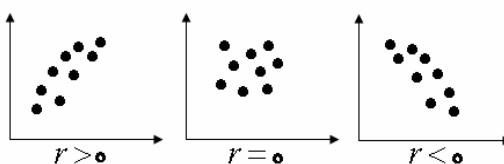
۳-۴-۱ ضریب همبستگی خطی

برای سنجش میزان وابستگی دو متغیر از معیاری به نام ضریب همبستگی خطی استفاده می شود که آنرا به صورت زیر تعریف می کنیم:

$$r_{xy} = r = \frac{s_{xy}}{\sqrt{s_x^2 s_y^2}} = \frac{s_{xy}}{s_x s_y}.$$

ضریب همبستگی خطی به مبدأ و واحد اندازه گیری داده ها بستگی ندارد، یعنی $r_{ax+b, cy+d} = \pm r_{xy}$ می باشد.

با رسم مشاهدات، (x_i, y_i) ها، می توان وضعیت رابطه های خطی بین دو متغیر را تا حدودی پیش بینی کرد. شکل های زیر این مطلب را بهتر نشان می دهند.



ضریب همبستگی خطی r در فاصله زیر تغییر می کند:

$$-1 \leq r \leq 1.$$

اگر $r = \pm 1$ باشد همبستگی خطی کامل است. دو متغیر با $r = 0$ ، همبستگی معکوس دارند یعنی تغییرات X و Y در خلاف جهت یکدیگر و با $r > 0$ همبستگی مستقیم دارد یعنی تغییرات X و Y همجهت هستند. اگر $r < -1$ - همبستگی ناقص است که با $r < 0$ همبستگی ناقص و معکوس و با $r > 0$ همبستگی ناقص و مستقیم است و اگر $r = 0$ باشد همبستگی خطی وجود ندارد. بنابراین هر چه اندازه r به $+1$ یا -1 - نزدیک باشد همبستگی خطی شدیدتر و هر چه r به صفر نزدیک‌تر شود همبستگی ضعیفتر می‌شود.

مثال ۲۵-۱. دنباله‌ی مثال ۲۴-۱ مقدار ضریب همبستگی برابر است با

$$r = \frac{8/141}{(1/88) \times (5/02)} = 0.86$$

این کمیت حاکی از همبستگی ناقص و مستقیم است. این مقدار با خروجی نرم‌افزار نیز قابل حصول است.

MTB > CORR C16 C17
Correlations (Pearson)
Correlation of C16 and C17 = 0.859

تمرین

۱-۲۲. برای اطلاعات مثال ۲۳-۱ ضریب همبستگی را به‌دست آورید.

فصل دوّم

قوانين شمارش

۱-۲ مقدمه

ساده‌ترین نوع محاسبه، شمارش تعداد اعضای یک مجموعه‌ی پایان‌دار است که انسان به طور طبیعی از طریق انگشتان دست یا اشیای دیگر، برای اوّلین بار به آن توجه کرده است. با این حال اگر این مجموعه‌ی پایان‌دار خیلی پر عضو باشد، به‌طوری که نتوانیم یک یک اعضای آن را نشان دهیم، عمل شمارش نه تنها از عهده‌ی انسان بلکه از عهده‌ی کامپیوتر هم ساخته نیست. همچنین محاسبه‌ی احتمال‌ها که در فصل بعدی خواهیم دید، اغلب منجر به شمارش حالات ممکن مختلف می‌شود. خوب‌بختانه به کمک اندیشه‌ی ریاضی می‌توان فرمول‌هایی ابداع کرد که از دشواری کار شمارش بکاهد. فن پیداکردن این فرمول‌ها بر اصولی بدیهی به‌نام اصول شمارش بنا شده است.

۲-۱ اصول شمارش

در اینجا «عمل» نمایشگر هر نوع شیوه، فرایند، یا روش انتخاب است. فرض کنید عمل A را به n_1 طریق به نام‌های x_1, x_2, \dots, x_{n_1} و عمل B را به n_2 راه با نام‌های y_1, y_2, \dots, y_{n_2} بتوان انجام داد.

اصل اوّل شمارش. اگر انجام عمل L منوط به انجام کار A یا B باشد، آنگاه عمل L را می‌توان به $n_1 + n_2$ راه به نام‌های $x_1, x_2, \dots, x_{n_1}, y_1, y_2, \dots, y_{n_2}$ انجام داد. این اصل را «اصل جمع» می‌گویند و در آن تکیه بر روی «یا» می‌باشد.

مثال ۲-۱. فرض کنید از دانشکده به منزل (عمل L)، با تاکسی (عمل A) از سه راه یا با اتوبوس (عمل B) از دو راه امکان داشته باشد. بنابراین به پنج طریق با تاکسی یا با اتوبوس می‌توان از دانشکده به منزل رفت.

تمرین

- ۱-۱. اگر A یک مجموعه با عناصر x_1, x_2, \dots, x_{n_1} و B مجموعه‌ای دیگر با عناصر y_1, y_2, \dots, y_{n_2} باشد، عمل L همتای کدام عمل روی دو مجموعه A و B است.
- ۲-۲. چند عدد یک رقمی که مضرب ۲ یا ۷ هستند، وجود دارد؟

اصل دوم شمارش. اگر انجام عمل L طی دو مرحله انجام‌پذیر باشد به‌طوری که در مرحله‌ی اول (عمل A) به n_1 راه و مرحله‌ی دوم (عمل B) به n_2 راه صورت گیرد، آنگاه عمل مذبور به $n_1 n_2$ راه انجام می‌گیرد.

برای توجیه این قضیه، زوج مرتب (x_i, y_i) را برآمدی تعریف می‌کنیم که وقتی مرحله‌ی اول به امکان x_i و مرحله دوم به امکان y_i متنج شود، رخ می‌دهد. در این صورت، مجموعه‌ی تمام برآمدهای ممکن، مرکب از $n_1 n_2$ زوج زیر است.

$$\begin{aligned} & (x_1, y_1), (x_1, y_2), \dots, (x_1, y_{n_2}) \\ & (x_2, y_1), (x_2, y_2), \dots, (x_2, y_{n_2}) \\ & \vdots \quad \vdots \quad \ddots \quad \vdots \\ & (x_{n_1}, y_1), (x_{n_1}, y_2), \dots, (x_{n_1}, y_{n_2}) \end{aligned}$$

این اصل را «اصل ضرب» می‌گویند و در آن تکیه بر روی «و» می‌باشد.

مثال ۲-۲. فرض کنید از دانشکده به منزل (عمل L) با تاکسی از دانشکده تا پارک شهر (عمل A) از سه راه و با اتوبوس از پارک شهر تا منزل (عمل B) از دو راه امکان داشته باشد. بنابراین به شش طریق با تاکسی و با اتوبوس می‌توان از دانشکده به منزل رفت.

| | x_1 | x_ν | x_μ |
|---------|----------------|------------------|------------------|
| y_1 | (x_1, y_1) | (x_ν, y_1) | (x_μ, y_1) |
| y_ν | (x_1, y_ν) | (x_ν, y_ν) | (x_μ, y_ν) |

تمرین

۲-۳. اگر A یک مجموعه با عناصر x_1, x_ν, \dots, x_n و B مجموعه‌ای دیگر با عناصر

y_1, y_ν, \dots, y_n باشد، عمل L همتای کدام عمل روی دو مجموعه‌ی A و B است.

۲-۴. به چند طریق می‌توانید یکی از سه جفت کفش را با یکی از چهار جفت جوراب

پوشید؟

تعمیم اصل اوّل شمارش. اگر در انجام عملی بتوانیم یکی از k عمل را انتخاب کنیم به‌طوری که عمل اوّل بتواند به n_1 راه انجام شود، عمل دوّم به n_ν راه انجام شود، و الی آخر، کل عمل به $n_1 + n_\nu + \dots + n_k$ راه انجام می‌گیرد.

تعمیم اصل دوّم شمارش. اگر عملی مرکب از k مرحله باشد، به‌طوری که مرحله‌ی اوّل بتواند به n_1 راه انجام شود، و برای هر یک از این راه‌ها، مرحله‌ی دوّم بتواند به n_ν راه صورت گیرد، و برای هریک از راه‌های این دو مرحله نخستین، مرحله‌ی سوم بتواند به n_k راه انجام گیرد و الی آخر، آنگاه کل عمل به $n_1 n_\nu \dots n_k$ راه صورت می‌پذیرد.

مثال ۲-۳. به چند راه می‌توان به آزمونی چهار گزینه‌ای با بیست پرسش پاسخ داد؟ حل. در پاسخ هر پرسش ۴ راه وجود دارد پس جواب بنابر تعمیم اصل دوّم شمارش $4^{\times} 4^{\times} \dots \times 4^{\times} = 4^m$ است.

مثال ۲-۴. از میان ۴ سوپ مختلف، ۳ نوع ساندویچ، ۵ نوع دسر و ۴ نوع نوشابه، چند ناهار مختلف که شامل یک سوپ، یک ساندویچ، یک دسر و یک نوشابه است می‌توان انتخاب کرد؟

حل. بنابر تعمیم اصل دوم شمارش جواب $= 2^4 \times 3 \times 5 \times 4 = 480$ است.

تمرین

- ۵-۲. اعضای شورای دانشجویی یک دانشگاه متشکل از ۳ نفر دانشجوی سال اول، ۴ نفر دانشجوی سال دوم، ۵ نفر دانشجوی سال سوم و ۲ نفر دانشجوی سال چهارم هستند. اگر بخواهیم یک شورای فرعی ۴ نفری را که دانشجویان سال‌های مختلف در آن هستند از این شورا انتخاب کنیم، چند شورای فرعی می‌توان انتخاب کرد؟
- ۶-۲. چند پلاک نمره‌ی اتومبیل با هفت علامت که ۳ علامت اول آن از حروف الفبای لاتین و ۴ علامت آخر آن از اعداد ۰ تا ۹ تشکیل شده است می‌توان تهیه کرد؟
- ۷-۲. چند تابع را می‌توان روی n نقطه تعریف نمود، اگر هر تابع بتواند فقط مقادیر ۰ و ۱ را اختیار کند؟

۳-۲ جایگشت

هر آرایش از اشیای متمایز را یک «جایگشت» می‌نامیم. فرض کنید n شیء متمایز داشته باشیم، می‌توانیم این اشیاء را به صورت‌های مختلف آرایش دهیم.

مثال ۵-۲. چند جایگشت از سه حرف a , b , و c وجود دارد؟
حل. آرایش‌های ممکن عبارت‌اند از:

$$abc \quad acb \quad bac \quad bca \quad cab \quad cba$$

بنابراین تعداد جایگشت‌های متمایز، ۶تاست. می‌توانستیم بدون فهرست کردن جایگشت‌های مختلف، از «تعمیم اصل دوم شمارش» استفاده کنیم و بگوییم: در مکان اول برای جایگاه حرف سمت چپ(راست) سه انتخاب برای گزینش یک حرف وجود دارد، و سپس دو انتخاب برای مکان دوم موجود است. فقط یک حرف برای مکان سوم باقی می‌ماند، لذا تعداد کل جایگشت‌ها $= 3 \times 2 \times 1 = 6$ است.

با تعمیم استدلالی که در مثال ۲-۵ به کار رفت، در می‌یابیم که n شیء متمایز را می‌توان به $n \times (n-1) \times \dots \times 2 \times 1$ راه آرایش داد. این حاصل ضرب را با نماد $n!$ که « n فاکتوریل» خوانده می‌شود نمایش می‌دهیم. پس

$$1! = 1$$

$$2! = 2 \times 1$$

$$3! = 3 \times 2 \times 1$$

:

$$n! = n \times (n-1) \times \dots \times 2 \times 1$$

بنابر تعریف $1! = 1$.

قضیه ۱-۲. تعداد جایگشت‌های متمایز n شیء متمایز برابر $n!$ است.

مثال ۲-۶. تعداد حالت‌هایی که پنج نفر در یک صف قرار می‌گیرند برابر با $5! = 120$ است.

تمرین

۸-۲ با ارقام ۹, ۸, ۷, ۶, ۵ چند عدد نه رقمی که در آن ارقام تکراری نباشند می‌توانند بسازید؟

۹-۲ یک کلاس نظریه‌ی احتمال دارای ۶ دانشجوی پسر و ۴ دانشجوی دختر است. پس از برگزاری یک امتحان، نمرات آنها مرتب می‌شوند. با فرض اینکه هیچ دو دانشجویی نمره یکسان کسب نکنند،

الف) چند حالت ممکن برای مرتب کردن نمرات وجود دارد؟

ب) اگر مرتب کردن دانشجویان پسر و دانشجویان دختر بین خودشان به‌طور جداگانه انجام گیرد، در این صورت چند حالت ممکن وجود دارد؟

۱۰-۲. شخصی ۱۰ کتاب دارد و می‌خواهد در قفسه‌ی کتابخانه خود قرار دهد. از ۱۰ کتاب، تعداد ۴ کتاب ریاضی، ۳ کتاب شیمی، ۲ کتاب تاریخ و یک کتاب زبان هستند. اگر این شخص بخواهد کتاب‌هایی با موضوع یکسان کنار هم قرار گیرند، آنگاه چند ترتیب ممکن برای قرار دادن کتاب‌ها وجود دارد؟

۲-۴ ترتیب

هرگاه از n شیء متمایز، r تا را برگزیده، $n \geq r \leq 1$ ، و تشکیل یک جایگشت دهیم، حاصل یک ترتیب r تایی از n شیء متمایز است. تعداد جایگشت‌های r تایی از n شیء متمایز را بدین طریق پیدا می‌کنیم: برای ساختن یک جایگشت r تایی از n شیء متمایز باید r تا از این n شیء را یک به یک انتخاب کرده مثلاً از چپ به راست پهلوی هم قرار دهیم. این کار معادل با انجام پیاپی r عمل $A_r A_{r-1} A_2 A_1$ به ترتیب، به طریق $n-r+1, n-r+2, \dots, n-1, n$ طریق می‌باشد. بنابر تعمیم اصل دوم شمارش، تعداد جایگشت‌های r تایی از n شیء متمایز می‌شود:

$$n \times (n-1) \times \dots \times (n-r+2) \times (n-r+1).$$

این حاصل ضرب را با ${}_n P_r$ نشان می‌دهند. بنا به تعریف ترتیب، قرار می‌دهیم ${}_n P_0 = 1$.

قضیه‌ی ۲-۲. تعداد ترتیب‌های r تایی از n شیء متمایز برای $r=0, 1, 2, \dots, n$ عبارت است از

$${}_n P_r = \frac{n!}{(n-r)!}.$$

برهان. فرمول ${}_n P_r = n(n-1)\dots(n-r+1)$ را نمی‌توان برای $r=0$ به کار برد، اما فرمول زیر را می‌توان به کار برد.

$${}_n P_0 = \frac{n!}{(n-0)!} = 1.$$

برای $n, r=0, 1, 2, \dots, n$ داریم

$$\begin{aligned} {}_n P_r &= n(n-1)(n-2)\dots(n-r+1) \\ &= \frac{n(n-1)(n-2)\dots(n-r+1) \times (n-r)!}{(n-r)!} = \frac{n!}{(n-r)!}. \end{aligned}$$

مثال ۲-۷. از فهرست نام ۲۴ عضو یک باشگاه، برای انتخاب رئیس، نایب رئیس، خزانه‌دار، و منشی ۴ نام استخراج می‌شود. به‌چند راه مختلف این کار را می‌توان انجام داد؟

حل. تعداد جایگشت‌های ۲۴ شیء متمایز که ۴ به ۴ اختیار می‌شوند برابر است با

$$P_4 = \frac{24!}{20!} = 24 \times 23 \times 22 \times 21 = 255024.$$

بسته‌های نرم‌افزاری آماری، و بسیاری از ماشین حساب‌های دستی برای محاسبه‌ی فرمول‌هایی با نمایش فاکتوریل برنامه‌ریزی شده‌اند. برای مثال با استفاده از مینیتوب مقدارهای ${}_4P_4$ ، و ${}_4P_{19}$ را به دست می‌آوریم.

```
MTB > SET C1
DATA> 1:19
DATA> end
MTB > PARPROD C1 C2
MTB > LET K1=C2(19)
MTB > LET K2=C2(19)/C2(15)
MTB > PRINT K1 K2
Data Display
K1      2.027418E+16
K2      93024.0
```

به کمک مینیتوب، شمارش ویژگی‌های مشخص برای اطلاعات وارد شده به نرم‌افزار امکان‌پذیر است. برای مثال در یک سری داده، تعداد مشاهداتی که کوچکتر، بزرگتر، و یا بین دو عدد قرار دارند قابل تعیین است. همچنین نسبت دادن یک آرایش تصادفی به اعداد، از توانمندی‌های این نرم‌افزار به شمار می‌آید. توضیحات بیشتر در این خصوص را در فصل هفتم آورده شده است.

تمرین

۱۱-۲. در یک مسابقه‌ی شطرنج از ۱۰ بازیکن، ۴ نفر روسی، ۳ نفر آمریکایی، ۲ نفر انگلیسی، و یک نفر بروزیلی هستند. اگر نتیجه مسابقه براساس مرتب کردن ملیت بازیکن‌ها اعلام شود، چند نتیجه ممکن وجود دارد؟

۱۲-۲. چند علامت مختلف را که هر کدام شامل قرارگرفتن ۹ پرچم در یک خط هستند به وسیله‌ی ۴ پرچم سفید، ۳ پرچم قرمز، ۲ پرچم آبی می‌توان تهیه نمود؟ پرچم‌های همنگ یکسان هستند.

۱۳-۲. یک شعبه‌ی محلی انجمن ریاضی به چند طریق می‌تواند برای سه گردهمایی مختلف، سه سخنران را تعیین کند، به شرطی که هر پنج سخنران برای هر یک از تاریخ‌های ممکن، حاضر به انجام سخنرانی باشند؟

۵-۲ جایگشت دوری

جایگشت‌های اشیا روی دایره‌ای مرتب شده‌اند جایگشت‌های دوری خوانده می‌شوند. اگر اشیا متناظر در دو آرایش، اشیای قبل و بعد یکسانی داشته باشند دو آرایش را مختلف تلقی نمی‌کنیم (و آنها را یک بار به حساب می‌آوریم). برای مثال اگر چهار نفر دور میزی گرد نشسته باشند وقتی هرکس به صندلی سمت چپ خود متقل شود، جایگشتی مختلف به وجود نمی‌آید. بنابر این اگر شخصی را به دلخواه در مکان ثابتی درنظر بگیریم، سه نفر دیگر را طی سه مرحله می‌توانیم بنشانیم که حاصل تعداد این سه مرحله، شش جایگشت دوری است.

با تعمیم استدلالی که در این مثال ارائه شد، نتیجه‌ای را که در قضیه‌ی ۲-۳ بیان شده است، به دست می‌آوریم.

قضیه‌ی ۲-۳. تعداد جایگشت‌های n شیء متمایز که روی یک دایره مرتب شده‌اند برابر است با $!(n-1)$.

مثال ۲-۸. به چند طریق می‌توان پنج آگهی تبلیغاتی را دور یک میدان نصب کرد؟ حل. براساس تعداد جایگشت‌های دوری، پاسخ $= 5! = 120$ است.

تمرین

۱۴-۲. چند جایگشت دوری از ۴ نفر که دور یک میز گرد نشسته‌اند وجود دارد؟

۱۵-۲. اگر در یک گروه شش تایی دو شیء کاملاً مشابه یکدیگر بودند، چند جایگشت دوری خواهیم داشت؟

۶-۲ ترکیب

اگر n شیء متمایز داشته باشیم و بخواهیم از این n شیء فقط r شیء را انتخاب کنیم به طوری که ترتیب r شیء مهم نباشد مسلماً به صورت‌های مختلف می‌توانیم این r شیء را برگزینیم. هر یک از این صورت‌ها را یک ترکیب r تایی از n شیء می‌نامیم.

تعداد ترکیب‌های r تایی از n شیء، تنها یک ترکیب r تایی است. بنابراین تعداد

ترکیب‌های r تایی از n شیء متمایز که معمولاً آن را با ${}_n C_r$ نشان می‌دهند، برابر می‌شود با

$${}_n C_r = \binom{n}{r} = \frac{n!}{r!(n-r)!}.$$

قضیه ۲-۴. تعداد ترکیب‌های r شیء از n شیء متمایز برای $n = 0, 1, 2, \dots$ عبارت است از

$$\binom{n}{r} = \frac{n!}{r!(n-r)!}.$$

مثال ۲-۹. در ۶ پرتاب یک سکه به چند راه ۲ شیر و ۴ خط ظاهر می‌شوند؟ حل. می‌توان گفت در چند راه مختلف ۲ پرتاب را که در آن شیر ظاهر می‌شود وجود دارد:

$$\binom{6}{2} = \frac{6!}{2!4!} = 15$$

این نتیجه را می‌توانستیم با فرایند پرزحمت‌تر با شمارش حالت‌های مختلف، یعنی

$$HHTTTT, HTHTTT, \dots$$

که در آن H نمایش شیر و T نمایش خط است به دست آوریم.

مثال ۱۰-۲. می خواهیم از یک گروه ۴۰ نفری یک شورای ۳ نفره را تشکیل دهیم. به چند طریق مختلف این کار امکان پذیر است؟ حل. تعداد حالت ممکن برابر است با

$$\binom{40}{3} = \frac{40 \times 39 \times 38}{3 \times 2 \times 1} = 1140.$$

تمرین

۱۶-۲. از یک گروه متشکل از ۵ زن و ۷ مرد، چند شورای مختلف ۵ عضوی متشکل از ۲ زن و ۳ مرد می توان انتخاب نمود؟ اگر ۲ نفر از مرد ها با یکدیگر خصوصیت داشته و نخواهند با هم در شورا انتخاب شوند، آنگاه چند شورا را می توان انتخاب نمود؟

۱۷-۲. در یک آزمون که شامل ۱۰ سؤال دو گزینه‌ای است به چند طریق می توان، با انتخاب تصادفی ۱۴ پاسخ صحیح داد؟

۷-۲ تعمیم ترکیب

یک ترکیب r تایی از n شیء دارای این تعبیر است که مجموعه‌ای از n شیء را به دو زیرمجموعه‌ی r شیء و $n-r$ شیء افزای می کنیم. پس تعداد ترکیب‌های r تایی از n شیء یعنی تعداد راههایی که می توان یک مجموعه‌ی n شیء را به دو زیرمجموعه‌ی r شیء و $n-r$ شیء افزای کرد. حال این تعبیر را تعمیم می دهیم. می خواهیم تعیین کنیم که یک مجموعه‌ی n شیء را به چند راه می توانیم به k زیرمجموعه که در اوّلی n_1 شیء، در دومی n_2 شیء، ...، و در امی n_k شیء وجود دارد افزای کنیم.

قضیه ۵-۲. تعداد ترکیب‌های n شیء که n_1 تای آنها از نوع اوّل، n_2 تای آنها از نوع دوم، ...، n_k تای آنها از نوع k است، و $n_1 + n_2 + \dots + n_k = n$ ، برابر است با

$$\binom{n}{n_1, n_2, \dots, n_k} = \frac{n!}{n_1! n_2! \dots n_k!}.$$

مثال ۱۱-۲. به چند طریق می‌توان ۲۰ دانشجو را در کلاسی که در ردیف اوّل آن ۵ صندلی، در ردیف دوّم آن ۶ صندلی، و در ردیف سوم آن ۹ صندلی دارد، نشانید؟ حل. در واقع مجموعه‌ای ۲۰ نفری را به چند راه می‌توان به زیرمجموعه‌های ۵ نفری، ۶ نفری و ۹ نفری افزایش کرد:

$$\binom{20}{5, 6, 9} = \frac{20!}{5! 6! 9!} = 189878780.$$

مثال ۱۲-۲. یک مرکز پلیس در شهری کوچک دارای ۱۰ افسر پلیس است. اگر برنامه‌ی مرکز این باشد که ۵ افسر وظیفه نظارت بر خیابان‌ها، ۲ افسر پاسداری از مرکز، و ۳ افسر به صورت ذخیره باشند، به چند حالت ممکن می‌توان ۱۰ افسر را به ۳ گروه تقسیم نمود؟

$$\text{حل. تعداد حالت‌های ممکن برابر است با } \frac{10!}{5! 2! 3!} = 2520.$$

تمرین

۱۸-۲. مجموعه‌ای از چهار شیء را به چند راه می‌توان به ۳ زیرمجموعه که به ترتیب شامل دو، یک، و یک شیء باشد افزایش کرد؟

۸-۲ ضرایب دو جمله‌ای

اگر n عدد صحیح مثبتی باشد و $(x+y)^n$ را جمله‌به‌جمله درهم ضرب کنیم، هر جمله حاصل‌ضربی از x ها و y هاست، به طوری که از هریک از n عامل $(x+y)$ ، یک x یا یک y در این جمله می‌آید. مثلاً بسط

$$\begin{aligned}(x+y)^3 &= (x+y)(x+y)(x+y) \\&= xxx + xxy + xyx + xyy + yxx + yxy + yyx + yyy \\&= x^3 + 3x^2y + 3xy^2 + y^3\end{aligned}$$

جملاتی به شکل x^3 , xy^2 , x^2y , y^3 را نتیجه می‌دهد. ضرایب این جملات ۱، ۳، ۳، ۱ هستند، و برای مثال ضریب xy^2 ، برابر $\binom{3}{2}$ ، یعنی تعداد راههایی است که

می‌توان دو عاملی که x و y را می‌دهند انتخاب کرد. همین‌طور ضریب $x^r y^s$ برابر $\binom{n}{r,s}$ ، یعنی تعداد راه‌هایی است که می‌توان یک عامل فراهم‌کننده‌ی x را انتخاب کرد، و ضرایب x^r و y^s به ترتیب $\binom{n}{r}$ و $\binom{n}{s}$ هستند.

کلی‌تر آن‌که اگر n عدد صحیح و مثبتی باشد و $(x+y)^n$ را جمله‌به‌جمله درهم ضرب کنیم، ضریب $x^{n-r} y^r$ برابر $\binom{n}{r}$ ، یعنی تعداد راه‌هایی است که می‌توان r عامل فراهم‌کننده‌ی y را انتخاب کرد. از این‌رو $\binom{n}{r}$ را ضریب دوچمله‌ای می‌نامیم. حال می‌توانیم قضیه‌ی زیر را بیان کنیم.

قضیه‌ی ۶-۲. برای هر عدد صحیح مثبت n

$$(x+y)^n = \sum_{r=0}^n \binom{n}{r} x^{n-r} y^r.$$

محاسبه‌ی ضرایب دوچمله‌ای را اغلب با استفاده از سه قضیه‌ی زیر می‌توان ساده کرد.

قضیه‌ی ۷-۲. برای هر عدد صحیح مثبت n

$$\binom{n}{r} = \binom{n}{n-r}; \quad r = 0, 1, 2, \dots, n.$$

برهان. می‌توانیم استدلال کنیم که وقتی زیرمجموعه‌ای از n شیء را از مجموعه‌ی n شیء متمایز انتخاب می‌کنیم، زیر مجموعه‌ای با $n-r$ شیء باقی می‌ماند، و بنابر این راه‌های باقی‌ماندن (یا انتخاب کردن) $n-r$ شیء برابر تعداد راه‌های انتخاب r شیء است. برای این‌که قضیه را به صورت جبری ثابت کنیم، می‌نویسیم:

$$\begin{aligned} \binom{n}{n-r} &= \frac{n!}{(n-r)! [n-(n-r)]!} \\ &= \frac{n!}{(n-r)! r!} = \frac{n!}{r!(n-r)!} = \binom{n}{r}. \end{aligned}$$

از قضیه‌ی ۷-۲ نتیجه می‌شود که اگر ضرایب دو جمله‌ای را وقتی n زوج است

برای $\frac{n}{\mu}, \dots, 1, 0, r = 0, 1, \dots, n$ فرد است برای $r = 0, 1, \dots, n$ محاسبه کنیم، ضرایب دو جمله‌ای باقی‌مانده را می‌توان با استفاده از این قضیه به دست آورد.

مثال ۱۳-۲ به فرض اینکه $\binom{4}{1} = 1$ ، $\binom{4}{0} = 1$ ، $\binom{4}{2} = 6$ ، $\binom{4}{3} = 4$ ، $\binom{4}{4} = 1$ را بباید.
 $\binom{4}{4} = \binom{4}{4-4} = \binom{4}{0} = 1$ و $\binom{4}{3} = \binom{4}{4-3} = \binom{4}{1} = 4$. حل.

قضیه‌ی ۸-۲ برای هر عدد صحیح مثبت n ،

$$\binom{n}{r} = \binom{n-1}{r} + \binom{n-r}{r-1}; \quad r = 1, 2, \dots, n-1.$$

برهان. در $(x+y)^n$ قرار می‌دهیم $x=1$ ، و می‌نویسیم

$$(1+y)^n = (1+y)(1+y)^{n-1} = (1+y)^{n-1} + y(1+y)^{n-1}.$$

ضریب y^r در $(1+y)^n$ را با ضریب y^r در $(1+y)^{n-1}$ برابر می‌گیریم.

چون ضریب y^r در $(1+y)^n$ برابر با $\binom{n}{r}$ و ضریب y^r در $(1+y)^{n-1}$ برابر با $\binom{n-1}{r-1}$ است، به دست می‌آوریم.

$$\binom{n}{r} = \binom{n-1}{r} + \binom{n-1}{r-1}$$

یعنی $\binom{n-1}{r-1}$ است، به دست می‌آوریم.

$$\binom{n}{r} = \binom{n-1}{r} + \binom{n-1}{r-1}$$

که برهان را کامل می‌کند.

به روشهای دیگر، هر یک از n شیء را اختیار می‌کنیم، اگر یک شیء مشخص را در

میان r شیء منظور نکنیم $\binom{n-1}{r}$ راه برای انتخاب r شیء وجود دارد، اگر آن شیء

مشخص را در میان r شیء منظور کنیم $\binom{n-1}{r-1}$ راه برای انتخاب $1-r$ شیء دیگر وجود دارد. بنابراین $\binom{n-1}{r} + \binom{n-1}{r-1}$ راه برای انتخاب r شیء موجود است.

برای بیان سومین قضیه درباره ضرایب دو جمله‌ای، تعریف می‌کنیم که هرگاه $\binom{n}{r} = 0$ ، یک عدد صحیح مثبت و r ، عددی صحیح و مثبت و بزرگتر از n باشد، (بدیهی است امکان ندارد بتوانیم زیرمجموعه‌ای از مجموعه‌ای انتخاب کنیم که تعداد عناصرش از خود مجموعه بیشتر باشد).

قضیه ۹-۲. برای هر k , m , و n صحیح مثبت،

$$\sum_{r=0}^k \binom{m}{r} \binom{n}{k-r} = \binom{m+n}{k}.$$

برهان. این قضیه را با برابر گرفتن ضرایب y^k در عبارت‌های دو طرف معادله

$$(1+y)^{m+n} = (1+y)^m (1+y)^n$$

ثابت می‌کنیم. ضریب y^k در $(1+y)^{m+n}$ برابر با $\binom{m+n}{k}$ است و ضریب y^k در $(1+y)^m (1+y)^n = [\binom{m}{0} + \binom{m}{1}y + \dots + \binom{m}{m}y^m] \times [\binom{n}{0} + \binom{n}{1}y + \dots + \binom{n}{n}y^n]$

مجموع حاصل ضرب‌هایی است که از ضرب جمله‌ی ثابت اوّلین عامل در ضریب y^k عامل دوّم، ضرب ضریب y^k اوّلین عامل در ضریب y^{k-1} عامل دوّم، ...، و ضرب ضریب y^k اوّلین عامل در جمله‌ی ثابت عامل دوّم به دست می‌آید. بنابر این

ضریب y^k در $(1+y)^{m+n}$ برابر است با

$$\binom{m}{k} \binom{n}{0} + \binom{m}{1} \binom{n}{k-1} + \binom{m}{2} \binom{n}{k-2} + \dots + \binom{m}{k} \binom{n}{0} = \sum_{r=0}^k \binom{m}{r} \binom{n}{k-r}$$

و برهان کامل می‌شود.

با استفاده از فرمول تعمیم ترکیب، می‌توان ضرایب بسط $(x_1 + x_2 + \dots + x_n)^n$ را تعیین کرد. ضریب چندجمله‌ای، جمله‌ی $x_1^{r_1} x_2^{r_2} \dots x_k^{r_k}$ در بسط $(x_1 + x_2 + \dots + x_n)^n$ برابر است با

$$\binom{n}{r_1, r_2, \dots, r_k} = \frac{n!}{r_1! r_2! \dots r_k!},$$

در صورتی که شرط $n = r_1 + r_2 + \dots + r_k$ برقرار باشد.

مثال ۱۴-۲. در بسط $(x_1 + x_2 + x_3)^6$ ضریب x^6 چیست؟

حل. چون $n = 6$ ، $r_1 = 3$ ، $r_2 = 2$ ، $r_3 = 1$ ، پس

$$\frac{6!}{3! 2! 1!} = 60.$$

تمرین

۱۹-۲. در فرمول قضیه‌ی ۶-۲، با قراردادن مقادیر مناسب به جای x و y نشان دهید که

$$\sum_{r=0}^n \binom{n}{r} (a-1)^r = a^n \quad (\text{ج}) \quad \sum_{r=0}^n (-1)^r \binom{n}{r} = 0 \quad (\text{ب}) \quad \sum_{r=0}^n \binom{n}{r} = 2^n \quad (\text{الف})$$

۲۰-۲. با به کار بردن مکرر قضیه‌ی ۲-۸ نشان دهید که

$$\binom{n}{r} = \sum_{i=1}^{r+1} \binom{n-i}{r-i+a}.$$

۲۱-۲. با به کار بردن قضیه‌ی ۲-۹ نشان دهید که

$$\sum_{r=0}^n \binom{n}{r}^r = \binom{rn}{n}.$$

۲۲-۲. با قراردادن $x = 1$ در رابطه‌ی قضیه‌ی ۶-۲، و سپس مشتق‌گیری از عبارت‌های

$$\sum_{r=0}^n r \binom{n}{r} = n \cdot 2^{n-1}$$

دو طرف نسبت به y و بالاخره قرار دادن $x = y$ ، نشان دهید که

۲۳-۲. اگر α ، عدد صحیح مثبت یا صفر نباشد، بسط دو جمله‌ای $(y+1)^\alpha$ به ازای

$y < 1$ به صورت سری نامتناهی

$$1 + \binom{\alpha}{1}y + \binom{\alpha}{2}y^2 + \binom{\alpha}{3}y^3 + \dots + \binom{\alpha}{r}y^r + \dots$$

تعریف می‌شود،

$$\binom{\alpha}{r} = \frac{\alpha(\alpha-1)\dots(\alpha-r+1)}{r!}$$

این تعریف تعمیم یافته ضرایب دوجمله‌ای را برای محاسبه موارد زیر به کار برد.

$$\text{الف) } \cdot \binom{-\frac{m}{3}}{3} \text{ و } \binom{\frac{1}{4}}{4}$$

$$\text{ب) } \sqrt{5}, \text{ با نوشتن } \sqrt{5} = \sqrt{(1 + \frac{1}{4})^{\frac{1}{2}}}, \text{ و با استفاده از چهار جمله‌ای اوّل بسط دوجمله‌ای } \sqrt{1 + \frac{1}{4}}.$$

۲۴-۲. با رجوع به تعریف تعمیم یافته ضرایب دوجمله‌ای تمرین قبل، نشان دهید که

$$\text{الف) } \binom{-1}{r} = (-1)^r$$

$$\text{ب) به ازای } n > 0, \binom{-n}{r} = (-1)^r \binom{n+r-1}{r}$$

۲۵-۲. در بسط $(x+y+z)^4$ ، ضریب $x^3y^2z^3$ را بیابید.

۲۶-۲. در بسط $(x+3y+4z+w)^9$ ، ضریب $x^3y^2z^3w^4$ را بیابید.

۲۷-۲. با بیان تمام ضرایب دوجمله‌ای بر حسب فاکتوریل‌ها و ساده کردن آنها به طور

جبری، نشان دهید که

$$\binom{n}{n_1, n_2, \dots, n_k} = \binom{n-1}{n_1-1, n_2, \dots, n_k} + \binom{n-1}{n_1, n_2-1, \dots, n_k} + \dots + \binom{n-1}{n_1, n_2, \dots, n_k-1}.$$

فصل سوّم

احتمال

۱-۳ مقدمه

در این فصل سعی می‌شود به اختصار مفاهیم مقدماتی احتمال شامل پیشامد، چگونگی تخصیص احتمال به آنها، روابط پیشامدها، قواعدی از احتمال، احتمال شرطی، و قضیه‌ی بیز را بیان کرده مثال‌هایی ارائه می‌کنیم.

۲-۳ آزمایش تصادفی

در بیشتر شاخه‌های علوم، آزمایش‌ها ابزارهای حیاتی هستند. آزمایش‌ها به دو گروه با نتیجه‌ی قطعی یا نتیجه‌ی تصادفی تقسیم می‌شوند. در آمار و احتمال با آزمایش‌های تصادفی سروکار داریم. در این گونه آزمایش‌ها قبل از انجام آنها نمی‌توانیم نتایج را با یقین کامل پیش‌بینی کنیم.

- در بین ۱۰۰ شیء همانند را که با ماشینی ساخته شده‌اند، نمی‌توانیم تعداد اشیای سالم یا معیوب را از قبل دقیقاً معین نماییم.

- در پرتتاب تاس همگن نمی‌توانیم با اطمینان بگوییم که کدام یک از ارقام ۱ تا ۶ ظاهر می‌شوند.

- در انتخاب مهره از ظرفی که محتوی ۱۰ مهره‌ی همانند به شماره‌های ۱ تا ۱۵ است، نمی‌توان نتیجه‌ی انتخاب را از قبل مشخص کرد.

آزمایش‌هایی که در بالا درباره‌ی آنها صحبت کردیم، دارای ویژگی‌های مشترک

خاصی هستند. برای هر آزمایش، ما از قبل تمام برآمدهای ممکن را نمی‌دانیم و به این خاطر است که بعد از انجام هر آزمایش، هیچ برآمد رخ داده‌ای شگفت‌انگیز نیست. اما با این وجود، ما در حین انجام آزمایش نمی‌دانیم که برآمد معین چه خواهد بود یا به عبارت دیگر در زمان انجام آزمایش یک عدم اطمینان درباره‌ی نتیجه وجود دارد. علاوه براین، آزمایش می‌تواند تحت شرایط یکسان، تکرار شود. این ویژگی‌ها شرح یک آزمایش تصادفی (یا آماری) هستند.

تعريف ۱-۳. یک آزمایش تصادفی (یا آماری) آزمایشی است که در آن
 الف) تمام برآمدهای ممکن آزمایش از قبل مشخص می‌باشند،
 ب) هر بار انجام آزمایش، متنج به نتیجه‌ای می‌شود که از قبل قطعی نیست،
 ج) آزمایش می‌تواند تحت شرایط یکسان تکرار شود.

از این به بعد مراد ما از آزمایش، همان آزمایش تصادفی است و به هر فرآیند مشاهده یا اندازه‌گیری یا شمارش، آزمایش می‌گوییم.

۳-۳ فضای نمونه‌ای و برآمد

نتایجی که از یک آزمایش به دست می‌آیند، برآمدهای آزمایش نامیده می‌شوند. مجموعه‌ی تمام برآمدهای ممکن آزمایش را فضای نمونه‌ای می‌خوانند و معمولاً آن را با حرف S نشان می‌دهند. در فضای نمونه‌ای، هر برآمد را یک عنصر فضای نمونه‌ای یا صرفاً یک نقطه‌ی نمونه‌ای می‌نامند. اگر فضای نمونه‌ای دارای تعداد متناهی از عناصر باشد، می‌توان این عناصر را با نماد معمولی مجموعه، فهرست کرد.

مثال ۳-۱. فضای نمونه‌ای برای برآمدهای ممکن پرتاپ یک سکه را می‌توان به صورت $S = \{H, T\}$ نوشت، که در آن H و T به ترتیب نمایش شیر و خط را می‌دهند.

مجموعه‌ی حاصل از برآمدها، یعنی S ، ممکن است متناهی، نامتناهی شمارا یا نامتناهی باشد.

مثال ۳-۲. فضای نمونه‌ای از آزمایش پرتاپ یک سکه و یک تاس عبارت است از

$$S = \{(1, H), (1, T), (2, H), \dots, (6, T)\}.$$

این مجموعه ۱۲ عضو دارد، پس یک فضای نمونه‌ای متناهی است.

مثال ۳-۳. سکه‌ای را آنقدر پرتاب می‌کنیم تا برای اوّلین بار شیر مشاهده شود. فضای نمونه‌ای این آزمایش دارای تعدادی متناهی عنصر نیست.

$$S = \{H, TH, TTH, \dots\}.$$

این یک مثال از فضای نمونه‌ای نامتناهی شمارا است. اعضای این مجموعه را می‌توان با مجموعه‌ی اعداد طبیعی متناظر نمود.

مثال ۳-۴. مدت زمان اقامت شهروندان در وسایل نقلیه در طول یک روز می‌تواند بازه‌ی [۰, ۲۴] برحسب ساعت باشد. بنابراین $S = [0, 24]$ یک فضای نمونه‌ای نامتناهی است.

تمرین

۱-۳. فضای نمونه‌ای حاصل از پرتاب ۲ تاس را بنویسید.

۲-۳. بر روی قوطی کنسرو لوبيا، عبارت «وزن 250 ± 5 گرم» چاپ شده است، فضای نمونه‌ای وزن قوطی‌های کنسرو را بنویسید.

۳-۳. فضای نمونه‌ای را برای تعداد تصادف‌های روزانه در جاده‌ی آبدان-اهواز را بنویسید.

۴-۳. برای تعداد بازدیدکنندگان هر گروه از اثر باستانی پاسارگاد در ایام نوروز، فضای نمونه‌ای مناسبی بنویسید.

۵-۳. فضای نمونه‌ای نوع خودروهایی را که از مرز زمینی وارد شهر مقدس کربلا می‌شوند، را مشخص نمایید.

۶-۳. سکه‌ای را یک بار پرتاب می‌کنیم، اگر شیر بیاید، تاسی را یک بار می‌ریزیم، اگر خط بیاید، سکه را دوبار دیگر پرتاب می‌کنیم. فضای نمونه‌ای مربوط را بنویسید.

۴-۳ پیشامد

یک پیشامد زیرمجموعه‌ای از فضای نمونه‌ای است و معمولاً با حروف A, B, C, \dots نشان می‌دهیم. گوییم پیشامد A رخ می‌دهد، هرگاه برآمد آزمایش عضوی از آن باشد.

مثال ۳-۵. در پرتاب یک تاس، $S = \{1, 2, 3, 4, 5, 6\}$ ، پیشامد A عبارت است از مشاهده‌ی عددی فرد. بنابراین A دارای عناصر ۱، ۳، ۵ است.

$$A = \{1, 3, 5\} \subset S.$$

هر برآمدی مانند ۱ یا ۳ یا ۵ منجر به رخدادن A می‌شود. پیشامد B که نتیجه‌ی پرتاب تاس عددی بخش‌پذیر بر ۳ باشد به صورت $B = \{3, 6\}$ است.

مثال ۳-۶. صفحه‌ی ساعتی بهصورت دایره است. اگر بر اثر تصادف تیری به سمت صفحه ساعت اصابت شود، اصابت تیر به ناحیه‌ی بین دو عقربه که ساعت ۱۲:۱۵ را تعیین می‌کند یک پیشامد محسوب می‌شود.

مثال ۳-۷. در میان بسته‌های کبریت پیشامد آنکه تعداد نخ‌های یک قوطی کبریت کمتر از ۳۰ باشد.

اگر پیشامد A تنها یک عنصر داشته باشد پیشامد ساده می‌نامند.

مثال ۳-۸. در پرتاب یک تاس، آمدن عدد ۶ یک پیشامد ساده است.

$$A = \{6\}.$$

اگر مجموعه‌ی A که زیرمجموعه‌ای از فضای نمونه‌ای است، تنهی باشد، آن را پیشامد محال می‌نامند و با \emptyset نشان می‌دهند و اگر A برابر S باشد آن را پیشامد حتمی می‌نامند.

مثال ۳-۹. در پرتاب یک تاس، مشاهده‌ی عدد $2/5$ یک پیشامد محال و پیشامد ظاهرشدن یک عدد طبیعی کوچکتر از ۷ یک پیشامد حتمی است.

اگر هر عضو که در A است در B نیز باشد، می‌گویند A زیرپیشامد B است و $A \subset B$ می‌نویسند.

مثال ۱۰-۳. خودروهایی را که در یک بزرگراه در حال حرکت هستند در نظر بگیرید. ملاحظه می‌نمایید که A زیرپیشامد B است.

$$A = \{\text{خودروهایی که سرعت آنها حداقل ۶۰ کیلومتر بر ساعت}\}$$

$$B = \{\text{خودروهایی که سرعت آنها حداقل ۴۵ کیلومتر بر ساعت}\}$$

اگر A زیرپیشامد B و B زیرپیشامد A باشد می‌گویند A و B پیشامدهای برابرند و می‌نویسند $A = B$.

مثال ۱۱-۳. در کتابخانه‌ای در طبقه دوام قفسه‌ایی به کتاب‌های تخصصی رشته‌ی آمار اختصاص داده‌اند. دو پیشامد زیر معادلنده:

$$A = \{\text{كتب طبقه دوام قفسه}\},$$

$$B = \{\text{كتب تخصصی رشته‌ی آمار}\}.$$

اجتماع دو پیشامد از تمام عنصرهای A یا B یا هردو تشکیل می‌شود و آن را با $A \cup B$ نشان داده، پیشامد « A » یا « B » می‌نامند. این پیشامد موقعی رخ می‌دهد که لااقل یکی از دو پیشامد A و B رخ دهند.

مثال ۱۲-۳. در آزمایش تصادفی از پرتاب دو تاس، اگر A پیشامد مجموع دو پرتاب عدد هفت، و B پیشامد نتیجه‌ی یکی از تاس‌ها وجه سه باشد، پیشامد اجتماع این دو عبارت است از:

$$A = \{(1,6), (2,5), (3,4), (4,3), (5,2), (6,1)\}$$

$$B = \{(3,1), (3,2), (3,3), (3,5), (3,6), (1,3), (2,3), (4,3), (5,3), (6,3)\}$$

$$A \cup B = \{(1,6), (2,5), (5,2), (6,1), (3,1), (3,2), (3,3), (3,4), (3,5), (3,6), (1,3), (2,3), (4,3), (5,3), (6,3)\}.$$

اشتراك پيشامد از تمام پيشامدهای ساده که در A و در B تشکيل می شود و آنرا با $A \cap B$ نشان داده پيشامد « A » و « B » می نامند. اين پيشامد موقعی رخ می دهد که A و B هر دو با هم رخدمند.

مثال ۱۳-۳. در مثال ۱۲-۳، ملاحظه می گردد:

$$A \cap B = \{(3,4), (4,3)\}.$$

اگر A و B دو پيشامد از S باشند و $A \cap B = \emptyset$ آنگاه A و B را دو پيشامد «جدا» یا ناسازگار می نامند. در اين حال رخدمند هر دو با هم محال می باشد و عموماً می گويند A و B ناسازگارند.

مثال ۱۴-۳. در پرتاب سه سکه سالم پيشامدهای زير همگي ناسازگارند:

$$A = \{HHH, HHT\}; B = \{HTH, THH, TTH\}; C = \{TTT, HTT, THT\}.$$

(آيا شرحی برای اين سه پيشامد داريد؟)

تفاضل A از پيشامد B از تمام عناصر که در A بوده ولی در B نباشد تشکيل می شود، و آن را با $A - B$ نشان داده، پيشامد « A نه B » می نامند. اگر $A \subset B$ باشد و $A - B = \emptyset$ و به $B - A$ را يك تفاضل واقعی می نامند زيرا B تمام A را دربردارد.

مثال ۱۵-۳. در پرتاب دو تاس، A پيشامد آنکه در بار دوم عدد شش مشاهده شود و B پيشامد آنکه در بار اوّل عدد شش مشاهده شود آنگاه

$$A - B = \{(1,6), (2,6), (3,6), (4,6), (5,6)\},$$

به عبارت ديگر $A - B$ ، پيشامد آن است که فقط در بار دوم عدد شش ظاهر می شود.

تفاضل پيشامد A از فضای نمونه، يعني $S - A$ را متمم A می گويند و آنرا با A' نشان داده، پيشامد «نه A » می نامند. واضح است

$$A \cup A' = S \quad , \quad A \cap A' = \emptyset.$$

مثال ۱۶-۳. دو پیشامد زیر در پرتاب n بار یک سکه، متمم یکدیگرند:

A : پیشامد آنکه حداقل در یکی از پرتابها شیر مشاهده شود،

B : پیشامد آنکه نتیجه تمامی پرتابها خط باشد.

تفاضل متقارن دو پیشامد A و B عناصر آن یا عضوی از A ، یا عضوی از B ، ولی نه عضو هردو تشکیل می‌شود و آن را با $A\Delta B$ نشان می‌دهند و پیشامد «یا A یا B » می‌نامند. واضح است که $A\Delta B$ و $B\Delta A$ با هم برابرند و از این‌رو کلمه‌ی متقارن را به کار می‌برند.

مثال ۱۷-۳. در مثال ۱۵-۳، دو پیشامد ذکر شده را در نظر بگیرید:

$$A\Delta B = \{(1,6), (2,6), (3,6), (4,6), (5,6), (6,6), (6,5), (6,4), (6,3), (6,2), (6,1)\},$$

می‌توان از نتیجه این عملیات متوجه شد که هم‌زمان هر دو تاس عدد شش مشاهده نمی‌شود.

تمرین

۷-۳. اگر $B = \{x | 5 < x < 10\}$ ، $A = \{x | 3 < x \leq 8\}$ ، $S = \{x | 0 < x < 10\}$ ، و پیشامدهای زیر را تعیین کنید.

$$\text{الف) } A \cup B \quad \text{ب) } A \cap B \quad \text{ج) } A \cap B' \quad \text{د) } A' \cup B$$

۸-۳ از گروه ۲۰۰ نفری دانشجویان یک دانشکده، ۱۳۸ نفر درس آمار، ۱۱۵ نفر درس جامعه‌شناسی، و ۹۱ فرد در هر دو درس ثبت‌نام کرده‌اند. چند نفر از این دانشجویان در هیچ یک از این دروس ثبت‌نام نکرده‌اند؟

۹-۳. در شکل ۱-۳، E ، T ، و N پیشامدهایی به ترتیب هستند که خودروهایی که به تعمیرگاهی آورده‌اند نیاز به تعمیر کامل موتور، تعمیر جعبه‌دنده، یا تایرهای نو دارند. پیشامدهایی را که با

الف) ناحیه ۱،

ب) ناحیه ۳،

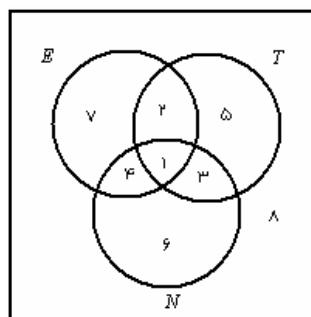
ج) ناحیه ۷،

د) نواحی ۱ و ۴ باهم،

ت) نواحی ۲ و ۵ باهم،

و) نواحی ۳، ۵، و ۶ باهم،

نشان داده می‌شوند در قالب کلمات بیان کنید.



شکل ۱-۳. شکل مورد نیاز برای تمرین ۳-۹

- ۱۰-۳. با رجوع به تمرین قبل و شکل مربوطه، ناحیه یا ترکیبی از نواحی را فهرست کنید که معرف پیشامدهایی برای خودروهایی توصیف شده باشد.
- الف) تعمیر جعبه دنده، ولی نه تعمیر کامل موتور و نه نیاز به تایر نو،
ب) تعمیر کامل موتور و تعمیر جعبه دنده،
ج) تعمیر جعبه دنده یا نیاز به تایر نو و لی نه تعمیر کامل موتور،
د) نیاز به تایر نو.

۳-۵ احتمال

احتمال‌ها، مقادیر یکتابع مجموعه‌ای‌اند که اندازه‌ی احتمال هم خوانده می‌شوند، زیرا همان‌طور که خواهیم دید، این تابع، اعداد حقیقی را به زیرمجموعه‌های مختلف فضای نمونه‌ای S نسبت می‌دهد. اصول موضوع احتمال که ما در اینجا فرمول‌بندی می‌کنیم فقط وقتی به کار می‌روند که فضای نمونه‌ای S گستته باشد به عبارت دیگر فضای نمونه‌ای که تعداد عناصر آن متناهی، $\{O_1, O_2, \dots, O_n\} = S$ ، یا شمارش پذیر نامتناهی،

$S = \{O_1, O_2, O_3, \dots\}$ باشد. علامت $P(A)$ برای نشان دادن احتمال پیشامد A به کار می‌رود.

اصل موضوع ۱. احتمال یک پیشامد، عددی حقیقی نامنفی است، یعنی برای هر زیرمجموعه‌ی A از S ، $P(A) \geq 0$.

اصل موضوع ۲. $P(S) = 1$.

اصل موضوع ۳. اگر A_1, A_2, A_3, \dots ، دنباله‌ای متناهی یا نامتناهی از پیشامدهای دو به دو ناسازگار در S باشند، آنگاه

$$P(A_1 \cup A_2 \cup A_3 \cup \dots) = P(A_1) + P(A_2) + P(A_3) + \dots$$

اصول موضوع ذاتاً نیازی به برهان ندارند، اما اگر بخواهیم نظریه‌ی حاصل را به کار ببریم، وقتی به احتمال‌ها یک مفهوم «حقیقی» می‌دهیم باید نشان دهیم که اصول موضوع صادق هستند.

به طور شهودی، ابتدا احتمال منسوب به یک پیشامد را به عنوان معیاری عددی در نظر می‌گیریم که با آن، میزان انتظار برای وقوع پیشامد را وقتی آزمایش اجرا می‌شود اندازه می‌گیرند. به منظور تعیین این میزان، طبیعی است که نسبت دفعات وقوع پیشامد در تکرار آزمایش را در نظر گیریم. تصور شهودی ما از اندازه‌ی عددی احتمال یک پیشامد، نسبت دفعاتی است که انتظار می‌رود آن پیشامد وقتی که آزمایش تحت شرایط یکسان تکرار گردد، رخ بدهد.

فرآیند مناسب برای تعیین احتمال پیشامدها، بستگی به طبیعت آزمایش و فضای نمونه مربوط است. در بعضی موارد، می‌توان نسبت دفعاتی را که انتظار می‌رود هر برآمد مقدماتی رخ بدهد، با استنتاج منطقی و بدون آنکه آزمایش عملاً اجرا شود، تعیین کرد. در موارد دیگر، لازم است آزمایش را برای تعداد زیادی از دفعات تکرار کرد تا اطلاعاتی درمورد فراوانی وقوع برآمدهای گوناگون به دست آید. تشریح این دو نوع حالت، ما را به سوی تعریف رسمی احتمال راهنمایی می‌کند.

چون نسبت‌ها همواره مثبت یا صفرند، اوّلین اصل موضوع با تعبیر فراوانی در هماهنگی کامل است. اصل موضوع دوم به‌طور غیرمستقیم بیان می‌کند که حتمیت با

احتمال ۱، یکی است، روی هم رفته همیشه می‌پذیریم که باید یکی از امکانات موجود در S رخ دهد، و به این پیشامد که حتمی است احتمال ۱ را نسبت می‌دهیم. تا آنجا که به تعبیر فراوانی مربوط است، احتمال ۱ اشاره براین دارد که پیشامد مورد بحث در ۱۰۰ درصد موقع رخ خواهد داد، یا به عبارت دیگر، این پیشامد مطمئناً رخ می‌دهد.

با در نظر گرفتن سومین اصل موضوع در ساده‌ترین حالت، یعنی برای دو پیشامد ناسازگار A_1 و A_2 ، می‌توان به سادگی دید که این اصل موضوع از نظر تعبیر فراوانی برآورده می‌شود. اگر پیشامدی مثلاً در ۲۸ درصد موقع، و پیشامد دیگری در ۳۹ درصد موقع رخ دهد، و هردو پیشامد نتوانند به طور همزمان رخ دهند (یعنی، دو بهدو ناسازگار باشند)، آنگاه یکی یا دیگری در $39 + 28 = 67$ درصد موقع رخ خواهد داد. بنابراین سومین اصل موضوع صادق است و وقتی بیش از دو پیشامد دو بهدو ناسازگار وجود داشته باشند، همین نوع استدلال به کار می‌رود.

قبل از اینکه برخی نتایج فوری اصول موضوع احتمال را مطالعه کنیم، بر این نکته تأکید می‌کنیم که این سه اصل موضوع به ما نمی‌گویند چگونه احتمال‌ها را به پیشامدها تخصیص دهیم، بلکه فقط راه‌های انجام این کار را محدود می‌کنند.

مثال ۱۸-۳. در یک آزمایش، فضای نمونه‌ای ممکن به چهار پیشامد دو بهدو ناسازگار A ، B ، C ، و D را افزای شده است. برای هریک از موارد زیر، توضیح دهید که چرا راهی معجاز برای تخصیص احتمال‌ها وجود ندارد.

$$\text{الف) } P(D) = -0/20, P(C) = 0/45, P(B) = 0/63, P(A) = 0/12$$

$$\text{ب) } P(D) = \frac{46}{120}, P(C) = \frac{27}{120}, P(B) = \frac{45}{120}, P(A) = \frac{9}{120}$$

حل. الف) $P(D) = -0/20$ ناقض اصل موضوع ۱ است.

$$\text{ب) } 1. P(S) = P(A \cup B \cup C \cup D) = \frac{9}{120} + \frac{45}{120} + \frac{27}{120} + \frac{46}{120} = \frac{127}{120} \neq 1$$

فضای نمونه مقدار احتمالی بزرگتر از یک دارد.

البته، در کاربرد واقعی، احتمال‌ها بر مبنای تجارت گذشته، بر مبنای تحلیل محتاطانه‌ی تمام شرایط زیربنایی، بر مبنای داوری‌های ذهنی، یا بر مبنای مفروضات - گاهی با فرض هم‌شанс بودن (یا به معنای دیگر، برابری احتمال‌ها برای) تمام برآمدهای ممکن، نسبت داده می‌شوند.

در تخصیص احتمال به پیشامدها در یک فضای نمونه‌ای، ضروری نیست که احتمال هر زیرمجموعه‌ی ممکن را توصیف کنیم، و این جای خوشیختی است، زیرا یک فضای نمونه‌ای با فقط 2^m برآمد ممکن، $= 10^{48576}$ زیرمجموعه دارد و تعداد زیرمجموعه‌ها وقتی 5^m برآمد ممکن، $= 10^m$ برآمد ممکن، یا بیشتر موجود باشند، بسیار سریع افزایش می‌یابد. اغلب به جای فهرست کردن احتمال‌های تمام زیرمجموعه‌های ممکن، احتمال‌های پیشامدهای، یا عناصر فضای نمونه‌ای S را فهرست می‌کنیم، و آنگاه قضیه‌ی زیر را به کار می‌بریم.

قضیه ۳-۱. اگر A پیشامدی از فضای نمونه‌ای گستته‌ی S باشد، آنگاه $P(A)$ برابر است با مجموع احتمال‌های پیشامدهای ساده‌ی تشکیل دهنده‌ی A .
برهان. فرض می‌کنیم O_1, O_2, O_3, \dots دنباله‌ای متناهی یا نامتناهی از پیشامدهای ساده باشد که پیشامد A را تشکیل می‌دهند. پس

$$A = O_1 \cup O_2 \cup O_3 \cup \dots$$

و چون پیشامدهای ساده، O_i ‌ها، بر حسب تعریف دو به دو ناسازگارند، از اصل سوم احتمال نتیجه می‌شود که

$$P(A) = P(O_1) + P(O_2) + P(O_3) + \dots$$

و این رابطه، برهان را کامل می‌کند.

برای استفاده از این قضیه، باید قادر به تخصیص احتمال‌ها به پیشامدهای ساده آزمایش باشیم. اینکه چگونه می‌توان این کار را در برخی حالات‌های خاص انجام داد، با مثال‌های زیر تشریح می‌شود.

مثال ۳-۱۹. اگر سکه‌ی سالمی را دوبار پرتاب کنیم، احتمال به دست آوردن حداقل یک شیر چقدر است؟
حل. فضای نمونه‌ای برای این آزمایش عبارت است از

$$S = \{HH, HT, TH, TT\}$$

که در آن، H و T به ترتیب شیر و خط را نشان می‌دهند. چون سکه متعادل است، می‌پذیریم که وقوع این برآمدها هم‌شانس‌اند، و بنابراین به هر نقطه‌ی فضای نمونه‌ای احتمال $\frac{1}{4}$ را نسبت می‌دهیم، اگر A پیشامدی باشد که حداقل یک شیر به دست آید، آنگاه $A = \{HH, HT, TH\}$ ، و

$$\begin{aligned} P(A) &= P(\{HH\}) + P(\{HT\}) + P(\{TH\}) \\ &= \frac{1}{4} + \frac{1}{4} + \frac{1}{4} = \frac{3}{4}. \end{aligned}$$

مثال ۳-۲۰. تاسی را که احتمال وقوع هر عدد فرد با آن، دو برابر احتمال وقوع هر عدد زوج است در نظر بگیرید. اگر در یکبار ریختن این تاس، G پیشامد وقوع عددی بزرگتر از ۳ باشد، $P(G)$ را بیابید.

حل. فضای نمونه‌ای عبارت است از $S = \{1, 2, 3, 4, 5, 6\}$. اگر به هر عدد زوج، احتمال w و به هر عدد فرد، احتمال $2w$ را نسبت دهیم، مطابق اصل موضوع ۲ داریم:

$$2w + w + 2w + w + 2w + w = 9w = 1$$

$$\text{پس } w = \frac{1}{9}, \text{ و}$$

$$P(G) = \frac{1}{9} + \frac{2}{9} + \frac{1}{9} = \frac{4}{9}.$$

اگر فضای نمونه‌ای، گسسته و لی نامتناهی باشد، احتمال‌ها را باید به وسیله‌ی یک قاعده‌ی ریاضی ترجیحاً به وسیله‌ی فرمول یا معادله‌ای، به برآمدهای فردی نسبت داد.

مثال ۳-۲۱. اگر O_1, O_2, \dots, O_n ، دنباله‌ای نامتناهی از پیشامدهای ساده‌ی یک آزمایش مفروض باشد، تحقیق کنید که

$$P(O_i) = \left(\frac{1}{\mu}\right)^i; \quad i = 1, 2, 3, \dots,$$

واقعاً یک اندازه‌ی احتمال است.

حل. چون احتمال‌ها همگی مثبت‌اند، فقط باید نشان‌دهیم که $P(S) = 1$. داریم

$$P(S) = \frac{1}{2} + \frac{1}{4} + \frac{1}{8} + \frac{1}{16} + \dots$$

اگر فرمول تعیین مجموع جملات یک تصاعد هندسی نامتناهی را به‌کار بریم، به دست آوریم

$$P(S) = \frac{\frac{1}{2}}{1 - \frac{1}{2}} = 1.$$

در رابطه با مثال ۲۱-۳، کلمه‌ی «مجموع» در قضیه‌ی ۳-۱، باید به‌قسمی تعبیر شود که شامل مقدار یک سری نامتناهی نیز باشد.

اندازه‌ی احتمال مثال ۲۱-۳، مثلاً اگر O_i این پیشامد باشد که شخصی در پرتاب سکه‌ای همگن برای اوّلین بار در پرتاب نام سکه، خط بیاورد، اندازه‌ی احتمال مناسبی است. پس، احتمال آنکه اوّلین خط در سومین، چهارمین، یا پنجمین پرتاب سکه ظاهر شود برابر است با

$$\left(\frac{1}{2}\right)^3 + \left(\frac{1}{2}\right)^4 + \left(\frac{1}{2}\right)^5 = \frac{7}{32}$$

و احتمال آنکه اوّلین خط در تعداد فردی از پرتاب‌ها رخدده برابر است با

$$\left(\frac{1}{2}\right)^1 + \left(\frac{1}{2}\right)^3 + \left(\frac{1}{2}\right)^5 + \dots = \frac{\frac{1}{2}}{1 - \frac{1}{2}} = \frac{1}{3}.$$

اگر مانند مثال ۱۹-۳، آزمایش چنان باشد که بتوانیم برای نقاط فضای نمونه‌ای S ، احتمال‌هایی برابر در نظر گیریم، می‌توانیم از حالت خاص قضیه‌ی ۳-۱ که به صورت زیر است، استفاده کنیم.

قضیه‌ی ۲-۳. اگر نتیجه‌ی آزمایشی بتواند به یکی از N برآمد مختلف هم‌شانس باشد، و اگر n تا از این برآمدها با هم پیشامد A را تشکیل دهند، آنگاه احتمال پیشامد A برابر است با

$$P(A) = \frac{n}{N}.$$

برهان. فرض می‌کنیم O_1, O_2, \dots, O_n ، پیشامدهای ساده، یا پیشامدهایی با اعضای برآمدهای فردی S را نشان دهنده احتمال هر کدام $\frac{1}{N}$ است. اگر پیشامد A ، اجتماع n تا از این برآمدهای دو به دو ناسازگار بوده، و مهم نباشد که کدامیک از آنها این پیشامد را تشکیل دهند، آنگاه

$$\begin{aligned} P(A) &= P(O_1 \cup O_2 \cup \dots \cup O_n) \\ &= P(O_1) + P(O_2) + \dots + P(O_n) \\ &= \underbrace{\frac{1}{N} + \frac{1}{N} + \dots + \frac{1}{N}}_{\#n} \\ &= \frac{n}{N}. \end{aligned}$$

مشاهده کنید که فرمول $P(A) = \frac{n}{N}$ در قضیه‌ی ۲-۳، با مفهوم احتمال کلاسیک یکی است (در برخی متون از قضیه‌ی ۲-۳ به عنوان تعریفی از مفهوم احتمال کلاسیک استفاده می‌کنند (بازرگان ۱۳۸۴)). البته آنچه در اینجا نشان داده‌ایم آن است که مفهوم احتمال کلاسیک با اصول موضوع احتمال سازگار است - این مطلب در مورد خاصی که برآمدهای فردی هم‌شانس‌اند از اصول موضوع احتمال نتیجه می‌شود.

مثال ۲۲-۳. در یک بازی با دسته کارت ۵۲ تایی (کارت‌ها از شماره‌های یک الی سیزده در چهار رنگ بسته‌بندی شده‌اند)، پنج کارت به تصادف اختیار می‌کنیم، اگر سه کارت دارای یک شماره و دو کارت دیگر هم دارای یک شماره باشند برنده‌ی بازی هستیم. احتمال برنده شدن در این بازی چقدر است؟

حل. تعداد راههایی که می‌توان سه کارت با یک شماره و دو کارت با یک شماره دیگر، مثلاً سه کارت با شماره‌ی ۷ و دو کارت با شماره‌ی ۱ انتخاب کرد برابر است با $\binom{14}{3} \times \binom{14}{2}$. چون ۱۳ شماره از هر رنگ داریم ۱۳ راه انتخاب سه کارت با یک شماره و سپس ۱۲ راه انتخاب دو کارت با یک شماره وجود دارند. کل راههای ممکن انتخاب چنین پنج کارتی برابر

$$n = 13 \times 12 \times \binom{14}{3} \times \binom{14}{2}$$

است. از طرفی تعداد راههای ممکن انتخاب پنج کارت از ۵۲ کارت

$$N = \binom{52}{5}$$

است و لذا از قضیه‌ی ۲-۳ نتیجه می‌شود که احتمال برد بازی عبارت است از

$$P(A) = \frac{n}{N} = \frac{13 \times 12 \times \binom{14}{3} \binom{14}{2}}{\binom{52}{5}} = 0.0014.$$

۶-۳ برخی قواعد احتمال

با استفاده از سه اصل موضوع احتمال، می‌توانیم بسیاری از قاعده‌های دیگر را که کاربردهای مهمی دارند، نتیجه بگیریم. بین پیامدهای فوری اصل موضوع، قضایای زیر را ثابت می‌کنیم.

قضیه‌ی ۳-۳. همواره داریم $P(\phi) = 0$.

برهان. چون پیشامدهای S و ϕ ناسازگارند و مطابق تعریف مجموعه‌ی تهی ϕ ، $S \cup \phi = S$ نتیجه می‌شود که

$$\begin{aligned} P(S) &= P(S \cup \phi) \\ &= P(S) + P(\phi) \quad (\text{بنای اصل موضوع ۳}) \end{aligned}$$

و در نتیجه $P(\phi) = 0$.

باید به این مطلب مهم توجه کرد که از $P(A) = 0$ نتیجه نمی‌شود که الزاماً A مجموعه‌ای تهی است. در عمل اغلب به پیشامدهایی که، به زبان محاوره‌ای، در یک میلیون سال هم رخ نمی‌دهند، احتمال صفر را نسبت می‌دهیم. برای نمونه، مثال کلاسیکی وجود دارد که ما به این پیشامد که میمونی "جمهوری افلاطون" را کلمه‌به‌کلمه بدون خطابا فشار روی دکمه‌ها، تایپ کند، احتمال صفر را نسبت می‌دهیم. این واقعیت شایسته‌ی توجه است که $P(A) = 0$ ، رابطه‌ی $A = \emptyset$ را، به ویژه در فضای نمونه‌ای ناشمارا، نتیجه نمی‌دهد.

قضیه‌ی ۳-۴. برای هر دو پیشامد ناسازگار از S مانند A و B داریم

$$P(A \cup B) = P(A) + P(B).$$

برهان. با توجه به رابطه‌ی ... $\phi \cup \emptyset \cup \dots = \emptyset$ و اصل موضوع ۳،

$$\begin{aligned} P(A \cup B) &= P(A \cup B \cup \phi \cup \emptyset \cup \dots) \\ &= P(A) + P(B) + P(\phi) + P(\emptyset) + \dots \\ &= P(A) + P(B). \end{aligned}$$

یک نتیجه‌ی فوری که می‌توان از قضیه‌ی ۳-۴ گرفت آن است که از ناسازگاری A' و A و اینکه $S = A \cup A'$ ، استفاده کنیم و رابطه‌ی زیر را به دست آوریم.

$$P(A') = 1 - P(A).$$

در رابطه با تعبیر فراوانی، این نتیجه اشاره بر آن دارد که اگر پیشامدی، مثلاً در ۳۷ درصد موقع رخ دهد، در ۶۳ درصد موقع رخ نمی‌دهد.

قضیه‌ی ۳-۵. اگر A و B پیشامدهایی در فضای نمونه‌ای S باشند و $A \subset B$ ، آنگاه

$$P(A) \leq P(B)$$

برهان. چون $A \subset B$ ، می‌توانیم بنویسیم

$$B = B \cap S = B \cap (A \cup A') = (B \cap A) \cup (B \cap A') = A \cup (B \cap A')$$

که درستی روابط آخر به آسانی به وسیله‌ی یک نمودار ون تحقیق می‌شود. در این صورت A و $A' \cap B$ ناسازگارند، و در نتیجه

$$\begin{aligned} P(B) &= P(A) + P(A' \cap B) && \text{(بنابر اصل موضوع ۳)} \\ &\geq P(A). && \text{(بنابر اصل موضوع ۱)} \end{aligned}$$

با استفاده از کلمات، این قضیه بیان می‌کند که اگر پیشامد A زیرمجموعه‌ای از پیشامد B باشد، آنگاه $P(A)$ بزرگتر از $P(B)$ نیست. مثلاً احتمال کشیدن یک کارت سیز از یک دسته کارت ۵۲ تایی، بزرگتر از احتمال کشیدن یک کارت سیز یا قرمز نیست. در واقع احتمال $\frac{1}{4}$ با احتمال $\frac{1}{2}$ مقایسه می‌شود.

قضیه ۶-۳. برای هر پیشامد A ، $0 \leq P(A) \leq 1$.

برهان. با به کار بردن قضیه ۳-۵ و این واقعیت که برای هر پیشامد A در S

$$\phi \subset A \subset S$$

$$P(\phi) \leq P(A) \leq P(S)$$

چون $P(\phi) = 0$ و $P(S) = 1$ خواهیم داشت

$$0 \leq P(A) \leq 1.$$

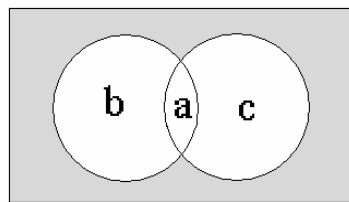
سوئین اصل موضوع احتمال اغلب به یک قاعده‌ی جمع خاص اطلاق می‌شود. قاعده‌ی مزبور از این نظر خاص است که پیشامدهای A_1, A_2, \dots, A_n ، باید دو به دو ناسازگار باشند. برای هر دو پیشامد A و B ، قاعده‌ی جمع عام زیر موجود است.

قضیه ۷-۳. اگر A و B دو پیشامد در فضای نمونه‌ای S باشند، آنگاه

$$P(A \cup B) = P(A) + P(B) - P(A \cap B).$$

برهان. همان‌طور که در نمودار ون شکل ۲-۳ دیده می‌شود به پیشامدهای دو به دو ناسازگار $A \cap B$ ، $A' \cap B$ ، $A \cap B'$ ، $A' \cap B'$ ، به ترتیب احتمال‌های a ، b ، c را نسبت می‌دهیم و به دست می‌آوریم

$$\begin{aligned}
 P(A \cup B) &= a + b + c \\
 &= (a + b) + (c + a) - a \\
 &= P(A) + P(B) - P(A \cap B).
 \end{aligned}$$



شکل ۲-۳: نمودار ون برای اثبات قضیه‌ی ۷-۳

مثال ۲۳-۳. اگر برای خانواده‌ای (که به منظور یک بررسی نمونه‌ای در یک ناحیه‌ی بزرگ شهری، به تصادف انتخاب شده است) احتمال دارا بودن یک دستگاه تلویزیون رنگی، یک دستگاه کامپیوتر و یا هردو وسیله به ترتیب $0/86$, $0/35$, و $0/29$ باشد، احتمال اینکه این خانواده یکی از دو نوع یا هر دو نوع دستگاه (حداقل یکی از دو نوع دستگاه) را داشته باشند چقدر است؟

حل. اگر A , این پیشامد باشد که خانواده‌ی مزبور دارای یک دستگاه تلویزیون رنگی است، و B , این پیشامد باشد که دارای یک دستگاه کامپیوتر است، داریم $P(B) = 0/35$, $P(A) = 0/86$, $P(A \cap B) = 0/29$ قضیه‌ی ۷-۳، قراردهیم، نتیجه می‌شود

$$P(A \cup B) = 0/86 + 0/35 - 0/29 = 0/92.$$

مثال ۲۴-۳. خودرویی که در کنار بزرگراهی متوقف شده است با احتمال $0/23$ ترمزهای معیوب و با احتمال $0/24$ فرسودگی شدید تایر دارد. همچنین با احتمال $0/38$ ترمزهای معیوب یا فرسودگی شدید تایر یا هردو را دارد. احتمال اینکه این اتومبیل، ترمزهایش معیوب بوده و فرسودگی شدید تایر داشته باشد، چقدر است؟

حل. اگر B این پیشامد باشد که اتومبیل مزبور ترمزهایش معیوب است، و T این پیشامد باشد که تایرهاش فرسودگی شدید دارند، داریم

$$P(B) = ۰/۲۳, \quad P(T) = ۰/۲۴, \quad P(B \cup T) = ۰/۳۸.$$

از قرار دادن این مقادیر در فرمول قضیه‌ی ۷-۳ نتیجه می‌شود

$$۰/۳۸ = ۰/۲۳ + ۰/۲۴ - P(B \cap T).$$

از حل این معادله نسبت به $P(B \cap T)$, به دست می‌آوریم

$$P(B \cap T) = ۰/۲۳ + ۰/۲۴ - ۰/۳۸ = ۰/۰۹.$$

با به کار بردن مکرر قضیه‌ی ۷-۳، می‌توان قاعده‌ی جمع را به‌قسمی تعمیم داد که برای هر تعداد از پیشامدها قابل اجرا باشد. مثلاً برای سه پیشامد به دست می‌آوریم:

قضیه‌ی ۸-۳. اگر A , B , و C , سه پیشامد دلخواه در فضای نمونه‌ای در S باشند، آنگاه

$$\begin{aligned} P(A \cup B \cup C) &= P(A) + P(B) + P(C) - P(A \cap B) - P(A \cap C) \\ &\quad - P(B \cap C) + P(A \cap B \cap C). \end{aligned}$$

برهان. $A \cup B \cup C$ را به صورت $A \cup (B \cup C)$ می‌نویسیم و دوبار قضیه‌ی ۷-۳ را به کار می‌بریم، یک بار برای $P(A \cup (B \cup C))$ و یک بار برای $P(B \cup C)$, به دست می‌آوریم

$$\begin{aligned} P(A \cup B \cup C) &= P(A \cup (B \cup C)) \\ &= P(A) + P(B \cup C) - P(A \cap (B \cup C)) \\ &= P(A) + P(B) + P(C) - P(B \cap C) - P(A \cap (B \cup C)). \end{aligned}$$

از قانون توزیع پذیری اشتراک نسبت به اجتماع استفاده نموده و داریم

$$\begin{aligned} P(A \cap (B \cup C)) &= P((A \cap B) \cup (A \cap C)) \\ &= P(A \cap B) + P(A \cap C) - P((A \cap B) \cap (A \cap C)) \\ &= P(A \cap B) + P(A \cap C) - P(A \cap B \cap C), \end{aligned}$$

و نتیجه آنکه

$$\begin{aligned} P(A \cup B \cup C) &= P(A) + P(B) + P(C) - P(A \cap B) - P(A \cap C) \\ &\quad - P(B \cap C) + P(A \cap B \cap C). \end{aligned}$$

مثال ۲۵-۳. فرض می‌کنیم شخصی که به دندانپزشک خود مراجعه می‌کند احتمال اینکه دندان‌هایش را جرم‌گیری کند 0.44 ، احتمال اینکه دندانی برای پرکردن داشته باشد 0.21 ، احتمال آنکه دندان‌هایش را جرم‌گیری و دندانی را پرکند 0.08 ، احتمال اینکه دندان‌هایش را جرم‌گیری کند و دندانی را بکشد 0.11 ، احتمال اینکه دندانی برای پرکردن و دندانی برای کشیدن داشته باشد 0.07 و احتمال اینکه دندان‌هایش را جرم‌گیری و دندانی را پرکند و دندانی را بکشد 0.03 باشد. احتمال اینکه دندانپزشک حداقل یکی از سه مورد را برای او انجام دهد چقدر است؟

حل. اگر C این پیشامد باشد که شخص دندان‌هایش را جرم‌گیری کند،

F این پیشامد باشد که شخص دندانی را پرکند،

E این پیشامد باشد که شخص دندانی را بکشد،

آنگاه $P(C \cap F) = 0.08$ ، $P(E) = 0.21$ ، $P(F) = 0.44$ ، $P(C) = 0.44$

$P(C \cap F \cap E) = 0.03$ ، $P(F \cap E) = 0.07$ ، $P(C \cap E) = 0.11$

فرمول، نتیجه می‌شود که

$$P(C \cup F \cup E) = 0.44 + 0.21 + 0.08 - 0.11 - 0.07 + 0.03 = 0.66.$$

تمرین

۱۱-۳. اگر A و B دو پیشامد در فضای نمونه‌ای S باشند و $P(A) = 0.2$ ، $P(B) = 0.7$ و $P(A' \cup B') = 0.4$ ، مقادیر $P(A \cap B)$ ، $P(A' \cap B)$ ، $P(A \cup B)$ و $P(A' \cap B')$ را به دست آورید.

۱۲-۳. فرمولی برای محاسبه $P(A \Delta B)$ بر حسب $P(A \cap B)$ ، $P(B)$ ، $P(A)$ و $P(B)$ ارائه دهید (احتمال اینکه دقیقاً یکی از دو پیشامد رخ دهد).

۱۳-۳. اگر A و B دو پیشامد در فضای نمونه‌ای S باشند، احتمال پیشامدهای زیر را بر حسب $P(A \cap B)$ ، $P(B)$ ، $P(A)$ و $P(A \cap B)$ به دست آورید:

الف) دقیقاً i تا، $0, 1, 2, \dots, i$ ، از پیشامدهای A و B رخ دهند.

ب) حداقل i تا، $i=0,1,2$ ، از پیشامدهای A و B رخ دهند.

ج) حداقل i تا، $i=0,1,2$ ، از پیشامدهای A و B رخ دهند.

۱۴-۳. جعبه‌ای دارای ۲۴ مهره است که از ۱ تا ۲۴ شماره‌گذاری شده است. یک مهره به تصادف از این جعبه انتخاب می‌کنیم. اگر A پیشامد مشاهده‌ی عددی بخش‌پذیر بر 2 و B پیشامد مشاهده‌ی عددی بخش‌پذیر بر 3 باشد، مقدار $P(A \cap B)$ را به دست آورید.

۱۵-۳. برای A و B دو پیشامد در فضای نمونه‌ای S ، نشان دهید

$$\text{الف) } P(A \cap B) \leq P(A) + P(B)$$

$$\text{ب) } P(A \cap B) \geq P(A) + P(B) - 1$$

۱۶-۳. به کمک استقرا ثابت کنید که برای هر دنباله‌ی متناهی از پیشامدهای

$$E_n, E_{n-1}, \dots, E_1$$

$$P(E_1 \cup E_2 \cup \dots \cup E_n) \leq \sum_{i=1}^n P(E_i).$$

۱۷-۳. بخت رخدادن یک پیشامد بر حسب نسبت احتمال وقوع آن به احتمال عدم وقوع آن داده می‌شود، به شرط آنکه هیچ یک از دو احتمال صفر نباشد. بخت را عموماً بر حسب دو عدد صحیح مثبت که عامل مشترک ندارند، بیان می‌کنند. اگر بخت اینکه

پیشامدی رخ دهد برابر a به b باشد، نشان دهید که احتمال این پیشامد است.

۱۸-۳. توضیح دهید که چرا در هریک از احکام زیر اشتباهی وجود دارد:

الف) احتمال اینکه فرهاد در امتحان وکالت قبول شود ۶۶٪ و احتمال اینکه رد شود ۳۴٪ است.

ب) احتمال اینکه یک منشی در تایپ گزارشی مرتکب ۱۰، ۱۴، ۲۰، ۳۰، ۴۰ یا بیش از ۵ اشتباه شود به ترتیب ۱۲٪، ۲۵٪، ۳۶٪، ۴۰٪ و ۵۷٪ است.

ج) احتمال اینکه بانکی در روزی مفروض $0.1, 0.2, 0.3$ یا بیش از 0.3 چک بدون محل دریافت کند به ترتیب $0.08, 0.21, 0.29, 0.45$ است.

۱۹-۳. احتمال اینکه نحوه بهره‌برداری از یک دستگاه جدید اشعه‌ای ایکس، به خیلی مشکل، مشکل، متوسط، آسان و خیلی آسان درجه‌بندی شود به ترتیب $0.17, 0.34, 0.58, 0.29, 0.05$ است. پیدا کنید احتمال اینکه نحوه بهره‌برداری دستگاه به یکی از صورت‌های زیر درجه‌بندی شود:

الف) مشکل یا خیلی مشکل،

ب) نه خیلی مشکل و نه خیلی آسان،

ج) متوسط یا بدتر،

د) متوسط یا بهتر.

۲۰-۳. شرکت ایران خودرو برای تأمین لاستیک، می‌تواند با احتمال‌های $0.15, 0.24, 0.5, 0.22, 0.08, 0.05$ به ترتیب از شرکت‌های A, B, C, D, E, F خریداری کند. احتمال خرید لاستیک را در هریک از حالت‌های زیر تعیین کنید.

الف) از شرکت B یا E باشد،

ب) از شرکت A, C یا E باشد،

ج) از شرکت C یا F باشد،

د) از شرکت A, C, D, E یا F باشد.

۲۱-۳. در کلاهی بیست تکه کاغذ سفید که از ۱ تا 20 ، ده تکه کاغذ قرمز که از ۱ تا 10 ، چهل تکه کاغذ زرد که از ۱ تا 40 ، و ده تکه کاغذ آبی که از ۱ تا 10 شماره‌گذاری شده‌اند ریخته شده است. اگر این 80 تکه کاغذ را کاملاً مخلوط کنیم به‌قسمی که تمام تکه کاغذها در موقع استخراج احتمال برابر داشته باشند، پیدا کنید احتمال استخراج یک تکه کاغذ را که

الف) آبی یا سفید باشد،

ب) به شماره‌ی $1, 2, 3, 4, 5$ باشد،

ج) قرمز یا زرد و با شماره‌ی ۱، ۲، ۳، ۴ باشد،

د) به شماره‌ی ۵، ۱۵، ۲۵، ۳۵ باشد،

ه) سفید و با شماره‌ای بزرگتر از ۱۲ یا زرد و با شماره‌ی بزرگتر از ۲۶ باشد.

۲۲-۳. یک استاد زیست‌شناسی در کارهای تحقیقاتی اش دو دستیار دارد. احتمال اینکه دستیار مسن‌تر در روز معینی غیبت کند 0.08 ، و احتمال اینکه دستیار جوان‌تر در روز معینی غیبت کند 0.05 ، و احتمال اینکه هر دو در روز معینی غایب باشند 0.02 است. پیدا کنید احتمال اینکه

الف) حداقل یکی از هر دو دستیار در روز معینی غایب باشند،

ب) حداقل یکی از دو دستیار در روز معینی غایب نباشد،

ج) فقط یکی از دو دستیار در روز معینی غایب باشد.

۷-۳ احتمال شرطی

در اکثر مواقع لازم می‌آید که احتمال پیشامدی مانند A ، که با پیشامدی مانند B مربوط باشد، بعد از اطلاع از وقوع یا عدم وقوع پیشامد B ، اصلاح گردد. به بیانی دیگر، ممکن است وقوع پیشامد B بر احتمال وقوع پیشامدی چون A تاثیر داشته باشد. بنابراین کسب اطلاعات درباره‌ی جنبه‌ای از نتایج آزمایش، ممکن است تجدیدنظر در احتمال پیشامدی را که مربوط به جنبه‌ی دیگری از نتایج است، ایجاد کند. احتمال تجدیدنظر شده‌ی A ، وقتی معلوم شود که B رخ داده است، احتمال شرطی A به شرط B نامیده و با $P(A|B)$ نشان داده می‌شود.

برخی از ایده‌هایی که به احتمال‌های شرطی مربوط می‌شوند، در مثال زیر تشریح شده‌اند.

مثال ۲۶-۳. در یک تحقیق از عملکرد بانکی، پنجاه شعبه مورد مطالعه قرار گرفته‌اند و نتایج در جدول زیر خلاصه شده‌اند:

| | | سرویس‌دهی به مشتریان | |
|--------------|--------------|----------------------|------|
| | | خوب | ضعیف |
| نوع پذیرش | حداقل ۱۰ سال | ۱۶ | ۴ |
| | | کمتر از ۱۰ سال | ۱۰ |

اگر فردی به تصادف یکی از این شعب را انتخاب کند، احتمال اینکه شعبه سرویس‌دهی خوبی به مشتریان ارائه دهد چقدر است؟ همچنین اگر شخصی به تصادف یکی از شعب را انتخاب کند که حداقل بیش از ده سال سابقه داشته باشد، احتمال اینکه شعبه‌ی انتخابی به مشتریان خوب سرویس‌دهی کند چقدر است؟ منظور از «تصادف» این است که در هر حالت، تمام انتخاب‌های ممکن هم‌شانس‌اند.

حل. با فرض همشانسی همه‌ی مشتریان، فرمول قضیه‌ی ۲-۳ را به کار می‌بریم. اگر G ، انتخاب شعبه را نشان دهد که سرویس خوب ارائه می‌دهد و اگر فرض کنیم $n(G)$ تعداد اعضای G را نشان دهد، و $n(S)$ تعداد اعضای فضای نمونه‌ای باشد، به دست می‌آوریم

$$P(G) = \frac{n(G)}{n(S)} = \frac{۱۶+۱۰}{۵۰} = ۰/۵۶.$$

این پاسخ سؤال اوّل است.

برای پاسخ سؤال دوّم، ما خود را به فضای نمونه‌ای کوچکتری که شامل سطر اوّل جدول است، یعنی به $۱۶+۱۰=۲۰$ شعب که حداقل ده سال سابقه دارند محدود می‌کنیم. از ۲۰ شعبه، ۱۶ شعبه سرویس خوب ارائه می‌دهند، و به دست می‌آوریم

$$P(G|T) = \frac{۱۶}{۲۰} = ۰/۸۰,$$

که در آن T ، انتخاب شعبه‌ای را نشان می‌دهد که حداقل ده سال سابقه دارد. این پاسخ سؤال دوّم است و همان‌طور که انتظار داشتیم، $P(G|T)$ خیلی از $P(G)$ بزرگتر است. چون صورت $P(G|T) = ۱۶$ برابر $P(G \cap T) = n(T \cap G)$ ، یعنی تعداد شعبی است که حداقل ده سال سابقه دارند و سرویس‌دهی خوبی دارند، و مخرج برابر $n(T)$ ، یعنی شعبی است که حداقل ده سال سابقه‌ی کار دارند. در واقع می‌توانیم آن را به صورت نمادی بنویسیم

$$P(G|T) = \frac{n(T \cap G)}{n(T)}.$$

پس، اگر صورت و مخرج را به $n(S)$ ، تعداد کل شعب موردنظر، تقسیم کنیم، به دست می‌آوریم

$$P(G | T) = \frac{\frac{n(T \cap G)}{n(S)}}{\frac{n(T)}{n(S)}} = \frac{P(T \cap G)}{P(T)},$$

ولذا احتمال شرطی $P(G | T)$ را برحسب دو احتمالی که برای تمام فضای نمونه‌ای S تعریف شده‌اند بیان کرده‌ایم.

با تعمیم این مثال، اکنون تعریف احتمال شرطی را به صورت زیر ارائه می‌دهیم.

تعریف ۲-۳. اگر A و B دو پیشامد دلخواه در فضای نمونه‌ای S باشند و $P(A) \neq 0$ ، احتمال شرطی B به شرط A به صورت زیر تعریف می‌شود

$$P(B | A) = \frac{P(A \cap B)}{P(A)}.$$

مثال ۲۷-۳. با مراجعه به مثال ۲۶-۳، احتمال اینکه یکی از شعب با سابقه‌ی کمتر از ده سال، سرویس خوبی به مشتریان ارائه دهد چقدر است؟

حل. چون $P(T') = \frac{10+20}{50} = 0/20$ و $P(T' \cap G) = \frac{10}{50} = 0/20$ ، اگر مقادیر را در فرمول قراردهیم نتیجه می‌دهد که

$$P(G|T') = \frac{P(T' \cap G)}{P(T')} = \frac{0/20}{0/20}.$$

گرچه ما $P(B | A)$ را با مثالی که در آن تمام حالت‌ها هم‌شansas بودند معرفی کردیم، اما این مطلب، شرطی برای به کار بردن تعریف نیست.

مثال ۲۸-۳. با رجوع به تاس ناهمگن مثال ۲۰-۳، احتمال اینکه شماره‌ای که ظاهر

می‌شود مربع کامل باشد چقدر است؟ همچنین به فرض اینکه این شماره بزرگتر از ۳ باشد، احتمال اینکه مربع کامل باشد چقدر است؟

حل. اگر A این پیشامد باشد که شماره‌ی ظاهر شده‌ی بزرگتر از ۳ است، و B این پیشامد باشد که شماره‌ی ظاهر شده مربع کامل است، داریم $\{4, 5, 6\} = A$ ، $\{1, 2\} = B$ ، و

چون احتمال‌های ظاهر شدن ۱، ۲، ۳، ۴، ۵، ۶ با تاس موردنظر $A \cap B = \{4\}$

به ترتیب $\frac{2}{9}, \frac{1}{9}, \frac{1}{9}, \frac{1}{9}, \frac{1}{9}, \frac{1}{9}$ هستند، پاسخ اویین سوال را به صورت

$$P(B) = \frac{2}{9} + \frac{1}{9} = \frac{1}{3}$$

پیدا می‌کنیم. برای تعیین $P(B | A)$ ، ابتدا

$$P(A) = \frac{1}{9} + \frac{2}{9} + \frac{1}{9} = \frac{4}{9} \quad \text{و} \quad P(A \cap B) = \frac{1}{9}$$

را محاسبه می‌کنیم. لذا با قراردادن این دو مقدار در فرمول تعریف ۲-۳، به دست می‌آوریم

$$P(B | A) = \frac{P(A \cap B)}{P(A)} = \frac{\frac{1}{9}}{\frac{4}{9}} = \frac{1}{4}.$$

برای اینکه تحقیق کنیم فرمول ۲-۳ در مثال قبل جوابی «درست» می‌دهد، تنها کافی است در فضای نمونه‌ای کوچک شده‌ی A ، به دو عدد زوج احتمال ۰ و به عدد فرد احتمال ۱ را نسبت دهیم، به قسمی که مجموع این سه احتمال مساوی ۱ باشد. در

این صورت داریم $0 = 0 + 0 + 0$ یا $\frac{1}{4} = 0 + \frac{1}{2} + \frac{1}{4}$ ، و لذا مثل قبل $P(B | A) = \frac{1}{4}$

با ضرب طرفین فرمول تعریف ۲-۳ در $P(A)$ ، قاعده‌ی ضرب زیر را به دست می‌آوریم.

قضیه‌ی ۹-۳. اگر A و B دو پیشامد دلخواه در فضای نمونه‌ای S باشند و $P(A) \neq 0$ ، آنگاه

$$P(A \cap B) = P(A) P(B | A).$$

با استفاده از کلمات، احتمال اینکه A و B با هم رخده‌ند برابر حاصل ضرب احتمال A در احتمال شرطی B به شرط A است. در صورتی دیگر، اگر $P(B) = 0$ ، احتمال آنکه A و B ، هر دو رخده‌ند، برابر با حاصل ضرب احتمال B و احتمال شرطی A به شرط B است. به صورت نمادی

$$P(A \cap B) = P(B) P(A | B).$$

برای به دست آوردن این قاعده حاصل ضرب، در فرمول ۹-۳ جای A و B را عوض و از واقعیت $A \cap B = B \cap A$ استفاده می کنیم.

مثال ۲۹-۳. اگر از محموله‌ی ۲۴۰ صفحه‌ی کلید کامپیوتر، که ۱۵ تایی آن معیوب است، به تصادف و به توالی دو تای آنها را برداریم، چقدر احتمال دارد که هردو معیوب باشند؟ حل. اگر احتمال انتخاب‌ها را برابر فرض کنیم احتمال اینکه اولین صفحه کلید معیوب باشد $\frac{15}{240}$ ، و احتمال اینکه دومین صفحه کلید، به شرط معیوب بودن اولی، معیوب باشد برابر $\frac{14}{239}$ است. لذا احتمال اینکه هردو معیوب باشند $= \frac{15}{240} \times \frac{14}{239} = \frac{7}{1912}$ است. البته فرض بر این است که انتخاب‌ها بدون جایگذاری است، یعنی اینکه قبل از انتخاب صفحه کلید دوم، صفحه کلید اول به جای خود گذاشته نمی شود.

مثال ۳۰-۳. پیدا کنید احتمال اینکه به تصادف، دو یک به توالی از یک دسته کارت ۵۲ تایی کشیده شوند.

- الف) اگر بدون جایگذاری انتخاب کنیم،
- ب) اگر با جایگذاری انتخاب کنیم.

حل. الف) اگر قبل از کشیدن کارت دوم، کارت اول را به جای خود برنگردانیم، احتمال به دست آوردن متوالی دو یک برابر است با

$$\frac{4}{52} \times \frac{3}{51} = \frac{1}{221}.$$

ب) اگر قبل از کشیدن کارت دوم، کارت اول را به جای خود برگردانیم، احتمال متناظر برابر است با

$$\frac{4}{52} \times \frac{4}{51} = \frac{1}{169}.$$

در وضعیت‌هایی که در دو مثال قبل توصیف شدند، بین دو پیشامد A و B ، یک ترتیب زمانی قطعی وجود دارد. به طور کلی، وقتی می‌نویسیم $P(B|A)$ یا $P(A|B)$ نیاز به چنین ترتیبی نیست. برای مثال، می‌توانستیم احتمال آن را بخواهیم که اوئین کارت کشیده شده عدد یک باشد به شرط آنکه دو مین کارت کشیده شده (بدون جایگذاری) نیز عدد یک باشد - پاسخ به این سوال نیز $\frac{3}{51}$ خواهد داد.

قضیه‌ی ۹-۳ را می‌توان به سهولت تعمیم داد، به قسمی که بتوان آن را برای بیش از دو پیشامد به کار برد. مثلاً برای سه پیشامد داریم:

قضیه‌ی ۱۰-۳. اگر A ، B ، و C سه پیشامد دلخواه در فضای نمونه‌ای S باشند، به قسمی که $P(A \cap B) \neq 0$ و $P(A) \neq 0$ ، آنگاه

$$P(A \cap B \cap C) = P(A) P(B|A) P(C|A \cap B).$$

برهان. با نوشتن $A \cap B \cap C$ به صورت $(A \cap B) \cap C$ و با دوبار استفاده از فرمول ۹-۳، به دست می‌آوریم

$$\begin{aligned} P(A \cap B \cap C) &= P((A \cap B) \cap C) \\ &= P(A \cap B) P(C|A \cap B) \\ &= P(A) P(B|A) P(C|A \cap B). \end{aligned}$$

مثال ۳۱-۳. بسته‌ای حاوی ۲۰ فیوز است که ۵ تای آنها معیوب‌اند. اگر به تصادف ۳ فیوز متوالیاً و بدون جایگذاری از بسته‌ی مزبور انتخاب کنیم، احتمال اینکه هر ۳ فیوز معیوب باشند چقدر است؟

حل. اگر A پیشامد معیوب‌بودن فیوز اول، B پیشامد معیوب‌بودن فیوز دوم، و C پیشامد معیوب‌بودن فیوز سوم باشد، آنگاه $P(A) = \frac{5}{20}$ ، $P(B|A) = \frac{4}{19}$ ، $P(C|A \cap B) = \frac{3}{18}$ ، و اگر آن را در فرمول قراردهیم نتیجه می‌شود

$$P(A \cap B \cap C) = \frac{5}{20} \times \frac{4}{19} \times \frac{3}{18} = \frac{1}{114}.$$

تعمیم بیشتر قضیه‌های ۹-۳ و ۱۰-۳ به k پیشامد نیز امکان‌پذیر است.

تمرین

۲۳-۳. نشان دهید که احتمال‌های شرطی در سه اصل موضوع احتمال صادقند، به عبارت دیگر نشان دهید که اگر $P(B) \neq 0$ آنگاه

$$\text{الف)} P(A|B) \geq 0$$

$$\text{ب)} P(B|B) = 1$$

ج) برای هر دنباله از پیشامدهای دوبهدو ناسازگار A_1, A_2, \dots, A_p

$$P(A_1 \cup A_2 \cup \dots \cup A_p | B) = P(A_1 | B) + P(A_2 | B) + \dots + P(A_p | B).$$

۲۴-۳. به کمک مثال‌های عدی نشان دهید که $P(B|A) + P(B|A') = 1$.

الف) ممکن است مساوی ۱ باشد،

ب) لازم نیست مساوی ۱ باشد.

۲۵-۳. نشان دهید که به شرط $P(A \cap B \cap C) \neq 0$

$$P(A \cap B \cap C) = P(A)P(B|A)P(C|A \cap B)P(D|A \cap B \cap C).$$

۲۶-۳. برای سه پیشامد A, B, C ، با داشتن شرایط

$$P(C|A \cap B) = P(C|B) \quad \text{و} \quad P(A \cap B \cap C) \neq 0.$$

نشان دهید که $P(A|B \cap C) = P(A|B)$

۲۷-۳. برای تصدی شغلی در بخش خبری یک ایستگاه تلویزیون، ۹۰ داوطلب وجود دارند. بعضی از آنها فارغ‌التحصیل دانشگاه‌اند، و برخی نیستند، برخی از آنها حداقل سه سال سابقه کار دارند و برخی ندارند. جدولی تفکیکی به صورت زیر است:

| | فارغ‌التحصیل دانشگاه | دانشگاه نرفته |
|-----------------------|----------------------|---------------|
| حداقل با سه سال تجربه | ۱۸ | ۹ |
| کمتر از سه سال تجربه | ۳۶ | ۲۷ |

اگر ترتیب مصاحبه‌ی داوطلبان به وسیله‌ی مدیر ایستگاه، تصادفی باشد، G پیشامدی است که اوئین داوطلب که مصاحبه می‌شود فارغ‌التحصیل دانشگاه است، و T پیشامدی است که اوئین داوطلبی که مصاحبه می‌شود حداقل سه سال سابقه‌ی کار دارد.

هر یک از احتمال‌های زیر را مستقیماً از درایه‌ها و حاصل‌جمع‌های سطري و ستونی جدول تعیین کنید.

الف) $P(G)$

ب) $P(T')$

ج) $P(G \cap T)$

د) $P(G' \cap T')$

ه) $P(T | G)$

و) $P(G' | T')$

۲۸-۳. نتایج تمرین قبل را برای تحقیق درستی تساوی‌های زیر به کار ببرید.

$$\text{الف)} P(T | G) = \frac{P(G \cap T)}{P(G)}$$

$$\text{ب)} P(G' | T') = \frac{P(G' \cap T')}{P(T')}$$

۲۹-۳. نشان دهید که اگر $P(A | B) = P(A)$ و $P(B | A) = P(B)$ آنگاه $P(B) \neq 0$.

۳۰-۳. جعبه‌ای شامل ۱۴ مهره‌ی سفید و ۸ مهره‌ی سیاه است. ۵ مهره به تصادف و بدون جایگذاری از این جعبه و بدون نگاه کردن به رنگ مهره‌ها کنار گذاشته می‌شود.

الف) اگر یک مهره مجدداً از این جعبه انتخاب شود، احتمال اینکه مهره‌ی انتخابی سفید باشد چقدر است؟

ب) اگر دو مهره مجدداً به تصادف از این جعبه انتخاب شود، احتمال اینکه دقیقاً شامل یک مهره‌ی سفید باشد چقدر است؟

۳۱-۳. جعبه‌ای شامل ۸ مهره‌ی سفید و ۱۴ مهره‌ی سیاه است. یک نمونه‌ی تصادفی ۱۴ تایی بدون جایگذاری از این جعبه انتخاب می‌شود.

الف) احتمال این را که مهره‌های دوام و سوم سفید باشند در حالی که بدانیم نمونه دقیقاً شامل ۳ مهره‌ی سفید است به دست آورید.

ب) احتمال این را که مهره‌ی اوّل سفید باشد در حالی که بدانیم نمونه شامل ۲ مهره‌ی سفید است به دست آورید.

۳۲-۳. سه جعبه را در نظر بگیرید، به طوری که در جعبه‌ی اوّل ۳ مهره‌ی سفید و ۵ مهره‌ی سیاه، در جعبه‌ی دوّم ۶ مهره‌ی سفید و ۴ مهره‌ی سیاه و در جعبه سوم ۲ مهره‌ی سفید و ۴ مهره‌ی سیاه است. اگر یک مهره به تصادف از هر جعبه انتخاب شود، احتمال اینکه مهره‌ی انتخابی از جعبه‌ی اوّل سفید باشد به شرط آن‌که دقیقاً ۲ مهره‌ی سفید انتخاب شده باشد، چقدر است؟

۳-۸ پیشامدهای مستقل

دو پیشامد A و B را مستقل می‌گویند اگر رخدادن یا رخدادن هریک از آنها در احتمال رخدادن دیگری تأثیری نداشته باشد. برای مثال در انتخاب‌های باجایگذاری در هر مرحله نتایج از هم مستقل هستند. به طور نمادی، دو پیشامد A و B مستقل‌اند اگر

$$P(A|B)=P(A), \quad P(B|A)=P(B),$$

و موقعی که هر دو احتمال شرطی موجود باشند، یعنی وقتی $P(A)$ و $P(B)$ صفر نباشند، می‌توان نشان داد که هر یک از این برابری‌ها، دیگری را نتیجه می‌دهد. حال اگر در فرمول قضیه‌ی ۳-۹، $P(B|A)$ را به جای $P(B|A)$ قرار دهیم، به دست می‌آوریم

$$P(A \cap B) = P(A) P(B|A) = P(A) P(B),$$

و این نتیجه را به عنوان تعریف صوری استقلال به کار خواهیم برد.

تعریف ۳-۳. دو پیشامد B و C مستقل‌اند اگر و تنها اگر

$$P(A \cap B) = P(A) P(B).$$

اگر دو پیشامد مستقل نباشند، می‌گویند وابسته‌اند. در به دست آوردن فرمول تعریف ۳-۳ فرض کردیم که $P(B|A)$ تعریف شده است و لذا فرض کردیم که $P(A) \neq 0$.

مثال ۳۲-۳. سکه‌ای همگن را سه بار پرتاب می‌کنیم فضای نمونه‌ای عبارت است از

$$S = \{TTT, TTH, THT, HTT, THH, HTH, HHT, HHH\}.$$

فرض می‌کنیم که کلیه‌ی عناصر S هم‌شانس هستند. اگر A پیشامد رخدادن شیر در هر یک از دو پرتاب اول، B پیشامد رخدادن خط در پرتاب سوم، و C پیشامد رخدادن دقیقاً دو خط در سه پرتاب باشد، نشان دهید که

الف) پیشامدهای A و B مستقل‌اند،

ب) پیشامدهای B و C وابسته‌اند.

حل.

$$A = \{HHH, HHT\},$$

$$B = \{HHT, HTT, THT, TTT\},$$

$$C = \{HTT, THT, TTH\},$$

$$A \cap B = \{HHT\},$$

$$B \cap C = \{HTT, THT\}.$$

فرض هم‌شانس بودن همه‌ی هشت عنصر نتیجه می‌دهد که

$$P(B) = \frac{1}{2}, P(A) = \frac{1}{4}, P(B \cap C) = \frac{1}{4}, P(A \cap B) = \frac{1}{8}, P(C) = \frac{3}{8}$$

الف) چون $P(A)P(B) = \frac{1}{4} \times \frac{1}{2} = \frac{1}{8} = P(A \cap B)$ پیشامدهای A و B مستقل‌اند.

ب) چون $P(B)P(C) = \frac{1}{2} \times \frac{3}{8} = \frac{3}{16} \neq P(B \cap C)$ پیشامدهای B و C مستقل نیستند.

می‌توان نشان داد که اگر A و B مستقل باشند، آنگاه A و B' ، A' و B ، و A' و B' هم مستقل‌اند. یکی از حالت‌های ذکر شده را اثبات می‌کنیم.

قضیه‌ی ۱۱-۳. اگر دو پیشامد A و B مستقل باشند، آنگاه A و B' نیز مستقل‌اند. برهان. از آنجا که $A \cap B = (A \cap B') \cup (A \cap B')$ ناسازگارند و چون A و B بنابر فرض مستقل‌اند، داریم

$$\begin{aligned} P(A) &= P((A \cap B) \cup (A \cap B')) \\ &= P(A \cap B) + P(A \cap B') \\ &= P(A)P(B) + P(A)P(B'). \end{aligned}$$

نتیجه می‌شود که

$$\begin{aligned} P(A \cap B') &= P(A) - P(A)P(B) \\ &= P(A)(1 - P(B)) \\ &= P(A)P(B'), \end{aligned}$$

ولذا A و B' مستقل‌اند.

در تمرین‌ها خواسته شده است که نشان دهید که اگر A و B مستقل باشند، آنگاه

A' و B' نیز مستقل‌اند، و اگر A و B وابسته باشند، آنگاه A' و B' وابسته‌اند.

برای تعمیم مفهوم استقلال به بیش از دو پیشامد، تعریف زیر را ارائه می‌دهیم.

تعریف ۳-۴. پیشامدهای A_1, A_2, \dots, A_k مستقل‌اند اگر و تنها اگر احتمال اشتراک هر $2, 3, \dots, k$ تا از این پیشامدها مساوی حاصل‌ضرب احتمال‌های مربوط به هر پیشامد باشد.

برای حالتی که سه پیشامد A ، B و C مطرح است شرایط استقلال به صورت زیر خواهد بود:

$$P(A \cap B) = P(A)P(B),$$

$$P(A \cap C) = P(A)P(C),$$

$$P(B \cap C) = P(B)P(C),$$

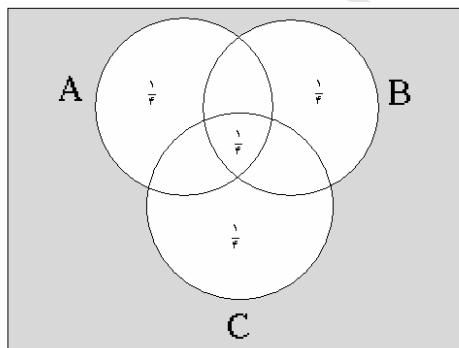
و

$$P(A \cap B \cap C) = P(A)P(B)P(C).$$

قابل توجه است که سه یا چند پیشامد می‌توانند دو به دو مستقل باشند بدون آنکه کلاً مستقل باشند.

مثال ۳-۳. سه پیشامد با احتمال‌های مشخص شده در نمودار ون شکل ۳-۳، آمده است. تحقیق کنید که A و B مستقل‌اند، A و C مستقل‌اند، B و C مستقل‌اند، اما A ، B ، و C مستقل نیستند.

$$\begin{aligned} P(A) &= P(B) = P(C) = \frac{1}{2} && \text{همان‌طور که از نمودار دیده می‌شود،} \\ \text{پس } P(A \cap B \cap C) &= \frac{1}{4} && P(A \cap B) = P(A \cap C) = P(B \cap C) = \frac{1}{4} \\ P(A)P(C) &= \frac{1}{4} = P(A \cap C) && P(A)P(B) = \frac{1}{4} = P(A \cap B) \\ P(A)P(B)P(C) &= \frac{1}{8} \neq P(A \cap B \cap C) && \text{اما } P(B)P(C) = \frac{1}{4} = P(B \cap C) \end{aligned}$$



شکل ۳-۳. نمودار ون برای مثال ۳-۳

ممکن است که $P(A \cap B \cap C) = P(A)P(B)P(C)$ بدون آنکه A ، B ، و C دو به دو مستقل باشند. مثال زیر را ببینید.

مثال ۳-۴. مطلوب است احتمال به دست آوردن
 الف) سه شیر در سه پرتاب تصادفی یک سکه‌ی همگن،
 ب) چهار شش و سپس عددی دیگر در پنج ریختن یک تاس ناهمگن.
 حل. الف) از ضرب احتمال‌های مربوط، به دست می‌آوریم

$$\frac{1}{2} \times \frac{1}{2} \times \frac{1}{2} = \frac{1}{8}.$$

ب) از ضرب احتمال‌های مربوط، به دست می‌آوریم

$$\frac{1}{6} \times \frac{1}{6} \times \frac{1}{6} \times \frac{1}{6} \times \frac{5}{6} = \frac{5}{7776}.$$

تمرین

۳۳-۳. نشان دهید که اگر دو پیشامد A و B مستقل باشند، آنگاه

الف) دو پیشامد A' و B مستقل اند،

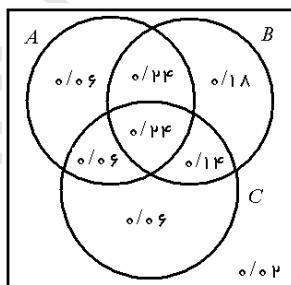
ب) دو پیشامد A' و B' مستقل اند.

۳۴-۳. نشان دهید که اگر پیشامدهای A و B وابسته باشند، آنگاه پیشامدهای A' و B' وابسته اند.

۳۵-۳. با رجوع به شکل ۴-۳ نشان دهید که $P(A \cap B \cap C) = P(A)P(B)P(C)$ الزاماً نتیجه نمی دهد که پیشامدهای A , B , و C همگی دوبه دو مستقلند.

۳۶-۳. با رجوع به شکل ۴-۳ نشان دهید که اگر B از A مستقل و از C مستقل باشد، آنگاه B الزاماً از C مستقل نیست.

۳۷-۳. با رجوع به شکل ۴-۳ نشان دهید که اگر A از B مستقل و A از C مستقل باشد، آنگاه A الزاماً از $B \cup C$ مستقل نیست.



شکل ۴-۳. شکل مورد نیاز برای تمرین های ۳۵-۳، ۳۶-۳، ۳۷-۳

۳۸-۳. اگر سه پیشامد A , B , و C مستقل باشند، نشان دهید که

الف) $A \cap B \cap C$ مستقل اند،

ب) $A \cup B \cup C$ مستقل اند،

۳۹-۳. نشان دهید که برای اینکه k پیشامد مستقل باشند باید $1 - k^k$ شرط برقرار باشد.

۴۰-۳. سکه‌ای به قسمی دستکاری شده است که احتمال آمدن شیر و خط به ترتیب $52/0$ و $48/0$ است. اگر این سکه سه بار پرتاب شود، احتمال به دست آوردن الف) سه شیر،

ب) دو خط و یک شیر با همین ترتیب.

۴۱-۳. اگر پنج کامیون از ده کامیون حمل بار شرکتی، استانداردهای مربوط به میزان دود اگزوز را نداشته باشند و سه تا از آنها برای بررسی انتخاب شوند، احتمال آنکه هیچ از یک کامیون‌های منتخب استانداردهای مذکور را نداشته باشد چقدر است؟

۴۲-۳. اگر شخصی از ۱۵ سکه‌ی طلای موجودی یک صراف که ۶ تای آنها تقلبی است ۴ سکه به تصادف انتخاب کند، احتمال آنکه سکه‌های منتخب همه تقلبی باشند چقدر است؟

۴۳-۳. یک فروشگاه بزرگ، ماهیانه، صورت حساب بدھی مشتریانی را که حق نسیه بردن دارند برای آنها می‌فرستند. فروشگاه دریافتہ است مشتری که بلاfacسله در یک ماه بدھی خود را می‌پردازند با احتمال $9/0$ ماه بعد نیز بلاfacسله بدھی خود را می‌پردازند، اما اگر در یک ماه بدھی خود را بلاfacسله نپردازد احتمال اینکه در ماه بعد بلاfacسله بدھی اش را بپردازد $4/0$ است.

الف) مطلوب است احتمال آنکه مشتری که در ماهی بلاfacسله بدھی خود را پرداخته است در سه ماه بعد نیز بلاfacسله بدھی اش را بپردازد؟

ب) مطلوب است احتمال آنکه مشتری که در ماهی بلاfacسله بدھی خود را نپرداخته است در دو ماه بعد نیز بدھی اش را فوراً نپردازد و آنگاه ماه بعد از آن فوراً بدھی خود را بپردازد.

۹-۳ قضیه‌ی بیز

مسائل زیادی وجود دارند که در آنها برآمد نهایی یک آزمایش به آنچه در مراحل میانی

مختلف رخ می‌دهند بستگی دارد. مثال زیر، مثالی ساده است که در آن، مرحله‌ای میانی وجود دارد که متشکل از دو شق است.

مثال ۳۵-۳. کار تکمیل بنای بزرگراهی ممکن است به علت اعتصاب به تأخیر افتد. فرض کنید احتمال این که اعتصابی رخ دهد $0/60$ ، احتمال این که اگر اعتصابی نباشد کار به موقع انجام شود $0/85$ ، و احتمال این که اگر اعتصابی باشد کار به موقع انجام شود، $0/35$ باشد. احتمال این که کار به موقع انجام شود چقدر است؟ حل. اگر A پیشامد تکمیل بهموقع کار، و B پیشامد وقوع اعتصاب باشد، اطلاعات داده شده را می‌توان چنین نوشت:

$$P(B) = 0/60, \quad P(A|B) = 0/35, \quad P(A|B') = 0/85.$$

با استفاده از قضیه‌های احتمال، و این واقعیت که $A \cap B$ و $A \cap B'$ ناسازگارند، و صورت دیگر قاعده‌ی ضرب، می‌توانیم بنویسیم

$$\begin{aligned} P(A) &= P((A \cap B) \cup (A \cap B')) \\ &= P(A \cap B) + P(A \cap B') \\ &= P(B) P(A|B) + P(B') P(A|B'). \end{aligned}$$

در این صورت، با قراردادن مقادیر عددی، نتیجه می‌شود که

$$P(A) = (0/60)(0/35) + (0/60)(0/85) = 0/55.$$

تعمیمی از این نوع وضعیت، موردی است که در آن مرحله‌ی میانی دارای k شق مختلف است (که رخدادن آنها را با B_1, B_2, \dots, B_k نشان می‌دهیم). این تعیم، به قضیه‌ی زیر که گاهی آن را قاعده‌ی احتمال کل یا قاعده‌ی حذف می‌نامند نیاز دارد.

قضیه‌ی ۳۶-۳. اگر پیشامدهای B_1, B_2, \dots, B_k افزایی از فضای نمونه‌ای S را تشکیل دهند و به ازای هر i ، $P(B_i) \neq 0$ ، آنگاه برای هر پیشامد A در S

$$P(A) = \sum_{i=1}^k P(B_i) P(A|B_i).$$

در اینجا B_i ها افزایی از فضای نمونه‌ای را تشکیل می‌دهند، اگر دو بهدو ناسازگار بوده و اجتماع آنها مساوی S باشد.

مثال ۳-۳. اعضای یک شرکت مشاور از سه آژانس، اتومبیل کرایه می‌کنند: از آژانس ۱ به میزان ۵۰ درصد، از آژانس ۲ به میزان ۴۰ درصد، و از آژانس ۳ به میزان ۱۰ درصد. اگر ۹ درصد از اتومبیل‌های آژانس ۱، ۲۰ درصد از اتومبیل‌های آژانس ۲، و ۶ درصد از اتومبیل‌های آژانس ۳ به تنظیم موتور نیاز داشته باشند، احتمال اینکه یک اتومبیل کرایه‌ای که به شرکت مشاور تحویل شده است به تنظیم موتور نیاز داشته باشد چقدر است؟

حل. اگر A پیشامد نیاز داشتن اتومبیلی به تنظیم موتور و B_1, B_2, B_3 به ترتیب پیشامدهای تعلق یک اتومبیل کرایه شده به آژانس‌های ۱، ۲، ۳ باشند، داریم

$$\begin{aligned} P(B_1) &= 0.10, \quad P(B_2) = 0.40, \quad P(B_3) = 0.50, \\ P(A|B_1) &= 0.06, \quad P(A|B_2) = 0.20, \quad P(A|B_3) = 0.09. \end{aligned}$$

اگر این مقادیر را در فرمول قضیه‌ی ۱۲-۳ قراردهیم، نتیجه می‌شود که

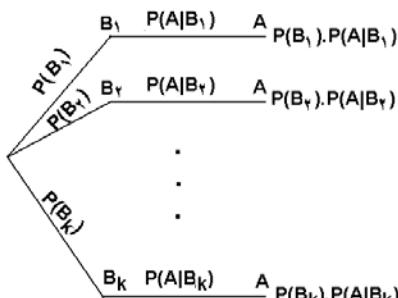
$$P(A) = (0.06)(0.10) + (0.20)(0.40) + (0.09)(0.50) = 0.14.$$

پس، ۱۴ درصد از اتومبیل‌هایی که به شرکت تحویل می‌شوند به تنظیم موتور نیاز دارند. با رجوع به مثال قبلی، فرض کنیم سوال زیر مورد نظر باشد: اگر یک اتومبیل کرایه که به شرکت مشاور تحویل شده است نیاز به تنظیم موتور داشته باشد، چقدر احتمال دارد که متعلق به آژانس ۲ باشد؟ برای پاسخ دادن به این نوع سوال‌ها، به قضیه‌ی زیر که قضیه‌ی بیز نامیده می‌شود نیاز داریم:

قضیه‌ی ۱۳-۳. اگر پیشامدهای B_1, B_2, \dots, B_k افزایی از فضای نمونه‌ای S را تشکیل دهند و به ازای هر i ، $P(B_i) \neq 0$ ، $i = 1, 2, \dots, k$ ، آنگاه برای هر پیشامد A در S ، به قسمی که $P(A) \neq 0$ ،

$$P(B_r | A) = \frac{P(B_r) P(A | B_r)}{\sum_{i=1}^k P(B_i) P(A | B_i)}, \quad r = 1, 2, \dots, k.$$

با استفاده از کلمات، احتمال آنکه از راه شاخه‌ی r نمودار درختی شکل ۵-۳ به پیشامد A بررسیم، به شرط آنکه از راه یکی از k شاخه‌ی نمودار به A رسیده باشیم، برابر نسبت احتمال مربوط به r این شاخه به مجموع احتمال‌های مربوط به کل k شاخه‌ی درخت است.



شکل ۵-۳. نمودار درختی برای قضیه‌ی بیز

برهان. مطابق تعریف احتمال شرطی، می‌نویسیم $P(A \cap B_r) / P(A) = P(B_r | A) \cdot P(A)$. لازم است فقط به جای $P(A \cap B_r)$ مقدار $P(B_r)P(A | B_r)$ ، و به جای $P(A)$ عبارت فرمول قضیه‌ی ۱۲-۳ را قرار دهیم.

مثال ۳۷-۳. با مراجعه به مثال ۳۶-۳، اگر اتومبیل کرایه‌ای که به شرکت مشاور تحويل شده است به تنظیم موتور نیاز داشته باشد، احتمال اینکه متعلق به آژانس ۲ باشد چقدر است؟

حل. با توجه به فرمول قضیه‌ی ۱۳-۳، قرار می‌دهیم و داریم

$$\begin{aligned} P(B_2 | A) &= \frac{(0/30)(0/20)}{(0/60)(0/90) + (0/30)(0/20) + (0/10)(0/06)} \\ &= \frac{0/060}{0/120} = 0/5. \end{aligned}$$

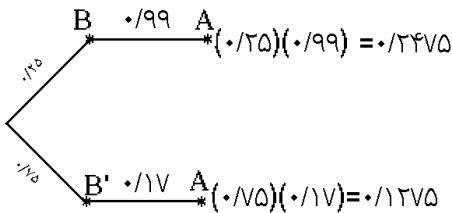
توجه کنید که گرچه فقط ۳۰ درصد اتومبیل‌هایی که به شرکت کرایه داده می‌شوند متعلق به آژانس ۲ است، ولی ۵۰ درصد اتومبیل‌هایی که به تنظیم نیاز دارند متعلق به این آژانس‌اند.

مثال ۳۸-۳. در استان معینی، دودزایی ۲۵ درصد تمام اتومبیل‌ها بیش از حد عادی است. اگر در آزمون دودزایی وسائل نقلیه‌ی استان، اتومبیلی که بیش از حد دود تولید می‌کند با احتمالی برای ۹۹٪ مردود شود و اگر اتومبیلی که دودزایی آن بیش از حد عادی نیست با احتمال ۷۱٪ در این آزمون رد شود، احتمال اینکه اتومبیلی که در آزمون رد می‌شود واقعاً اتومبیلی باشد که دودزایی آن بیش از حد عادی است چقدر است؟

حل. از شکل ۳-۴ درمی‌یابیم که احتمال‌های مربوط به دو شاخه‌ی نمودار درختی عبارت‌اند از $0/2475 = 0/99$ و $0/1275 = 0/017$ ، پس احتمال اینکه اتومبیلی که در آزمون رد می‌شود واقعاً اتومبیلی باشد که بیش از حد دود ایجاد می‌کند برابر است با

$$\frac{0/2475}{0/2475 + 0/1275} = 0/66.$$

البته می‌توانستیم این نتیجه را از قراردادن مستقیم در فرمول قضیه‌ی بیز، بدون نمودار درختی نیز به دست آوریم.



شکل ۳-۴. نمودار درختی برای مثال ۳۸-۳

گرچه قضیه‌ی بیز از اصول احتمال و تعریف احتمال شرطی نتیجه می‌شود، ولی این قضیه موضوع مباحثه‌های دامنه‌داری بوده است. درباره‌ی اعتبار قضیه‌ی بیز نمی‌توان سوالی مطرح کرد، اما بحث‌های مفصلی درباره‌ی تعبیر احتمال‌های پیشین ($P(B_i)$) مطرح شده‌اند. همچنین مقدار نسبتاً زیادی از هاله‌ی اسراری که قضیه‌ی بیز را احاطه کرده‌است، ناشی از این واقعیت است که این قضیه نظیر مثال ۳-۸، نوعی استدلال «قهرمانی» یا «معکوس»، یعنی استدلال «از معلول به علت» را موجب می‌شود. مثلاً در مثال ۳-۸، رد شدن از آزمون، یک معلول، و دوزایی بیش از حد، یک علت ممکن است.

تمرین

۴۴-۳. یک کارخانه خودروسازی، بلبرینگ‌های موردنیاز خود را از سه تولیدکننده به شرح: ۵۰% از تولیدکننده ۱، ۳۰% از تولیدکننده ۲ و ۲۰% از تولیدکننده ۳، تهیه می‌کند. براساس تجربه‌ی گذشته شرکت خودروسازی می‌داند که استاندارد کترل کیفیت این سه تولیدکننده یکسان نیست، به‌طوری که ۲% از محصولات تولیدی ۱، ۳% از محصولات تولیدی ۲، و ۴% از محصولات تولیدی ۳، معیوب هستند. چه نسبتی از بلبرینگ‌های خریداری‌شده‌ی کارخانه خودروسازی، معیوب هستند.

۴۵-۳. در تمرین ۴۴-۳، اگر بلبرینگ انتخابی معیوب باشد، احتمال اینکه بلبرینگ از محصولات تولیدی ۲ باشد چقدر است؟

۴۶-۳. جعبه‌ی I شامل ۳ مهره سفید و ۵ مهره سیاه است و جعبه‌ی II شامل یک مهره‌ی سفید و یک مهره‌ی سیاه است. دو مهره به تصادف از جعبه‌ی I انتخاب و بدون نگاه‌کردن به جعبه‌ی II منتقل می‌کنیم و یک مهره به تصادف از جعبه‌ی II انتخاب می‌کنیم.

الف) احتمال اینکه مهره‌ی سفید از جعبه II استخراج شده باشد چقدر است؟

ب) احتمال انتقال مهره‌های سفید از جعبه‌ی I به II ، به شرط مشاهده‌ی مهره‌ی سفید از جعبه‌ی II ، چقدر است؟

۴۷-۳. جعبه‌ی I شامل ۲ مهره‌ی سفید و ۴ مهره‌ی سیاه است و جعبه‌ی II شامل ۳ مهره‌ی سفید و ۲ مهره‌ی سیاه است. یک جعبه به تصادف انتخاب و دو مهره از این جعبه انتخاب می‌شود.

الف) احتمال اینکه هر دو مهره سفید باشند، چقدر است؟

ب) احتمال اینکه جعبه‌ی II انتخاب شود به شرط آنکه هر دو مهره سفید باشند چقدر است؟

فصل چهارم

متغیرهای تصادفی، توزیع‌ها، و چگالی‌های احتمال

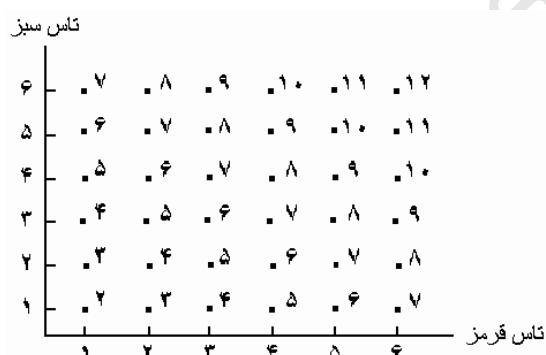
۴-۱ متغیر تصادفی

در بیشتر مسأله‌های متضمن احتمال‌ها، ما تنها به جنبه‌ای خاص (یا به دو یا چند جنبه‌ی خاص) برآمدهای آزمایش‌ها توجه داریم. برای مثال، وقتی یک جفت تاس را می‌ریزیم، معمولاً فقط مجموع دو شماره‌ای که ظاهر می‌شوند مورد توجه است، و نه برآمد هر تاس، وقتی با زن و شوهری که به تصادف انتخاب شده‌اند مصاحبه می‌کنیم ممکن است تعداد افراد خانواده و درآمد توأم آنها مورد توجه باشند، و نه تعداد سال‌هایی که با هم ازدواج کرده‌اند یا پسنداز کل آنها، و بالاخره وقتی از لامپ‌های روشنایی که در سطح انبوه تولید می‌شوند نمونه می‌گیریم ممکن است دوام یا میزان روشنایی آنها مورد توجه باشند و نه بهای آنها.

در هریک از مثال‌های بالا توجه ما به اعدادی است که با برآمدهای یک آزمایش شناسی همراه‌اند، یعنی توجه ما به مقادیری است که آنچه اصطلاحاً متغیرهای تصادفی خوانده‌می‌شوند اختیار می‌کنند. تعداد افراد خانواده‌ی زوجی که به تصادف انتخاب شده‌اند و درآمد توأم آنها متغیرهای تصادفی‌اند، و دوام و میزان روشنایی لامپی که به تصادف برای بررسی انتخاب می‌شود نیز متغیرهای تصادفی‌اند.

برای تصریح بیشتر، شکل ۱-۳ را در نظر بگیرید که فضای نمونه‌ای را برای آزمایش پرتاب یک جفت تاس همگن تصویر می‌کند، و فرض کنید که هریک از ۳۶

برآمد ممکن دارای احتمال $\frac{1}{6^6}$ است. ولی توجه کنید که در شکل ۱-۳ به هر نقطه، عددی وابسته کرده‌ایم. برای مثال، به نقطه‌ی (۱,۱) عدد ۲، به نقطه‌ی (۱,۵) عدد ۶، به نقطه‌ی (۶,۲) عدد ۸، و قس‌علی‌هذا. به وضوح، ما به هر نقطه مقداری از یک متغیر تصادفی، یعنی مجموع متناظر با دو عددی را که در ریختن دو تاس ظاهر می‌شوند، نسبت می‌دهیم.



شکل ۱-۴. مجموع دو عددی که در ریختن یک جفت تاس ظاهر می‌شوند

چون «نسبت دادن یک عدد به هر نقطه‌ی (عنصر) فضای نمونه‌ای» صرفاً راه دیگری برای بیان «تعریف تابعی روی نقاط یک فضای نمونه‌ای» است، لذا تعریف زیر را ارائه می‌دهیم:

تعریف ۱-۴. اگر S یک فضای نمونه‌ای و X یک تابع حقیقی - مقدار باشد که روی عناصر S تعریف شده است، آنگاه X **متغیر تصادفی** نامیده می‌شود.

ما در این کتاب، متغیرهای تصادفی را همیشه با حروف بزرگ و مقادیر آنها را با حروف کوچک متناظرشان نشان می‌دهیم. مثلاً برای نشان دادن مقداری از متغیر X ، می‌نویسیم x .

با رجوع به مثال قبل و شکل ۱-۳ مشاهده می‌کنید که متغیر تصادفی X در زیر مجموعه‌ی

$$\{(6,3), (5,4), (4,5), (3,6)\}$$

از فضای نمونه‌ای S ، مقدار ۹ را اختیار می‌کند و می‌نویسیم $X = 9$. لذا، $X = 9$ به

عنوان مجموعه‌ای از عناصر S تعبیر می‌شود که مجموع دو مؤلفه‌ی آنها ۹ است، و یا کلی‌تر، $X = x$ به عنوان مجموعه‌ای از عناصر فضای نمونه‌ای تعبیر می‌شود که برای آنها متغیر تصادفی X مقدار x را اختیار می‌کند. این مطلب ممکن است مبهم به نظر آید، ولی ریاضیدانی را به یاد می‌آورد که به جای اینکه بگوید « $f(x)$ ، مقدار تابع به ازای x است» می‌گوید « $f(x)$ ، تابع x است».

مثال ۴-۱. از جعبه‌ای که محتوی پنج مهره‌ای قوهای و سه مهره‌ی سبز است، به تصادف و متزالیاً دو مهره انتخاب می‌شود. فهرست عناصر فضای نمونه‌ای، احتمال‌های متناظر آنها، و مقادیر متناظر، W ، از متغیر تصادفی W را که تعداد مهره‌های قوهای انتخاب شده را نشان می‌دهد، تهیه کنید.

حل. اگر B و G به ترتیب برای نشان دادن مهره‌ی قوهای و سبز به کار روند، احتمال‌های BB ، GB ، BG ، و GG به ترتیب عبارت‌اند از $\left(\frac{5}{14}\right)\left(\frac{5}{14}\right)$ ، $\left(\frac{5}{14}\right)\left(\frac{3}{14}\right)$ ، $\left(\frac{3}{14}\right)\left(\frac{5}{14}\right)$ ، و $\left(\frac{3}{14}\right)\left(\frac{3}{14}\right)$. نتایج در جدول زیر نشان داده شده‌اند.

| عنصر فضای نمونه‌ای | احتمال | W |
|--------------------|-----------------|-----|
| BB | $\frac{5}{14}$ | ۲ |
| BG | $\frac{15}{56}$ | ۱ |
| GB | $\frac{15}{56}$ | ۱ |
| GG | $\frac{3}{56}$ | ۰ |

همچنین مثلاً، برای احتمال این پیشامد که متغیر تصادفی W مقدار ۲ را اختیار کند، می‌توانیم بنویسیم $P(W=2) = \frac{5}{14}$ و یا احتمال این پیشامد که متغیر تصادفی W مقدار یک را اختیار کند برابر است با $P(W=1) = \frac{15}{56} + \frac{15}{56} = \frac{30}{56}$.

مثال ۴-۲. یک سکه همگن چهار بار پرتاب می‌شود. عناصر فضای نمونه‌ای را که بنا به فرض هم‌شانس‌اند، (منظور از همگن بودن سکه همین است)، و مقادیر متناظر x از متغیر تصادفی X را که تعداد کل شیرهاست فهرست کنید.

حل. اگر H و T به ترتیب برای نمایش دادن شیر و خط به کار روند، نتایج در

جدول زیر نشان داده شده‌اند. مثلاً، برای احتمال این پیشامد که متغیر تصادفی X مقدار ۳ را اختیار کند می‌توانیم بنویسیم $P(X = ۳) = \frac{۴}{۱۶}$.

| عنصر فضای نمونه‌ای | احتمال | x |
|--------------------|----------------|-----|
| HHHH | $\frac{۱}{۱۶}$ | ۴ |
| HHHT | $\frac{۱}{۱۶}$ | ۳ |
| HHTH | $\frac{۱}{۱۶}$ | ۳ |
| HTHH | $\frac{۱}{۱۶}$ | ۳ |
| THHH | $\frac{۱}{۱۶}$ | ۳ |
| HHTT | $\frac{۱}{۱۶}$ | ۲ |
| HTHT | $\frac{۱}{۱۶}$ | ۲ |
| THHT | $\frac{۱}{۱۶}$ | ۲ |
| HTTH | $\frac{۱}{۱۶}$ | ۲ |
| THTH | $\frac{۱}{۱۶}$ | ۲ |
| TTHH | $\frac{۱}{۱۶}$ | ۲ |
| HTTT | $\frac{۱}{۱۶}$ | ۱ |
| THTT | $\frac{۱}{۱۶}$ | ۱ |
| TTHT | $\frac{۱}{۱۶}$ | ۱ |
| TTTH | $\frac{۱}{۱۶}$ | ۱ |
| TTTT | $\frac{۱}{۱۶}$ | ۰ |

این واقعیت که تعریف ۱-۴ به توابع حقیقی محدود شده است هیچ محدودیتی را تحمیل نمی‌کند. اگر اعدادی که می‌خواهیم به برآمدهای یک آزمایش نسبت دهیم اعداد مختلط باشند، همیشه می‌توانیم قسمت‌های حقیقی و موهومی را جدا از هم به عنوان مقادیری که دو متغیر تصادفی اختیار می‌کنند در نظر بگیریم. همچنین، اگر بخواهیم برآمدهای یک آزمایش را به صورت کمی توصیف کنیم، مثلاً برای معین کردن رنگ موی یک شخص می‌توانیم با کدگذاری رنگ‌های مختلف، شاید با نمایش آنها به

صورت اعداد ۱، ۲، ۳، و غیره، به دلخواه به توصیف‌های کیفی، مقادیر حقیقی بدھیم. در همهٔ مثال‌های این بخش، بحث خود را به فضاهای نمونه‌ای گسترش می‌دهیم. نتیجه به متغیرهای تصادفی گسترش می‌دهیم، یعنی متغیرهایی که برآ آنها متناهی یا نامتناهی شماراست محدود کرده‌ایم. متغیرهای تصادفی پیوسته که روی فضاهای نمونه‌ای پیوسته تعریف شده‌اند در بخش ۴-۴ مورد بحث قرار می‌گیرند.

۴-۲ توزیع‌های احتمال

در مثال‌های ۱-۴ و ۲-۴ دیدیم، اندازه‌های احتمالی که روی یک فضای نمونه‌ای گسترش تعریف می‌شود، خود به خود احتمال هر مقدار مفروضی را که متغیر تصادفی در برداش اختیار می‌کند در اختیار می‌گذارد.

مثال ۳-۴. (دباله‌ی مثال ۲-۴) با نسبت دادن احتمال $\frac{1}{36}$ به هر عنصر، بی‌درنگ در می‌یابیم که متغیر تصادفی X ، مجموع اعداد حاصل از ریختن دو تاس، مقدار ۹ را با احتمال $\frac{4}{36}$ اختیار می‌کند. زیرا $X = 9$ شامل ۴ نقطه از ۳۶ نقطه‌ی هم‌شانس فضای نمونه‌ای است. احتمال‌های مربوط به همهٔ مقادیر ممکن X در جدول زیر نشان داده شده‌اند.

| x | $P(X = x)$ |
|-----|----------------|
| ۲ | $\frac{1}{36}$ |
| ۳ | $\frac{2}{36}$ |
| ۴ | $\frac{3}{36}$ |
| ۵ | $\frac{4}{36}$ |
| ۶ | $\frac{5}{36}$ |
| ۷ | $\frac{6}{36}$ |
| ۸ | $\frac{5}{36}$ |
| ۹ | $\frac{4}{36}$ |
| ۱۰ | $\frac{3}{36}$ |
| ۱۱ | $\frac{2}{36}$ |
| ۱۲ | $\frac{1}{36}$ |

به جای اینکه احتمال‌های وابسته به مقادیر یک متغیر تصادفی را مانند مثال قبل در یک جدول نشان دهیم، معمولاً بهتر آن است که فرمولی ارائه کنیم، یعنی احتمال‌ها را به وسیله‌ی تابعی نشان دهیم که مقادیر آن، یعنی $f(x)$ ، بهازای هر x ی که در برد متغیر تصادفی است برابر $P(X = x)$ باشد.

مثال ۴-۴. (دباله‌ی مثال ۴-۳) برای مجموع دو عددی که در ریختن دو تاس ظاهر می‌شوند، می‌توانیم بنویسیم

$$f(x) = \frac{6 - |x - 7|}{36}, \quad x = 2, 3, \dots, 12$$

که درستی آن را می‌توان به آسانی تحقیق کرد. به وضوح

$$f(2) = \frac{6 - |2 - 7|}{36} = \frac{6 - 5}{36} = \frac{1}{36}$$

$$f(3) = \frac{6 - |3 - 7|}{36} = \frac{6 - 4}{36} = \frac{2}{36}$$

⋮

$$f(12) = \frac{6 - |12 - 7|}{36} = \frac{6 - 5}{36} = \frac{1}{36}$$

و تمام این مقادیر با آنچه در جدول قبل نشان دادیم مطابقت دارند.

تعریف ۴-۲. اگر X یک متغیر تصادفی گستته باشد، تابعی که برای هر مقدار x در برد X با $f(x) = P(X = x)$ داده می‌شود، توزیع احتمال X نامیده می‌شود.

برمبنای اصول موضوع احتمال، بی‌درنگ نتیجه می‌شود که

قضیه ۴-۱. تابعی را می‌توان وقتی و فقط وقتی به عنوان توزیع احتمال یک متغیر تصادفی گستته‌ی X به کاربرد که مقادیر آن، $f(x)$ ، در شرایط زیر صدق کنند:

الف) برای هر مقدار حوزه‌ی تابع، $f(x) \geq 0$ ،

ب) $\sum_x f(x) = 1$ ، که در آن مجموع یابی روی تمام مقادیر حوزه‌ی تابع صورت می‌گیرد.

مثال ۴-۵. فرمولی برای توزیع احتمال تعداد کل شیرهایی که در n پرتاب یک سکه‌ی همگن به دست می‌آیند پیدا کنید.

حل. بر پایه‌ی احتمال‌های جدول مثال ۲-۴، درمی‌باییم که $P(X=0) = \frac{1}{16}$ ، $P(X=1) = \frac{4}{16}$ ، $P(X=2) = \frac{6}{16}$ ، $P(X=3) = \frac{4}{16}$ ، $P(X=4) = \frac{1}{16}$. از صورت‌های این پنج کسر، یعنی ۰، ۱، ۲، ۳، ۴، و ۵ که ضرایب دوجمله‌ای

$$\binom{4}{0}, \binom{4}{1}, \binom{4}{2}, \binom{4}{3}, \binom{4}{4}$$

هستند، متوجه می‌شویم که می‌توان برای توزیع احتمال، فرمولی به صورت

$$f(x) = \binom{4}{x}, \quad x = 0, 1, 2, 3, 4$$

نوشت. توجیهی نظری برای این فرمول، و بحثی کلی‌تر برای n پرتاب سکه‌ی همگن، در فصل پنجم، در توزیع‌های دوجمله‌ای داده خواهد شد.

مثال ۴-۶. آیا تابعی به صورت

$$f(x) = \frac{x+4}{25}, \quad x = 1, 2, 3, 4, 5$$

می‌تواند به عنوان تابع احتمال یک متغیر تصادفی گستته به کار رود؟

حل. اگر مقادیر مختلف x را در $f(x)$ قراردهیم، به دست می‌آوریم:

$$f(1) = \frac{1}{25}, \quad f(2) = \frac{2}{25}, \quad f(3) = \frac{3}{25}, \quad f(4) = \frac{4}{25}, \quad f(5) = \frac{5}{25}$$

چون تمام این مقادیر نامنفی‌اند، اوّلین شرط قضیه‌ی ۴-۱ برقرار است، و چون

$$f(1) + f(2) + f(3) + f(4) + f(5) = \frac{1}{25} + \frac{2}{25} + \frac{3}{25} + \frac{4}{25} + \frac{5}{25} = 1$$

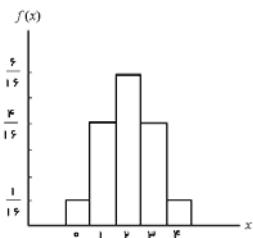
شرط دوّم قضیه‌ی ۴-۱ نیز صادق است. پس تابع داده شده، می‌تواند به عنوان توزیع احتمال یک متغیر تصادفی در برد $\{1, 2, 3, 4, 5\}$ به کار رود. البته این‌که آیا X متغیر تصادفی مفروضی واقعاً دارای این توزیع احتمال باشد یا نه، کلاً مطلبی دیگر است.

در بعضی از مسائل، ارائه‌ی توزیع‌های احتمال به صورت نموداری موردنظر است.

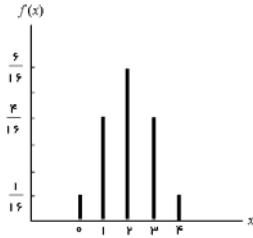
دو نوع نمایش نموداری که برای این منظور به کار می‌روند در شکل‌های ۲-۴ و ۳-۴ نشان داده شده‌اند. نمایشی که در شکل ۲-۴ نشان داده شده است بافت‌نمای احتمال خوانده می‌شود و توزیع احتمال مثال ۳-۴ را نمایش می‌دهد. ارتفاع هر مستطیل برابر

احتمال این است که X مقداری متناظر با طول نقطه‌ی وسط قاعده‌ی آن مستطیل اختیار کند. با نمایشی بازه‌ی از $5/0$ تا $5/0$ با 5 ، بازه‌ی از $5/0$ تا $1/5$ با 1 ، ...، بازه‌ی از $3/5$ تا $5/5$ با 4 ، مقادیر متغیر تصادفی گستته مفروض را به اصطلاح روی مقیاسی پیوسته «گسترده‌ایم». چون هر مستطیل بافت‌نمای شکل ۲-۴ دارای پهنه‌ی واحد است، می‌توانیم بگوییم مساحت‌های مستطیل‌ها، به جای ارتفاع‌های آن‌ها، برابر احتمال‌های متناظرند. برابر دانستن مساحت‌های مستطیل‌ها با احتمال‌ها، مزیت‌هایی دارد: برای مثال، وقتی می‌خواهیم نمودار توزیع احتمال گستته را با خمی پیوسته تقریب بزنیم از این مطلب استفاده می‌کنیم: این کار را، حتی وقتی مستطیل‌های بافت‌نمای پهنه‌ی برابر واحد هم ندارند، می‌توان با تعدیل ارتفاع‌های مستطیل‌ها یا با اصلاح مقیاس محور قائم انجام داد.

نمودار شکل ۳-۴، نمودار میله‌ای نامیده می‌شود. مانند شکل ۲-۴، ارتفاع هر مستطیل، یا میله، مساوی احتمال مقدار متناظر متغیر تصادفی است، اما وانمود نمی‌شود که مقیاس پیوسته‌ای وجود دارد. گرچه در این کتاب، در موقعی از چنین نمودارهایی استفاده خواهیم کرد، اما بافت‌نمایها و نمودارهای میله‌ای عمدتاً در آمار توصیفی برای ارائه‌ی بصری اطلاعاتی که به وسیله‌ی توزیع احتمال یا توزیع واقعی داده‌ها فراهم می‌شوند به کار می‌روند.



شکل ۲-۴. بافت‌نمای احتمال



شکل ۳-۴. نمودار میله‌ای

در مسائل زیادی، دانستن احتمال اینکه مقداری از متغیر تصادفی کوچکتر از یک مقدار حقیقی x یا برابر با آن باشد مورد توجه است. لذا، احتمال این را که X مقداری کوچکتر از x یا برابر آن اختیار کند به صورت $F(x) = P(X \leq x)$ می‌نویسیم، و این تابع را که برای تمام اعداد حقیقی x تعریف شده است، تابع توزیع تجمعی یا توزیع تجمعی متغیر تصادفی X می‌نامیم.

تعریف ۴-۳. اگر X یک متغیر تصادفی گستته باشد، تابعی که با

$$F(x) = P(X \leq x) = \sum_{t \leq x} f(t), \quad -\infty < x < +\infty,$$

داده می‌شود، و در آن $f(t)$ مقدار احتمال X به ازای t است، تابع توزیع یا توزیع تجمعی X خوانده می‌شود.

بنابر اصول موضوع احتمال و بعضی از پیامدهای فوری آن، نتیجه می‌شود که قضیه‌ی ۴-۲، مقادیر $F(x)$ ، تابع توزیع متغیر تصادفی گستته‌ی X ، در شرایط زیر صدق می‌کنند:

$$1. F(-\infty) = 0, \quad F(+\infty) = 1$$

۲. به ازای هر دو عدد حقیقی a و b ، اگر $a < b$ ، آنگاه $F(a) \leq F(b)$.

اگر توزیع احتمال متغیر تصادفی گستته‌ی مفروضی را داشته باشیم، تابع توزیع متناظر آن عموماً ساده است.

مثال ۴-۷. تابع توزیع تعداد کل شیرهایی را که در چهار پرتاب یک سکه‌ی همگن به دست می‌آیند بیابید.

حل. از مثال ۴-۵، با دانستن

$$f(4) = \frac{1}{16}, \quad f(3) = \frac{4}{16}, \quad f(2) = \frac{6}{16}, \quad f(1) = \frac{4}{16}, \quad f(0) = \frac{1}{16}$$

نتیجه می‌شود که

$$\begin{aligned}
 F(\circ) &= f(\circ) = \frac{1}{15}, \\
 F(\text{I}) &= f(\circ) + f(\text{I}) = \frac{5}{15}, \\
 F(\text{r}) &= f(\circ) + f(\text{I}) + f(\text{r}) = \frac{11}{15}, \\
 F(\text{w}) &= f(\circ) + f(\text{I}) + f(\text{r}) + f(\text{w}) = \frac{16}{15}, \\
 F(\text{e}) &= f(\circ) + f(\text{I}) + f(\text{r}) + f(\text{w}) + f(\text{e}) = \frac{15}{15} = 1.
 \end{aligned}$$

بنابر این تابع توزیع با

$$F(x) = \begin{cases} 0 & x < \circ \\ \frac{1}{15} & \circ \leq x < \text{I} \\ \frac{5}{15} & \text{I} \leq x < \text{r} \\ \frac{11}{15} & \text{r} \leq x < \text{w} \\ \frac{16}{15} & \text{w} \leq x < \text{e} \\ 1 & \text{e} \leq x \end{cases}$$

داده می‌شود. ملاحظه کنید که این تابع توزیع نه تنها برای مقادیری که متغیر تصادفی دارد، مقدار اختیار می‌کند، بلکه برای تمام اعداد حقیقی تعریف شده است. مثلاً می‌توانیم بنویسیم $F(\text{I}/\text{r}) = 1 - F(\text{r}/\text{I}) = 1 - \frac{11}{15} = \frac{4}{15}$ و $F(\text{w}/\text{e}) = 1 - F(\text{e}/\text{w}) = 1 - \frac{16}{15} = \frac{1}{15}$. گرچه احتمال‌های به دست آوردن «حداکثر ۷/۱ شیر» یا «حداکثر ۱۰۰ شیر» در چهار پرتاب یک سکه‌ی همگن ممکن است هیچ معنایی واقعی نداشته باشد.

مثال ۴-۸ تابع توزیع متغیر تصادفی W ای مثال ۴-۱ را بباید و نمودار آن را رسم کنید.

حل. می‌توانیم بنویسیم

$$f(\text{r}) = \frac{5}{14}, \quad f(\text{I}) = \left(\frac{1}{15}\right) + \left(\frac{1}{14}\right) = \left(\frac{1}{14}\right), \quad f(\circ) = \frac{3}{14}$$

به طوری که

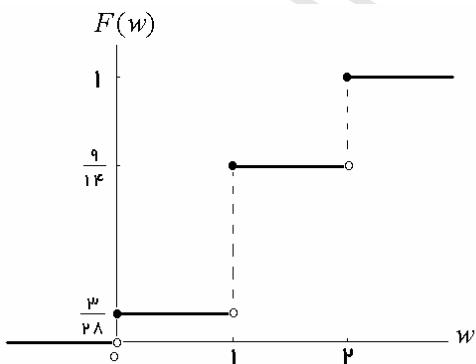
$$\begin{aligned}
 F(\circ) &= f(\circ) = \frac{3}{14}, \\
 F(\text{I}) &= f(\circ) + f(\text{I}) = \frac{9}{14}, \\
 F(\text{r}) &= f(\circ) + f(\text{I}) + f(\text{r}) = 1.
 \end{aligned}$$

بنابر این تابع توزیع W با

$$F(w) = \begin{cases} 0 & w < 0 \\ \frac{3}{28} & 0 \leq w < 1 \\ \frac{9}{14} & 1 \leq w < 2 \\ 1 & 2 \leq w \end{cases}$$

داده می‌شود.

نمودار این تابع توزیع که در شکل ۴-۴ نشان داده شده است، بدین طریق به دست می‌آید که اول نمایش نقاط $(w, F(w))$ را به ازای $w=0, 1, 2$ مشخص می‌کنیم سپس تابع پله‌ای را همان‌طور که نشان داده‌ایم کامل می‌کنیم. باید توجه داشته باشیم که این تابع در تمام نقاط ناپیوستگی، از دو مقدار، آن را که بزرگ‌تر است اختیار می‌کند.



شکل ۴-۴. نمودار تابع توزیع مثال ۸-۴

می‌توانیم فرایندی را که در دو مثال قبل توضیح دادیم در جهت عکس نیز انجام دهیم، یعنی مقادیر توزیع احتمال متغیر تصادفی را از روی تابع توزیع به دست آوریم. برای این منظور قضیه‌ی زیر را به کار می‌بریم.

قضیه‌ی ۳-۴. اگربرد متغیر تصادفی X ، متشکل از مقادیر $x_1 < x_2 < x_3 < \dots < x_n$ باشد، آنگاه $f(x_i) = F(x_i) - F(x_{i-1})$ و

$$f(x_i) = F(x_i) - F(x_{i-1}), \quad i = 1, 2, 3, \dots, n.$$

مثال ۹-۴. اگر تابع توزیع X به صورت

$$F(x) = \begin{cases} 0 & x < 2 \\ \frac{1}{36} & 2 \leq x < 3 \\ \frac{3}{36} & 3 \leq x < 4 \\ \frac{6}{36} & 4 \leq x < 5 \\ \frac{10}{36} & 5 \leq x < 6 \\ \frac{15}{36} & 6 \leq x < 7 \\ \frac{11}{36} & 7 \leq x < 8 \\ \frac{9}{36} & 8 \leq x < 9 \\ \frac{3}{36} & 9 \leq x < 10 \\ \frac{3}{36} & 10 \leq x < 11 \\ \frac{3}{36} & 11 \leq x < 12 \\ 1 & x \geq 12 \end{cases}$$

باشد، توزیع احتمال این متغیر تصادفی را بیابید.

حل. با استفاده از قضیه‌ی ۳-۴ به دست می‌آوریم $f(3) = \frac{3}{36} - \frac{1}{36} = \frac{2}{36}$ ، $f(4) = \frac{1}{36}$ ، $f(5) = \frac{1}{36}$ ، $f(6) = 1 - \frac{35}{36} = \frac{1}{36}$ ، $f(7) = \frac{1}{36}$ ، $f(8) = \frac{1}{36}$ ، $f(9) = \frac{1}{36}$ ، $f(10) = \frac{1}{36}$ ، $f(11) = \frac{1}{36}$ ، $f(12) = \frac{1}{36}$ و مقایسه‌ی این‌ها با احتمال‌های مثال ۳-۴، آشکار می‌کند که متغیر تصادفی که در اینجا با آن سروکار داریم همان مجموع اعدادی است که در ریختن یک جفت تاس ظاهر می‌شوند.

تمرین

۴-۱. تحقیق کنید که آیا تابع زیر یکتابع احتمال است یا نه؟

$$f(x) = \begin{cases} cx^r & x = 1, 2, 3 \\ 0 & \text{سایر جاهای} \end{cases}$$

۴-۲. مقدار c را طوری پیدا کنید که تابع زیر یکتابع احتمال باشد.

$$f(x) = \begin{cases} c\left(\frac{1}{6}\right)^{x-1} & x = 1, 2, 3, \dots \\ 0 & \text{سایر جاهای} \end{cases}$$

۴-۳. برای توابع تعریف شده X در زیر نشان دهید X یک متغیر تصادفی گستته است و تابع احتمال آن را به دست آورید.

الف) یک سکه‌ی همگن را سه بار پرتاب می‌کنیم. X نمایانگر تعداد خط‌های مشاهده

شده است.

ب) یک تاس همگن را دو مرتبه پرتاب می‌کنیم. X نمایانگر قدر مطلق تفاضل دو خال مشاهده شده است.

ج) یک تاس همگن را آن قدر پرتاب می‌کنیم تا برای اوّلین بار شیر بیاید. X نمایانگر پرتاب‌های لازم قبل از مشاهده اوّلین شیر است.

د) یک تاس را آن قدر پرتاب می‌کنیم تا دو خال ۶ مشاهده شود. X تعداد پرتاب‌های لازم برای مشاهده دو خال شش تعریف می‌شود.

۴-۴. جعبه‌ای شامل ۵ مهره است که از ۱ تا ۵ شماره‌گذاری شده‌اند. دو مهره به تصادف بدون جایگذاری از این جعبه انتخاب می‌کنیم.

الف) X نمایانگر مаксیمم دو عدد مشاهده شده است،تابع احتمال X را به‌دست آورید.

ب) اگر X نمایانگر مینیمم دو عدد مشاهده شده باشد، تابع احتمال X را به‌دست آورید.

ج) اگر X نمایانگر مجموع دو عدد مشاهده شده باشد، تابع احتمال X را به‌دست آورید.

۴-۵. فرض کنید ۵ نفر به همراه شما و دوستان به تصادف در یک ردیف ایستاده باشید. اگر X نمایانگر تعداد افراد بین شما و دوست شما باشد، تابع احتمال X را به دست آورید.

۴-۶. جعبه‌ی I شامل ۲ مهره‌ی سفید و ۲ مهره‌ی سیاه و جعبه‌ی II شامل ۳ مهره‌ی سفید و ۳ مهره‌ی سیاه است. دو مهره به تصادف و بدون جایگذاری از هر جعبه انتخاب می‌کنیم اگر X نمایانگر مهره‌های سفید انتخابی از جعبه‌ی I و Y نمایانگر مهره‌های سفید انتخابی از جعبه‌ی II باشیم، مطلوب است

الف) تابع احتمال X ،

ب) تابع احتمال Y تابع احتمال $Z = X + Y$

۴-۷. جعبه‌ای شامل ۲ مهره‌ی سفید و ۳ مهره‌ی سیاه است. دو مهره به تصادف و بدون جایگذاری از این جعبه انتخاب می‌کنیم. اگر X نمایانگر تعداد مهره‌های سفید در نمونه‌ی تصادفی باشد، نشان دهید X یک متغیر تصادفی گسسته است و تابع احتمال

آن را به دست آورید. اگر انتخاب نمونه با جایگذاری باشد چطور؟

- ۴-۸. اگر X یک متغیر تصادفی گسسته با تابع احتمال زیر باشد، مقدار c را به دست آورید.

$$f(x) = c \frac{\theta^x}{x} \quad x = 1, 2, \dots, \quad 0 < \theta < 1$$

- ۴-۹. اگر x_1, x_2, \dots نمایانگر اعداد گویا در بازه‌ی $[1, \infty)$ و X یک متغیر تصادفی گسسته با تابع احتمال زیر باشد، تکیه‌گاه متغیر تصادفی X کدام است؟ چرا؟

$$P(X = x_n) = 2^{-n} \quad n = 1, 2, \dots$$

- ۴-۱۰. اگر N یک عدد صحیح مثبت و X دارای تابع احتمال زیر باشد:

$$f(x) = c 2^{-x}, \quad n = 1, 2, \dots, N.$$

الف) c را تعیین کنید.

ب) تابع توزیع X را به شکل بسته به دست آورید.

- ۴-۱۱. اگر X دارای تابع احتمال زیر باشد، مقادیر $P(X \leq 1)$ ، $P(X \geq 2)$ و $P(|X - 1| \leq 1)$ را به دست آورید.

| | | | | |
|--------|-------|-------|-------|-------|
| x | ۰ | ۱ | ۲ | ۳ |
| $f(x)$ | $5/3$ | $5/4$ | $5/2$ | $1/1$ |

- ۴-۱۲. از بین اعداد ۱ تا ۹، چهار عدد را به تصادف و با هم انتخاب می‌کنیم. اگر X نمایانگر عددی باشد که از کوچکترین عدد انتخابی نمونه بزرگتر و از دو تای دیگر کوچکتر باشد، مقدار $P(X = 4)$ را به دست آورید.

- ۴-۱۳. ظرفی شامل ۵ مهره است که X تای آن سفید و بقیه سیاه‌اند. می‌دانیم که $P(X = 1) = \frac{1}{5}$ و $P(X = 2) = \frac{2}{5}$. از این ظرف یک مهره به تصادف انتخاب می‌کنیم، ملاحظه می‌شود که رنگ آن سفید است. احتمال آنرا که در ظرف ۳ مهره سفید وجود داشته باشد به دست آورید.

- ۴-۱۴. اگر تابع توزیع متغیر تصادفی X به صورت

$$F(x) = \begin{cases} 0 & x < 0 \\ \frac{1}{5} & 0 \leq x < 1 \\ \frac{3}{5} & 1 \leq x < 2/5 \\ \frac{5}{5} & 2/5 \leq x < 4 \\ \frac{11}{14} & 4 \leq x < 5/5 \\ 1 & 5/5 \leq x \end{cases}$$

باشد، مطلوب است

الف) تابع احتمال X ,

ب) $P(1 \leq X < 2/5)$,

ج) $P(0 < X \leq 2/5)$,

د) $P(0 < X < 2/5)$,

ه) $P(1 \leq X \leq 2/5)$.

۴-۱۵. تحقیق کنید که $f(x) = \frac{2x}{k(k+1)}$ به ازای $k = 1, 2, 3, \dots$ را می‌توان به عنوان

توزیع احتمال یک متغیر تصادفی به کار برد.

۴-۱۶. نشان دهید که برای c نمی‌توان مقادیری یافت که به ازای آنها

$$f(x) = \frac{c}{x},$$

را بتوان به عنوان مقادیر توزیع احتمال یک متغیر تصادفی با برد نامتناهی شمارا
به کار برد.

۴-۱۷. برای هر یک از توزیع‌های احتمال زیر بافت‌نمای احتمال را رسم کنید.

$$\text{الف) } f(x) = \binom{2}{x} \binom{4}{3-x} / \binom{6}{3}, \quad x = 0, 1, 2,$$

$$\text{ب) } f(x) = \binom{5}{x} \left(\frac{1}{5}\right)^x \left(\frac{4}{5}\right)^{5-x}, \quad x = 0, 1, 2, 3, 4, 5$$

۴-۱۸. برای هر یک از موارد زیر تعیین کنید که آیا مقادیر داده شده را می‌توان به عنوان

مقادیر تابع توزیع یک متغیر تصادفی با برد $\{0, 1, 2, 3, 4\}$ به کار برد یا نه.

الف) $F(4) = 0/2$, $F(3) = 0/5$, $F(2) = 0/8$, $F(1) = 0/3$

ب) $F(4) = 1/0$, $F(3) = 0/7$, $F(2) = 0/4$, $F(1) = 0/5$

ج) $F(4) = 1/0$, $F(3) = 0/8$, $F(2) = 0/6$, $F(1) = 0/25$

۴-۳ متغیرهای تصادفی پیوسته

در بخش ۱-۴ متغیر تصادفی را به صورت تابعی حقیقی معرفی کردیم که روی نقاط فضای نمونه‌ای دارای یک اندازه‌ی احتمال تعریف شده بود، و در شکل ۱-۴، با تخصیص مجموع اعدادی که در ریختن یک جفت تاس ظاهر می‌شوند به هر یک از ۳۶ نقطه‌ی هم‌شانس فضای نمونه‌ای، این مفهوم را تشریح کردیم. در حالت پیوسته، که متغیرهای تصادفی می‌توانند مقادیری را روی مقیاسی پیوسته اختیار کنند، شیوه‌ی کار به میزان زیاد همانند حالت گستته است. برآمدهای آزمایش‌ها با نقاط روی پاره‌خط‌ها یا خط‌ها نمایش داده می‌شوند، و مقادیر متغیرهای تصادفی اعدادی هستند که به گونه‌ای مناسب به وسیله‌ی قاعده‌ها یا معادله‌ها B_x به این نقاط نسبت داده می‌شوند. وقتی مقدار متغیر تصادفی مستقیماً با یک اندازه یا مشاهده داده می‌شود، عموماً به خود زحمت نمی‌دهیم که بین مقدار متغیر تصادفی (اندازه‌ای که به دست می‌آوریم) و برآمد آزمایش (نقطه‌ی متناظر روی محور حقیقی) تمایزی قائل شویم. مثلاً، اگر یک آزمایش عبارت از تعیین محتوای واقعی یک شیشه ۲۳۵ گرمی قهقهه باشد، نتیجه‌ی حاصل، مثلاً $\frac{۲۲۵}{۳}$ گرم، مقدار متغیر تصادفی است که موردنظر ماست، و حقیقتاً نیازی نیست اضافه کنیم که فضای نمونه‌ای عبارت از بازه‌ی پیوسته‌ی معینی از نقاط روی محور اعداد حقیقی است.

مسئله‌ی تعریف احتمال‌ها در رابطه با فضاهای نمونه‌ای پیوسته و متغیرهای تصادفی پیوسته متضمن پیچیدگی‌هایی است. برای توضیح مطلب وضعیت زیر را در نظر می‌گیریم.

مثال ۴-۱۰. فرض کنید به احتمال سانحه‌ای که در بزرگراهی به طول ۲۰۰ کیلومتر رخ دهد علاقه‌مندیم، و احتمال اینکه تصادف در محل مشخصی، یا شاید روی قطعه‌ی معینی از جاده رخ دهد، مورد توجه ماست. فضای نمونه‌ای این «آزمایش»، متشکل از پیوستاری از نقاط است، نقاطی که در بازه‌ی ۰ تا ۲۰۰ کیلومتر قرار دارند، و ما برای ارائه‌ی استدلال، فرض می‌کنیم احتمال اینکه تصادفی در هر بازه‌ای به طول d رخ دهد برابر $\frac{d}{200}$ باشد، که در آن d بر حسب کیلومتر اندازه گرفته می‌شود. توجه کنید که این تخصیص احتمال‌ها با اصل‌های موضوع ۱ و ۲ سازگار است، زیرا احتمال‌های $\frac{d}{200}$,

همگی نامنفی‌اند و $P(S) = \frac{۲۰۰}{۳۰۰}$. تا اینجا این تخصیص احتمال‌ها فقط در مورد بازه‌ی $۰\text{--}۲۰۰$ انجام شد، اما اگر اصل موضوع ۳ را به کار ببریم، می‌توانیم احتمال‌ها را برای اجتماع هر تعداد متناهی یا نامتناهی شمارای دنباله‌ای از بازه‌های جدا از هم نیز به کار ببریم. به عنوان مثال، احتمال این‌که تصادفی روی هر کدام از دو بازه‌ی جدا از هم به طول‌های $d_۱$ و $d_۲$ رخ دهد برابر

$$\frac{d_۱ + d_۲}{۳۰۰}$$

است، و احتمال این‌که تصادفی روی یکی از بازه‌های متعلق به دنباله‌ای نامتناهی شمارا از بازه‌های جدا از هم به طول‌های $d_۱, d_۲, \dots$ رخ دهد برابر است با

$$\frac{d_۱ + d_۲ + d_۳ + \dots}{۳۰۰}$$

در این صورت، اگر قضیه‌ی $۳-۷$ را به کار ببریم، می‌توانیم تخصیص احتمال را به اجتماع بازه‌هایی که از هم جدا نیستند گسترش دهیم، و چون اشتراک دو بازه، یک بازه، و متمم یک بازه نیز یک بازه، یا اجتماع دو بازه است، می‌توانیم تخصیص احتمال را به هر زیرمجموعه‌ای از فضای نمونه‌ای تعمیم دهیم که بتوان آن زیرمجموعه را از تشکیل اجتماع‌ها یا اشتراک‌های تعداد نامتناهی یا نامتناهی شمارا از بازه‌ها، یا با تشکیل متمم‌ها به دست آورد.

پس در توسعی مفهوم احتمال به حالت پیوسته، باز اصل‌های $۱, ۲$ ، و ۳ را به کار بردہ‌ایم، اما برای انجام این کار در حالت کلی، باید از تعریفی که برای «پیشامد» کردیم، تمام زیرمجموعه‌هایی از فضای نمونه‌ای را که نمی‌توان با ساختن اجتماع‌ها یا اشتراک‌های تعدادی نامتناهی شمارای بازه‌ها، یا ساختن متمم‌ها به دست آورد مستثنی کنیم. از نظر عملی، این استثناء اهمیتی ندارد، زیرا ما صرفاً احتمال‌هایی به این نوع مجموعه‌های پیچیده نسبت نمی‌دهیم.

با مراجعه به مثال $۴-۱۰$ ، همچنین ملاحظه کنید که احتمال وقوع تصادفی در یک بازه‌ی خیلی کوتاه، مثلاً بازه‌ای به طول یک سانتی‌متر فقط برابر ۵×۱0^{-۸} است که خیلی کوچک است. وقتی طول بازه به صفر می‌گراید، احتمال این‌که تصادفی روی آن رخ دهد نیز به صفر می‌گراید؛ البته در حالت پیوسته، ما به نقطه‌های تنها احتمال صفر را نسبت می‌دهیم. این بدین معنا نیست که پیشامدهای متناهی نمی‌توانند رخ دهند،

بالاخره وقتی سانحه‌ای در طول ۲۰۰ کیلومتر جاده رخ می‌دهد، الزاماً این تصادف در نقطه‌ای رخ می‌دهد، ولو اینکه هر نقطه احتمال صفر داشته باشد.

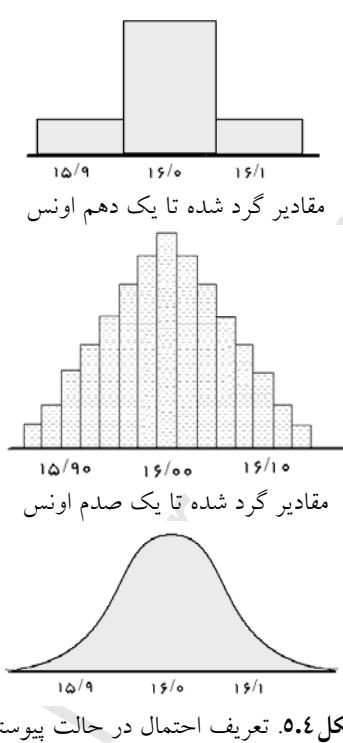
۴-۴ تابع‌های چگالی احتمال

راهی که در مثال ۱۰-۴ برای تخصیص احتمال‌ها به کار بردهیم حالت بسیار خاصی است، و ذاتاً شبیه راهی است که طی آن، احتمال‌های مساوی به شش وجه یک تاس، به شیرها و خطها و به ۵۲ کارت از یک نوع دسته کارت معمولی وغیره نسبت می‌دهیم. برای بحث در مسئله‌ی نسبت دادن احتمال‌ها به متغیرهای تصادفی به‌طور کلی تر، فرض می‌کنیم که برای متصدی قسمت پرکردن بطری‌های آب معدنی، مقدار واقعی آب معدنی که ماشین بطری پرکن در بطری‌های یک لیتری می‌ریزد مطرح باشد. به وضوح مقدار نوشابه از یک بطری به بطری دیگر تغییر می‌کند، و در واقع این مقدار، یک متغیر تصادفی پیوسته است. اما اگر متصدی قسمت، مقادیر آب معدنی بطری‌ها را تا یک دهم لیتر گردکند با یک متغیر تصادفی گستته سروکار دارد که دارای توزیع احتمال است، و این توزیع احتمال را می‌توان به صورت بافت‌نمایی که در آن احتمال‌ها با مساحت‌های مستطیل‌ها داده شده‌اند، مثلاً نظیر نمودار قسمت بالای شکل ۵-۴، نمایش داد. اگر او مقادیر آب معدنی را تا یک صدم لیتر گرد کند، باز با متغیر تصادفی گستته‌ای (متفاوت با اویلی) که دارای یک توزیع احتمال است سروکار دارد، و این توزیع احتمال را می‌توان به صورت بافت‌نمایی که در آن احتمال‌ها با مساحت‌های مستطیل‌ها داده شده‌اند، مثلاً نظیر نمودار قسمت وسط شکل ۵-۴، نمایش داد.

واضح است که اگر او مقادیر نوشابه را تا یک هزارم اونس یا تا یک ده هزارم اونس گرد کند، بافت‌نمای توزیع‌های احتمال متغیرهای تصادفی گستته‌ی متناظر، به منحنی پیوسته‌ای که در قسمت پایین شکل ۵-۴ نشان داده‌ایم میل می‌کند، و مجموع مساحت‌های مستطیل‌هایی که احتمال قرار گرفتن مقادیر آب معدنی در هر بازه‌ی معینی را نمایش می‌دهند، به مساحت سطح متناظر زیر منحنی میل می‌کند.

در واقع، تعریف احتمال در حالت پیوسته، برای هر متغیر تصادفی، وجود تابعی را که تابع چگالی احتمال نامیده می‌شود مفروض می‌گیرد، به قسمی که مساحت‌های زیر منحنی این تابع، احتمال‌های مربوط به بازه‌های متناظر در طول محور افقی را مشخص

می‌کنند. به بیان دیگر، انتگرال یک تابع چگالی از a تا b ($a \leq b$) احتمال این را که متغیر تصادفی متناظر، مقداری را در بازه‌ی a تا b اختیار کند به دست می‌دهد.



شکل ۵.۴. تعییف احتمال در حالت پیوسته

تعريف ۴-۴. تابع حقیقی مقدار f ، که روی مجموعه‌ی تمام اعداد حقیقی تعییف شده است، تابع چگالی احتمال متغیر تصادفی پیوسته‌ی X نامیده می‌شود اگر و تنها اگر به ازای هر دو مقدار حقیقی ثابت a و b با شرط $a \leq b$

$$P(a \leq X \leq b) = \int_a^b f(x) dx.$$

به توابع چگالی به اختصار چگالی‌های احتمال، توابع چگالی، یا ت.چ.ا. نیز اطلاق می‌شود.

توجه کنید که $f(c)$ ، مقدار تابع چگالی به ازای c ، مانند حالت گسسته،

$P(X=c)$ را نمی‌دهد. در ارتباط با متغیرهای تصادفی پیوسته، احتمال‌ها همیشه به بازه‌ها نسبت داده می‌شوند و به ازای هر مقدار حقیقی ثابت c ، $P(X=c)=0$. این مطلب با تعریف ۴-۴ با $a=b=c$ مستقیماً نتیجه می‌شود.

به دلیل این ویژگی، مقدارتابع چگالی احتمال را می‌توان به ازای برعی از مقادیر متغیر تصادفی تغییر داد، بدون اینکه هیچ یک از احتمال‌ها تغییر کنند، و به همین دلیل در تعریف ۴-۴ گفتیم که $f(x)$ مقدار یکی از چگالی‌های احتمال متغیر تصادفی X به ازای x است و نه چگالی احتمال آن به معنی مطلق. به دلیل همین ویژگی نیز، مهم نیست که نقاط دو سر بازه‌ی a تا b را در محاسبه‌ی احتمال منظور کنیم یا نکنیم. به صورت نمادی در قضیه‌ی ۴-۴، اگر X ، متغیر تصادفی پیوسته و a و b دو عدد حقیقی ثابت با شرط $a \leq b$ باشند، آنگاه

$$P(a \leq X \leq b) = P(a \leq X < b) = P(a < X \leq b) = P(a < X < b).$$

مشابه با قضیه‌ی ۴-۱، ویژگی‌های زیر از توابع چگالی احتمال را که باز هم مستقیماً از اصول موضوع احتمال نتیجه می‌شوند بیان می‌کنیم.

قضیه‌ی ۴-۵. تابعی را می‌توان به عنوان تابع چگالی احتمال یک متغیر تصادفی پیوسته X به کاربرد اگر مقادیر آن، $f(x)$ ، در شرایط ۱. به ازای $-\infty < x < +\infty$ ، $f(x) \geq 0$ ۲. $\int_{-\infty}^{+\infty} f(x) dx = 1$ صدق کند.

مثال ۴-۱۱. اگر X دارای چگالی احتمال

$$f(x) = \begin{cases} k e^{-\frac{x}{3}} & x > 0 \\ 0 & \text{سایر جاهای} \end{cases}$$

باشد، k و $P(0 \leq X \leq 1) = 0.5$ را بیابید.

حل. برای اینکه دو مین شرط قضیه‌ی ۴-۵ برقرار باشد باید داشته باشیم

$$\int_{-\infty}^{+\infty} f(x) dx = \int_0^{+\infty} k e^{-\frac{x}{3}} dx = k \lim_{t \rightarrow +\infty} \left[-e^{-\frac{x}{3}} \right]_0^t = k = 1$$

و نتیجه می‌شود که $k = 3$. برای احتمالی که خواسته‌ایم، به دست می‌آوریم

$$P(0/5 \leq X \leq 1) = \int_{0/5}^1 e^{-3x} dx = -e^{-3x} \Big|_{0/5}^1 = -e^{-3} + e^{-1/5} = 0.173.$$

گرچه متغیر تصادفی مثال قبل نمی‌تواند مقادیر منفی اختیار کند، ولی ما در حوزه‌ی چگالی احتمال آن را به گونه‌ای تصنیع بسط دادیم تا تمام اعداد حقیقی را شامل شود. این شیوه‌ای است که در تمام این کتاب از آن پیروی خواهیم کرد.

مثل حالت گسسته، مسائل زیادی وجود دارند که علاقه‌مندیم در آنها احتمال این را که مقدار متغیر تصادفی پیوسته از عدد حقیقی x ، کوچکتر یا مساوی یا آن باشد بدانیم. لذا تعریف زیر را که مشابه تعریف ۴-۳ است ارائه می‌دهیم.

تعریف ۴-۵. اگر X ، متغیر تصادفی پیوسته‌ای با تابع چگالی احتمال $f(t)$ باشد، تابع

$$F(x) = P(X \leq x) = \int_{-\infty}^x f(t) dt, \quad -\infty < x < +\infty$$

تابع توزیع یا توزیع تجمعی X نامیده می‌شود.

ویژگی‌های توابع توزیع که در قضیه‌ی ۴-۲ ارائه شدند، برای حالت پیوسته نیز برقرارند؛ یعنی $F(+\infty) = 1$ ، $F(-\infty) = 0$ ، $F(a) \leq F(b)$ ، $a < b$. به علاوه از تعریف ۴-۵ مستقیماً نتیجه می‌شود که

قضیه‌ی ۴-۶. اگر $f(x)$ و $F(x)$ ، به ترتیب مقدارهای تابع توزیع احتمال و تابع توزیع X در x باشند، آنگاه به ازای هر دو مقدار حقیقی و ثابت a و b با شرط $a \leq b$

$$P(a \leq X \leq b) = F(b) - F(a),$$

و

$$f(x) = \frac{dF(x)}{dx},$$

و این مشتق موجود است.

مثال ۴-۱۲. تابع توزیع متغیر تصادفی X مربوط به مثال ۱۱-۴ را بیابید. این تابع توزیع را برای محاسبه‌ی مجدد $P(0/5 \leq X \leq 1)$ نیز به کار ببرید.
حل. به ازای $x > 0$,

$$F(x) = \int_{-\infty}^x f(t) dt = \int_0^x e^{-\mu t} dt = -e^{-\mu t} \Big|_0^x = 1 - e^{-\mu x}$$

و چون بهازای $x \leq 0$ ، $F(x) = 0$ می‌توانیم بنویسیم

$$F(x) = \begin{cases} 0 & x \leq 0 \\ 1 - e^{-\mu x} & x > 0 \end{cases}$$

برای یافتن $P(0/5 \leq X \leq 1)$ از فرمول قسمت قضیه‌ی ۶-۴ استفاده می‌کنیم، که به دست می‌دهد

$$\begin{aligned} P(0/5 \leq X \leq 1) &= F(1) - F(0/5) \\ &= (1 - e^{-\mu}) - (1 - e^{-1/5}) \\ &= 0/173 \end{aligned}$$

این جواب با نتیجه‌ای که با استفاده از تابع چگالی احتمال در مثال ۱۱-۴ مستقیماً به دست آمد مطابقت دارد.

مثال ۱۳-۴. یک تابع چگالی احتمال، برای متغیر تصادفی که تابع توزیع اش

$$F(x) = \begin{cases} 0 & x < 0 \\ x & 0 \leq x < 1 \\ 1 & 1 \leq x \end{cases}$$

است، پیدا و نمودار آن را رسم کنید.

حل. چون تابع توزیع داده شده همه‌جا جز در $x = 0$ و $x = 1$ ، مشتق‌پذیر است. از تابع توزیع برای $x < 0$ ، $x > 0$ و $0 < x < 1$ مشتق می‌گیریم و 0 ، 1 را به دست می‌آوریم.

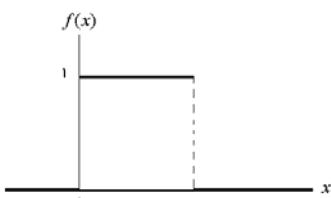
پس بنابر قسمت دوم از قضیه‌ی ۶-۴، می‌توانیم بنویسیم

$$f(x) = \begin{cases} 0 & x < 0 \\ 1 & 0 < x < 1 \\ 0 & x > 1 \end{cases}$$

و برای پرکردن دو رخنه در $x = 0$ و $x = 1$ ، مقادیر $f(0)$ و $f(1)$ را برابر صفر می‌گیریم. در واقع مهم نیست که چگالی احتمال در این دو نقطه چگونه تعریف شده است، اما انتخاب مقادیر به این طریق که تابع چگالی احتمال روی یک بازه‌ی باز صفر نباشد مزیت‌هایی دارد. پس می‌توان تابع چگالی احتمال متغیر تصادفی اصلی را به صورت

$$f(x) = \begin{cases} 1 & 0 < x < 1 \\ 0 & \text{سایر جاهای دیگر} \end{cases}$$

نوشت. نمودار این تابع را در شکل ۶ نشان داده‌ایم.



شکل ۶-۴. نمودار تابع چگالی احتمال مثال ۱۳-۴

تمرین

۱۹-۴. چگالی احتمال متغیر تصادفی پیوسته X به صورت

$$f(x) = \begin{cases} \frac{1}{5} & 2 < x < 7 \\ 0 & \text{سایر جاهای} \end{cases}$$

داده شده است.

الف) نمودار آن رارسم و تحقیق کنید که کل مساحت زیر منحنی (بالای محور x) برابر با یک است.

ب) مقدار $P(5 < X < 3)$ را بیابید.

۲۰-۴. چگالی احتمال متغیر تصادفی پیوسته Y به صورت

$$f(y) = \begin{cases} \frac{1}{8}(y+1) & -2 < y < 4 \\ 0 & \text{سایر جاهای} \end{cases}$$

داده شده است. مطلوب است $P(\frac{1}{9} < Y < \frac{3}{2})$ و $P(Y > \frac{1}{2})$.

۲۱-۴. تابع چگالی احتمال متغیر تصادفی X به صورت

$$f(x) = \begin{cases} \frac{c}{\sqrt{x}} & 0 < x < 4 \\ 0 & \text{سایر جاهای} \end{cases}$$

داده شده است مطلوب است

الف) مقدار c ، ب) $P(X > 1)$ و $P(X < \frac{1}{4})$

۲۲-۴. چگالی احتمال متغیر تصادفی پیوسته X به صورت

$$f(z) = \begin{cases} k z e^{-z^r} & z > 0 \\ 0 & z \leq 0 \end{cases}$$

داده شده است. k را بیابید و نمودار این چگالی احتمال را رسم کنید.

۴-۲۳. چگالی احتمال متغیر تصادفی پیوسته X به صورت

$$g(x) = \begin{cases} 6x(1-x) & 0 < x < 1 \\ 0 & \text{سایر جاهای} \end{cases}$$

داده شده است. مطلوب است $P(X > \frac{1}{4})$ و $P(X < \frac{1}{2})$

۴-۲۴.تابع توزیع متغیر تصادفی X را که چگالی احتمال آن به صورت

$$f(x) = \begin{cases} \frac{1}{\pi} & 0 < x < 1 \\ \frac{1}{\pi} & 2 < x < 4 \\ 0 & \text{سایر جاهای} \end{cases}$$

داده شده است بیابید. نمودارهای تابع چگالی احتمال و تابع توزیع را نیز رسم کنید.

۴-۲۵. تابع توزیع متغیر تصادفی X را که چگالی احتمال آن به صورت

$$f(x) = \begin{cases} x & 0 < x < 1 \\ 2-x & 1 \leq x < 2 \\ 0 & \text{سایر جاهای} \end{cases}$$

داده شده است بیابید. نمودارهای تابع چگالی احتمال و تابع توزیع را نیز رسم کنید.

۴-۲۶. تابع توزیع متغیر تصادفی X را که چگالی احتمال آن به صورت

$$f(x) = \begin{cases} \frac{x}{2} & 0 < x \leq 1 \\ \frac{1}{2} & 1 < x \leq 2 \\ \frac{2-x}{2} & 2 < x \leq 3 \\ 0 & \text{سایر جاهای} \end{cases}$$

داده شده است بیابید. نمودارهای تابع چگالی احتمال و تابع توزیع را نیز رسم کنید.

۴-۲۷. تابع توزیع متغیر تصادفی X به صورت

$$F(x) = \begin{cases} 0 & x < -1 \\ \frac{x+1}{2} & -1 \leq x < 1 \\ 1 & 1 \leq x \end{cases}$$

داده شده است. مطلوب است $P(-\frac{1}{2} < X < \frac{1}{2})$ و $P(-\frac{1}{2} < X < \frac{1}{2})$

۴-۲۸. تابع توزیع متغیر تصادفی X به صورت

$$F(y) = \begin{cases} 1 - \frac{9}{y^3} & y > 3 \\ 0 & y \leq 3 \end{cases}$$

داده شده است. مطلوب است $P(Y > 8)$ و $P(Y \leq 5)$

۴-۲۹. تابع توزیع متغیر تصادفی X به صورت

$$F(x) = \begin{cases} 1 - (1+x)e^{-x} & x > 0 \\ 0 & x \leq 0 \end{cases}$$

داده شده است. مطلوب است $P(X > 2)$, $P(1 < X < 3)$, $P(X \leq 4)$ و $P(X > 6)$

۴-۳۰. تعداد دقایقی که پروازی از شهر A به شهر B زودتر یا دیرتر انجام می‌شود متغیری تصادفی است که چگالی احتمال آن به صورت

$$f(x) = \begin{cases} \frac{1}{228}(36-x^3) & -6 < x < 6 \\ 0 & \text{سایر جاهای} \end{cases}$$

است، که در آن مقادیر منفی، معرف زود انجام شدن پرواز و مقادیر مثبت، معرف تأخیر در پروازند. پیدا کنید احتمال آن را که یکی از این پروازها

الف) حداقل ۲ دقیقه زودتر، ب) حداقل یک دقیقه دیرتر،

ج) مدتی از ۱ تا ۳ دقیقه زودتر، د) دقیقاً ۵ دقیقه دیرتر،
انجام گیرد.

۴-۳۱. مدت زمان سالم ماندن (برحسب ساعت) یک ماده‌ی غذایی بسته‌بندی شده‌ی فاسد نشدنی، متغیری تصادفی است که تابع احتمال آن به صورت

$$f(x) = \begin{cases} \frac{40000}{(x+100)^3} & x > 0 \\ 0 & x \leq 0 \end{cases}$$

است. پیدا کنید احتمال آن را که مدت زمان سالم ماندن یکی از این بسته‌ها

الف) حداقل ۲۰۰ ساعت،

ب) حداقل ۱۰۰ ساعت،

ج) بین ۸۰ تا ۱۲۰ ساعت.

۴-۳۲. میزان فرسودگی تاییر (برحسب ۱۰۰۰ کیلومتر) برای دارندگان اتومبیل با نوع خاصی تاییر، متغیری تصادفی است که تابع چگالی احتمال آن به صورت

$$f(x) = \begin{cases} \frac{1}{30} e^{-\frac{x}{30}} & x > 0 \\ 0 & x \leq 0 \end{cases}$$

داده شده است. پیدا کنید احتمال اینکه یکی از این لاستیک‌ها
 الف) حداقل ۱۸۰۰۰ کیلومتر،
 ب) از ۲۷۰۰۰ تا ۳۶۰۰۰ کیلومتر،
 ج) حداقل ۴۸۰۰۰ کیلومتر.

۴-۳۳. در شهری معین، مصرف روزانه‌ی آب (برحسب میلیون لیتر)، متغیری تصادفی است که چگالی احتمال آن به صورت

$$f(x) = \begin{cases} \frac{1}{9} x e^{-\frac{x}{9}} & x > 0 \\ 0 & \text{سایر جاهای} \end{cases}$$

است. احتمال اینکه در روز معینی،
 الف) مصرف آب در این شهر بیش از ۶ میلیون نباشد،
 ب) ذخیره آب کافی نباشد، در صورتی که ظرفیت روزانه این شهر ۹ میلیون لیتر باشد،
 چقدر است؟

۴-۳۴. طول عمر (برحسب سال) سگ‌های پنج ساله‌ی نژاد خاصی، متغیری است تصادفی که تابع توزیع آن به صورت

$$F(x) = \begin{cases} 0 & x \leq 5 \\ 1 - \frac{25}{x^3} & x > 5 \end{cases}$$

است. پیدا کنید احتمال اینکه چنین سگ پنج ساله‌ای
 الف) بیش از ده سال، ب) کمتر از هشت سال،
 ج) بین ۱۲ تا ۱۵ سال، عمر کند.

۴-۵ توزیع‌های چندمتغیره

در آغاز این فصل، متغیرتصادفی را به صورت تابعی حقیقی‌مقدار بر فضایی نمونه‌ای تعریف کردیم، و منطقی است که متغیرهای تصادفی زیادی را بتوان روی فضای نمونه‌ای واحدی تعریف کرد. برای مثال، در ارتباط با فضای نمونه‌ای شکل ۱-۴، ما فقط متغیری تصادفی را در نظر گرفتیم که مقادیرش مجموع دو عدد بودند که در ریختن

یک جفت تاس ظاهر می‌شوند، اما می‌توانستیم متغیری تصادفی را در نظر بگیریم که مقادیرش حاصل ضرب دو عددی باشند که در ریختن یک جفت تاس ظاهر می‌شوند، یا متغیر تصادفی که مقادیرش $0, 1$ یا 2 بوده و بیانگر تعداد تاس‌هایی باشند که برای آنها عدد 2 بیایید، و قس‌علی‌هذا. در ارتباط نزدیکتر با زندگی روزمره، ممکن است آزمایشی عبارت از انتخاب تصادفی تعدادی از 345 دانش‌آموزی باشد که به یک مدرسه‌ی ابتدایی می‌روند، و رئیس مدرسه به تعیین بهره‌ی هوشی آنها، مراقب بهداشتی مدرسه به وزن آنها، معلمان مدرسه به تعداد روزهایی که غیبت کردند و ... علاقه‌مند. در این بخش ابتدا به حالت دومتغیره می‌پردازیم، یعنی به وضعیت‌هایی با یک جفت متغیر تصادفی که هم‌زمان روی یک فضای نمونه‌ای توأم تعریف شده‌اند. بعدها، این بحث را به حالت چند متغیره، که هر تعداد متناهی از متغیرهای تصادفی را شامل می‌شود، تعمیم می‌دهیم.

اگر X و Y دو متغیر تصادفی گسسته باشند، احتمال این را که X مقدار x و y مقدار y را اختیار کند به صورت $P(X=x, Y=y)$ می‌نویسیم؛ بنابر این $P(X=x, Y=y)$ ، احتمال اشتراک پیشامدهای $X=x$ و $Y=y$ است. مانند حالت یک متغیره که با یک متغیر تصادفی سروکار داشتیم و می‌توانستیم احتمال‌های مربوط به همه‌ی مقادیر X را به وسیله‌ی یک جدول نمایش دهیم. اینک می‌توانیم در حالت دومتغیره نیز احتمال‌های مربوط به همه‌ی جفت‌های مقادیر X و Y را به وسیله‌ی یک جدول نمایش دهیم.

مثال ۴-۴. دو مهره به تصادف از ظرفی که محتوی 3 مهره‌ی قرمز، 2 مهره‌ی آبی و 4 مهره‌ی سبز است، انتخاب می‌کنیم. اگر X و Y به ترتیب تعداد مهره‌های قرمز و مهره‌های آبی باشند که بین دو مهره‌ی منتخب از ظرف وجود دارند، احتمال‌های مربوط به همه‌ی جفت‌های مقادیر ممکن X و Y را بیایید.

حل. جفت‌های ممکن عبارت‌اند از $(0,0)$ ، $(0,1)$ ، $(1,0)$ ، $(1,1)$ ، $(0,2)$ ، و $(2,0)$. برای پیدا کردن احتمال‌های مربوط به $(1,0)$ ، مشاهده می‌کنیم که با پیشامد به دست آوردن یک مهره از سه مهره‌ی قرمز، 0 مهره از دو مهره‌ی آبی، و در نتیجه یک مهره از چهار مهره سبز سروکار داریم. تعداد راههایی که می‌توان این کار را انجام داد برابر

است، و تعداد کل راههایی که می‌توان ۲ مهره از ۹ مهره را انتخاب کرد برابر $\binom{9}{3} = 84$ است. چون این امکانات بنابر فرض تصادفی بودن انتخاب ها هم‌شانس‌اند از قضیه‌ی ۳-۲، نتیجه می‌شود که احتمال مربوط به $(1, 0)$ برابر $\frac{1}{84}$ است. همین‌طور، احتمال مربوط به $(0, 1)$ برابر

$$\frac{\binom{3}{1}\binom{2}{1}\binom{4}{0}}{84} = \frac{6}{84} = \frac{1}{14}$$

است و با ادامه‌ی این راه، مقادیری را که در جدول زیر نشان داده‌ایم به دست می‌آوریم.

| | | x | | |
|---|---|----------------|---------------|----------------|
| | | ۰ | ۱ | ۲ |
| y | ۰ | $\frac{1}{6}$ | $\frac{1}{3}$ | $\frac{1}{14}$ |
| | ۱ | $\frac{2}{9}$ | $\frac{1}{6}$ | ۰ |
| | ۲ | $\frac{1}{36}$ | ۰ | ۰ |

درواقع مانند حالت یک متغیره، عموماً بهتر است که چنین احتمال‌هایی را به وسیله‌ی فرمولی ارائه کنیم. به عبارت دیگر بهتر است که احتمال‌ها را به وسیله‌ی تابعی با مقادیر $f(x, y) = P(X = x, Y = y)$ ، برای هر جفت از مقادیر (x, y) در بردا متغیرهای تصادفی X و Y بیان کنیم. به عنوان نمونه، می‌توانیم برای جفت متغیرهای تصادفی مثال ۴-۴ به ازای $x = 0, 1, 2$ ، $y = 0, 1, 2$ ، $P(X = x, Y = y)$ ، بنویسیم

$$f(x, y) = \frac{\binom{3}{x}\binom{2}{y}\binom{4}{2-x-y}}{\binom{9}{3}}.$$

تعریف ۶-۴. اگر X و Y متغیرهای تصادفی گستته باشند، تابعی که با $f(x, y) = P(X = x, Y = y)$ برای هر جفت مقدار (x, y) در بردا X ، Y داده می‌شود، توزیع احتمال توأم X و Y خوانده می‌شود.

شبیه قضیه‌ی ۴-۱، از اصول موضوع احتمال نتیجه می‌شود که

قضیه‌ی ۷-۴. تابعی دو متغیره وقتی و تنها وقتی می‌تواند به عنوان توزیع احتمال توأم یک جفت متغیر تصادفی X و Y به کار رود، که مقادیر آن، $f(x,y)$ ، در شرایط زیر صدق کنند.

۱. به ازای هر جفت مقدار (x,y) در حوزه‌ی مربوط، $f(x,y) \geq 0$ ؛
۲. $\sum_x \sum_y f(x,y) = 1$ ، که مجموع یابی دوگانه به ازای تمام جفت (x,y) ‌های ممکن در حوزه‌ی مربوط انجام می‌شود.

مثال ۷-۴. مقدار k را طوری تعیین کنید که تابع

$$f(x,y) = kxy, \quad x=1,2,3; \quad y=1,2,3$$

را بتوان به عنوان توزیع احتمال توأم به کار برد.

حل. اگر مقادیر مختلف x و y را در تابع قراردهیم، به دست می‌آوریم $f(1,1) = k$ ، $f(2,1) = 2k$ ، $f(3,1) = 3k$ ، $f(1,2) = 2k$ ، $f(2,2) = 4k$ ، $f(3,2) = 6k$ ، $f(1,3) = 3k$ ، $f(2,3) = 6k$ ، $f(3,3) = 9k$. برای برقراری اوّلین شرط قضیه‌ی ۷-۴ ثابت k نباید منفی باشد، و برای برقراری دوّمین شرط

$$k + 2k + 3k + 4k + 5k + 6k + 7k + 8k + 9k = 1$$

$$\text{به قسمی که } 1 = \frac{1}{36}k \text{ و } k = 36$$

مانند حالت یک متغیره، در مسائل زیادی دانستن احتمال اینکه مقادیر دو متغیر تصادفی برابر با اعداد حقیقی x و y یا کوچکتر از آنها باشند، مورد توجه است.

تعریف ۷-۵. اگر X و Y متغیرهای تصادفی گستته باشند، تابعی که به صورت به ازای $-\infty < y < +\infty$ و $-\infty < x < +\infty$

$$F(x,y) = P(X \leq x, Y \leq y) = \sum_{s \leq x} \sum_{t \leq y} f(s,t)$$

داده می‌شود، و در آن $f(s,t)$ مقدار توزیع احتمال توأم X و Y در (s,t) است، تابع توزیع توأم، یا توزیع تجمعی توأم X و Y خوانده می‌شود.

مثال ۷-۶. با مراجعه به مثال ۷-۱۴، $F(1,1)$ را پیدا کنید.
حل.

$$\begin{aligned}
 F(1,1) &= P(X \leq 1, Y \leq 1) \\
 &= f(0,0) + f(0,1) + f(1,0) + f(1,1) \\
 &= \frac{1}{6} + \frac{2}{9} + \frac{1}{3} + \frac{1}{6} \\
 &= \frac{8}{9}.
 \end{aligned}$$

مانند حالت یک متغیره، تابع توزیع توانم دو متغیرتصادفی به ازای تمام اعداد حقیقی تعریف شده است؛ به عنوان نمونه برای مثال ۱۴-۴، $F(3/7, 4/5) = P(X \leq 3/7, Y \leq 4/5) = 1$ و $F(-2, 1) = P(X \leq -2, Y \leq 1) = 0$.

حال مفاهیمی را که تاکنون در این بخش معرف کردہ‌ایم به حالت پیوسته تعمیم می‌دهیم.

تعریف ۴-۸ یک تابع دو متغیره با مقادیر $f(x,y)$ ، که روی صفحه‌ی xy ، تعریف شده است، تابع چگالی احتمال توانم متغیرهای تصادفی X و Y خوانده می‌شود، اگر و تنها اگر، برای هر ناحیه‌ی A از صفحه‌ی xy

$$P((X,Y) \in A) = \iint_A f(x,y) dx dy.$$

قضیه‌ی ۴-۸ تابعی دو متغیره را می‌توان به عنوان تابع چگالی احتمال توانم یک جفت متغیر تصادفی پیوسته‌ی X و Y به کار برد که مقادیر آن $(f(x,y))$ ، در شرایط زیر صدق کنند:

۱. به ازای $f(x,y) \geq 0$ ؛ $-\infty < y < +\infty$ ، $-\infty < x < +\infty$ ،

$$2. \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f(x,y) dx dy = 1.$$

مثال ۴-۱۷. اگر تابع چگالی احتمال توانم دو متغیر تصادفی X و Y به صورت

$$f(x,y) = \begin{cases} \frac{3}{5}x(y+x) & 0 < x < 1, 0 < y < 2 \\ 0 & \text{سایر جاهات} \end{cases}$$

باشد، $P((X,Y) \in A)$ را که در آن A ، ناحیه‌ی $\{(x,y) | 0 < x < \frac{1}{2}, 1 < y < 2\}$ است

باید.

حل.

$$\begin{aligned} P[(X,Y) \in A] &= P(0 < X < \frac{1}{\sqrt{2}}, 1 < Y < 2) \\ &= \int_0^{\frac{1}{\sqrt{2}}} \int_{\frac{y}{\sqrt{2}}}^{\frac{1}{\sqrt{2}}x} x(y+x) dx dy = \int_0^{\frac{1}{\sqrt{2}}} \left(\frac{yx^2}{2} + \frac{x^3}{3} \right) \Big|_{x=\frac{y}{\sqrt{2}}} dy \\ &= \int_0^{\frac{1}{\sqrt{2}}} \left(\frac{y^3}{4} + \frac{1}{4} \right) dy = \frac{y^4}{16} + \frac{y}{4} \Big|_0^{\frac{1}{\sqrt{2}}} = \frac{11}{80}. \end{aligned}$$

تمرین

۴-۳۵. اگر توزیع توأم (X,Y) به صورت

$$f(x,y) = \begin{cases} \frac{3x+2y}{50} & x=1, 2; y=1, 2, 3 \\ 0 & \text{سایر جاهای} \end{cases}$$

باشد، مقدار $F(1, 2)$ چقدر است؟

۴-۳۶. اگر تابع چگالی توأم X و Y به صورت

$$f(x,y) = \begin{cases} 1 & x > 0, y > 0, x+y < 1 \\ 0 & \text{سایر جاهای} \end{cases}$$

باشد، $P(\frac{1}{3} < X < \frac{2}{3}, \frac{1}{8} < Y < \frac{1}{3})$ را به دست آورید.

۴-۳۷. جعبه‌ای را در نظر گیرید که r مهره‌ی یکسان با شماره‌های $1, 2, \dots, r$ در آن

باشدند. یک نمونه‌ی تصادفی n تایی از این جعبه

الف) بدون جایگذاری،

ب) با جایگذاری،

انتخاب می‌کنیم. اگر X نمایانگر کوچک‌ترین عدد انتخابی در نمونه و X_r نمایانگر

بزرگ‌ترین عدد انتخابی در نمونه باشد، توزیع توأم (X, X_r) را باید.

۴-۳۸. مقدار k را به نحوی باید تا تابع زیر یک تابع چگالی توأم باشد.

$$f(x,y) = \begin{cases} \frac{k}{y} & 0 < x < y, 0 < y < 1 \\ 0 & \text{سایر جاهای} \end{cases}$$

۴-۶ توزیع‌های حاشیه‌ای

برای معرفی مفهوم توزیع حاشیه‌ای، ابتدا مثال زیر را در نظر می‌گیریم.

مثال ۴-۱۸. در مثال ۴-۱۴ توزیع احتمال توان متفاوت‌های تصادفی X و Y ، یعنی تعداد مهره‌های قرمز و تعداد مهره‌های آبی موجود بین مهره‌ای را که به تصادف از ظرفی محتوی ۳ مهره‌ی قرمز، ۲ مهره‌ی آبی، و ۴ مهره‌ی سبز بیرون می‌آوریم، به دست آوردیم. توزیع احتمال X و توزیع احتمال Y را به طور جداگانه پیدا کنید.
حل. نتایج مثال ۴-۱۴، همراه با مجموع‌های حاشیه‌ای، یعنی مجموع‌های مربوط به هر سطر و هر ستون را در جدول زیر نشان داده‌ایم.

| | | x | | |
|-----|----------------|---------------|----------------|----------------|
| | | ۰ | ۱ | ۲ |
| y | ۰ | $\frac{1}{6}$ | $\frac{1}{3}$ | $\frac{1}{12}$ |
| | ۱ | $\frac{2}{9}$ | $\frac{1}{6}$ | $\frac{7}{18}$ |
| ۲ | $\frac{1}{36}$ | ۰ | ۰ | $\frac{1}{36}$ |
| | $\frac{5}{12}$ | $\frac{1}{2}$ | $\frac{1}{12}$ | |

مجموع‌های ستونی احتمال‌هایی هستند که X مقادیر ۰، ۱، و ۲ را اختیار می‌کند. به عبارت دیگر این مجموع‌ها، مقادیر

$$g(x) = \sum_{y=0}^2 f(x,y) \quad x = 0, 1, 2$$

مربوط به توزیع احتمال X هستند. به همین ترتیب، مجموع‌های سطري، مقادیر

$$h(y) = \sum_{x=0}^2 f(x,y) \quad y = 0, 1, 2$$

مربوط به توزیع احتمال Y هستند.

لذا به تعریف زیر هدایت می‌شویم.

تعریف ۴-۹. اگر X و Y متغیرهای تصادفی گسسته، و $f(x,y)$ مقدار توزیع احتمال توان آنها در (x,y) باشد، تابعی که برای هر x در برد X به صورت

$$g(x) = \sum_y f(x,y),$$

داده می‌شود، توزیع حاشیه‌ای X نام دارد. متناظرً تابعی که برای هر y در برد Y به صورت

$$h(y) = \sum_x f(x, y),$$

داده می‌شود، توزیع حاشیه‌ای Y نام دارد.

وقتی X و Y متغیرهای تصادفی پیوسته‌اند، چگالی‌های احتمال به جای توزیع‌های احتمال، و انتگرال‌ها به جای مجموع‌ها قرار می‌گیرند، و به دست می‌آوریم:

تعريف ۴-۱۰. اگر X و Y متغیرهای تصادفی پیوسته و $f(x, y)$ مقدار چگالی احتمال توأم آنها در (x, y) باشد، تابعی که به صورت

$$g(x) = \int_{-\infty}^{+\infty} f(x, y) dy, \quad -\infty < x < +\infty$$

داده می‌شود، چگالی حاشیه‌ای X نام دارد. متناظرً تابعی که به صورت

$$h(y) = \int_{-\infty}^{+\infty} f(x, y) dx, \quad -\infty < y < +\infty$$

داده می‌شود، چگالی حاشیه‌ای Y نام دارد.

مثال ۴-۱۹. چگالی توأم

$$f(x, y) = \begin{cases} \frac{1}{\pi}(x + 2y) & 0 < x < 1, 0 < y < 1 \\ 0 & \text{سایر جاهای} \end{cases}$$

داده شده است، چگالی‌های حاشیه‌ای X و Y را بیابید.
حل. با انجام انتگرال‌گیری‌های لازم برای $x < 1 < y$ ، به دست می‌آوریم

$$g(x) = \int_{-\infty}^{+\infty} f(x, y) dy = \int_0^1 \frac{1}{\pi}(x + 2y) dy = \frac{1}{\pi}(x + 1),$$

و برای سایر مقادیر x داریم $g(x) = 0$. به همین طریق برای $y < 1$ ،

$$h(y) = \int_{-\infty}^{+\infty} f(x, y) dx = \int_0^1 \frac{1}{\pi}(x + 2y) dx = \frac{1}{\pi}(1 + 2y),$$

و برای سایر مقادیر y ، $h(y) = 0$.

وقتی با بیش از دو متغیر تصادفی سروکار داریم، نه تنها می‌توانیم از توزیع‌های حاشیه‌ای تک تک متغیرهای تصادفی صحبت کنیم، بلکه می‌توانیم از توزیع‌های حاشیه‌ای توأم چند متغیر تصادفی هم صحبت به میان آوریم. اگر توزیع احتمال توأم

متغیرهای تصادفی گستته‌ی X_1, X_2, \dots, X_n دارای مقادیر $f(x_1, x_2, \dots, x_n)$ باشد، توزیع حاشیه‌ای X_1 به تنها یی برای تمام مقادیر واقع در برد X_1 به صورت

$$g(x_1) = \sum_{x_2} \cdots \sum_{x_n} f(x_1, x_2, \dots, x_n),$$

داده می‌شود. توزیع حاشیه‌ای توأم X_1, X_2 برای همه مقادیر در برد X_1, X_2 به صورت

$$m(x_1, x_2, x_3) = \sum_{x_4} \cdots \sum_{x_n} f(x_1, x_2, \dots, x_n),$$

داده می‌شود، و توزیع‌های حاشیه‌ای دیگر را نیز می‌توان به همین طریق تعریف کرد. برای حالت پیوسته، چگالی‌های احتمال به جای توزیع‌های احتمال و انتگرال‌ها به جای مجموع‌ها قرار می‌گیرند، و اگر چگالی احتمال توأم متغیرهای تصادفی پیوسته‌ی X_1, X_2, \dots, X_n مقدار $f(x_1, x_2, \dots, x_n)$ را داشته باشند، چگالی حاشیه‌ای X_1 تنها یی برای $-\infty < x_1 < +\infty$ به صورت

$$h(x_1) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f(x_1, x_2, \dots, x_n) dx_2 dx_3 \dots dx_n,$$

داده می‌شود، چگالی حاشیه‌ای توأم X_1 و X_n برای $-\infty < x_1 < +\infty$ ، $-\infty < x_n < +\infty$ به صورت

$$\varphi(x_1, x_n) = \int_{-\infty}^{+\infty} \cdots \int_{-\infty}^{+\infty} f(x_1, x_2, \dots, x_n) dx_2 dx_3 \dots dx_{n-1},$$

داده می‌شود، و قس علی‌هذا.

مثال ۴-۲۰. چگالی احتمال سه متغیره‌ی

$$f(x_1, x_2, x_3) = \begin{cases} (x_1 + x_2)e^{-x_3} & 0 < x_1 < 1, 0 < x_2 < 1, x_3 > 0 \\ 0 & \text{سایر جاهای} \end{cases}$$

را در نظر می‌گیریم. چگالی حاشیه‌ای توأم X_1 و X_2 و چگالی حاشیه‌ای X_1 را بیابید. حل. با انجام انتگرال‌گیری‌های لازم، چگالی حاشیه‌ای توأم X_1 و X_2 را برای $0 < x_1 < 1$ و $0 < x_2 < 1$ به صورت زیر پیدا می‌کنیم:

$$m(x_1, x_2) = \int_0^1 (x_1 + x_2)e^{-x_3} dx_3 = (x_1 + \frac{1}{x_2}) e^{-x_3},$$

و برای سایر مقادیر داریم $m(x_1, x_2) = 0$. با استفاده از این نتیجه، متوجه می‌شویم که چگالی حاشیه‌ای X_1 تنها، برای $0 < x_1 < 1$ ، به صورت

$$\begin{aligned}
 g(x_1) &= \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f(x_1, x_2, x_3) dx_2 x_3 = \int_{-\infty}^{+\infty} m(x_1, x_3) dx_3 \\
 &= \int_{-\infty}^{+\infty} \left(x_1 + \frac{1}{\mu} \right) e^{-x_3} dx_3 = x_1 + \frac{1}{\mu} \\
 \text{داده می‌شود و به ازای سایر مقادیر } g(x_1) &= 0.
 \end{aligned}$$

تمرین

۴-۳۹. برای تمرین ۴-۳۵، توابع چگالی حاشیه‌ای را به دست آورید.

۴-۴۰. برای تمرین ۴-۳۶، توابع چگالی حاشیه‌ای را به دست آورید.

۴-۴۱. توابع چگالی حاشیه‌ای را از روی توزیع توأم زیر به دست آورید

| x | y | ۱ | ۲ | ۳ | $f(x)$ |
|--------|-----|---------------|---------------|---------------|--------|
| ۱ | | $\frac{1}{9}$ | ۰ | ۰ | |
| ۲ | | $\frac{2}{9}$ | $\frac{1}{9}$ | ۰ | |
| ۳ | | $\frac{2}{9}$ | $\frac{2}{9}$ | $\frac{1}{9}$ | |
| $f(y)$ | | | | | ۱ |

۴-۴۲. اگر $F(x,y)$ مقدار تابع توزیع توأم دو متغیر تصادفی X و Y باشد، برحسب کمیت‌های $P(a < X \leq b, c < Y \leq d)$ ، $F(b,d)$ ، $F(b,c)$ ، $F(a,d)$ ، $F(a,c)$ را بیابید.

۷-۴ متغیرهای تصادفی مستقل

وقتی با دو یا چند متغیر تصادفی سروکار داریم، مسائل استقلال معمولاً اهمیت زیادی دارند.

تعريف ۷-۴. اگر $f(x_1, x_2, \dots, x_n)$ مقدار توزیع احتمال توأم n متغیر تصادفی X_1, X_2, \dots, X_n در (x_1, x_2, \dots, x_n) و $f_i(x_i)$ مقدار توزیع حاشیه‌ای در X_i به ازای $x_i = i, 1, 2, \dots, n$ باشد، این متغیرها مستقل‌اند اگر و تنها اگر به ازای کلیه‌ی (x_1, x_2, \dots, x_n) در برداشان

$$f(x_1, x_2, \dots, x_n) = f_1(x_1) f_2(x_2) \dots f_n(x_n).$$

برای ارائه‌ی تعریفی متناظر در مورد متغیرهای تصادفی پیوسته، تنها کلمه‌ی چگالی را به جای کلمه‌ی توزیع قرار می‌دهیم.

با این تعریف برای استقلال، به آسانی می‌توان تحقیق کرد که سه متغیر تصادفی مثل X_1, X_2, X_3 مستقل نیستند، اما دو متغیر تصادفی X_1 و X_2 همچنین دو متغیر تصادفی X_2 و X_3 دو به دو مستقل‌اند.

مثال‌های زیر برای نشان‌دادن نحوه‌ی استفاده از تعریف ۴-۱۱ در یافتن احتمال‌های مربوط به چندین متغیر تصادفی مستقل، به کار می‌روند.

مثال ۴-۲۱. n پرتاب مستقل یک سکه‌ی همگن را در نظر می‌گیریم. فرض کنید X_i تعداد (۰ یا ۱) شیر حاصل در نامین پرتاب، $i=1, 2, \dots, n$ باشد. توزیع احتمال توأم این n متغیر تصادفی را بباید.

حل. چون هر یک از متغیرهای تصادفی X_i ، به ازای $i=1, 2, \dots, n$ ، دارای توزیع احتمال

$$f_i(x_i) = \begin{cases} \frac{1}{2} & x_i = 0, 1 \\ 0 & \text{سایر جاهای} \end{cases}$$

است، و n متغیر تصادفی مستقل‌اند، توزیع احتمال توأم آنها با

$$f(x_1, x_2, \dots, x_n) = f_1(x_1) f_2(x_2) \dots f_n(x_n)$$

$$= \frac{1}{2} \cdot \frac{1}{2} \cdots \frac{1}{2} = \left(\frac{1}{2}\right)^n,$$

داده می‌شود، که در آن هر x_i ، به ازای $i=1, 2, \dots, n$ ، برابر ۰ یا ۱ است.

مثال ۴-۲۲. متغیرهای تصادفی مستقل X_1, X_2, X_3 با چگالی‌های احتمال

$$f_1(x_1) = \begin{cases} e^{-x_1} & x_1 > 0 \\ 0 & \text{سایر جاهای} \end{cases}$$

$$f_2(x_2) = \begin{cases} 2e^{-2x_2} & x_2 > 0 \\ 0 & \text{سایر جاهای} \end{cases}$$

$$f_3(x_3) = \begin{cases} 3e^{-3x_3} & x_3 > 0 \\ 0 & \text{سایر جاهای} \end{cases}$$

داده شده‌اند، تابع چگالی احتمال توأم آنها را بباید، و آن را برای محاسبه‌ی $P(X_1 + X_2 \leq 1, X_3 > 1)$ به کار برد.

حل. مقادیر چگالی احتمال توأم به ازای $x_1 > 0, x_2 > 0, x_3 > 0$ به صورت

$$\begin{aligned} f(x_1, x_2, x_3) &= f_1(x_1)f_2(x_2)f_3(x_3) \\ &= e^{-x_1} \cdot 2e^{-2x_2} \cdot 3e^{-3x_3} = 6e^{-x_1 - 2x_2 - 3x_3}, \end{aligned}$$

داده می‌شود، و به ازای سایر مقادیر $f(x_1, x_2, x_3) = 0$. پس

$$\begin{aligned} P(X_1 + X_2 \leq 1, X_3 > 1) &= \int_{-\infty}^{\infty} \int_0^1 \int_0^{1-x_2} 6e^{-x_1 - 2x_2 - 3x_3} dx_3 dx_2 dx_1 \\ &= (1 - 2e^{-1} + e^{-2})e^{-3} \\ &= 0.020. \end{aligned}$$

تمرین

۴-۴۳. برای تمرین ۴-۳۵ تعیین کنید، متغیرهای تصادفی مستقل‌اند یا مستقل نیستند.

۴-۴۴. اگر توزیع احتمال توأم سه متغیر تصادفی گستته X, Y ، و Z به صورت زیر باشد

$$f(x, y, z) = \frac{(x+y)z}{6^3}, \quad x = 1, 2; \quad y = 1, 2, 3; \quad z = 1, 2$$

الف) آیا سه متغیر تصادفی از هم مستقل‌اند؟

ب) مقدار $P(X = 2, Y + Z \leq 3)$ را بباید.

۴-۴۵. توابع چگالی حاشیه‌ای را از روی توزیع توأم زیر به دست آورید، سپس استقلال دو متغیر تصادفی را مورد بررسی قرار دهید.

| x | y | ۱ | ۲ | ۳ | $f(x)$ |
|-----|---------------|---------------|---------------|---|--------|
| | $f(y)$ | | | | ۱ |
| ۱ | ۰ | $\frac{1}{6}$ | $\frac{1}{6}$ | | |
| ۲ | $\frac{1}{6}$ | ۰ | $\frac{1}{6}$ | | |
| ۳ | $\frac{1}{6}$ | $\frac{1}{6}$ | ۰ | | |

۴-۸ توزیع‌های شرطی

در فصل سوم، احتمال شرطی پیشامد A را وقتی که پیشامد B معلوم است، به شرط $P(B) \neq 0$ به صورت

$$P(A|B) = \frac{P(A \cap B)}{P(B)},$$

تعریف کردیم. فرض کنید که اکنون A و B پیشامدهای $X = x$ و $Y = y$ هستند،

بنابراین می‌توانیم به شرط $P(Y = y) = h(y) \neq 0$ بنویسیم

$$P(X = x | Y = y) = \frac{P(X = x, Y = y)}{P(Y = y)} = \frac{f(x, y)}{h(y)},$$

که در آن $f(x, y)$ مقدار توزیع احتمال توأم X و Y در (x, y) ، و $h(y)$ مقدار توزیع حاشیه‌ای Y در y است. احتمال شرطی را به صورت $f(x|y)$ می‌نویسیم تا نشان دهیم که x مقدار متغیر و y تثبیت شده است. اینک تعریف زیر را ارائه می‌دهیم:

تعریف ۴-۱۲. اگر $f(x, y)$ مقدار توزیع احتمال توأم متغیرهای تصادفی گسسته‌ی X و Y در (x, y) ، و $h(y)$ مقدار توزیع حاشیه‌ای Y در y باشد، تابعی که برای هر x در برد X به صورت

$$f(x|y) = \frac{f(x, y)}{h(y)}, \quad h(y) \neq 0$$

است، توزیع شرطی X به شرط $Y = y$ نامیده می‌شود. متناظرًا، اگر $g(x)$ ، مقدار توزیع حاشیه‌ای X در x باشد، تابعی که برای هر y در برد Y به صورت

$$w(y|x) = \frac{f(x, y)}{g(x)}, \quad g(x) \neq 0$$

است، توزیع شرطی Y به شرط $X = x$ نامیده می‌شود.

روشن است برای دو متغیر تصادفی مستقل داریم:

$$f(x|y) = g(x), \quad w(y|x) = h(y).$$

مثال ۴-۲۳. اگر توزیع توأم دو متغیر تصادفی به صورت

| x | $y = \circ$ | $y = ।$ | $y = ॥$ | $g(x)$ |
|--------|---------------|---------------|---------------|---------------|
| ۱ | ۰ | ۰ | $\frac{1}{6}$ | $\frac{1}{6}$ |
| ۲ | $\frac{1}{3}$ | ۰ | $\frac{1}{6}$ | $\frac{1}{2}$ |
| ۳ | ۰ | $\frac{1}{3}$ | ۰ | $\frac{1}{3}$ |
| $h(y)$ | $\frac{1}{3}$ | $\frac{1}{3}$ | $\frac{1}{3}$ | ۱ |

باشد، سه چگالی شرطی مطرح به شرط y عبارتند از:

| $(x y = \circ)$ | ۱ | ۲ | ۳ |
|-------------------|---|---|---|
| $f(x \circ)$ | $\frac{0}{\frac{1}{3}} = 0$ | $\frac{\frac{1}{3}}{\frac{1}{3}} = 1$ | $\frac{0}{\frac{1}{3}} = 0$ |
| $(x y = ।)$ | ۱ | ۲ | ۳ |
| $f(x ।)$ | $\frac{0}{\frac{1}{3}} = 0$ | $\frac{0}{\frac{1}{3}} = 0$ | $\frac{\frac{1}{3}}{\frac{1}{3}} = 1$ |
| $(x y = ॥)$ | ۱ | ۲ | ۳ |
| $f(x ॥)$ | $\frac{\frac{1}{6}}{\frac{1}{3}} = \frac{1}{2}$ | $\frac{\frac{1}{6}}{\frac{1}{3}} = \frac{1}{2}$ | $\frac{\frac{1}{6}}{\frac{1}{3}} = \frac{1}{2}$ |

| $(x y = \circ)$ | ۱ | ۲ | ۳ |
|-------------------|---|---|---|
| $f(x \circ)$ | $\frac{0}{\frac{1}{3}} = 0$ | $\frac{\frac{1}{3}}{\frac{1}{3}} = 1$ | $\frac{0}{\frac{1}{3}} = 0$ |
| $(x y = ।)$ | ۱ | ۲ | ۳ |
| $f(x ।)$ | $\frac{0}{\frac{1}{3}} = 0$ | $\frac{0}{\frac{1}{3}} = 0$ | $\frac{\frac{1}{3}}{\frac{1}{3}} = 1$ |
| $(x y = ॥)$ | ۱ | ۲ | ۳ |
| $f(x ॥)$ | $\frac{\frac{1}{6}}{\frac{1}{3}} = \frac{1}{2}$ | $\frac{\frac{1}{6}}{\frac{1}{3}} = \frac{1}{2}$ | $\frac{\frac{1}{6}}{\frac{1}{3}} = \frac{1}{2}$ |

تمرین

- ۴-۴۶. با توجه به اطلاعات مثال ۴-۲۳، چگالی‌های شرطی $w(y | x)$ را به دست آورید.
 ۴-۴۷. نشان دهید که $f(x | y)$ و $w(y | x)$ ویژگی‌های تابع چگالی را دارا می‌باشند.

فصل پنجم

امید ریاضی

۱-۵ مقدمه

مفهوم امید ریاضی، در ابتدا، در ارتباط با بازی‌های شانسی به وجود آمده است، و در ساده‌ترین صورتش، حاصل ضرب مبلغی است که بازیکن امکان برد آن را دارد در احتمال آنکه برنده شود. امروزه کاربرد آن فقط منحصر به این موضوع نیست. به سبب سادگی از مثالی که بسیار رایج است برای ارائه‌ی مفهوم امید ریاضی استفاده می‌کنیم: فروشگاهی از میان ۱۰۰۰۰ مشتری، جایزه‌ای به ارزش ۱۰۰ سکه‌ی طلا برای مشتری که به تصادف انتخاب شود، در نظر گرفته، پس امید ریاضی دریافت جایزه برای هر مشتری $\frac{1}{10000} \times 100\% = 1\%$ سکه‌ی طلا است. این رقم به مفهوم یک متوسط تعبیر می‌شود، بدین معنا که مقدار جایزه‌ی ۱۰۰۰۰ مشتری جمعاً ۱۰۰ سکه‌ی طلا، یا به‌طور متوسط به هر مشتری $\frac{100}{10000} \times 100\% = 1\%$ سکه‌ی طلا می‌رسد. اگر جایزه‌ی دوستی به ارزش ۵۰ سکه‌ی طلا و جایزه‌ی سومی به ارزش ۳۰ سکه‌ی طلا نیز وجود داشته باشند، استدلال می‌کنیم که مقدار جایزه‌ی برای ۱۰۰۰۰ مشتری جمعاً

$$100 + 50 + 30 = 180$$

سکه‌ی طلا، یا به‌طور متوسط مقدار جایزه‌ی هر مشتری $\frac{180}{10000} \times 100\% = 1.8\%$ سکه‌ی طلا است. اگر به‌طریق دیگری به مطلب نگاه کنیم، می‌توانیم استدلال نمائیم که اگر قرعه‌کشی به دفعات زیادی تکرار شود، در 99.97% درصد از دفعات (یا با احتمال 99.97%) جایزه‌ای نصیب‌مان نمی‌شود و در 0.03% درصد از دفعات (یا با احتمال

یکی از سه جایزه را می‌بریم. پس به‌طور متوسط خواهیم برد که مجموع حاصل ضرب‌های حاصل از ضرب هر مبلغ در احتمال متناظر با آن است.

۲-۵ مقدار امید ریاضی یک متغیر تصادفی

در مثال بخش قبل، مبلغی که امکان برد آن را داشتیم یک متغیر تصادفی گستته بود، و امید ریاضی این متغیر تصادفی مجموع حاصل ضرب‌های حاصل از ضرب هر مقدار متغیر تصادفی در احتمال متناظر با آن بود. لذا با نامیدن امید ریاضی یک متغیر تصادفی صرفاً به عنوان مقدار مورد انتظار، و تعمیم این تعریف به حالت پیوسته، و با قرار دادن عمل انتگرال‌گیری به جای مجموع یابی، داریم:

تعریف ۱-۵. اگر X یک متغیر تصادفی گستته و f توزیع احتمال آن باشد، مقدار امید ریاضی X برابر است با

$$E(X) = \sum_x x f(x),$$

متناظراً، اگر X متغیر تصادفی پیوسته و f تابع چگالی احتمال آن باشد، مقدار امید ریاضی X برابر است با

$$E(X) = \int_{-\infty}^{+\infty} x f(x) dx.$$

البته، در این تعریف فرض شده است که مجموع یا انتگرال موجود است؛ در غیر این صورت، امید ریاضی قابل تعریف نیست.

مثال ۱-۵. محموله‌ایی مرکب از ۱۲ دستگاه کامپیوتر، شامل دو کامپیوتر معیوب است. اگر سه دستگاه کامپیوتر را برای ارسال به یک هتل به تصادف انتخاب کنیم، وجود چند دستگاه معیوب را می‌توان انتظار داشت؟

حل. می‌توانیم x تا از ۲ کامپیوتر معیوب و $3-x$ تا از ۱۰ کامپیوتر سالم را به راه انتخاب کنیم، و می‌توانیم $\binom{2}{x}$ راه انتخاب کنیم، و $\binom{10}{3-x}$ راه انتخاب کنیم، و می‌توانیم $\binom{12}{3}$ کامپیوتر را به

برگزینیم. با فرض اینکه همهی $\binom{12}{3}$ امکان، هم شانس باشند، متوجه می‌شویم که توزیع احتمال X ، تعداد کامپیوترهای معیوبی که به یک مرکز آموزشی فرستاده می‌شود، به صورت

$$f(x) = \frac{\binom{2}{x} \binom{10}{3-x}}{\binom{12}{3}}, \quad x = 0, 1, 2$$

یا به صورت جدول

| | | | |
|--------|----------------|----------------|----------------|
| x | ۰ | ۱ | ۲ |
| $f(x)$ | $\frac{6}{11}$ | $\frac{9}{22}$ | $\frac{1}{22}$ |
| | ۱۱ | ۲۲ | ۲۲ |

است. حال

$$E(X) = 0 \times \frac{6}{11} + 1 \times \frac{9}{22} + 2 \times \frac{1}{22} = \frac{1}{2},$$

و چون امکان به دست آوردن یک نیمه کامپیوتر معیوب وجود ندارد، روشن است که اصطلاح «انتظار» به مفهوم محاوره‌ای آن مطرح نیست. در واقع این اصطلاح را باید به عنوان یک متوسط مربوط به تکرار ارسال محموله‌ها تحت شرایط داده شده تعبیر کرد.

مثال ۵-۲. اندازه‌گیری‌های کدگذاری شدهی خاصی از فاصله‌ی بین دو دندانهی پیچ در یک بست، دارای چگالی احتمال

$$f(x) = \begin{cases} \frac{4}{\pi(1+x^2)} & 0 < x < 1 \\ 0 & \text{سایر جاهات} \end{cases}$$

است. مقدار مورد انتظار این متغیر تصادفی را پیدا کنید.

حل. با استفاده از تعریف ۱-۵ داریم

$$E(X) = \int_0^1 x \frac{4}{\pi(1+x^2)} dx = \frac{4}{\pi} \int_0^1 \frac{x}{1+x^2} dx = \frac{\ln 4}{\pi} = 0.4413.$$

در بسیاری از مسائل آمار، نه تنها مقدار امید ریاضی یک متغیر تصادفی X ، بلکه مقادیر امید ریاضی متغیرهای تصادفی وابسته به X نیز مورد توجه‌اند. مثلاً ممکن است

متغیر تصادفی Y مورد توجه ما باشد که مقادیرش با مقادیر X از طریق معادله‌ی $y = g(x)$ در ارتباط‌اند. برای ساده‌کردن نمادگذاری، این متغیر تصادفی را به صورت $(X, g(X))$ نشان می‌دهیم؛ برای مثال $(X, g(X))$ ممکن است X باشد، به قسمی که وقتی X ، مقدار ۲ را اختیار می‌کند، $(X, g(X))$ ، مقدار $8 = 2^3$ را اختیار کند. اگر بخواهیم مقدار امید ریاضی چنین متغیر تصادفی $(X, g(X))$ را پیدا کنیم، می‌توانیم ابتدا توزیع احتمال یا چگالی احتمال آن را بیابیم، و آنگاه تعریف ۱-۴ را به کاربریم، اما معمولاً به کار بردن قضیه‌ی زیر آسان‌تر و سرراست‌تر است:

قضیه‌ی ۱-۴. اگر X متغیر تصادفی گستته، و $f(x)$ مقدار توزیع احتمال آن به ازای x باشد، مقدار امید ریاضی متغیر تصادفی $(X, g(X))$ ،

$$E[g(X)] = \sum_x g(x) f(x)$$

است. متناظرًا اگر X متغیر تصادفی پیوسته و $f(x)$ مقدار چگالی احتمال آن به ازای x باشد، مقدار امید ریاضی متغیر تصادفی $(X, g(X))$ ،

$$E[g(X)] = \int_{-\infty}^{+\infty} g(x) f(x) dx$$

است.

برهان. این قضیه را فقط برای حالتی که X گستته است و برد متناهی دارد اثبات می‌کنیم. چون $y = g(x)$ الزاماً یک تناظر یک‌به‌یک را تعریف نمی‌کند. فرض می‌کنیم که وقتی x مقادیر $x_{i_1}, x_{i_2}, \dots, x_{i_{n_i}}$ را اختیار می‌نماید، $(X, g(X))$ مقدار i را اختیار کند. در این صورت احتمال اینکه $(X, g(X))$ مقدار i را اختیار کند

$$P[g(X) = g_i] = \sum_{j=1}^{n_i} f(x_{ij})$$

و اگر $(x, g(x))$ مقادیر g_1, g_2, \dots, g_m را اختیار کند، نتیجه می‌شود که

$$\begin{aligned}
 E[g(X)] &= \sum_{i=1}^m g_i P[g(X)=g_i] \\
 &= \sum_{i=1}^m g_i \sum_{j=1}^{n_i} f(x_{ij}) \\
 &= \sum_{i=1}^m \sum_{j=1}^{n_i} g_i f(x_{ij}) \\
 &= \sum g(x) f(x)
 \end{aligned}$$

که در آن مجموع یابی روی همهٔ مقادیر X انجام می‌شود.

مثال ۳-۵. اگر X عددی باشد که در ریختن یک تاس سالم ظاهر می‌شود، مقدار امید ریاضی متغیر تصادفی $g(X) = 2X + 1$ را بیابید.

حل. چون هر برآمد ممکن دارای احتمال $\frac{1}{6}$ است، به دست می‌آوریم:

$$\begin{aligned}
 E[g(X)] &= \sum_{x=1}^6 (2x + 1) \frac{1}{6} \\
 &= (2 \times 1 + 1) \frac{1}{6} + \dots + (2 \times 6 + 1) \frac{1}{6} = \frac{94}{6}.
 \end{aligned}$$

مثال ۴-۵. اگر X دارای چگالی احتمال

$$f(x) = \begin{cases} e^{-x} & x > 0 \\ 0 & \text{سایر جاهای} \end{cases}$$

باشد، مقدار امید ریاضی متغیر تصادفی $g(X) = e^{\frac{x}{4}X}$ را پیدا کنید.

حل. بنابر قضیهٔ ۱-۵ داریم:

$$E[e^{\frac{x}{4}X}] = \int_0^{+\infty} e^{\frac{x}{4}x} e^{-x} dx = \int_0^{+\infty} e^{\frac{-1}{4}x} dx = 4.$$

تعیین امیدهای ریاضی را اغلب می‌توان با استفاده از قضیه‌های زیر ساده کرد. این قضیه‌ها ما را قادر می‌سازند که مقادیر امید را از روی امیدهای دیگری که معلوم‌اند و یا به راحتی قابل محاسبه‌اند، حساب کنیم. چون مراحل اثبات، چه برای متغیرهای تصادفی پیوسته و چه برای متغیرهای گسسته یکی هستند، برخی برهان‌ها را برای حالت گسسته

و یا برای حالت پیوسته خواهیم داد؛ سایر برهان‌ها را به عنوان تمرین به عهده‌ی خواننده می‌گذاریم.

قضیه‌ی ۲-۵. اگر a و b مقادیر ثابتی باشند، آنگاه $E(aX + b) = aE(X) + b$ برهان. با استفاده از قضیه‌ی ۴-۱، با $g(X) = aX + b$ ، به دست می‌آوریم

$$\begin{aligned} E(aX + b) &= \int_{-\infty}^{+\infty} (ax + b) f(x) dx \\ &= a \int_{-\infty}^{+\infty} x f(x) dx + b \int_{-\infty}^{+\infty} f(x) dx \\ &= aE(X) + b. \end{aligned}$$

اگر به ترتیب قرار دهیم $a = ۰$ ، $b = ۰$ ، از قضیه‌ی ۲-۵ نتیجه می‌شود که

نتیجه‌ی ۱. اگر a مقداری ثابت باشد، آنگاه $E(aX) = aE(X)$.

نتیجه‌ی ۲. اگر b مقداری ثابت باشد، آنگاه $E(b) = b$.

مالحظه کنید که اگر بنویسیم $E(b)$ ، ثابت b را می‌توان متغیری تصادفی تلقی کرد که همیشه مقدار b را اختیار می‌کند.

قضیه‌ی ۳-۵. اگر c_1, c_2, \dots, c_n مقادیری ثابت باشند، آنگاه

$$E\left[\sum_{i=1}^n c_i g_i(X)\right] = \sum_{i=1}^n c_i E[g_i(X)].$$

برهان. بنابر قضیه‌ی ۱.۵ با $g(X) = \sum_{i=1}^n c_i g_i(X)$

$$\begin{aligned} E\left[\sum_{i=1}^n c_i g_i(X)\right] &= \sum_x \left[\sum_{i=1}^n c_i g_i(x) \right] f(x) \\ &= \sum_{i=1}^n \sum_x c_i g_i(x) f(x) \\ &= \sum_{i=1}^n c_i \sum_x g_i(x) f(x) \\ &= \sum_{i=1}^n c_i E[g_i(X)]. \end{aligned}$$

مثال ۵-۵. با استفاده از این واقعیت که

$$E(X^r) = (1^r + 2^r + 3^r + 4^r + 5^r + 6^r) \times \frac{1}{6} = \frac{91}{6},$$

برای متغیر تصادفی مثال ۳-۴ مثال را مجدداً حل کنید.
حل.

$$E(2X^r + 1) = 2E(X^r) + 1 = 2 \times \frac{91}{6} + 1 = \frac{94}{3}.$$

مثال ۵-۶. اگر چگالی احتمال متغیر تصادفی X به صورت زیر باشد

$$f(x) = \begin{cases} 2(1-x) & 0 < x < 1 \\ 0 & \text{سایر جاهای} \end{cases}$$

الف) نشان دهید که $E(X^r) = \frac{2}{(r+1)(r+2)}$ ، ب) از این نتیجه استفاده کرده، $E[(2X+1)^r]$ را حساب کنید.
حل. الف)

$$\begin{aligned} E(X^r) &= \int_0^1 x^r \cdot 2(1-x) dx = 2 \int_0^1 (x^r + x^{r+1}) dx \\ &= 2 \left(\frac{1}{r+1} - \frac{1}{r+2} \right) = \frac{2}{(r+1)(r+2)}. \end{aligned}$$

ب) چون $1, r=2$ و $r=1$ در فرمول $E[(2X+1)^r] = 2E(X^r) + 2E(X) + 1$ قرار دادن، $E(X^r) = \frac{2}{3 \times 4} = \frac{1}{6}$ و $E(X) = \frac{2}{3 \times 3} = \frac{1}{3}$ با
 $E[(2X+1)^r] = 2 \times \frac{1}{6} + 2 \times \frac{1}{3} + 1 = 3$.

مثال ۵-۷. نشان دهید که

$$E[(aX+b)^n] = \sum_{i=0}^n \binom{n}{i} a^{n-i} b^i E(X^{n-i}).$$

حل. چون بنابر قضیه ۹-۲ داریم

$$(ax+b)^n = \sum_{i=0}^n \binom{n}{i} (ax)^{n-i} b^i$$

نتیجه می‌شود که

$$\begin{aligned} E[(aX+b)^n] &= E\left[\sum_{i=0}^n \binom{n}{i} (aX)^{n-i} b^i\right] \\ &= \sum_{i=0}^n \binom{n}{i} a^{n-i} b^i E(X^{n-i}). \end{aligned}$$

مفهوم امید ریاضی را به آسانی می‌توان به وضعیت‌هایی که شامل دو یا چند متغیر تصادفی‌اند تعیین داد. برای نمونه، اگر Z ، متغیری تصادفی باشد که مقادیرش به مقادیر دو متغیر تصادفی X و Y به وسیله‌ی معادله‌ی $z = g(x, y)$ مربوط است، آنگاه می‌توان نشان داد که

قضیه ۴-۵. اگر X و Y ، متغیرهای تصادفی گستته بوده و $f(x, y)$ مقدار توزیع احتمال توأم آنها در (x, y) باشد، مقدار امید ریاضی متغیر تصادفی $g(X, Y)$ با

$$E[g(X, Y)] = \sum_x \sum_y g(x, y) f(x, y)$$

داده می‌شود. متناظراً اگر X و Y متغیرهای تصادفی پیوسته بوده و $f(x, y)$ مقدار چکالی توأم آنها در (x, y) باشد، مقدار امید ریاضی متغیر تصادفی $g(X, Y)$ با

$$E[g(X, Y)] = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} g(x, y) f(x, y) dx dy$$

داده می‌شود، مشروط بر اینکه مجموع و انتگرال بالا موجود باشد.

واضح است که قضیه‌ی فوق برای توابعی از هر تعداد متناهی از متغیرهای تصادفی برقرار است.

مثال ۵-۸. با مراجعه به مثال ۴-۴، مقدار مورد انتظار $g(X, Y) = X + Y$ را پیدا کنید.
حل.

$$\begin{aligned} E(X + Y) &= \sum_{x=0}^2 \sum_{y=0}^2 (x + y) f(x, y) \\ &= (0+0)\frac{1}{6} + (0+1)\frac{2}{9} + (0+2)\frac{1}{36} + (1+0)\frac{1}{3} + (1+1)\frac{1}{6} + (2+0)\frac{1}{12} = \frac{10}{9}. \end{aligned}$$

مثال ۵-۹. اگر تابع چگالی توأم X و Y به صورت زیر باشد

$$f(x,y) = \begin{cases} \frac{1}{\pi}(x+2y) & 0 < x < 1, 1 < y < 2 \\ 0 & \text{سایر جاهای} \end{cases}$$

مقدار امید ریاضی $g(X,Y) = \frac{X}{Y^2}$ را بیابید.

حل.

$$E\left(\frac{X}{Y^2}\right) = \int_0^1 \int_1^2 \frac{x(x+2y)}{\pi y^4} dx dy = \frac{1}{\pi} \int_1^2 \left(\frac{1}{y^3} + \frac{2}{y} \right) dy = \frac{15}{8\pi}.$$

قضیه‌ی زیر، قضیه‌ی دیگری است که در مطالب آتی کاربردهای مفیدی دارد. این قضیه، تعمیمی از قضیه‌ی ۳-۵ است و برهانش نظری برهان همان قضیه است.

قضیه‌ی ۵-۵. اگر c_1, c_2, \dots, c_n مقادیری ثابت باشند، آنگاه

$$E\left[\sum_{i=1}^n c_i g_i(X_1, X_2, \dots, X_k)\right] = \sum_{i=1}^n c_i E[g_i(X_1, X_2, \dots, X_k)].$$

تمرین

۱-۵. قضیه‌ی ۲-۵ را برای متغیرهای گسسته اثبات کنید.

۲-۵. قضیه‌ی ۳-۵ را برای متغیرهای پیوسته اثبات کنید.

۳-۵. مقدار مورد انتظار متغیر تصادفی گسسته‌ی X را که دارای توزیع احتمال

$$f(x) = \frac{|x-2|}{\pi}, \quad x = -1, 0, 1, 2$$

است بیابید.

۴-۵. مقدار مورد انتظار متغیر تصادفی Y را که چگالی احتمال آن

$$f(y) = \begin{cases} \frac{1}{\pi}(y+1) & -1 < y < 1 \\ 0 & \text{سایر جاهای} \end{cases}$$

است بیابید.

۵-۵. مقدار مورد انتظار متغیر تصادفی X را که چگالی احتمال آن

$$f(x) = \begin{cases} x & 0 < x < 1 \\ 2-x & 1 \leq x < 2 \\ 0 & \text{سایر جاهات} \end{cases}$$

است بیابید.

۶-۵ الف) اگر متغیر تصادفی X ، مقادیر $۰, ۱, ۲, ۳$ را به ترتیب با احتمال‌های

$$\frac{۱}{۱۲۵}, \frac{۱۲}{۱۲۵}, \frac{۴۸}{۱۲۵} \text{ و } \frac{۶۴}{۱۲۵} \text{ اختیار کند } E(X) \text{ و } (E(X))^۳ \text{ را پیدا کنید.}$$

ب) از نتایج قسمت (الف) استفاده کرده، $E[(X^3 + 2)^3]$ را بیابید.

۷-۵ الف) اگر تابع چگالی متغیر تصادفی پیوسته X

$$f(x) = \begin{cases} \frac{۱}{x(\ln ۳)} & ۱ < x < ۳ \\ ۰ & \text{سایر جاهات} \end{cases}$$

باشد $E(X)$ و $(E(X))^3$ را بیابید.

ب) از نتایج قسمت (الف) استفاده کرده، $E(X^3 + 2X^2 - 3X + 1)$ را پیدا کنید.

۸-۵ اگر چگالی احتمال X به صورت

$$f(x) = \begin{cases} \frac{x}{۴} & ۰ < x \leq ۱ \\ \frac{۱}{۴} & ۱ < x \leq ۲ \\ \frac{۳-x}{۴} & ۲ < x < ۳ \\ ۰ & \text{سایر جاهات} \end{cases}$$

باشد، مقدار مورد انتظار $g(x) = X^3 - 5X + 3$ را بیابید.

۹-۵ مدیر یک شیرینی‌پزی می‌داند که تعداد کیک‌های شکلاتی که می‌تواند در روزی

معین بفروشد متغیری تصادفی است که دارای توزیع احتمال $f(x) = \frac{۱}{۵}$ به ازای

$x = ۰, ۱, ۲, ۳, ۴, ۵$ ، است. ضمناً می‌داند که هر کیکی که می‌فروشد ۱۰ تومان سود دارد

و هر کیکی که فروش نمی‌رود ضرری (ناشی از فاسد شدن) برابر ۴ تومان خواهد

داشت. به فرض اینکه هر کیک را فقط همان روزی که پخته است می‌توان فروخت،

سود مورد انتظار شیرینی‌پز را برای روزی تعیین کنید که او

الف) یک کیک بپزد،

ب) دو کیک بپزد،

- ج) سه کیک بپزد،
د) چهار کیک بپزد،
ه) پنج کیک بپزد.

چند کیک باید بپزد تا سود مورد انتظار را ماکسیمم کند؟

۱۰-۵. سود مقاطعه کاری در یک کار ساختمانی را می‌توان متغیر تصادفی پیوسته‌ای در نظر گرفت که دارای چگالی احتمال

$$f(x) = \begin{cases} \frac{1}{18}(x+1) & -1 < x < 5 \\ 0 & \text{سایر مقادیر} \end{cases}$$

است، که در آن، واحد x برابر ۱۰۰۰۰ تومان است. سود مورد انتظار مقاطعه کار چقدر است؟

۳-۵ گشتاورها

امیدهای ریاضی که در اینجا و در تعریف ۴-۵ ارائه خواهند شد به گشتاورهای توزیع متغیر تصادفی یا صرفاً گشتاورهای متغیر تصادفی موسوم‌اند و در آمار از اهمیتی خاص برخوردارند.

تعریف ۲-۵. r امین گشتاور حول مبدأ متغیر تصادفی X ، که با μ'_r نشان داده می‌شود، امید X^r است؛ به صورت نمادی برای $\dots, 3, 2, 1, 0, r$ ، اگر X متغیر تصادفی گستته باشد،

$$\mu'_r = E(X^r) = \sum_x x^r f(x),$$

و اگر X متغیر تصادفی پیوسته باشد،

$$\mu'_r = E(X^r) = \int_{-\infty}^{+\infty} x^r f(x) dx.$$

شایان ذکر است که اصطلاح «گشتاور» مربوط به علم فیزیک است: اگر کمیت‌های $f(x)$ در حالت گستته جرم‌هایی نقطه‌ایی باشند که بر نقاط محور x ، واقع در فواصل x از مبدأ، به طور قائم عمل کنند، μ'_r مختص x مرکز ثقل است، یعنی اوئین گشتاور تقسیم بر $\sum f(x) = 1$ ، و μ'_r گشتاور اینرسی است، این مطلب همچنین توضیح

می‌دهید که چرا گشتاورهای μ' ، گشتاورهای حول مبدأ نام دارند در قیاس با فیزیک، طول بازوی اهرم در این حالت، فاصله‌ی تا مبدأست. این قیاس در حالت پیوسته نیز به کار می‌آید که در آن μ' و μ'' باید مختص x مرکز ثقل و گشتاور اینرسی یک میله با چگالی متغیر باشد.

وقتی $r=0$ ، بنابر نتیجه‌ی ۲ از قضیه‌ی ۲-۵، داریم $\mu' = E(X^\circ) = E(1) = \mu$ ، و این نتیجه همان‌طور که انتظار می‌رود با قضیه‌های ۱-۴ و ۵-۴ مطابقت دارد. وقتی $r=1$ ، داریم $\mu'' = E(X)$ ، که درست همان مقدار امید خود متغیر تصادفی X است، که از نظر اهمیت آن در آمار، به آن نماد خاص و اسم خاصی می‌دهیم.

تعریف ۳-۵. μ' ، میانگین توزیع X ، یا صرفاً میانگین X نامیده می‌شود و آن را با μ نشان می‌دهیم.

گشتاورهای خاصی که اینک تعریف می‌کنیم در آمار اهمیت دارند، زیرا در توصیف شکل توزیع متغیر تصادفی، یعنی شکل نمودار توزیع احتمال یا چگالی احتمال به کار می‌روند.

تعریف ۴-۵. گشتاور r حول میانگین متغیر تصادفی X ، که آن را با μ_r نشان می‌دهیم، مقدار امید $E[(X - \mu)^r]$ است؛ به صورت نمادی برای $r = 0, 1, 2, 3, \dots$ ، اگر X گسسته باشد،

$$\mu_r = E[(X - \mu)^r] = \sum_x (x - \mu)^r f(x),$$

و اگر X پیوسته باشد،

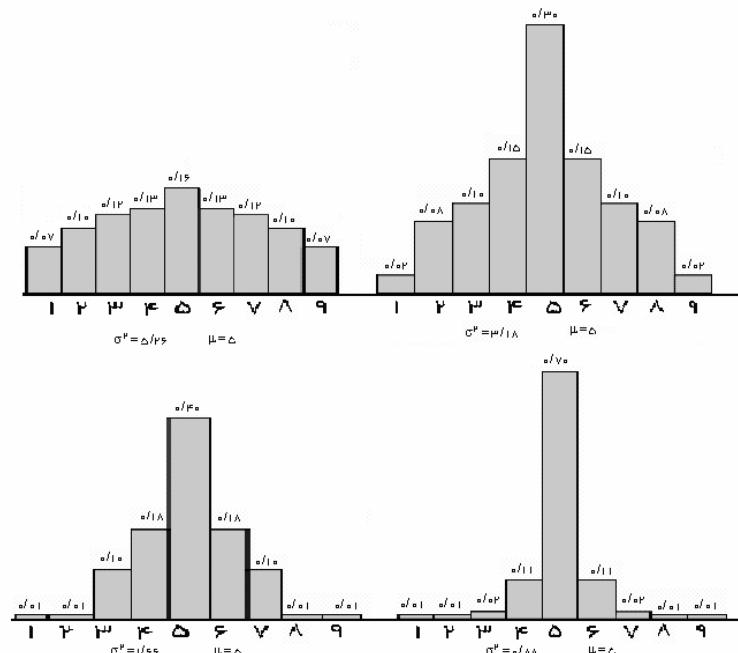
$$\mu_r = E[(X - \mu)^r] = \int_{-\infty}^{+\infty} (x - \mu)^r f(x) dx.$$

توجه کنید که برای هر متغیر تصادفی که μ برای آن وجود دارد، $\mu_0 = 1$ و $\mu_1 = 0$.

دوّمین گشتاور حول میانگین در آمار اهمیت خاصی دارد، زیرا نشان دهنده‌ی پراکندگی توزیع متغیر تصادفی است، لذا به آن نماد خاص و نام خاصی داده شده است.

تعريف ۵-۵. μ را واریانس توزیع X ، یا صرفاً واریانس X می‌نامند، و آن را با σ^2 ، $\text{var}(X)$ ، یا $V(X)$ نشان می‌دهند؛ σ ، ریشه‌ی دوم مثبت واریانس را انحراف معیار می‌نامند.

شکل ۱-۵ نشان می‌دهد که چگونه واریانس، منعکس کننده‌ی پراکندگی توزیع متغیر تصادفی است. در این شکل ما بافت‌نمایه‌ای توزیع‌های احتمال چهار متغیر تصادفی با میانگین یکسان $\mu = 5$ ، ولی با واریانس‌هایی مساوی $5/26$ ، $5/18$ ، $3/18$ ، و $5/88$ را نشان‌داده‌ایم. همان‌طور که دیده می‌شود، یک مقدار کوچک σ^2 ، این نکته را القا می‌کند که به دست آوردن مقداری نزدیک میانگین متحمل‌تر است، و یک مقدار بزرگ σ^2 ، این نکته را القا می‌کند که به دست آوردن مقداری که نزدیک میانگین نیست احتمال زیادی دارد. این مطلب در بخش ۵-۴، بیشتر مورد بحث قرار می‌گیرد. بحث مختصری از اینکه چگونه μ ، سومین گشتاور حول میانگین، تقارن یا چولگی (عدم تقارن) توزیع را توصیف می‌کند، که در فصل اوّل، آمار توصیف ارائه شده است.



شکل ۱-۵. توزیع‌های با پراکندگی مختلف

در بسیاری از موارد، گشتوارهای حول میانگین، ابتدا با محاسبه‌ی گشتوارهای حول مبدأ و سپس با بیان μ' بر حسب μ ، به دست می‌آیند. در اینجا صرفاً فرمول محاسباتی زیر را برای σ^2 به دست می‌آوریم.

$$\text{قضیه ۵-۶.} \quad \sigma^2 = \mu'_p - \mu^2 \quad \text{برهان.}$$

$$\begin{aligned}\sigma^2 &= E[(X - \mu)^2] \\ &= E(X^2 - 2\mu X + \mu^2) \\ &= E(X^2) - 2\mu E(X) + E(\mu^2) \\ &= E(X^2) - 2\mu \mu + \mu^2 \\ &= \mu'_p - \mu^2.\end{aligned}$$

مثال ۱۰-۵. با استفاده از قضیه‌ی ۵-۶، واریانس X را حساب کنید. متغیر تصادفی X معرف تعداد خالهایی است که در ریختن یک تاس همگن ظاهر می‌شود.

حل. ابتدا میانگین X را حساب می‌کنیم

$$\mu = E(X) = 1 \times \frac{1}{6} + 2 \times \frac{1}{6} + 3 \times \frac{1}{6} + 4 \times \frac{1}{6} + 5 \times \frac{1}{6} + 6 \times \frac{1}{6} = \frac{7}{2}.$$

حال

$$\mu'_p = E(X^2) = 1^2 \times \frac{1}{6} + 2^2 \times \frac{1}{6} + 3^2 \times \frac{1}{6} + 4^2 \times \frac{1}{6} + 5^2 \times \frac{1}{6} + 6^2 \times \frac{1}{6} = \frac{91}{6}$$

و نتیجه می‌شود که

$$\sigma^2 = \frac{91}{6} - \left(\frac{7}{2} \right)^2 = \frac{35}{12}.$$

مثال ۱۱-۵. با رجوع به مثال ۲-۵، انحراف معیار متغیر تصادفی X را تعیین کنید.

حل. در مثال ۲-۵، نشان دادیم که $\mu = E(X) = 0/4413$. حال

$$\mu'_p = E(X^2) = \frac{\pi}{\pi} \int_0^1 \frac{x^2}{1+x^2} dx = \frac{\pi}{\pi} \int_0^1 \left(1 - \frac{1}{1+x^2} \right) dx = \frac{\pi}{\pi} - 1 = 0/2732.$$

و نتیجه می‌شود که

$$\sigma^2 = 0/2732 - (0/4413)^2 = 0/0785.$$

$$\sigma = \sqrt{0/0785} = 0/2802.$$

قضیه‌ی زیر، قضیه‌ی دیگری است که در کارهای مربوط به انحراف معیارها یا واریانس‌ها اهمیت دارد.

قضیه‌ی ۷-۵. اگر واریانس X برابر σ^2 باشد، آنگاه $\text{var}(aX + b) = a^2\sigma^2$.

اثبات این قضیه به عهده‌ی خواننده واگذار می‌شود، اما به فرعهای زیر اشاره می‌کنیم: برای $a=1$ و $b=0$ ، متوجه می‌شویم که اضافه‌کردن مقداری ثابت به متغیر تصادفی، که نتیجه‌ی آن انتقال تمام مقادیر به چپ یا به راست است، به هیچ وجه اثری بر پراکندگی توزیع آن ندارد؛ برای $a=0$ ، متوجه می‌شویم که اگر مقادیر متغیر تصادفی را در ثابتی ضرب کنیم، واریانس در مربع آن ثابت ضرب می‌شود، که موجب تغییر متناظری در پراکندگی توزیع می‌شود.

تمرین

- ۱۱-۵. در تمرین ۳-۵ مقدار $\text{var}(X)$ را به دست آورید.
- ۱۲-۵. در تمرین ۴-۵ مقدار $\text{var}(Y)$ را به دست آورید.
- ۱۳-۵. در تمرین ۵-۵ مقدار $\text{var}(eX)$ را به دست آورید.
- ۱۴-۵. در تمرین ۶-۵ مقدار $\text{var}(2X)$ را به دست آورید.
- ۱۵-۵. در تمرین ۷-۵ مقدار $\text{var}(1-X)$ را به دست آورید.
- ۱۶-۵. برای متغیر تصادفی X که دارای توزیع احتمال $f(x) = \frac{1}{2}$ ، به ازای $x=-2$ و $x=2$ است، μ و μ' و σ^2 را بیابید.

- ۱۷-۵. برای متغیر تصادفی X که دارای چگالی احتمال

$$f(x) = \begin{cases} \frac{x}{2} & 0 < x < 2 \\ 0 & \text{سایر جاهای} \end{cases}$$

است، μ و μ' و σ^2 را بیابید.

- ۱۸-۵. اگر متغیر تصادفی X دارای میانگین μ و واریانس σ^2 باشد، نشان دهید که برای متغیر تصادفی Z ، که مقادیرش به مقادیر X با معادله $Z = \frac{X-\mu}{\sigma}$ را اعمال

می‌کنیم، دارای میانگین صفر و واریانس است. گوییم توزیع X را استاندارد کردیم.

۱۹-۵. اگر چگالی احتمال متغیر تصادفی X به صورت

$$f(x) = \begin{cases} 2x^{-3} & x > 1 \\ 0 & \text{سایر جاهات} \end{cases}$$

باشد، بررسی کنید که آیا میانگین و واریانس آن موجودند؟

۲۰-۵. تابع مولد گشتاورهای متغیر تصادفی گستته‌ی X را که دارای توزیع احتمال

$$f(x) = \left(\frac{1}{\mu}\right)^x, \quad x = 1, 2, 3, \dots$$

است بیابید، و آن را برای تعیین مقادیر μ و σ^2 به کار ببرید.

۲۱-۵. تابع مولد گشتاورهای متغیر تصادفی پیوسته X را که چگالی احتمال آن به صورت

$$f(x) = \begin{cases} 1 & 0 < x < 1 \\ 0 & \text{سایر جاهات} \end{cases}$$

داده شده است پیدا کنید و آن را برای تعیین μ , σ^2 , و σ^4 به کار ببرید.

۴-۵ قضیه‌ی چبیشف

برای نشان دادن اینکه چگونه σ یا σ^2 بر پراکندگی توزیع متغیر تصادفی دلالت دارد، قضیه‌ی زیر را که به یاد چبیشف ریاضیدان روسی قرن نوزدهم، قضیه‌ی چبیشف می‌نامند ثابت می‌کنیم. ما آن را در اینجا فقط برای حالت پیوسته ثابت می‌کنیم، و اثبات آن را برای حالت گستته به عنوان تمرین واگذار می‌کنیم.

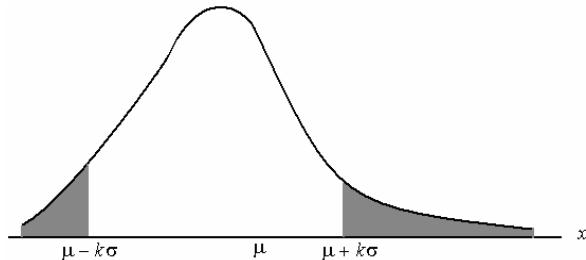
قضیه‌ی ۴-۵ (قضیه‌ی چبیشف). اگر μ و σ به ترتیب میانگین و انحراف معیار متغیر تصادفی X باشند، آنگاه برای هر ثابت مثبت k ، احتمال اینکه X ، مقداری با فاصله‌ی کمتر از $k\sigma$ از میانگین اختیار کند حداقل $\frac{1}{k^2} - 1$ است؛ به صورت نمادی

$$P(|X - \mu| < k\sigma) \geq 1 - \frac{1}{k^2}.$$

برهان. با توجه به تعریف‌های ۴-۵ و ۵-۵، می‌نویسیم

$$\sigma^2 = E[(X - \mu)^2] = \int_{-\infty}^{+\infty} (x - \mu)^2 f(x) dx$$

و با تفکیک انتگرال به سه بخش، همان‌طور که در شکل ۲-۵ نشان داده‌ایم، به‌دست می‌آوریم



شکل ۲-۵. نموداری برای اثبات قضیه‌ی چبیشف

$$\sigma^2 = \int_{-\infty}^{\mu-k\sigma} (x-\mu)^2 f(x) dx + \int_{\mu-k\sigma}^{\mu+k\sigma} (x-\mu)^2 f(x) dx + \int_{\mu+k\sigma}^{+\infty} (x-\mu)^2 f(x) dx.$$

چون $(x-\mu)^2 f(x)$ ، عبارت زیر انتگرال، نامنفی است، می‌توانیم نابرابری

$$\sigma^2 \geq \int_{-\infty}^{\mu-k\sigma} (x-\mu)^2 f(x) dx + \int_{\mu+k\sigma}^{+\infty} (x-\mu)^2 f(x) dx,$$

را با حذف انتگرال دوم تشکیل دهیم. حال چون برای $x \leq \mu - k\sigma$ یا $x \geq \mu + k\sigma$ ، $(x-\mu)^2 \geq k^2\sigma^2$

$$\sigma^2 \geq \int_{-\infty}^{\mu-k\sigma} k^2\sigma^2 f(x) dx + \int_{\mu+k\sigma}^{+\infty} k^2\sigma^2 f(x) dx,$$

و بنابراین به شرط $\sigma^2 \neq 0$

$$\frac{1}{k^2} \geq \int_{-\infty}^{\mu-k\sigma} f(x) dx + \int_{\mu+k\sigma}^{+\infty} f(x) dx.$$

چون مجموع دو انتگرال در این نابرابری، احتمال این را نمایش می‌دهد که X مقداری نایبیشتر از $\mu - k\sigma$ یا ناکمتر از $\mu + k\sigma$ را اختیار کند، لذا نشان داده‌ایم که

$$P(|X - \mu| \geq k\sigma) \leq \frac{1}{k^2}$$

و نتیجه می‌شود که

$$P(|X - \mu| \geq k\sigma) \leq 1 - \frac{1}{k^2}.$$

برای مثال، احتمال اینکه X مقداری در فاصله‌ی 2σ از میانگین اختیار کند حداقل $\frac{1}{2} = \frac{3}{4}$ است، احتمال اینکه X مقداری در فاصله‌ی 3σ از میانگین اختیار کند حداقل $\frac{1}{3} = \frac{8}{9}$ است و احتمال اینکه X مقداری در فاصله‌ی 5σ از میانگین اختیار کند حداقل $\frac{1}{5} = \frac{24}{25}$ است. این بدان معناست که 5 پراکندگی توزیع متغیر تصادفی را کنترل می‌کند. به وضوح، احتمالی که به وسیله‌ی قضیه‌ی چبیشف داده می‌شود فقط یک کران پایینی است؛ درباره‌ی اینکه وقتی متغیر تصادفی مقداری با فاصله‌ی کمتر از $k\sigma$ از میانگین اختیار می‌کند احتمالش واقعاً بزرگتر از $\frac{1}{k^2} = \frac{1}{1}$ است یا نه، و اگر بزرگتر باشد به چه اندازه، نمی‌توانیم چیزی بگوییم، اما قضیه‌ی چبیشف ما را مطمئن می‌سازد که این احتمال نمی‌تواند کمتر از $\frac{1}{k^2} = \frac{1}{1}$ باشد. تنها وقتی می‌توانیم احتمال دقیق را محاسبه کنیم که توزیع متغیر تصادفی معلوم باشد.

مثال ۱۲-۵. اگر چگالی احتمال متغیر تصادفی X به صورت

$$f(x) = \begin{cases} 630x^4(1-x)^4 & 0 < x < 1 \\ 0 & \text{سایر جاهات} \end{cases}$$

باشد، احتمال این را که X مقداری در فاصله‌ی 2σ از میانگین اختیار کند پیدا کنید، و آن را با کران پایینی که از روی قضیه‌ی چبیشف حاصل می‌شود مقایسه نمایید. انتگرال‌گیری مستقیم نشان می‌دهد که $\mu = \frac{1}{2}$ و $\sigma = \sqrt{1/14}$ ، به قسمی که حل میانگین اختیار کند برابر احتمال آن است که X مقداری بین $0/20$ و $0/80$ اختیار کند، یعنی

$$P(0/20 < X < 0/80) = \int_{0/20}^{0/80} 630x^4(1-x)^4 dx = 0/96.$$

ملاحظه کنید که حکم «احتمال برابر $96/100$ است» خیلی قوی‌تر است از حکم «احتمال حداقل برابر $75/100$ است» که از قضیه‌ی چبیشف به دست می‌آید.

تمرین

۲۲-۵. قضیه‌ی چبیشف را برای متغیرهای گستته اثبات کنید.

۲۳-۵. نشان دهید که اگر X متغیری تصادفی با میانگین μ باشد که برای آن به ازای $x < 0$ ، $f(x) = 0$ ، آنگاه برای هر ثابت مثبت a

$$P(X \geq a) \leq \frac{\mu}{a}.$$

این نابرابری را نابرابری مارکوف می‌نامند، و آن را در اینجا بیشتر به این دلیل آوردیم که استدلال نسبتاً ساده‌تر دیگری برای قضیه‌ی چبیشف منجر می‌شود.

۲۴-۵. نابرابری تمرین ۲۳-۵ را برای اثبات قضیه‌ی چبیشف به کار بردیم.
(راهنمایی: به جای X مقدار $(\mu - X)$ را قرار دهید).

۲۵-۵. در قضیه‌ی چبیشف، وقتی احتمال اینکه متغیر تصادفی مقداری بین $\mu - k\sigma$ و $\mu + k\sigma$ اختیار کند

الف) حداقل $95/0$ ، ب) حداقل $99/0$

است، کمترین مقدار k چه خواهد بود؟

۲۶-۵. اگر در قضیه‌ی چبیشف، قرار دهیم $c = k\sigma$ ، این قضیه‌درباره‌ی احتمال اینکه متغیر تصادفی مقداری بین $\mu - c$ و $\mu + c$ اختیار نماید چه حکمی می‌کند؟

۲۷-۵. توضیح دهید چرا نمی‌توان متغیر تصادفی یافت که برای آن $M_X(t) = \frac{t}{1-t}$
۲۸-۵. نشان دهید که اگر متغیری تصادفی دارای چگالی احتمال

$$f(x) = \frac{1}{\mu} e^{-|x|/\mu}, \quad -\infty < x < +\infty$$

باشد، تابع مولد گشتاورهایش به صورت $M_X(t) = \frac{1}{1-t^2}$ است.

۲۹-۵. تابع مولد گشتاورهای متغیر تصادفی X به صورت $M_X(t) = e^{3t+8t^2}$ داده شده است، تابع مولد گشتاورهای متغیر تصادفی Z به صورت $(X - 3)/\mu = Z$ را پیدا کنید، و برای تعیین میانگین و واریانس Z ، به کار بردیم.

۳۰-۵. مطالعه‌ای درباره‌ی ارزش غذایی نوع یک نان نشان می‌دهد که میزان تیامین (ویتامین B_1)، در یک قرص این نوع نان را می‌توان به صورت متغیری تصادفی با $\mu = ۰/۲۶۰$ میلی‌گرم و $\sigma = ۰/۰۰۵$ میلی‌گرم در نظر گرفت. بنابر قضیه‌ی چبیشف

مقدار تیامین موجود در

الف) حداقل $\frac{۳۵}{۶۴}$ تمام قرص‌های این نوع نان،

ب) حداقل $\frac{۱۴۳}{۱۴۴}$ تمام قرص‌های این نوع نان،

بین کدام دو مقدار باید باشد؟

۵-۵ توابع مولد گشتاورها

گرچه گشتاورهای بیشتر توزیع‌ها را می‌توان مستقیماً با محاسبه‌ی انتگرال‌ها یا مجموعه‌های لازم معین کرد، ولی شیوه‌ی دیگری نیز وجود دارد که اغلب تسهیلات قابل ملاحظه‌ای را در اختیار می‌گذارد. در این شیوه، از توابع مولد گشتاورها استفاده می‌شود.

تعریف ۵-۶. تابع مولد گشتاورهای متغیر تصادفی X ، در صورت وجود، اگر X گسسته باشد، با تابع

$$M_X(t) = E(e^{tX}) = \sum_x e^{tx} f(x),$$

و اگر X پیوسته باشد، با تابع

$$M_X(t) = E(e^{tX}) = \int_{-\infty}^{+\infty} e^{tx} f(x) dx.$$

داده می‌شود، که در آن t عددی حقیقی است.

t شناسه‌ی مستقل است و معمولاً مقادیری از t در همسایگی 0 مورد نظرند. برای توضیح اینکه چرا به این تابع، عنوان تابع «مولد گشتاورها» را اطلاق می‌کنیم، به

جای e^{tx} بسط آن به سری ماکلورن را قرار می‌دهیم، یعنی

$$e^{tx} = 1 + tx + \frac{t^2 x^2}{2!} + \frac{t^3 x^3}{3!} + \dots + \frac{t^r x^r}{r!} + \dots$$

بنابراین، برای حالت گسسته به دست می‌آوریم

$$\begin{aligned}
M_X(t) &= \sum_x \left[1 + tx + \frac{t^r x^r}{r!} + \frac{t^r x^r}{r!} + \dots + \frac{t^r x^r}{r!} + \dots \right] f(x) \\
&= \sum_x f(x) + t \sum_x x f(x) + \frac{t^r}{r!} \sum_x x^r f(x) + \dots + \frac{t^r}{r!} \sum_x x^r f(x) + \dots \\
&= 1 + \mu t + \mu'_r \frac{t^r}{r!} + \dots + \mu'_r \frac{t^r}{r!} + \dots
\end{aligned}$$

و می‌توان دید که در سری ماکلورن تابع مولد گشتاورهای X ، ضریب $\frac{t^r}{r!}$ ، همان μ'_r ، r امین گشتاور حول مبدأ متغیر تصادفی X است. در حالت پیوسته، استدلال همین‌گونه است.

مثال ۱۳-۵. تابع مولد گشتاورهای متغیر تصادفی X را که چگالی احتمال آن به صورت

$$f(x) = \begin{cases} e^{-x} & x > 0 \\ 0 & \text{سایر جاهای} \end{cases}$$

است بیایید، و آن را برای تعیین به کار بردیم.

حل. بنابر تعریف

$$M_X(t) = E(e^{tX}) = \int_0^{+\infty} e^{tx} e^{-x} dx = \int_0^{+\infty} e^{-x(t-1)} dx = \frac{1}{t-1}, \quad |t| < 1$$

همان‌طور که می‌دانید، وقتی $|t| < 1$ سری ماکلورن برای این تابع مولد گشتاورها به صورت

$$\begin{aligned}
M_X(t) &= 1 + t + t^r + t^r + \dots + t^r + \dots \\
&= 1 + 1! \frac{t}{1!} + 2! \frac{t^2}{2!} + 3! \frac{t^3}{3!} + \dots + r! \frac{t^r}{r!} + \dots
\end{aligned}$$

است و بنابر این برای $r = 0, 1, 2, \dots$.

مشکل اصلی در به کار بردن سری ماکلورن تابع مولد گشتاورها برای تعیین گشتاورهای یک متغیر تصادفی، معمولاً پیدا کردن تابع مولد گشتاورها نیست، بلکه مشکل، بسط آن به صورت سری ماکلورن است. اگر فقط چند گشتاور اول متغیر تصادفی، مثلاً μ'_0 و μ'_1 مورد توجه باشند، عمل تعیین آنها را معمولاً می‌توان با استفاده

از قضیه‌ی زیر ساده‌تر کرد.

$$\text{قضیه‌ی ۹-۵} \cdot \left. \frac{d^r M_X(t)}{dt^r} \right|_{t=0} = \mu'_r$$

این قضیه از این واقعیت نتیجه می‌شود که اگر تابعی به صورت سری توانی بر حسب t بسط داده شود، ضریب $\frac{t^r}{r!}$ ، مشتق مرتبه‌ی r ام تابع نسبت به t به ازای $t=0$ است.

مثال ۱۴-۵. توزیع احتمال X ، برای $x=0, 1, 2, 3, \dots$ به صورت $f(x) = \binom{3}{x} \frac{1}{\lambda}$ است، تابع مولد گشتاورهای این متغیر تصادفی را پیدا کنید و آن را برای تعیین μ' و μ'' به کار ببرید.

حل. با جایگذاری توزیع احتمال در تعریف ۶-۵، به دست می‌آوریم

$$\begin{aligned} M_X(t) &= E(e^{tX}) = \sum_{x=0}^{\infty} e^{tx} \binom{3}{x} \\ &= \frac{1}{\lambda} (1 + 3e^t + 3e^{2t} + e^{3t}) = \frac{1}{\lambda} (1 + e^t)^3. \end{aligned}$$

پس، بنابر قضیه‌ی ۹-۵

$$\mu' = M'_X(0) = \left. \frac{1}{\lambda} (1 + e^t)^3 e^t \right|_{t=0} = 3,$$

و

$$\mu'' = M''_X(0) = \left. \frac{1}{\lambda} (1 + e^t)^3 e^{2t} + \frac{1}{\lambda} (1 + e^t)^3 e^t \right|_{t=0} = 6.$$

اغلب مسائل مربوط به کاربرد توابع مولد گشتاورها را می‌توان با استفاده از قضیه‌ی زیر ساده کرد.

قضیه‌ی ۱۰-۵. اگر a و b دو مقدار ثابت باشند، آنگاه

$$M_{X+a}(t) = E[e^{(X+a)t}] = e^{at} M_X(t).$$

$$M_{bX}(t) = E(e^{bXt}) = M_X(bt) \quad .\quad ۲$$

$$M_{\frac{X+a}{b}}(t) = E[e^{\frac{(X+a)}{b}t}] = e^{\frac{a}{b}t} M_X\left(\frac{t}{b}\right) \quad .\quad ۳$$

برهان. به عنوان تمرین به عهده‌ی خواننده واگذار شده است.

همان‌طور که بعداً خواهیم دید، قسمت اول این قضیه، وقتی $\mu = -a$ ، اهمیت خاصی دارد، و قسمت سوم وقتی $a = -\mu$ و $b = \sigma$ نیز دارای اهمیت خاصی است، در این حالت

$$M_{\frac{X-\mu}{\sigma}}(t) = e^{-\frac{\mu t}{\sigma}} M_X\left(\frac{t}{\sigma}\right).$$

تعريف ۵-۷. امین و سامین گشتاور حاصل‌ضربی حول مبدأ متغیرهای تصادفی X و Y ، که با $\mu'_{r,s}$ نشان داده می‌شود، مقدار مورد انتظار $X^r Y^s$ است؛ به صورت نمادی، برای $r, s = 0, 1, 2, 3, \dots$ و X و Y گسسته

$$\mu'_{r,s} = E(X^r Y^s) = \sum_x \sum_y x^r y^s f(x, y),$$

و با X و Y پیوسته،

$$\mu'_{r,s} = E(X^r Y^s) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} x^r y^s f(x, y) dx dy.$$

در حالت گسسته، مجموع یابی دوگانه روی تمام برد توأم دو متغیر تصادفی است. توجه کنید که $\mu'_{0,0} = E(X)$ ، که در اینجا با μ_X نشان داده می‌شود، و $\mu'_{0,1} = E(Y)$ ، که در اینجا با μ_y نشان داده می‌شود.

اکنون مشابه تعريف ۴-۴، تعريف زیر را از گشتاورهای حاصل‌ضربی حول میانگین‌های مربوطه ارائه می‌دهیم

تعريف ۵-۸. امین و سامین گشتاور حاصل‌ضربی حول مبدأ متغیرهای تصادفی X و Y ، که با $\mu_{r,s}$ نشان داده می‌شود، مقدار مورد انتظار $(X - \mu_X)^r (Y - \mu_Y)^s$ است؛ به صورت نمادی برای $r, s = 0, 1, 2, 3, \dots$ و

$$\begin{aligned}\mu_{r,s} &= E[(X - \mu_X)^r (Y - \mu_Y)^s] \\ &= \sum_x \sum_y (x - \mu_X)^r (y - \mu_Y)^s f(x,y),\end{aligned}$$

وقتی X و Y گسسته‌اند و

$$\begin{aligned}\mu_{r,s} &= E[(X - \mu_X)^r (Y - \mu_Y)^s] \\ &= \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} (x - \mu_X)^r (y - \mu_Y)^s f(x,y) dx dy,\end{aligned}$$

وقتی X و Y پیوسته‌اند.

در آمار، $\mu_{r,s}$ از اهمیت خاصی برخوردار است، زیرا بر رابطه بین مقادیر X و Y ، در صورت وجود، دلالت دارد؛ لذا به آن نماد و نام خاصی داده شده است.

تعريف ۹-۵. $\mu_{r,s}$ را کوواریانس X و Y می‌نامند و آن را با $\text{cov}(X,Y)$ ، یا σ_{XY} نشان می‌دهند.

مالحظه کنید که اگر احتمال زیادی وجود داشته باشد که مقادیر بزرگ X با مقادیر بزرگ Y و مقادیر کوچک X با مقادیر کوچک Y همراه باشند، کوواریانس مثبت خواهد بود؛ و اگر احتمال زیادی وجود داشته باشد که مقادیر بزرگ X با مقادیر کوچک Y و برعکس همراه باشند، کوواریانس منفی خواهد بود. با این مفهوم است که کوواریانس، رابطه یا پیوند بین مقادیر X و Y را اندازه می‌گیرد.

$$\text{قضیه ۱۱-۵. } \sigma_{XY} = \mu'_{1,1} - \mu_X \mu_Y$$

برهان. با استفاده از قضیه‌های مختلف درباره مقدار مورد انتظار، می‌توانیم بنویسیم

$$\begin{aligned}\sigma_{XY} &= E[(X - \mu_X)(Y - \mu_Y)] \\ &= E[XY - \mu_X Y - X \mu_Y + \mu_X \mu_Y] \\ &= E[XY] - \mu_X E[Y] - \mu_Y E[X] + \mu_X \mu_Y \\ &= E[XY] - \mu_X \mu_Y - \mu_Y \mu_X + \mu_X \mu_Y \\ &= \mu'_{1,1} - \mu_X \mu_Y.\end{aligned}$$

مثال ۱۵-۵. درمثال ۲۲-۴، احتمال‌های توأم و حاشیه‌ای X و Y ، که به ترتیب معروف تعداد مهره‌های قرمز و تعداد مهره‌های سبز بین دو مهره‌ای بودند که از ظرف محتوی ۳ مهره‌ی قرمز، ۲ مهره‌ی سبز، و ۴ مهره‌ی آبی به تصادف در آوریم، به صورت جدول زیر بودند.

| | | | X |
|-----|---|---------------|--|
| | | | ۰ ۱ ۲ |
| | | ۰ | $\frac{1}{6}$ $\frac{1}{3}$ $\frac{1}{12}$ $\frac{1}{6}$ $\frac{1}{3}$ $\frac{1}{12}$ |
| y | ۰ | $\frac{1}{6}$ | $\frac{1}{12}$ |
| | ۱ | $\frac{1}{9}$ | $\frac{1}{18}$ |
| | | ۲ | $\frac{1}{36}$ |

کوواریانس X و Y را بیابید.

حل. با استفاده از احتمال‌های توأمی که داده شده‌اند، به دست می‌آوریم

$$\begin{aligned} \mu'_{1,1} &= E(XY) \\ &= 0 \times 0 \times \frac{1}{6} + 0 \times 1 \times \frac{1}{9} + 0 \times 2 \times \frac{1}{36} + 1 \times 0 \times \frac{1}{6} + 1 \times 1 \times \frac{1}{9} + 2 \times 0 \times \frac{1}{12} \\ &= \frac{1}{6}, \end{aligned}$$

و با استفاده از احتمال‌های حاشیه‌ای، به دست می‌آوریم

$$\begin{aligned} \mu_X &= E(X) = 0 \times \frac{5}{12} + 1 \times \frac{1}{12} + 2 \times \frac{1}{12} = \frac{2}{3} \\ \mu_Y &= E(Y) = 0 \times \frac{7}{12} + 1 \times \frac{7}{18} + 2 \times \frac{1}{36} = \frac{4}{9}. \end{aligned}$$

نتیجه می‌شود که

$$\sigma_{XY} = \mu'_{1,1} - \mu_X \mu_Y = \frac{1}{6} - \frac{2}{3} \times \frac{4}{9} = -\frac{7}{54}.$$

نتیجه‌ی منفی القا می‌کند که هر چه مهره‌های قرمز که به دست می‌آوریم بیشتر باشند مهره‌های سبز کمترند و بر عکس، و البته این نکته‌ی معقولی است.

مثال ۱۶-۵. کوواریانس دو متغیر تصادفی را که چگالی توأم آنها به صورت

$$f(x,y) = \begin{cases} ۲ & x > ۰, y > ۰, x+y < ۱ \\ ۰ & \text{سایر جاهای} \end{cases}$$

است پیدا کنید.

حل. با محاسبه‌ی انتگرال‌های لازم، به دست می‌آوریم

$$\mu_X = \int_0^1 \int_0^{1-x} 2x \, dy \, dx = \frac{1}{3}$$

$$\mu_Y = \int_0^1 \int_0^{1-y} 2y \, dy \, dx = \frac{1}{3}$$

و

$$\mu'_{1,1} = \int_0^1 \int_0^{1-x} 2xy \, dy \, dx = \frac{1}{12}.$$

نتیجه می‌شود که

$$\sigma_{XY} = \mu'_{1,1} - \mu_X \mu_Y = \frac{1}{12} - \frac{1}{3} \times \frac{1}{3} = -\frac{1}{36}.$$

تاکنون آنچه گفتیم مربوط به رابطه‌ی بین X و Y بود، ملاحظه کنید که اگر X و Y مستقل باشند کوواریانس آنها صفر است؛ به صورت نمادی،

قضیه‌ی ۱۲-۵. اگر X و Y مستقل باشند، آنگاه $E(XY) = E(X)E(Y)$ ، و $\sigma_{XY} = 0$.

برهان. برای حالت گسسته، بنابر تعریف داریم $E(XY) = \sum_x \sum_y xy f(x,y)$ ، چون X و Y مستقل‌اند، می‌توانیم بنویسیم $f(x,y) = g(x)h(y)$ ، که در آن $g(x)$ و $h(y)$ به ترتیب مقادیر توزیع‌های حاشیه‌ای X و Y هستند، و به دست می‌آوریم

$$\begin{aligned} E(XY) &= \sum_x \sum_y xy g(x)h(y) \\ &= [\sum_x x g(x)] [\sum_y y h(y)] \\ &= E(X)E(Y). \end{aligned}$$

بنابراین

$$\sigma_{XY} = \mu'_{1,1} - \mu_X \mu_Y = E(XY) - E(X)E(Y) = 0.$$

درخور ذکر است که استقلال دو متغیر تصادفی، صفر بودن کوواریانس را ایجاب می‌کند، اما صفر بودن کوواریانس الزاماً استقلال آنها را نتیجه نمی‌دهد.

مثال ۱۷-۵. اگر توزیع احتمال توأم X و Y به صورت

| | | x | | |
|---|----|---------------|---------------|---------------|
| | | -1 | 0 | 1 |
| y | -1 | $\frac{1}{6}$ | $\frac{1}{3}$ | $\frac{1}{6}$ |
| | 0 | 0 | 0 | 0 |
| | 1 | $\frac{1}{6}$ | 0 | $\frac{1}{6}$ |
| | | $\frac{1}{3}$ | $\frac{1}{3}$ | $\frac{1}{3}$ |

باشد، نشان دهید با اینکه دو متغیر تصادفی مستقل نیستند، کوواریانس آنها صفر است.
حل. با استفاده از احتمال‌هایی که در حاشیه‌ها نشان داده‌ایم، به دست می‌آوریم

$$\mu_X = (-1) \times \frac{1}{3} + 0 \times \frac{1}{3} + 1 \times \frac{1}{3} = 0,$$

$$\mu_Y = (-1) \times \frac{1}{3} + 0 \times 0 + 1 \times \frac{1}{3} = -\frac{1}{3},$$

و

$$\mu'_{1,1} = (-1) \times (-1) \times \frac{1}{6} + 0 \times (-1) \times \frac{1}{6} + 1 \times (-1) \times \frac{1}{6} + 1 \times 1 \times \frac{1}{6} = 0.$$

پس، $\sigma_{XY} = 0 - 0 = 0$ و کوواریانس صفر است، اما دو متغیر تصادفی مستقل نیستند. مثلاً برای $x = -1$ و $y = -1$ ، $f(x, y) \neq g(x)h(y)$.

گشتاورهای حاصل ضریبی را می‌توان برای حالتی نیز که بیش از دو متغیر تصادفی وجود دارد تعریف کرد. در اینجا فقط قضیه‌ی مهم زیر را بیان می‌کنیم.

قضیه‌ی ۱۳-۵. اگر X_1, X_2, \dots, X_n مستقل باشند، آنگاه

$$E(X_1 X_2 \cdots X_n) = E(X_1) E(X_2) \cdots E(X_n).$$

این، تعمیمی از قسمت اوّل قضیه‌ی ۱۲-۵ است؛ درواقع، برهان این قضیه، که مبنی بر تعریف ۱۴-۴ است، اساساً نظری برهان قسمت اوّل قضیه‌ی ۱۲-۵ است.

۶-۵ گشتاورهای ترکیب‌های خطی متغیرهای تصادفی

در این بخش عبارت‌هایی برای میانگین و واریانس ترکیبی خطی از n متغیر تصادفی و کوواریانس دو ترکیب خطی از n متغیر تصادفی به دست می‌آوریم. کاربردهای این نتایج، در مبحث نظریه‌ی نمونه‌گیری و مسائل استنباط آماری مورد استفاده قرار خواهند گرفت.

قضیه‌ی ۵-۱۴. اگر X_1, X_2, \dots, X_n متغیرهای تصادفی و $Y = \sum_{i=1}^n a_i X_i$ باشند که در آن a_1, a_2, \dots, a_n مقادیری ثابت هستند، آنگاه $E(Y) = \sum_{i=1}^n a_i E(X_i)$ و $\text{var}(Y) = \sum_{i=1}^n a_i^2 \text{var}(X_i) + 2 \sum_{i < j} a_i a_j \text{cov}(X_i, X_j)$ ، که در آن مجموع یابی دوگانه روی تمام مقادیر i و j از 1 تا n با شرط $j < i$ انجام می‌شود.

برهان. از قضیه‌ی ۵-۵ با $g_i(X_1, X_2, \dots, X_n) = X_i$ به ازای $i = 1, 2, \dots, n$ ، بی‌درنگ نتیجه می‌شود که

$$E(Y) = E\left(\sum_{i=1}^n a_i X_i\right) = \sum_{i=1}^n a_i E(X_i),$$

و این، اوّلین قسمت قضیه را ثابت می‌کند. در به دست آوردن عبارتی برای واریانس، μ_i را به جای $E(X_i)$ قرار می‌دهیم، و در نتیجه به دست می‌آوریم

$$\begin{aligned} \text{var}(Y) &= E([Y - E(Y)]^2) \\ &= E\left\{\left[\sum_{i=1}^n a_i X_i - \sum_{i=1}^n a_i E(X_i)\right]^2\right\} \\ &= E\left\{\left[\sum_{i=1}^n a_i (X_i - \mu_i)\right]^2\right\}. \end{aligned}$$

حال با بسط این نتیجه بر حسب قضیه‌ی چندجمله‌ای که بنابر آن مثلاً $(a+b+c+d)^2 = a^2 + b^2 + c^2 + d^2 + 2ab + 2ac + 2ad + 2bc + 2bd + 2cd$ مساوی

است، و مراجعه‌ی دوباره به قضیه‌ی ۴-۵، به دست می‌آوریم

$$\begin{aligned} \text{var}(Y) &= \sum_{i=1}^n a_i^2 E[(X_i - \mu_i)^2] + 2 \sum_{i < j} a_i a_j E[(X_i - \mu_i)(X_j - \mu_j)] \\ &= \sum_{i=1}^n a_i^2 \text{var}(X_i) + 2 \sum_{i < j} a_i a_j \text{cov}(X_i, X_j). \end{aligned}$$

توجه کنید که ما به طور ضمنی از واقعیت $\text{cov}(X_i, X_j) = \text{cov}(X_j, X_i)$ استفاده کردایم.

چون وقتی X_i و X_j مستقل‌اند، $\text{cov}(X_i, X_j) = 0$ ، بی‌درنگ نتیجه می‌شود که

نتیجه‌ی ۳. اگر متغیرهای تصادفی X_1, X_2, \dots, X_n مستقل باشند و آنگاه

$$\text{var}(Y) = \sum_{i=1}^n a_i^2 \text{var}(X_i).$$

مثال ۱۸-۵. اگر متغیرهای تصادفی X, Y و Z دارای میانگین‌های $\mu_X = 2$ ، $\mu_Y = -3$ ، $\mu_Z = 4$ ، واریانس‌های $\sigma_X^2 = 5$ ، $\sigma_Y^2 = 2$ ، $\sigma_Z^2 = 1$ ، و کوواریانس‌های $\text{cov}(Y, Z) = 1$ ، $\text{cov}(X, Z) = -1$ ، $\text{cov}(X, Y) = -2$ باشند، میانگین و واریانس

$$W = 3X - Y + 2Z$$

را پیدا کنید.

حل. بنابر قضیه‌ی ۱۴-۵، به دست می‌آوریم

$$\begin{aligned} E(W) &= E(3X - Y + 2Z) \\ &= 3E(X) - E(Y) + 2E(Z) \\ &= 3 \times 2 - (-3) + 2 \times 4 = 17 \end{aligned}$$

و

$$\begin{aligned} \text{var}(W) &= \text{var}(3X - Y + 2Z) \\ &= 9\text{var}(X) + \text{var}(Y) + 4\text{var}(Z) \\ &\quad - 6\text{cov}(X, Y) + 12\text{cov}(X, Z) - 4\text{cov}(Y, Z) \\ &= 9 \times 5 + 2 + 4 \times 4 - 6(-2) + 12(-1) - 4 \times 1 = 18. \end{aligned}$$

قضیه‌ی زیر، قضیه‌ی مهم‌ دیگری درباره‌ی ترکیب‌های خطی متغیرهای تصادفی است. این قضیه مربوط به کوواریانس دو ترکیب خطی از n متغیر تصادفی است.

قضیه‌ی ۱۵-۵. اگر X_1, X_2, \dots, X_n متغیرهای تصادفی باشند و

$$Y_2 = \sum_{i=1}^n b_i X_i \quad \text{و} \quad Y_1 = \sum_{i=1}^n a_i X_i$$

که در آنها a_1, a_2, \dots, a_n و b_1, b_2, \dots, b_n مقادیر ثابت‌اند، آنگاه

$$\text{cov}(Y_1, Y_p) = \sum_{i=1}^n a_i b_i \text{var}(X_i) + p \sum_{i < j} (a_i b_j + a_j b_i) \text{cov}(X_i, X_j).$$

برهان. اثبات قضیه مشابه برهان قضیه ۱۴-۵ است.

چون وقتی X_i و X_j مستقل‌اند، $\text{cov}(X_i, X_j) = 0$. بی‌درنگ نتیجه می‌شود که

نتیجه‌ی ۴. اگر متغیرهای تصادفی X_1, X_2, \dots, X_n مستقل باشند،

$$\text{cov}(Y_1, Y_p) = \sum_{i=1}^n a_i b_i \text{var}(X_i), \quad Y_p = \sum_{i=1}^n b_i X_i$$

مثال ۱۹-۵. اگر متغیرهای تصادفی X, Y و Z دارای میانگین‌های $\mu_X = 3$ ، $\mu_Y = 2$ و $\mu_Z = 1$ ، واریانس‌های $\sigma_X^2 = 12$ ، $\sigma_Y^2 = 8$ و $\sigma_Z^2 = 5$ و کوواریانس‌های $\text{cov}(Y, Z) = 2$ ، $\text{cov}(X, Z) = -3$ ، $\text{cov}(X, Y) = 1$ باشند، مطلوب کوواریانس

$$W = X + 2Y + 2Z \quad \text{و} \quad V = 3X - Y - Z.$$

حل. بنابر قضیه‌ی ۱۵-۵

$$\begin{aligned} \text{cov}(U, W) &= \text{cov}(X + 2Y + 2Z, 3X - Y - Z) \\ &= 3\text{var}(X) - 2\text{var}(Y) - 2\text{var}(Z) \\ &\quad + 11\text{cov}(X, Y) + 6\text{cov}(X, Z) - 6\text{cov}(Y, Z) \\ &= 3 \times 8 - 2 \times 12 - 2 \times 18 + 11 \times 1 + 6 \times (-3) - 6 \times 2 = -76. \end{aligned}$$

تمرین

۳۱-۵. اگر X و Y دارای توزیع احتمال توأم $f(x, y) = \frac{1}{4}$ به ازای $x = -3, 1$ و $y = 5, 1, -1, -5$ باشند، مقدار $\text{cov}(X, Y)$ را پیدا کنید.

۳۲-۵. اگر توزیع احتمال توأم X و Y به صورت $f(-1, 1) = \frac{1}{4}$ ، $f(-1, 0) = 0$ ، $f(1, 1) = \frac{1}{4}$ ، $f(1, 0) = \frac{1}{4}$ ، $f(0, 1) = 0$ ، $f(0, 0) = \frac{1}{4}$ باشد، نشان دهید: الف) مقدار $\text{cov}(X, Y)$ صفر است.

ب) دو متغیر تصادفی مستقل نیستند.

۵-۳۳. اگر متغیرهای تصادفی مستقل X_1, X_2, \dots, X_m دارای میانگین‌های $4, 9, 3$ و واریانس‌های $3, 7, 5$ باشند، میانگین و واریانس

$$\text{الف)} Y = 2X_1 - 3X_2 + 4X_3$$

$$\text{ب)} Z = X_1 + 2X_2 - X_3$$

را بیابید.

۵-۳۴. $\text{cov}(X+Y, X-Y)$, $\text{var}(X+Y)$, $\text{var}(X-Y)$ را بحسب واریانس‌ها و کوواریانس‌های X و Y بیان کنید.

۵-۳۵. اگر $\text{cov}(X_1, X_2) = v$, $\text{var}(X_1) = v_1$, $\text{var}(X_2) = v_2$, $\text{cov}(X_1, X_3) = -v$, $\text{cov}(X_2, X_3) = -v$ و X_1, X_2, X_3 مستقل باشند، کوواریانس $Y_1 = X_1 + 2X_2 + 3X_3$ و $Y_2 = -2X_1 + 3X_2 + 4X_3$ را پیدا کنید.

۵-۳۶. اگر وقتی سکه‌ای را پرتاپ می‌کنیم، آمدن شیر یک موفقیت باشد، وقتی تاسی را می‌ریزیم آمدن ۶ یک موفقیت باشد، و در کشیدن کارتی از یک دسته‌ی ۵۲ کارتی آمدن تک، یک موفقیت باشد، مطلوب است میانگین و انحراف معیار تعداد کل موفقیت‌ها، وقتی

الف) سکه‌ی همگنی را پرتاپ کنیم، تاس همگنی را بربیزیم، و آنگاه کارتی را از دسته کارتی که خوب بُرخورده است بیرون بکشیم.

ب) سکه‌ی همگنی را سه بار پرتاپ کنیم، تاس همگنی را دو بار بربیزیم، و آنگاه کارتی را از دسته کارتی که خوب بُرخورده است بیرون بکشیم.

۵-۳۷. اگر متناوباً سکه‌ای همگن و سکه‌ای ناهمگن را که احتمال شیر آمدن آن 0.45 است پرتاپ کنیم، میانگین و واریانس تعداد شیرهایی که در ده پرتاپ این سکه‌ها به دست می‌آوریم چقدرند؟

فصل ششم

توزیع‌ها و چگالی‌های احتمال خاص

۱-۶ مقدمه

در این فصل بعضی توزیع‌های احتمال را که در نظریه‌ی آمار و کاربردها به صورتی بسیار چشم‌گیر ظاهر می‌شوند مطالعه می‌کنیم. همچنین پارامترهای این توزیع‌ها، یعنی کمیت‌هایی را که برای توزیع‌های خاص ثابت‌اند، ولی برای اعضای مختلف خانواده‌های توزیع‌های همنوع، مقادیر مختلفی دارند، مطالعه خواهیم کرد. متداول‌ترین پارامترها، گشتاورهای مراتب پایین، عمدتاً l و 5 هستند، و همان‌طور که در فصل قبل دیدیم اساساً دو راه وجود دارد که به وسیله‌ی آنها این گشتاورها به دست می‌آیند. می‌توانیم مجموعه‌ای لازم را مستقیماً محاسبه کنیم، و یا با تابع‌های مولد گشتاورها کار کنیم. گرچه به نظر منطقی می‌رسد که در هر مورد، روشی را که ساده‌تر است به کار ببریم، اما اغلب هر دو راه را به کار خواهیم گرفت. در برخی موارد این کار به دلیل آنکه نتایج حاصل بعداً مورد نیازند انجام خواهد شد؛ و در سایر موارد صرفاً برای دادن تجربه به خواننده در کاربرد تکنیک‌های ریاضی مربوط است. همچنین برای اینکه حدود این فصل را در حد معقول نگاه داریم، بسیاری از جزئیات به عنوان تمرین به عهده‌ی خواننده واگذار شده‌اند.

۲-۶ توزیع یکنواخت گسسته

اگر یک متغیر تصادفی بتواند k مقدار مختلف را با احتمال‌های برابر اختیار کند، گوییم که دارای توزیع یکنواخت گسسته است؛ به صورت نمادی

تعريف ۶-۱. متغیر تصادفی X دارای توزيع یکنواخت گسته است، و به آن، عنوان متغیر تصادفی یکنواخت گسته داده می‌شود، اگر و تنها اگر توزيع احتمال آن، برای

$x_i \neq x_j, i \neq j$ ، وقتی به صورت

$$f(x) = \frac{1}{k}, \quad x = x_1, x_2, \dots, x_k$$

باشد.

بنابر تعریف‌های ۲-۵ و ۵-۴، میانگین و واریانس این توزيع $\mu = \sum_{i=1}^k \frac{1}{k} x_i$ و $\sigma^2 = \sum_{i=1}^k \frac{1}{k} (x_i - \mu)^2$ هستند. این فرمول‌ها با میانگین و واریانسی که در فصل اول شرح داده شد کاملاً منطبق است.

در حالت خاصی که $x_i = i$ ، توزيع یکنواخت گسته به صورت $f(x) = \frac{1}{k}$ به ازای $x = 1, 2, \dots, k$ درمی‌آید، و این شکل توزيع، مثلاً، برای توزيع عددی که در ریختن یک تاس همگن ظاهر می‌شود به کار می‌رود. میانگین و واریانس این توزيع یکنواخت گسته وتابع مولد گشتاورهای آن در تمرین‌های ۱-۶ و ۲-۶ مورد بحث واقع می‌شوند.

۶-۳ توزيع برنولی

اگر آزمایشی دو برآمد داشته باشد، «موفقیت» و «شکست»، و احتمال آنها به ترتیب θ و $1-\theta$ باشند، آنگاه متغیر تصادفی که منطبق با دو وضعیت، یا ۰ و ۱ را اختیار می‌کند، توزيع برنولی دارد. به صورت نمادی

تعريف ۶-۲. متغیر تصادفی X توزيع برنولی دارد، و به آن عنوان متغیر تصادفی برنولی داده می‌شود، اگر و تنها اگر توزيع احتمالش به صورت

$$f(x_i; \theta) = \theta^x (1-\theta)^{1-x}, \quad x = 0, 1$$

باشد.

پس، $f(0; \theta) = (1-\theta)$ و $f(1; \theta) = \theta$. ملاحظه کنید که نماد $f(x; \theta)$ را به کار برده‌ایم

تا صریحاً نشان دهیم که توزیع برنولی دارای یک پارامتر θ است. چون توزیع برنولی حالتی خاص از توزیع بخش ۶-۴ است. در اینجا به تفضیل درباره‌ی آن بحث نخواهیم کرد.

در ارتباط با توزیع برنولی، موفقیت می‌تواند به دست آوردن شیر در پرتاپ یک سکه‌ی همگن، ابتلا به ذات الريه، قبول شدن (یا ردشدن) در یک آزمون، و باختن در یک مسابقه باشد. این ناسازگاری، یادگار زمانی است که نظریه‌ی احتمال فقط در مورد بازی‌های شانسی (که شکست یک بازیکن موفقیت بازیکن دیگر بود) به کار می‌رفت. همچنین به همین دلیل ما به آزمایشی که توزیع برنولی برای آن به کار می‌رود، عنوان آزمایش برنولی یا به صورت ساده آزمایش و به دنباله‌هایی از چنین آزمایش‌هایی، عنوان آزمایش‌های تکراری را می‌دهیم.

۶-۴ توزیع دوجمله‌ای

آزمایش‌های تکراری، نقش بسیار مهمی در احتمال و آمار بازی می‌کنند، خصوصاً وقتی تعداد آزمایش‌ها ثابت، پارامتر θ (احتمال موفقیت) برای تمام آزمایش‌ها برابر، و آزمایش‌ها همگی مستقل باشند. به طوری که خواهیم دید، چندین متغیر تصادفی وجود دارند که در رابطه با آزمایش‌های تکراری پیش می‌آیند. یکی از آنها که در اینجا مطالعه خواهیم کرد مربوط به تعداد کل موفقیت‌هاست؛ دیگر متغیرها را در بخش ۶-۵ ارائه می‌دهیم.

نظریه‌ای که در این بخش از آن بحث می‌کنیم کاربردهای زیادی دارد؛ به عنوان نمونه، اگر بخواهیم بدانیم که احتمال آوردن ۵ شیر در ۱۲ پرتاپ یک سکه، احتمال بهبود ۷ نفر از ۱۰ نفر مبتلا به یک بیماری گرم‌سیری، یا احتمال اینکه ۳۵ نفر از ۸۰ نفر به آگهی فروش یک قلم کالا پاسخ دهند چقدر است این نظریه به کار می‌رود. اما، این مطلب فقط وقتی درست است که هر ۱۰ نفر شناس بهبود یکسانی داشته و بهبود آنها مستقل از یکدیگر باشد. (در نمونه‌های طرح شده، که به وسیله‌ی پزشکان مختلف و در بیمارستان‌های مختلف درمان شوند، و به شرطی که احتمال دادن پاسخ به آگهی فروش کالا برای هریک از ۸۰ نفر یکسان بوده و استقلال نیز وجود داشته باشد، و هیچ دو نفری از آنها به یک خانوار متعلق نباشند).

به منظور تهیه‌ی فرمولی برای احتمال به دست آوردن « x » تا موفقیت در n

آزمایش» تحت شرایطی که بیان شدند، ملاحظه کنید که احتمال به دست آوردن x تا موفقیت و $n-x$ تا شکست در یک ترتیب مشخص برابر $\theta^x(1-\theta)^{n-x}$ است. برای هر موفقیت یک احتمال θ و برای هر شکست یک احتمال $1-\theta$ وجود دارد و بنابر فرض استقلال، x عامل θ و $n-x$ عامل $1-\theta$ در یکدیگر ضرب می‌شوند. چون این احتمال با هر دنباله‌ای از n آزمایش که در آن x موفقیت و $n-x$ شکست وجود دارند همراه است، تنها باید تعداد دنباله‌های از این نوع را بشماریم و سپس $\theta^x(1-\theta)^{n-x}$ را در این تعداد ضرب کنیم. روشن است تعداد راههایی که می‌توانیم x آزمایش را، که برآمد همه‌ی آنها موفقیت است، انتخاب کنیم برابر است با $\binom{n}{x}$ ، و نتیجه می‌شود که احتمال مطلوب برای « x موفقیت در n آزمایش» برابر $\binom{n}{x} \theta^x (1-\theta)^{n-x}$ است.

تعریف ۶-۳. متغیر تصادفی X توزیع دوجمله‌ای دارد، و به عنوان متغیر تصادفی دوجمله‌ای داده می‌شود، اگر و تنها اگر توزیع احتمال آن به صورت

$$b(x; n, \theta) = \binom{n}{k} \theta^x (1-\theta)^{n-x}, \quad x = 0, 1, 2, \dots, n$$

باشد.

پس، تعداد موفقیت‌ها در n آزمایش، متغیری تصادفی است که توزیع دوجمله‌ای با پارامترهای n و θ دارد. نام «توزیع دوجمله‌ای» از این واقعیت نتیجه می‌شود که مقادیر $b(x; n, \theta)$ به ازای $x = 0, 1, 2, \dots, n$ ، جملات متوالی بسط دوجمله‌ای $[(1-\theta) + \theta]^n$ هستند؛ این مطلب نشان می‌دهد که مجموع احتمال‌ها، همان‌گونه که باید، مساوی ۱ است.

اگر سعی کنیم سوئین احتمالی را که مربوط به پاسخ‌دادن به آگهی فروش یک قلم کالا، با قراردادن $x=35$ ، $n=80$ ، و مثلاً $\theta=0.15$ در فرمول توزیع دوجمله‌ای محاسبه کنیم، متوجه می‌شویم انجام این محاسبه مستلزم کار فوق العاده زیادی است. در عمل به ندرت احتمال‌های دوجمله‌ای مستقیماً محاسبه می‌شوند، زیرا آنها به صورتی جامع برای مقادیر مختلف θ و n جدول‌بندی شده‌اند و نرم‌افزارهای کامپیوتری

فراوانی وجود دارند که احتمال‌های دوجمله‌ای و همچنین احتمال‌های تجمعی متناظر آنها، یعنی

$$B(x; n, \theta) = \sum_{k=0}^x b(k; n, \theta)$$

را با دستورهای ساده به دست می‌دهند.

مثال ۶-۱. احتمال به دست آوردن ۵ شیر و ۷ خط را در ۱۲ پرتاب یک سکه‌ی همگن پیدا کنید.

حل. در فرمول توزیع دوجمله‌ای اگر قرار دهیم $x = ۵$ ، $n = ۱۲$ ، و $\theta = \frac{۱}{۳}$ ، به دست می‌آوریم

$$b(5; 12, \frac{1}{3}) = \binom{12}{5} \left(\frac{1}{3}\right)^5 \left(1 - \frac{1}{3}\right)^{12-5}$$

و با محاسبه‌ی مقدار $\binom{12}{5}$ نتیجه می‌گیریم که جواب $\frac{1}{3}^{12} \cdot \frac{1}{5!} \cdot 792$ است. با استفاده از جدول ۱، و رابطه‌ی $P(X=x) = F(x) - F(x^-)$ که در آن $F(x^-)$ نمایانگر مقدار حد چپ تابع توزیع به ازای مقدار x است، مقدار احتمال مربوط به دست می‌آید.

$$\begin{aligned} b(5; 12, \frac{1}{3}) &= P(X \leq 5) - P(X \leq 4) \\ &= B(5; 12, \frac{1}{3}) - B(4; 12, \frac{1}{3}) \\ &= 0.387 - 0.194 \\ &= 0.193. \end{aligned}$$

مثال ۶-۲. احتمال بهبد ۷ نفر از ۱۰ نفر از یک بیماری گرسیری را در صورتی که فرض استقلال برقرار و احتمال بهبد هریک از آنها ۰/۸۰ باشد پیدا کنید.

حل. اگر در فرمول توزیع دوجمله‌ای قرار دهیم $x = ۷$ ، $n = ۱۰$ ، و $\theta = ۰/۸۰$ ، به دست می‌آوریم

$$b(7; 10, 0.80) = \binom{10}{7} (0.80)^7 (1 - 0.80)^{10-7}$$

و با روشی مشابهی مثال ۶-۱، مقدار احتمال فوق برابر با $0.131 = 0.322 - 0.191$ است.

جدول ۱ در انتهای کتاب، مقادیر $B(x; n, \theta)$ را با سه رقم دهدی، برای $n=1$ تا $n=23$ ، و $\theta=0.05, 0.1, 0.2, \dots, 0.9, 0.95$ به دست می‌دهد. می‌توان جدول ۱ را خلاصه‌تر از وضعیت فعلی ارائه داد زیرا وقتی θ بزرگ‌تر از 0.5 است، اتحاد زیرکمیت‌های دیگر را به دست می‌دهد.

$$\text{قضیه ۶.۱.} . b(x; n, \theta) = b(n - x; n, 1 - \theta)$$

اثبات قضیه فوچ به عنوان تمرین به خواننده واگذار شده است. به عنوان مثال، برای تعیین $b(11; 18, 0.75)$ ، مقدار $b(7; 18, 0.25)$ را پیدا می‌کنیم که 0.1376 می‌شود. وقتی n بزرگ است، چندین راه مختلف وجود دارند که طی آنها می‌توان احتمال‌های دوجمله‌ای را تقریب زد. یکی از راه‌ها را در بخش ۶.۶، و یکی دیگر را در بخش ۶.۶ ذکر خواهیم کرد.

در زیر نحوه محاسبه مقدار احتمال وتابع توزیع در توزیع دوجمله‌ای را به وسیله‌ی نرم‌افزار نشان می‌دهیم.

```

MTB > LET K1=5
MTB > LET K2=0.33
MTB > LET K3=10
MTB > PDF K1 K4;
SUBC> Binomial K3 K2.
MTB > PRINT K4
Data Display
K4      0.133151
MTB > CDF K1 K5;
SUBC> Binomial K3 K2.
MTB > PRINT K5
Data Display
K4      0.926800

```

مقدار $K4$ برابر $(0.133151)^5$ و $K5$ برابر با $(0.926800)^5$ است. حال برای میانگین و واریانس توزیع دوجمله‌ای فرمول‌هایی پیدا می‌کنیم.

قضیه ۶.۲. میانگین و واریانس توزیع دوجمله‌ای برابرند با

$$\sigma^2 = n\theta(1-\theta), \quad \mu = n\theta.$$
برهان. برای تعیین میانگین، مستقیماً مجموع زیر را محاسبه می‌کنیم

$$\begin{aligned}\mu &= \sum_{x=0}^n x \binom{n}{x} \theta^x (1-\theta)^{n-x} \\ &= \sum_{x=1}^n \frac{n!}{(x-1)!(n-x)!} \theta^x (1-\theta)^{n-x}\end{aligned}$$

که در آن جمله‌ی متناظر با $x=0$ را که برابر صفر است حذف و عامل x را با موجود در $x!$ که در مخرج $\binom{n}{x}$ است ساده کرده‌ایم. سپس با خارج کردن عامل n که در $n! = n(n-1)!$ وجود دارد و عامل θ ، به دست می‌آوریم

$$\mu = n\theta \sum_{x=1}^n \binom{n-1}{x-1} \theta^{x-1} (1-\theta)^{n-x}$$

که اگر قرار دهیم $y = x - 1$ و $m = n - 1$ ، این عبارت به صورت

$$\mu = n\theta \sum_{y=0}^m \binom{m}{y} \theta^y (1-\theta)^{m-y} = n\theta$$

درمی‌آید، زیرا آخرین مجموع، مساوی مجموع تمام مقادیر توزیع دوجمله‌ای با پارامترهای m و θ است که در نتیجه برابر ۱ است.

در یافتن عبارتی برای μ' و سپس برای σ^2 ، واقعیت $E(X^r) = E[X(X-1)] + E(X)$ را به کار می‌بریم و ابتدا $E[(X(X-1))^2]$ را محاسبه می‌کنیم. برای تمام مقاصد عملی دقیقاً با تکرار مراحل بالا به دست می‌آوریم

$$\begin{aligned}E[(X(X-1))] &= \sum_{x=0}^n x(x-1) \binom{n}{x} \theta^x (1-\theta)^{n-x} \\ &= \sum_{x=2}^n \frac{n!}{(x-2)!(n-x)!} \theta^x (1-\theta)^{n-x} \\ &= n(n-1) \theta^2 \sum_{x=2}^n \binom{n-2}{x-2} \theta^{x-2} (1-\theta)^{n-x},\end{aligned}$$

که اگر قرار دهیم $y = x - 2$ و $m = n - 2$ ، این عبارت به صورت

$$\begin{aligned}E[X(X-1)] &= n(n-1) \theta^2 \sum_{y=0}^m \binom{m}{y} \theta^y (1-\theta)^{m-y} \\ &= n(n-1) \theta^2,\end{aligned}$$

درمی‌آید. بنابر این

$$\mu' = E[X(X-1)] + E[X] = n(n-1)\theta^2 + n\theta$$

و سرانجام

$$\begin{aligned}\sigma^2 &= \mu'_2 - \mu'_1 \\ &= n(n-1)\theta^2 + n\theta + n^2\theta^2 \\ &= n\theta(1-\theta).\end{aligned}$$

برهان دیگری از این قضیه که عملیات جبری کمتری دارد در تمرین ۶-۴ پیشنهاد شده است.

اینکه میانگین توزیع دوجمله‌ای برابر $n\theta$ است نباید شگفت‌انگیز باشد. بحثی ندارد که وقتی سکه‌ی همگنی ۲۰۰ بار پرتاب شود، انتظار داریم که (به مفهوم امید ریاضی)، $= 100 \times \frac{1}{2} = 100$ بار شیر و $100 \times \frac{1}{2} = 100$ بار خط بیاید، به‌همین نحو وقتی تاسی همگن 240 بار ریخته شود انتظار $= 240 \times \frac{1}{2} = 120$ وجه شش را داریم، و وقتی احتمال اینکه شخصی از فروشگاهی خریدی بکند $= 80 / 200 = 0.4$ باشد، انتظار داریم $= 320$ نفر از 400 نفری که به این فروشگاه مراجعه می‌کنند چیزی خریداری کنند.

فرمول واریانس توزیع دوجمله‌ای، که معیار تغییرات است، کاربردهای خیلی مهمی دارد، اما برای تأکید بر اهمیت آن، متغیر تصادفی $\frac{X}{n} = Y$ را در نظر بگیرید، که در آن X متغیر تصادفی است که توزیعی دوجمله‌ای با پارامترهای n و θ دارد. این متغیر تصادفی نسبت موفقیت‌ها در n آزمایش است. خواننده می‌تواند به عنوان تمرین، قضیه‌ی زیر را ثابت کند.

قضیه‌ی ۶-۳. اگر X توزیع دوجمله‌ای با پارامترهای n و θ داشته باشد و آنگاه

$$E(Y) = \theta, \quad \sigma_Y^2 = \frac{\theta(1-\theta)}{n}.$$

حال اگر قضیه‌ی چبیشف را با $k\sigma = c$ به کار ببریم، می‌توانیم حکم کنیم که برای هر ثابت مثبت c ، احتمال اینکه نسبت موفقیت‌ها در n آزمایش بین $c - \theta$ و $c + \theta$ قرار گیرد حداقل برابر است با

$$1 - \frac{\theta(1-\theta)}{n c^r}.$$

در نتیجه، وقتی $n \rightarrow +\infty$ ، احتمال اینکه اختلاف نسبت موفقیت‌ها با θ ، کوچکتر از هر عدد ثابت دلخواه c باشد به ۱ میل می‌کند. این نتیجه به نام قانون اعداد بزرگ موسوم است، و باید توجه‌داشت که این قانون درباره‌ی نسبت موفقیت‌ها به کار می‌رود و نه درباره‌ی تعداد واقعی آنها. اگر فرض کنیم که وقتی n بزرگ است تعداد موفقیت‌ها باید الزاماً نزدیک به $n\theta$ باشد، فرضی اشتباه خواهد بود. چون به دست آوردنتابع مولد گشتاورهای توزیع دوجمله‌ای آسان است، آن را پیدا می‌کنیم و برای تحقیق نتایج قضیه‌ی ۶-۲ به کار می‌بریم.

قضیه‌ی ۶-۴. تابع مولد گشتاور توزیع دوجمله‌ای به صورت

$$M_X(t) = [1 + \theta(e^t - 1)]^n$$

است.

برهان. بنابر تعریف‌های ۶-۵ و ۶-۳، به دست می‌آوریم

$$\begin{aligned} M_X(t) &= \sum_{x=0}^n e^{xt} \binom{n}{x} \theta^x (1-\theta)^{n-x} \\ &= \sum_{x=0}^n \binom{n}{x} (\theta e^t)^x (1-\theta)^{n-x}, \end{aligned}$$

و بنابر قضیه‌ی ۶-۲، به آسانی قابل تشخیص است که این مجموع، بسط دوجمله‌ای

$$[\theta e^t + (1-\theta)]^n = [1 + \theta(e^t - 1)]^n$$

است و این اثبات را کامل می‌کند.

اگر دوبار از $M_X(t)$ نسبت به t مشتق بگیریم، به دست می‌آوریم

$$\begin{aligned} M'_X(t) &= n\theta e^t [1 + \theta(e^t - 1)]^{n-1}, \\ M''_X(t) &= n\theta e^t [1 + \theta(e^t - 1)]^{n-1} + n(n-1)\theta^r e^{rt} [1 + \theta(e^t - 1)]^{n-r} \\ &= n\theta e^t (1-\theta + n\theta e^t) [1 + \theta(e^t - 1)]^{n-r}, \end{aligned}$$

و پس از قراردادن $t = 0$ در آنها، به دست می‌آوریم $\mu'_r = n\theta(1-\theta + n\theta)$ و $\mu'_1 = n\theta$

لذا، $\mu = n\theta$ و $\sigma^r = \mu'_r - \mu^r = n\theta(1-\theta + n\theta) - (n\theta)^r = n\theta(1-\theta)$ که با

فرمول‌هایی که در قضیه‌ی ۶-۲ داده شدند یکی هستند.

از آنچه در این بخش گفته شد که پیدا کردن گشتاورهای توزیع دو جمله‌ای از روی تابع مولد گشتاورها آسان‌تر از محاسبه آنها از راه مستقیم است، اما اگر بخواهیم، مثلاً μ' یا μ'' را معین کنیم آشکار است که مشتق‌گیری نسبتاً مشکل‌تر می‌شود. در واقع راهی آسان‌تر برای تعیین گشتاورهای توزیع دو جمله‌ای وجود دارد؛ راهی که مبتنی بر تابع مولد گشتاورهای عاملی است و در تمرین ۶-۹ توضیح داده شده است.

تمرین

۶-۱. اگر X دارای توزیع یکنواخت گسسته‌ی $f(x) = \frac{1}{k}$ ، به ازای $x = 1, 2, \dots, k$ باشد، نشان دهید که

$$\text{الف) میانگین آن } \mu = \frac{(k+1)}{k} \text{ است،}$$

$$\text{ب) واریانس آن } \sigma^2 = \frac{(k^2 - 1)}{k^2} \text{ است،}$$

۶-۲. اگر X دارای توزیع یکنواخت گسسته‌ی $f(x) = \frac{1}{k}$ به ازای $x = 1, 2, \dots, k$ باشد، نشان دهید که تابع مولد گشتاورهایش به صورت

$$M_X(t) = \frac{e^t (1 - e^{kt})}{k (1 - e^t)}$$

است. همچنین این توزیع را با محاسبه $\lim_{t \rightarrow \infty} M'_X(t)$ بیابید، و آن را با نتیجه‌ای که در قسمت (الف) تمرین ۶-۱ به دست آمده مقایسه کنید.

۶-۳. تحقیق کنید که الف) $b(x; n, \theta) = b(n-x; n, 1-\theta)$

همچنین نشان دهید که اگر برای $x = 0, 1, 2, \dots, n$ $B(x; n, \theta) = \sum_{k=0}^x b(k; n, \theta)$ آنگاه

$$\text{ب) } b(x; n, \theta) = B(x; n, \theta) - B(x-1; n, \theta)$$

$$\text{ج) } b(x; n, \theta) = B(n-x; n, 1-\theta) - B(n-x-1; n, 1-\theta)$$

$$\text{د) } B(x; n, \theta) = 1 - B(n-x-1; n, 1-\theta)$$

۶-۴. برهانی دیگر از قضیه ۶-۲ را می‌توان براساس این واقعیت استوار کرد: اگر X_1, X_2, \dots, X_n متغیرهای تصادفی مستقل و هم‌توزیع با توزیع برنولی با پارامتر θ باشند،

آنگاه $Y = X_1 + \dots + X_n$ متغیری تصادفی است که دارای توزیع دوجمله‌ای با پارامترهای n و θ است.

۶-۵. هنگام محاسبه تمام مقادیر توزیع دوجمله‌ای، می‌توان معمولاً ابتدا با محاسبه‌ی $b(x; n, \theta)$ و سپس استفاده از فرمول بازگشتی

$$b(x+1; n, \theta) = \frac{\theta(n-x)}{(x+1)(1-\theta)} b(x; n, \theta)$$

کار را ساده کرد. درستی این فرمول را تحقیق کنید و آن را برای محاسبه‌ی مقادیر توزیع دوجمله‌ای با $n=7$ و $\theta=0.25$ به کار ببرید.

۶-۶. فرمول بازگشتی تمرین قبل را به کار ببرده، نشان دهید که برای $\theta = \frac{1}{n}$ ، توزیع دوجمله‌ای دارای

الف) ماقسیممی در $x = \frac{n}{2}$ است، وقتی n زوج است،

ب) ماقسیمم‌هایی در $x = \frac{n-1}{2}$ و $x = \frac{n+1}{2}$ است، وقتی n فرد است.

۶-۷. اگر X ، متغیر تصادفی دوجمله‌ای باشد، به ازای چه مقدار θ ، احتمال $b(x; n, \theta)$ ماقسیمم است؟

۶-۸. در برهان قضیه‌ی ۲-۶، کمیت $E[X(X-1)]$ را که دوامین گشتاور عاملی نامیده می‌شود تعیین کردیم؛ به طور کلی، r امین گشتاور عاملی X به صورت

$$\mu'_{(r)} = E[X(X-1)\dots(X-r+1)]$$

است. μ'_1 ، μ'_2 ، و μ'_3 را بر حسب گشتاورهای عاملی بیان کنید.

۶-۹. تابع مولد گشتاورهای عاملی متغیر تصادفی گستته‌ی X به صورت

$$\Psi_X(t) = E(t^X) = \sum_x t^x f(x)$$

داده می‌شود. نشان دهید که r امین مشتق $(\Psi_X(t))'$ نسبت t به ازای $t=1$ برابر $\mu'_{(r)}$ ، همان r امین گشتاور عاملی است که در تمرین ۶-۸ تعریف شد.

۶-۱۰. یک آزمایش تستی چندگزینه‌ای شامل ۸ سئوال، و هر سوال شامل سه پاسخ است (که فقط یکی از آنها صحیح است). اگر دانشجویی با ریختن ناسی همگن به هر سوال پاسخ دهد، به این طریق که اگر ۱ یا ۲ بیاورد جواب اول، اگر ۳ یا ۴ بیاورد جواب دوم و اگر ۵ یا ۶ بیاورد جواب سوم را علامت بزند، احتمال اینکه دقیقاً به ۴ سوال جواب صحیح بدهد چقدر است؟

۱۱-۶. یک مهندس اینمی اتومبیل ادعا می‌کند که از هر ده تصادف یکی ناشی از خستگی راننده است. با استفاده از فرمول توزیع دوجمله‌ای و گردکردن نتیجه تا چهار رقم اعشار، احتمال اینکه حداقل ۳ تا از ۵ تصادف اتومبیل ناشی از خستگی راننده باشد چقدر است؟

۱۲-۶. اگر 40% درصد موش‌هایی که در یک آزمایشگاه هستند، در طول یک دقیقه بعد از تزریق یک داروی آزمایشی حالت تهاجمی شدید پیدا کنند، احتمال اینکه دقیقاً شش تا از پانزده موشی که دارو به آنها تزریق شده است در طول یک هفته حالت تهاجمی شدید پیدا کنند با استفاده از

الف) فرمول توزیع دوجمله‌ای،
ب) جدول ۱،
چقدر است؟

۱۳-۶. در شهری، دلیل قانونی 70% درصد از تمام موارد طلاق عدم توافق اخلاقی اعلام شده است. احتمال اینکه پنج تا از شش مورد طلاق بعدی که در این شهر ثبت می‌شود عدم توافق اخلاقی اعلام شود، با استفاده از

الف) فرمول توزیع دوجمله‌ای،
ب) جدول ۱،
چقدر است؟

۱۴-۶. در برنامه‌ریزی برای راه انداختن یک مدرسه‌ی جدید، یکی از اعضای هیئت امنای مدرسه ادعا می‌کند 14% نفر از 5 نفر از معلمان پاره‌وقت بیش از یک سال در مدرسه به کار خود ادامه می‌دهند، در حالی که عضو دیگر هیئت امنا ادعا می‌کند صحیح آن است که بگوییم، 33% تا از 5 نفر به کار خود ادامه می‌دهند. پیش‌بینی‌های این دو عضو در گذشته تقریباً به یک اندازه اعتبار داشته‌اند، به قسمی که بدون درنظر گرفتن اطلاعات دیگر، می‌توان قضاوت‌های این دو نفر را به یک اندازه معتبر دانست. اگر ادعای یکی از این دو درست باشد، در صورتی که دریابیم که 11 نفر از این 12 نفر معلمان پاره‌وقت بیش از یک سال به کار خود ادامه داده‌اند چه احتمال‌هایی باید به ادعاهای این دو عضو تخصیص دهیم؟

۱۵-۶. با استفاده از قضیه‌ی چیزیف و قضیه‌ی $2-6$ ، تحقیق کنید احتمالی حداقل برابر

$\frac{۳۵}{۳۶}$ وجود دارد که

- (الف) در ۹۰۰ پرتاب یک سکه‌ی همگن، نسبت شیرها بین $۴۰/۰$ و $۶۰/۰$ باشد،
 (ب) در ۱۰۰۰۰ پرتاب یک سکه‌ی همگن، نسبت شیرها بین $۴۷/۰$ و $۵۳/۰$ باشد،
 (ج) در ۱۰۰۰۰۰۰ پرتاب یک سکه‌ی همگن، نسبت شیرها بین $۴۹/۰$ و $۵۰/۰$ باشد.
 توجه کنید که این تمرین برای توضیح قانون اعداد بزرگ به کار می‌رود.

۶-۵ توزیع فوق‌هندسی

در فصل سوم به منظور تشریح قواعد ضرب برای پیشامدهای مستقل و وابسته، نمونه‌گیری با جایگذاری و بدون جایگذاری را به کار بردیم. برای به دست آوردن فرمولی شبیه فرمول توزیع دو جمله‌ای، که در آن نمونه‌گیری بدون جایگذاری است، و در نتیجه آزمایش‌ها مستقل نیستند، مجموعه‌ای از N عنصر را در نظر می‌گیریم که k تای آنها را که دارای یک نسخه‌ی خاص هستند به عنوان موفقیت و $N-k$ تای دیگر را شکست تلقی می‌کنیم. مثل حالت توزیع دو جمله‌ای، احتمال به دست آوردن x موفقیت در n آزمایش مورد توجه است، اما اینک بدون جایگذاری، n عنصر از N عنصر موجود در مجموعه را انتخاب می‌کنیم.

راه برای انتخاب x تا از k موفقیت و $\binom{N-k}{n-x}$ راه برای انتخاب x تا از $N-k$ شکست وجود دارد، و بنابراین $\binom{k}{x} \binom{N-k}{n-x}$ راه برای انتخاب x موفقیت و $n-x$ شکست موجود است. چون $\binom{N}{n}$ راه برای انتخاب n عنصر از N عنصر جامعه وجود دارد و ما فرض می‌کنیم تمام آنها هم‌شانس، لذا از قضیه‌ی ۲-۳ نتیجه می‌شود که احتمال « x موفقیت در n آزمایش» برابر است.

$$\frac{\binom{k}{x} \binom{N-k}{n-x}}{\binom{N}{n}}$$

تعریف ۶-۴. متغیر تصادفی X ، دارای توزیع فوق‌هندسی است و به آن، عنوان متغیر تصادفی فوق‌هندسی داده می‌شود، اگر و تنها اگر توزیع احتمال آن برای $n-x \leq N-k$ ، $x=0, 1, 2, \dots, n$ به صورت

$$h(x; n, N, k) = \frac{\binom{k}{x} \binom{N-k}{n-x}}{\binom{N}{n}}$$

باشد.

پس، برای نمونه‌گیری بدون جایگذاری، تعداد موفقیت‌ها در n آزمایش، متغیری تصادفی با توزیع فوق هندسی با پارامترهای n , N , و k دارد.

مثال ۳-۶. به عنوان بخشی از بررسی آلودگی هوا، بازرگانی تصمیم می‌گیرد مواد آلودکننده‌ای را که از آگروز ۶تا از ۲۴ ماشین شرکتی خارج می‌شوند آزمایش کند. اگر ۴تا از ماشین‌های شرکت به میزان زیادی ماده‌ی آلوده کننده‌ی هوا منتشر کنند، احتمال اینکه نمونه‌ی بررسی کننده شامل هیچ یک از آنها نباشد چقدر است؟ حل. اگر در فرمول توزیع فوق هندسی قرار دهیم $x=0$, $n=6$, $N=24$, و $k=4$, به دست می‌آوریم

$$h(0; 6, 24, 4) = \frac{\binom{4}{0} \binom{20}{6}}{\binom{24}{6}} = 0/2880.$$

روشی که با آن، میانگین و واریانس توزیع فوق هندسی را پیدا می‌کنیم شباهت زیاد به روشی دارد که در برهان ۲-۶ به کار رفت.

قضیه ۶-۵. میانگین و واریانس توزیع فوق هندسی عبارت‌اند از

$$\mu = \frac{n k}{N}, \quad \sigma^2 = \frac{n k (N-k) (N-n)}{N^2 (N-1)}.$$

برهان. برای تعیین میانگین، مجموع زیر را مستقیماً محاسبه می‌کنیم

$$\mu = \sum_{x=0}^n x \frac{\binom{k}{x} \binom{N-k}{n-x}}{\binom{N}{n}} = \sum_{x=1}^n \frac{k!}{(x-1)! (k-x)!} \frac{\binom{N-k}{n-x}}{\binom{N}{n}},$$

که در آن جمله‌ای را که متناظر با $x=0$, مساوی صفر است حذف، و x را با اوّلین

عامل $x! = x(x-1)\dots(1)$ است ساده کرده‌ایم. لذا با فاکتور گرفتن از

$\frac{k}{N}$ ، به دست می‌آوریم

$$\mu = \frac{k}{N} \sum_{x=1}^n \binom{k-1}{x-1} \binom{N-k}{n-x}.$$

که با قرار دادن $x = n - y$ و $m = n - 1$ به صورت

$$\mu = \frac{k}{N} \sum_{y=0}^m \binom{k-1}{y} \binom{N-k}{m-y}$$

در می‌آید. سرانجام با استفاده از قضیه‌ی ۲-۲، به دست می‌آوریم

$$\mu = \frac{k}{N} \binom{N-1}{m} = \frac{k}{N} \binom{N-1}{n-1} = \frac{n k}{N}$$

برای به دست آوردن فرمولی برای σ^2 ، مانند آنچه در برهان قضیه‌ی ۲-۶ آمد عمل می‌کنیم، یعنی ابتدا $E[X(X-1)]$ را محاسبه می‌کنیم، و سپس این واقعیت را که

$$E(X') = E(X(X-1)) + E(X)$$

به کار می‌بریم و با واگذاری اثبات

$$E[X(X-1)] = \frac{k(k-1)n(n-1)}{N(N-1)}$$

در تمرین ۶-۱۸ به عهده‌ی خواننده، به دست می‌آوریم

$$\begin{aligned} \sigma^2 &= \frac{k(k-1)n(n-1)}{N(N-1)} + \frac{n k}{N} - \left(\frac{n k}{N}\right)^2 \\ &= \frac{n k (N-k) (N-n)}{N^2 (N-1)}. \end{aligned}$$

وقتی N بزرگ و n در مقایسه با N نسبتاً کوچک است (قانون سرانگشتی معمولی آن است که n نباید از 5 درصد N تجاوز کند)، تفاوت چندانی بین نمونه‌گیری باجایگذاری و بدونجایگذاری وجود ندارد، و برای تقریب احتمال‌های فوق هندسی می‌توان فرمول توزیع دوجمله‌ای را با پارامترهای n و $\theta = \frac{k}{N}$ به کار برد.

مثال ۶-۴. بین ۱۲۰ نفر متقارضی شغلی فقط ۸۰ نفرشان واقعاً واجد شرایط‌اند. اگر ۵ نفر از متقارضیان بدونجایگذاری و به تصادف برای یک مصاحبه مفصل انتخاب شوند، با استفاده از الف) توزیع فوق هندسی، و

ب) توزیع دوجمله‌ای با $\theta = \frac{80}{120} = \frac{2}{3}$ به عنوان یک تقریب، پیدا کنید احتمال آن را که تنها ۲ نفر از ۵ نفر برای شغل مزبور واجد شرایط باشند.

حل. الف) اگر در فرمول توزیع فوق‌هندسی قرار دهیم $N = 120$ ، $n = 5$ ، $x = 2$ ، $k = 80$ ، با گردکردن نتیجه تا سه رقم اعشار به دست می‌آوریم

$$h(2; 5, 120, 80) = \frac{\binom{80}{2} \binom{40}{3}}{\binom{120}{5}} = 0.164$$

ب) اگر در فرمول توزیع دوجمله‌ای قرار دهیم $x = 2$ ، $n = 5$ ، و $\theta = \frac{2}{3}$ ، با گردکردن نتیجه تا سه رقم اعشار، به دست می‌آوریم

$$b(2; 5, \frac{2}{3}) = \binom{5}{2} \left(\frac{2}{3}\right)^2 \left(1 - \frac{2}{3}\right)^3 = 0.165.$$

همان‌طور که از این نتایج مشهود است، تقریب حاصل خیلی دقیق است.

تمرین

۶-۱۶. یک مهندس کنترل کیفیت، نمونه‌ای تصادفی مرکب از دو ماشین حساب جیبی را که از هر بسته‌ی ۱۸ تایی دریافت شده انتخاب شده است بررسی می‌کند، اگر هر دو خوب کار کنند بسته را می‌پذیرد؛ در غیر این صورت تمام بسته به خرج فروشنده

بررسی می‌شود. احتمال اینکه این بسته بدون بررسی بعدی پذیرفته شود چقدر است در صورتی که این بسته شامل

- الف) ۱۰ ماشین حساب باشد که خوب کار نکنند،
- ب) ۸ ماشین حساب باشد که خوب کار نکنند،
- ج) ۱۲ ماشین حساب باشد که خوب کار نکنند؟

۶-۱۷. بین ۱۶ متقارضی شغلی ده نفر تحصیلات دانشگاهی دارند. اگر سه متقارضی به-

تصادف برای مصاحبه انتخاب شوند، احتمال اینکه

- الف) هیچ یک تحصیلات دانشگاهی نداشته باشند،
 - ب) یکی تحصیلات دانشگاهی داشته باشد،
 - ج) دو نفر تحصیلات دانشگاهی داشته باشند،
 - د) هر سه تحصیلات دانشگاهی داشته باشند،
- چقدر است؟

۶-۱۸. میانگین و واریانس توزیع فوق هندسی را با $N=16$ ، $n=3$ ، و $k=10$ با

استفاده از

- الف) نتایج تمرین قبل،
 - ب) فرمول‌های قضیه‌ی ۵-۶،
- بیابید.

۶-۱۹. احتمال آنکه ممیز مالیاتی، از بین ۵ اظهارنامه‌ی مالیاتی فقط ۲ اظهارنامه با بخشودگی‌های غیرمجاز باید چقدر است، به‌شرطی که این ۵ اظهارنامه را به تصادف از بین ۱۵ اظهارنامه که شامل ۹ اظهارنامه با بخشودگی غیرمجاز است انتخاب کرده باشد؟

۶-۲۰. محموله‌ای مرکب از ۸۰ دستگاه شامل ۴ دستگاه معیوب است. اگر سه تا از

این دستگاه‌ها به تصادف انتخاب و برای یک مشتری ارسال شوند، با استفاده از

- الف) فرمول توزیع فوق‌هندسی،
 - ب) توزیع دوجمله‌ای به‌عنوان یک تقریب،
- احتمال این را پیدا کنید که مشتری یک دستگاه معیوب دریافت کند.

۶- توزیع پواسون

وقتی n بزرگ است، محاسبه احتمال‌های دوچمله‌ای با استفاده از فرمول تعریف ۳-۶ معمولاً مستلزم کار فوق‌العاده زیادی است. به عنوان مثال، برای محاسبه احتمال اینکه ۱۸ نفر از ۳۰۰۰ نفری که در یک روز گرم تابستان رژه‌ای را تماشا می‌کنند دچار گرم‌گی شوند، مجبوریم ابتدا $\binom{3000}{18}$ را تعیین کنیم، و اگر احتمال اینکه یکی از ۳۰۰۰ نفری که رژه را تماشا می‌کنند دچار گرم‌گی شود برابر 50% باشد، مجبوریم مقدار $^{995}(0/005)^{18}$ را نیز محاسبه کنیم.

در این بخش توزیع احتمالی را معرفی خواهیم کرد که می‌توان آن را برای تقریب زدن این نوع توزیع دوچمله‌ای به کار برد. خصوصاً صورتی حدی از توزیع دوچمله‌ای را وقتی $n \rightarrow +\infty$ ، $\theta \rightarrow 0$ و در عین حال $n\theta$ ثابت می‌ماند بررسی خواهیم کرد. اگر این مقدار ثابت را λ فرض کنیم، یعنی $n\theta = \lambda$ ، و در نتیجه $\frac{\lambda}{n} = \theta$ می‌توانیم بنویسیم

$$\begin{aligned} b(x; n, \theta) &= \binom{n}{x} \left(\frac{\lambda}{n}\right)^x \left(1 - \frac{\lambda}{n}\right)^{n-x} \\ &= \frac{n(n-1)(n-2)\cdots(n-x+1)}{x!} \left(\frac{\lambda}{n}\right)^x \left(1 - \frac{\lambda}{n}\right)^{n-x}. \end{aligned}$$

حال اگر هریک از x عامل n را که در $\left(\frac{\lambda}{n}\right)^x$ وجود دارد به ترتیب، مخرج هریک از عامل‌های حاصل ضرب $n(n-1)(n-2)\cdots(n-x+1)$ را به صورت

$$\begin{aligned} &\left[\left(1 - \frac{\lambda}{n}\right)^{-n} \right]^{-\lambda} \left(1 - \frac{\lambda}{n}\right)^{-x} \\ &\text{بنویسیم، به دست می‌آوریم} \\ &\frac{1 \left(1 - \frac{1}{n}\right) \left(1 - \frac{2}{n}\right) \cdots \left(1 - \frac{x-1}{n}\right)}{x!} (\lambda)^x \left[\left(1 - \frac{\lambda}{n}\right)^{-n} \right]^{-\lambda} \left(1 - \frac{\lambda}{n}\right)^{-x}. \end{aligned}$$

سرانجام اگر فرض کنیم $n \rightarrow +\infty$ و x و λ را ثابت نگه داریم، به دست می‌آوریم

$$(1 - \frac{1}{n})(1 - \frac{2}{n}) \cdots (1 - \frac{x-1}{n}) \rightarrow 1$$

$$\left(1 - \frac{\lambda}{n}\right)^{-x} \rightarrow 1$$

$$\left(1 - \frac{\lambda}{n}\right)^{\frac{-n}{\lambda}} \rightarrow e$$

و در نتیجه، توزیع حدی به صورت

$$p(x; \lambda) = \frac{\lambda^x e^{-\lambda}}{x!}, \quad x = 0, 1, 2, \dots$$

درمی‌آید.

تعريف ۵-۶. متغیر تصادفی X ، توزیع پواسون دارد، و به آن عنوان متغیر تصادفی پواسون داده می‌شود، اگر و تنها اگر توزیع احتمال آن به صورت

$$p(x; \lambda) = \frac{\lambda^x e^{-\lambda}}{x!}, \quad x = 0, 1, 2, \dots$$

باشد.

جدول ۶-۱. خروجی کامپیوتر برای توزیع دوجمله‌ای با $n=150$ و $\theta=0.05$ و توزیع پواسون $\lambda=7.5$

| x | b(x; 150, 0.05) | p(x; 7.5) | x | b(x; 150, 0.05) | p(x; 7.5) |
|----|-----------------|-----------|----|-----------------|-----------|
| 1 | 0.003596 | 0.004148 | 11 | 0.058211 | 0.058521 |
| 2 | 0.014102 | 0.015555 | 12 | 0.035488 | 0.036575 |
| 3 | 0.036616 | 0.038889 | 13 | 0.019828 | 0.021101 |
| 4 | 0.070823 | 0.072916 | 14 | 0.010212 | 0.011304 |
| 5 | 0.108843 | 0.109375 | 15 | 0.004873 | 0.005652 |
| 6 | 0.138441 | 0.136718 | 16 | 0.002164 | 0.002649 |
| 7 | 0.149891 | 0.146484 | 17 | 0.000898 | 0.001169 |
| 8 | 0.141016 | 0.137329 | 18 | 0.000349 | 0.000487 |
| 9 | 0.117101 | 0.114444 | 19 | 0.000128 | 0.000192 |
| 10 | 0.086901 | 0.08583 | 20 | 0.000044 | 0.000072 |

بنابراین، در حد، وقتی $n \rightarrow +\infty$ ، $\theta \rightarrow 0$ ، و $n\theta = \lambda$ ثابت باقی می‌ماند، تعداد موفقیت‌ها متغیری تصادفی است که توزیع پواسون با پارامتر λ دارد. این توزیع به افتخار ریاضی‌دان فرانسوی، سیمون پواسون (۱۸۴۰-۱۷۸۱) نام‌گذاری شده است.

به طور کلی وقتی $n \geq 20$ ، و $\theta / n \leq 0.05$ ، توزیع پواسون تقریبی خوب برای احتمال‌های دوچمله‌ای به دست می‌دهد. وقتی $n \geq 100$ و $n\theta > 10$ ، عموماً تقریب بسیار عالی است. برای به دست آوردن ایده‌ای درباره‌ی نزدیکی تقریب پواسون برای توزیع دوچمله‌ای، خروجی کامپیوتری شکل ۶-۱ را در نظر بگیرید که توزیع دوچمله‌ای را با $n = 150$ و $\theta = 0.05$ در بالای توزیع پواسون با $\lambda = 7/5 = 1.4$ نشان می‌دهد.

مثال ۶-۵. وقتی توزیع پواسون با $\lambda = 7/5$ را برای تقریب توزیع دوچمله‌ای با $n = 150$ و $\theta = 0.05$ به کار می‌بریم، با استفاده از جدول ۶-۱، مقدار x ‌ی (از ۱ تا ۲۰) را که برای آن میزان خطای بزرگترین قدر مطلق را دارد معین کنید.

حل. با محاسبه‌ی تفاضل‌های متناظر با $x = 1, 2, \dots, 6, x = 20, \dots, 25$ ، ماکسیمم خطای (از لحاظ عددی) برابر با 0.0036781 است که با $x = 8$ متناظر است.

مثال بعدی، تقریب پواسون را برای توزیع دوچمله‌ای نشان می‌دهند.

مثال ۶-۶. اگر دو درصد از کتاب‌های که در یک صحافی جلد شده‌اند، بد صحافی شده باشند، با استفاده از تقریب پواسون برای توزیع دوچمله‌ای، احتمال آن را تعیین کنید که ۵ جلد از ۴۰۰ کتاب جلد شده‌ی این صحافی بد صحافی شده باشند.

حل. با قرار دادن $\lambda = 8, x = 5$ در فرمول تعریف ۶-۷، به دست می‌آوریم

$$p(5; \lambda) = \frac{\lambda^5 e^{-\lambda}}{5!} = \frac{(32768)(0.00034)}{120} = 0.093$$

در عمل، احتمال‌های پواسون به ندرت با جانشینی مستقیم در فرمول تعریف ۶-۵ به دست می‌آیند. گاهی به جداول احتمال‌های پواسون، نظیر جدول ۲ آخر کتاب، یا به جداول جامع‌تر در کتاب‌های دستی جداول آماری مراجعه می‌کنیم، اما این روزها اغلب به نرم‌افزارهای کامپیوتری مناسب رجوع می‌کنیم. وقتی با احتمال‌های مربوط به چندین مقدار x سروکار داریم، استفاده از جدول‌ها یا کامپیوترها اهمیتی خاص پیدا می‌کند.

مثال ۶-۷. سوابق نشان می‌دهند که احتمال پنچرشدن اتومبیلی در حال عبور از پل

معین $5/0000$ است. با استفاده از توزیع پواسون برای تقریب احتمال‌های دو‌جمله‌ای، احتمال آن را بباید که بین 10000 اتومبیل که از این پل عبور می‌کنند

(الف) دقیقاً دو اتومبیل دچار پنچری شوند؛
ب) حداقل دو اتومبیل دچار پنچری شوند.

حل. الف) با رجوع به جدول ۲ برای $x=2$ ، و $\lambda=10000 \cdot 5/0000 = 50$ ، احتمال پواسون $0/0758$ را به دست می‌آوریم.

ب) با رجوع به جدول ۲، برای $x=1, 2$ و $\lambda=5/0$ احتمال‌های پواسون را به ترتیب برابر $0/0758$ ، $0/03033$ به دست می‌آوریم. پس، احتمال اینکه بین 10000 اتومبیل که از پل عبور می‌کنند حداقل دو اتومبیل دچار پنچری شوند برابر است با

$$0/09856 + 0/0758 = 0/06065.$$

حال که توزیع پواسون را به عنوان صورت حدی توزیع دو‌جمله‌ای به دست آورده‌یم، می‌توانیم فرمول‌های میانگین و واریانس آن را با به‌کاربردن همان شرایط حدی ($n \rightarrow +\infty$ ، $0 \rightarrow \theta$ و ثابت ماندن $n\theta = \lambda$) از روی میانگین و واریانس توزیع دو‌جمله‌ای به دست آوریم. برای میانگین $\mu = n\theta = \lambda$ ، و برای واریانس $\sigma^2 = n\theta(1-\theta) = \lambda(1-\lambda)$ حاصل می‌شود که وقتی $0 \rightarrow \theta$ ، به λ می‌گراید.

قضیه ۶-۶. میانگین و واریانس توزیع پواسون عبارت‌اند از

$$\mu = \lambda, \quad \sigma^2 = \lambda.$$

این نتایج را می‌توان با محاسبه‌ی مجموعه‌ای لازم یا از طریق تابع مولد گشتاورها نیز به دست آورد.

قضیه ۶-۷. تابع مولد گشتاورهای توزیع پواسون به صورت

$$M_X(t) = e^{\lambda(e^t - 1)}$$

است.

برهان. بنابر تعریف‌های ۶-۵ و ۶-۶

$$M_X(t) = \sum_{x=0}^{\infty} e^{xt} \frac{\lambda^x e^{-\lambda}}{x!} = e^{-\lambda} \sum_{x=0}^{\infty} \frac{(\lambda e^t)^x}{x!},$$

که در آن $z = \lambda e^t$ با $e^z = \sum_{x=0}^{+\infty} \frac{(\lambda e^t)^x}{x!}$ همان سری ماکلورن است، پس

$$M_X(t) = e^{-\lambda} e^{\lambda e^t} = e^{\lambda(e^t - 1)}.$$

در این صورت اگر از $M_X(t)$ دوبار نسبت به t مشتق بگیریم، به دست می‌آوریم

$$M'_X(t) = \lambda e^t e^{\lambda(e^t - 1)}$$

$$M''_X(t) = \lambda e^t e^{\lambda(e^t - 1)} + \lambda^2 e^t t e^{\lambda(e^t - 1)}$$

به قسمی که $\mu' = M''_X(0) = \lambda + \lambda^2$ و $\mu'' = M'_X(0) = \lambda$ پس، $\mu = \lambda$ و $\sigma^2 = \mu'' - \mu' = (\lambda + \lambda^2) - \lambda^2 = \lambda$ با قضیه ۶-۸ مطابقت دارد.

گرچه توزیع پواسون به صورت شکل حدی توزیع دوجمله‌ای حاصل شد، ولی کاربردهای فراوانی دارد که رابطه‌ی مستقیم با توزیع‌های دوجمله‌ای ندارند. مثلاً، توزیع پواسون را می‌توان به عنوان مدلی برای تعداد موفقیت‌هایی که در طول بازه‌ی زمانی

مفروض یا در ناحیه‌ای مشخص رخ می‌دهند به کار برد، به شرطی که

۱. تعداد موفقیت‌ها در بازه‌های زمانی یا در ناحیه‌های نامتناخل مستقل باشند؛

۲. احتمال رخداد تنها یک موفقیت در بازه‌های زمانی کوتاه یا در هر ناحیه‌ی کوچک، متناسب با طول بازه‌ی زمانی یا اندازه‌ی ناحیه باشد؛

۳. احتمال رخداد بیش از یک موفقیت در چنین بازه‌ی زمانی کوتاه با قرار گرفتن در چنین ناحیه‌ای کوچک ناچیز باشد.

بنابراین، توزیع پواسون ممکن است تعداد مکالمات تلفنی دریافتی یکی اداره را در یک ساعت، تعداد خطاهای تایپی را در یک صفحه، یا تعداد باکتری‌های یک کشت مفروض را، وقتی متوسط تعداد موفقیت‌ها، λ ، برای بازه‌ی زمانی مفروض یا ناحیه‌ای مشخص معلوم باشد توصیف کند.

مثال ۶-۸. می‌دانیم متوسط تعداد کامیون‌هایی که در هر روز به توقف‌گاه کامیون‌های شهری می‌رسند ۱۲ است. احتمال اینکه در روزی معین کمتر از ۹ کامیون به توقف‌گاه برسند چقدر است؟

حل. فرض کنید X تعداد کامیون‌هایی باشد که در روزی مفروض وارد توقف‌گاه می‌شوند. در این صورت، با استفاده از جدول ۲، با $\lambda = 12$ ، به دست می‌آوریم

$$P(X < 9) = \sum_{x=0}^8 p(x; 12) = 0.1550.$$

اگر در وضعیتی که برای آن شرایط قبلی برقرارند، موفقیت‌ها با میانگین نرخ α در واحد زمان یا در واحد ناحیه رخ دهند، آنگاه تعداد موفقیت‌ها در بازه‌ی t واحد زمان یا t واحد از ناحیه‌ی مشخص، متغیر تصادفی پواسون با میانگین $\lambda = \alpha t$ است. بنابر این، تعداد موفقیت‌ها X در بازه‌ی زمانی به طول t واحد یا ناحیه‌ی به اندازه‌ی t واحد، دارای توزیع پواسون

$$p(x; \alpha t) = \frac{e^{-\alpha t} (\alpha t)^x}{x!}, \quad x = 0, 1, 2, \dots$$

است.

مثال ۶-۹. نوع خاصی از یک صفحه‌ی فلزی، در هر ۱۰ فوت مربع به‌طور متوسط ۵ عیب دارد. اگر توزیع عیب‌ها را پواسون فرض کنیم، احتمال اینکه در ۱۵ فوت مربع یک صفحه‌ی فلزی، حداقل ۶ عیب وجود داشته باشد چقدر است؟

حل. فرض می‌کنیم X ، معرف تعداد عیب‌ها در ۱۵ فوت مکعب از صفحه‌ی فلز باشد. در این صورت چون واحد مساحت، برابر با ۱۰ فوت است، داریم

$$\lambda = \alpha t = (5)(1/5) = 1,$$

و در نتیجه $P(X \geq 6) = 1 - P(X \leq 5) = 1 - 0.7586 = 0.2414$. این محاسبات براساس خروجی برنامه‌ای از نرم‌افزار مینی‌تب استخراج شده است.

```
MTB > CDF 5 k6;
SUBC> Poisson 7.5.
MTB > let k7=1-k6
MTB > print k7
Data Display
K7 0.758564
```

تمرین

- ۲۱-۶. هنگام محاسبه‌ی همه‌ی مقادیر توزیع پواسون، غالباً می‌توان ابتدا با محاسبه‌ی $p(x; \lambda)$ و سپس با استفاده از فرمول بازگشته‌ی

$$p(x+1; \lambda) = \frac{\lambda}{x+1} p(x; \lambda)$$

انجام کار را تسهیل کرد. درستی این فرمول را تحقیق کنید و آن را با توجه به $e^{-\lambda} = 0.135^3$ برای تحقیق در درستی مقادیری که در جدول ۲ برای $\lambda = 2$ داده شده‌اند به کار ببرید.

- ۲۲-۶. با مشتق‌گیری از عبارت دو طرف معادله‌ی

$$\mu_r = \sum_{x=0}^{\infty} (x - \lambda)^r \frac{\lambda^x e^{-\lambda}}{x!}$$

نسبت به λ ، فرمول بازگشته‌ی

$$\mu_{r+1} = \lambda \left[r \mu_{r-1} + \frac{d\mu_r}{d\lambda} \right]$$

را به ازای $r = 1, 2, 3, \dots$ ، برای گشتاورهای حول میانگین توزیع پواسون به دست آورید.

- ۲۳-۶. تابع مولد گشتاورهای $Y = X - \lambda$ را بباید که در آن X متغیری تصادفی است که توزیع پواسون با پارامتر λ دارد. سپس آن را برای تحقیق $\sigma_x^2 = \lambda$ به کار ببرید.

- ۲۴-۶. وقتی می‌خواهیم توزیع پواسون را برای تقریب کردن احتمال‌های دوچمله‌ای به کار ببریم، در هر یک از موارد زیر بررسی کنید که آیا مقادیر n و θ در قاعده‌ی سرانگشته‌ی برای تقریب خوب، تقریب عالی، یا هیچکدام صدق می‌کنند یا نه

الف) $n = 125$ و $\theta = 0.10$

ب) $n = 25$ و $\theta = 0.04$

ج) $n = 120$ و $\theta = 0.05$

د) $n = 40$ و $\theta = 0.06$

- ۲۵-۶. تجربه نشان داده است که $1/14$ درصد از تلفن‌هایی که به یک تلفن‌خانه می‌شود

شماره‌های اشتباه‌اند. برای تعیین احتمال آنکه بین ۱۵۰ تلفن دریافتی دو نمره‌ی اشتباه وجود داشته باشد تقریب پواسون برای توزیع دوجمله‌ای را به کار برد.

۶-۲۶. سوابق موجود نشان می‌دهند شخصی که روزی را در یک نمایشگاه استانی می‌گذراند با احتمال 0.001^2 دچار مسمومیت غذایی می‌شود. با استفاده از تقریب پواسون برای توزیع دوجمله‌ای، مطلوب است احتمال اینکه بین ۱۰۰۰ بازدیدکننده‌ی نمایشگاه حداقل دو نفر دچار مسمومیت غذایی شوند.

۶-۲۷. در شهری معین، 4% درصد همه‌ی رانندگان مجاز، در هر سال حداقل در گیر یک تصادف اتومبیل هستند. با استفاده از تقریب پواسون برای توزیع دوجمله‌ای، مطلوب است احتمال آنکه بین ۱۵۰ راننده‌ی مجازی که به تصادف انتخاب شده‌اند

- الف) تنها ۵ نفر در سالی معین، حداقل یک تصادف اتومبیل داشته باشد،
- ب) حداقل 3^{rd} نفر در سالی معین، حداقل یک تصادف اتومبیل داشته باشند.

۶-۲۸. تعداد شکایت‌هایی که روزانه از یک مؤسسه‌ی خشکشویی به عمل می‌آید، متغیری تصادفی است که دارای توزیع پواسون با $\lambda = \frac{3}{3} = 1$ است. با استفاده از فرمول توزیع پواسون، احتمال آن را باید که در روزی معین، تنها دو شکایت انجام شود.

۶-۲۹. تعداد از کارافتادگی ماهیانه‌ی کامپیوتری، متغیری تصادفی است که توزیع پواسون با $\lambda = 1/8$ دارد. با استفاده از فرمول توزیع پواسون، احتمال آن را باید که این کامپیوتر در یک ماه

- الف) بدون از کارافتادگی،
- ب) تنها با یک از کارافتادگی،
- کار کند.

۶-۳۰. در یک ناحیه‌ی کویری تعداد افرادی که هر سال از خوردن یک نوع گیاه سمی شدیداً بیمار می‌شوند، متغیری است که توزیع پواسون با $\lambda = 5/2 = 2.5$ دارد. با استفاده از جدول ۲، مطلوب است احتمال آنکه

- الف) در سالی معین ۳ نفر،
- ب) در سالی معین حداقل ۱۰ نفر،
- ج) در سالی معین تعدادی از ۴ تا ۶ نفر،
به این بیماری مبتلا شوند.

۷-۶ چگالی یکنواخت

چگالی‌های احتمال مثال‌های ۱۰-۴ و ۱۳-۴ حالت‌هایی خاص از چگالی یکنواخت‌اند که نمودار آنها را می‌توان مثل شکل ۶-۴ رسم کرد.

تعریف ۶-۶. متغیر تصادفی X دارای چگالی یکنواخت است، و به آن متغیر تصادفی یکنواخت پیوسته اطلاق می‌شود، اگر و تنها اگر چگالی احتمال آن به صورت

$$f(x) = \begin{cases} \frac{1}{\beta-\alpha} & \alpha < x < \beta \\ 0 & \text{سایر جاهای} \end{cases}$$

باشد. پارامترهای α و β ثابت‌های حقیقی‌اند و $\alpha < \beta$.

در تمرین ۶-۳۱ از خواننده خواسته شده است که تحقیق کند که

قضیه ۶-۸. میانگین و واریانس چگالی یکنواخت عبارت‌اند از

$$\mu = \frac{\alpha + \beta}{2}, \quad \sigma^2 = \frac{1}{12}(\beta - \alpha)^2.$$

گرچه چگالی یکنواخت برخی کاربردهای مستقیم دارد، ولی ارزش اصلی‌اش آن است که به دلیل سادگی، به آسانی برای توضیح جنبه‌های مختلف نظریه‌ی آماری به کار می‌آید.

تمرین

۶-۳۱. نشان دهید که اگر X متغیری تصادفی دارای چگالی یکنواخت با پارامترهای α و β باشد،

الف) احتمال اینکه مقداری کمتر از $(\beta - \alpha)p + \alpha$ اختیار کند مساوی p است.

ب) امید ریاضی و واریانس X را بیابید.

۶-۳۲. نشان دهید که اگر متغیر تصادفی دارای چگالی یکنواخت با پارامترهای α و β باشد، ۲۰٪ میانگین آن برابر است با

الف) α ، وقتی r فرد است،

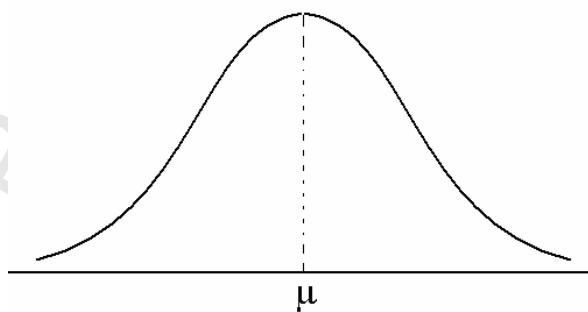
$$\text{ب) } \left(\frac{\beta - \alpha}{r+1} \right)^r, \text{ وقتی } r \text{ زوج است.}$$

۳۳-۶. نقطه‌ی D روی پاره‌خط AB که نقطه‌ی وسطش C و طولش a است انتخاب شده است. اگر X ، فاصله‌ی D از A ، متغیری تصادفی باشد که چگالی یکنواخت با $\alpha = 0$ و $\beta = a$ دارد، احتمال آنکه AD ، BD ، و AC تشکیل مشتمل بدهند چقدر است؟

۳۴-۶. در برخی از آزمایش‌ها، خطایی که در تعیین وزن مخصوص یک ماده، مرتکب می‌شوند متغیری تصادفی است که دارای چگالی یکنواخت با $\alpha = -0.015$ و $\beta = 0.015$ است. تعیین کنید، احتمال اینکه
 الف) چنین خطایی بین -0.002 و 0.003 باشد،
 ب) قدر مطلق چنین خطایی از 0.005 تجاوز کند.

۸-۶ توزیع نرمال

توزیع نرمال که در این بخش مطالعه می‌شود از جهات زیادی، شالوده‌ای برای نظریه‌ی آماری نوین است. این توزیع ابتدا در سده‌ی هجدهم، وقتی دانشمندان نظمی بسیار شگفت‌انگیز را در خطاهای اندازه‌گیری مشاهده کردند، مورد بررسی قرار گرفت. این دانشمندان دریافتند که الگوها (توزیع‌ها) یک را که مشاهده می‌کنند می‌توان به دقت با منحنی‌های پیوسته‌ای که آنها را «منحنی‌های نرمال خطاهای» نامیدند و به قوانین شانس نسبت دادند تقریب کرد. ویژگی‌های ریاضی چنین منحنی‌های نرمالی بدؤاً به وسیله‌ی آبراهام دموآور (۱۶۶۷-۱۷۴۵)، پیر لابلس (۱۷۴۹-۱۸۲۷)، و کارل گاووس (۱۷۷۷-۱۸۵۵) مطالعه شد.



شکل ۲-۶. نمودار توزیع نرمال

تعريف ۶-۷. متغیر تصادفی X دارای توزیع نرمال است، و به آن متغیر تصادفی نرمال اطلاق می‌شود، اگر و تنها اگر چگالی احتمال آن به صورت

$$n(x; \mu, \sigma) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}(\frac{x-\mu}{\sigma})^2} \quad -\infty < x < +\infty$$

باشد، که در آن $\mu > 0$ ، $\sigma > 0$.

نمودار توزیع نرمال که شکلی شبیه مقطع عرضی یک زنگ دارد در شکل ۶-۶ نشان داده شده است.

نمادگذاری که در اینجا به کار بردهیم صریحاً نشان می‌دهد که دو پارامتر توزیع نرمال μ و σ هستند. اما آنچه باقی می‌ماند آن است که بینیم پارامتر μ در واقع همان $E(X)$ و پارامتر σ در واقع همان $\sqrt{\text{var}(X)}$ است یا نه، که در آنها متغیری تصادفی است که دارای توزیع نرمال با این دو پارامتر است.

با این حال ابتدا نشان می‌دهیم که فرمول تعریف ۶-۶ را می‌توان به عنوان یک چگالی احتمال به کار برد. واضح است مقادیر $n(x; \mu, \sigma)$ مادامی که $\sigma > 0$ مثبت‌اند، پس باید نشان دهیم که مساحت زیر منحنی مساوی ۱ است. اگر از $-\infty$ تا $+\infty$

$$\begin{aligned} \text{انتگرال‌گیری کنیم و قرار دهیم } z = \frac{x-\mu}{\sigma}, \text{ به دست می‌آوریم} \\ \int_{-\infty}^{\infty} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}(\frac{x-\mu}{\sigma})^2} dx = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{1}{2}z^2} dz = \frac{1}{\sqrt{2\pi}} \int_0^{\infty} e^{-\frac{1}{2}z^2} dz. \end{aligned}$$

پس، چون انتگرال سمت راست مساوی $\frac{\Gamma(\frac{1}{2})}{\sqrt{2}} = \frac{\sqrt{\pi}}{\sqrt{2}}$ است، نتیجه می‌شود که مساحت کل زیر منحنی، برابر $1 = \frac{\sqrt{\pi}}{\sqrt{2\pi}} \cdot \frac{\sqrt{\pi}}{\sqrt{2}}$ است. اینک نشان می‌دهیم که

قضیه ۶-۹.تابع مولد گشتاورهای توزیع نرمال عبارت است از

$$M_X(t) = e^{\mu t + \frac{1}{2}\sigma^2 t^2}.$$

برهان. بنابر تعریف

$$M_X(t) = \int_{-\infty}^{+\infty} e^{xt} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} dx \\ = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{-\frac{1}{2}\left[\frac{(x-\mu)^2}{\sigma^2} + xt\right]} dx,$$

و اگر عبارت داخل کروشه را به صورت مریع کامل درآوریم، یعنی اتحاد $-xt\sigma^2 + (x-\mu)^2 = [x - (\mu + t\sigma)]^2 - 2\mu t\sigma^2 - t^2\sigma^2$

را به کار ببریم، به دست می‌آوریم

$$M_X(t) = e^{\mu t + \frac{1}{2}t^2\sigma^2} \left\{ \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{-\frac{1}{2}\left[\frac{(x-\mu-t\sigma)^2}{\sigma^2}\right]} dx \right\}.$$

چون مقدار داخل دو ابرو، انتگرال از $-\infty$ تا $+\infty$ یک چگالی نرمال با پارامترهای $\mu + t\sigma$ و σ و لذا مساوی ۱ است، نتیجه می‌شود که $M_X(t) = e^{\mu t + \frac{1}{2}t^2\sigma^2}$. حال آمده‌ی تحقیق درستی آنیم که در تعریف ۶-۷، پارامترهای μ و σ ، در واقع میانگین و انحراف معیار توزیع نرمال‌اند. اگر دوبار از $M_X(t)$ نسبت به t مشتق بگیریم، به دست می‌آوریم

$$M'_X(t) = (\mu + \sigma^2 t) M_X(t)$$

و

$$M''_X(t) = [(\mu + \sigma^2 t)^2 + \sigma^2] M_X(t)$$

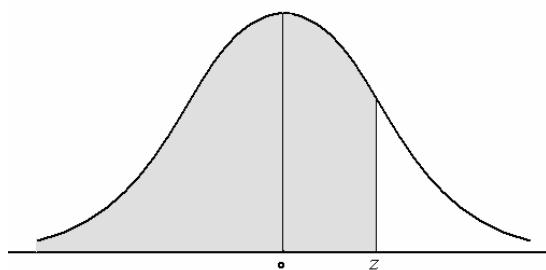
به قسمی که $E(X) = \mu$ و $M'_X(0) = \mu + \sigma^2$ و $M''_X(0) = (\mu + \sigma^2)^2 + \sigma^2$.

چون توزیع نرمال نقشی پایه‌ای در آمار دارد و از چگالی آن نمی‌توان مستقیماً انتگرال‌گیری کرد، مساحت‌هایش برای حالتی خاص که در آن $\mu = 0$ و $\sigma = 1$ ، جدول-بندی شده‌اند.

تعريف ۸-۶ توزيع نرمال با $\mu=0$ و $\sigma=1$ ، توزيع نرمال استاندارد نامیده می‌شود.

درایه‌های جدول ۳ که مساحت هاشورخوردهی شکل ۳-۶ نمایش داده می‌شوند،

$$\text{مقادیر } \int_0^z \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2} dx \text{ هستند.}$$



شکل ۳.۶. مساحت‌های جدول‌بندی شدهی زیرمنحنی توزيع نرمال استاندارد

مثال ۱۰-۴. پیدا کنید احتمال آنکه متغیری تصادفی با توزيع نرمال استاندارد مقداری

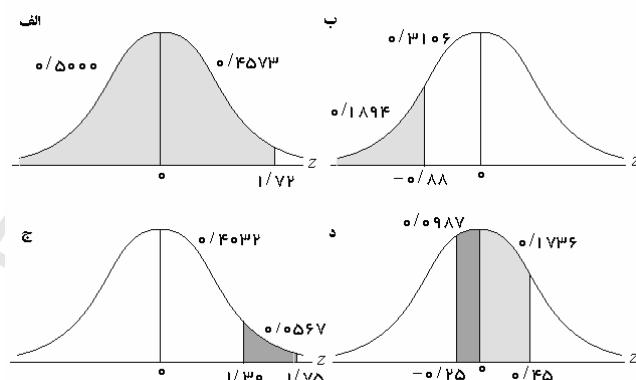
الف) کمتر از ۱/۷۲،

ب) کمتر از -۰/۸۸،

ج) بین ۰/۳۰ و ۱/۷۵،

د) بین ۰/۲۵ و ۰/۴۵،

اختیار کند.



شکل ۱۰-۶. نمودارهای مثال ۱۰-۶

حل. الف) در جدول ۳، درایه‌ی متناظر با $z=1/72$ را پیدا می‌کنیم، جواب برابر با 0.9573% است.

ب) مشابه قسمت الف، درایه‌ی متناظر با $z=-0.88$ را در جدول ۳ برابر با 0.1894% است.

ج) در جدول ۳ درایه‌های متناظر با $z=1/75$ و $z=1/30$ را پیدا و دوّمی را از اوّلی کم می‌کنیم (شکل ۶-۴ را ببینید)، و به دست می‌آوریم 0.0567% .

د) در جدول ۳ درایه‌های متناظر با $z=-0.25$ و $z=0.45$ را پیدا و آنها را از هم کم می‌کنیم (شکل ۶-۴ را ببینید)، و به دست می‌آوریم 0.2723% .

گاهی نیاز داریم مقداری از z را بباییم که متناظر با احتمالی مشخص است که بین دو مقدار فهرست شده در جدول ۳ قرار می‌گیرند. در این موارد، برای راحتی، همیشه مقداری از z را انتخاب می‌کنیم که احتمال متناظر با آن در جدول، نزدیکترین مقدار به احتمال مشخص باشد. اما اگر احتمال داده شده وسط دو احتمال متولی در جدول باشد، مقدار z را انتخاب می‌کنیم که وسط مقادیر متناظر z بیفتند.

مثال ۱۱-۶. با رجوع به جدول ۳، مقادیر z متناظر با درایه‌های

الف) 0.8512% ،

ب) 0.7533% ،

را ببایید.

حل. الف) چون 0.8512% بین 0.8508% و 0.8531% که متناظر با $z=1/04$ و $z=1/05$ هستند، می‌افتد، و چون 0.8512% به 0.8508% نزدیکتر است تا به 0.8531% ، پس مقدار $z=1/04$ را انتخاب می‌کنیم.

ب) چون 0.7533% وسط دو مقدار 0.7517% و 0.7549% است که با $z=0.68$ و $z=0.69$ متناظرند، پس مقدار $z=0.685$ را انتخاب می‌کنیم.

با وجود اینکه جدول ۳ در آخر کتاب تهیه شده است اما همواره نمی‌تواند همه‌ی مسائل را با دقت بالا حل کند. مشابهی سایر توزیع‌ها، مینیتب قادر است مقدار تابع توزیع، و چندک‌های توزیع نرمال را محاسبه کند.

```

MTB > CDF 1.72 k9;
SUBC> Normal 0.0 1.0.
MTB > InvCDF 0.8512 k10;
SUBC> Normal 0.0 1.0.
MTB > print k9 k10
Data Display
K9    0.957284
K10   1.04159

```

برای تعیین احتمال‌های مربوط به متغیرهای تصادفی که توزیعی نرمال غیر از توزیع نرمال استاندارد دارند از قضیه‌ی زیر استفاده می‌کنیم.

قضیه‌ی ۶-۱۰. اگر X دارای توزیع نرمال با میانگین μ و انحراف معیار σ باشد، آنگاه

$$Z = \frac{X - \mu}{\sigma}$$

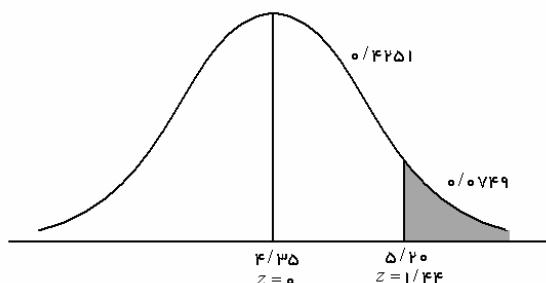
توزیع نرمال استاندارد دارد.

برهان. چون بستگی بین مقادیر X و Z خطی است، وقتی X مقداری بین x_l و x_r اختیار می‌کند، Z باید مقداری بین $\frac{x_r - \mu}{\sigma}$ و $\frac{x_l - \mu}{\sigma} = z_r$ را اختیار نماید. بنابراین می‌توانیم بنویسیم

$$P(x_l < X < x_r) = \int_{x_l}^{x_r} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx = \int_{z_l}^{z_r} \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} dz = P(z_l < Z < z_r)$$

که در آن دیده می‌شود Z متغیری تصادفی است که توزیع نرمال استاندارد دارد. بنابراین در ارتباط با هر متغیر تصادفی که توزیع نرمال دارد، برای استفاده از جدول ۳ به تعویض مقیاس $\frac{x-\mu}{\sigma} = z$ دست می‌زنیم.

مثال ۶-۱۲. فرض کنید میزان تشعشع فضایی که فردی در یک پرواز درعرض آن قرارگیرد، متغیری تصادفی است که توزیع نرمال با میانگین $14/35$ میلی‌رم و انحراف معیار $5/59$ میلی‌رم دارد. احتمال اینکه شخصی در چنین پروازی درعرض میزانی بیش از $5/20$ میلی‌رم تشعشع فضایی قرارگیرد چقدر است؟ حل. در جدول ۳، درایه‌ی متناظر با $1/144 = 5/20 - 4/35 = 0/59$ را پیدا می‌کنیم و آن را از ۱ کم می‌کنیم (شکل ۶-۵ را ببینید)، و به دست می‌آوریم $0/0749 = 0/9251 = 1 - 0/0$.



شکل ۶-۵. نمودار مثال ۱۲-۶

با مینیتب محاسبه‌ی فوق را مرور می‌کنیم.

```
MTB > CDF 5.2 k11;
SUBC> Normal 4.35 0.59.
MTB > let k12=1-k11
MTB > print k12
Data Display
K12      0.0748379
```

تمرین

۳۵-۶. نشان دهید که توزیع نرمال دارای

الف) ماکسیممی نسبی در $x = \mu$ است،

ب) نقطه‌ی عطفی در $x = \mu - \sigma$ و $x = \mu + \sigma$ است.

۳۶-۶. اگر X متغیری تصادفی باشد که دارای توزیع نرمال با میانگین μ و انحراف معیار σ است، با استفاده از قسمت سوم قضیه‌ی ۶-۴ و قضیه‌ی ۶-۶، نشان دهید که

تابع مولد گشتاورهای $Z = \frac{X - \mu}{\sigma}$ ، تابع مولد گشتاورهای توزیع نرمال استاندارد است.

۳۷-۶. اگر Z متغیری تصادفی با توزیع نرمال استاندارد باشد، پیدا کنید احتمال آنکه

این متغیر مقداری

الف) بزرگتر از $1/14$ ،

ب) بزرگتر از $-0/36$ ،

ج) بین $-0/46$ و $-0/09$ ،

د) بین $-0/85$ و $1/12$.

اختیار کند.

۶-۳۸. اگر Z متغیری تصادفی با توزیع نرمال باشد، مطلوب است

$$\text{الف) } P(Z < 1/33)$$

$$\text{ب) } P(Z \leq -0.79)$$

$$\text{ج) } P(0.55 < Z < 1/22)$$

$$\text{د) } P(-1/90 \leq Z \leq 0/44)$$

۶-۳۹. مقدار z را بیابید اگر مساحت زیر منحنی توزیع نرمال استاندارد

$$\text{الف) بین ۰ و } z, \text{ برابر } 0.4726 \text{ باشد،}$$

$$\text{ب) در چپ } z, \text{ برابر } 0.9868 \text{ باشد،}$$

$$\text{ج) در راست } z, \text{ برابر } 0.1314 \text{ باشد،}$$

$$\text{د) بین } -z \text{ و } z, \text{ برابر } 0.8502 \text{ باشد.}$$

۶-۴۰. اگر Z متغیری تصادفی با توزیع نرمال استاندارد باشد، مطلوب است مقادیر z_1, z_2, z_μ و z_σ به قسمی که

$$\text{الف) } P(0 < Z < z_1) = 0.4306$$

$$\text{ب) } P(Z \geq z_2) = 0.7704$$

$$\text{ج) } P(Z \geq z_\mu) = 0.2912$$

$$\text{د) } P(-z_\sigma \leq Z < z_\mu) = 0.9700$$

۶-۴۱. اگر X متغیری تصادفی با توزیع نرمال باشد، احتمال اینکه مقداری

الف) به فاصله‌ی یک انحراف معیار از میانگین،

ب) به فاصله‌ی دو انحراف معیار از میانگین،

ج) به فاصله‌ی سه انحراف معیار از میانگین،

د) به فاصله‌ی چهار انحراف معیار از میانگین،

اختیار کند چقدر است؟

۶-۴۲. اگر z_α به صورت $\int_{z_\alpha}^{+\infty} n(z; 0, 1) dz = \alpha$ مقدار آن را به ازای

$$\text{الف) } \alpha = 0.05$$

$$\text{ب) } \alpha = 0.025$$

$$\text{ج) } \alpha = 0.01$$

د) $\alpha = 0/005$

بیابید.

۶-۴۳. فرض کنید در طول حالت خلسه، کاهش مصرف اکسیژن یک فرد، متغیری تصادفی باشد که توزیع نرمال با میانگین $\mu = ۳۷/۶$ سانتی‌متر مکعب در دقیقه و $\sigma = ۴/۶$ سانتی‌متر مکعب در دقیقه دارد. پیدا کنید احتمال آنکه در طول مدت دوره‌ی

حالت خلسه، مصرف اکسیژن به

الف) حداقل $۴۴/۵$ سانتی‌متر مکعب در دقیقه،

ب) حداکثر $۳۵/۰$ سانتی‌متر مکعب در دقیقه،

ج) مقداری از $۳۵/۰$ تا $۴۰/۵$ سانتی‌متر مکعب در دقیقه،

کاهش پیدا کند.

۶-۴۴. در کار عکاسی، زمان ظهر عکس را می‌توان متغیری تصادفی گرفت که دارای توزیع نرمال با $\mu = ۱۵/۰$ ثانیه و $\sigma = ۰/۴۸$ ثانیه است. احتمال آن را بیابید که زمان ظهر عکس

الف) حداقل $۱۶/۰۰$ ثانیه،

ب) حداکثر $۱۴/۲۰$ ثانیه،

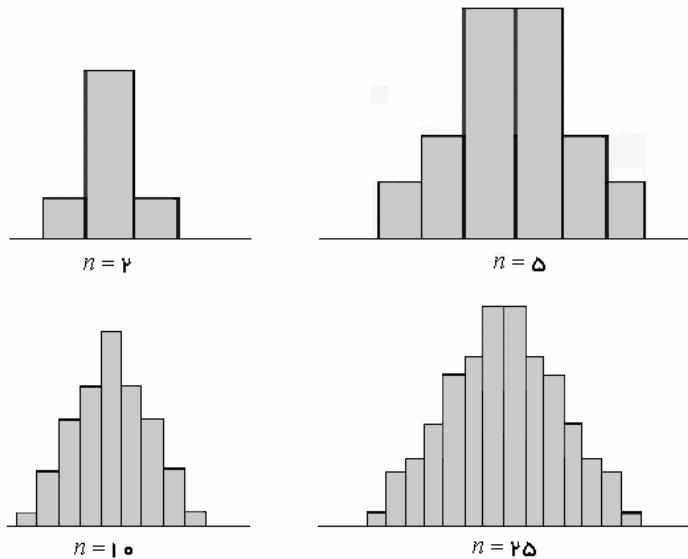
ج) زمانی که بین $۱۵/۰۰$ تا $۱۵/۸۰$ ثانیه،

باشد.

۶-۴۵. متغیری تصادفی دارای توزیع نرمال با $\sigma = ۵$ است. اگر احتمال اینکه متغیر تصادفی مقداری کمتر از $۸۲/۵$ اختیار کند برابر با $۸۲۱۲/۰$ باشد، احتمال اینکه مقداری بزرگتر از $۵۸/۳$ اختیار کند چقدر است؟

۶-۹ تقریب نرمال برای توزیع دوجمله‌ای

توزیع نرمال گاهی به عنوان توزیع پیوسته‌ای معرفی می‌شود که برای توزیع دوجمله‌ای، وقتی n ، تعداد آزمایش‌ها خیلی بزرگ است، و θ ، احتمال موفقیت در یک آزمایش نزدیک $\frac{1}{\mu}$ است تقریب خوبی فراهم می‌کند. شکل ۶-۶ بافت‌نمایه‌ای توزیع‌های دوجمله‌ای با $\theta = \frac{1}{\mu}$ و $n = ۲, ۵, ۱۰, ۲۵$ را نشان می‌دهد، و می‌توان دید که افزایش n ، این توزیع‌ها به الگوی زنگ شکل متقارن توزیع نرمال میل می‌کنند.



شکل ۶-۶. توزیع‌های دوجمله‌ای با $\theta = \frac{1}{3}$

برای فراهم کردن یک پایه‌ی نظری برای این بحث، ابتدا قضیه‌ی زیر را مطرح می‌کنیم.

قضیه‌ی ۶-۱۱. اگر X متغیری تصادفی با توزیع دوجمله‌ای با پارامترهای n و θ باشد آنگاه تابع مولد گشتاورهای

$$Z = \frac{X - n\theta}{\sqrt{n\theta(1-\theta)}}$$

وقتی $n \rightarrow +\infty$ ، به تابع مولد گشتاورهای توزیع نرمال استاندارد میل می‌کند.

برهان. با استفاده از قضیه‌های ۵-۶ و ۱۰-۵ می‌توانیم بنویسیم

$$M_Z(t) = M_{\frac{X-\mu}{\sigma}}(t) = e^{\frac{-\mu t}{\sigma} [1 + \theta(e^{\frac{t}{\sigma}} - 1)]^n}$$

که در آن $\sigma = \sqrt{n\theta(1-\theta)}$ و $\mu = n\theta$. حال اگر لگاریتم بگیریم و به جای $e^{\frac{t}{\sigma}}$ از سری ماکلورن آن استفاده کنیم، به دست می‌آوریم

$$\begin{aligned}\ln M_{\frac{x-\mu}{\sigma}}(t) &= -\frac{\mu t}{\sigma} + n \ln \left[1 + \theta(e^{\frac{t}{\sigma}} - 1) \right] \\ &= -\frac{\mu t}{\sigma} + n \ln \left[1 + \theta \left\{ \left(\frac{t}{\sigma} \right)^1 + \frac{1}{2} \left(\frac{t}{\sigma} \right)^2 + \frac{1}{6} \left(\frac{t}{\sigma} \right)^3 + \dots \right\} \right]\end{aligned}$$

و با استفاده از سری نامتناهی $|x| < 1$, $\ln(1+x) = x - \frac{1}{2}x^2 + \frac{1}{3}x^3 + \dots$, که به ازای $1 < |x| < 1$ همگرا است، برای بسط این لگاریتم، نتیجه می‌شود که

$$\begin{aligned}\ln M_{\frac{X-\mu}{\sigma}}(t) &= -\frac{\mu t}{\sigma} + n \theta \left[\left(\frac{t}{\sigma} \right)^1 + \frac{1}{2} \left(\frac{t}{\sigma} \right)^2 + \frac{1}{6} \left(\frac{t}{\sigma} \right)^3 + \dots \right] \\ &\quad - \frac{n \theta^2}{2} \left[\left(\frac{t}{\sigma} \right)^1 + \frac{1}{2} \left(\frac{t}{\sigma} \right)^2 + \frac{1}{6} \left(\frac{t}{\sigma} \right)^3 + \dots \right]^2 \\ &\quad + \frac{n \theta^3}{3!} \left[\left(\frac{t}{\sigma} \right)^1 + \frac{1}{2} \left(\frac{t}{\sigma} \right)^2 + \frac{1}{6} \left(\frac{t}{\sigma} \right)^3 + \dots \right]^3 \\ &\quad + \dots.\end{aligned}$$

با گردآوری توان‌های همانند t , به دست می‌آوریم

$$\begin{aligned}\ln M_{\frac{X-\mu}{\sigma}}(t) &= \left(-\frac{\mu}{\sigma} + \frac{n\theta}{\sigma} \right) t + \left(\frac{n\theta^2}{2\sigma^2} - \frac{n\theta^3}{3\sigma^3} \right) t^2 + \left(\frac{n\theta^3}{3\sigma^3} - \frac{n\theta^4}{4\sigma^4} + \frac{n\theta^5}{5\sigma^5} \right) t^3 + \dots \\ &= \frac{1}{\sigma^2} \left(\frac{n\theta - n\theta^2}{2\sigma^2} \right) t^2 + \frac{n}{\sigma^3} \left(\frac{\theta - 2\theta^2 + 3\theta^3}{6\sigma^3} \right) t^3 + \dots\end{aligned}$$

زیرا $n\theta = \mu$. در این صورت اگر قرار دهیم $\sigma = \sqrt{n\theta(1-\theta)}$, پیدا می‌کنیم که

$$\ln M_{\frac{X-\mu}{\sigma}}(t) = \frac{1}{2} t^2 + \frac{n}{\sigma^3} \left(\frac{\theta + 2\theta^2 + 2\theta^3}{6} \right) t^3 + \dots$$

که در آن به ازای $r > 2$, ضریب t^r مضرب ثابتی از $\frac{n}{\sigma^r}$ است که وقتی $n \rightarrow +\infty$, به صفر می‌گراید. نتیجه می‌شود که

$$\lim_{n \rightarrow +\infty} \ln M_{\frac{X-\mu}{\sigma}}(t) = \frac{1}{2} t^2,$$

و چون حد یک لگاریتم برابر لگاریتم حد است (به شرط آنکه هر دو حد وجود داشته

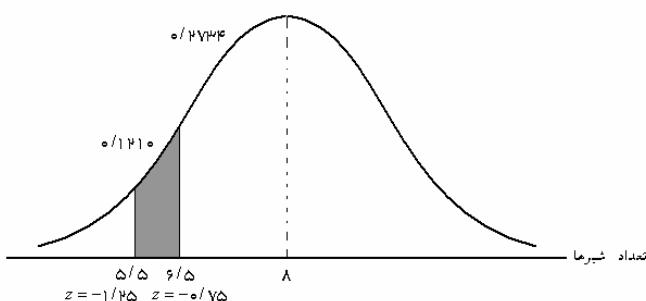
باشند)، نتیجه می‌گیریم که

$$\lim_{n \rightarrow +\infty} \ln M_{\frac{X-\mu}{\sigma}}(t) = e^{\frac{1-t^2}{2}}$$

که تابع مولد گشتاورهای قضیه‌ی ۶-۸ با $\mu=0$ و $\sigma=1$ است. این برهان قضیه‌ی ۶-۱۰ را تکمیل می‌کند، اما آیا نشان داده‌ایم که وقتی $n \rightarrow +\infty$ توزیع Z ، متغیر تصادفی دوچمله‌ای استاندارد شده، به توزیع نرمال استاندارد می‌گراید؟ نه کاملاً. برای این منظور باید به دو قضیه‌ای که در اینجا بدون اثبات بیان می‌شوند مراجعه کنیم:

۱. بین تابع‌های مولد گشتاورها و توزیع‌های احتمال (چگالی‌ها) وقتی تابع‌های مولد وجود دارند تناظری یک به یک موجود است.
۲. اگر تابع مولد گشتاورهای یک متغیر تصادفی به تابع مولد گشتاورهای متغیر تصادفی دیگری میل کند، آنگاه توزیع (چگالی) متغیر تصادفی اوّل تحت همان شرایط حدی به توزیع (چگالی) متغیر تصادفی دوم میل می‌کند.

اگر دقیق‌تر صحبت کنیم، این قضایا وقتی صادق‌اند که $n \rightarrow +\infty$ ، اما توزیع نرمال اغلب حتی وقتی n نسبتاً کوچک است برای تقریب احتمال‌های دوچمله‌ای به کار می‌رود. قانون سرانگشتی خوبی آن است که این تقریب را وقتی $n\theta$ و $n(1-\theta)$ هر دو بزرگ‌تر از ۵ هستند به کار ببریم.



شکل ۷-۶. نمودار مثال ۱۳-۶

مثال ۶-۱۳. با استفاده از تقریب نرمال برای توزیع دوجمله‌ای احتمال آن را که در ۱۶ پرتاب یک سکه‌ی همگن، ۶ شیر و ۱۰ خط به دست می‌آوریم تعیین کنید.

حل. برای یافتن این تقریب باید از تصحیح پیوستگی استفاده کنیم که بنابر آن، هر عدد صحیح نامنفی k با بازه‌ی $k - \frac{1}{2} \leq k + \frac{1}{2} \leq n$ نشان داده می‌شود. با رجوع به شکل ۶-۶، باید مساحت زیرمنحنی را بین $\mu - \frac{\sigma}{\sqrt{n}}$ و $\mu + \frac{\sigma}{\sqrt{n}}$ تعیین کنیم، و چون $\mu = 16 \times \frac{1}{2} = 8$ و

$$\sigma = \sqrt{16 \times \frac{1}{2}} = 2$$

$$z = \frac{\mu - 8}{\sigma} = -1/25 \quad \text{و} \quad z = \frac{8 - 8}{\sigma} = 0/75$$

تعیین کنیم. چون در جدول ۳، درایه‌های متناظر با $z = 1/25$ و $z = 0/75$ به ترتیب 0.8944 و 0.7734 هستند، درمی‌یابیم که تقریب نرمال برای احتمال داشتن «۶ شیر و ۱۰ خط» برابر $0.1210 = 0.7734 - 0.8944$ است. چون مقدار متناظر در جدول ۲ برابر با 0.1222 است، درمی‌یابیم که خطای تقریب برابر با 0.0012 است، و قدر مطلق خطای درصد برابر با $0.0012 \times 100 = 0.1222$ درصد است.

تمرین

۶-۴. بررسی کنید که برای هر یک از موارد زیر می‌توان، طبق قاعده‌ی سرانگشتی، تقریب نرمال را برای توزیع دوجمله‌ای به کار برد یا نه؟

الف) $n = 16$ و $\theta = 0/20$ ،

ب) $n = 65$ و $\theta = 0/10$ ،

ج) $n = 120$ و $\theta = 0/98$ ،

۶-۵. فرض کنید می‌خواهیم تقریب نرمال برای توزیع دوجمله‌ای را به کاربرده، مقدار $b(1; 150, 0/05)$ را تعیین کنیم.

الف) مبتنی بر قاعده‌ی سرانگشتی، آیا توجیهی برای استفاده از این تقریب داریم؟

ب) تقریب را انجام دهید و آن را تا چهار رقم اعشار گرد کنید.

ج) اگر خروجی کامپیوتر نشان دهد که $b(1; 150, 0/05) = 0.0036$ که تا چهار رقم گردشده است، درصد خطای تقریب حاصل در قسمت (ب) چقدر است؟

۶-۶. برای تعیین (تا چهار رقم دهدی) احتمال به دست آوردن ۷ شیر و ۷ خط را در ۱۴ بار پرتاب یک سکه‌ی همگن، تقریب نرمال برای توزیع دوجمله‌ای را به

کاربرید. همچنین با رجوع به جدول ۲، خطای این تقریب را بیابید.

۶-۴۹. اگر ۲۳درصد از تمام بیمارانی که فشار خون بالا دارند دچار عوارض جانبی ناشی از نوعی دارو باشد، برای پیدا کردن احتمال آنکه بین ۱۲۰ نفر بیمار با فشار خون بالا که با این دارو معالجه می‌شوند بیش از ۳۲نفر دچار عوارض جانبی شوند، تقریب نرمال را به کار برد.

۶-۵۰. برای تشریح قانون اعداد بزرگ، تقریب نرمال را برای توزیع دوجمله‌ای به کار برد و احتمال آن را بیابید که وقتی سکه‌ای همگن

الف) ۱۰۰ بار،

ب) ۱۰۰۰ بار،

ج) ۱۰۰۰۰ بار،

پرتاپ شود نسبت شیرها عددی بین ۵۱٪ و ۴۹٪ باشد.

۶-۱۰ توزیع نرمال دو متغیره

بین چگالی‌های چند متغیره، توزیعی که از اهمیت خاصی برخوردار است، توزیع نرمال چند متغیره است که تعییمی از توزیع نرمال یک متغیره است. چون بهتر است که این توزیع برای حالت چند متغیره با نمادگذاری ماتریسی نمایش داده شود، ما در اینجا فقط حالت دو متغیره را ارائه می‌دهیم.

تعریف ۶-۹. جفت متغیرهای تصادفی X و Y دارای توزیع نرمال دو متغیره‌اند و به آنها متغیرهای تصادفی که توامان به صورت نرمال توزیع شده‌اند اطلاق می‌شوند، اگر و تنها اگر چگالی احتمال توأم‌شان برای $x < +\infty$ و $y < +\infty$ به صورت زیر باشد

$$f(x,y) = \frac{e^{-\frac{1}{2(1-\rho^2)} \left[\left(\frac{x-\mu_1}{\sigma_1}\right)^2 - 2\rho \left(\frac{x-\mu_1}{\sigma_1}\right) \left(\frac{y-\mu_2}{\sigma_2}\right) + \left(\frac{y-\mu_2}{\sigma_2}\right)^2 \right]}}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}},$$

که در آن $\mu_1 > 0$ ، $\sigma_1 > 0$ و $\mu_2 > 0$ و $\sigma_2 > 0$ و $-1 < \rho < 1$.

نشان می‌دهیم که μ_1 ، μ_2 ، σ_1 ، σ_2 و ρ به ترتیب، میانگین‌ها و انحراف معیارهای

متغیرهای تصادفی X و Y هستند. اگر برای بهدست آوردن چگالی حاشیه‌ای X ، از $f(x,y)$ نسبت به y از $-\infty$ تا $+\infty$ انتگرال بگیریم، بهدست می‌آوریم

$$g(x) = \frac{e^{-\frac{1}{r(1-\rho^r)} \left(\frac{x-\mu_1}{\sigma_1}\right)^r}}{\sqrt[4]{2\pi\sigma_1\sigma_r\sqrt{1-\rho^r}}} \int_{-\infty}^{\infty} e^{-\frac{1}{r(1-\rho^r)} \left[\left(\frac{y-\mu_r}{\sigma_r}\right)^r - r\rho \left(\frac{x-\mu_1}{\sigma_1}\right)\left(\frac{y-\mu_r}{\sigma_r}\right)\right]} dy.$$

اگر برای ساده‌کردن نمادگذاری موقتاً قراردهیم $u = \frac{x-\mu_1}{\sigma_1}$ و با قراردادن $v = \frac{y-\mu_r}{\sigma_r}$ متغیر انتگرال‌گیری را تعویض کنیم، بهدست می‌آوریم

$$g(x) = \frac{e^{-\frac{1}{r(1-\rho^r)} u^r}}{\sqrt[4]{2\pi\sigma_1\sqrt{1-\rho^r}}} \int_{-\infty}^{\infty} e^{-\frac{1}{r(1-\rho^r)} (v^r - r\rho u v)} dv,$$

و با قراردادن $v^r - r\rho u v = (v - pu)^r - p^r u^r$ برای اینکه مربع کاملی حاصل شود و با جمع‌آوری جملات، نتیجه می‌شود که

$$g(x) = \frac{e^{-\frac{1}{r} u^r}}{\sigma_1 \sqrt[4]{2\pi}} \left\{ \frac{1}{\sqrt[4]{2\pi\sqrt{1-\rho^r}}} \int_{-\infty}^{\infty} e^{-\frac{1}{r} \left(\frac{v-pu}{\sqrt{1-\rho^r}}\right)^r} dv \right\}.$$

سرانجام، با تشخیص اینکه کمیت داخل پرانتز همان انتگرال چگالی نرمال از $-\infty$ تا $+\infty$ ، و بنابراین مساوی ۱ است، به ازای $x < +\infty$ بهدست می‌آوریم

$$g(x) = \frac{e^{-\frac{1}{r} u^r}}{\sigma_1 \sqrt[4]{2\pi}} = \frac{1}{\sigma_1 \sqrt[4]{2\pi}} e^{-\frac{1}{r} \left(\frac{x-\mu_1}{\sigma_1}\right)^r}.$$

با بررسی نتیجه می‌شود که چگالی حاشیه‌ای X ، توزیعی نرمال با میانگین μ_1 و انحراف معیار σ_1 است، و بنابر تقارن، چگالی حاشیه‌ای Y ، توزیعی نرمال با میانگین μ_r و انحراف معیار σ_r است.

توزیع نرمال دومتغیره چندین ویژگی مهم دارد که برخی جنبه‌ی آماری و برخی جنبه‌ی ریاضی محض دارند. بین ویژگی‌های آماری، ویژگی زیر را داریم که اثبات آن

در تمرین ۶-۵۱ از خواننده خواسته شده است.

قضیه‌ی ۶-۱۲. دو متغیر تصادفی که دارای توزیع نرمال دو متغیره‌اند مستقل‌اند اگر و تنها اگر $\rho = 0$.
اگر $\rho = 0$ ، متغیرهای تصادفی را ناهمبسته گویند.

تمرین

۶-۵۱. برای اثبات قضیه‌ی ۶-۱۱، نشان دهید که اگر X و Y توزیع نرمال دو متغیره داشته باشند، آنگاه

(الف) استقلال آنها نتیجه می‌دهد که $\rho = 0$ ، (ب) $\rho = 0$ نتیجه می‌دهد که آنها مستقل‌اند.

۶-۵۲. اگر نمای e در چگالی نرمال دو متغیره به صورت

$$\frac{-1}{10^2}[(x+2)^3 - 2/8(x+2)(y-1) + 4(y-1)^3]$$

باشد، پیدا کنید $\mu_1, \mu_2, \sigma_1, \sigma_2$ و ρ .

۶-۵۳. اگر X و Y دارای توزیع نرمال دو متغیره داشته باشند، و $U = X + Y$ و $V = X - Y$ ، برای ضریب همبستگی U و V عبارتی بیابید.

۶-۵۴. مرکز یک هدف، به عنوان مبدأ یک دستگاه مختصات قائم اختیار می‌شود و نسبت به این دستگاه، نقطه‌ی اصابت یک گلوله دارای مختصات X و Y است. اگر X و Y دارای چگالی نرمال دو متغیره با $\mu_1 = 0, \mu_2 = 0, \sigma_1 = 120, \sigma_2 = 120$ پا باشد، پیدا کنید احتمال آنکه نقطه‌ی اصابت

(الف) درون مربعی به ضلع ۱۸۰ پا باشد که مرکزش در مبدأ بوده و اضلاع‌اش موازی محورهای مختصات‌اند،

(ب) درون دایره‌ای به شعاع ۷۵ پا و به مرکز مبدأ باشد.

۶-۵۵. اگر X و Y توزیع نرمال دو متغیره با $\mu_1 = 12, \mu_2 = 14, \sigma_1 = \sigma_2 = 5$ داشته باشند، پیدا کنید

(الف) احتمال به دست آوردن نقطه‌ی (x, y) در داخل دایره‌ی $x^2 + y^2 = 36$ ،
(ب) مقدار c که به ازای آن، احتمال به دست آوردن نقطه‌ی (x, y) در داخل دایره‌ی $x^2 + y^2 = c^2$ برابر ۸۰٪ باشد.

فصل هفتم

محاسبات با نرم افزار

۱-۷ مقدمه

روند تغییرات در نگارش یک نرم افزار آنقدر سریع می باشد که مجالی به آموزش از طریق منوهای نرم افزار با یک دست نوشته ثابت، را نمی دهد. علاوه بر آن قابلیت فهرست ها(منوها) همواره پاسخ گوی نیازها و حل مسائل آماری نمی باشد و همواره نوشتن برنامه را باید جزئی از آموختنی ها برای دانشجویان رشته‌ی آمار در نظر گرفت. لذا در این فصل با در نظر گرفتن این اهداف برای دانشجویان درس آمار و احتمال تهیه گردیده است. امید است با به کار گیری و ادامه آن در دروس تخصصی بعدی در تفهیم مفاهیم آماری نقش سازنده‌ای داشته باشد.

۲-۷ ورود و خروج اطلاعات در مینیتب

در این بخش آنچه برای وارد کردن و یا ویرایش داده‌ها نیاز است را ارائه داده‌ایم. هرگاه وارد نرم افزار مینیتب می شوید با دو محیط رو برو هستید:

۱. محیط DATA

۲. محیط SESSION

دستورات را فقط در محیط دوم می‌توان نوشت و تغییرات را در محیط اول دید. نکته دیگر اینکه دستورات اصلی را باید جلوی MTB> تایپ کرد و این در حالی است که برای زیر برنامه‌ها باید نماد SUBC> ظاهر گردد.

دستور Read

```
READ C...C
FILE 'filename'
FORMAT (specification)
NOBS=K
```

در دستور فوق، داده‌ها سطر به سطر خوانده می‌شود. اگر نام فایلی مشخص باشد داده‌ها را از آن پرونده می‌گیرد. در غیر این صورت داده‌ها را زیر فرمان Read می‌نویسیم. حالت ورودی آزاد داده‌ها طی مثال زیر نشان داده شده است. در آخر دستور END را برای نشان دادن پایان داده‌ها به کار می‌بریم.

مثال ۱-۷:

```
MTB> READ C2 C3
DATA> 2 4
DATA> 3.5 27
DATA> 1 12
DATA>END
```

نرم افزار داده‌های ۲، ۳/۵، و ۱ را در ستون دوم و داده‌های ۴، ۲۷، و ۱۲ را در ستون سوم قرار می‌دهد.

نکات:

- ۱- خطوط داده‌ها می‌توانند شامل حروف و کلمات باشد اما نمی‌توانند با نامی که دستوری هستند در نرم افزار شروع می‌شود. حروف خوانده نشده حذف می‌شوند. اگر داده‌ای حرفی (الفبایی) داشته باشیم باید از فرمت فرترن استفاده کنیم.
- ۲- داده می‌تواند به صورت اعشاری یا بدون اعشار نوشته شود. همچنین می‌توانیم به وسیله‌ی نماد علمی (نماد نمایی) داده‌ها را وارد کنیم مثلاً

- 1.234E2 1.234E+2 123.4
- ۳- خطوط داده‌ها میتواند شامل یک & در آخر خط به معنای ادامه داشتن داده‌ها، باشد.
 - ۴- از نماد “*” برای نمایش داده‌ی گم شده استفاده کنید.
 - ۵- نماد (input width) IW در خطوط داده‌ها با تعریف کردن رقم‌های داده‌ها به کار می‌رود. برای مثال اگر بنویسیم IW72 داده‌ای که تعداد رقم‌هایش متجاوز از ۷۲ باشد، رقم‌های بعدی را نمی‌خواند.
 - ۶- دستور READ برای وارد کردن یک ماتریس نیز به کار می‌رود.
 - ۷- اگر در یک خط END نوشته شود یعنی داده‌ها تمام شده است.
 - ۸- ثابت‌های ذخیره شده از طریق دستور READ خوانده نمی‌شود بلکه باید دستور SET کار برد.

زیربرنامه‌ها در دستور READ

FORMAT (specification)

داده‌ها در ستون‌ها با آرایشی مشخص وارد می‌شوند.

| | |
|------|------------------------------|
| FW | برای خواندن اعداد |
| FW.D | برای خواندن اعداد با اعشار |
| AW | داده حروفی - الفبایی |
| X | برای تعریف فضای خالی - فاصله |
| Tn | طول کل فضای داده‌ها در یک خط |
| N | تکرار عامل |
| (| پرانتز باز |
|) | پرانتز بسته |
| , | فرمت‌ها را جدا می‌کند |
| / | برای رفتن به خط جدید داده‌ها |

NOBS = K

هرگاه داده‌ها از یک فایل خوانده شود با دستور فوق فقط K تای آنرا وارد نرم‌افزار می‌کند.

مثال ۲-۷. اعداد زیر با فرمت F3.1 وارد کامپیوتر کنید و سپس چگونگی خواندن آن توسط نرم‌افزار را مورد بررسی قرار دهید.

۱۲۱۲۳۴۳۴۵۵۶۶۷۸۸۹

: حل

```
MTB > READ C1;
SUBC> FORMAT(F3.1).
DATA> 1212343455667889
DATA> END
      1 rows read.
MTB > PRINT C1
Data Display
C1
12.1
```

دستور Set

```
SET   C
FORMAT (specification)
NOBS=K
```

وارد کردن داده‌ها به کمک دستور SET را فقط برای یک ستون می‌توان به کار برد. اگر نام پرونده مشخص باشد داده‌ها را می‌توان از آن طریق خوانده، در غیر این صورت در خط زیر دستور مربوطه، داده‌ها را بنویسید.

.**مثال ۳-۷**

```
MTB > SET C7
DATA> 2      7      9
DATA> 3.5    22
DATA> END
```

در مثال فوق داده‌ها در ستون هفتم جای می‌گیرند.

کاربرد دیگر این دستور، ساختن داده‌های الگودار (PATTERNED DATA) است. به کمک دستور SET می‌توانیم داده‌هایی که از الگویی معین تبعیت می‌کنند وارد کامپیوتر نمائیم. برای مثال

| | |
|---------|------------------------------|
| 1:4 | 1, 2, 3, 4 |
| 4:1 | 4, 3, 2, 1 |
| 1:3/0.5 | 1, 1/5, 2, 2/5, 3 |
| 3(1) | 1, 1, 1 |
| 3(1:3) | 1, 2, 3, 1, 2, 3, 1, 2, 3 |
| (1:3)2 | 1, 1, 2, 2, 3, 3 |
| 2(1:3)2 | 1, 2, 2, 1, 3, 3, 1, 2, 3, 3 |

عدد جلوی پرانتز هر عدد را تکرار می‌کند، عدد قبل از پرانتز هر دنباله را تکرار می‌کند.
توجه داشته باشید از پرانتزهای تودرتو نمی‌توان استفاده کرد.
مقادیر ثابت ذخیره شده در ... k1, k2 ... را می‌توان در دستور SET به کار برد.
این ثابت‌ها به سایر دستورات نرم افزار ایجاد شده و یا ما با دستور LET آنرا تعریف می‌کنیم.

مثال ۷-۴

```
MTB > LET K2=3
MTB > LET K3=10.2
MTB > LET K1=2
MTB > SET C1
DATA> 2(K1:20) K2 K3 10
DATA> END
```

به کمک دستورات فوق داده‌های زیر را در نرم افزار وارد کرده‌ایم:

| C1 | 2.0 | 2.0 | 2.0 | 3.0 | 3.0 | 3.0 | 4.0 | 4.0 | 4.0 | 5.0 | 5.0 |
|----|-----|-----|------|------|------|------|------|------|------|------|------|
| | 5.0 | 6.0 | 6.0 | 6.0 | 7.0 | 7.0 | 7.0 | 8.0 | 8.0 | 8.0 | 9.0 |
| | 9.0 | 9.0 | 10.0 | 10.0 | 10.0 | 11.0 | 11.0 | 11.0 | 12.0 | 12.0 | 12.0 |

| | | | | | | | | | | |
|------|------|------|------|------|------|------|------|------|------|------|
| 13.0 | 13.0 | 13.0 | 14.0 | 14.0 | 14.0 | 15.0 | 15.0 | 15.0 | 16.0 | 16.0 |
| 16.0 | 17.0 | 17.0 | 17.0 | 18.0 | 18.0 | 18.0 | 19.0 | 19.0 | 19.0 | 20.0 |
| 20.0 | 20.0 | 2.0 | 2.0 | 2.0 | 3.0 | 3.0 | 3.0 | 4.0 | 4.0 | 4.0 |
| 5.0 | 5.0 | 5.0 | 6.0 | 6.0 | 6.0 | 7.0 | 7.0 | 7.0 | 8.0 | 8.0 |
| 8.0 | 9.0 | 9.0 | 9.0 | 10.0 | 10.0 | 10.0 | 11.0 | 11.0 | 11.0 | 12.0 |
| 12.0 | 12.0 | 13.0 | 13.0 | 13.0 | 14.0 | 14.0 | 14.0 | 15.0 | 15.0 | 15.0 |
| 16.0 | 16.0 | 16.0 | 17.0 | 17.0 | 17.0 | 18.0 | 18.0 | 19.0 | 19.0 | 19.0 |
| 19.0 | 20.0 | 20.0 | 20.0 | 10.2 | 10.0 | | | | | |

همانند دستور Read زیر برنامه های NOBS و FORMAT قابل اجرا می باشند.

مثال ۵-۷.

```
MTB > SET C1;
SUBC> FORMAT(A11,T16,A2).
DATA> ABCDEFGHIJK  LM
DATA> abcdefghijk  lm
DATA> end
```

مطابق فرمت تعریف شده باید طول اولین داده ۱۶ باشد. به عبارت دیگر ۱۶ فضا برای عنصر اول در هر سطر در نظر می گیرد. چهار داده حرفی در ستون C1 تعریف می شود.

در آخر دستورات INSERT READ SET از فرمان END استفاده می کنیم. در صورتی که نام فایلی مشخص نشده باشد.

Name دستور
NAME C='name' ... C='name'

به هر ستون نامی را نسبت می دهد. نام گذاری نباید از هشت مشخصه بیشتر باشد. از علائمی مانند ”#“ یا ”#“ نباید استفاده کنید. اسم ها نباید تکراری باشد. نام هر ستون را می توان به وسیله دستور Name تغییر داد. برای پاک کردن نام یک ستون از فاصله استفاده کنید برای مثال

```
MTB > name c1= ''
```

Print دستور

PRINT E... E;
FORMAT (specification)

این دستور داده‌ها را صفحه چاپ می‌کند. یک ستون یا بیشتر را می‌توانید چاپ کنید. در این دستور فرمت قابل اجرا می‌باشد. برخی نکات قابل توجه می‌باشند:

دستور فرمت برای ثابت‌ها و ماتریس‌ها به کار نمی‌رود.

اگر در دستور چاپ ماتریس‌ها، ثابت‌ها و ستون‌ها را همزمان بخواهیم چاپ کند، نرم‌افزار ابتدا ماتریس‌ها و سپس ثابت‌ها و بعد ستون‌ها را چاپ می‌کند.

اگر عدد محاسبه شده در نرم‌افزار ارقام زیادی داشته باشد در دستور چاپ از نماد نمایی استفاده می‌کند.

Write دستور

WRITE ‘filename’ E... E;
FORMAT (specification)

داده‌ها را در یک فایل قرار می‌دهد. فایلی که از طریق این دستور ایجاد می‌شود از طریق سایر نرم‌افزارها قابل استفاده است. اکثر برنامه‌ها، دارای پسوند DAT باشند تا خوانده شوند.

با دستور Read می‌توانید داده‌های ذخیره شده توسط Write را خواند. اگر یک ستون باشد با دستور Set می‌خوانید. فایلی که با دستور Write به وجود آورده‌اید می‌توانید در کامپیوتر تحت DOS ویرایش کنید ولی نمی‌توانید این فایل را به وسیله‌ی دستور SAVE به وجود آورید. هرگاه ستون‌ها دارای داده‌های برابر نباشند minitab در صورت نیاز، با قراردادن داده‌های گم شده ستون‌ها را مساوی می‌کند.

Delete دستور

DELETE rows K... K of C...C

سطرهای ستون‌های تعیین شده را حذف می‌کند.

MTB> DELETE 2 5 6 C1 C2

در عبارت فوق سطرهای دوم، پنجم و ششم از ستون‌های اوّل و دوّم حذف می‌گردد.
اگر بخواهیم تمامی محتوی یک ستون یا چند ستون را پاک کنیم از دستور
استفاده می‌کنیم.

Insert دستور

INSERT 'filename' [Between k k] E...E;
FORMAT (specification)
NOBS=K

بین سطر، ستون‌ها اطلاعات جدیدی می‌توانیم وارد کنیم. داده‌ها می‌توانند از پرونده
دیگری خوانده شوند. برای روشن شدن چگونگی عملکرد این دستور به مثال‌های زیر
توجه کنید.

اگر در بالای ستون‌های C1-C4 بخواهیم داده وارد کنیم:

MTB>INSERT 0 1 C1-C4

در بین ردیف‌های پنجم و ششم داده وارد می‌کنیم:

MTB>INSERT 5 6 C1-C4

اگر معین نکردیم کدام ردیف با دستور زیر داده‌ها را در آخر ستون‌ها وارد می‌کنند:

MTB>INSERT 2 3 C1-C4

با توجه به دستور داده شده بعد و قبل از وارد کردن داده‌های جدید ملاحظه بفرمایید.

MTB>INSERT 2 3 C1-C3

DATA>62 105 0.4

DATA>63 120 0.7

DATA>END.

| C1 | C2 | C3 | C4 | C1 | C2 | C3 | C4 |
|----|-----|-----|----|----|-----|-----|----|
| 61 | 96 | 0.5 | 14 | 61 | 96 | 0.5 | 14 |
| 65 | 115 | 0.3 | 12 | 65 | 115 | 0.3 | 12 |
| 67 | 131 | 0.8 | 13 | 62 | 105 | 0.4 | 13 |
| 64 | 125 | 0.5 | 17 | 63 | 120 | 0.7 | 17 |
| | | | | 67 | 131 | 0.8 | |
| | | | | 64 | 125 | 0.5 | |

Copy دستور

```
COPY C...C into C...C;
COPY C...C into k...k;
    USE rows k...k.
    USE rows C=k...k.
    OMIT k...k.
    OMIT C=k...k.
COPY k...k   k...k
COPY k...k   C
```

به کمک این دستور، ستون ها را می توان در ستون های دیگر یا ثابت های دیگر و بالعکس کپی کرد. زیر برنامه های دستور کپی، USE و OMIT هستند. برای مشخص کردن سطرهای خاص در کپی از زیر برنامه USE، برای تعیین اینکه کدام سطر را کپی نکند از OMIT استفاده می شود. به مثال زیر توجه کنید:

```
MTB > copy 'name' 'ht' 'wt' c12-c14;
SUBC> use 'sex'=1.
```

| | name | sex | ht | wt | C12 | C13 | C14 |
|-------------|------|-----|-----|----|--------|-----|-----|
| Joan Arnold | 2 | 66 | 135 | | Henry | 70 | 155 |
| Henry | 1 | 70 | 155 | | James | 64 | 145 |
| Marie | 2 | 64 | 125 | | Martin | 69 | 160 |
| Susan | 2 | 65 | 115 | | | | |
| Jane | 2 | 63 | 108 | | | | |
| James | 1 | 64 | 145 | | | | |
| Martin | 1 | 69 | 160 | | | | |

نکات:

- ۱- در حالت کلی می توانید هر فهرستی از مقادیر در دامنه خاصی به وسیله دو زیر برنامه فوق اجرا نمایید.

```
MTB > COPY C1-C3 C11-C13;
SUBC> USE 'ht'=0:64 70.
```

این دستور داده هایی را کپی می کند که در فاصله هی صفر تا ۶۴ یا برابر ۷۰ باشند.

- ۲- می توان کپی از سطرهای مورد نظر را تعریف کرد:

```
MTB>COPY C1-C3 C11-C13;
SUBC>USE 1  3:5.
```

سطرهای اول، سوم، چهارم و پنجم را کپی می کند.

- ۳- می تواند سطرهای خاص را در کپی کردن حذف شود.

```
MTB>COPY C1-C3 C11-C13;
SUBC>OMIT 2 6.
MTB>COPY C1-C3 C11-C13;
SUBC>OMIT C1='*'.
```

دستور **Code**

CODE (K...K) K,...,(K...K) to K for C... C put in C...C

با مثال‌های زیر، ملاحظه می‌کنید که چگونه برخی داده‌ها را به داده‌های دیگر مبدل ساخت.

داده‌های در فاصله‌ی تعریف شده را به عدد ۱۰۰ تغییر داد.

```
MTB > code (-1 -2) 100 c1 c2
MTB > print c1 c2
Data Display
Row   C1    C2
1    -0.1  -0.1
2     3.0   3.0
3     1.1   1.1
4    -1.0  100.0
5     0.0   0.0
6    -2.0  100.0
7     2.4   2.4
```

- یک فاصله را می‌تواند به یک عدد مبدل کند.

MTB>CODE (1:1.5, 2) 5 C1-C3 C11-C13

- کلیه‌ی مقادیر در (۱/۵ ۱) و مقدار ۲ را به عدد پنج مبدل می‌کند.

MTB>CODE (1:5) 10 (6:10) 20 C1-C5 C11-C15

- دستور فوق مقادیر (۱۰ ۵) را به (۱۰ ۲۰) و مقادیر (۶ ۱) را به ۲۰ مبدل می‌سازد.

MTB>CODE (10:20) 1 (20:30) 2 C1 C2

- در این وضعیت عدد ۲۰ را به یک مبدل می‌کند.

MTB>CODE (-99) '*' C1-C10 C1-C10

- عدد -۹۹ را به داده گمشده مبدل می‌کند.

تمرین

۱-۷. یکی از مسائل کتاب در قسمت توصیفی، که شامل حداقل ۲۰ داده است را انتخاب کرده و سپس به کمک دستور فوق، داده‌ها را طبقه‌بندی نمایید.

Stack دستور
STACK (E...E) on (E...E) put in (C...C);
SUBSCRIPTS put in C

این دستور قابلیت متصل کردن ستون‌ها به هم را دارد می‌باشد.

MTB > STACK (C3 C4) (C1 C2) (C5 C6)

MTB > PRINT C1-C6

Data Display

| ROW | C1 | C2 | C3 | C4 | C5 | C6 |
|-----|----|-----|----|-----|-----|-----|
| 1 | 66 | 130 | 70 | 155 | 70 | 155 |
| 2 | 64 | 125 | 66 | 145 | 66 | 145 |
| 3 | 65 | 115 | 69 | 160 | 69 | 160 |
| 4 | | | | 66 | 130 | |
| 5 | | | | 64 | 125 | |
| 6 | | | | 65 | 115 | |

اگر پرانتز را به کار نبریم:

MTB > STACK C1 C2 67 63 C10

MTB > PRINT C1 C2 C10

Data Display

| Row | C1 | C2 | C10 |
|-----|----|-----|-----|
| 1 | 66 | 130 | 66 |
| 2 | 64 | 125 | 64 |
| 3 | 65 | 115 | 65 |
| 4 | | 130 | |
| 5 | | 125 | |
| 6 | | 115 | |
| 7 | | 67 | |
| 8 | | 63 | |

آخرین ستون، ستون هدف است.

یک ستون C از زیرنویس SUBSCRIPTS C زیربرنامه

MTB > STACK (C3 C4) (C1 C2) (C5 C6);

SUBC> SUBSCRIPTS C7.

MTB > PRINT C1-C7

Data Display

| Row | C1 | C2 | C3 | C4 | C5 | C6 | C7 |
|-----|----|-----|----|-----|----|-----|----|
| 1 | 66 | 130 | 70 | 155 | 70 | 155 | 1 |
| 2 | 64 | 125 | 66 | 145 | 66 | 145 | 1 |
| 3 | 65 | 115 | 69 | 160 | 69 | 160 | 1 |
| 4 | | | | | 66 | 130 | 2 |
| 5 | | | | | 64 | 125 | 2 |
| 6 | | | | | 65 | 115 | 2 |

Unstack دستور

UNSTACK (C...C) into (E...E) ... (E...E);
SUBSCRIPTS are in C.

ستون‌های را بر اساس زیربرنامه به ستون‌های دیگر تجزیه می‌کند.

MTB > unstack (c1 c2) (c11 c12) (c13 c14) (c15 c16);
SUBC> subscripts c3.

MTB > print c1-c3 c11-c16

Data Display

| Row | C1 | C2 | C3 | C11 | C12 | C13 | C14 | C15 | C16 |
|-----|----|-----|----|-----|-----|-----|-----|-----|-----|
| 1 | 24 | 126 | 2 | 35 | 142 | 24 | 126 | 27 | 134 |
| 2 | 27 | 134 | 3 | 31 | 122 | 21 | 119 | 32 | 148 |
| 3 | 32 | 148 | 3 | | 33 | 142 | | | |
| 4 | 35 | 142 | 1 | | | | | | |
| 5 | 21 | 119 | 2 | | | | | | |
| 6 | 33 | 142 | 2 | | | | | | |
| 7 | 31 | 122 | 1 | | | | | | |

اگر یک ستون را بخواهیم تفکیک کنیم نوشتن پرانتز الزامی نیست.

تمرین

۲-۷. به کمک چهار کد ایجاد شده برای داده‌های مفروض، داده‌های خود را برحسب طبقه‌ی مربوطه در چهار ستون تفکیک کنید.

Convert دستور

CONVERT using conversion table in C...C, convert C to C

داده های حرفی را می توان به عددی تغییر داد. ابتدا ستون هایی که کدها در آن قرار دارند را تعریف می کنیم.

```
MTB > READ C1 C2;
SUBC> FORMAT(F1.0, 2X, A6).
DATA> 1 RED
DATA> 2 YELLOW
DATA> END
```

به دو فرم می توان دستور را در مبدل سازی به کار برد.

```
MTB > CONVERT C1 C2 C10 C11.
* NOTE * 1 other values converted to blank
MTB > CONVERT C2 C1 C12 C13
* NOTE * 1 other values converted to missing
MTB > PRINT C1 C2 C10-C13
```

Data Display

| Row | C1 | C2 | C10 | C11 | C12 | C13 |
|-----|----|--------|-----|--------|--------|-----|
| 1 | 1 | RED | 1 | RED | RED | 1 |
| 2 | 2 | YELLOW | 2 | YELLOW | RED | 1 |
| 3 | | | 1 | RED | YELLOW | 2 |
| 4 | | | 3 | | RED | 1 |
| 5 | | | 2 | YELLOW | BLUE | * |

Concatenate دستور

CONCATENATE C...C put in C

اگر در تعداد حروف در یک ستون محدودیت داشته باشیم از این دستور استفاده می کنیم.

```
MTB > READ C1-C3;
SUBC> FORMAT(A4,A3,A4).
DATA> 194-42-7777
DATA> 123-45-6789
DATA> END
2 rows read.
```

MTB > CONCATENATE C1-C3 C4

MTB > PRINT C1-C4

Data Display

| Row | C1 | C2 | C3 | C4 |
|-----|------|-----|------|-------------|
| 1 | 194- | 42- | 7777 | 194-42-7777 |
| 2 | 123- | 45- | 6789 | 123-45-6789 |

۳-۷ محاسبات در میتب

در این بخش دستوراتی که به کمک آنها محاسبات ریاضی و آماری را ساده‌تر می‌نماید ارائه می‌دهیم. در برخی موارد دستورات دارای زیر دستوراتی می‌باشند که برای سادگی این فصل از نوشتمن و توضیح دادن خودداری شده است.

دستور Let
LET E=expression

این دستور، به کمک یک عبارت عملیات محاسباتی را انجام می‌دهد. عبارت باید حداقل شامل یکی از عملگرهای حسابی، یک مقایسه‌ای، منطقی و یا توابع باشند. در ابتدا می‌توانید تعدادی عدد وارد یک ستون نمایید. شناسه می‌تواند ستون‌ها ثابت‌های ذخیره شده باشند. توجه داشته باشید که از ماتریس‌ها نمی‌توان استفاده کرد. نمی‌توان در دستور LET متن به کار برد مگر متن را در آخر دستور، بعد از علامت # بنویسیم. در زیر چند مثال آورده شده است.

```
MTB > LET C1=(C2+C3)*10-60
MTB > LET C1=MEAN(C1)
MTB > LET K1=0.5
MTB > LET K2=MEAN(C1)/STDEV(C1)
MTB > LET C5=(C1<5)
MTB > LET K2=C1(28)
MTB > LET C1(28)=2.35
```

حداکثر پرانتزها در دستور LET، نه (۹) می‌باشد وحداکثر عبارتی که می‌توان در LET نوشت ۷۸ کارکتردارد. در صورت نیاز می‌توانید عبارت را به زیر عبارتی دیگر خرد کرد.

(ARITHMETIC OPERATION) عملگرهای محاسباتی

| | |
|----|-------|
| ** | توان |
| * | ضرب |
| / | تقسیم |
| + | جمع |
| - | منها |

کلیه‌ی عملیات به صورت سطری انجام می‌گیرد. در صورت وجود عضو گم شده نتیجه به صورت گم شده نمایش می‌دهد. اگر عمل‌گری امکان انجام محاسبه را نداشته باشد (مانند تقسیم بر صفر) نتیجه عضوی گم شده می‌شود. در زیر چند مثال پرداخته‌ایم.

```
MTB > let c10=c1+c2
MTB > let c11=c2**2
MTB > let c12=c1/c2
MTB > let c12=c1/c2
J
*** Values out of bounds during operation at J
Missing returned 1 times
```

```
MTB > let c13=c1+c2+c3+100
MTB > print c1-c3 c10-c13
Data Display
Row C1 C2 C3 C10 C11 C12 C13
1 5 2 2 7 4 2.5 109
2 3 6 2 9 36 0.5 111
3 1 -1 * 0 1 -1.0 *
4 6 0 1 6 0 * 107
```

عملگرهای مقایسه‌ای (COMPRISON OPERATION)

ساختار یک عبارت منطقی به کمک عملگرهای زیر انجام می‌گیرد.

| | | | |
|----|----|----|-----------------|
| = | یا | EQ | برابر |
| ~= | یا | NE | نابرابر |
| < | یا | LT | کمتر |
| > | یا | GT | بزرگتر |
| <= | یا | LE | کمتر یا مساوی |
| >= | یا | GE | بزرگتر یا مساوی |

عملیات به صورت سطری انجام می‌گیرد. اگر عبارت درست باشد مقدار یک، در غیر این صورت مقدار صفر حاصل می‌شود. با عملگرهای GE, GT, LT, LE اگر سطری عضو گم شده داشته باشد آنگاه نتیجه نیز به صورت گم شده خواهد شد. اما با

عملگرهای NE, EQ مقایسه را می‌توان با عضو گم شده نیز انجام داد. به مثال‌های زیر توجه کنید.

MTB > LET C3=(C1<C2)

MTB > LET C4=(C1='*')

MTB > LET C5=2*(C1=1)+5*(C1~=1)

MTB > LET C6=(C1 EQ C2)

MTB > PRINT C1-C6

Data Display

| Row | C1 | C2 | C3 | C4 | C5 | C6 |
|-----|----|----|----|----|----|----|
| 1 | 4 | 8 | 1 | 0 | 5 | 0 |
| 2 | 4 | 4 | 0 | 0 | 5 | 1 |
| 3 | 1 | -2 | 0 | 0 | 2 | 0 |
| 4 | * | 5 | * | 1 | 5 | 0 |

عملگرهای منطقی (LOGICAL OPERATION)

| | | |
|---|-----|----|
| & | AND | و |
| | OR | یا |
| ~ | NOT | نه |

عملیات سطری انجام می‌گیرد. با ارزش هر عبارت بر حسب درستی یا نادرستی مقدار یک یا صفر نسبت می‌دهد. در صورتی که محاسبه انجام‌پذیر نباشد، علامت * را نشان می‌دهد. اگر هر دو عبارت در طرفین AND درست باشد ارزش آن یک در غیر این-صورت صفر خواهد شد. اگر یکی از عبارت‌های طرفین OR درست باشد ارزش آن یک است. اگر عبارتی درست باشد، با دستور NOT تغییر به نادرست می‌شود.

مثال ۶-۷

MTB > LET C3=C1&C2

MTB > LET C4=C1|C2

MTB > LET C5=~C2

MTB > LET C6=(C1 GT 2)OR(C2 EQ 1)

MTB > PRINT C1-C6

Data Display

| Row | C1 | C2 | C3 | C4 | C5 | C6 |
|-----|----|----|----|----|----|----|
| 1 | 1 | 1 | 1 | 1 | 0 | 1 |
| 2 | 1 | 0 | 0 | 1 | 1 | 0 |
| 3 | 0 | 0 | 0 | 0 | 1 | 0 |

| | | | | | | |
|---|---|---|---|---|---|---|
| 4 | 1 | * | * | 1 | * | 0 |
| 5 | 0 | * | 0 | * | * | 0 |
| 6 | 0 | 5 | 0 | 1 | 0 | 0 |

توابع و عملگرهای ستونی (FUNCTIONS AND COLUMN OPERATIONS)

توابع زیر به صورت سطحی عملیات را انجام می‌دهند.

| | | | |
|---------|-------------------------|----------|---------------------|
| SQRT | ریشه‌ی دوم | ABSOLUTE | قدر مطلق |
| LOGE | لگاریتم در مبنای عدد پی | LOGTEN | لگاریتم در مبنای ۱۰ |
| ANTILOG | تابع 10^x | EXPONENT | تابع نمایی e^x |
| ASIN | ارک سینوس | ROUND | گرد |
| ACOS | ارک کسینوس | SIN | سینوس |
| ATAN | ارک تانژانت | COS | کسینوس |
| PARPROU | ضرب جزئی | TAN | تانژانت |
| | | PARSUM | جمع جزئی |

عملگرهای زیر، بر روی ستون‌ها محاسبات را انجام می‌دهند و یک عدد به دست می‌آورند.

| | | | |
|--------|--------------|-------|-----------------------|
| MEAN | میانگین | COUNT | تعداد داده‌های معلوم |
| MEDIAN | میانه | NMISS | تعداد داده‌های گم شده |
| STDEV | انحراف معیار | N | تعداد کل داده‌ها |
| SSQ | مجموع مربعات | MAX | ماکسیمم داده‌ها |
| | | MIN | مینیمم داده‌ها |

سه عملگر زیر بر روی ستون‌ها اثر می‌گذارند و به شکل یک ستون پاسخ می‌دهند.
عملگرهای زیر خاص محاسبات آماری می‌باشند.

- از کوچک به بزرگ داده‌ها را مرتب می‌کند
- به عناصر یک ستون رتبه‌ای را نسبت می‌دهد
- لگ یا تأخیر

SORT

RANK

LAG

توجه داشته باشید که نوشتن پرانتز جلوی عملگرها ضروری است. توابع و عملگرهای ستونی را می‌توان با چهار حرف اوّل آن به کار گرفت.

مثال ۷-۷.

```
MTB > LET C10=SQRT(C1)
MTB > LET C11=ABSO(C2)+10
MTB > LET K1=SUM(C1)
MTB > LET C12=C1-MEAN(C1)
MTB > LET C13=K1*C2
MTB > PRINT C1-C2 C10-C13
Data Display
Row C1 C2 C10 C11 C12 C13
 1   1  4 1.00000 14  -3   48
 2   9  -3 3.00000 13    5  -36
 3   2   0 1.41421 10   -2    0
MTB > PRINT K1
Data Display
K1      12.0000
```

می‌توان عناصر مشخص از یک ستون را به کمک دستور LET تغییر داد. برای مثال:

```
MTB > LET C1=C2(3)*C10
MTB > LET C2(3)=5+K1
MTB > LET C2=C1(COUNT(C2)-10)*C11
```

در ستون اوّل، ستون دهم را در سومین مقدار ستون دوم ضرب کرده و نتیجه را در ستون اوّل قرار می‌دهد. سومین عنصر از ستون دوم را برابر جمع دو عدد معلوم جایگزین می‌کند.

تذکر: چنانچه دستور را به شکل زیر بنویسیم

LET C1(15)=28

اگر ستون اوّل، ده عنصر داشته باشد آنگاه به جز عنصر پانزدهم که مقدار ۲۸ را نسبت می‌دهد از عنصر ۱۱ تا ۱۴ علامت ستاره می‌زند.

دستورات جبری ADD, SUBTRACT, MULTIPLY, DIVIDE, RAISE

ADD E to E, ..., to E put into E
SUBTRACT E from E

MULTIPLY E by E, ..., E put into E
 DIVIDE E by E
 RAISE E to the power E put into E

پنج دستور سطر به سطر عملیات جبری روی ستون ها را انجام می دهد.

مثال ۸-۷

```
MTB > ADD C1 C2 C3
MTB > ADD 5 C1 C4
MTB > RAISE C2 3 C5
MTB > PRINT C1-C5
Data Display
Row   C1   C2   C3   C4   C5
  1     1.1   2   3.1   6.1   8
  2     2.3   1   3.3   7.3   1
  3    -5.0   2   -3.0   0.0   8
```

تذکر: توابع تعریف شده قبلی نیز به همین شیوه قابل انجام شدن هستند.

دستورات سطري

```
RCOUNT   E... E put into C
RN      E... E put into C
RNMISS   E... E put into C
RSUM    E... E put into C
RSTDEV   E... E put into C
RMEAN    E... E put into C
RMEDIAN   E... E put into C
RMIN     E... E put into C
RMAX     E... E put into C
RSSQ     E... E put into C
```

عملیات را به صورت سطري برای حداقل ۱۰ ستون، انجام می دهند.

مثال ۹-۷

```
MTB > RSUM C1-C4 C5
MTB > RMEAN C1-C4 C6
MTB > RN C1-C4 C7
MTB > PRINT C1-C7
Data Display
Row   C1   C2   C3   C4   C5   C6   C7
  1     2     4     7     3    16    4.0     4
```

| | | | | | | | |
|---|---|---|---|---|----|-----|---|
| 2 | 5 | 1 | 2 | 2 | 10 | 2.5 | 4 |
| 3 | 2 | * | 1 | 3 | 6 | 2.0 | 3 |

Describe دستور

MTB> DESCRIBE C1 C2 ...

به کمک این دستور می‌توان برخی معیارهای تمرکز و معیارهای پراکندگی را به دست آورد.

مثال ۱۰-۷.

MTB > SET C1

DATA > 10 13 15 9 17 20 15 17 8 11 * 12 14

DATA > END

MTB > DESC C1

Descriptive Statistics

| Variable | N | N* | Mean | Median | TrMean | StDev | SEMean |
|----------|----|----|-------|--------|--------|-------|--------|
| C1 | 12 | 1 | 13.42 | 13.50 | 13.30 | 3.60 | 1.04 |

| Variable | Min | Max | Q1 | Q3 |
|----------|------|-------|-------|-------|
| C1 | 8.00 | 20.00 | 10.25 | 16.50 |

ترجمه‌ی کلمات و حروف به کار رفته

| | | | |
|----------------|-----------------------|--------|---|
| N | تعداد کل داده‌ها | MEAN | میانگین |
| N* | تعداد داده‌های گم شده | MEDIAN | میانه |
| MIN | مینیمم داده‌ها | TRMEAN | میانگین پیراسته٪ ۱۰ |
| MAX | ماکسیمم داده‌ها | STDEV | انحراف معیار $\sqrt{\frac{\sum(X_i-\bar{X})^2}{n-1}}$ |
| Q ₁ | چارک اول | SEMEAN | $stdev/\sqrt{n}$ |
| Q ₃ | چارک سوم | | |

Table دستور

از این دستور برای ساختن جدول در داده‌های گسسته به کار گرفته می‌شود. به مثال‌های زیر توجه کنید.

.۱۱-۷ مثال

```
MTB > set c1  
DATA> 1 1 2 1 1 0 2 2 1  
DATA> end  
MTB > TABLE C1  
Tabulated Statistics  
ROWS: C1  
    COUNT  
    0      1  
    1      5  
    2      3  
ALL     9
```

.۱۲-۷ مثال

```
MTB > READ C1 C2  
DATA> 1 1  
DATA> 1 0  
DATA> 1 1  
DATA> 0 1  
DATA> 0 1  
DATA> 0 0  
DATA> END  
    6 rows read.  
MTB > TABLE C1 C2.  
Tabulated Statistics  
ROWS: C1    COLUMNS: C2  
    0      1      ALL  
    0      1      2      3  
    1      1      2      3  
ALL     2      4      6  
CELL CONTENTS -- COUNT
```

Covariance دستور

MTB> COVARINCE C1 C2

این دستور مقدار کواریانس را با ضریب $1 - n$ محاسبه می کند. با درنظر گرفتن داده های

مثال ۱۲-۷ داریم:

```
MTB > COVARINCE C1 C2  
Covariances
```

| | C1 | C2 |
|----|----------|----------|
| C1 | 0.300000 | |
| C2 | 0.000000 | 0.266667 |

مقدار $\frac{300000}{300000} = 1$ واریانس ستون اول، مقدار $\frac{266667}{266667} = 1$ واریانس ستون دوم، و مقدار $\frac{0.000000}{0.000000} = 0$ کوواریانس بین ستون اول و دوم است. مقدار کوواریانس با ضریب n را می‌توانید با دستور LET بسازید.

```
MTB > LET k1=sum((C1-MEAN(C1))*(C2-MEAN(C2)))/N(C1)
MTB > PRINT k1
```

تمرین

- ۳-۷. عملیات لازم همراه با مثالی عددی، برای محاسبهٔ نما را بنویسید.
 ۷-۴. به کمک دستورات، مقدماتی چارک اول و چارک سوم هر سری داده را به دست آورید.

۷-۵. نشان دهید برای سری داده فرضی $\sum_{i=1}^n (x_i - \bar{x})$ برابر صفر است.

۷-۶. داده‌های زیر را در نظر بگیرید:

۱ ۴ ۳ ۲ ۵ ۸ ۹ ۵ ۴ ۳ ۲ ۶ ۵ ۳ ۲ ۱

الف) نشان دهید عبارت $\sum_{i=1}^n |x_i - a|$ به ازای a برابر میانه مینیمم می‌شود (با سه مقدار: کوچکتر از میانه، میانه، بزرگتر از میانه نشان دهید).

ب) نشان دهید عبارت $b = \bar{x}$ به ازای $\sum_{i=1}^n (x_i - b)^2$ مینیمم می‌شود.

ج) گشتاور مرکزی مرتبه‌های سوم و چهارم را محاسبه کنید.

$$M_p = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^p, \quad M_4 = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^4.$$

۷-۷. چند درصد داده‌های سؤال ۷-۶، از میانگین بزرگتر هستند.

۸-۷ نشان دهید هرگاه از داده‌های سؤال ۶-۷، میانه را کم کنیم تغییری در مقدار واریانس به وجود نمی‌آید.

۹-۷ حداقل پنجاه داده را برداشته (از کتاب استفاده نمایید) آنرا به کمک دستورات نرم افزار طبقه‌بندی نمایید و سپس میانگین، واریانس، میانه، چارک اول و چارک سوم و نما را به کمک داده‌های طبقه‌بندی شده محاسبه نمایید.

۱۰-۷ به کمک نرم افزار کمیت‌های زیر را محاسبه کنید:

$$\log(\frac{3}{4}) \quad e^{1/5} \quad 3^{\sqrt[3]{3}}$$

$$(1+2)(3+4)(8+9)(11+10)(2+7)$$

۱۱-۷ به کمک عبارت‌های منطقی، ریشه‌های معادله‌ی $x^3 + 3x + 2 = 0$ را بیابید.

۱۲-۷ به دستورات نرم افزار، !۱۲ را محاسبه کنید.

۱۳-۷ دستورات لازم جهت محاسبه‌ی دامنه‌ی تغییرات را بنویسید.

۷-۴ رسم نمودار

هنگامی که دستورات رسم نمودارها را اجرا می‌کنید به استثناء نمودار نقطه‌ای، محیط دیگری برای نشان دادن نمودارها فراهم می‌گردد، که می‌توان در آن محیط نمودارها را ذخیره نمود.

Histogram دستور

```
HISTOGRAM C ... C;
INCREMENT =K;
START =K [end =K];
BY C;
SAME scale for all columns
```

به طور جداگانه برای هر یک از ستون‌ها نمودار هستیوگرام را رسم می‌کند. به کمک زیر برنامه‌های آمده می‌توان طول طبقات، نقطه‌ی شروع نمودار و سایر مقیاس‌بندی‌ها را

برای ستون‌ها به کار برد. ساده‌ترین شکل آن است که هیچیک از زیر برنامه‌ها را به کار نبرده تا نرم‌افزار با رعایت فواصل مساوی نمودار را رسم کند.

مثال ۱۳-۷.

```
MTB> SET C1
DATA> 12 11 15 21 15 14 17 18 19 14 12 13 15
DATA> END
MTB> HISTOGRAM C1
```

Dotplot دستور

```
DOTPLOT C
INCREMENT =K;
START =K [end =K];
BY C;
SAME scale for all columns
```

آنچه که یک نمودار نقطه‌ای نشان می‌دهد، مشابهی یک هسیتوگرام است اما محورها به طبقات بیشتری تقسیم و هر مشاهده به وسیله‌ی یک نقطه نمایش داده می‌شود.

مثال ۱۴-۷. با داده‌های مثال فوق نمودار نقطه‌ای را رسم می‌کنیم

```
MTB > DOTPLOT C1
Character Dotplot
```



Plot دستور

```
PLOT C versus C
TITLE='text'
FOOTNOTE='text'
XLABEL='text'
YLABEL='text'
SYMBOL='text'
XINCREMENT =K;
XSTART =K [end =K];
YINCREMENT =K;
YSTART =K [end =K]
```

ستون اوّل را روی محور عرض‌ها و ستون دوم را روی محور طول‌ها قرار می‌دهد. معمولاً هر نقطه را با نماد \times نمایش می‌دهد هرگاه دو یا چند نقطه روی هم قرار گیرند تعداد آنها در کنار نقطه‌ی مشترک ذکر می‌شود اگر تعداد آنها از نه تا بیشتر شد از علامت + استفاده می‌کند. زیربرنامه‌ها را برای اضافه کردن متن‌هایی در نمودار به کار می‌گیرند. به کمک SYMBOL می‌توان علامت نمایشی نمودار را تغییر داد.

مثال ۱۵-۷

```
MTB > SET C1  
DATA> 1:10/0.1  
DATA> END  
MTB > LET C2=C1**2  
MTB > PLOT C2*C1
```

این نرم افزار نمودارهای متعدد دیگری را رسم می‌کند منجمله STEM-AND-LEAF, PIE CHART, ... که می‌توانید با انتخاب منوها و یا با نوشتن دستورات آنها را انجام دهید.

تمرین

۱۴-۷. به کمک دستورات فوق داده‌های کتاب را انتخاب کرده و سپس انواع نمودارها را بسته به نوع آنها، رسم کنید.

حل تمرین‌ها

حل تمرین‌های فصل اول

۱-۴. الف)

| طبقه | تعداد | فراوانی | تجمعی | فراوانی |
|--------|-------|---------|-------|---------|
| ۵-۶- | ۸ | ۰/۸ | ۸ | |
| ۶-۷- | ۱۹ | ۰/۱۹ | ۳۷ | |
| ۷-۸- | ۳۵ | ۰/۳۵ | ۶۲ | |
| ۸-۹- | ۱۵ | ۰/۱۵ | ۷۷ | |
| ۹-۱۰- | ۱۴ | ۰/۱۴ | ۹۱ | |
| ۱۰-۱۱- | ۹ | ۰/۹ | ۱۰۰ | |

۱-۵

| طبقه | فراوانی | تجمعی | درصد | فراوانی | کرانهای ردیدها |
|---------|---------|-------|-------------|---------|----------------|
| ۱۵-۲۰/۹ | ۰/۱۲ | ۳ | ۱۴/۹۵-۲۰/۹۵ | | |
| ۲۱-۲۶/۹ | ۰/۳۲ | ۱۱ | ۲۰/۹۵-۲۶/۹۵ | | |
| ۲۷-۳۲/۹ | ۰/۳۶ | ۲۰ | ۲۶/۹۵-۳۲/۹۵ | | |
| ۳۳-۳۸/۹ | ۰/۵۸ | ۲۲ | ۳۲/۹۵-۳۸/۹۵ | | |
| ۳۹-۴۴/۹ | ۰/۱۲ | ۲۵ | ۳۸/۹۵-۴۴/۹۵ | | |

۱-۶. ابتدا داده‌ها را مرتب می‌کنیم:

۰ ۰ ۰ ۱ ۱ ۱ ۱ ۲ ۲ ۳ ۳ ۳ ۴ ۴ ۵ ۶ ۷ ۸

$$\tilde{x} = \frac{۲+۳}{۴} = ۲/۵$$

در اینجا دو عدد ۳ و ۱ هر کدام چهار بار تکرار شده‌اند. لذا هر دو مد هستند. برای

محاسبه‌ی چارک اول $\frac{n}{f} = \frac{۳۰}{۴} = ۵$. پس چارک اول پنجمین عدد مرتب شده است،

یعنی $Q_1 = 1$.

$$\bar{x}_H = \left(\frac{1}{\mu} + \frac{1}{\sigma^2} + \frac{1}{\lambda} \right)^{-1} = \left(\frac{1}{\mu} + \frac{\gamma}{\lambda} \right)^{-1} = \frac{\mu + \gamma}{\lambda} \quad .7-1$$

۸-۱ به ازای هر a ای،

$$\begin{aligned} \sum_{i=1}^n (x_i - a)^r &= \sum_{i=1}^n (x_i - \bar{x} + \bar{x} - a)^r \\ &= \sum_{i=1}^n (x_i - \bar{x})^r + \sum_{i=1}^n (\bar{x} - a)^r + r \sum_{i=1}^n (x_i - \bar{x})(\bar{x} - a). \end{aligned}$$

از طرفی عبارت $(\bar{x} - a)^r$ به اندیس i بستگی ندارد، پس داریم:

$$\sum_{i=1}^n (\bar{x} - a)^r = n(\bar{x} - a)^r$$

و چون می‌دانیم که $\sum_{i=1}^n (x_i - \bar{x}) = 0$ ، خواهیم داشت:

$$\sum_{i=1}^n (x_i - \bar{x})(\bar{x} - a) = (\bar{x} - a) \sum_{i=1}^n (x_i - \bar{x}) = 0$$

در نهایت خواهیم داشت

$$\sum_{i=1}^n (x_i - a)^r = \sum_{i=1}^n (x_i - \bar{x})^r + n(\bar{x} - a)^r.$$

۹-۱. خیر زیرا

$$\text{مجموع وزن افراد} = 8 \times 123 + 5 \times 174 = 984 + 870 = 1854 < 2000$$

.10-۱

$$\bar{x} = \frac{575 \times 120000 + 1410 \times 117000 + 1425 \times 121000}{575 + 1410 + 1425}$$

$$= \frac{168395000}{1410} = 119 / 1429.$$

.11-۱

$$\sigma^r = r, \quad \bar{x} = \mu, \quad \sum_{i=1}^N x_i^r = \Delta \Delta,$$

$$\sigma^2 = \frac{\sum_{i=1}^N x_i^2 - N \bar{x}^2}{N} = \frac{55 - N \times 11}{N} = \frac{55}{N} - 11 = 2.$$

$$5N = 55 \rightarrow N = 11.$$

۱۲-۱. میانگین داده‌های جدید $\bar{x} + 1$ است.

۱۳-۱. میانگین داده‌های جدید به صورت زیر محاسبه می‌شود:

$$\bar{x}_{new} = \frac{\sum_{i=1}^9 (x_i + i)}{9} = \frac{\sum_{i=1}^9 x_i}{9} + \frac{\sum_{i=1}^9 i}{9} = \bar{x} + \frac{9 \times 10}{9} = 11 + 5 = 15.$$

۱۴-۱. میانگین داده‌های جدید به صورت زیر محاسبه می‌شود:

$$\bar{x}_{new} = \frac{\sum_{i=1}^{10} (x_i + 10i)}{10} = \frac{\sum_{i=1}^{10} x_i}{10} + \frac{\sum_{i=1}^{10} 10i}{10} = \bar{x} + \frac{10 \times 11}{10} = 11 + 10 = 21.$$

.۱۵-۱

$$\bar{x} = \frac{\sum_{i=1}^{10} x_i}{10} = 15 \rightarrow \sum_{i=1}^{10} x_i = 10 \times 15 = 150,$$

$$\bar{x}_{new} = \frac{\sum_{i=1}^{13} x_i + 15 + 15 + 16}{13} = \frac{150 + 15 + 15 + 16}{13} = \frac{186}{13} = 14.3.$$

۱۶-۱. بنابر صفر بودن انحراف معیار داریم: $x_1 = x_2 = x_3 = x_4 = 25$ ، پس

$$\frac{\sum_{i=1}^4 x_i + 50}{5} = \frac{4 \times 25 + 50}{5} = 30.$$

۱۷-۱. تغییری ایجاد نمی‌شود.

۱۸-۱. بنا براین داده‌های مرتب شده عبارتند از $x_1 = x_2 = \dots = x_{15}$.

$$17, 19, 31, 31, \dots, 31$$

در نتیجه $\tilde{x} = 31$.

۱۹-۱. اعداد جدید مضرب $\frac{1}{4}$ اعداد قدیمی‌اند. لذا واریانس اعداد جدید $\frac{1}{16}$ اعداد

قدیمی و در نتیجه انحراف معیار آنها $\frac{1}{4}$ اعداد قدیمی است.

$$\sigma_{ax+b} = |a| \sigma_x.$$

پس انحراف معیار اعداد جدید برابر است با $\frac{1}{4} \times 9/8 = 2/45$.

(الف) ۲۰-۱

| ردی | کرانهای ردتها | f_i | m_i | $f_i m_i$ | فراوانی تجمعی | $f_i m_i^r$ |
|------------|---------------|-------|-------|-----------|---------------|-------------|
| ۱۴/۵_۵۱/۵ | ۱۴/۲۵_۵۱/۷۵ | ۵ | ۴۹/۵ | ۲۴۷/۵ | ۵ | ۱۲۲۵۱ |
| ۵۱/۵_۶۱/۵ | ۵۱/۷۵_۶۱/۷۵ | ۱۵ | ۵۹/۵ | ۲۶۷۷/۵ | ۵۰ | ۱۵۹۳۱۱ |
| ۶۱/۵_۷۱/۵ | ۶۱/۷۵_۷۱/۷۵ | ۱۳ | ۶۹/۵ | ۲۹۸۷/۵ | ۹۳ | ۲۰۷۷۰۰ |
| ۷۱/۵_۸۱/۵ | ۷۱/۷۵_۸۱/۷۵ | ۱۹ | ۷۹/۵ | ۱۵۱۰/۵ | ۱۰۲ | ۱۲۰۰۸۴ |
| ۸۱/۵_۹۱/۵ | ۸۱/۷۵_۹۱/۷۵ | ۷ | ۸۹/۵ | ۶۲۶/۵ | ۱۱۹ | ۵۶۰۷۱ |
| ۹۱/۵_۱۰۱/۵ | ۹۱/۷۵_۱۰۱/۷۵ | ۱ | ۹۹/۵ | ۹۹/۵ | ۱۲۰ | ۹۹۰۰ |
| | | ۱۲۰ | | ۸۱۵۰ | | ۵۶۵۳۱۷ |

$$\bar{x} = \frac{\sum_{i=1}^n f_i m_i}{120} = \frac{8150}{120} = 67.5,$$

$$\tilde{x} = 67.5 / 75 + \frac{60 - 50}{45} \times 10 = 67.5 / 75 + 2 / 22 = 66 / 97,$$

$$s^r = \frac{n(\sum_{i=1}^n m_i^r f_i) - (\sum_{i=1}^n f_i m_i)^r}{n(n-1)} = \frac{120(565317) - (8150)^r}{120(119)} = 99.$$

(ب)

٢٧٥ حل تمرین‌ها

$$\begin{aligned} \frac{1}{\varsigma} \times n = ۳۰ &\rightarrow Q_1 = \omega \nu / \nu \omega + \frac{\nu \circ - \omega}{\nu \omega} \times ۱۰ = \omega \nu / \nu \omega + \omega / \omega \omega = \omega / \nu, \\ \frac{۳}{\varsigma} \times n = ۹۰ &\rightarrow Q_{\nu} = \omega \nu / \nu \omega + \frac{\nu \circ - \omega \circ}{\nu \omega} \times ۱۰ = \nu \nu / \omega \omega, \\ SK = \frac{\nu(\bar{x} - Md)}{S} &= \frac{\nu(\omega \nu / \nu - \omega \nu / \nu)}{\nu / \nu \nu} = \circ / \nu \nu. \end{aligned}$$

.۲۱-۱

$$SK = \frac{\nu(\bar{x} - Md)}{S} = \frac{\nu \times (\nu / \omega - \nu / \nu \nu)}{\nu / \nu \omega} = \circ / \circ \nu \nu.$$

حل تمرین‌های فصل دوم

$$. A \cup B . ۱-۲$$

۲-۲. اعداد ۴، ۶، ۸ و ۲ مضارب عدد ۲ و عدد ۷ مضرب عدد ۷ هستند. لذا مجموعه‌ی ۳، ۴، ۶، ۷ و ۸ مضارب اعداد ۲ یا ۷ هستند.

$$. A \times B . ۳-۲$$

۴-۲. به ${}^4 \times {}^3 = 12$ حالت می‌توان با انتخاب یکی از ۴ جفت کفشهای جوراب و یکی از ۳ جفت کفشهای جوراب و کفشهای پوشید.

$$. {}^3 \times {}^4 \times {}^5 \times {}^2 = 120 . ۵-۲$$

$$. ۶-۲ \quad ۳^2 \times ۳^2 \times ۱^0 \times ۱^0 \times ۱^0 \times ۱^0 = ۳^2 \times ۳^2 \times ۱^0 = ۳^2 \times ۹^2 = ۸۱^2 = ۶۵۶$$

۷-۲. به ازای هر نقطه دو حالت، ۰ و ۱ داریم پس n تابع می‌توان ساخت.

$$. ب) !^4 \cdot ا) !^4 . ۹-۲$$

$$. ا) !^4 . ۱۰-۲$$

$$. ۱۱-۲ \quad \frac{10!}{4!3!2!}$$

$$. ۱۲-۲ \quad \frac{9!}{4!3!2!}$$

$$. ۱۳-۲ \quad {}^5 P_3 = \frac{5!}{(5-3)!} = 5 \times 4 \times 3 = 60$$

$$. ۱۴-۲ \quad (4-1)! = 3!$$

حل تمرین‌ها ۲۷۷

$$\frac{(5-i)!}{i!} = \frac{4!}{2} = 12 \quad .15-2$$

$$\binom{5}{2} \binom{4}{3} = \frac{5!}{3!2!} \times \frac{4!}{4!3!} = 10 \times 35 = 350 \quad .16-2$$

در حالت دوم، هر دو مرد را که با هم تفاهمندانند در یک گروه در نظر گرفته و از آن گروه فقط یک نفر را انتخاب می‌کنیم و از ۵ مرد باقیمانده دو نفر انتخاب می‌کنیم

$$\binom{5}{2} \binom{3}{1} \binom{5}{2} = \frac{5!}{3!2!} \times 2 \times \frac{5!}{3!2!} = 200.$$

$$\binom{10}{5} = \frac{10!}{5!5!} = 210 \quad .17-2$$

$$\frac{4!}{2!1!1!} = 12 \quad .18-2$$

۱۹-۲. الف) اگر $x=1$ و $y=1$

$$\sum_{r=0}^n \binom{n}{r} x^r y^{n-r} = \sum_{r=0}^n \binom{n}{r} = (1+1)^n = 2^n.$$

ب) اگر $x=-1$ و $y=1$

$$\sum_{r=0}^n \binom{n}{r} x^r y^{n-r} = \sum_{r=0}^n \binom{n}{r} (-1)^r 1^{n-r} = \sum_{r=0}^n \binom{n}{r} (-1)^r = (-1+1)^n = 0.$$

ج) اگر $x=a-1$ و $y=1$

$$\sum_{r=0}^n \binom{n}{r} x^r y^{n-r} = \sum_{r=0}^n \binom{n}{r} (a-1)^r 1^{n-r} = \sum_{r=0}^n \binom{n}{r} (a-1)^r = (a-1+1)^n = a^n.$$

۲۰-۲. طبق قضیه‌ی ۸-۲ داریم:

$$\binom{n}{r} = \binom{n-1}{r} + \binom{n-1}{r-1},$$

از قضیه‌ی ۸-۲ به جای $\binom{n-1}{r-1}$ داریم:

$$\binom{n-1}{r-1} = \binom{n-1-1}{r-1} + \binom{n-1-1}{r-1-1}$$

پس می‌توان نوشت

$$\begin{aligned}
 \binom{n}{r} &= \binom{n-1}{r} + \binom{n-1}{r-1} \\
 &= \binom{n-1}{r} + \binom{n-1-1}{r-1} + \binom{n-1-1}{r-1-1} \\
 &= \binom{n-1}{r-1+1} + \binom{n-1-1}{r-1} + \binom{n-1-1}{r-1-1} + \binom{n-1-1-1}{r-1-1-1} + \dots \\
 &= \sum_{i=1}^{r+1} \binom{n-i}{r-i+1}.
 \end{aligned}$$

۲۱-۲. طبق قضیه‌ی ۹-۲ داریم

$$\sum_{r=0}^k \binom{m}{r} \binom{n}{k-r} = \binom{m+n}{k}.$$

حال اگر $k = n$ و $m = n$ خواهیم داشت

$$\binom{mn}{n} = \sum_{r=0}^n \binom{n}{r} \binom{n}{n-r} = \sum_{r=0}^n \binom{n}{r} \binom{n}{r} = \sum_{r=0}^n \binom{n}{r}^2.$$

۲۲-۲

$$\sum_{r=0}^n \binom{n}{r} x^{n-r} y^r = (x+y)^n \xrightarrow{x=1} \sum_{r=0}^n \binom{n}{r} y^r = (1+y)^n.$$

از طرفین نسبت به y مشتق می‌گیریم، خواهیم داشت

$$\sum_{r=0}^n r \binom{n}{r} y^{r-1} = n(1+y)^{n-1},$$

حال اگر قرار دهیم $y = 1$ ، خواهیم داشت

$$\sum_{r=0}^n r \binom{n}{r} = n!^{n-1}.$$

۲۳-۲. الف)

$$\binom{\frac{1}{r}}{r} = \frac{\frac{1}{r}(\frac{1}{r}-1)(\frac{1}{r}-2)(\frac{1}{r}-3)}{r!} = \frac{-15}{384}.$$

٢٧٩ حل تمرین‌ها

(ب)

$$\binom{-m}{m} = \frac{-m(-m-1)(-m-2)}{m!} = \frac{-60}{5} = -12.$$

٢٤-٢ (الف)

$$\binom{-1}{r} = \frac{-1(-r)(-m)\cdots(-1-(r-1))}{r!} = \frac{(-1)^r (1 \times 2 \times m \times \cdots \times r)}{r!} = (-1)^r.$$

(ب)

$$\binom{-n}{r} = \frac{-n(-n-1)(-n-2)\cdots(-n-r+1)}{r!}$$

$$= \frac{(-1)^r (n(n+1)\cdots(n+r-1))}{r!}$$

$$= \frac{(-1)^r (n+r-1)!}{r! (n-1)!}$$

$$= (-1)^r \binom{n+r-1}{r}.$$

$$\frac{\lambda!}{r! m! n!} = 280. \quad ٢٥-٢$$

$$r^r \times m^r \times n^r \times \frac{q!}{r! \times m! \times n!} = r^r \times q!. \quad ٢٦-٢$$

حل تمرین‌های فصل سوم

$$. S = \{HH, HT, TH, TT\} . ۱-۳$$

$$S = \{245, 246, 247, 248, 249, 250, 251, 252, 253, 254, 255\} . ۲-۳$$

اندازه‌گیری می‌توان فضای نمونه‌ای آن را به صورت $S = [245, 255]$ اعلام کرد.

$$S = \{0, 1, 2, 3, \dots\} . ۳-۳$$

$$S = \{0, 1, 2, 3, \dots\} . ۴-۳$$

$$S = \{\text{تریلی}, \text{کامیون}, \text{اتوبوس}, \text{سواری}\} . ۵-۳$$

$$. ۶-۳$$

$$S = \{(H, i), (H, r), (H, s), (H, t), (H, u), (H, v), \\(T, T, T), (T, T, H), (T, H, T), (T, H, H)\}.$$

$$A \cap B = \{x | 5 < x < 8\} \quad \text{ب) } A \cup B = \{x | 2 < x < 10\} . ۷-۳$$

$$A' \cup B = \{x | 0 < x < 3, 5 < x < 10\} \quad \text{د) } A \cap B' = \{x | 3 < x < 5\} . ۸-۳$$

$$. ۸-۳$$

$$\#(B) = 115 = \text{تعداد جامعه شناس}$$

$$\#(A) = 138 = \text{تعداد آماردان}$$

$$\#(A \cap B) = \text{تعداد آمار و جامعه شناس} = 91$$

$$\#(A \cup B) = \#(A) + \#(B) - \#(AB) = 138 + 115 - 91 = 162.$$

تعدادی که در هیچ‌کدام ثبت‌نام نکرده‌اند

$$\#(A \cup B)' = 200 - 162 = 38.$$

۹-۳. الف) خودروهایی که به تعمیر موتور، جعبه‌دنده و تعویض تایر نیاز دارند.

ب) خودروهایی که به تایرهای نو موتور و جعبه‌دنده نیاز دارند ولی نیازی به تعمیر کامل موتور ندارند.

ج) خودروهایی که فقط به تعمیر کامل موتور نیاز دارند.

د) خودروهایی که به تعمیر کامل موتور و تایرهای نو نیاز دارند.

ت) خودروهایی که فقط به تعمیر کامل موتور و جعبه‌دنده نیاز دارند یا خودروهایی که فقط به تعمیر جعبه‌دنده نیاز دارند.

و) خودروهایی که فقط به تعمیر جعبه‌دنده یا فقط به تعویض تایر و یا فقط به تعمیر جعبه‌دنده و تعویض تایر نیاز دارند.

۱۰-۳. الف) ناحیه‌ی ۲

ج) ناحیه‌های ۶، ۵، ۳ د) ناحیه‌ی ۶

۱۱-۳

$$P(A \cup B) = P(A) + P(B) - P(A \cap B) = ۰/۴ + ۰/۶ - P(A \cap B) = ۰/۵,$$

$$P(A \cap B) = ۰/۱,$$

$$P(A' \cap B) = P(B) - P(A \cap B) = ۰/۶ - ۰/۱ = ۰/۵,$$

$$P(A' \cap B') = ۱ - P(A \cup B) = ۱ - ۰/۵ = ۰/۴,$$

$$P(A' \cup B') = ۱ - P(A \cap B) = ۱ - ۰/۱ = ۰/۹.$$

۱۲-۳

$$P(A \Delta B) = (A \cap B') \cup (B \cap A') = P(A - B) + P(B - A).$$

۱۳-۳. الف)

$$P(A \cap B); \quad i = ۲ \quad P(A \Delta B); \quad i = ۱ \quad ۱ - P(A \cup B); \quad i = ۰$$

(ب)

$$۱ - P(A \Delta B); \quad i = ۲ \quad P(A \cup B); \quad i = ۱ \quad ۱ - P(A \cup B); \quad i = ۰$$

(ج)

$$۱ - P(A \cup B) + P(A \Delta B); \quad i = ۱ \quad ۱ - P(A \cup B); \quad i = ۰$$

$$P(A \cap B) + P(A \Delta B) + ۱ - P(A \cup B); \quad i = ۲$$

۱۴-۳

$$A = \{۲, ۴, ۶, ۸, ۱۰, ۱۲, ۱۴, ۱۶, ۱۸, ۲۰, ۲۲, ۲۴\}, \quad B = \{۳, ۶, ۹, ۱۲, ۱۵, ۱۸, ۲۱, ۲۴\}$$

$$\#(A \cap B) = ۴$$

$$P(A \cap B) = \frac{۴}{۲۴} = \frac{۱}{۶}.$$

۱۵-۳. الف)

$$P(A \cup B) = P(A) + P(B) - P(A \cap B).$$

پس

$$P(A \cap B) = P(A) + P(B) - P(A \cup B) \leq P(A) + P(B),$$

یا

$$P(A \cap B) \leq P(A) + P(B).$$

(ب)

$$A \cup B \subseteq S$$

$$P(S) \geq P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

$$\geq P(A) + P(B) - P(A \cap B)$$

$$P(A \cap B) \geq P(A) + P(B) - 1.$$

۱۶-۳. برای $n=1$

$$P(E_1) \leq P(E_p).$$

فرض می‌کنیم برای $n=k+1$ ثابت می‌کنیم $P(E_1) \leq P(E_p)$ برقرار است پس برای $n=k+1$

$$\begin{aligned} P\left(\bigcup_{i=1}^{k+1} E_i\right) &= P\left(\bigcup_{i=1}^{k+1} E_i \cup E_{k+1}\right) \\ &= P\left(\bigcup_{i=1}^k E_i\right) + P(E_{k+1}) - P\left(E_{k+1} \cap \bigcup_{i=1}^k E_i\right) \\ &\leq \sum_{i=1}^k P(E_i) + P(E_{k+1}) = \sum_{i=1}^{k+1} P(E_i). \end{aligned}$$

$$\frac{P(A)}{1-P(A)} = \frac{a}{b}. \quad ۱۷-۳$$

$$bP(A) = a - P(A)$$

$$(a+b)P(A) = a$$

$$P = P(A) = \frac{a}{a+b}$$

۱۸-۳. الف) چون هیچ‌گاه مقدار احتمال منفی نیست.

ب) چون مجموع احتمال‌ها روی کلیه‌ی موارد از یک بیشتر است.

$$0/12 + 0/25 + 0/36 + 0/14 + 0/09 + 0/07 = 1/03 > 1$$

$$0/08 + 0/21 + 0/29 + 0/40 = 0/98 \neq 1 \quad \text{ج)$$

.۱۹-۳

$$0/29 + 0/34 + 0/17 = 0/8 \quad \text{ب)} \quad 0/17 + 0/12 = 0/29 \quad \text{الف)}$$

$$0/08 + 0/29 + 0/34 = 0/71 \quad \text{د)} \quad 0/34 + 0/17 + 0/12 = 0/63 \quad \text{ج)}$$

.۲۰-۳

$$0/15 + 0/03 + 0/22 = 0/4 \quad \text{ب)} \quad 0/24 + 0/22 = 0/46 \quad \text{الف)}$$

$$0/15 + 0/03 + 0/28 + 0/22 = 0/68 \quad \text{د)} \quad 0/03 + 0/08 = 0/11 \quad \text{ج)}$$

.۲۱-۳

$$\frac{4 \times 5}{80} = \frac{1}{4} \quad \text{ب)} \quad \frac{10 + 20}{80} = \frac{30}{80} \quad \text{الف)}$$

$$\frac{4 + 2 + 1 + 1}{80} = \frac{1}{10} \quad \text{د)} \quad \frac{2 \times 4}{80} = \frac{1}{10} \quad \text{ج)}$$

$$\frac{8 + 14}{40} = \frac{22}{40} = \frac{11}{20} \quad \text{ه)}$$

۲۲-۳. اگر غیبت دستیار مسن را با A و جوان را با B نشان دهیم، داریم:

$$P(A' \cup B') = P(A') + P(B') - P(A' \cap B') = 0/92 + 0/95 - 0/98 = 0/95 \quad \text{الف)}$$

$$P(A \cup B) = P(A) + P(B) - P(AB) = 0/08 + 0/05 - 0/02 = 0/11 \quad \text{ب)}$$

$$P(A \Delta B) = P(A) + P(B) - 2P(A \cap B) = 0/08 + 0/05 - 2 \times 0/02 = 0/09 \quad \text{ج)}$$

$$P(A | B) = \frac{P(A \cap B)}{P(B)} \geq 0, \quad A, B \quad \text{الف) برای هر}$$

$$P(B | B) = \frac{P(B \cap B)}{P(B)} = \frac{P(B)}{P(B)} = 1 \quad \text{ب)}$$

(ج)

$$\begin{aligned} \frac{P((A_1 \cup A_2 \cup \dots) \cap B)}{P(B)} &= \frac{P((A_1 \cap B) \cup (A_2 \cap B) \cup \dots)}{P(B)} \\ &= \frac{P(A_1 \cap B)}{P(B)} + \frac{P(A_2 \cap B)}{P(B)} + \dots \\ &= P(A_1|B) + P(A_2|B) + \dots \end{aligned}$$

٢٤-٣ .الف) $B = \{1, 2, 5\}$ ، $A' = \{3, 4, 6\}$ ، $A = \{1, 2, 3\}$ ، $S = \{1, 2, 3, 4, 5, 6\}$

$$P(B|A) + P(B|A') = \frac{\frac{1}{6}}{\frac{3}{6}} + \frac{\frac{1}{6}}{\frac{3}{6}} = \frac{1}{3} + \frac{1}{3} = 1.$$

ب) بدون تغییر دادن A و S ، قرار دهید $\{B\}$ در این صورت

$$P(B|A) + P(B|A') = \frac{1}{3} + 0 = \frac{1}{3}.$$

.٢٥-٣

$$\begin{aligned} P(A)P(A|B)P(C|A \cap B) &= P(A) \frac{P(A \cap B)}{P(A)} \frac{P(A \cap B \cap C)}{P(A \cap B)} \\ &= P(A \cap B \cap C). \end{aligned}$$

.٢٦-٣

$$\begin{aligned} \frac{P(B \cap C)}{P(B)} &= P(C|B) = P(C|A \cap B) = \frac{P(A \cap B \cap C)}{P(A \cap B)} \\ \frac{P(B \cap C)}{P(B)} &= \frac{P(A \cap B \cap C)}{P(A \cap B)} \\ P(A|B \cap C) &= \frac{P(A \cap B \cap C)}{P(B \cap C)} = \frac{P(A \cap B)}{P(B)} = P(A|B). \end{aligned}$$

.٢٧-٣

$$P(T') = \frac{36 + 27}{90} = \frac{63}{90} \quad (\text{ب}) \quad P(G) = \frac{18 + 36}{90} = \frac{54}{90} \quad (\text{الف})$$

$$P(G' \cap T') = \frac{18}{90} \quad (\text{د}) \quad P(G \cap T) = \frac{18}{90} \quad (\text{ج})$$

$$P(G'|T') = \frac{18}{54} \quad (\text{و}) \quad P(T|G) = \frac{18}{54} \quad (\text{ه})$$

۲۸-۳

$$P(T|G) \frac{P(T \cap G)}{P(G)} = \frac{\frac{18}{90}}{\frac{54}{90}} = \frac{18}{54}, \quad P(G'|T') = \frac{P(T' \cap G')}{P(T')} = \frac{\frac{18}{90}}{\frac{63}{90}} = \frac{18}{63}$$

۲۹-۳. اگر $P(B|A) = P(B)$, آنگاه داریم:

$$\frac{P(A \cap B)}{P(A)} = P(B) \Rightarrow P(A \cap B) = P(A)P(B) \quad (1)$$

حال

$$P(A|B) = \frac{P(A \cap B)}{P(B)} \xrightarrow{(1)} \frac{P(A)P(B)}{P(B)} = P(A).$$

۳۱-۳. الف)

(از چهارتا انتخاب سه تا سفید باشد | مهره دوم و سوم سفید) P (از دو مهره اول یکی سفید و یکی سیاه باشد | مهره دوم و سوم سفید باشد) $= P$

$$= \frac{7}{10} \times \frac{6}{9}.$$

ب)

$$P(\text{اولی سفید باشد}) = \frac{8}{12} = (\text{از چهارتا انتخاب دو تا سفید باشد | اولی سفید باشد})$$

۳۲-۳

(از سه جعبه دو تا سفید انتخاب شده باشد | انتخاب از جعبه اول سفید) P (انتخاب از جعبه دوم سفید و انتخاب از جعبه سوم سیاه | انتخاب از جعبه اول سفید) $= P$ (انتخاب از جعبه دوم سیاه و انتخاب از جعبه سوم سفید | انتخاب از جعبه اول سفید) $+ P$

$$= \frac{3}{8} \times \frac{6}{10} \times \frac{4}{6} + \frac{3}{8} \times \frac{4}{10} \times \frac{2}{6}.$$

توجه داشته باشید که انتخاب جعبه‌ها از هم مستقل هستند.

۳۳-۳. اگر A از B مستقل باشد، داریم $P(A \cap B) = P(A)P(B)$ یا

$$P(A|B) = P(A)$$

(الف)

$$\begin{aligned}
 P(A'|B) &= \frac{P(A' \cap B)}{P(B)} = \frac{P(B - A)}{P(B)} \\
 &= \frac{P(B) - P(A \cap B)}{P(B)} = \frac{P(B)}{P(B)} - \frac{P(A \cap B)}{P(B)} \\
 &= 1 - \frac{P(A)P(B)}{P(B)} = 1 - P(A) = P(A').
 \end{aligned}$$

(ب)

$$\begin{aligned}
 P(A'|B') &= \frac{P(A' \cap B')}{P(B')} = \frac{1 - P(A \cup B)}{P(B')} = \frac{1 - P(A) - P(B) + P(A \cap B)}{P(B')} \\
 &= \frac{1 - P(A) - P(B) + P(A)P(B)}{P(B')} = \frac{P(B') - P(A)(1 - P(B))}{P(B')} \\
 &= \frac{P(B')(1 - P(A))}{P(B')} = 1 - P(A) = P(A').
 \end{aligned}$$

۳۵-۳. مطابق شکل ۴-۳ داریم: $P(C) = ۰/۵$ و $P(B) = ۰/۶$ ، $P(A) = ۰/۴$. پس

$$P(A \cap B \cap C) = ۰/۱۴ = ۰/۶ \times ۰/۴ \times ۰/۵ = P(A)P(B)P(C)$$

$$P(A \cap B) = ۰/۴۸ = ۰/۶ \times ۰/۴ = P(A)P(B)$$

از B مستقل است،

$$P(A \cap C) = ۰/۲۴ = ۰/۶ \times ۰/۵ = P(A)P(C)$$

از C مستقل است، و

$$P(B \cap C) = ۰/۳۸ \neq ۰/۴ \times ۰/۵ = P(B)P(C) = ۰/۲۴$$

از C مستقل نیست.

ولی در کل هر سه دو به دو مستقل نیستند.

۳۶-۳. در حل ۳۵-۳ مشاهده می‌کنیم که A از B مستقل است و لی B از C مستقل نیست.

۳۷-۳. در تمرین ۳۵-۳ نشان داده شد که A از B و A از C مستقل است اما

$$P(A \cap (B \cup C)) = ۰/۵۵ \neq ۰/۶ \times ۰/۹۴ = P(A)P(B \cup C).$$

پس A از $B \cup C$ مستقل نیستند.

۳۸-۳. وقتی که سه پیشامد A ، B و C مستقل هستند پس

$$P(A \cap B) = P(A)P(B) \quad (1)$$

$$P(A \cap C) = P(A)P(C) \quad (2)$$

$$P(B \cap C) = P(B)P(C) \quad (3)$$

$$P(A \cap B \cap C) = P(A)P(B)P(C) \quad (4)$$

(الف)

$$P(A \cap (B \cap C)) = P(A \cap B \cap C)$$

$$\xrightarrow{(r)} = P(A)P(B)P(C)$$

$$\xrightarrow{(r)} = P(B)P(B \cap C).$$

(ب)

$$\begin{aligned} P(A \cap (B \cup C)) &= P((A \cap B) \cup (A \cap C)) \\ &= P(A \cap B) + P(A \cap C) - P(A \cap B \cap C) \\ &= P(A)P(B) + P(A)P(C) - P(A)P(B)P(C) \\ &= P(A)(P(B) + P(C) - P(B \cap C)) \\ &= P(A)P(B \cup C). \end{aligned}$$

۳۹-۳. تعداد شرط‌هایی که دو پیشامد در آن شرکت دارند

تعداد شرط‌هایی که سه پیشامد در آن شرکت دارند

\vdots

تعداد شرط‌هایی که k پیشامد در آن شرکت دارند

$$\text{پس جواب است که حاصل آن برابر با } \sum_{i=r}^k \binom{k}{i} = M$$

$$\sum_{i=0}^k \binom{k}{i} = M + 1 + \binom{k}{r} \Rightarrow r^k - r - k = M.$$

$$P(\{T\}) = 0/48, P(\{H\}) = 0/52 \quad .40-3$$

$$P(\{HHH\}) = 0/52 \times 0/52 \times 0/52 = 0/144 \quad (\text{الف})$$

$$P(\{TTT\}) = 0/48 \times 0/48 \times 0/48 = 0/120 \quad (\text{ب})$$

$$\frac{5 \times 4 \times 3}{10 \times 9 \times 8} = \frac{\binom{5}{3}}{\binom{10}{3}} \quad .41-3$$

$$\frac{6}{15} \times \frac{5}{14} \times \frac{4}{13} \times \frac{3}{12} = \frac{360}{32760} = \frac{1}{91} \quad .42-3$$

$$0/9 \times 0/9 \times 0/9 = 0/729 \quad .43-3$$

$$0/4 \times 0/1 \times 0/4 + 0/4 \times 0/6 \times 0/4 = 0/16 + 0/144 = 0/160 \quad (\text{ب})$$

$$0/6 \times 0/6 \times 0/4 = 0/144 \quad (\text{پ})$$

$$P(A) = 0/4 \times 0/1 \times 0/4 + 0/6 \times 0/6 \times 0/4 = 0/016 + 0/144 = 0/16. \quad (\text{پیشامد مطلوب})$$

٤٤-٣. پیشامد معیوب بودن E و A_1, A_2, A_3 به ترتیب خطوط تولیدی ١ و ٢ و ٣ است.

$$P(A_1) = 0/4, \quad P(A_2) = 0/3, \quad P(A_3) = 0/5,$$

$$P(E | A_1) = 0/02, \quad P(E | A_2) = 0/03, \quad P(E | A_3) = 0/04,$$

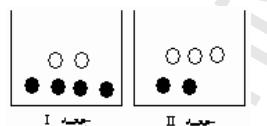
$$P(E) = \sum_{i=1}^3 P(A_i)P(E | A_i) = 0/4 \times 0/02 + 0/3 \times 0/03 + 0/5 \times 0/04 = 0/027. \quad .45-3$$

$$P(E | A_2) = \frac{P(A_2)P(E | A_2)}{P(E)} = \frac{0/3 \times 0/03}{0/027} = \frac{0/009}{0/027} = \frac{1}{3}. \quad .46-3$$



$$\begin{aligned}
 A: & \text{ انتخاب مهره‌ی سفید از جعبه I} \quad P(A) = \frac{۳}{۸} \\
 A': & \text{ انتخاب مهره‌ی سفید از جعبه I} \quad P(A') = \frac{۵}{۸} \\
 B: & \text{ انتخاب مهره‌ی سفید از جعبه I} \quad P(B | A) = \frac{۲}{۷} \\
 B': & \text{ انتخاب مهره‌ی سفید از جعبه I} \quad P(B | A') = \frac{۶}{۷} \\
 P(B) &= P(A)(B | A) + P(A')P(B | A') = \frac{۳}{۸} \times \frac{۲}{۷} + \frac{۵}{۸} \times \frac{۶}{۷} = \frac{۱۱}{۲۸} \quad (\text{الف}) \\
 P(A | B) &= \frac{P(A)P(B | A)}{P(B)} = \frac{\frac{۳}{۸} \times \frac{۲}{۷}}{\frac{۱۱}{۲۸}} = \frac{۶}{۱۱} \quad (\text{ب})
 \end{aligned}$$

۴۷-۳. الف)



$$\begin{aligned}
 P(A_1) &= \frac{۱}{۵} && \text{پیشامد انتخاب جعبه‌ی I} \quad A_1 \\
 P(A_r) &= \frac{۴}{۵} && \text{پیشامد انتخاب جعبه‌ی II} \quad A_r \\
 P(B) &= \text{پیشامد انتخاب دو مهره‌ی سفید}
 \end{aligned}$$

احتمال انتخاب هر جعبه $\frac{۱}{۲}$ است پس

$$\begin{aligned}
 P(B) &= P(A_1)P(B | A_1) + P(A_r)P(B | A_r) \\
 &= \frac{۱}{۵} \times \frac{۲}{۵} \times \frac{۱}{۵} + \frac{۴}{۵} \times \frac{۳}{۵} \times \frac{۲}{۴} = \frac{۲}{۱۲۵} + \frac{۶}{۱۰۰} = \frac{۲۲}{۱۲۵}. \quad (\text{ب})
 \end{aligned}$$

$$P(A_r | B) = \frac{P(A_r)P(B | A_r)}{P(B)} = \frac{\frac{۴}{۵} \times \frac{۳}{۵} \times \frac{۲}{۴}}{\frac{۲۲}{۱۲۵}} = \frac{۱۸}{۲۲}.$$

حل تمرین‌های فصل چهارم

۴-۱. برای اینکه این تابع یک تابع احتمال باشد باید $0 \leq f(x) \leq 1$ که در آن صورت

$$\text{حتماً باید مقدار } c \geq 0 \text{ و نیز از طرفی } 1 = \sum_x f(x)$$

$$\sum_x f(x) = c(1)^x + c(2)^x + c(3)^x = 1 \quad \forall c$$

و که نتیجه می‌دهد که باید $c = \frac{1}{14}$ و $f(x)$ یک تابع احتمال خواهد بود.

۴-۲. شرط اول $0 \leq f(x) \leq 1$ به ازای x برقرار است و برای شرط دوّم

$$\sum_x f(x) = \sum_x c\left(\frac{1}{6}\right)^{x-1} = c \sum_x \left(\frac{1}{6}\right)^{x-1} = c \frac{1}{1 - \frac{1}{6}} = c \frac{6}{5} = \frac{6}{5}c$$

به ازای $c = \frac{5}{6}$ تابع $f(x)$ تابع احتمال خواهد بود.

۴-۳. الف) فضای نمونه‌ی آزمایش عبارت است از

$$S = \{HHH, HHT, HTH, THH, TTH, THT, HTT, TTT\}.$$

با توجه به تعریف X ، متغیر X مقادیر $1, 2, 3, 4, 5$ را می‌گیرد و تابع احتمال آن به صورت زیر است.

| | | | | |
|--------|---------------|---------------|---------------|---------------|
| x | ۰ | ۱ | ۲ | ۳ |
| $f(x)$ | $\frac{1}{8}$ | $\frac{3}{8}$ | $\frac{3}{8}$ | $\frac{1}{8}$ |
| | $\frac{1}{8}$ | $\frac{3}{8}$ | $\frac{3}{8}$ | $\frac{1}{8}$ |

ب) فضای نمونه‌ای 36 عضو دارد و طبق تعریف متغیر تصادفی X مقادیری که می‌گیرد عبارتند از $1, 2, 3, 4, 5$ و تابع احتمال آن به صورت زیر است:

| | | | | | | |
|--------|----------------|-----------------|----------------|----------------|----------------|----------------|
| x | ۰ | ۱ | ۲ | ۳ | ۴ | ۵ |
| $f(x)$ | $\frac{1}{36}$ | $\frac{10}{36}$ | $\frac{8}{36}$ | $\frac{6}{36}$ | $\frac{4}{36}$ | $\frac{2}{36}$ |
| | $\frac{1}{36}$ | $\frac{10}{36}$ | $\frac{8}{36}$ | $\frac{6}{36}$ | $\frac{4}{36}$ | $\frac{2}{36}$ |

ج) مقادیری که X می‌گیرد عبارتند از $1, 2, \dots$ و تابع احتمال آن عبارت است از

| | | | | | |
|--------|-------------------|-------------------|-------------------|-----|-------------------|
| x | ۱ | ۲ | ۳ | ... | x |
| $f(x)$ | $(\frac{1}{2})^1$ | $(\frac{1}{2})^2$ | $(\frac{1}{2})^3$ | ... | $(\frac{1}{2})^x$ |

به طور کلی

$$f(x) = \left(\frac{1}{\mu}\right)^x, \quad x = 1, 2, \dots$$

د) شبیه به حالت ج است با این تفاوت که احتمال موفقیت (ظاهر شدن دو خال ۶)

برابر با $\frac{1}{3^6}$ است.

$$f(x) = \left(\frac{1}{3^6}\right)\left(\frac{3^5}{3^6}\right)^{x-1}, \quad x = 1, 2, \dots$$

۴-۴. الف) مقادیری که X می‌گیرد ۲، ۳، ۴، ۵ هستند وتابع احتمال آن به صورت

| | | | | |
|--------|----------------|----------------|----------------|----------------|
| x | ۲ | ۳ | ۴ | ۵ |
| $f(x)$ | $\frac{2}{20}$ | $\frac{4}{20}$ | $\frac{6}{20}$ | $\frac{8}{20}$ |

است. به عنوان مثال برای $x = 2$ ، یک وضعیت آن است که مهره‌ی اول ۲ و مهره‌ی

دوم، یک مشاهده شود که احتمال این پیشامد $\frac{1}{5} \times \frac{1}{4}$ است ولی می‌تواند مهره‌ی اول ۱

و دومی ۲ باشد که در کل احتمال $= \frac{1}{5} \times \frac{1}{4} = \frac{2}{20}$ است.

| | | | | |
|--------|----------------|----------------|----------------|----------------|
| x | ۱ | ۲ | ۳ | ۴ |
| $f(x)$ | $\frac{8}{20}$ | $\frac{6}{20}$ | $\frac{4}{20}$ | $\frac{2}{20}$ |

ب)

| | | | | | | | |
|--------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|
| x | ۳ | ۴ | ۵ | ۶ | ۷ | ۸ | ۹ |
| $f(x)$ | $\frac{2}{20}$ | $\frac{2}{20}$ | $\frac{4}{20}$ | $\frac{3}{20}$ | $\frac{4}{20}$ | $\frac{2}{20}$ | $\frac{2}{20}$ |

ج)

| | | | | | | |
|--------|-----------------|-----------------|----------------|----------------|----------------|----------------|
| x | ۰ | ۱ | ۲ | ۳ | ۴ | ۵ |
| $f(x)$ | $\frac{12}{42}$ | $\frac{10}{42}$ | $\frac{8}{42}$ | $\frac{6}{42}$ | $\frac{4}{42}$ | $\frac{2}{42}$ |

۵-۴

| | | | | | |
|--------|----------------|-----------------|-----------------|-----------------|----------------|
| z | ۰ | ۱ | ۲ | ۳ | ۴ |
| $f(z)$ | $\frac{3}{90}$ | $\frac{21}{90}$ | $\frac{42}{90}$ | $\frac{21}{90}$ | $\frac{3}{90}$ |

۶-۴

| | | | |
|--------|----------------|----------------|----------------|
| y | ۰ | ۱ | ۲ |
| $f(y)$ | $\frac{3}{15}$ | $\frac{9}{15}$ | $\frac{3}{15}$ |

| | | | |
|--------|---------------|---------------|---------------|
| x | ۰ | ۱ | ۲ |
| $f(x)$ | $\frac{1}{6}$ | $\frac{4}{6}$ | $\frac{1}{6}$ |

۶-۷. با جایگذاری

| | | | |
|--------|----------------|-----------------|----------------|
| x | ۰ | ۱ | ۲ |
| $f(x)$ | $\frac{9}{25}$ | $\frac{12}{25}$ | $\frac{4}{25}$ |

| | | | |
|--------|----------------|----------------|----------------|
| y | ۰ | ۱ | ۲ |
| $f(y)$ | $\frac{3}{10}$ | $\frac{5}{10}$ | $\frac{1}{10}$ |

$$1 = C \sum_{x=1}^{\infty} \frac{\theta^x}{x} \quad \text{A-4}$$

$$S_x = \{r_1, r_2, \dots\} \quad .9-4$$

$$c = \frac{1}{\sum_{x=1}^N r^x} = \frac{1}{\frac{1}{r} - \left(\frac{1}{r}\right)^{N+1}} \quad .10-4$$

$$F(x) = \begin{cases} \sum_{i=1}^{[x]} r^{-i-1} & 1 \leq x < N \\ 1 & x \geq N \end{cases} \quad .11-4$$

$$\begin{aligned} P(X \geq 2) &= 0/3, & P(0 \leq X \leq 2) &= 0/6 \\ P(|X - 1| \leq 1) &= P(-1 \leq X - 1 \leq 1) = P(0 \leq X \leq 2) = 0/9. \end{aligned}$$

۱۲-۴. اگر این چهار عدد را با $X_{(1)}, X_{(2)}, X_{(3)}, X_{(4)}$ نشان دهیم به طوری که $X_{(1)} < X_{(2)} < X_{(3)} < X_{(4)}$ ، آنگاه حالت‌های انتخاب ۴ از بین ۹ عدد برابر $\binom{9}{4} = 126$ که کل حالت‌های ممکن است. از طرفی حالت‌های مساعد آن‌هایی هستند که برای آنها $X_{(2)} = 4$ و یا $X_{(1)} < 4 < X_{(3)}$ برای $X_{(4)}$ فقط ۳ حالت (یا عدد) ۱, ۲, ۳ ممکن است و برای $X_{(3)}$ و $X_{(4)}$ اگر $5 = 5$ ، مقدار $X_{(4)}$ چهار حالت مساعد خواهد داشت. وقتی مقدار $X_{(3)} = 6$ ، $X_{(4)} = 7$ فقط ۳ حالت مساعد خواهد داشت و وقتی که $X_{(3)} = 8$ ، $X_{(4)} = 9$ فقط یک حالت مساعد خواهد داشت. در کل

$$X_{(1)} = \text{حالت‌های مساعد } 3$$

$$X_{(3)} = \text{حالت‌های مساعد } X_{(4)} \text{ و } \binom{5}{2} = \frac{5!}{2!3!} = \frac{5 \times 4}{2} = 10$$

$$30 = 10 \times 3 \text{ تعداد حالت‌های مساعد}$$

پس احتمال مورد نظر $\frac{30}{126}$ است.

۱۳-۴. B : پیشامد مشاهده‌ی مهره‌ی سفید

$$P(X = \omega | B) = \frac{P(X = \omega).P(B | X = \omega)}{P(B)}$$

اما

$$P(B) = P(X = \omega)P(B | X = \omega) + P(X = \mu)P(B | X = \mu)$$

$$= \frac{1}{3} \left(\frac{2}{5} \right) + \frac{2}{3} \left(\frac{1}{5} \right) = \frac{1}{15}$$

پس

$$P(X = \omega | B) = \frac{\frac{2}{3} \times \frac{1}{5}}{\frac{1}{15}} = \frac{2}{18}.$$

٤-١٤. الف)

| | | | | | |
|--------|----------------|----------------|----------------|----------------|----------------|
| x | ٠ | ١ | $2/5$ | 4 | ٥ |
| $f(x)$ | $\frac{2}{15}$ | $\frac{4}{15}$ | $\frac{2}{15}$ | $\frac{1}{15}$ | $\frac{3}{15}$ |

$$P(1 \leq X < 2/5) = \frac{4}{15}$$

$$P(0 < X \leq 2/5) = \frac{10}{15} \quad (ج)$$

$$P(0 < X < 2/5) = \frac{4}{15} \quad (د)$$

$$P(1 \leq X \leq 2/5) = \frac{1}{15} \quad (ه)$$

٤-١٥. دو شرط توزیع احتمال

الف) $f(x) \leq 0$ که برقرار است و ب می‌بایست $\sum_x f(x) = 1$

$$\sum_x f(x) = \sum_{x=1}^k \frac{px}{k(k+1)} = \frac{p}{k(k+1)} \sum_{x=1}^k x = \frac{p}{k(k+1)} \times \frac{k(k+1)}{2} = 1.$$

.١٦-٤

$$\sum_{x=1}^{\infty} f(x) = \sum_{x=1}^{\infty} \frac{c}{x} = c \sum_{x=1}^{\infty} \frac{1}{x}$$

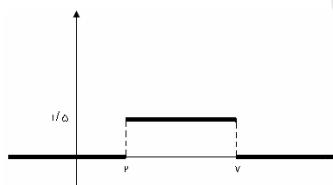
اماً $\sum_x f(x) = 1$ بدهست یک سری واگر است. پس مقدار c هایی که از رابطه $\sum_{x=1}^{\infty} \frac{1}{x}$ می آید، وجود ندارد.

۱۸-۴. الف) $F(4) = 1/2 > 1$ که با شرط (الف) تناقض دارد.

$$x_1 \leq x_2 \Rightarrow F(x_1) \leq F(x_2)$$

ج) کلیه‌ی ویژگی‌هایتابع توزیع احتمال را دارا است.

۱۹-۴. الف)



و

$$\int_a^b f(x)dx = \int_2^7 \frac{1}{5} dx = \frac{1}{5}(7 - 2) = 1$$

$$P(3 < x < 5) = \int_3^5 f(x)dx = \int_3^5 \frac{1}{5} dx = \frac{5-3}{5} = \frac{2}{5}.$$

.۲۰-۴

$$\begin{aligned} P(Y < 3/2) &= \int_2^{3/2} \frac{1}{8} (y+1) dy \\ &= \left[\frac{1}{8} \left(\frac{y^2}{2} + y \right) \right]_2^{3/2} = \frac{1}{8} (8/32 - 4) = 0/54. \end{aligned}$$

$$P(9/4 < Y < 3/2) = P(0 < Y < 2) + P(9/4 < Y < 3/2)$$

$$= 0 + \int_2^{3/2} \frac{1}{8} (y+1) dy = 0/54.$$

۲۱-۴. الف)

$$\int_0^4 f(x)dx = \int_0^4 \frac{c}{\sqrt{x}} dx = c \int_0^4 (x)^{-1/2} dx = 2c\sqrt{x} \Big|_0^4 = 4c = 1 \Rightarrow c = \frac{1}{4}.$$

ب)

$$P(X > 1) = \int_{1}^{\infty} \frac{1}{\pi \sqrt{x}} dx = \frac{\sqrt{x}}{\pi} \Big|_1^{\infty} = \left(1 - \frac{1}{\pi}\right) = \frac{1}{\pi}$$

$$P(X < \frac{1}{\pi}) = \int_{0}^{\frac{1}{\pi}} \frac{1}{\pi \sqrt{x}} dx = \left(\frac{\sqrt{x}}{\pi} \right) \Big|_0^{\frac{1}{\pi}} = \frac{1}{\pi}.$$

٤-٢٢. با توجه به تغییر متغیر

$$u = -z^{\pi} \Rightarrow du = -\pi z dz$$

داریم

$$1 = \int_{0}^{\pm\infty} f(z) dz = k \int_{0}^{\pm\infty} z e^{-z^{\pi}} dz = \left(-\frac{k}{\pi} e^{-z^{\pi}} \right) \Big|_0^{\pm\infty} = \frac{k}{\pi} (-e^{-\infty} + 1),$$

$$k = \pi.$$

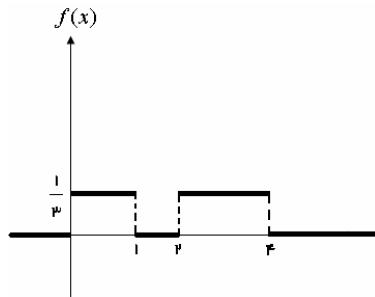
.٢٣-٤

$$\begin{aligned} P(X < \frac{1}{\pi}) &= \int_{0}^{\frac{1}{\pi}} \pi x(1-x) dx = \pi \int_{0}^{\frac{1}{\pi}} (x - x^{\pi}) dx \\ &= \left(\pi \left(\frac{x^{\pi}}{\pi} - \frac{x^{\pi}}{\pi^{\pi}} \right) \right) \Big|_0^{\frac{1}{\pi}} = \pi \left[\frac{1}{\pi^{\pi}} - \frac{1}{\pi \times \pi^{\pi}} \right] = \pi \left(\frac{\Delta}{\pi^{\pi}} \right) = \frac{\pi^{\pi}}{\pi^{\pi}}. \end{aligned}$$

$$\begin{aligned} P(X > \frac{1}{\pi}) &= 1 - P(X < \frac{1}{\pi}) = \int_{0}^{\frac{1}{\pi}} \pi x(1-x^{\pi}) dx \\ &= \left(\pi \left(\frac{x^{\pi}}{\pi} - \frac{x^{\pi}}{\pi^{\pi}} \right) \right) \Big|_0^{\frac{1}{\pi}} = \pi \left(\frac{1}{\pi} - \frac{1}{\pi^{\pi}} \right) = \pi \left(\frac{1}{\pi^{\pi}} \right) = \frac{1}{\pi}. \end{aligned}$$

.٢٤-٤

$$F(x) = \begin{cases} 0 & x < 0 \\ \frac{x}{\pi} & 0 \leq x < 1 \\ \frac{1}{\pi} & 1 \leq x < \pi \\ \frac{1}{\pi} + \frac{x-\pi}{\pi} & \pi \leq x < \pi \\ 1 & \pi \leq x \end{cases}$$



.٢٥-٤

$$f(x) = \begin{cases} 0 & x < 0 \\ \frac{x^2}{2} & 0 \leq x < 1 \\ \frac{1}{2} + \int_1^x (Q-t) dt & 1 \leq x < 2 \\ 1 & 2 \leq x \end{cases}$$

.٢٦-٤

$$F(x) = \begin{cases} 0 & x < 0 \\ \frac{x^2}{4} & 0 \leq x < 1 \\ \frac{1}{4} + \frac{(x-1)}{2} & 1 \leq x < 2 \\ \frac{5}{4} + \frac{(3-x)^2}{2} & 2 \leq x < 3 \\ 1 & 3 \leq x \end{cases}$$

.٢٧-٤

$$f(x) = \frac{1}{2}, \quad 0 \leq x < 1$$

$$P\left(-\frac{1}{2} < X < \frac{1}{2}\right) = \frac{1}{2}, \quad P(Q < X < 3) = 0.$$

.٢٨-٤

$$f(y) = \frac{18}{y^3} \Rightarrow P(Y < 5) = \frac{144}{225}, \quad P(Y > 8) = \frac{495}{576}.$$

.٢٩-٤

$$P(X < 2) = F(2) - F(0) = 1 - 3e^{-2},$$

$$P(1 < X < 3) = F(3) - F(1) = 2e^{-1} - 4e^{-3},$$

$$\begin{aligned} P(X > r) &= 1 - P(X < r) = 1 - P(0 < X < r) = 1 - (F(r) - F(0)) \\ &= 1 - (1 - \alpha e^{-r}) = \alpha e^{-r}. \end{aligned}$$

٤-٣٠. الف)

$$\begin{aligned} P(-s < X < -t) &= \int_{-s}^{-t} \frac{1}{228} (3s - x^3) dx = \left[\frac{1}{228} \left(3sx - \frac{x^4}{4} \right) \right]_{-s}^{-t} \\ &= \frac{1}{228} \left[3s(-t) - \frac{t^4}{4} + 3s(-s) - \frac{s^4}{4} \right] = \frac{5s^4 - t^4}{684}. \end{aligned}$$

(ب)

$$\begin{aligned} P(1 < X < s) &= \int_1^s \frac{1}{228} (3s - x^3) dx = \left[\frac{1}{228} \left(3sx - \frac{x^4}{4} \right) \right]_1^s \\ &= \frac{1}{228} \left[3s(s) - \frac{s^4}{4} + 3s(1) - \frac{1^4}{4} \right] = \frac{5s^4 - 1^4}{684}. \end{aligned}$$

(ج)

$$\begin{aligned} P(-m < X < -1) &= \int_{-1}^{-m} \frac{1}{228} (3s - x^3) dx = \left[\frac{1}{228} \left(3sx - \frac{x^4}{4} \right) \right]_{-1}^{-m} \\ &= \frac{1}{228} \left[-3s(-m) + \frac{1}{4} + 3s(-1) - \frac{1^4}{4} \right] = \frac{190}{648}. \end{aligned}$$

P(X = \omega) = 0) د)

٤-٣٢. الف)

$$P(X < 1000) = P(0 < X < 1000) = \int_0^{1000} \frac{1}{\mu_0} e^{-\frac{x}{\mu_0}} dx = \left[e^{-\frac{x}{\mu_0}} \right]_0^{1000} = 1 - e^{-\frac{1000}{\mu_0}}.$$

(ب)

$$P(27000 < X < 36000) = \left[-e^{-\frac{x}{\mu_0}} \right]_{27000}^{36000} = -e^{-\frac{36000}{\mu_0}} + e^{-\frac{27000}{\mu_0}}.$$

(ج)

$$P(\infty < X < \infty) = 1 - P(0 < X < \infty) = \left(1 - e^{-\frac{x}{\lambda}}\right)^{\infty} \\ = 1 - (1 - e^{-\lambda}) = e^{-\lambda}.$$

٤-٣٣. الف)

$$P(X > 1) = 1 - P(0 < X < 1) = 1 - (F(1) - F(0)) \\ = 1 - \left[1 - \frac{e^{-\lambda}}{1} - 1 + \frac{e^{-\lambda}}{2}\right] = 1/2.$$

$$P(0 < X < \lambda) = F(\lambda) - F(0) = 1 - \frac{e^{-\lambda}}{1} - 1 + \frac{e^{-\lambda}}{2} = \frac{e^{-\lambda}}{2}. \quad (\text{ب})$$

$$P(1 < X < 1.5) = F(1.5) - F(1) = 1 - \frac{e^{-1.5}}{1} - 1 + \frac{e^{-1.5}}{2} = \frac{e^{-1.5}}{2}. \quad (\text{ج})$$

٤-٣٥

$$F(1,2) = f(1,1) + f(1,2) = \frac{1}{1} + \frac{1}{2} = \frac{1}{1}.$$

٤-٣٦

$$P\left(\frac{1}{2} < X < \frac{1}{2}, \frac{1}{2} < Y < \frac{1}{2}\right) = \int_{\frac{1}{2}}^{\frac{1}{2}} \int_{\frac{1}{2}}^{\frac{1}{2}} dy dx = \frac{1}{12}.$$

٤-٣٧

$$\int_0^1 \int_0^y \frac{k}{y} dx dy = \int_0^1 \left(\left(\frac{x}{y} k \right) \right)_0^y dy = \int_0^1 k dy = (ky)_0^1 = 1 \Rightarrow k = 1.$$

٤-٣٨

$$f(x) = \sum_{y=1}^x f(x,y) = \frac{xy+1}{1} + \frac{xy+2}{1} + \frac{xy+3}{1} = \frac{9x+1}{1}, \quad x = 1,2$$

$$f(y) = \sum_{x=1}^y f(x,y) = \frac{xy+1}{1} + \frac{xy+2}{1} = \frac{4y+1}{1}, \quad y = 1,2$$

٤-٣٩

$$f(x) = \int_y^1 f(x,y) dy = \int_y^{1-x} dy = y \Big|_y^{1-x} = 1-x, \quad 0 < x < 1.$$

۲۹۹ حل تمرین‌ها

$$f(y) = \int_x^1 f(x, y) dx = \int_x^1 y dx = yx \Big|_x^1 = y(1 - x), \quad 0 < y < 1.$$

.۴۱-۴

| x | y | ۱ | ۲ | ۳ | $f(x)$ |
|--------|---------------|---------------|---------------|---|---------------|
| ۱ | $\frac{1}{9}$ | ۰ | ۰ | ۰ | $\frac{1}{9}$ |
| ۲ | $\frac{2}{9}$ | $\frac{1}{9}$ | ۰ | ۰ | $\frac{2}{9}$ |
| ۳ | $\frac{3}{9}$ | $\frac{2}{9}$ | $\frac{1}{9}$ | ۰ | $\frac{5}{9}$ |
| $f(y)$ | $\frac{5}{9}$ | $\frac{3}{9}$ | $\frac{1}{9}$ | ۰ | ۱ |

.۴۲-۴

$$P(a < X \leq b, c < Y \leq d) = F(b, d) - F(a, d) - F(b, c) + F(a, c).$$

۴-۳. شرط استقلال این است که برای هر x و y

$$f(x, y) = f(x)f(y)$$

$$f(x, y) = \frac{xy + y}{51}, \quad f_X(x) = \frac{9x + 14}{51}, \quad f_Y(y) = \frac{9 + 4y}{51}.$$

حال مثال نقض

$$f(1, 1) = \frac{5}{51}, \quad f_X(1) = \frac{14}{51}, \quad f_Y(1) = \frac{13}{51}$$

$$f_X(1)f_Y(1) = \frac{14}{51} \times \frac{13}{51} = \frac{182}{2601} \neq \frac{5}{51} = f(1, 1)$$

نشان می‌دهد که استقلال وجود ندارد.

۴-۴. الف) شرط استقلال سه متغیره این است که برای هر x_1, x_2, x_3

$$f(x_1, x_2, x_3) = f(x_1)f(x_2)f(x_3)$$

$$\begin{aligned} f(x) &= \sum_{z=1}^3 \sum_{y=1}^3 \frac{(x+y)z}{6^3} \\ &= \frac{x+1}{6^3} + \frac{x+2}{6^3} + \frac{x+3}{6^3} + \frac{2(x+1)}{6^3} + \frac{2(x+2)}{6^3} + \frac{2(x+3)}{6^3} = \frac{9x+18}{6^3}, \quad x = 1, 2 \end{aligned}$$

$$\begin{aligned}
 f(y) &= \sum_{z=1}^3 \sum_{y=1}^3 \frac{(x+y)z}{6^3} \\
 &= \frac{1+y}{6^3} + \frac{2+y}{6^3} + \frac{3+y}{6^3} + \frac{4+y}{6^3} = \frac{5y+9}{6^3}, \quad y=1,2,3 \\
 f(z) &= \sum_{x=1}^3 \sum_{y=1}^3 \frac{(x+y)z}{6^3} \\
 &= \frac{(1+1)z}{6^3} + \frac{2z}{6^3} + \frac{3z}{6^3} + \frac{4z}{6^3} + \frac{5z}{6^3} = \frac{15z}{6^3}, \quad z=1,2,3 \\
 f(1,1,1) &= \frac{1}{6^3} \neq f_X(x)f_Y(y)f_Z(z) = \frac{1}{6^3} \times \frac{1}{6^3} \times \frac{1}{6^3}.
 \end{aligned}$$

پس استقلال وجود ندارد.
(ب)

$$\begin{aligned}
 P(X = 1, Y + Z \leq 2) &= f(X = 1, Y = 1, Z = 1) + f(X = 1, Y = 1, Z = 2) \\
 &\quad + f(X = 1, Y = 2, Z = 1) \\
 &= \frac{1}{6^3} + \frac{1}{6^3} + \frac{1}{6^3} = \frac{1}{6^3}.
 \end{aligned}$$

.۴۵-۴

| x | y | ۱ | ۲ | ۳ | $f(x)$ |
|--------|-----|---------------|---------------|---------------|---------------|
| ۱ | ۱ | ۰ | $\frac{1}{6}$ | $\frac{1}{6}$ | $\frac{1}{6}$ |
| ۲ | ۱ | $\frac{1}{6}$ | ۰ | $\frac{1}{6}$ | $\frac{1}{6}$ |
| ۳ | ۱ | $\frac{1}{6}$ | $\frac{1}{6}$ | ۰ | $\frac{1}{6}$ |
| $f(y)$ | | $\frac{1}{6}$ | $\frac{1}{6}$ | $\frac{1}{6}$ | ۱ |

$$f_X(1)f_Y(1) = \frac{1}{6} \times \frac{1}{6} \neq 0 = f(1,1).$$

.۴۶-۴

$$\begin{aligned}
 f(Y = 1 | X = 1) &= 1 \\
 f(y|x=1) &= \frac{y}{6}
 \end{aligned}$$

٣٠١ حل تمرین‌ها

$$f(y|X) = \begin{array}{c|ccc} y|x=p & \circ & 1 & p \\ \hline f(y|p) & \frac{1}{\mu} & & \frac{1}{\mu} \\ & \frac{1}{\mu} & \circ & \frac{1}{\mu} \\ & \frac{1}{\mu} & & \end{array}$$

$$f(Y=0|X=p) = \frac{1}{\mu}, \quad f(Y=1|X=p) = \frac{1}{\mu}$$

$$f(Y=1|X=\mu) = 1 \quad f(y|\mu) = \begin{array}{c|ccc} y|x=\mu & \circ & 1 & \mu \\ \hline f(y|\mu) & \circ & \frac{1}{\mu} & \circ \\ & & \frac{1}{\mu} & \end{array}$$

حل تمرین‌های فصل پنجم

.۱-۵

$$\begin{aligned}
 E(aX + b) &= \sum_x (ax + b)f(x) = \sum_x axf(x) + \sum_x bf(x) \\
 &= a \sum_x xf(x) + b \sum_x f(x) \\
 &= aE(X) + b.
 \end{aligned}$$

.۲-۶

$$\begin{aligned}
 E\left(\sum_{i=1}^n c_i g_i(x)\right) &= \int_x \left(\sum_{i=1}^n c_i g_i(x)\right) f(x) dx = \sum_{i=1}^n \int_x c_i g_i(x) f(x) dx \\
 &= \sum_{i=1}^n c_i \int_x g_i(x) f(x) dx = \sum_{i=1}^n c_i E(g_i(X)).
 \end{aligned}$$

.۳-۷

$$E(X) = \sum_x xf(x) = -1 \times \frac{1}{4} + 0 \times \frac{1}{4} + 1 \times \frac{1}{4} + 2 \times \frac{1}{4} = \frac{1}{2}.$$

.۴-۸

$$E(Y) = \int_1^4 \frac{1}{3} y(1+y) dy = \frac{1}{3} \left(y + \frac{y^2}{2} \right) \Big|_1^4 = \frac{15}{2}.$$

.۵-۹

$$E(X) = \int_0^1 x^2 dx + \int_1^2 x(2-x) dx = \frac{x^3}{3} \Big|_0^1 + \left(x^2 - \frac{x^3}{3} \right) \Big|_1^2 = 1.$$

(۶-۱۰. الف)

$$E(X) = 0 \times \frac{1}{125} + 1 \times \frac{12}{125} + 2 \times \frac{48}{125} + 3 \times \frac{96}{125} = \frac{300}{125}.$$

$$E(X^2) = 0 \times \frac{1}{125} + 1 \times \frac{12}{125} + 2 \times \frac{48}{125} + 3 \times \frac{96}{125} = \frac{780}{125}.$$

(ب)

$$\begin{aligned} E(\mu X + \nu) &= E(\alpha X^\mu + \nu + \beta X) = \alpha E(X^\mu) + \nu + \beta E(X) \\ &= \alpha \times \frac{78}{125} + \nu + \beta \times \frac{30}{125} = \frac{111}{125}. \end{aligned}$$

(٧-الف)

$$\begin{aligned} E(X) &= \frac{1}{\ln \mu} \int_1^{\mu} \frac{1}{x} x dx = \frac{1}{\ln \mu} \int_1^{\mu} dx = \frac{1}{\ln \mu}, \\ E(X^\mu) &= \frac{1}{\ln \mu} \int_1^{\mu} \frac{1}{x} x^\mu dx = \frac{1}{\ln \mu} \int_1^{\mu} x dx \\ &= \left[\frac{1}{\ln \mu} \left(\frac{x^\mu}{\mu} \right) \right]_1^{\mu} = \frac{1}{\ln \mu} \left(\frac{\mu}{\mu} - \frac{1}{\mu} \right) = \frac{1}{\ln \mu}, \\ E(X^\nu) &= \frac{1}{\ln \mu} \int_1^{\mu} \frac{1}{x} x^\nu dx = \frac{1}{\ln \mu} \int_1^{\mu} x^\nu dx \\ &= \left[\frac{1}{\ln \mu} \left(\frac{x^\nu}{\nu} \right) \right]_1^{\mu} = \frac{1}{\ln \mu} \left(\frac{\mu}{\nu} - \frac{1}{\nu} \right) = \frac{\mu - 1}{\nu \ln \mu}. \end{aligned}$$

(ب)

$$\begin{aligned} E(X^\mu + \nu X^\mu - \mu X + \lambda) &= E(X^\mu) + \nu E(X^\mu) - \mu E(X) + \lambda \\ &= \frac{1}{\ln \mu} \left(\frac{\mu}{\mu} + \lambda - \mu \right) + \lambda = \frac{\lambda}{\mu \ln \mu} + \lambda. \end{aligned}$$

(جـ)

$$\begin{aligned} E(X) &= \int_0^1 x \frac{x}{\mu} dx + \int_1^{\mu} x \frac{x}{\mu} dx + \int_{\mu}^{\infty} x \frac{\mu - x}{\mu} dx \\ &= \left[\frac{x^2}{2\mu} \right]_0^1 + \left[\frac{x^2}{2\mu} \right]_1^{\mu} + \left[\frac{1}{\mu} \left(\frac{\mu x^2}{2} - \frac{x^2}{2} \right) \right]_{\mu}^{\infty} = \frac{1}{2\mu} + \frac{\mu}{2} + \frac{1}{\mu} = \frac{\mu + 1 + 2}{2\mu}, \\ E(X^\mu) &= \int_0^1 x^\mu \frac{x}{\mu} dx + \int_1^{\mu} x^\mu \frac{x}{\mu} dx + \int_{\mu}^{\infty} x^\mu \frac{\mu - x}{\mu} dx \\ &= \left[\frac{x^{\mu+1}}{\mu+1} \right]_0^1 + \left[\frac{x^{\mu+1}}{\mu+1} \right]_1^{\mu} + \left[\frac{1}{\mu} \left(x^{\mu+1} - \frac{x^{\mu+1}}{\mu+1} \right) \right]_{\mu}^{\infty} = \frac{1}{\mu+1} + \frac{\mu}{\mu+1} + \frac{1}{\mu} = \frac{\mu+2}{\mu}, \\ E(g(X)) &= \frac{\lambda}{\mu} - \omega \left(\frac{\mu \gamma}{1 \mu} \right) + \nu = -\frac{\mu \gamma}{\mu}. \end{aligned}$$

۴ آمار و احتمال

۹-۵. فرض کنید Z_i ها سود ناشی از پخت i تا کیک است.

| X | ۰ | ۱ | ۲ | ۳ | ۴ | ۵ | $E(Z_i)$ |
|--------|---------------|---------------|---------------|---------------|---------------|---------------|-----------------|
| $f(x)$ | $\frac{1}{6}$ | $\frac{1}{6}$ | $\frac{1}{6}$ | $\frac{1}{6}$ | $\frac{1}{6}$ | $\frac{1}{6}$ | |
| Z_1 | -۴ | ۱۰ | ۱۰ | ۱۰ | ۱۰ | ۱۰ | $\frac{۱۶}{۶}$ |
| Z_2 | -۸ | -۴+۱۰ | ۲۰ | ۲۰ | ۲۰ | ۲۰ | $\frac{۷۸}{۶}$ |
| Z_3 | -۱۲ | -۸+۱۰ | -۸+۲۰ | ۳۰ | ۳۰ | ۳۰ | $\frac{۹۶}{۶}$ |
| Z_4 | -۱۶ | -۱۲+۱۰ | -۸+۲۰ | -۴+۳۰ | ۴۰ | ۴۰ | $\frac{۱۰۰}{۶}$ |
| Z_5 | -۲۰ | -۱۶+۱۰ | -۱۲+۲۰ | -۸+۳۰ | -۴+۴۰ | ۵۰ | $\frac{۹۰}{۶}$ |

با توجه به مقادیر امید ریاضی سودها، با پختن ۴ کیک به ماکسیمم انتظار سود خواهیم رسید.

۱۰-۵

$$E(X) = \int_{-1}^{\Delta} x \frac{1}{18}(x+1) dx = \left(\frac{1}{18} \left(\frac{x^3}{3} + \frac{x^2}{2} \right) \right)_{-1}^{\Delta} = \frac{1}{18} \left(\frac{12\Delta}{3} + \frac{2\Delta}{2} + \frac{1}{3} - \frac{1}{2} \right) = \Delta.$$

۱۱-۵

$$E(X^p) = (-1)^p \times \frac{1}{V} + (0)^p \times \frac{1}{V} + (1)^p \times \frac{1}{V} + (2)^p \times \frac{1}{V} = \frac{1}{V},$$

$$\text{var}(X) = E(X^p) - (E(X))^p = \frac{1}{V} - \left(\frac{1}{V}\right)^p = \frac{9}{144}.$$

۱۲-۵

$$\begin{aligned} E(Y^p) &= \int_{-1}^{\Delta} y^p \frac{1}{18}(y+1) dy = \frac{1}{18} \int_{-1}^{\Delta} (y^p + y^{p+1}) dy = \left(\frac{1}{18} \left(\frac{y^{p+1}}{p+1} + \frac{y^{p+2}}{p+2} \right) \right)_{-1}^{\Delta} \\ &= \frac{1}{18} \left(\frac{\Delta^{p+1}}{p+1} + \frac{\Delta^{p+2}}{p+2} - \frac{(-1)^{p+1}}{p+1} - \frac{(-1)^{p+2}}{p+2} \right) = \frac{59}{144}, \end{aligned}$$

$$\text{var}(Y) = \frac{59}{144} - (1)^p = \frac{59}{144}.$$

۱۳-۵

$$\begin{aligned}
 E(X^r) &= \int_0^1 x^r x dx + \int_1^r x^r (r-x) dx \\
 &= \int_0^1 x^r dx + \int_1^r (rx^r - x^r) dx \\
 &= \left(\frac{x^{r+1}}{r+1} \right)_0^1 + \left(\left(\frac{rx^r}{r} - \frac{x^{r+1}}{r+1} \right) \right)_1^r = \frac{r}{r+1},
 \end{aligned}$$

$$\text{var}(X) = \frac{r}{r+1} - (1)^r = \frac{1}{r+1},$$

$$\text{var}(sX) = s^2 \text{var}(X) = r \times \frac{1}{r+1} = \frac{r}{r+1}.$$

.١٤-٥

$$\begin{aligned}
 \text{var}(X) &= E(X^r) - (E(X))^r = \frac{\frac{r(r+1)}{2} - \left(\frac{r(r+1)}{2}\right)^r}{r(r+1)} = \frac{r(r+1)(1-r^r)}{2(r(r+1))}, \\
 \text{var}(rX) &= r^2 \text{var}(X) = r^2 \times \frac{r(r+1)(1-r^r)}{2(r(r+1))} = \frac{r(r+1)(1-r^r)}{2}.
 \end{aligned}$$

.١٥-٦

$$\begin{aligned}
 E(X) &= \frac{r}{\ln r}, \quad E(X^r) = \frac{r}{\ln r}, \quad \text{var}(X) = \frac{r}{\ln r} - \left(\frac{r}{\ln r}\right)^r = \frac{r \ln r - r}{(\ln r)^r}, \\
 \text{var}(1-X) &= \text{var}(X) = \frac{r \ln r - r}{(\ln r)^r}.
 \end{aligned}$$

.١٦-٧

$$\begin{aligned}
 \mu &= E(X) = \frac{1}{r}(-r) + \frac{1}{r}(r) = 0, \quad \mu'_r = E(X^r) = \frac{1}{r}(-r)^r + \frac{1}{r}(r)^r = r, \\
 \sigma^r &= \text{var}(X) = \mu'_r - \mu^r = r - (0)^r = r.
 \end{aligned}$$

.١٧-٨

$$\begin{aligned}
 \mu &= E(X) = \int_0^r x \frac{x}{r} dx = \frac{1}{r} \int_0^r x^r dx = \left(\frac{x^{r+1}}{r+1} \right)_0^r = \frac{r}{r+1}, \\
 \mu'_r &= E(X^r) = \frac{1}{r} \int_0^r x^r dx = \left(\frac{1}{r} x^r \right)_0^r = r, \\
 \sigma^r &= \mu'_r - \mu^r = r - \left(\frac{r}{r+1} \right)^r = \frac{r}{r+1}.
 \end{aligned}$$

.۱۸-۵

$$E(Z) = E\left(\frac{X - \mu}{\sigma}\right) = \frac{1}{\sigma}E(X - \mu) = \frac{1}{\sigma}(E(X) - \mu) = \frac{1}{\sigma}(\mu - \mu) = 0,$$

$$\text{var}(Z) = \text{var}\left(\frac{X - \mu}{\sigma}\right) = \frac{1}{\sigma^2}\text{var}(X - \mu) = \frac{1}{\sigma^2}\text{var}(X) = \frac{1}{\sigma^2} \times \sigma^2 = 1.$$

.۱۹-۶

$$\mu = E(X) = \int_1^\infty rx x^{-r} dx = \int_1^\infty rx^{-r} dx = \left(-rx^{-1}\right)_1^{+\infty} = r,$$

$$E(X^r) = \int_1^\infty rx^r x^{-r} dx = \int_1^\infty rx^{-1} dx = (r \ln x)_1^{+\infty} = +\infty.$$

پس واریانس آن نیز موجود نمی باشد.

.۲۰-۶

$$M_X(t) = E(e^{tX}) = \sum_{x=1}^{\infty} r \left(\frac{1}{\mu}\right)^x e^{tx} = r \sum_{x=1}^{\infty} \left(\frac{e^t}{\mu}\right)^x = r \frac{\frac{e^t}{\mu}}{1 - \frac{e^t}{\mu}} = \frac{re^t}{\mu - e^t},$$

$$M'_X(0) = \left. \left(\frac{re^t(\mu - e^t) + e^t(re^t)}{(\mu - e^t)^2} \right) \right|_{t=0} = \left. \left(\frac{re^t}{(\mu - e^t)^2} \right) \right|_{t=0} = \frac{\mu}{r},$$

$$M''_X(0) = \left. \left(\frac{re^t(\mu - e^t)^2 + r(\mu - e^t)e^t(re^t)}{(\mu - e^t)^3} \right) \right|_{t=0} = \mu.$$

.۲۱-۶

$$M_X(t) = E(e^{tX}) = \int_0^1 e^{tx} dx = \left[\frac{e^{tx}}{t} \right]_0^1 = \frac{e^t - e^0}{t} = \frac{e^t - 1}{t},$$

$$M'_X(0) = \left. \left(\frac{te^t - (e^t - 1)}{t^2} \right) \right|_{t=0} = \left. \left(\frac{te^t - e^t + 1}{t^2} \right) \right|_{t=0} = \frac{1}{r}.$$

.۲۲-۶

$$\sigma^2 = \text{var}(X) = \sum_i (x_i - \mu)^2 p(x_i),$$

برای x_i هایی که خارج بازه‌ی $[\mu - k\sigma, \mu + k\sigma]$ می‌افتد، داریم:

$$\sigma^r \geq \sum_{i: |x_i - \mu| > k\sigma} (x_i - \mu)^r P(x_i),$$

حال اگر تمام $|x_i - \mu|$ را برابر $k\sigma$ بگیریم باز نابرابری بالا برقرار است.

$$\sigma^r \geq \sum_{i: |x_i - \mu| \geq k\sigma} P_i k^r \sigma^r \Leftrightarrow \sigma^r \geq k^r \sigma^r \sum_{i: |x_i - \mu| \geq k\sigma} P_i$$

اما $\sum_{i: |x_i - \mu| \geq k\sigma} P_i = P(|X - \mu| \geq k\sigma)$ پس خواهیم داشت:

$$\sigma^r \geq k^r \sigma^r P(|X - \mu| \geq k\sigma)$$

با حذف عامل σ^r از طرفین

$$P(|X - \mu| \geq k\sigma) \leq \frac{1}{k^r}$$

یا

$$P(|X - \mu| \leq k\sigma) \geq 1 - \frac{1}{k^r}.$$

۲۳-۵. چون X فقط مقادیر مثبت را اختیار می‌کند پس

$$E(X) = \int_0^\infty xf(x)dx = \int_0^a xf(x)dx + \int_a^\infty xf(x)dx \geq \int_a^\infty xf(x)dx.$$

حال اگر به جای متغیر X کوچکترین مقدار آن یعنی a را قرار دهیم باز هم نابرابری بالا برقرار است. پس

$$E(X) \geq \int_a^{+\infty} af(x)dx = a \int_a^{+\infty} f(x)dx = aP(X \geq a)$$

در نتیجه

$$P(X \geq a) \leq \frac{E(X)}{a}.$$

۲۴-۵. چون از نابرابری مارکف X فقط مقادیر مثبت را اختیار می‌کند پس با استفاده

از تابع قدرمطلق برای مقادیر مثبت k داریم:

$$\begin{aligned}
 P(|X - \mu| \geq k\sigma) &= P(|X - \mu|^r \geq k^r \sigma^r) \\
 &= P((X - \mu)^r \geq k^r \sigma^r) \\
 &\leq \frac{E(X - \mu)^r}{k^r \sigma^r} \\
 &= \frac{\text{var}(X)}{k^r \sigma^r} = \frac{1}{k^r},
 \end{aligned}$$

يعنى

$$P(|X - \mu| \geq k\sigma) \leq \frac{1}{k^r},$$

يا

$$P(|X - \mu| \leq k\sigma) \geq 1 - \frac{1}{k^r}.$$

$$1 - \frac{1}{k^r} = 0.95 \rightarrow k^r = \frac{1}{0.05} = 20 \rightarrow k = \sqrt{20} \quad (25-5 \text{ الف})$$

$$1 - \frac{1}{k^r} = 0.99 \rightarrow k^r = \frac{1}{0.01} = 100 \rightarrow k = 10 \quad (ب)$$

$$P(|X - \mu| \geq c) \leq \frac{1}{k^r} = \frac{\sigma^r}{c^r} \quad (26-5)$$

$$27-5. \text{ خير. زیرا } M_X(\circ) = \frac{\circ}{1-\circ} = 0 \neq 1 \text{ است.}$$

28-5

$$\begin{aligned}
 M_X(t) &= E(e^{tX}) = \int_{-\infty}^{\infty} \frac{1}{\pi} e^{tx} e^{-|x|} dx \\
 &= \frac{1}{\pi} \int_{-\infty}^0 e^{tx} e^{+x} dx + \frac{1}{\pi} \int_0^{+\infty} e^{tx} e^{-x} dx \\
 &= \frac{1}{\pi} \int_{-\infty}^0 e^{x(1+t)} dx + \frac{1}{\pi} \int_0^{+\infty} e^{-x(1-t)} dx \\
 &= \left(\frac{e^{x(1+t)}}{1+t} \right) \Big|_{-\infty}^0 - \left(\frac{e^{-x(1-t)}}{1-t} \right) \Big|_0^{+\infty} = \frac{1}{1-t}.
 \end{aligned}$$

.٢٩-٥

$$M_Z(t) = M_{\frac{1}{r}X - \frac{\mu}{r}}(t) = e^{-\frac{\mu}{r}t} M_X\left(\frac{t}{r}\right) = e^{-\frac{\mu}{r}t} e^{\frac{\mu}{r}t + \frac{\lambda t^r}{r}} = e^{\frac{t^r}{r}}.$$

$$E(Z) = \frac{\partial M_Z(t)}{\partial t} \Big|_{t=0} = t e^{\frac{t^r}{r}} \Big|_{t=0} = 0,$$

$$E(Z^r) = \text{var}(Z) = \frac{\partial^r M_Z(t)}{\partial t^r} \Big|_{t=0} = e^{\frac{t^r}{r}} + t^r e^{\frac{t^r}{r}} \Big|_{t=0} = 1.$$

(الف) ٣٠-٥

$$1 - \frac{1}{k^r} = \frac{35}{36} = 0/99,$$

$$k^r = \frac{1}{0/03} = 33/3,$$

$$k = \sqrt{33/3} = 5/3,$$

$$(0/25 - 5/77(0/005)) , 0/25 + 5/77(0/005)).$$

(ب)

$$1 - \frac{1}{k^r} = \frac{143}{144} = 0/99,$$

$$k^r = \frac{1}{0/01} = 100,$$

$$k = 10,$$

$$(0/25 - 10(0/005)) , 0/25 + 10(0/005)).$$

.٣١-٥

| | | | | |
|---------|----------------|----------------|---------------|---------------|
| x | -3 | -1 | 1 | 3 |
| $f(x)$ | $\frac{1}{r}$ | $\frac{1}{r}$ | $\frac{1}{r}$ | $\frac{1}{r}$ |
| $xf(x)$ | $\frac{-3}{r}$ | $\frac{-1}{r}$ | $\frac{1}{r}$ | $\frac{3}{r}$ |
| y | -5 | -1 | 1 | 5 |
| $f(y)$ | $\frac{1}{r}$ | $\frac{1}{r}$ | $\frac{1}{r}$ | $\frac{1}{r}$ |
| $yf(y)$ | $\frac{-5}{r}$ | $\frac{-1}{r}$ | $\frac{1}{r}$ | $\frac{5}{r}$ |

$$E(X) = \sum_x xf(x) = 0$$

$$E(Y) = \sum_y yf(y) = 0$$

$$\begin{aligned}
 E(XY) &= \sum_x \sum_y xyf(x,y) \\
 &= \frac{1}{4}(-\omega)(-\Delta) + \frac{1}{4}(-1)(-1) + \frac{1}{4}(1)(1) + \frac{1}{4}(\omega)(\Delta) = \frac{\omega\Delta}{4}, \\
 \text{cov}(X,Y) &= E(XY) - E(X)E(Y) = \frac{\omega\Delta}{4}.
 \end{aligned}$$

٣٢-٥. (الف)

$$\begin{array}{c|ccc}
 x & -1 & 0 & 1 \\
 \hline
 f(x) & \frac{\omega}{12} & \frac{\Delta}{12} & \frac{\gamma}{12} \\
 xf(x) & \frac{-\omega}{12} & 0 & \frac{\gamma}{12} \\
 \hline
 y & 0 & 1 \\
 \hline
 f(y) & \frac{\omega}{12} & \frac{\Delta}{12} \\
 yf(y) & 0 & \frac{\Delta}{12} \\
 \hline
 \end{array}$$

$$\begin{aligned}
 E(X) &= \sum_x xf(x) = \frac{\Delta}{12} \\
 E(Y) &= \sum_y yf(y) = \frac{\Delta}{12} \\
 E(XY) &= \sum_x \sum_y xyf(x,y) \\
 &= 0(-1)(0) + \frac{1}{4}(-1)(1) + \frac{1}{4}(0)(0) + 0(0)(1) + \frac{1}{4}(1)(0) + \frac{1}{4}(1)(1) = \frac{1}{4}, \\
 \text{cov}(X,Y) &= E(XY) - E(X)E(Y) = \frac{1}{4} - \frac{\Delta}{12} \times \frac{\Delta}{12} = 0.
 \end{aligned}$$

ب) مثال نقض:

$$f(0,1) = 0 \neq f_X(0)f_Y(1) = \frac{\omega}{12} \times \frac{\Delta}{12}$$

٣٣-٦

$$\begin{aligned}
 E(X_1) &= \omega, & E(X_\rho) &= \Delta, & E(X_\mu) &= \gamma, \\
 \text{var}(X_1) &= \omega, & \text{var}(X_\rho) &= \Delta, & \text{var}(X_\mu) &= \gamma.
 \end{aligned}$$

(الف)

$$\begin{aligned}
 E(Y) &= \omega E(X_1) - \Delta E(X_\rho) + \gamma E(X_\mu) = \omega(\omega) - \Delta(\Delta) + \gamma(\gamma) = -\Delta, \\
 \text{var}(Y) &= \omega \text{var}(X_1) + \Delta \text{var}(X_\rho) + \gamma \text{var}(X_\mu) = \omega(\omega) + \Delta(\Delta) + \gamma(\gamma) = 12\omega.
 \end{aligned}$$

ب)

$$\begin{aligned}
 E(Z) &= E(X_1) + \omega E(X_\rho) - E(X_\mu) = \omega + \omega(\Delta) - \omega = 1\omega, \\
 \text{var}(Z) &= \text{var}(X_1) + \omega \text{var}(X_\rho) + \text{var}(X_\mu) = \omega + \omega\Delta + \omega = 3\omega.
 \end{aligned}$$

٣٤-٦

٣١١ حل تمرین‌ها

$$\begin{aligned}\text{var}(X+Y) &= \text{var}(X) + \text{var}(Y) + 2\text{cov}(X,Y), \\ \text{var}(X-Y) &= \text{var}(X) + \text{var}(Y) - 2\text{cov}(X,Y), \\ \text{cov}(X+Y, X-Y) &= \text{cov}(X,X) + \text{cov}(X,-Y) + \text{cov}(X,Y) + \text{cov}(Y,-Y) \\ &= \text{var}(X) - \text{cov}(X,Y) + \text{cov}(X,Y) - \text{var}(Y) \\ &= \text{var}(X) - \text{var}(Y).\end{aligned}$$

.٣٥-٥

$$\begin{aligned}\text{cov}(Y_1, Y_2) &= -2\text{var}(X_1) + 3\text{cov}(X_1, X_2) + 2\text{cov}(X_1, X_3) \\ &\quad - 2\text{cov}(X_1, X_4) + 2\text{var}(X_2) + 3\text{cov}(X_2, X_3) \\ &\quad - 3\text{cov}(X_2, X_4) + 3\text{cov}(X_3, X_4) + 2\text{var}(X_4) \\ &= -2(5) + 3(3) + 2(-2) - 2(3) + 2(2) + 3(0) - 3(-2) + 3(0) + 2(7) \\ &= 9.\end{aligned}$$

٣٦-٥. الف) X تعداد پیروزی‌ها

| x | ٠ | ١ | ٢ | ٣ |
|------------|-------------------|-------------------|-------------------|------------------|
| $f(x)$ | $\frac{240}{624}$ | $\frac{308}{624}$ | $\frac{72}{624}$ | $\frac{4}{624}$ |
| $xf(x)$ | ٠ | $\frac{308}{624}$ | $\frac{144}{624}$ | $\frac{12}{624}$ |
| $x^2 f(x)$ | ٠ | $\frac{308}{624}$ | $\frac{288}{624}$ | $\frac{36}{624}$ |

$$E(X) = \sum_x xf(x) = 0/624, \quad E(X^2) = \frac{624}{624},$$

$$\text{var}(X) = \frac{624}{624} - (0/624)^2 = 0/46, \quad \sigma = 0/68$$

$$P(X=0) = \frac{1}{2} \times \frac{5}{6} \times \frac{48}{52} = \frac{240}{624},$$

$$P(X=1) = \frac{1}{2} \times \frac{5}{6} \times \frac{48}{52} + \frac{1}{2} \times \frac{1}{6} \times \frac{48}{52} + \frac{1}{2} \times \frac{5}{6} \times \frac{4}{52} = \frac{308}{624},$$

$$P(X=2) = \frac{1}{2} \times \frac{1}{6} \times \frac{48}{52} + \frac{1}{2} \times \frac{5}{6} \times \frac{4}{52} + \frac{1}{2} \times \frac{1}{6} \times \frac{4}{52} = \frac{72}{624},$$

$$P(X=3) = \frac{1}{2} \times \frac{1}{6} \times \frac{4}{52} = \frac{4}{624}.$$

ب) به روش مشابه قسمت الف، قابل محاسبه است.

$$E(X) = \mu = 1/91, \quad \sigma = 1/05.$$

حل تمرین‌های فصل ششم

۱-۶. با استفاده از برابری‌های $\sum_{i=1}^k i^r = \frac{k(k+1)(2k+1)}{6}$ و $\sum_{i=1}^k i = \frac{k(k+1)}{2}$ داریم

(الف)

$$\begin{aligned}\mu &= E(X) = \sum_{x=1}^k xf(x) = \frac{1}{k} \sum_{x=1}^k x = \frac{k+1}{2}, \\ \mu'_r &= E(X^r) = \sum_{x=1}^k x^r f(x) = \frac{1}{k} \sum_{x=1}^k x^r = \frac{(k+1)(2k+1)}{6}, \\ \sigma^r &= \mu'_r - \mu^r = \frac{(k+1)(2k+1)}{6} - \left(\frac{k+1}{2}\right)^r = \frac{k^r - 1}{12}.\end{aligned}$$

.۲-۷

$$\begin{aligned}M_X(t) &= E(e^{tX}) = \sum_{x=1}^k \frac{e^{tx}}{k} = \frac{1}{k} \sum_{x=1}^k (e^t)^x \\ &= \frac{e^t}{k} \sum_{x=1}^k (e^t)^{x-1} = \frac{e^t}{k} \times \frac{1-e^{kt}}{1-e^t} = \frac{e^t - e^{kt+1}}{k(1-e^t)}, \\ M'_X(t) &= \left(\frac{1}{k} \sum_{x=1}^k e^{tx} \right)' = \frac{1}{k} \sum_{x=1}^k x e^{tx}, \\ M'_X(0) &= \frac{1}{k} \sum_{x=1}^k x e^0 = \frac{1}{k} \sum_{x=1}^k x = \frac{1}{k} \times \frac{k(k+1)}{2} = \frac{k+1}{2}.\end{aligned}$$

۳-۷. نمادگذاری $b(x; n, \theta)$ را در نظر بگیرید، و توجه کنید که

$$\binom{n}{x}, \text{آنگاه } \binom{n}{x} = \binom{n}{n-x}$$

$$\begin{aligned}b(n-x; n, 1-\theta) &= \binom{n}{n-x} (1-\theta)^{n-x} (1-(1-\theta))^{n-(n-x)} \\ &= \binom{n}{n-x} (1-\theta)^{n-x} \theta^x = \binom{n}{x} \theta^x (1-\theta)^{n-x} \\ &= b(x; n, \theta).\end{aligned}$$

(ب)

$$\begin{aligned}
 B(x; n, \theta) - B(x-1; n, \theta) &= \sum_{k=0}^x b(k; n, \theta) - \sum_{k=0}^{x-1} b(k; n, \theta) \\
 &= b(x; n, \theta) + \sum_{k=0}^{x-1} b(k; n, \theta) - \sum_{k=0}^{x-1} b(k; n, \theta) \\
 &= b(x; n, \theta).
 \end{aligned}$$

(ج)

$$\begin{aligned}
 B(n-x; n, 1-\theta) - B(n-x-1; n, 1-\theta) &= \sum_{k=0}^{n-x} b(k; n, 1-\theta) - \sum_{k=0}^{n-x-1} b(k; n, 1-\theta) \\
 &= b(n-x; n, 1-\theta) + \sum_{k=0}^{n-x-1} (b(k; n, 1-\theta) - b(k; n, 1-\theta)) \\
 &= b(n-x; n, 1-\theta) \\
 &= b(x; n, \theta).
 \end{aligned}$$

(د)

$$\begin{aligned}
 B(x; n, \theta) + B(n-x-1; n, 1-\theta) &= \sum_{k=0}^x \binom{n}{k} \theta^k (1-\theta)^{n-k} + \sum_{k=x+1}^n \binom{n}{k} \theta^k (1-\theta)^{n-k} \\
 &= \sum_{k=0}^n \binom{n}{k} (1-\theta)^{n-k} = (\theta + (1-\theta))^n = 1.
 \end{aligned}$$

توجه داریم که

$$B(n-x-1; n, 1-\theta) = \sum_{k=0}^{n-x-1} \binom{n}{k} (1-\theta)^k \theta^{n-k}, \quad k = n-i$$

با تغییر متغیر

$$\begin{aligned}
 B(n-x-1; n, 1-\theta) &= \sum_{i=x+1}^n \binom{n}{n-i} (1-\theta)^{n-i} \theta^i \\
 &= \sum_{i=x+1}^n \binom{n}{i} \theta^i (1-\theta)^{n-i}.
 \end{aligned}$$

۶-۴. X_i ها مستقل و هم توزیع‌اند بنابراین واریانس‌ها و میانگین‌های آنها برابرند:

$$E(Y) = E(X_1) + \cdots + E(X_n) = nE(X) = n\theta,$$

$$\text{var}(Y) = \text{var}(X_1) + \cdots + \text{var}(X_n) = n \text{var}(X) = n\theta(1-\theta).$$

۵-۶. از برابری $\binom{n}{x+1} = \frac{n-x}{x+1} \binom{n}{x}$ استفاده می‌کنیم و بنابراین

$$\begin{aligned} b(x+1; n, \theta) &= \binom{n}{x+1} \theta^{x+1} (1-\theta)^{n-x-1} \\ &= \frac{n-x}{x+1} \times \frac{\theta}{1-\theta} \binom{n}{x} \theta^x (1-\theta)^{n-x} \\ &= \frac{\theta(n-x)}{(x+1)(1-\theta)} b(x; n, \theta). \\ b(o; v, o/v) &= \binom{v}{o} (o/v)^o (v/o)^v = o/v \cdot v, \\ b(1; v, o/v) &= \frac{o/v}{v} b(o; v, o/v) = o/v \cdot v, \\ &\vdots \\ b(v; v, o/v) &= o/v. \end{aligned}$$

۶-۶. در تمرین قبل فرض کنیم $\theta = \frac{1}{2}$ بنابراین

$$b(x+1; n, \frac{1}{2}) = \frac{n-x}{x+1} b(x; n, \frac{1}{2})$$

$$n-x > x+1 \rightarrow x < \frac{n}{2} - \frac{1}{2}: b(x+1; n, \frac{1}{2}) > b(x; n, \frac{1}{2}),$$

$$n-x < x+1 \rightarrow x > \frac{n}{2} - \frac{1}{2}: b(x+1; n, \frac{1}{2}) < b(x; n, \frac{1}{2}),$$

برای n زوج، اگر $x < \frac{n}{2}$ ، $b(x; n, \frac{1}{2})$ اکیداً صعودی و اکیداً نزولی است اگر و تنها

اگر $x > \frac{n}{2}$. بنابراین $x = \frac{n}{2}$ نقطه‌ی ماقسیمم $b(x; n, \frac{1}{2})$ است.

ب) همانند حالت قبل وقتی $x < \frac{n}{2} - \frac{1}{2}$ تابع اکیداً نزولی و $x > \frac{n}{2} - \frac{1}{2}$ فرد باشد

تابع اکیدا صعودی است، $b(x+1; n, \frac{1}{2}) = b(x; n, \frac{1}{2})$ چون پس ماقسیمم در

$x = \frac{n+1}{2}$ ، $x = \frac{n-1}{2}$ اتفاق می‌افتد.

۷-۶. چون $1 \leq \theta \leq n$ پس تابع $f(\theta) = \binom{n}{x} \theta^x (1-\theta)^{n-x}$ از θ است.

$$\begin{aligned} f'(\theta) &= \binom{n}{x} \left[x\theta^{x-1}(1-\theta)^{n-x} + (n-x)(1-\theta)^{n-x-1}\theta^x \right] \\ &= \binom{n}{x} \theta^{x-1}(1-\theta)^{n-x-1} [x(1-\theta) + (x-n)\theta] \\ &= 0, \end{aligned}$$

جواب معادلهٔ بالا برابر با $\theta = \frac{x}{n}$ ، خواهد شد از طرفی $f''\left(\frac{x}{n}\right) < 0$ ، پس جواب مشتق اول، تابع را مаксیمم می‌کند.

۸-۶

$$\mu'(r) = E(X(X-1)) = E(X^r) - E(X) = \mu'_r - \mu'_1,$$

$$\mu'(m) = E(X(X-1)(X-r)) = E(X^m) - mE(X^r) + rE(X) = \mu'_m - m\mu'_r + r\mu'_1,$$

$$\begin{aligned} \mu'(k) &= E(X(X-1)(X-r)(X-m)) = E(X^k) - kE(X^r) + mrE(X^m) - kmE(X^r) \\ &= \mu'_k - k\mu'_r + mr\mu'_m - km\mu'_r. \end{aligned}$$

۹-۶

$$\Psi'_X(t) = E'(X^t) = E(X t^{X-1}),$$

$$\Psi''_X(t) = E(X(X-1)t^{X-2}).$$

در حالت کلی

$$\Psi_X^{(r)}(t) = E(X(X-1)(X-r)\cdots(X-r+1)t^{X-r}),$$

$$\Psi_X^{(r)}(0) = E(X(X-1)\cdots(X-r+1)).$$

۱۰-۶. شانس پاسخ صحیح برای هر سؤال برابر با $\frac{1}{m}$ است، و تعداد سؤال‌ها $n = 8$

تعداد موفقیت‌ها $x = 4$ است. بنابر توزیع دوجمله‌ای

$$P(X = 4) = \binom{8}{4} \left(\frac{1}{m}\right)^4 \left(\frac{m-1}{m}\right)^4 = \frac{1}{1024}.$$

۱۱-۶. بنابر توزیع دوجمله‌ای $n = 5$ ، $\theta = \frac{1}{10}$

$$P(X \geq n) = 1 - P(X \leq n) = 0.0085.$$

$$\theta = 0/4, \quad n = 15, \quad x = 6. \quad .12-7$$

$$P(X = 6) = \binom{15}{6} (0/4)^6 (0/4)^9 = 0.2022 \quad (\text{الف})$$

$$P(X = 6) = P(X \leq 6) - P(X \leq 5) = 0.2022 \quad (\text{ب})$$

$$\theta = 0/4, \quad n = 6, \quad x = 5. \quad .13-7$$

$$P(X = 5) = \binom{6}{5} (0/4)^5 (0/4)^1 = 0.3025 \quad (\text{الف})$$

$$P(X = 5) = P(X \leq 5) - P(X \leq 4) = 0.3025 \quad (\text{ب})$$

$$.14-7$$

$$P(X = 11) = \binom{14}{11} (0/4)^{11} (0/4)^3 = 0.0923$$

$$P(X = 11) = \binom{14}{11} (0/4)^{11} (0/4)^3 = 0.0923 \quad (\text{الف})$$

$$P(\theta - c < X < \theta + c) \geq 1 - \frac{\theta(1-\theta)}{nc^2}. \quad .15-7$$

$$\begin{cases} \theta - c = 0/4 \\ \theta + c = 0/4 \end{cases} \Rightarrow \left(\theta = \frac{1}{2}, \quad c = 0/1 \right) \xrightarrow{n=900} 1 - \frac{\theta(1-\theta)}{nc^2} = \frac{35}{36} \quad (\text{الف})$$

$$\begin{cases} \theta - c = 0/4 \\ \theta + c = 0/5 \end{cases} \Rightarrow \left(\theta = \frac{1}{2}, \quad c = 0/0.5 \right) \xrightarrow{n=10000} 1 - \frac{\theta(1-\theta)}{nc^2} = \frac{35}{36} \quad (\text{ب})$$

$$\begin{cases} \theta - c = 0/4 \\ \theta + c = 0/50 \end{cases} \Rightarrow \left(\theta = \frac{1}{2}, \quad c = 0/0.05 \right) \xrightarrow{n=100000} 1 - \frac{\theta(1-\theta)}{nc^2} = \frac{35}{36} \quad (\text{ج})$$

هرچه n بزرگتر شود احتمال موردنظر به یک نزدیکتر می‌گردد و c نسبت موقیت‌ها

به $\frac{1}{2}$ میل می‌کند.

$$.16-7. \text{ با استفاده از تعریف فوق هندسی الف) } h(0; 2, 18, 4) = \frac{\binom{4}{0} \binom{14}{2}}{\binom{18}{2}} = 0.595$$

$$h(o; 2, 1\lambda, 1\alpha) = \frac{\binom{1\alpha}{o} \binom{6}{2}}{\binom{1\alpha}{2}} = o/504 \quad (ج)$$

$$h(o; 2, 1\lambda, \lambda) = \frac{\binom{\lambda}{o} \binom{1\alpha}{2}}{\binom{1\alpha}{2}} = o/294 \quad (ب)$$

١٧-٦. با استفاده از تعریف فوق هندسی

$$h(1; 3, 1\alpha, 1\alpha) = \frac{\binom{1\alpha}{1} \binom{6}{2}}{\binom{1\alpha}{3}} = \frac{1\alpha}{56} \quad (ب)$$

$$h(o; 3, 1\alpha, 1\alpha) = \frac{\binom{1\alpha}{o} \binom{6}{3}}{\binom{1\alpha}{3}} = \frac{1}{28} \quad (الف)$$

$$h(3; 3, 1\alpha, 1\alpha) = \frac{\binom{1\alpha}{3} \binom{6}{o}}{\binom{1\alpha}{3}} = \frac{3}{14} \quad (د)$$

$$h(2; 3, 1\alpha, 1\alpha) = \frac{\binom{1\alpha}{2} \binom{6}{1}}{\binom{1\alpha}{3}} = \frac{27}{56} \quad (ج)$$

١٨-٦. (الف)

| x | ۰ | ۱ | ۲ | ۳ |
|--------|----------------|----------------------|-----------------|----------------|
| $f(x)$ | $\frac{1}{28}$ | $\frac{1\alpha}{56}$ | $\frac{27}{56}$ | $\frac{6}{28}$ |

$$\mu = E(X) = \sum_{x=0}^3 xf(x) = \frac{1\alpha}{8},$$

$$\mu' = E(X') = \sum_{x=0}^3 xf(x) = \frac{33}{8},$$

$$\sigma' = \mu' - \mu = \frac{39}{144}.$$

ب) از فرمول‌ها

$$\mu = \frac{nk}{N} = \frac{3 \times 1\alpha}{1\alpha} = \frac{1\alpha}{8},$$

$$\sigma' = \frac{nk(N-k)(N-n)}{N'(N-1)} = \frac{39}{144}.$$

$$h(2; 5, 1\alpha, 9) = \binom{9}{2} \binom{6}{3} / \binom{1\alpha}{5}$$

١٩-٦. توزیع فوق هندسی

۲۰-۶. الف) مستقیماً از توزیع استفاده می‌کنیم

$$h(1; 3, 8, 0, 1) = \frac{\binom{1}{1} \binom{7}{2}}{\binom{8}{3}} = 0/1 = 0.$$

ب) بنابر توزیع دو جمله‌ای $\theta = \frac{k}{N} = \frac{1}{8} = 0.125$

$$h(1; 3, 0.125, 0, 1) = \binom{3}{1} (0.125)^1 (0.875)^2 = 0.135.$$

۲۱-۷

$$p(x+1; \lambda) = \frac{e^{-\lambda} \lambda^{x+1}}{(x+1)!} = \frac{\lambda}{x+1} \times \frac{e^{-\lambda} \lambda^x}{x!} = \frac{\lambda}{x+1} p(x; \lambda).$$

$$p(0; 2) = \frac{e^{-2} 2^0}{0!} = 0.135,$$

$$p(1; 2) = \frac{2}{0+1} = p(0; 2) = 0.135.$$

۲۲-۷

$$\mu_r = E(X^r) = \sum_{x=0}^{+\infty} (x-\lambda)^r \frac{\lambda^x e^{-\lambda}}{x!}.$$

$$\begin{aligned} \frac{\partial \mu_r}{\partial \lambda} &= \sum_{x=0}^{+\infty} \left[r(x-\lambda)^{r-1} \frac{\lambda^x e^{-\lambda}}{x!} + (x-\lambda)^r \left\{ \frac{x \lambda^{x-1} e^{-\lambda}}{x!} - \frac{\lambda^x e^{-\lambda}}{x!} \right\} \right] \\ &= r \sum_{x=0}^{+\infty} (x-\lambda)^{r-1} \frac{\lambda^x e^{-\lambda}}{x!} + \sum_{x=1}^{+\infty} \frac{\lambda^x e^{-\lambda}}{x!} \left\{ (x-\lambda)^r \left(\frac{x}{\lambda} - 1 \right) \right\} \\ &= r \mu_{r-1} + \frac{1}{\lambda} \sum_{x=0}^{+\infty} (x-\lambda)^{r+1} \frac{\lambda^x e^{-\lambda}}{x!} = r \mu_{r-1} + \frac{1}{\lambda} \mu_{r+1}, \end{aligned}$$

در نتیجه

$$\mu_{r+1} = \lambda [r \mu_{r-1} + \frac{\partial \mu_r}{\partial \lambda}].$$

۲۳-۷

$$M_Y(t) = M_{X-\lambda}(t) = e^{-\lambda t} M_X(t) = e^{-\lambda t} e^{\lambda(e^t - 1)} = e^{\lambda(e^t - 1 - t)}.$$

$$\begin{aligned}\frac{\partial M_Y(t)}{\partial t} &= \lambda(e^t - 1)e^{\lambda(e^t - 1-t)}, \\ \left. \frac{\partial^r M_Y(t)}{\partial t^r} \right|_{t=0} &= \left[\lambda^r (e^t - 1)^{r-1} e^{\lambda(e^t - 1-t)} + \lambda e^t e^{\lambda(e^t - 1-t)} \right]_{t=0} \\ &= E(Y^r) = \text{var}(X) = \lambda.\end{aligned}$$

- ۲۴-۶. الف) در حالت کلی تقریبی که توصیه می‌شود برای حالت $n \geq ۲۰$ ، و $0.05 \leq \theta \leq 0.10$ است اما اگر $n \geq ۱۰۰$ ، و $n\theta < ۱۰$ برقار باشد تقریب دقیق‌تر خواهد شد.
- الف) برای $n\theta = ۱۲/۵ > ۲۰$ و $n = ۲۵$ ، تقریب خوب است.
- ب) در شرایط $n > ۱۰۰$ ، و $\theta = ۰.۰۵ < ۰.۱$ ، $n\theta < ۱۰$ تقریب خوب است.
- د) چون $0.05 > \theta$ ، تقریب مناسب نیست.

۲۵-۶. طبق صورت تمرین داریم: $\lambda = n\theta = ۲/۱ < ۱۰$ ، $n = ۱۵۰ > ۱۰۰$

$$P(r; \mu) = \frac{e^{-\mu} (\mu/r)^r}{r!} = ۰.۲۷$$

- ۲۶-۶. میانگین مسمومیت غذایی بازدیدکنندگان از یک نمایشگاه استانی، برابر است با: $\lambda = n\theta = ۱/۲$.

$$\begin{aligned}P(X \geq r) &= ۱ - P(X \leq r) = ۱ - P(X = ۰) - P(X = ۱) \\ &= ۱ - \frac{e^{-1/2} (1/2)^0}{0!} - \frac{e^{-1/2} (1/2)^1}{1!} = ۰.۹۰۲.\end{aligned}$$

۲۷-۶. میانگین تصادف‌های رانندگان مجاز، $\lambda = n\theta = ۶$

$$P(X = \mu) = \frac{e^{-\mu} \times \mu^\mu}{\mu!} = ۰.۱۶۰۶.$$

$$P(X = \mu) = \sum_{x=0}^{\mu} p(x; \mu) = \sum_{x=0}^{\mu} \frac{e^{-\mu} \mu^x}{x!} = P(X \leq \mu) = ۰.۱۵۱۲.$$

۲۸-۶

$$P(X = \mu) = \frac{e^{-\mu/\mu} (\mu/\mu)^\mu}{\mu!} = ۰.۲۰۱۵.$$

۲۹-۶

$$P(X=0) = \frac{e^{1/\lambda} (1/\lambda)^0}{0!} = 0.1653.$$

$$P(X=1) = \frac{e^{-1/\lambda} (1/\lambda)^1}{1!} = 0.2975.$$

۳۰-۶

$$P(X=2) = p(2; \alpha, \beta) = 0.1293.$$

$$P(X \geq 1) = 1 - \sum_{x=0}^1 p(x; \alpha, \beta) = 1 - P(X \leq 1) = 0.0396.$$

$$P(\epsilon \leq X \leq \delta) = P(X=\epsilon) + P(X=\alpha) + P(X=\delta) = 0.14944.$$

۳۱-۶. برای توزیع یکنواخت $f(x) = \frac{1}{\beta-\alpha}$ ، $\alpha < x < \beta$ احتمال موردنظر عبارت

است از:

$$\begin{aligned} P(X < \alpha + p(\beta - \alpha)) &= \int_{\alpha}^{\alpha + p(\beta - \alpha)} \frac{1}{\beta - \alpha} dx \\ &= \frac{1}{\beta - \alpha} [\alpha + p(\beta - \alpha)] - \frac{1}{\beta - \alpha} [\alpha] = p. \end{aligned}$$

۳۲-۷. امین گشتاور حول میانگین، $\mu = \frac{\alpha + \beta}{2}$ برابر است با:

$$\begin{aligned} \mu_r &= E(X - \mu)^r \\ &= \int_{\alpha}^{\beta} (x - \mu)^r \frac{1}{\beta - \alpha} dx = \frac{1}{\beta - \alpha} \left[\frac{(x - \mu)^{r+1}}{r+1} \right]_{\alpha}^{\beta} \\ &= \frac{(\beta - \mu)^{r+1} - (\alpha - \mu)^{r+1}}{(\beta - \alpha)(r+1)} = \frac{(\beta - \frac{\alpha + \beta}{2})^{r+1} - (\alpha - \frac{\alpha + \beta}{2})^{r+1}}{(\beta - \alpha)(r+1)} \\ &= \frac{(\frac{\beta - \alpha}{2})^{r+1} - (\frac{\alpha - \beta}{2})^{r+1}}{(\beta - \alpha)(r+1)} = \frac{(\beta - \alpha)^{r+1} - (\alpha - \beta)^{r+1}}{2^{r+1}(\beta - \alpha)(r+1)}, \end{aligned}$$

به ازای $r+1$ های زوج یا r های فرد مقدار کمیت فوق برابر با ۰ است و برای r های زوج گشتاور مرتبه r حول میانگین برابر با

۳۲۱ حل تمرین‌ها

$$\mu_r = \frac{(\beta - \alpha)^{r+1} + (\beta - \alpha)^{r+1}}{r^{r+1}(\beta - \alpha)(r+1)} = \frac{r(\beta - \alpha)^{r+1}}{r^{r+1}(\beta - \alpha)(r+1)} = \frac{1}{r+1} \left(\frac{\beta - \alpha}{r} \right)^r.$$

۳۳-۶. خاصیتی مشهور که در هر مثلث حاکم است (مجموع هر دو ضلع مثلث از طول ضلع دیگر بیشتر است، مشهور به نابرابری مثلث) را به کار می‌بریم:

$$|AD| < |BD| + |AD| \\ |BD| < |AD| + |AC| \\ |AC| < |AD| + |BC| \rightarrow \begin{cases} x < a - x + \frac{a}{2} \\ a - x < x + \frac{a}{2} \\ \frac{a}{2} < x - a - x \end{cases}$$

غایق

با توجه به توزیع متغیر تصادفی X

$$f(x) = \begin{cases} \frac{1}{a} & 0 \leq x \leq a \\ 0 & \text{سایر جاهات} \end{cases}$$

داریم

$$P\left(\frac{a}{4} < X < \frac{3a}{4}\right) = \int_{\frac{a}{4}}^{\frac{3a}{4}} \frac{1}{a} dx = \frac{1}{a} \left(\frac{3a}{4} - \frac{a}{4}\right) = \frac{1}{2}.$$

۳۴-۷. مقدار ارتکاب خطأ در تعیین وزن مخصوص یک ماده است.

$$f(x) = \begin{cases} \frac{1}{0.015 - (-0.015)} = \frac{1}{0.03} & -0.015 < x < 0.015 \\ 0 & \text{سایر جاهات} \end{cases}$$

(الف)

$$P(-0.002 < X < 0.003) = \int_{-0.002}^{0.003} \frac{1}{0.03} dx = \frac{0.005}{0.03} = \frac{1}{6}.$$

(ب)

$$P(X > 0.005) = \int_{0.005}^{0.015} \frac{1}{0.03} dx = \frac{1}{3}.$$

.۳۵-۷

$$n(x; \mu, \sigma) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}(\frac{x-\mu}{\sigma})^2}, \quad -\infty < x < \infty$$

$$n'(x; \mu, \sigma) = -\frac{1}{\sigma\sqrt{2\pi}} \left(\frac{x-\mu}{\sigma} \right) \left(\frac{1}{\sigma} \right) e^{-\frac{1}{2}(\frac{x-\mu}{\sigma})^2} = 0$$

$$\frac{x-\mu}{\sigma} = 0$$

$$x = \mu$$

$$n''(x; \mu, \sigma) = -\frac{1}{\sigma\sqrt{2\pi}} \left[\frac{1}{\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}} - \frac{1}{\sigma} \left(\frac{(x-\mu)^2}{2\sigma^2} \right) e^{-\frac{(x-\mu)^2}{2\sigma^2}} \right]$$

$$= -\frac{1}{\sigma^3\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \left[1 - \left(\frac{x-\mu}{\sigma} \right)^2 \right] = 0,$$

$$x = \mu \pm \sigma.$$

چون $\circ < n''(x; \mu, \sigma)$, پس نقطه‌ی ماقسیمم برابر است با $x = \mu$. $n''(x; \mu, \sigma)$ به ازای $x = \mu \pm \sigma$ صفر می‌شود.

.۳۶-۶

$$M_Z(t) = E(e^{tZ}) = E\left(e^{\frac{t(X-\mu)}{\sigma}}\right)$$

$$= e^{-\frac{\mu t}{\sigma}} M_X\left(\frac{t}{\sigma}\right) = e^{-\frac{\mu t}{\sigma}} e^{\mu\left(\frac{t}{\sigma}\right) + \frac{1}{2}\sigma^2\left(\frac{t}{\sigma}\right)^2} = e^{\frac{1}{2}t^2}.$$

.۳۷-۷

$$P(Z \geq 1/14) = 1 - P(Z < 1/14) = 0.1271 \quad (\text{الف})$$

$$P(Z > -0/35) = 1 - P(Z < 0/35) = 0.6406 \quad (\text{ب})$$

$$P(-0/45 < Z < -0/09) = P(Z < -0/09) - P(Z < -0/45) = 0.1413 \quad (\text{ج})$$

$$P(-0/58 < Z < 1/12) = P(Z < 1/12) - P(Z < 0/58) = 0.5876 \quad (\text{د})$$

.۳۸-۱

$$P(Z < 1/33) = 0.9082 \quad \text{(الف)}$$

$$P(Z \leq -0.79) = 0.2148 \quad \text{(ب)}$$

$$P(0.55 < Z < 1.22) = P(Z < 1.22) - P(Z < 0.55) = 0.1 \quad \text{(ج)}$$

$$P(-1.9 \leq Z \leq 0.44) = P(Z < -1.9) - P(Z < 0.44) = 0.6413 \quad \text{(د)}$$

.۳۹-۱

$$P(0 < Z < z) = 0.4726 \rightarrow z = 1/92 \quad \text{(الف)}$$

$$P(Z < z) = 0.9868 \quad \text{جواب از روی جدول نرمال} \quad z = 2/22 \quad \text{(ب)}$$

$$P(Z < z) = 0.3686 \quad \text{برابر} \quad P(Z > z) = 0.1314 \quad \text{است، پس جواب از روی جدول} \quad z = 1/12 \quad \text{(ج)}$$

$$P(-z < Z < z) = 0.4251 \quad \text{برابر} \quad P(Z < -z) = 0.8502 \quad \text{در نتیجه} \quad z = 1/44 \quad \text{(د)}$$

.۴۰-۱

$$z_1 = 1/48 \quad \text{(الف)}$$

$$z_2 = -0.74 \quad \text{(ب)}$$

$$z_3 = 0.55 \quad \text{(ج)}$$

$$z_4 = 2/17 \quad \text{(د)}$$

$$P(\mu - k\sigma < X < \mu + k\sigma) = P(-k < Z < k) = P(Z < k) - P(Z < -k) \quad .41-۱$$

$$P(\mu - \sigma < X < \mu + \sigma) = 0.6826 \quad \text{(الف)}$$

$$P(\mu - 2\sigma < X < \mu + 2\sigma) = 0.9544 \quad \text{(ب)}$$

$$P(\mu - 3\sigma < X < \mu + 3\sigma) = 0.9974 \quad \text{(ج)}$$

$$P(\mu - 4\sigma < X < \mu + 4\sigma) = 0.99994 \quad \text{(د)}$$

.۴۲-۱

$$\alpha = P(Z > Z_\alpha) = 1 - P(Z < Z_\alpha)$$

$$P(Z < Z_\alpha) = 1 - \alpha$$

$$Z_{0.05} = 1.645 \quad \text{الف)$$

$$Z_{0.025} = 1.96 \quad \text{ب)}$$

$$Z_{0.01} = 2.22 \quad \text{ج)}$$

$$Z_{0.005} = \frac{2/58 + 2/57}{2} = 2/575 \quad \text{د)}$$

.٤٣-٦

$$P(X \geq 14/5) = P\left(Z \geq \frac{14/5 - 37/5}{4/5}\right) = P(Z \geq 1/5) = 0.068 \quad \text{الف)$$

$$P(X \leq 35) = P\left(Z \leq \frac{35 - 37/5}{4/5}\right) = P(Z \leq 0/565) = 0.786 \quad \text{ب)}$$

$$P(30 < X < 15) = P\left(\frac{30 - 37/5}{4/5} < Z < \frac{15 - 37/5}{4/5}\right) = 0.649 \quad \text{ج)}$$

٦-٤. متغير استاندارد تمرین به صورت $Z = \frac{X - 15/4}{0/48}$ است.

$$P(X \geq 16) = P\left(Z \geq \frac{16 - 15/4}{0/48}\right) = P(Z \geq 1/25) = 0.1056 \quad \text{الف)$$

$$P(X \leq 14/2) = P\left(Z < \frac{14/2 - 15/4}{0/48}\right) = P(Z < -1/5) = 0.0062 \quad \text{ب)}$$

$$P(15 < X < 15/8) = P(-0.833 < Z < 0.833) = 0.5934 \quad \text{ج)}$$

٦-٥. با توجه به مجھول بودن میانگین متغير تصادفی داریم:

$$P(X < 28.2/5) = P\left(Z < \frac{28.2/5 - \mu}{0.48}\right) = 0.8212,$$

از روی اعداد جدول نتیجه می شود:

$$\frac{28.2/5 - \mu}{0.48} = 0.9, \quad \mu = 73/3.$$

احتمال مورد نظر برابر است با

$$P(X > 58/5) = P\left(Z > \frac{58/5 - 73/3}{0.48}\right) = P(Z > -1/5) = 0.9332.$$

٤-٦. در صورتی که $n\theta > 5$ و $n(1-\theta) > 5$ برقرار باشد تقریب مناسب است.

الف) $n\theta = 3/2 < 5$ تقریب مناسب نیست.

ب) $n\theta = 6/5 > 5$ تقریب مناسب است.

ج) $n(1-\theta) = 58/5 > 5$ تقریب مناسب است.

د) $n(1-\theta) = 2/42 < 5$ تقریب مناسب نیست.

٤-٧. الف) مشابهی تمرین ٤-٦، چون

$$n\theta = v/5 > 5, \quad n(1-\theta) = 142/5 > 5$$

برقرار است پس تقریب خوبی است.

ب) برای تقریب به نرمال میانگین و انحراف معیار به ترتیب برابرند با $\mu = n\theta = v/5$,

$$\sigma = \sqrt{n\theta(1-\theta)} = 2/67$$

$$\begin{aligned} b(v; 150, 5/100) &= P(-v/5 < X < v/5) = P(-v/5 < X < 1/5) \\ &= P\left(\frac{-v/5 - v/5}{2/67} < Z < \frac{1/5 - v/5}{2/67}\right) \\ &= P(-2/67 < Z < -2/25) = 0.0078. \end{aligned}$$

ج) مقدار خطای نسبی برابر است با

$$\frac{0.0078 - 0.0036}{0.0036} \times 100 = 116/67.$$

٤-٨. میانگین و واریانس عبارتند از

$$\mu = n\theta = 14 \times \frac{1}{2} = v, \quad \sigma = \sqrt{n\theta(1-\theta)} = 1/8 v.$$

احتمال موردنظر با انجام تصحیح پیوستگی برابر است با:

$$\begin{aligned} P(X=v) &= P(v/5 < X < v/5) = P(-v/14 < Z < v/14) \\ &= P\left(\frac{-v/14 - v}{1/8 v} < Z < \frac{v/14 - v}{1/8 v}\right) = 0.2128. \end{aligned}$$

در حالی که مقدار دقیق این احتمال معادل است با: $b(v, 14, 1/2) = 0.2095$, در نتیجه

مقدار خطای $0.2128 - 0.2095 = 0.0033$.

۴۹-۶. برای استفاده از تقریب نرمال، میانگین و انحراف معیار را به دست می‌آوریم:

$$\mu = n\theta = 120(0/23) = 27/6, \quad \sigma = \sqrt{n\theta(1-\theta)} = 4/6.$$

احتمال مورد نظر برابر است با

$$P(X > 32) = P(X > 32/5) = P\left(Z > \frac{32/5 - 27/6}{4/6}\right) = P(Z > 1/0.6) = 0.1446.$$

۵۰-۶. الف) نسبت شیرها $\frac{X}{100}$ ، دارای انحراف معیار و میانگین ۵ و $\mu = n\theta = 50$ است.

$$\begin{aligned} P(0.49 < \frac{X}{100} < 0.51) &= P(49 < X < 51) = P(X = 50) \\ &= P(49/5 < X < 50/5) \\ &= P\left(\frac{49/5 - 50}{5} < Z < \frac{50/5 - 50}{5}\right) \\ &= P(-0.1 < Z < 0.1) = 0.796. \end{aligned}$$

ب) برای ۱۰۰۰۰ ابار پرتتاب سکه داریم:

$$\mu = n\theta = 500, \quad \sigma = \sqrt{n\theta(1-\theta)} = 15.$$

$$\begin{aligned} P(0.49 < \frac{X}{100} < 0.51) &= P(490 < X < 510) = P(490 \leq X \leq 500) \\ &= P(490/5 < X < 500/5) \\ &= P\left(\frac{490/5 - 500}{15} < Z < \frac{500/5 - 500}{15}\right) \\ &= P(-0.63 < Z < 0.63) = 0.4714. \end{aligned}$$

ج) برای ۱۰۰۰۰۰ ابار پرتتاب

$$\mu = n\theta = 100000 \times \frac{1}{2} = 50000, \quad \sigma = 50.$$

$$\begin{aligned} P\left(0.49 < \frac{X}{100000} < 0.51\right) &= P(4900 < X < 5100) = P(4901 \leq X \leq 5099) \\ &= P\left(\frac{4901 - 5000}{50} \leq Z \leq \frac{5099 - 5000}{50}\right) = 0.9522. \end{aligned}$$

۵۱-۶. الف) اگر X و Y از هم مستقل باشند آنگاه $\text{cov}(X, Y) = 0$ ، در نتیجه

$$\rho = \frac{\text{cov}(X, Y)}{\sigma_1 \sigma_2} = \frac{0}{\sigma_1 \sigma_2} = 0.$$

ب) بنابر چگالی توأم نرمال دو متغیره با صفر بودن ρ ، داریم:

$$\begin{aligned} f(x, y) &= \frac{1}{2\pi\sigma_1\sigma_2} e^{-\frac{1}{2}\left[\left(\frac{x-\mu_1}{\sigma_1}\right)^2 + \left(\frac{y-\mu_2}{\sigma_2}\right)^2\right]} \\ &= \frac{1}{\sigma_1\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu_1}{\sigma_1}\right)^2} \frac{1}{\sigma_2\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{y-\mu_2}{\sigma_2}\right)^2} \\ &= f_1(x)f_2(y). \end{aligned}$$

۵۲-۶. از متحده ساختن عبارت

$$\frac{-1}{2(1-\rho^2)} \left\{ \left(\frac{x-\mu_1}{\sigma_1} \right)^2 - 2\rho \left(\frac{x-\mu_1}{\sigma_1} \right) \left(\frac{y-\mu_2}{\sigma_2} \right) + \left(\frac{y-\mu_2}{\sigma_2} \right)^2 \right\}$$

با عبارت

$$-\frac{1}{1-\rho^2} \left[(x+\rho)^2 - 2/\lambda(x+\rho)(y-1) + \rho(y-1)^2 \right]$$

داریم:

$$x^2 \text{ ضرب } = \frac{-1}{2\sigma_1^2(1-\rho^2)} = \frac{-1}{10^2}$$

$$y^2 \text{ ضرب } = \frac{-1}{2\sigma_2^2(1-\rho^2)} = \frac{-4}{10^2}$$

$$xy \text{ ضرب } = \frac{\rho}{\sigma_1\sigma_2(1-\rho^2)} = \frac{2/\lambda}{10^2}$$

دستگاهی با سه معادله و سه مجهول که از حل آن نتیجه می‌شود

$$\rho = 0/7, \quad \sigma_1 = 10, \quad \sigma_2 = 5.$$

در ادامه می‌توانید مقادیر دو میانگین را به دست آورید

$$\mu_1 = -2, \quad \mu_2 = 1.$$

۵۳-۷

$$\begin{aligned} \text{cov}(X+Y, X-Y) &= \text{cov}(X, X) - \text{cov}(X, Y) + \text{cov}(Y, X) - \text{cov}(Y, Y) \\ &= \sigma_X^2 - \sigma_Y^2, \end{aligned}$$

$$\text{var}(X + Y) = \text{var}(X) + \text{var}(Y) + \rho \text{cov}(X, Y) = \sigma_X^2 + \sigma_Y^2 + \rho \sigma_X \sigma_Y,$$

$$\text{var}(X - Y) = \text{var}(X) + \text{var}(Y) - \rho \text{cov}(X, Y) = \sigma_X^2 + \sigma_Y^2 - \rho \sigma_X \sigma_Y,$$

$$\rho_{u,v} = \frac{\text{cov}(u,v)}{\sigma_u \cdot \sigma_v} = \frac{\sigma_X^2 - \sigma_Y^2}{\sqrt{(\sigma_X^2 + \sigma_Y^2)^2 - 4\rho^2 \sigma_X^2 \sigma_Y^2}}.$$

(٦-٥٤. الف)

$$\begin{aligned} P(-90 < X < 90, -90 < Y < 90) &= \int_{-90}^{90} \int_{-90}^{90} f(x,y) dx dy \\ &= \int_{-90}^{90} f(x) dx \times \int_{-90}^{90} f(y) dy \\ &= [P(-90 < X < 90)]^2 \\ &= P\left(\frac{-90 - 0}{120} < Z < \frac{90 - 0}{120}\right) \\ &= 0.294. \end{aligned}$$

ب) با تغییر مختصات دکارتی به مختصات قطبی انتگرال زیر حل می شود.

$$\begin{aligned} P(X^2 + Y^2 < 4\pi^2) &= \int_0^{2\pi} \int_0^{2\pi} \frac{r}{2\pi(120)} e^{-\frac{r^2}{120}} dr d\theta \\ &= \int_0^{2\pi} \frac{1}{2\pi} \left[-e^{-\frac{r^2}{120}} \right]_{r=0}^{r=4\pi} d\theta = 0.1776. \end{aligned}$$

(٦-٥٤. الف)

$$\begin{aligned} P(X^2 + Y^2 < 4\pi^2) &= \int_0^{2\pi} \int_0^{\pi} \frac{r}{\pi 120} e^{-\frac{r^2}{120}} dr d\theta = \int_0^{2\pi} \frac{1}{2\pi} d\theta \times \int_0^{\pi} \frac{r}{\pi 120} e^{-\frac{r^2}{120}} dr \\ &= 1 \times \left(-e^{-\frac{r^2}{120}} \right)_0^\pi = 0.1175. \end{aligned}$$

(ب)

$$\begin{aligned} P(X^2 + Y^2 < c^2) &= 1 - e^{-\frac{c^2}{120}} = 0.1, \\ c &= 21/\sqrt{3}. \end{aligned}$$

سوالات تستی

۹. در سؤال هفت نما چقدر است؟

- الف) ۲۲/۵ ب) ۲۰ ج) ۲۱/۷ د) ۵۶/۲۰

۱۰. در توزیع‌هایی که چولگی زیاد ندارند کدام مشخص کننده‌ی مرکزی مناسب‌تر است؟

- الف) میانه ب) نما ج) میانگین حسابی د) میانگین هندسی

۱۱. اگر $x^3 = 9$ و $x+2=y$ ، انحراف معیار y چقدر است؟

- الف) ۹ ب) ۳ ج) ۵ د) ۱۱

۱۲. اگر $|S_K| = 0.1$ ، (ضریب چولگی پی‌رسون) کدامیک از عبارات زیر صحیح است؟

الف) توزیع چوله به چپ است

ج) توزیع چوله به راست است

۱۳. اگر میانگین پنج داده برابر ۱۰ و مجموع مجذورات آنها برابر ۵۲۰ باشد، واریانس آنها کدام است؟

- الف) ۹۲ ب) ۹۴ ج) ۹۳ د) ۹۵

۱۴. میانگین وزن ۱۰ نفر ۶۵ کیلوگرم است، دو نفر به وزن ۱۴۲ کیلوگرم به این افراد اضافه می‌شوند، میانگین جدید بر حسب گرم کدام است؟

- الف) ۶۹۰۰۰ ب) ۶۷۰۰۰ ج) ۶۶۰۰۰ د) ۶۵۰۰۰

۱۵. اگر میانگین اعداد از ۱ تا ۱۰۰ برابر ۵۰/۵ باشد، میانگین اعداد زوج از ۲ تا ۲۰۰ کدام است؟

- الف) ۵۰/۵ ب) ۱۰۰ ج) ۱۰۱ د) ۱۵

۱۶. اگر $\sum_{i=1}^5 x_i = 40$ و $\sum_{i=1}^5 x_i^2 = 10$ ، انحراف معیار کدام است؟

- الف) ۵ ب) ۲۵ ج) ۴ د) ۲

۱۷. میانگین ۵ داده برابر ۱۵ و میانگین ۱۰ داده‌ی آماری دیگر برابر ۴۵ است. میانگین کل داده‌ها برابر است با

- الف) ۱۲ ب) ۲۵ ج) ۳۵ د) ۴۵

۱۸. اگر داده‌ها نسبت به میانگین متقارن باشند، کدامیک از موارد زیر، برای گشتاور

مرکزی فرد برقرار است؟

الف) برابر صفر است

ب) وجود ندارد

ج) برابر میانگین است

د) نصف میانگین است

۱۹. اگر در یک سری نمونه‌ی تصادفی میانه برابر $\frac{8}{3}$ ، مد برابر $\frac{3}{2}$ ، و میانگین برابر $\frac{4}{3}$ باشد چوگنگی این توزیع چگونه است؟

الف) چوله به راست

ب) متقارن

ج) چوله به چپ

د) نامشخص

۲۰. به ازای چه مقداری از a ، عبارت $\sum_{i=1}^n |x_i - a|$ مینیمم می‌شود.

الف) میانه

ب) مد

ج) میانگین

د) انحراف معیار

۲۱. فرض کنید میانه ۵۰، دهک اول ۱۰، صدک نودم ۹۰، باشد ضریب چوگنگی کدام است؟

الف) ۵٪

ب) ۱۴٪

ج) ۰٪

د) ۲٪

۲۲. کدام گزینه صحیح است؟

الف) مهمترین پارامتر مرکزی میانه است

ب) همه‌ی توزیع‌های آماری دارای مد هستند

ج) همیشه نیم‌دامنه از دامنه تغییرات بیشتر است

د) مجموع انحراف‌ها از میانگین صفر است

۲۳. کدام یک از شاخص‌های زیر معیار مرکزی نیست؟

الف) میانگین

ب) میانه

ج) مد

د) دامنه

۲۴. در داده‌های آماری داده‌ای که ۲۵٪ داده‌ها از آن کمتر باشد و ۷۵٪ داده‌ها از آن بزرگتر باشد چه نامیده می‌شود؟

الف) چارک سوم

ب) چارک دوم

ج) چارک اول

د) صدک هفتاد و پنج

۲۵. در رسم نمودار دایره‌ای فراوانی یک رده برابر با ۱۵ و تعداد داده‌ها برابر ۹۰ می‌باشد، زاویه‌ی قطاع متناظر با این رده چند درجه است؟

الف) ۶۰

ب) ۸۰

ج) ۹۰

د) ۷۰

۲۶. حاصل عبارت $\frac{P_r}{\binom{n}{n+1} P_{r+1}}$ کدام است؟

$$\text{الف) } \frac{r-1}{n+1} \quad \text{ب) } \frac{1}{(n+1)!} \quad \text{ج) } \frac{r}{n}$$

۲۷. با ارقام ۵، ۴، ۳، ۲، ۱، ۰ چند عدد سه رقمی بزرگتر از ۳۰۰، بدون تکرار ارقام می‌شود نوشته؟

$$\text{الف) } ۱۲۰ \quad \text{ب) } ۸۰ \quad \text{ج) } ۱۴۰$$

۲۸. یک تاس و یک سکه را با هم پرتاب می‌کنیم. تعداد صورت‌هایی که در آن عدد زوج آمده، کدام است؟

$$\text{الف) } ۲ \quad \text{ب) } ۳ \quad \text{ج) } ۴ \quad \text{د) } ۱$$

۲۹. سه کتاب ریاضی و دو کتاب اقتصاد را که با هم متفاوتند، به چند طریق می‌توان در یک قفسه کنار هم قرار داد به طوری که کتاب‌های هم موضوع، همواره کنار هم باشند؟

$$\text{الف) } ۶۰ \quad \text{ب) } ۱۲۰ \quad \text{ج) } ۱۲ \quad \text{د) } ۱$$

۳۰. اگر ترکیب $\binom{a+b}{b} = m$ باشد مقدار ترکیب $\binom{a+b}{a}$ کدام است؟

$$\text{الف) } m \quad \text{ب) } am \quad \text{ج) } bm \quad \text{د) } (a+m)m$$

۳۱. شخصی از میان ۱۰ کتاب خود، می‌خواهد دو کتاب را انتخاب کرده، به یکی از دوستانش هدیه کند، این عمل به چند طریق ممکن است؟

$$\text{الف) } ۴۰ \quad \text{ب) } ۴۵ \quad \text{ج) } ۹۰ \quad \text{د) } ۲۰$$

۳۲. مجموعه $\{a, b, c, d, e\}$ چند زیرمجموعه دو عضوی دارد؟

$$\text{الف) } ۲۵ \quad \text{ب) } ۲۰ \quad \text{ج) } ۱۵ \quad \text{د) } ۱۰$$

۳۳. بین ۵ معلم و ۴ دانش‌آموز چند کمیته ۵ نفری مرکب از ۳ معلم و ۲ دانش‌آموز می‌توان تشکیل داد؟

$$\text{الف) } ۶ \quad \text{ب) } ۱۰ \quad \text{ج) } ۱۵ \quad \text{د) } ۶۰$$

۳۴. اگر $P_n = ۲۰$ ، مقدار n کدام است؟

$$\text{الف) } ۲ \quad \text{ب) } ۵ \quad \text{ج) } ۱۰ \quad \text{د) } ۲۰$$

۳۵. جواب معادله $C_x^3 = ۲x$ کدام است؟

$$\text{الف. . } ۲ \quad \text{ب) } ۳ \quad \text{ج) } ۴ \quad \text{د) } ۵$$

۳۶. به چند طریق می‌توان ۱۰ سبب نامتمایز را بین ۳ نفر تقسیم کرد به‌طوری‌که لااقل یک سبب به هر نفر برسد؟

$$\text{الف) } \binom{10}{3} \quad \text{ب) } \binom{10}{2} \quad \text{ج) } \binom{10}{1}$$

۳۷. به چند طریق می‌توان m مهره‌ی نامتمایز سفید و n مهره‌ی نامتمایز سیاه را در یک ردیف قرار داد که هیچ دو مهره‌ی سفیدی کنار هم نباشند؟ (m و n در شرایط مناسب صدق می‌کنند)

$$\text{الف) } \binom{n}{m} \quad \text{ب) } \binom{n+1}{m} \quad \text{ج) } \binom{m+1}{n}$$

۳۸. به چند طریق می‌توان ۲۰ مهره‌ی نامتمایز را در ۱۵ ظرف متمایز قرار داد به‌طوری که ظرف شماره‌ی ۷ درست ۴ مهره داشته باشد؟

$$\text{الف) } \binom{20}{14} \quad \text{ب) } \binom{20}{13} \quad \text{ج) } \binom{20}{12}$$

۳۹. ۴ زوج (زن و شوهر) به چند طریق می‌توانند روی یک ردیف صندلی بنشینند به‌طوری‌که هر زوج کنار هم باشند؟

$$\text{الف) } 2 \times 2 \times 2 \times 2 \quad \text{ب) } 2 \times 2 \times 2 \times 2 \quad \text{ج) } 2 \times 2 \times 2 \times 2$$

۴۰. تعداد جواب‌های صحیح و نامنفی نامعادله‌ی $x_1 + x_2 + x_3 + x_4 \leq 7$ کدام است؟

$$\text{الف) } \binom{7}{4} \quad \text{ب) } \binom{10}{3} \quad \text{ج) } \binom{9}{2}$$

۴۱. اگر در یک سیستم الکترونیکی هر کلید فقط بتواند برای ارسال دو پیام به کار رود و کلیدها مستقل از هم باشد برای مبادله‌ی ۲۱۹۷ پیام به چند کلید احتیاج است؟

$$\text{الف) } 10 \text{ کلید} \quad \text{ب) } 11 \text{ کلید} \quad \text{ج) } 12 \text{ کلید} \quad \text{د) } 13 \text{ کلید}$$

۴۲. شهری ۴ تعمیرکار تلویزیون دارد. اگر ۴ دستگاه تلویزیون خراب شود به چند طریق ممکن است فقط و فقط دو نفر از این تعمیرکارها انتخاب شوند؟

$$\text{الف) } \binom{4}{2} \quad \text{ب) } \binom{4}{3} \quad \text{ج) } \binom{4}{1}$$

۴۳. کدام عبارت صحیح است؟

$$\sum_{k=1}^n (-1)^k \binom{n}{k} = 0 \quad \text{الف) } \sum_{k=0}^n \binom{n}{k}^2 = \binom{2n}{n}$$

$$\binom{m+n-k}{n+k-m} = \binom{m+n}{n+k} \quad (d) \quad \binom{r+2}{n} = \binom{r+1}{n-1} + \binom{r}{n-1} \quad (c)$$

۴۴. در یک لیگ مسابقات حذفی فوتبال ۱۵ تیم شرکت دارند، به چند طریق این پانزده تیم می‌توانند در دوره‌های اول و دوم با هم بازی کنند؟

$$(f) \frac{15!}{2^7} \quad (b) \frac{15!}{2^7} \quad (c) \frac{15!}{2^8} \quad (d) \frac{15!}{2^8}$$

۴۵. ۱۸ تخته سیاه نامتمایز را به چند طریق می‌توان به ۸ مدرسه تقسیم کرد؟

$$(a) 8^{18} \quad (b) \binom{18+8-1}{18} \quad (c) \binom{18+8-1}{8} \quad (d) \binom{18+8-1}{18}$$

۴۶. حاصل عبارت $\binom{n+2}{r-1} + \binom{n+2}{r}$ کدام است؟

$$(a) \binom{n+1}{r-1} \quad (b) \binom{n+3}{r} \quad (c) \binom{n+3}{r-1} \quad (d) \binom{n+1}{r}$$

۴۷. ۱۵ سیب نامتمایز را به چند طریق می‌توان بین سه نفر تقسیم کرد؟

$$(a) \binom{14}{3} \quad (b) \binom{15}{3} \quad (c) \binom{17}{3} \quad (d) \binom{17}{2}$$

۴۸. به چند طریق ۶ زوج (زن و شوهر) می‌توانند در یک ردیف بنشینند به‌قسمی که زن‌ها کنار هم باشند؟

$$(a) 12 \quad (b) 2^{(12)} \quad (c) 7! \quad (d) 12!$$

۴۹. در سؤال قبل چند حالت هیچ دو زنی کنار هم نیستند؟

$$(a) 2^{(12)} \quad (b) 3^{(12)} \quad (c) 7! \quad (d) 12!$$

۵۰. به چند طریق می‌توان به ۵ سؤال ۴ جوابی پاسخ داد؟

$$(a) 4^5 \quad (b) 5^4 \quad (c) \binom{4+5-1}{4} \quad (d) \binom{4+5-1}{5}$$

۵۱. به چند طریق می‌توان نمره‌ی ماشین ۷ رقمی تهیه کرد اگر بخواهیم دو رقم اول از بین ۲۶ حرف انگلیسی و بقیه از ارقام صفر تا ۹ انتخاب شود؟

$$(a) 26 \times 25 \times 10^5 \quad (b) 26 \times 10^5 \quad (c) 26 \times 25 \times 10^3 \quad (d) \binom{26}{2} \binom{10}{5}$$

۵۲. تعداد اعداد چهار رقمی بدون تکرار با استفاده از ارقام ۰، ۱، ۲، ۳، ۴، ۵، ۶، ۷، ۸، ۹ که بر ۵ قابل قسمت باشد چقدر است؟

الف) ۱۸۰ ب) ۲۲۰ ج) ۵۵ د) ۵۰

۵۳. به چند طریق می‌توان ۱۰ نفر را به دو تیم ۵ نفره تقسیم نمود که با هم بازی کنند؟

$$P_{10} = \frac{\binom{10}{5}}{\binom{10}{5}} = \frac{10!}{5!5!}$$

الف) $\binom{10}{5}$
ب) $\binom{10}{5}$

۵۴. از بین ۵ مرد و ۸ زن یک گروه سه نفری انتخاب می‌شوند به چند طریق یک مرد خاص در این گروه هست؟

الف) ۵۰ ب) ۶۶ ج) ۱۴۰ د) ۲۸۶

۵۵. چند عدد سه رقمی زوج می‌توان به وسیلهٔ ارقام ۹, ۶, ۵, ۴, ۲, ۱ ساخت که رقم تکراری نداشته باشدند؟

الف) ۶ ب) ۱۲ ج) ۲۴ د) ۳۶

۵۶. به چند طریق می‌توان ۸ سبب را بین ۴ نفر تقسیم کرد که اولاً همه سبب‌ها تقسیم شوند؛ ثانیاً بریدن سبب مجاز نباشد؛ ثالثاً به هر نفر حداقل یک سبب برسد؟

$$\binom{7}{3} = \binom{8}{4}$$

الف) $\binom{7}{3}$
ب) $\binom{8}{4}$

۵۷. کدام تساوی درست نیست؟

$$\binom{r-3}{k+1} = \binom{r-2}{k} + \binom{r-4}{k-1}$$

الف) $\sum_{k=1}^n \binom{n}{k} = 2^n - n - 1$

$$\sum_{r=0}^n \binom{n}{r}^2 = \binom{2n}{n}$$

ج) $\sum_{k=0}^n k \binom{n}{k} = n 2^{n-1}$

۵۸. به چند طریق می‌توان ده نفر را در پنج اتاق دو نفره نامماییز جا داد؟

$$\binom{10}{5} = \frac{10!}{(2!)^5}$$

الف) $\frac{10!}{(2!)^5}$

۵۹. به چند طریق پنج زوج می‌توانند دور یک میز بشینند به‌طوری که هر زوج کنار هم باشند؟

$$\frac{9!}{5!}$$

الف) $\frac{9!}{5!}$

۶۰. به چند طریق می‌توان صد مهره‌ی نامماییز را در هفت جعبهٔ جای داد به‌طوری که جعبه‌ی i ام حداقل $(i\text{زوج})^2$ مهره داشته باشد؟

$$\text{الف) } \binom{16}{7} \quad \text{ب) } \binom{16}{6} \quad \text{ج) } \binom{16}{7}$$

۶۱. در صفحه‌ی شطرنج چند مربع وجود دارد؟

$$\text{الف) } 201 \quad \text{ب) } 202 \quad \text{ج) } 203 \quad \text{د) } 204$$

۶۲. به چند طریق می‌توان ۵ نامه را به تصادف در پاکت‌های آنها جای داد به‌طوری‌که هیچ نامه‌ای در پاکت خود نباشد؟

$$\text{الف) } 44 \quad \text{ب) } 45 \quad \text{ج) } 46 \quad \text{د) } 47$$

۶۳. عدد طبیعی $N = P_1^{n_1} P_2^{n_2} \dots P_k^{n_k}$ که در آن P_1, P_2, \dots, P_k اعداد اول متمایزند.

$$\prod_{i=1}^k (n_i + 1) \quad \text{د) } \quad \prod_{i=1}^k (n_i - 1) \quad \text{ج) } \quad \prod_{i=1}^k (1 + n_i) \quad \text{ب) } \quad \prod_{i=1}^k n_i \quad \text{الف) }$$

۶۴. هفت پسر و هشت دختر روی ۱۵ صندلی در یک ردیف می‌نشینند در چند حالت صندلی شماره‌ی ۱۱ توسط یک پسر اشغال می‌شود؟

$$\text{الف) } 15! \quad \text{ب) } 14!(14!) \quad \text{ج) } 8(14!) \quad \text{د) } 8-15!$$

۶۵. احتمال آنکه خانواده‌ای که سه فرزند دارند، حداقل دارای دو فرزند پسر باشد کدام است؟

$$\text{الف) } \frac{5}{8} \quad \text{ب) } \frac{1}{8} \quad \text{ج) } \frac{3}{8} \quad \text{د) } \frac{7}{8}$$

۶۶. اگر A و B دو پیشامد ناسازگار باشند و $P(A) = \frac{1}{3}$, $P(B) = \frac{1}{4}$ حاصل $P(A \cup B)$ برابر است با

$$\text{الف) } \frac{4}{12} \quad \text{ب) } \frac{3}{12} \quad \text{ج) } \frac{1}{2} \quad \text{د) } \frac{7}{12}$$

۶۷. در کیسه‌ای ۵ مهره‌ی سفید و ۴ مهره‌ی آبی وجود دارد. اگر دو مهره به تصادف خارج کنیم احتمال اینکه دو مهره همنگ باشند، کدام است؟

$$\text{الف) } \frac{2}{9} \quad \text{ب) } \frac{5}{9} \quad \text{ج) } \frac{7}{9} \quad \text{د) } \frac{4}{9}$$

۶۸. یک تیرانداز به‌طور متوسط نصف تیرهای خود را به هدف می‌زند. در صورتی که این تیرانداز ۵ بار شلیک نماید احتمال اینکه ۳ بار تیر را به هدف بزند چقدر است؟

$$\text{الف) } \frac{3}{16} \quad \text{ب) } \frac{5}{16} \quad \text{ج) } \frac{7}{16} \quad \text{د) } \frac{9}{16}$$

۶۹. یک تاس را دو بار پرتاب می‌کنیم احتمال بهدست آمدن ۱ و ۶ کدام است؟

- الف) $\frac{7}{36}$ ب) $\frac{5}{36}$ ج) $\frac{11}{36}$ د) $\frac{6}{36}$

۷۰. یک سکه را ۵ بار پرتاب می‌کنیم، احتمال بهدست آمدن حداقل یک شیر کدام است؟

- الف) $\frac{3}{32}$ ب) $\frac{1}{32}$ ج) $\frac{15}{32}$ د) $\frac{31}{32}$

۷۱. اگر $P(A \cup B)$, $P(B)$, $P(A|B)$ کدام است؟

- الف) $\frac{1}{12}$ ب) $\frac{1}{36}$ ج) $\frac{1}{3}$ د) $\frac{5}{12}$

۷۲. ۱۰ کارت را به شماره‌های ۱ تا ۱۰ در یک جعبه قرار می‌دهیم. از این جعبه ۳ کارت را پی‌درپی بیرون می‌آوریم. احتمال اینکه شماره‌های این کارت به ترتیب فرد و زوج و فرد باشد، کدام است؟

- الف) $\frac{1}{720}$ ب) $\frac{5}{36}$ ج) $\frac{5}{30}$ د) $\frac{5}{12}$

۷۳. اگر A و B دو پیشامد باشند مقدار $P(A \cap B) + P(A - B)$ کدام است؟

- الف) $P(A \cup B)$ ب) $P(B')$ ج) $P(A)$ د) $P(B)$

۷۴. دو مکعب پرتاب می‌شود احتمال اینکه مجموع دو مکعب بیشتر از ۱۰ باشد در صورتی که مکعب اول ۵ باشد کدام است؟

- الف) $\frac{1}{3}$ ب) $\frac{1}{6}$ ج) $\frac{3}{5}$ د) $\frac{2}{5}$

۷۵. دو رقم به‌طور تصادفی از بین ارقام ۱ تا ۶ انتخاب می‌کنیم. فرض کنیم مجموع دو رقم زوج باشد. احتمال اینکه هر دو رقم فرد باشد کدام است؟

- الف) $\frac{1}{3}$ ب) $\frac{1}{4}$ ج) $\frac{2}{3}$ د) $\frac{2}{5}$

۷۶. احتمال قبول شدن محمد در کنکور به قبول نشدنش $\frac{1}{3}$ است، احتمال قبول شدن محمد در کنکور کدام است؟

- الف) $\frac{1}{2}$ ب) $\frac{1}{3}$ ج) $\frac{1}{4}$ د) $\frac{2}{3}$

۷۷. دو مکعب را با هم می‌اندازیم اگر بدانیم مجموع ۵ آمده است، احتمال آنکه هر دو عدد اول باشد کدام است؟

- الف) صفر ب) $\frac{1}{3}$ ج) $\frac{1}{9}$

۷۸. در یک جعبه ۳ مهره‌ی قرمز، ۴ مهره‌ی سفید، و ۵ مهره‌ی آبی وجود دارد. ۳ مهره به تصادف از جعبه بیرون می‌آوریم، احتمال آنکه هر کدام از یک رنگ باشند، کدام است؟

- الف) $\frac{5}{11}$ ب) $\frac{5}{12}$ ج) $\frac{5}{13}$

۷۹. یک سکه را پنج بار پرتاب می‌کنیم احتمال آنکه سه بار خط بیاید کدام است؟

- الف) $\frac{3}{22}$ ب) $\frac{1}{32}$ ج) $\frac{5}{16}$

۸۰. اگر $P(A|B)=\frac{1}{3}$ ، $P(B)=\frac{1}{5}$ ، $P(A)=\frac{1}{3}$ باشد، مقدار $P(A \cup B)$ کدام است؟

- الف) $\frac{1}{2}$ ب) $\frac{6}{15}$ ج) $\frac{7}{15}$

۸۱. اگر $P(A)=0/5$ ، $P(B)=0/7$ ، $P(A \cup B)=0/7$ و A و B مستقل باشند، مقدار $P(B)$ کدام است؟

- الف) $\frac{2}{17}$ ب) $0/4$ ج) $\frac{2}{7}$

۸۲. سه پیچ و سه مهره در یک جعبه قرار دارند اگر ۲ شیء از این ۶ شیء را به تصادف از جعبه بیرون آوریم، احتمال اینکه یک پیچ و یک مهره به دست آید، کدام است؟

- الف) $\frac{1}{3}$ ب) $\frac{2}{9}$ ج) $\frac{3}{5}$

۸۳. سه شناگر A ، B ، و C ، در یک مسابقه شنا شرکت کردند. در صورتی که احتمال برد A با B برابر باشد و احتمال برد A و B هر کدام دو برابر احتمال برد C باشد احتمال برد شناگر A کدام است؟

- الف) $\frac{1}{5}$ ب) $\frac{3}{10}$ ج) $\frac{2}{5}$

۸۴. دو عدد حقیقی به تصادف در بازه‌ی $[0, 2]$ انتخاب می‌کنیم. احتمال اینکه مجموع مربعات آن دو عدد کوچکتر یا مساوی 4 باشد، کدام است؟

- الف) $\frac{1}{2}$ ب) $\frac{\pi}{4}$ ج) $\frac{1}{4}$

۸۵. عددی به تصادف از فضای نمونه‌ای $\{1, 2, 3, \dots, 9\}$ انتخاب می‌کنیم. اگر بدانیم عدد انتخاب شده، مضرب سه است، احتمال آنکه عدد رو شده زوج باشد، چقدر است؟

- الف) $\frac{1}{3}$ ب) $\frac{1}{9}$ ج) $\frac{4}{9}$

۸۶. اگر A و B ، دو پیشامد مستقل باشند و $P(A \cup B) = 0.4$ ، $P(A) = 0.2$ ، مقدار $P(B)$ کدام است؟

- الف) 0.3 ب) 0.2 ج) 0.25

۸۷. از ظرفی با ۳ مهره‌ی سفید و ۲ مهره‌ی سیاه، ۲ مهره با هم بیرون می‌کشیم، احتمال آنکه حداقل یک مهره سیاه بیرون آمده باشد، کدام است؟

- الف) $\frac{9}{10}$ ب) $\frac{7}{10}$ ج) $\frac{6}{10}$

۸۸. توزیع درصد لامپ‌های ۶۰ و ۱۰۰ وات از دو کارخانه (الف) و (ب) به صورت زیر است. احتمال آنکه لامپی از کارخانه (الف) ۱۰۰ وات باشد کدام است؟

| نوع کارخانه | الف | ب |
|-------------|-----|----|
| | ۱۰۰ | ۲۵ |
| | ۶۰ | ۱۵ |

الف) $\frac{4}{7}$ ب) $\frac{5}{18}$ ج) $\frac{7}{20}$

۸۹. در پرتاب دو تاس A و B همانند، مشروط بر آنکه بدانیم مجموع دو عدد مضرب ۶ باشند، احتمال آنکه هر دو عدد زوج باشد کدام است؟

- الف) $\frac{2}{5}$ ب) $\frac{3}{5}$ ج) $\frac{5}{8}$

۹۰. اگر بخواهیم از بین ۷ گلوله‌ی سفید و ۸ گلوله‌ی قرمز یکی را انتخاب کنیم، تعداد حالت‌های ممکن چندتاست؟

الف) ۵۶ ب) ۶۵ ج) ۷ د) ۸

۹۱. به چند طریق می‌توان از بین ۱۰ نفر یک رئیس، یک معاون، یک خزانه‌دار انتخاب کرد؟

$$\text{الف) } \frac{1}{10} \cdot P_{\binom{10}{3}} = \frac{1}{10} \cdot \frac{10!}{3!(10-3)!}$$

۹۲. اگر A و B دو پیشامد ناسازگار باشد کدام گزاره درست است؟

$$P(A \cup B) = P(A) + P(B) \quad \text{الف) } P(A \cap B) = P(A) \cdot P(B)$$

$$P(A|B) = \frac{P(A \cap B)}{P(B)} \quad \text{د) } P(A \cup B) = P(A) + P(B) - P(A \cap B) \quad \text{ج) }$$

$$P(A|B) = P(A \cap B) = 0/05, P(B) = 0/25, P(A) = 0/2 \quad \text{آنگاه مقدار}$$

۹۳. اگر A کدام است؟

$$\text{الف) } 0/2 \quad \text{ب) } 0/25 \quad \text{ج) } 0/02 \quad \text{د) } 0/75$$

۹۴. در جعبه‌ی شماره‌ی ۱، تعداد ۴ مهره‌ی قرمز، ۳ مهره‌ی سفید و در جعبه‌ی شماره‌ی ۲ تعداد ۴ مهره‌ی قرمز و ۵ مهره‌ی سفید قرار دارد. یکی از جعبه‌ها را به تصادف انتخاب و سپس مهره‌ای از آن به تصادف انتخاب می‌کنیم. احتمال آنکه مهره‌ی انتخاب شده قرمز باشد برابر است با

$$\text{الف) } \frac{38}{45} \quad \text{ب) } \frac{19}{45} \quad \text{ج) } \frac{4}{9} \quad \text{د) } \frac{3}{7}$$

۹۵. در سؤال قبل اگر مهره قرمز باشد چقدر احتمال دارد که از ظرف شماره‌ی ۱

انتخاب شده باشد؟

$$\text{الف) } \frac{8}{19} \quad \text{ب) } \frac{9}{19} \quad \text{ج) } \frac{10}{19} \quad \text{د) } \frac{7}{19}$$

۹۶. اگر A و B دو پیشامد مستقل باشند کدام نادرست است؟

$$P(B|A) = P(B) \quad \text{الف) } P(A|B) = P(A)$$

$$P(A \cap B) = P(A) \cdot P(B) \quad \text{د) } P(A \cup B) = P(A) \cdot P(B)$$

۹۷. مقدار $P(A|A \cup B)$ برابر است با

$$\frac{P(A)}{P(A \cup B)} \text{ د) } \quad P(A) \text{ ج) } \quad 1 + P(A|B) \text{ ب) } \quad \frac{1}{P(A \cup B)} \text{ الف) }$$

۹۸. A و B دو پیشامد از فضای نمونه‌ای S اند به طوری که $P(A) = 1$, در این صورت

$$A - B = B^c \text{ ب) } \quad A \cup B \text{ دو پیشامد ناسازگارند الف)}$$

$$A = S \text{ الف) } \quad A \text{ از } B \text{ مستقل است ج) }$$

۹۹. فرض کنید برای دو پیشامد A و B داشته باشیم: $P(B) = 0/3$, $P(A) = 0/59$,

$$P(A \cap B) = 0/21. P(A'|B) = 0/21 \text{ برابر است با}$$

$$\text{الف) } 0/69 \quad \text{ب) } 0/46 \quad \text{ج) } 0/3 \quad \text{د) } 0/21$$

۱۰۰. A و B دو پیشامد مستقل اند و $P(B) = 0/2$, $P(A) = 0/3$. حاصل $P(A \cup B)$ با

برابر است با

$$\text{الف) } 0/4 \quad \text{ب) } 0/5 \quad \text{ج) } 0/44 \quad \text{د) } 0/56$$

۱۰۱. A و B مستقل اند، $P(A) = \frac{1}{3}$, $P(A \cap B) = \frac{1}{8}$. مقدار $P(B)$ کدام است؟

$$\text{الف) } \frac{1}{4} \quad \text{ب) } \frac{1}{16} \quad \text{ج) } \frac{3}{8} \quad \text{د) } \frac{5}{8}$$

۱۰۲. از ظرفی با دو مهره‌ی سفید و ۳ مهره‌ی سیاه، دو مهره با هم بیرون می‌آوریم.

احتمال آنکه حداقل یک مهره‌ی سیاه بیرون آمده باشد، کدام است؟

$$\text{الف) } \frac{3}{8} \quad \text{ب) } \frac{6}{10} \quad \text{ج) } \frac{7}{8} \quad \text{د) } \frac{9}{10}$$

۱۰۳. امید ریاضی تابع توزیع زیر کدام است؟

| | | | | |
|--------|---------------|---------------|---------------|---------|
| x | ۲ | ۳ | ۵ | |
| $f(x)$ | a | $\frac{1}{2}$ | $\frac{1}{6}$ | |
| | $\frac{1}{2}$ | $\frac{1}{6}$ | | |
| | $\frac{1}{6}$ | $\frac{5}{6}$ | $\frac{1}{3}$ | الف) ۱۰ |

۱۰۴. اگر جدول توزیع X به صورت

| | | | | | |
|------------|----------------|----------------|----------------|----------------|----------------|
| $X = x$ | ۱ | $\frac{3}{2}$ | ۲ | $\frac{2}{5}$ | ۳ |
| $P(X = x)$ | $\frac{4}{25}$ | $\frac{4}{25}$ | $\frac{9}{25}$ | $\frac{4}{25}$ | $\frac{4}{25}$ |

باشد، $E(X - 2)$ برابر است با

الف) ۰ ب) ۱ ج) ۲ د) ۴

۱۰۵. در صورتی که $f(x) = \begin{cases} Kx & 0 \leq x \leq 8 \\ 0 & \text{سایرجاها} \end{cases}$ یک تابع چگالی احتمال از متغیر پیوسته X باشد، مقدار K کدام است؟

الف) $\frac{1}{16}$ ب) $\frac{1}{8}$ ج) $\frac{1}{32}$ د) $\frac{1}{4}$

۱۰۶. متغیر تصادفی X دارای میانگین ۵ و واریانس ۴ است. متغیر تصادفی Y از رابطه‌ی $Y = 2X - 3$ به دست می‌آید. میانگین و انحراف معیار متغیر به ترتیب برابر است با

الف) ۸, ۱۰ ب) ۷, ۸ ج) ۴, ۷ د) ۴, ۱۰

۱۰۷. تابع چگالی متغیر تصادفی X به صورت زیر است.

| | | | |
|--------|---------------|---------------|---------------|
| x | -1 | 1 | 2 |
| $f(x)$ | $\frac{1}{2}$ | $\frac{1}{3}$ | $\frac{1}{6}$ |
| | ۳ | ۲ | ۱ |

مقدار واریانس X برابر است با

الف) $\frac{1}{36}$ ب) $\frac{5}{36}$ ج) $\frac{13}{36}$ د) $\frac{25}{6}$

۱۰۸. به ازای چه مقداری از A تابع زیر یک تابع چگالی احتمال است؟

| | | | |
|--------|-----------------|---------------|-----------------|
| x | a | b | c |
| $f(x)$ | $\frac{1-A}{3}$ | $\frac{1}{3}$ | $\frac{1+A}{3}$ |
| | ۳ | ۳ | ۳ |

۰ < A < ۱ ۱ < A < ۲ -۲ < A < ۱ -۱ < A < ۱ الف) ۱ ب) ۱ ج) ۲ د) ۱

۱۰۹. میانه‌ی متغیر تصادفی با تابع توزیع $F(x) = \begin{cases} ۰ & x < ۰ \\ x^r & ۰ \leq x < ۱ \\ ۱ & ۱ \leq x \end{cases}$ چقدر است؟

الف) $\frac{1}{3}$ ب) $\frac{\sqrt{3}}{3}$ ج) $\frac{4}{21}$ د) $\frac{15}{21}$

۱۱۰. اگر $P(X + Y < 1)$ ، آنگاه $f(x, y) = \frac{1}{\pi}$ برابر است با

الف) $\frac{1}{2}$ ب) $\frac{1}{2\pi}$ ج) $\frac{1}{3} + \frac{1}{3\pi}$ د) $\frac{3}{4}$

۱۱۱. اگر $P(X + Y \geq 3) = f(x, y) = \frac{x+y}{21}$ و $x = 1, 2$ و $y = 1, 2, 3$ برابر است با

الف) ۱ ب) $\frac{19}{21}$ ج) $\frac{6}{21}$ د) $\frac{15}{21}$

۱۱۲. اگر $Y = \ln - X$ ، تابع مولد $M_X(t) = \frac{1}{\lambda}(1 + e^t)$ چقدر است؟

- (الف) $e^{-\lambda t}$ (د) $\frac{1}{\lambda}(1 + e^t)$ (ج) $e^{\lambda t}(1 + e^t)^{-1}$ (ب) $M_X(t)$

۱۱۳. اگر $\text{cov}(X, Y) = 4$ و $\text{var}(X + Y) = 6$ آنگاه مقدار $\text{var}(X - Y)$ برابر است با

- (الف) $\frac{1}{2}$ (د) $-\frac{3}{4}$ (ج) $\frac{3}{4}$ (ب) $\frac{1}{2}$

۱۱۴. اگر $E(X) = 2$ و $E(X^2) = 6$ دوّمین گشتاور عامل برابر است با

- (الف) ۶ (ب) ۴ (ج) ۲ (د) ۸

۱۱۵. اگر بدانیم که $X + Y = 4$ آنگاه $E(X|Y=4)$ برابر است با

- (الف) ۴ (ب) ۸ (ج) ۴ (د) صفر

۱۱۶. اگر $f(x,y) = \begin{cases} kxy & x=1,2,3 \\ 0 & y=1,2,3 \\ \text{سایر جاهای} & \end{cases}$ احتمال توأم X و Y باشد مقدار

$f(1,1)$ کدام است؟

- (الف) $\frac{1}{36}$ (د) $\frac{4}{36}$ (ج) $\frac{3}{36}$ (ب) $\frac{2}{36}$

۱۱۷. اگر $\text{cov}(3X + 2, 2Y - 1) = 10$ ، $\text{var}(Y) = 25$ ، $\text{var}(X) = 26$ مقدار $\text{cov}(X, Y)$ برابر است با

و $\text{var}(X - Y)$ برابر است با

- (الف) ۲۰، -۶۰ (د) ۱۱، -۴۰ (ج) ۴۶، -۱۲۰ (ب) ۲۰، -۶۰

۱۱۸. برای توزیع احتمال جدول زیر، مقدار $E(2X - 8)$ برابر است با

| | | | | | |
|--------|-------|-------|-------|-------|-------|
| x | ۲ | ۳ | ۴ | ۵ | ۶ |
| $f(x)$ | $0/1$ | $0/3$ | $0/3$ | $0/2$ | $0/1$ |
| | $5/2$ | $4/2$ | $3/5$ | $2/5$ | |

- (الف) $2/5$ (د) $5/2$ (ج) $3/5$ (ب) $4/2$

۱۱۹. توزیع احتمال متغیر تصادفی X به صورت زیر است.

| | | | | |
|--------|---------------|---------------|---------------|---------------|
| x | -1 | 0 | 1 | 2 |
| $P(x)$ | $\frac{1}{4}$ | $\frac{1}{4}$ | $\frac{3}{8}$ | $\frac{1}{8}$ |

واریانس متغیر تصادفی X است

- (الف) $\frac{36}{64}$ (د) $\frac{5}{16}$ (ب) صفر (ج) $\frac{1}{16}$

۱۲۰. عبارت $E(X - \mu[X - \mu])$ برابر است با

- (الف) صفر (ب) امید ریاضی X

ج) امید ریاضی انحراف‌ها د) واریانس X

۱۲۱. ۷۰ درصد تولیدات کارخانه‌ای سالم است. یک نمونه‌ی ۲۰ تایی از محصولات این کارخانه انتخاب می‌کنیم، انتظار می‌رود چند کالا معیوب باشند؟

الف) ۸ ب) ۶ ج) ۱۲ د) ۱۴

۱۲۲. احتمال آنکه یک دانه‌ی لوبيا جوانه بزند ۸/۰ است. اگر چهار دانه از اين لوبيا را بكاريم احتمال آنکه فقط ۳ تاي آن جوانه بزند چقدر است؟

الف) ۰/۲۴۹۶ ب) ۰/۴۰۹۶ ج) ۰/۸۳۳۳ د) ۰/۱۰۳۳

۱۲۳. نوعی واکسن با احتمال ۹۰ درصد برای پرندگان تأثیر مثبت دارد. اگر ۵ مورد از اين واکسن را به کار رود با کدام احتمال فقط ۳ مورد آن تأثیر مثبت خواهد داشت؟

الف) ۰/۰۷۲۹ ب) ۰/۵۸۱ ج) ۰/۷۲۹ د) ۸۱

۱۲۴. انحراف معیار توزیع دوجمله‌ای کدام است؟

الف) \sqrt{npq} ب) \sqrt{np} ج) $n(1-q)$ د) np

۱۲۵. اگر سطح زير منحنی نرمال استاندارد بین صفر و یک ۳۴٪ باشد سمت راست عدد ۱ برابر است با

الف) ۰/۳۴ ب) ۰/۱۶ ج) ۰/۶۸ د) ۰/۸۴

۱۲۶. در سئوال قبل چپ ۱ برابر است با

الف) ۰/۱۶ ب) ۰/۳۴ ج) ۰/۸۴ د) ۰/۶۸

۱۲۷. برای تابع چگالی سایر جاهای $f_X(x) = \begin{cases} \lambda e^{-\lambda x} & x \geq 0 \\ 0 & \text{سایر جاهای} \end{cases}$ میانه‌ی توزیع کدام است؟

الف) $\ln^{\frac{1}{\lambda}} \lambda$ ب) $\frac{\ln^{\frac{1}{\lambda}} \lambda}{\lambda}$ ج) $\frac{-\ln^{\frac{1}{\lambda}} \lambda}{\lambda}$ د) λ

۱۲۸. تعداد کلمات غلط ماشین شده در یک صفحه از یک کتاب از توزیع پواسون با

پارامتر $\frac{1}{\mu} = \lambda$ تبعیت می‌کند. احتمال اینکه در دو صفحه حداقل یک غلط وجود داشته باشد، چقدر است؟

الف) $e^{-\frac{1}{\mu}}$ ب) $1 - e^{-\frac{1}{\mu}}$ ج) $e^{-\frac{1}{\mu}}$ د) $e^{-\frac{1}{\mu}}$

۱۲۹. فرض کنید $f_X(x)$ چگالی پواسون باشد. کدام یک از روابط زیر صحیح است:

$$f_X(x+1) = \frac{\lambda}{\lambda+1} f_X(x) \quad \text{ب) } \quad f_X(x+1) = \frac{\lambda+1}{\lambda} f_X(x) \quad \text{الف) }$$

$$f_X(x+1) = \frac{\lambda}{x+1} f_X(x) \quad \text{د) } \quad f_X(x+1) = \frac{\lambda}{x} f_X(x) \quad \text{ج) }$$

۱۳۰. متغیر تصادفی X دارای تابع چگالی احتمال نرمال به صورت زیر است. میانگین و واریانس توزیع کدام است؟

$$f_X(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2-2x+1}{2}}, \quad x \in R.$$

$$\sigma^2 = 4, \mu = 2 \quad \text{الف) } \quad \sigma^2 = 4, \mu = 1 \quad \text{ب) }$$

$$\sigma^2 = 1, \mu = 1 \quad \text{د) } \quad \sigma^2 = 2, \mu = 2 \quad \text{ج) }$$

۱۳۱. اگر X دارای توزیع یکنواخت روی بازه $(0, 2)$ باشد امید ریاضی و واریانس X برابر است با

$$\text{الف) } 2, \frac{1}{3} \quad \text{ب) } 1, \frac{1}{3} \quad \text{ج) } 1, \frac{1}{4} \quad \text{د) } 1, \frac{1}{2}$$

۱۳۲. در توزیع پواسون با پارامتر λ نسبت $\frac{E(X)}{\text{var}(X)}$ چقدر است؟

$$\text{الف) } 1 \quad \text{ب) } \lambda \quad \text{ج) } \frac{1}{\lambda} \quad \text{د) } 1+\lambda$$

۱۳۳. اگر میانگین و واریانس متغیر تصادفی X به ترتیب μ و σ^2 باشد انحراف استاندارد متغیر تصادفی $Z = \frac{X-\mu}{\sigma}$ کدام است؟

$$\text{الف) صفر} \quad \text{ب) } \sigma \quad \text{ج) } \sigma^2 \quad \text{د) یک}$$

۱۳۴. برای توزیع $P(X=x) = \binom{r}{x} \left(\frac{1}{r}\right)^x \left(\frac{r-1}{r}\right)^{rx}$; $x = 0, 1, 2, 3, 4$ ، واریانس متغیر تصادفی X کدام است؟

$$\text{الف) } 1 \quad \text{ب) } \frac{3}{4} \quad \text{ج) } \frac{4}{3} \quad \text{د) } 4$$

۱۳۵. احتمال آنکه بازیکن توپی را وارد سبد کند $\frac{3}{10} = p$ است، اگر این بازیکن ۵ بار توپ را به طرف سبد رها کند انتظار می‌رود چند بار موفق شود؟

$$\text{الف) ۵ بار} \quad \text{ب) یک بار} \quad \text{ج) ۸ بار} \quad \text{د) ۲ بار}$$

۱۳۶. فرض کنید متغیر تصادفی X دارای توزیع نرمال با $\mu = 50$ و $\delta = 4$ باشد هرگاه

$P(X > b) = 0/5$ آنگاه مقدار b کدام است؟

$$\begin{array}{ccccccc} \text{الف)} & \frac{50}{4} & \text{ب)} & \frac{50}{2} & \text{ج)} & \frac{50}{16} & \text{د)} \\ \end{array}$$

۱۳۷. میانگین و واریانس نرمال استاندارد به ترتیب برابرند با

$$\begin{array}{ccccccc} \text{الف)} & \text{صفر و یک} & \text{ب)} & \text{یک و صفر} & \text{ج)} & \text{یک و یک} & \text{د)} \end{array}$$

۱۳۸. اگر تعداد تلفن‌های مزاحم به یک مرکز تلفن به‌طور متوسط در هر ساعت 3

تلفن فرض شود احتمال دقیقاً دو تلفن مزاحم در ساعتی معین زده شود کدام است؟

$$\begin{array}{ccccccc} \text{الف)} & 8e^{-3} & \text{ب)} & \frac{-9}{2}e^{-3} & \text{ج)} & 4e^{-3} & \text{د)} \\ \frac{1}{2}e^{-3} & & & & & & \end{array}$$

جدول ۱. احتمال‌های تجمعی دو جمله‌ای

$$P[X \leq c] = \sum_{x=0}^c \binom{n}{x} p^x (1-p)^{n-x}$$

| | | p | | | | | | | | | | |
|-------------|---|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|
| | | 0.۰۵ | 0.۱ | 0.۲ | 0.۳ | 0.۴ | 0.۵ | 0.۶ | 0.۷ | 0.۸ | 0.۹ | 0.۹۵ |
| n=۱ | ۰ | 0/۹۰۰ | 0/۹۰۰ | 0/۸۰۰ | 0/۷۰۰ | 0/۶۰۰ | 0/۵۰۰ | 0/۴۰۰ | 0/۳۰۰ | 0/۲۰۰ | 0/۱۰۰ | 0/۰۰۰ |
| | ۱ | 1/۰۰۰ | 1/۰۰۰ | 1/۰۰۰ | 1/۰۰۰ | 1/۰۰۰ | 1/۰۰۰ | 1/۰۰۰ | 1/۰۰۰ | 1/۰۰۰ | 1/۰۰۰ | 1/۰۰۰ |
| n=۲ | ۰ | 0/۹۰۳ | 0/۸۱۰ | 0/۷۴۰ | 0/۶۹۰ | 0/۶۰ | 0/۵۰ | 0/۴۰ | 0/۳۰ | 0/۲۰ | 0/۱۰ | 0/۰۰۳ |
| | ۱ | 0/۹۹۸ | 0/۹۹۰ | 0/۹۷۰ | 0/۹۱۰ | 0/۸۴۰ | 0/۷۵۰ | 0/۶۴۰ | 0/۵۱۰ | 0/۳۶۰ | 0/۱۹۰ | 0/۰۹۸ |
| n=۳ | ۰ | 0/۸۰۷ | 0/۷۲۹ | 0/۶۱۲ | 0/۴۳۳ | 0/۲۱۶ | 0/۱۲۵ | 0/۰۷۴ | 0/۰۲۷ | 0/۰۰۸ | 0/۰۰۱ | 0/۰۰۰ |
| | ۱ | 0/۹۹۳ | 0/۹۷۲ | 0/۸۹۶ | 0/۷۸۴ | 0/۶۴۸ | 0/۵۰۰ | 0/۳۵۲ | 0/۲۱۶ | 0/۱۰۴ | 0/۰۲۸ | 0/۰۰۷ |
| n=۴ | ۰ | 0/۷۰۰ | 0/۶۰۰ | 0/۴۹۰ | 0/۳۶۰ | 0/۲۵۰ | 0/۱۶۰ | 0/۰۹۰ | 0/۰۴۰ | 0/۰۱۰ | 0/۰۰۰ | 0/۰۰۰ |
| | ۱ | 0/۹۹۰ | 0/۹۷۰ | 0/۹۶۰ | 0/۹۱۰ | 0/۸۴۰ | 0/۷۵۰ | 0/۶۴۰ | 0/۵۱۰ | 0/۳۶۰ | 0/۱۹۰ | 0/۰۹۸ |
| n=۵ | ۰ | 0/۶۰۰ | 0/۵۰۰ | 0/۴۰۰ | 0/۳۰۰ | 0/۲۰۰ | 0/۱۰۰ | 0/۰۶۰ | 0/۰۲۰ | 0/۰۰۸ | 0/۰۰۱ | 0/۰۰۰ |
| | ۱ | 0/۹۸۷ | 0/۹۴۸ | 0/۸۱۹ | 0/۶۰۲ | 0/۴۷۵ | 0/۳۱۳ | 0/۱۷۹ | 0/۰۸۴ | 0/۰۲۷ | 0/۰۰۴ | 0/۰۰۰ |
| n=۶ | ۰ | 0/۵۰۰ | 0/۴۰۰ | 0/۳۰۰ | 0/۲۰۰ | 0/۱۰۰ | 0/۰۶۰ | 0/۰۲۶ | 0/۰۰۸ | 0/۰۰۲ | 0/۰۰۰ | 0/۰۰۰ |
| | ۱ | 0/۹۸۶ | 0/۹۴۶ | 0/۸۷۳ | 0/۷۱۶ | 0/۶۲۱ | 0/۴۸۸ | 0/۳۲۵ | 0/۲۴۸ | 0/۱۸۱ | 0/۰۵۲ | 0/۰۱۴ |
| n=۷ | ۰ | 0/۴۰۰ | 0/۳۰۰ | 0/۲۰۰ | 0/۱۰۰ | 0/۰۶۰ | 0/۰۲۶ | 0/۰۰۸ | 0/۰۰۲ | 0/۰۰۰ | 0/۰۰۰ | 0/۰۰۰ |
| | ۱ | 0/۹۸۵ | 0/۹۴۵ | 0/۸۷۲ | 0/۷۱۵ | 0/۶۲۰ | 0/۴۸۰ | 0/۳۲۰ | 0/۲۴۰ | 0/۱۷۱ | 0/۰۴۳ | 0/۰۰۰ |
| n=۸ | ۰ | 0/۳۰۰ | 0/۲۰۰ | 0/۱۰۰ | 0/۰۶۰ | 0/۰۲۶ | 0/۰۰۸ | 0/۰۰۲ | 0/۰۰۰ | 0/۰۰۰ | 0/۰۰۰ | 0/۰۰۰ |
| | ۱ | 0/۹۸۴ | 0/۹۴۴ | 0/۸۷۱ | 0/۷۱۴ | 0/۶۱۳ | 0/۴۷۹ | 0/۳۱۷ | 0/۲۳۳ | 0/۱۵۸ | 0/۰۴۰ | 0/۰۰۰ |
| n=۹ | ۰ | 0/۲۰۰ | 0/۱۰۰ | 0/۰۶۰ | 0/۰۲۶ | 0/۰۰۸ | 0/۰۰۲ | 0/۰۰۰ | 0/۰۰۰ | 0/۰۰۰ | 0/۰۰۰ | 0/۰۰۰ |
| | ۱ | 0/۹۸۳ | 0/۹۴۳ | 0/۸۷۰ | 0/۷۱۳ | 0/۶۱۲ | 0/۴۷۸ | 0/۳۱۷ | 0/۲۳۳ | 0/۱۵۱ | 0/۰۴۲ | 0/۰۰۰ |
| n=۱۰ | ۰ | 0/۱۰۰ | 0/۰۶۰ | 0/۰۲۶ | 0/۰۰۸ | 0/۰۰۲ | 0/۰۰۰ | 0/۰۰۰ | 0/۰۰۰ | 0/۰۰۰ | 0/۰۰۰ | 0/۰۰۰ |
| | ۱ | 0/۹۸۲ | 0/۹۴۲ | 0/۸۶۹ | 0/۷۱۲ | 0/۶۱۱ | 0/۴۷۷ | 0/۳۱۶ | 0/۲۳۲ | 0/۱۴۶ | 0/۰۴۱ | 0/۰۰۰ |
| n=۱۱ | ۰ | 0/۰۶۰ | 0/۰۲۶ | 0/۰۰۸ | 0/۰۰۲ | 0/۰۰۰ | 0/۰۰۰ | 0/۰۰۰ | 0/۰۰۰ | 0/۰۰۰ | 0/۰۰۰ | 0/۰۰۰ |
| | ۱ | 0/۹۸۱ | 0/۹۴۱ | 0/۸۶۸ | 0/۷۱۱ | 0/۶۱۰ | 0/۴۷۶ | 0/۳۱۵ | 0/۲۳۱ | 0/۱۴۵ | 0/۰۴۰ | 0/۰۰۰ |
| n=۱۲ | ۰ | 0/۰۲۶ | 0/۰۰۸ | 0/۰۰۲ | 0/۰۰۰ | 0/۰۰۰ | 0/۰۰۰ | 0/۰۰۰ | 0/۰۰۰ | 0/۰۰۰ | 0/۰۰۰ | 0/۰۰۰ |
| | ۱ | 0/۹۸۰ | 0/۹۴۰ | 0/۸۶۷ | 0/۷۱۰ | 0/۶۱۰ | 0/۴۷۵ | 0/۳۱۴ | 0/۲۳۰ | 0/۱۴۴ | 0/۰۳۹ | 0/۰۰۰ |
| n=۱۳ | ۰ | 0/۰۰۸ | 0/۰۰۲ | 0/۰۰۰ | 0/۰۰۰ | 0/۰۰۰ | 0/۰۰۰ | 0/۰۰۰ | 0/۰۰۰ | 0/۰۰۰ | 0/۰۰۰ | 0/۰۰۰ |
| | ۱ | 0/۹۷۹ | 0/۹۳۹ | 0/۸۶۶ | 0/۷۰۹ | 0/۶۱۰ | 0/۴۷۴ | 0/۳۱۳ | 0/۲۲۹ | 0/۱۴۳ | 0/۰۳۸ | 0/۰۰۰ |
| n=۱۴ | ۰ | 0/۰۰۷ | 0/۰۰۲ | 0/۰۰۰ | 0/۰۰۰ | 0/۰۰۰ | 0/۰۰۰ | 0/۰۰۰ | 0/۰۰۰ | 0/۰۰۰ | 0/۰۰۰ | 0/۰۰۰ |
| | ۱ | 0/۹۷۸ | 0/۹۳۸ | 0/۸۶۵ | 0/۷۰۸ | 0/۶۱۰ | 0/۴۷۳ | 0/۳۱۲ | 0/۲۲۸ | 0/۱۴۲ | 0/۰۳۷ | 0/۰۰۰ |
| n=۱۵ | ۰ | 0/۰۰۶ | 0/۰۰۲ | 0/۰۰۰ | 0/۰۰۰ | 0/۰۰۰ | 0/۰۰۰ | 0/۰۰۰ | 0/۰۰۰ | 0/۰۰۰ | 0/۰۰۰ | 0/۰۰۰ |
| | ۱ | 0/۹۷۷ | 0/۹۳۷ | 0/۸۶۴ | 0/۷۰۷ | 0/۶۱۰ | 0/۴۷۲ | 0/۳۱۱ | 0/۲۲۷ | 0/۱۴۱ | 0/۰۳۶ | 0/۰۰۰ |
| n=۱۶ | ۰ | 0/۰۰۵ | 0/۰۰۲ | 0/۰۰۰ | 0/۰۰۰ | 0/۰۰۰ | 0/۰۰۰ | 0/۰۰۰ | 0/۰۰۰ | 0/۰۰۰ | 0/۰۰۰ | 0/۰۰۰ |
| | ۱ | 0/۹۷۶ | 0/۹۳۶ | 0/۸۶۳ | 0/۷۰۶ | 0/۶۱۰ | 0/۴۷۱ | 0/۳۱۰ | 0/۲۲۶ | 0/۱۴۰ | 0/۰۳۵ | 0/۰۰۰ |
| n=۱۷ | ۰ | 0/۰۰۴ | 0/۰۰۲ | 0/۰۰۰ | 0/۰۰۰ | 0/۰۰۰ | 0/۰۰۰ | 0/۰۰۰ | 0/۰۰۰ | 0/۰۰۰ | 0/۰۰۰ | 0/۰۰۰ |
| | ۱ | 0/۹۷۵ | 0/۹۳۵ | 0/۸۶۲ | 0/۷۰۵ | 0/۶۱۰ | 0/۴۷۰ | 0/۳۰۹ | 0/۲۲۵ | 0/۱۳۹ | 0/۰۳۴ | 0/۰۰۰ |
| n=۱۸ | ۰ | 0/۰۰۳ | 0/۰۰۲ | 0/۰۰۰ | 0/۰۰۰ | 0/۰۰۰ | 0/۰۰۰ | 0/۰۰۰ | 0/۰۰۰ | 0/۰۰۰ | 0/۰۰۰ | 0/۰۰۰ |
| | ۱ | 0/۹۷۴ | 0/۹۳۴ | 0/۸۶۱ | 0/۷۰۴ | 0/۶۱۰ | 0/۴۶۹ | 0/۳۰۸ | 0/۲۲۴ | 0/۱۳۸ | 0/۰۳۳ | 0/۰۰۰ |
| n=۱۹ | ۰ | 0/۰۰۲ | 0/۰۰۲ | 0/۰۰۰ | 0/۰۰۰ | 0/۰۰۰ | 0/۰۰۰ | 0/۰۰۰ | 0/۰۰۰ | 0/۰۰۰ | 0/۰۰۰ | 0/۰۰۰ |
| | ۱ | 0/۹۷۳ | 0/۹۳۳ | 0/۸۶۰ | 0/۷۰۳ | 0/۶۱۰ | 0/۴۶۸ | 0/۳۰۷ | 0/۲۲۳ | 0/۱۳۷ | 0/۰۳۲ | 0/۰۰۰ |
| n=۲۰ | ۰ | 0/۰۰۱ | 0/۰۰۲ | 0/۰۰۰ | 0/۰۰۰ | 0/۰۰۰ | 0/۰۰۰ | 0/۰۰۰ | 0/۰۰۰ | 0/۰۰۰ | 0/۰۰۰ | 0/۰۰۰ |
| | ۱ | 0/۹۷۲ | 0/۹۳۲ | 0/۸۵۹ | 0/۷۰۲ | 0/۶۱۰ | 0/۴۶۷ | 0/۳۰۶ | 0/۲۲۲ | 0/۱۳۶ | 0/۰۳۱ | 0/۰۰۰ |

ادامه‌ی جدول ۱

جداول ٣٤٩

ادامهی جدول ۱

ادامه‌ی جدول ۱

| | | P | | | | | | | | | | |
|----|------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|-------|
| | | 0/0 | 0/1 | 0/2 | 0/3 | 0/4 | 0/5 | 0/6 | 0/7 | 0/8 | 0/9 | 0/90 |
| c | n | 0/463 | 0/706 | 0/030 | 0/000 | 0/000 | 0/000 | 0/000 | 0/000 | 0/000 | 0/000 | 0/000 |
| 0 | n=10 | 0/829 | 0/049 | 0/167 | 0/030 | 0/005 | 0/000 | 0/000 | 0/000 | 0/000 | 0/000 | 0/000 |
| 1 | | 0/974 | 0/816 | 0/398 | 0/127 | 0/027 | 0/004 | 0/000 | 0/000 | 0/000 | 0/000 | 0/000 |
| 2 | | 0/990 | 0/944 | 0/648 | 0/297 | 0/091 | 0/018 | 0/002 | 0/000 | 0/000 | 0/000 | 0/000 |
| 3 | | 0/999 | 0/987 | 0/836 | 0/510 | 0/217 | 0/059 | 0/009 | 0/001 | 0/000 | 0/000 | 0/000 |
| 4 | | 0/1000 | 0/998 | 0/939 | 0/722 | 0/403 | 0/101 | 0/034 | 0/004 | 0/000 | 0/000 | 0/000 |
| 5 | | 0/1000 | 0/1000 | 0/982 | 0/879 | 0/610 | 0/304 | 0/095 | 0/010 | 0/001 | 0/000 | 0/000 |
| 6 | | 0/1000 | 0/1000 | 0/996 | 0/900 | 0/7887 | 0/500 | 0/213 | 0/050 | 0/004 | 0/000 | 0/000 |
| 7 | | 0/1000 | 0/1000 | 0/999 | 0/980 | 0/905 | 0/796 | 0/390 | 0/121 | 0/018 | 0/000 | 0/000 |
| 8 | | 0/1000 | 0/1000 | 0/999 | 0/999 | 0/980 | 0/905 | 0/796 | 0/390 | 0/121 | 0/018 | 0/000 |
| 9 | | 0/1000 | 0/1000 | 0/1000 | 0/996 | 0/976 | 0/849 | 0/597 | 0/278 | 0/061 | 0/002 | 0/000 |
| 10 | | 0/1000 | 0/1000 | 0/1000 | 0/999 | 0/991 | 0/941 | 0/783 | 0/480 | 0/164 | 0/013 | 0/001 |
| 11 | | 0/1000 | 0/1000 | 0/1000 | 0/1000 | 0/998 | 0/982 | 0/909 | 0/703 | 0/302 | 0/056 | 0/005 |
| 12 | | 0/1000 | 0/1000 | 0/1000 | 0/1000 | 0/996 | 0/973 | 0/873 | 0/602 | 0/184 | 0/036 | |
| 13 | | 0/1000 | 0/1000 | 0/1000 | 0/1000 | 0/1000 | 0/990 | 0/960 | 0/833 | 0/401 | 0/171 | |
| 14 | | 0/1000 | 0/1000 | 0/1000 | 0/1000 | 0/1000 | 0/1000 | 0/990 | 0/960 | 0/794 | 0/537 | |
| 15 | | 0/1000 | 0/1000 | 0/1000 | 0/1000 | 0/1000 | 0/1000 | 0/1000 | 0/1000 | 0/1000 | 0/1000 | |
| 0 | n=16 | 0/440 | 0/180 | 0/028 | 0/003 | 0/000 | 0/000 | 0/000 | 0/000 | 0/000 | 0/000 | 0/000 |
| 1 | | 0/811 | 0/010 | 0/141 | 0/026 | 0/003 | 0/000 | 0/000 | 0/000 | 0/000 | 0/000 | 0/000 |
| 2 | | 0/975 | 0/789 | 0/302 | 0/99 | 0/018 | 0/002 | 0/000 | 0/000 | 0/000 | 0/000 | 0/000 |
| 3 | | 0/993 | 0/932 | 0/098 | 0/246 | 0/060 | 0/011 | 0/001 | 0/000 | 0/000 | 0/000 | 0/000 |
| 4 | | 0/999 | 0/983 | 0/998 | 0/450 | 0/167 | 0/038 | 0/005 | 0/000 | 0/000 | 0/000 | 0/000 |
| 5 | | 0/1000 | 0/997 | 0/918 | 0/660 | 0/329 | 0/105 | 0/019 | 0/002 | 0/000 | 0/000 | 0/000 |
| 6 | | 0/1000 | 0/1000 | 0/973 | 0/825 | 0/027 | 0/227 | 0/058 | 0/007 | 0/000 | 0/000 | 0/000 |
| 7 | | 0/1000 | 0/1000 | 0/993 | 0/926 | 0/716 | 0/402 | 0/142 | 0/026 | 0/001 | 0/000 | 0/000 |
| 8 | | 0/1000 | 0/1000 | 0/999 | 0/974 | 0/508 | 0/058 | 0/284 | 0/074 | 0/007 | 0/000 | 0/000 |
| 9 | | 0/1000 | 0/1000 | 0/1000 | 0/993 | 0/942 | 0/773 | 0/473 | 0/170 | 0/027 | 0/001 | 0/000 |
| 10 | | 0/1000 | 0/1000 | 0/1000 | 0/998 | 0/981 | 0/890 | 0/671 | 0/340 | 0/082 | 0/003 | 0/000 |
| 11 | | 0/1000 | 0/1000 | 0/1000 | 0/1000 | 0/990 | 0/962 | 0/833 | 0/500 | 0/202 | 0/117 | 0/001 |
| 12 | | 0/1000 | 0/1000 | 0/1000 | 0/1000 | 0/999 | 0/989 | 0/930 | 0/704 | 0/402 | 0/268 | 0/007 |
| 13 | | 0/1000 | 0/1000 | 0/1000 | 0/1000 | 0/998 | 0/982 | 0/901 | 0/648 | 0/211 | 0/043 | |
| 14 | | 0/1000 | 0/1000 | 0/1000 | 0/1000 | 0/1000 | 0/1000 | 0/997 | 0/974 | 0/809 | 0/480 | 0/189 |
| 15 | | 0/1000 | 0/1000 | 0/1000 | 0/1000 | 0/1000 | 0/1000 | 0/1000 | 0/997 | 0/972 | 0/810 | 0/560 |
| 16 | | 0/1000 | 0/1000 | 0/1000 | 0/1000 | 0/1000 | 0/1000 | 0/1000 | 0/1000 | 0/1000 | 0/1000 | |
| 0 | n=17 | 0/418 | 0/167 | 0/023 | 0/002 | 0/000 | 0/000 | 0/000 | 0/000 | 0/000 | 0/000 | 0/000 |
| 1 | | 0/792 | 0/482 | 0/118 | 0/19 | 0/02 | 0/000 | 0/000 | 0/000 | 0/000 | 0/000 | 0/000 |
| 2 | | 0/900 | 0/792 | 0/310 | 0/77 | 0/12 | 0/01 | 0/000 | 0/000 | 0/000 | 0/000 | 0/000 |
| 3 | | 0/991 | 0/917 | 0/049 | 0/202 | 0/047 | 0/007 | 0/000 | 0/000 | 0/000 | 0/000 | 0/000 |
| 4 | | 0/999 | 0/978 | 0/708 | 0/389 | 0/126 | 0/025 | 0/003 | 0/000 | 0/000 | 0/000 | 0/000 |
| 5 | | 0/1000 | 0/990 | 0/894 | 0/597 | 0/264 | 0/072 | 0/011 | 0/001 | 0/000 | 0/000 | 0/000 |
| 6 | | 0/1000 | 0/999 | 0/762 | 0/750 | 0/448 | 0/167 | 0/035 | 0/003 | 0/000 | 0/000 | 0/000 |
| 7 | | 0/1000 | 0/1000 | 0/989 | 0/890 | 0/641 | 0/310 | 0/092 | 0/013 | 0/000 | 0/000 | 0/000 |
| 8 | | 0/1000 | 0/1000 | 0/997 | 0/970 | 0/801 | 0/500 | 0/199 | 0/040 | 0/003 | 0/000 | 0/000 |
| 9 | | 0/1000 | 0/1000 | 0/1000 | 0/987 | 0/908 | 0/680 | 0/309 | 0/100 | 0/11 | 0/000 | 0/000 |
| 10 | | 0/1000 | 0/1000 | 0/1000 | 0/997 | 0/970 | 0/834 | 0/502 | 0/220 | 0/038 | 0/001 | 0/000 |

۳۵۱ جداول

ادامه‌ی جدول ۱

| | p | | | | | | | | | | | |
|------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|--|
| c | ./۰۰ | ./۱ | ./۲ | ./۳ | ./۴ | ./۵ | ./۶ | ./۷ | ./۸ | ./۹ | ./۹۵ | |
| ۱۱ | ۱/۰۰۰ | ۱/۰۰۰ | ۱/۰۰۰ | ۰/۹۹۹ | ۰/۹۸۹ | ۰/۹۲۸ | ۰/۷۳۶ | ۰/۴۰۳ | ۰/۱۰۶ | ۰/۰۰۰ | ۰/۰۰۰ | |
| ۱۲ | ۱/۰۰۰ | ۱/۰۰۰ | ۱/۰۰۰ | ۱/۰۰۰ | ۰/۹۹۷ | ۰/۹۷۰ | ۰/۸۷۴ | ۰/۶۱۱ | ۰/۲۴۲ | ۰/۰۲۲ | ۰/۰۰۱ | |
| ۱۳ | ۱/۰۰۰ | ۱/۰۰۰ | ۱/۰۰۰ | ۱/۰۰۰ | ۱/۰۰۰ | ۰/۹۹۴ | ۰/۹۰۴ | ۰/۷۹۸ | ۰/۴۰۱ | ۰/۰۸۳ | ۰/۰۰۹ | |
| ۱۴ | ۱/۰۰۰ | ۱/۰۰۰ | ۱/۰۰۰ | ۱/۰۰۰ | ۱/۰۰۰ | ۰/۹۹۹ | ۰/۹۸۸ | ۰/۹۲۳ | ۰/۷۹۰ | ۰/۲۳۸ | ۰/۰۰۵ | |
| ۱۵ | ۱/۰۰۰ | ۱/۰۰۰ | ۱/۰۰۰ | ۱/۰۰۰ | ۱/۰۰۰ | ۱/۰۰۰ | ۰/۹۹۸ | ۰/۹۸۱ | ۰/۸۸۲ | ۰/۰۱۸ | ۰/۲۰۸ | |
| ۱۶ | ۱/۰۰۰ | ۱/۰۰۰ | ۱/۰۰۰ | ۱/۰۰۰ | ۱/۰۰۰ | ۱/۰۰۰ | ۰/۹۹۸ | ۰/۹۷۷ | ۰/۸۳۳ | ۰/۰۸۲ | | |
| ۱۷ | ۱/۰۰۰ | ۱/۰۰۰ | ۱/۰۰۰ | ۱/۰۰۰ | ۱/۰۰۰ | ۱/۰۰۰ | ۱/۰۰۰ | ۱/۰۰۰ | ۱/۰۰۰ | ۱/۰۰۰ | ۱/۰۰۰ | |
| n=۱۸ | . | ۰/۳۹۷ | ۰/۱۰۰ | ۰/۰۱۸ | ۰/۰۰۲ | ۰/۰۰۰ | ۰/۰۰۰ | ۰/۰۰۰ | ۰/۰۰۰ | ۰/۰۰۰ | ۰/۰۰۰ | |
| | ۱ | ۰/۷۷۴ | ۰/۴۰۰ | ۰/۰۹۹ | ۰/۰۱۴ | ۰/۰۰۱ | ۰/۰۰۰ | ۰/۰۰۰ | ۰/۰۰۰ | ۰/۰۰۰ | ۰/۰۰۰ | |
| | ۲ | ۰/۹۴۲ | ۰/۷۳۴ | ۰/۲۷۱ | ۰/۰۷۰ | ۰/۰۰۸ | ۰/۰۰۱ | ۰/۰۰۰ | ۰/۰۰۰ | ۰/۰۰۰ | ۰/۰۰۰ | |
| | ۳ | ۰/۹۸۹ | ۰/۹۰۲ | ۰/۰۰۱ | ۰/۱۶۰ | ۰/۰۳۳ | ۰/۰۰۴ | ۰/۰۰۰ | ۰/۰۰۰ | ۰/۰۰۰ | ۰/۰۰۰ | |
| | ۴ | ۰/۹۹۸ | ۰/۹۷۲ | ۰/۷۱۶ | ۰/۳۳۳ | ۰/۰۹۴ | ۰/۰۱۰ | ۰/۰۰۱ | ۰/۰۰۰ | ۰/۰۰۰ | ۰/۰۰۰ | |
| | ۵ | ۱/۰۰۰ | ۰/۹۹۴ | ۰/۸۷۷ | ۰/۰۳۴ | ۰/۲۰۹ | ۰/۰۴۸ | ۰/۰۰۶ | ۰/۰۰۰ | ۰/۰۰۰ | ۰/۰۰۰ | |
| | ۶ | ۱/۰۰۰ | ۰/۹۹۹ | ۰/۹۴۹ | ۰/۷۲۲ | ۰/۳۷۴ | ۰/۱۱۹ | ۰/۰۲۰ | ۰/۰۰۱ | ۰/۰۰۰ | ۰/۰۰۰ | |
| | ۷ | ۱/۰۰۰ | ۱/۰۰۰ | ۰/۹۸۴ | ۰/۸۰۹ | ۰/۰۶۳ | ۰/۲۴۰ | ۰/۰۵۸ | ۰/۰۰۷ | ۰/۰۰۰ | ۰/۰۰۰ | |
| | ۸ | ۱/۰۰۰ | ۱/۰۰۰ | ۰/۹۹۷ | ۰/۹۴۰ | ۰/۰۳۷ | ۰/۰۴۷ | ۰/۱۳۰ | ۰/۰۲۱ | ۰/۰۰۱ | ۰/۰۰۰ | |
| | ۹ | ۱/۰۰۰ | ۱/۰۰۰ | ۰/۹۹۹ | ۰/۹۷۹ | ۰/۸۶۰ | ۰/۰۹۳ | ۰/۲۶۳ | ۰/۰۶۰ | ۰/۰۰۴ | ۰/۰۰۰ | |
| | ۱۰ | ۱/۰۰۰ | ۱/۰۰۰ | ۱/۰۰۰ | ۰/۹۹۴ | ۰/۹۴۲ | ۰/۷۶۰ | ۰/۴۳۷ | ۰/۱۴۱ | ۰/۰۱۶ | ۰/۰۰۰ | |
| | ۱۱ | ۱/۰۰۰ | ۱/۰۰۰ | ۱/۰۰۰ | ۰/۹۹۹ | ۰/۹۸۰ | ۰/۸۸۱ | ۰/۲۶۶ | ۰/۲۷۸ | ۰/۰۵۱ | ۰/۰۰۱ | |
| | ۱۲ | ۱/۰۰۰ | ۱/۰۰۰ | ۱/۰۰۰ | ۱/۰۰۰ | ۰/۹۹۶ | ۰/۹۰۲ | ۰/۷۹۱ | ۰/۴۶۶ | ۰/۱۳۳ | ۰/۰۰۷ | |
| | ۱۳ | ۱/۰۰۰ | ۱/۰۰۰ | ۱/۰۰۰ | ۱/۰۰۰ | ۰/۹۹۹ | ۰/۹۸۰ | ۰/۰۹۶ | ۰/۶۶۷ | ۰/۲۸۴ | ۰/۰۲۸ | |
| | ۱۴ | ۱/۰۰۰ | ۱/۰۰۰ | ۱/۰۰۰ | ۱/۰۰۰ | ۱/۰۰۰ | ۰/۹۹۷ | ۰/۹۶۷ | ۰/۸۳۰ | ۰/۴۹۹ | ۰/۰۹۸ | |
| | ۱۵ | ۱/۰۰۰ | ۱/۰۰۰ | ۱/۰۰۰ | ۱/۰۰۰ | ۱/۰۰۰ | ۰/۹۹۹ | ۰/۹۹۲ | ۰/۹۴۰ | ۰/۷۲۹ | ۰/۲۶۶ | |
| | ۱۶ | ۱/۰۰۰ | ۱/۰۰۰ | ۱/۰۰۰ | ۱/۰۰۰ | ۱/۰۰۰ | ۰/۹۹۹ | ۰/۹۸۶ | ۰/۹۰۱ | ۰/۰۵۰ | ۰/۲۲۶ | |
| | ۱۷ | ۱/۰۰۰ | ۱/۰۰۰ | ۱/۰۰۰ | ۱/۰۰۰ | ۱/۰۰۰ | ۰/۹۹۸ | ۰/۹۸۲ | ۰/۸۰۰ | ۰/۰۸۰ | ۰/۶۰۳ | |
| | ۱۸ | ۱/۰۰۰ | ۱/۰۰۰ | ۱/۰۰۰ | ۱/۰۰۰ | ۱/۰۰۰ | ۰/۹۹۹ | ۰/۹۰۰ | ۰/۰۰۰ | ۰/۰۰۰ | ۰/۰۰۰ | |
| n=۱۹ | . | ۰/۳۷۷ | ۰/۱۳۰ | ۰/۰۱۴ | ۰/۰۰۱ | ۰/۰۰۰ | ۰/۰۰۰ | ۰/۰۰۰ | ۰/۰۰۰ | ۰/۰۰۰ | ۰/۰۰۰ | |
| | ۱ | ۰/۷۵۰ | ۰/۴۲۰ | ۰/۰۸۳ | ۰/۰۱۰ | ۰/۰۰۱ | ۰/۰۰۰ | ۰/۰۰۰ | ۰/۰۰۰ | ۰/۰۰۰ | ۰/۰۰۰ | |
| | ۲ | ۰/۹۳۳ | ۰/۷۰۰ | ۰/۲۳۷ | ۰/۰۴۶ | ۰/۰۰۵ | ۰/۰۰۰ | ۰/۰۰۰ | ۰/۰۰۰ | ۰/۰۰۰ | ۰/۰۰۰ | |
| | ۳ | ۰/۹۸۷ | ۰/۸۸۰ | ۰/۴۰۵ | ۰/۱۳۳ | ۰/۰۲۳ | ۰/۰۰۲ | ۰/۰۰۰ | ۰/۰۰۰ | ۰/۰۰۰ | ۰/۰۰۰ | |
| | ۴ | ۰/۹۹۸ | ۰/۹۶۰ | ۰/۶۷۳ | ۰/۲۸۲ | ۰/۰۷۰ | ۰/۰۱۰ | ۰/۰۰۱ | ۰/۰۰۰ | ۰/۰۰۰ | ۰/۰۰۰ | |
| | ۵ | ۱/۰۰۰ | ۰/۹۹۱ | ۰/۰۳۷ | ۰/۰۷۴ | ۰/۱۳۳ | ۰/۰۳۲ | ۰/۰۰۳ | ۰/۰۰۰ | ۰/۰۰۰ | ۰/۰۰۰ | |
| | ۶ | ۱/۰۰۰ | ۰/۹۹۸ | ۰/۹۳۲ | ۰/۶۶۶ | ۰/۳۰۸ | ۰/۰۸۴ | ۰/۰۱۲ | ۰/۰۰۱ | ۰/۰۰۰ | ۰/۰۰۰ | |
| | ۷ | ۱/۰۰۰ | ۰/۹۷۷ | ۰/۸۱۸ | ۰/۴۸۸ | ۰/۱۸۰ | ۰/۰۳۰ | ۰/۰۰۳ | ۰/۰۰۰ | ۰/۰۰۰ | ۰/۰۰۰ | |
| | ۸ | ۱/۰۰۰ | ۰/۹۹۳ | ۰/۹۱۶ | ۰/۶۶۷ | ۰/۳۲۴ | ۰/۰۸۸ | ۰/۰۱۱ | ۰/۰۰۰ | ۰/۰۰۰ | ۰/۰۰۰ | |
| | ۹ | ۱/۰۰۰ | ۰/۹۹۸ | ۰/۹۶۷ | ۰/۱۱۴ | ۰/۰۰۰ | ۰/۱۸۶ | ۰/۰۳۳ | ۰/۰۰۲ | ۰/۰۰۰ | ۰/۰۰۰ | |
| | ۱۰ | ۱/۰۰۰ | ۱/۰۰۰ | ۱/۰۰۰ | ۰/۹۸۹ | ۰/۹۱۲ | ۰/۶۷۶ | ۰/۳۳۳ | ۰/۰۸۴ | ۰/۰۰۷ | ۰/۰۰۰ | |
| | ۱۱ | ۱/۰۰۰ | ۱/۰۰۰ | ۱/۰۰۰ | ۰/۹۹۷ | ۰/۹۶۰ | ۰/۸۲۰ | ۰/۰۱۲ | ۰/۱۸۲ | ۰/۰۲۳ | ۰/۰۰۰ | |
| | ۱۲ | ۱/۰۰۰ | ۱/۰۰۰ | ۱/۰۰۰ | ۰/۹۹۹ | ۰/۹۸۸ | ۰/۹۱۶ | ۰/۶۹۲ | ۰/۳۳۴ | ۰/۰۶۸ | ۰/۰۰۲ | |
| | ۱۳ | ۱/۰۰۰ | ۱/۰۰۰ | ۱/۰۰۰ | ۱/۰۰۰ | ۰/۹۹۷ | ۰/۹۶۸ | ۰/۸۳۷ | ۰/۰۲۶ | ۰/۱۶۳ | ۰/۰۰۹ | |
| | ۱۴ | ۱/۰۰۰ | ۱/۰۰۰ | ۱/۰۰۰ | ۱/۰۰۰ | ۰/۹۹۹ | ۰/۹۹۰ | ۰/۹۳۰ | ۰/۷۱۸ | ۰/۳۲۷ | ۰/۰۰۲ | |
| | ۱۵ | ۱/۰۰۰ | ۱/۰۰۰ | ۱/۰۰۰ | ۱/۰۰۰ | ۱/۰۰۰ | ۰/۹۹۸ | ۰/۹۷۷ | ۰/۸۶۷ | ۰/۰۴۵ | ۰/۱۱۰ | |
| | ۱۶ | ۱/۰۰۰ | ۱/۰۰۰ | ۱/۰۰۰ | ۱/۰۰۰ | ۱/۰۰۰ | ۰/۹۹۰ | ۰/۹۰۴ | ۰/۷۶۳ | ۰/۲۹۰ | ۰/۰۶۷ | |
| | ۱۷ | ۱/۰۰۰ | ۱/۰۰۰ | ۱/۰۰۰ | ۱/۰۰۰ | ۱/۰۰۰ | ۰/۹۹۹ | ۰/۹۹۰ | ۰/۹۱۷ | ۰/۰۵۰ | ۰/۲۴۵ | |
| | ۱۸ | ۱/۰۰۰ | ۱/۰۰۰ | ۱/۰۰۰ | ۱/۰۰۰ | ۱/۰۰۰ | ۰/۹۹۹ | ۰/۹۸۶ | ۰/۸۶۰ | ۰/۶۲۳ | | |

ادامه‌ی جدول ۱

٣٥٣ جداول

ادامهی جدول ۱

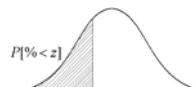
| | c | ٠/٠٥ | ٠/١ | ٠/٢ | ٠/٣ | ٠/٤ | p ٠/٥ | ٠/٦ | ٠/٧ | ٠/٨ | ٠/٩ | ٠/٩٥ |
|-------|---|-------|-------|-------|-------|-------|----------|-------|-------|-------|-------|-------|
| n=٢٢ | ٠ | ٠/٣٢٤ | ٠/٠٩٨ | ٠/٠٠٧ | ٠/٠٠٠ | ٠/٠٠٠ | ٠/٠٠٠ | ٠/٠٠٠ | ٠/٠٠٠ | ٠/٠٠٠ | ٠/٠٠٠ | ٠/٠٠٠ |
| ١ | | ٠/٦٩٨ | ٠/٣٣٩ | ٠/٠٤٨ | ٠/٠٠٤ | ٠/٠٠٠ | ٠/٠٠٠ | ٠/٠٠٠ | ٠/٠٠٠ | ٠/٠٠٠ | ٠/٠٠٠ | ٠/٠٠٠ |
| ٢ | | ٠/٩٠٥ | ٠/٦٢٠ | ٠/١٥٤ | ٠/٠٢١ | ٠/٠٠٢ | ٠/٠٠٠ | ٠/٠٠٠ | ٠/٠٠٠ | ٠/٠٠٠ | ٠/٠٠٠ | ٠/٠٠٠ |
| ٣ | | ٠/٩٧٨ | ٠/٨٢٨ | ٠/٣٣٢ | ٠/٠٦٨ | ٠/٠٠٨ | ٠/٠٠٠ | ٠/٠٠٠ | ٠/٠٠٠ | ٠/٠٠٠ | ٠/٠٠٠ | ٠/٠٠٠ |
| ٤ | | ٠/٩٩٦ | ٠/٩٣٨ | ٠/٥٤٣ | ٠/١٦٥ | ٠/٠٢٧ | ٠/٠٠٢ | ٠/٠٠٠ | ٠/٠٠٠ | ٠/٠٠٠ | ٠/٠٠٠ | ٠/٠٠٠ |
| ٥ | | ٠/٩٩٩ | ٠/٩٨٢ | ٠/٧٣٣ | ٠/٣١٣ | ٠/٠٧٢ | ٠/٠٠٨ | ٠/٠٠٠ | ٠/٠٠٠ | ٠/٠٠٠ | ٠/٠٠٠ | ٠/٠٠٠ |
| ٦ | | ١/٠٠٠ | ٠/٩٩٦ | ٠/٨٧٧ | ٠/٤٩٤ | ٠/١٥٨ | ٠/٠٢٦ | ٠/٠٠٢ | ٠/٠٠٠ | ٠/٠٠٠ | ٠/٠٠٠ | ٠/٠٠٠ |
| ٧ | | ١/٠٠٠ | ٠/٩٩٩ | ٠/٩٤٤ | ٠/٦٧١ | ٠/٢٩٠ | ٠/٠٧ | ٠/٠٠٧ | ٠/٠٠٠ | ٠/٠٠٠ | ٠/٠٠٠ | ٠/٠٠٠ |
| ٨ | | ١/٠٠٠ | ١/٠٠٠ | ٠/٩٨٠ | ٠/٨١٤ | ٠/٤٥٤ | ٠/١٤٣ | ٠/٠٢١ | ٠/٠٠١ | ٠/٠٠٠ | ٠/٠٠٠ | ٠/٠٠٠ |
| ٩ | | ١/٠٠٠ | ١/٠٠٠ | ٠/٩٩٤ | ٠/٩٠٨ | ٠/٦٢٤ | ٠/٢٦٢ | ٠/٠٥٥ | ٠/٠٠٤ | ٠/٠٠٠ | ٠/٠٠٠ | ٠/٠٠٠ |
| ١٠ | | ١/٠٠٠ | ١/٠٠٠ | ٠/٩٩٨ | ٠/٩٦١ | ٠/٧٧٢ | ٠/٤١٦ | ٠/٠٢١ | ٠/٠١٤ | ٠/٠٠٠ | ٠/٠٠٠ | ٠/٠٠٠ |
| ١١ | | ١/٠٠٠ | ١/٠٠٠ | ١/٠٠٠ | ٠/٩٨٦ | ٠/٨٧٩ | ٠/٥٨٤ | ٠/٢٢٨ | ٠/٠٣٩ | ٠/٠٠٢ | ٠/٠٠٠ | ٠/٠٠٠ |
| ١٢ | | ١/٠٠٠ | ١/٠٠٠ | ١/٠٠٠ | ٠/٩٩٦ | ٠/٩٤٥ | ٠/٧٣٨ | ٠/٣٧ | ٠/٠٩٢ | ٠/٠٠٦ | ٠/٠٠٠ | ٠/٠٠٠ |
| ١٣ | | ١/٠٠٠ | ١/٠٠٠ | ١/٠٠٠ | ٠/٩٩٩ | ٠/٩٧٩ | ٠/٨٥٧ | ٠/٥٤٦ | ٠/١٨٦ | ٠/٠٢٠ | ٠/٠٠٠ | ٠/٠٠٠ |
| ١٤ | | ١/٠٠٠ | ١/٠٠٠ | ١/٠٠٠ | ١/٠٠٠ | ٠/٩٩٣ | ٠/٩٣٣ | ٠/٧١٠ | ٠/٣٢٩ | ٠/٠٥٦ | ٠/٠٠١ | ٠/٠٠٠ |
| ١٥ | | ١/٠٠٠ | ١/٠٠٠ | ١/٠٠٠ | ١/٠٠٠ | ٠/٩٩٨ | ٠/٩٧٤ | ٠/٨٤٢ | ٠/٥٠٦ | ٠/١٣٣ | ٠/٠٠٤ | ٠/٠٠٠ |
| ١٦ | | ١/٠٠٠ | ١/٠٠٠ | ١/٠٠٠ | ١/٠٠٠ | ٠/٩٩٢ | ٠/٩٢٨ | ٠/٦٨٧ | ٠/٢٦٧ | ٠/٠١٨ | ٠/٠٠١ | |
| ١٧ | | ١/٠٠٠ | ١/٠٠٠ | ١/٠٠٠ | ١/٠٠٠ | ٠/٩٩٨ | ٠/٩٧٣ | ٠/٨٣٥ | ٠/٤٥٧ | ٠/٠٦٢ | ٠/٠٠٤ | |
| ١٨ | | ١/٠٠٠ | ١/٠٠٠ | ١/٠٠٠ | ١/٠٠٠ | ٠/٩٩٢ | ٠/٩٣٢ | ٠/٦٦٨ | ٠/١٧٢ | ٠/٠٢٢ | | |
| ١٩ | | ١/٠٠٠ | ١/٠٠٠ | ١/٠٠٠ | ١/٠٠٠ | ٠/٩٧٩ | ٠/٨٤٦ | ٠/٧٨٠ | ٠/٠٩٥ | | | |
| ٢٠ | | ١/٠٠٠ | ١/٠٠٠ | ١/٠٠٠ | ١/٠٠٠ | ٠/٩٦١ | ٠/٩٥٢ | ٠/٧٦١ | ٠/٣٢ | | | |
| ٢١ | | ١/٠٠٠ | ١/٠٠٠ | ١/٠٠٠ | ١/٠٠٠ | ٠/٩٩٣ | ٠/٩٠٢ | ٠/٧٧٦ | | | | |
| ٢٢ | | ١/٠٠٠ | ١/٠٠٠ | ١/٠٠٠ | ١/٠٠٠ | ٠/٩٩٣ | ٠/٩٠٠ | ٠/٧٠٠ | ٠/٠٠٠ | ٠/٠٠٠ | | |
| <hr/> | | | | | | | | | | | | |
| n=٢٣ | ٠ | ٠/٣٠٧ | ٠/٠٨٩ | ٠/٠٠٦ | ٠/٠٠٠ | ٠/٠٠٠ | ٠/٠٠٠ | ٠/٠٠٠ | ٠/٠٠٠ | ٠/٠٠٠ | ٠/٠٠٠ | ٠/٠٠٠ |
| ١ | | ٠/٦٧٩ | ٠/٣١٥ | ٠/٠٤٠ | ٠/٠٠٣ | ٠/٠٠٠ | ٠/٠٠٠ | ٠/٠٠٠ | ٠/٠٠٠ | ٠/٠٠٠ | ٠/٠٠٠ | ٠/٠٠٠ |
| ٢ | | ٠/٨٩٥ | ٠/٥٩٢ | ٠/١٣٣ | ٠/٠١٦ | ٠/٠٠١ | ٠/٠٠٠ | ٠/٠٠٠ | ٠/٠٠٠ | ٠/٠٠٠ | ٠/٠٠٠ | ٠/٠٠٠ |
| ٣ | | ٠/٩٧٤ | ٠/٨٠٧ | ٠/٢٩٧ | ٠/٠٥٤ | ٠/٠٠٥ | ٠/٠٠٠ | ٠/٠٠٠ | ٠/٠٠٠ | ٠/٠٠٠ | ٠/٠٠٠ | ٠/٠٠٠ |
| ٤ | | ٠/٩٩٠ | ٠/٩٢٧ | ٠/٥٠١ | ٠/١٣٦ | ٠/٠١٩ | ٠/٠٠١ | ٠/٠٠٠ | ٠/٠٠٠ | ٠/٠٠٠ | ٠/٠٠٠ | ٠/٠٠٠ |
| ٥ | | ٠/٩٩٩ | ٠/٩٧٧ | ٠/٦٩٥ | ٠/٢٦٩ | ٠/٠٥٤ | ٠/٠٠٥ | ٠/٠٠٠ | ٠/٠٠٠ | ٠/٠٠٠ | ٠/٠٠٠ | ٠/٠٠٠ |
| ٦ | | ١/٠٠٠ | ٠/٩٩٤ | ٠/٨٤٠ | ٠/٤٤٠ | ٠/١٢٤ | ٠/٠١٧ | ٠/٠٠١ | ٠/٠٠٠ | ٠/٠٠٠ | ٠/٠٠٠ | ٠/٠٠٠ |
| ٧ | | ١/٠٠٠ | ٠/٩٩٩ | ٠/٩٢٨ | ٠/٧١٨ | ٠/٢٣٧ | ٠/٠٤٧ | ٠/٠٤ | ٠/٠٠٠ | ٠/٠٠٠ | ٠/٠٠٠ | ٠/٠٠٠ |
| ٨ | | ١/٠٠٠ | ١/٠٠٠ | ٠/٩٧٣ | ٠/٧٧١ | ٠/٢٨٨ | ٠/١٠ | ٠/٠١٣ | ٠/٠٠١ | ٠/٠٠٠ | ٠/٠٠٠ | ٠/٠٠٠ |
| ٩ | | ١/٠٠٠ | ١/٠٠٠ | ٠/٩٩١ | ٠/٨٨٠ | ٠/٥٥٦ | ٠/٢٠٢ | ٠/٠٣٥ | ٠/٠٠٢ | ٠/٠٠٠ | ٠/٠٠٠ | ٠/٠٠٠ |
| ١٠ | | ١/٠٠٠ | ١/٠٠٠ | ٠/٩٩٧ | ٠/٩٤٥ | ٠/٧١٣ | ٠/٣٣٩ | ٠/٠٨١ | ٠/٠٠٧ | ٠/٠٠٠ | ٠/٠٠٠ | ٠/٠٠٠ |
| ١١ | | ١/٠٠٠ | ١/٠٠٠ | ٠/٩٩٩ | ٠/٩٧٩ | ٠/٨٣٦ | ٠/٠٠٠ | ٠/٦٤ | ٠/٠٢١ | ٠/٠٠١ | ٠/٠٠٠ | ٠/٠٠٠ |
| ١٢ | | ١/٠٠٠ | ١/٠٠٠ | ١/٠٠٠ | ٠/٩٩٣ | ٠/٩١٩ | ٠/٦٦١ | ٠/٧٨٧ | ٠/٠٥٠ | ٠/٠٣ | ٠/٠٠٠ | ٠/٠٠٠ |
| ١٣ | | ١/٠٠٠ | ١/٠٠٠ | ١/٠٠٠ | ٠/٩٩٨ | ٠/٩٧٥ | ٠/٧٤٤ | ٠/١٢٠ | ٠/٠٠٩ | ٠/٠٠٠ | ٠/٠٠٠ | ٠/٠٠٠ |
| ١٤ | | ١/٠٠٠ | ١/٠٠٠ | ١/٠٠٠ | ٠/٩٩٩ | ٠/٩٨٧ | ٠/٨٩٥ | ٠/٦١٢ | ٠/٢٢٩ | ٠/٠٢٧ | ٠/٠٠٠ | ٠/٠٠٠ |
| ١٥ | | ١/٠٠٠ | ١/٠٠٠ | ١/٠٠٠ | ١/٠٠٠ | ٠/٩٩٦ | ٠/٩٥٣ | ٠/٧٦٣ | ٠/٣٨٢ | ٠/٠٧٢ | ٠/٠٠١ | ٠/٠٠٠ |
| ١٦ | | ١/٠٠٠ | ١/٠٠٠ | ١/٠٠٠ | ١/٠٠٠ | ٠/٩٩٩ | ٠/٨٨٣ | ٠/٧٨٧ | ٠/٥٦٠ | ٠/١٦٠ | ٠/٠٠٦ | ٠/٠٠٠ |
| ١٧ | | ١/٠٠٠ | ١/٠٠٠ | ١/٠٠٠ | ١/٠٠٠ | ٠/٩٩٥ | ٠/٩٤٦ | ٠/٧٣١ | ٠/٣٥٠ | ٠/٢٠٥ | ٠/٠٢٣ | ٠/٠٠١ |
| ١٨ | | ١/٠٠٠ | ١/٠٠٠ | ١/٠٠٠ | ١/٠٠٠ | ٠/٩٩٩ | ٠/٩٩٩ | ٠/٨٦٤ | ٠/٤٩٩ | ٠/٠٧٣ | ٠/٠٠٥ | |
| ١٩ | | ١/٠٠٠ | ١/٠٠٠ | ١/٠٠٠ | ١/٠٠٠ | ٠/٩٩٥ | ٠/٩٤٦ | ٠/٧٣١ | ٠/٣٥٠ | ٠/٢٠٥ | ٠/٠٢٣ | |
| ٢٠ | | ١/٠٠٠ | ١/٠٠٠ | ١/٠٠٠ | ١/٠٠٠ | ٠/٩٩٩ | ٠/٩٩٩ | ٠/٨٦٤ | ٠/٤٩٩ | ٠/٠٧٣ | ٠/٠٢٦ | |
| ٢١ | | ١/٠٠٠ | ١/٠٠٠ | ١/٠٠٠ | ١/٠٠٠ | ٠/٩٩٧ | ٠/٩٩٧ | ٠/٨٦٧ | ٠/٤٠٨ | ٠/١٩٣ | ٠/٠٢٦ | |
| ٢٢ | | ١/٠٠٠ | ١/٠٠٠ | ١/٠٠٠ | ١/٠٠٠ | ٠/٩٩٤ | ٠/٩٩٤ | ٠/٨٦٥ | ٠/٤٠٨ | ٠/١٩٣ | ٠/٠٢٦ | |
| ٢٣ | | ١/٠٠٠ | ١/٠٠٠ | ١/٠٠٠ | ١/٠٠٠ | ٠/٩٩٣ | ٠/٩٩٣ | ٠/٨٦٤ | ٠/٤٠٨ | ٠/١٩٣ | ٠/٠٢٦ | |

جدول ۲. احتمال‌های پواسون

٣٥٥ جداول

ادامه‌ی جدول ۲

جدول ۳ احتمال‌های توزیع نرم‌افزار استاندارد



ادامه‌ی جدول ۳

خواننده محترم

این پرسشنامه به منظور ارتقای کیفیت کتاب‌های درسی و رفع نواقص آنها تهیه شده است. دقت شما در پاسخگویی به این پرسشنامه در پایان هر نیمسال ما را در تحقق این هدف یاری خواهد کرد.

نام کتاب نام مؤلف/مترجم سال انتشار
 پاسخگو: عضو علمی پایمانور عضو علمی سایر دانشگاهها رشته تخصصی سابقه تدریس
 دانشجویی پایمانور دانشجوی سایر دانشگاهها رشته تحصیلی ورودی سال

| سؤال | نهایی | بازگشته | بازگشته | بازگشته | بازگشته |
|---|-------|---------|---------|---------|---------|
| ۱. آیا از زمان تحویل و نحوه دسترسی به کتاب راضی بودید؟ | | | | | |
| ۲. آیا حجم کتاب با توجه به تعداد واحد مناسب بود؟ | | | | | |
| ۳. آیا راهنمایی‌های لازم برای مطالعه کتاب منتظر شده بود؟ | | | | | |
| ۴. آیا در ترتیب مطالب کتاب سلسه مراتب شناختی (آسان به مشکل) رعایت شده بود؟ | | | | | |
| ۵. آیا تقسیم‌بندی مطالب در فصل‌ها یا بخش‌ها مناسب و بجا بود؟ | | | | | |
| ۶. آیا متن کتاب روان و ساده و جمله‌ها قابل فهم بود؟ | | | | | |
| ۷. آیا بهروزبودن مطالب و آمارها رعایت شده بود؟ | | | | | |
| ۸. آیا مطالب تکراری داشت؟ | | | | | |
| ۹. آیا پیوستگی مطالب با درس‌های پیش‌نیاز رعایت شده بود؟ | | | | | |
| ۱۰. آیا مثال‌ها، شکل‌ها، نمودارها، جدول‌ها و ... گویا بودند و در فهم مطلب تأثیر داشتند؟ | | | | | |
| ۱۱. مطالعه هدف‌های کلی، آموزشی / رفتاری تا چه اندازه به درک بهتر شما کمک کرد؟ | | | | | |
| ۱۲. آیا خودآزمایی‌های کتاب به گونه‌ای بود که تمام مطالب درسی را شامل شود؟ | | | | | |
| ۱۳. آیا پاسخ خودآزمایی‌ها و تمرین‌ها کامل و گویا بود؟ | | | | | |
| ۱۴. چقدر با غلط‌های املایی و اشکال‌های چاپی مواجه شدید؟ | | | | | |
| ۱۵. آیا از کیفیت چاپ و صحافی کتاب راضی بودید؟ | | | | | |
| ۱۶. آیا طرح روی جلد کتاب با مطالب کتاب تناسب داشت؟ | | | | | |
| ۱۷. چنانچه دانشگاه وسائل کمک‌آموزشی از قبیل نوار، فیلم، لوح فشرده و ... در اختیارتان گذارد، آیا به درک بهتر شما کمک کرده‌اند؟ | | | | | |
| ۱۸. تا چه اندازه این کتاب شما را از حضور در کلاس بینیاز کرد؟ | | | | | |

در مجموع کتاب را چگونه ارزیابی می‌کنید؟ عالی خوب متوسط ضعیف بسیار ضعیف
 لطفاً چنانچه با اشکال‌های تابیی یا محتوایی و مطالب تکراری مواجه شده‌اید، فهرستی از آن‌ها را با ذکر شماره صفحه ضمیمه کنید. در صورت تمایل سایر پیشنهادهای خود را نیز بنویسید.

این پرسشنامه را پس از تکمیل از کتاب جدا کنید و به قسمت آموزش مرکز تحویل دهید یا مستقیماً به نشانی تهران، صندوق پستی ۱۴۳۳۵-۳۳۳، مدیریت تولید مواد و تجهیزات آموزشی کتاب ارسال فرمایید. آدرس وبگاه ما www.pnu.ac.ir است. با ورود به وبگاه، مسیر زیر را طی نمایید: ساختار دانشگاه / معاونت‌ها / فناوری اطلاعات / مدیریت تولید مواد و تجهیزات آموزشی.

با تشکر
 مدیریت تولید مواد و تجهیزات آموزشی