www.engclubs.net

A site for All Engineers

آمار و احتمال مهندسی

دکتر کاوه کریمی

جلسه اول:

نقش آمار (Statistics) در مهندسی

مقدمه:

آمار علمی است که به ما کمک می کند که در حضور طیف متغیرهای مختلف تصمیم گیری و نتیجه گیری کنیم.

- **برای مثال: ظرفیت سیستم بزرگراهی در مهندسی عمران**
  - **تعداد سفرهای غیرکاری و کاری از مبدا خانه**
  - **تعداد افراد هر خانوار**
  - **تعداد ماشین هر خانوار**
- **هدف ساخت یک مدل که تعداد سفرها را به پارامترهای فوق ربط دهد**
- **روش رگرسیون** (Regression)

مقدمه:

آمار علمی است که به ما کمک می کند که در حضور طیف متغیر های مختلف تصمیم گیری و نتیجه گیری کنیم.

مثال: بخش اورژانس در بیمارستانها

- **فرآیند رسیدن بیماران به اورژانس بسیار متغیر است**
  - **ساعت روز**
  - **روز**
- **فرآیند سرویس دهی به بیماران بسیار متغیر است**
  - **نوع سرویس مورد نیاز**
  - **تعداد بیماران در اورژانس**
  - **تعداد پرسنل و سازمان اورژانس**

❖ **هدف ساخت یک مدل که متوسط زمان انتظار بیماران و تعداد بیمارانی که سرویس نمی گیرند را دهد**

- **مدلهای احتمالاتی**

روش مهندسی و تفکر آماری

مقدمه:

- مهندس کسی است مسایل مورد علاقه جامعه را با کاربرد موثر اصول علمی حل می کند.

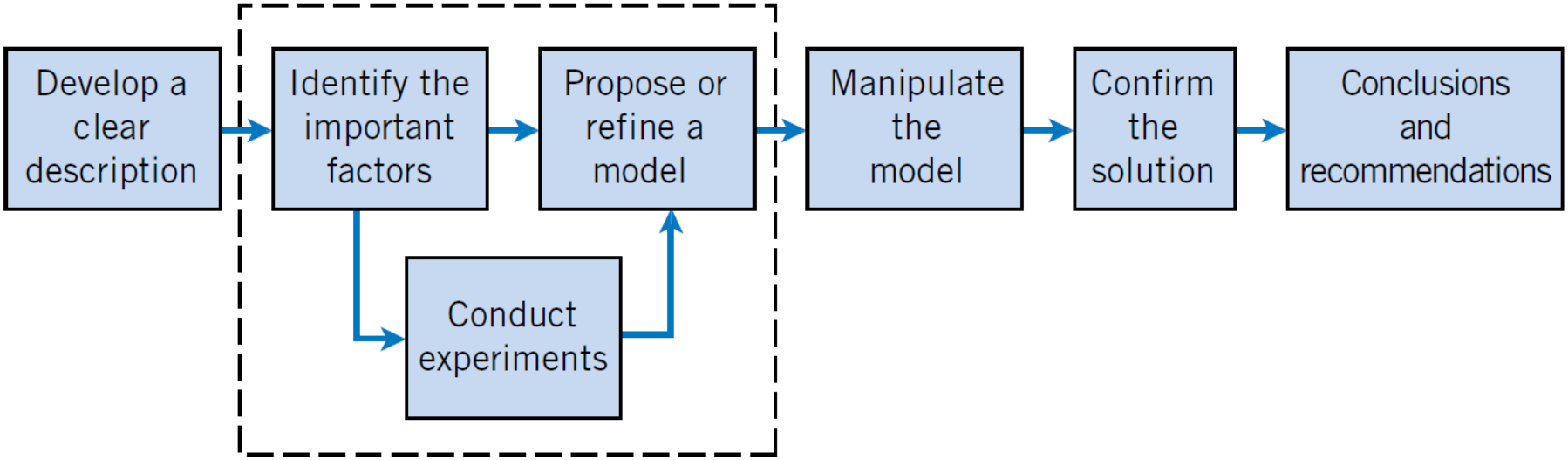
- این کار یا با اصلاح یک محصول موجود انجام می شود یا با طراحی یک محصول یا فرآیند جدید
- روش مهندسی(یا علمی)رویکردی است برای فرمول بندی و حل اینگونه مسایل

مقدمه:

- گامهای روش مهندسی:

1. ایجاد یک تعریف دقیق و واضح از مساله
2. مشخص کردن فاکتورهای مهمی که بر این مساله تاثیر گذارند یا نقشی در حل آن ایفا می کنند
3. پیشنهاد مدلی برای مساله با استفاده از دانش علمی یا مهندسی از پدیده با بیان محدودیت ها و فرضیات مدل
4. انجام آزمایش های مختلف و جمع آوری داده برای آزمودن و اعتبار سنجی مدل و نتیجه گیری از گامهای ۲ و ۳
5. اصلاح مدل بر مبنای داده های مشاهده شده
6. دستکاری مدل برای رسیدن به راه حلی برای مساله
7. انجام آزمایش مناسب برای تایید موثر بودن و کارایی راه حل
8. نتیجه گیری یا ارایه پیشنهادات بر مبنای راه حل

❖ گامهای روش مهندسی:



- مهندسین باید بدانند که چگونه آزمایشهای کارایی را طراحی کنند ، به جمع آوری داده بپردازند ، داده ها را آنالیز و تفسیر کنند و بفهمند که چگونه داده های مشاهده شده به مدل پیشنهادی آنها برای مساله ارتباط دارد.

مقدمه:

- **آمار** به جمع آوری ، ارایه ، تحلیل و استفاده از داده ها برای تصمیم گیری ، حل مسایل و طراحی محصول ها و فرآیندها می پردازد.

- به زبان ساده آمار **علم داده ها** است.

- روش های آماری برای توصیف و درک «طیف متغیرها» (variability) به کار می روند.
  - منظور از طیف متغیرها (variability) مشاهدات پی در پی یک سیستم یا پدیده است که نتایج دقیقاً یکسانی تولید نمی کند.

- همه ما در زندگی روزمره با طیف متغیرها سر و کار داریم ، تفکر آماری کمک می کند که راهی مفید برای به کار بردن این طیف متغیرها در تصمیم گیری های خود بیابیم.

مقدمه:

- **مثال: مصرف بنزین ماشین**
- آیا همیشه ماشین شما مصرف ثابت و یکسانی به ازای هر کیلومتر دارد ؟
- البته که نه!
- بعضی اوقات مقدار مصرف بسیار متفاوت است.
- این طیف متغیر (variability)در مصرف سوخت ، وابسته به فاکتورهای زیادی است: استیل رانندگی ،

  رانندگی در شهر(ترافیک سنگین) یا در بزرگراه (ترافیک روان) ،

  تغییر شرایط ماشین با گذر زمان(تغییر باد لاستیک ها ، استهلاک موتور ، استهلاک سوپاپ ها و ...)

  نوع بنزین و اکتان آن

  حتی شرایط جوی

  و ...
- این عوامل منابعی برای وجود طیف متغیر در سیستم مساله ما هستند.

مقدمه:

- آمار کمک می کند که چارچوبی برای توصیف این طیف متغیرها بیابیم و کمک می کند که یاد بگیریم کدام یک از این عوامل مهمتر هستند یا بیشترین تاثیر را بر کارایی مصرف سوخت دارند

مقدمه:

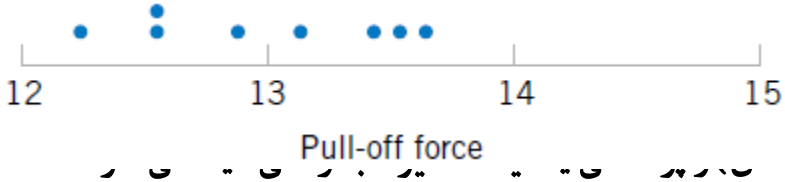
- **مثال: طراحی یک اتصال پلاستیکی برای موتور ماشین**
- مهندسی می خواهد ضخامت این اتصال را تعیین کند وی ضخامت ۳/۳۲ اینچ را در نظر گرفته اما در تاثیر این ضخامت بر تحمل نیروی کششی این اتصال تردید دارد.
- ساخت ۸ نمونه از این اتصال و امتحان آن
- رسیدن به نتایج مقابل: ۱۲.۶ ، ۱۲.۹ ، ۱۳.۴ ، ۱۲.۳ ، ۱۳.۶ ، ۱۳.۵ ، ۱۲.۶ و ۱۳.۱ پوند
- همانطور که انتظار داشتیم همه نمونه ها نتایج یکسانی ندارند
- به این وجود طیف متغیر در آزمایش نیروی کششی می گوییم.
- با توجه به این که آزمایشهای نیروی کششی طیف متغیری را نشان دادند نیروی کششی را یک **متغیر تصادفی** در نظر می گیریم.
- یک راه برای مدل کردن این متغیر تصادفی که آن را X می نامیم توصیف آن به شکل زیر است:

$$X = \mu + \epsilon$$

- در این رابطه μ ثابت است و ϵ اغتشاش تصادفی است.

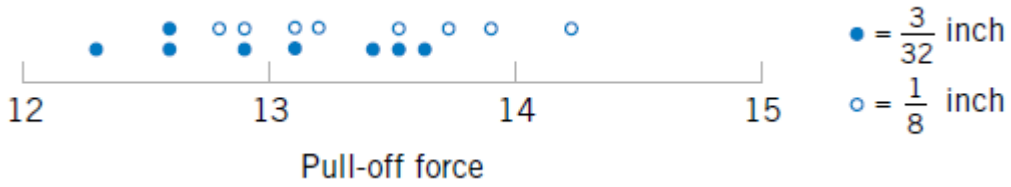
مقدمه:

- بخش ثابت در هر آزمایش بدون تغییر می ماند اما تغییرات کوچک در محیط ، تغییرات در دستگاه آزمایش ، تغییرات در خود اجزا و عوامل دیگر باعث می شوند که مقدار ϵ تغییر کند.
- اگر اغتشاشی در سیستم نباشد مقدار ϵ صفر خواهد بود که چنین چیزی در دنیای واقعی رخ نمی دهد.
- اغلب در مهندسی لازم است که این طیف متغیر را توصیف کنیم ، محاسبه کنیم(کمی سازی) و درنهایت آن را کاهش دهیم.
- نمودار نقطه ای داده های آزمایش:
- این شکل نمایش برای داده ها
- با این نمایش محل دادها (یا و



مقدمه:

- وقتی تعداد مشاهده ها کم باشد سخت است که یک الگوی مشخصی را برای طیف متغیر پیدا کنیم ، هرچند که نمودار نقطه ای برای دیدن داده های غیر عادی مناسب است.
- نیاز به تفکر آماری معمولاً در حل مسایل مهندسی لازم است.
- طراحی اتصال را در نظر بگیرید.
- با آزمایش نمونه های ساخته شده معلوم است که متوسط نیروی کششی ۱۳.۰ پوند است.
- اگر این نیرو کم باشد لازم است ضخامت را افزایش دهیم(مثلاً به ۱/۸ اینچ)
- اگر دوباره نمونه جدیدی بسازیم و در آزمایش نتایج زیر حاصل شود:
- ۱۲.۹ ، ۱۳.۷ ، ۱۲.۸ ، ۱۳.۹ ، ۱۴.۲ ، ۱۳.۲ ، ۱۳.۵ و ۱۳.۱ پوند
- متوسط این مقدار ۱۳.۴ است.



مقدمه:

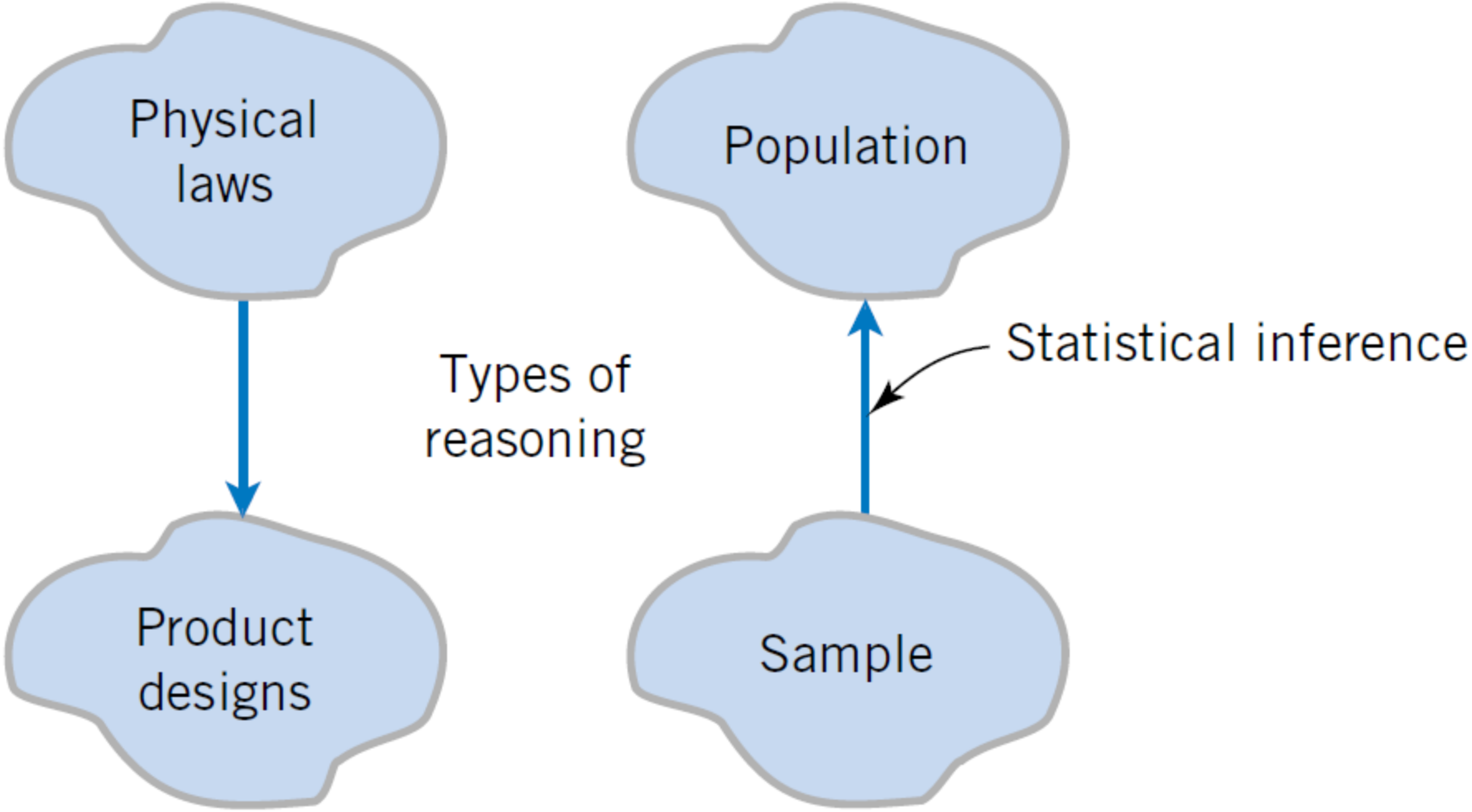
- این ، باعث می شود که نتیجه گیری کنیم که افزایش ضخامت منجر به افزایش نیروی کششی میشود.
- در نتیجه ممکن است سوالات زیر به ذهنمان بیایند؟
  - چطور می شود فهمید که نمونه های دیگر نتایج متفاوتی به دست نمی دهند؟
  - آیا تعداد ۸ نمونه برای رسیدن به نتایج قابل اطمینان کافی هستند؟
  - اگر از این نتیجه آزمایش که افزایش ضخامت باعث افزایش مقاومت می شود استفاده کنیم چه ریسک هایی در پی دارد؟

(برای مثال آیا ممکن است افزایش مقاومت مشاهده شده در نمونه ضخیم تر تنها به خاطر طیف متغیر ذاتی سیستم باشد و افزایش ضخامت نمونه (و افزایش هزینه) در عمل تاثیری در استحکام نداشته باشد؟)

مقدمه:

- معمولا از قوانین فیزیکی برای طراحی محصولات و فرآیندها کمک گرفته می شود.
- ما با این نوع استنتاج از قوانین فیزیکی برای حالتهای خاص آشنا هستیم.
- اما ضروری است که از مجموعه ای از آزمایشهای مشخص نیز بتوانیم برای پاسخ به سوالات قبلی به حالت های کلی برسیم.
- این استنتاج از **نمونه** (مثلا این ۸ اتصال) به یک **جمعیت** (مثلا همه اتصالاتی که به مشتری ها باید فروخته شود) است.
- به این نوع استدلال کردن **«استنتاج آماری»** (Statistical Inference) می گویند.

انواع استدلال



مقدمه:

- واضح است که نتیجه گیری بر مبنای آزمایش از موضوع مورد آزمایش برای همه موضوع ها می تواند منجر به خطا (خطای نمونه گیری) شود.
- اگر نمونه خوب انتخاب شود این ریسک می تواند محاسبه شود و سایز مناسبی از نمونه تخمین زده شود.

مقدمه:

فرض كنيد مي خواهيم اطلاعاتي درباره موارد زير بدست آوريم:

- **ميزان تحصيلات افراد ساكن در شهر تهران**
- **طول عمر لامپ هاي ساخته شده توسط يك كارخانه**
- **ميزان بارندگي يا تصادفات در سال جاري**
- **وزن دانش آموزان يك كلاس**

تعريف:

آمار:

علم آمار به مجموعه روش هاي علمي اطلاق مي شود كه براي جمع آوري اطلاعات اوليه ، مرتب و خلاصه كردن ، طبقه بندي و تجزيه و تحليل اطلاعات اوليه و تفسير آنها به كار مي رود.

- **آمار توصيفي** (Descriptive Statistics)

قسمتي از روش هاي آماري كه تنها به توصيف و تجزيه و تحليل گروه معيني ، بدون تعميم نتايج حاصله به گروه بزرگتر از آن محدود مي گردد ، آمار توصيفي ناميده مي شود.

- **آمار استنباطي** (Inferential Statistics)

به قسمتي از آمار كه مي تواند نتايج حاصل از تجزيه و تحليل نمونه را به جامعه تعميم دهد ، آمار استنباطي گفته مي شود.

فهرست مطالب:

مفاهيم اساسي

جامعه آماري (Statistical Population):

به مجموعه اي از اشياء يا افراد كه حداقل داراي يك ويژگي مشترك باشند ، جامعه آماري مي گوييم.

مثال:

كارمندان شركت ايكس

نمونه (Sample):

به هر بخش از جامعه آماري يك نمونه گفته مي شود.

مثال:

كارشناسان بخش y مي توانند به عنوان يك نمونه از جامعه آماري كارمندان شركت ايكس در نظر گرفته شوند.

متغير(Variable):

ويژگي هايي از اعضاء يك جامعه كه از عضوي به عضو ديگر تغيير مي كند متغير گفته مي شود.

۱- متغير كمي (گسسته ، پيوسته)

قابل اندازه گيري هستند مانند تعداد كارمندان يك شركت ، قد و وزن دانش آموزان ، حجم ، فشار و ...

۲- متغير كيفي

مستقيماً توسط اعداد و ارقام قابل اندازه گيري نيستند مانند رنگ چشم افراد ، گروه خوني و ...

نمونه گيري
(Sampling)

دلايل نمونه گيري:

- هزينه (Cost)
- زمان (Time)
- به روز بودن (Update)
- درستي و صحت (Accuracy)

انواع روش هاي نمونه گيري:

- نمونه گيري تصادفي (Random Sampling)
- نمونه گيري منظم (Regular Sampling)
- نمونه گيري خوشه اي (Cluster Sampling)
- نمونه گيري مرحله اي (Stage Sampling)

نمونه گيري تصادفي (Random Sampling):

در اين حالت افراد يا اشياء بطور تصادفي از ميان عناصر جامعه انتخاب مي شوند و هر يك از آنها براي انتخاب شدن شانس مساوي دارند. (مثلاً انتخاب ۵ لامپ از بين ۱۰۰ لامپ توليد شده توسط يك كارخانه)

نمونه گيري منظم (Regular Sampling):

در اين روش يك نقطه از فهرست افراد يا اشياء جامعه را به طور تصادفي انتخاب مي كنيم و بعد از آن نمونه هاي مورد نظر را به صورت منظم پشت سر هم انتخاب مي كنيم. (مثلاً انتخاب پنج دانشجو بر اساس شماره دانشجويي)

نمونه گيري خوشه اي (Cluster Sampling):

هرگاه جامعه مورد نظر خيلي وسيع و گسترده باشد مانند بررسي وضعيت معاش كارمندان يك كشور ، براي نمونه گيري ابتدا يك سازمان يا اداره را بصورت تصادفي انتخاب مي كنند سپس از بين آنها كارمنداني را بصورت تصادفي انتخاب كرده تا بررسي بيشتري روي آنها انجام شود. (در اين جا واحد نمونه گيري خوشه اي سازمان مي باشد.)

نمونه گيري مرحله اي (Stage Sampling):

شكل گسترش يافته نمونه گيري خوشه اي است. در اين حالت نمونه گيري از جامعه طي چند مرحله انجام مي گيرد.

فراواني

انواع فراواني ها:

۱- فراواني مطلق (Frequency/ Absolute Frequency):

تعداد دفعات تکرار هر داده را فراواني آن داده مي گويند که آن را با f_i نشان مي دهند.

۲- فراواني نسبي (Relative Frequency):

نسبت فراواني مطلق هر داده به مجموع فراواني مطلق تمامي داده ها فراواني نسبي ناميده مي شود که آن را با r_i نشان مي دهند.

۳- فراواني تجمعي (Cumulative Frequency):

حاصل جمع فراواني هر داده (طبقه) با داده هاي (طبقات) قبل از آن را فراواني تجمعي مي نامند که با g_i نمايش داده مي شود.

۴- فراواني تجمعي نسبي (Relative Cumulative Frequency):

نسبت فراواني تجمعي هر داده (طبقه) به مجموع فراواني تجمعي تمامي داده ها که با S_i نمايش داده مي شود.

مثال (داده هاي گسسته):

داده هاي زير تعداد فرزندان تحت پوشش بيمه در ۳۰ خانواده را نشان مي دهد. جدول فراواني آن را بدست آوريد.

۳، ۰، ۱، ۲، ۵، ۲، ۳، ۱، ۱، ۰، ۱، ۴، ۴، ۲، ۲، ۲، ۲، ۰، ۰، ۵، ۲، ۳، ۳، ۳، ۱، ۱، ۱، ۲، ۴، ۵

پاسخ:

ابتدا داده ها را از كوچك به بزرگ براي تشكيل جدول فراواني مرتب مي كنيم.

۵، ۵، ۵، ۴، ۴، ۴، ۴، ۳، ۳، ۳، ۳، ۲، ۲، ۲، ۲، ۲، ۲، ۲، ۲، ۱، ۱، ۱، ۱، ۱، ۱، ۱، ۰، ۰، ۰، ۰، ۰، ۰

	فراوانی	فراوانی نسبی	فراوانی تجمعی	فراوانی تجمعی نسبی
x_i	f_i	r_i	g_i	s_i
۰	۴	۰/۱۳	۴	۰/۱۳
۱	۷	۰/۲۳	۱۱	۰/۳۶
۲	۸	۰/۲۶	۱۹	۰/۶۲
۳	۴	۰/۱۳	۲۳	۰/۷۵
۴	۴	۰/۱۳	۲۷	۰/۸۸
۵	۳	۰/۱	۳۰	۱
جمع	۳۰	۱		

مثال (داده هاي پيوسته):

داده هاي زير طول عمر ۴۰ عدد لامپ را نشان مي دهند كه به نزديك ترين عدد صحيح گرد شده اند. جدول فراواني آن را بدست آوريد.

۱۱	۹	۱۲	۱۵	۲۰	۱۳	۱۴	۱۷	۲۳	۲۲
۸	۱۶	۱۷	۲۱	۱۱	۱۸	۲۱	۱۲	۱۱	۱۰
۱۴	۱۳	۱۹	۱۶	۱۵	۱۷	۲۰	۸	۷	۱۳
۱۵	۱۷	۱۶	۱۴	۲۲	۱۲	۱۱	۹	۱۸	۱۹

پاسخ:

زماني كه با داده هاي پيوسته كار مي كنيم داده ها را به رده هاي (فاصله ها) با طول مساوي تقسيم مي كنيم و فراواني داده ها را در هر رده بدست مي آوريم.

ردەها	خط و نشان	x_i	فراواني f_i	فراواني نسبي r_i	فراواني تجمعي g_i	فراواني تجمعي نسبي s_i
6/5–9/5	𝍸	٨	٥	0/125	٥	0/125
9/5–12/5	𝍸 |||	١١	٨	0/2	١٣	0/325
12/5–15/5	𝍸 ||||	١٤	٩	0/225	٢٢	0/55
15/5–18/5	𝍸 ||||	١٧	٩	0/225	٣١	0/775
18/5–21/5	𝍸 |	٢٠	٦	0/15	٣٧	0/925
21/5–24/5	|||	٢٣	٣	0/075	٤٠	1/0
جمع			٤٠			

شاخص هاي مركزي

۱- میانگین

الف-میانگین حسابی

ب- میانگین هندسی

پ-میانگین هارمونیک

ت- میانگین پیراسته

الف- میانگین حسابي (Arithmetic Mean)

$$\mu = \overline{X} = \frac{1}{N}\sum_{i=1}^{N} X_i = \frac{1}{N}(X_1 + X_2 + \ldots + X_N)$$

مثال:

فرض كنيد داده هاي زير ، نمرات دروس يك دانش آموز باشند ميانگين نمرات آن را حساب كنيد.

۱۲ ، ۸ ، ۱۴ ، ۱۷ ، ۱۶ ، ۱۵ ، ۹ ، ۱۱ ، ۱۳ ، ۱۴

پاسخ:

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

$$= \frac{1}{10}(12+8+14+17+16+15+9+11+13+14) = \frac{129}{10} = 12.9$$

ب- ميانگين هندسي (Geometric Mean)

اگر داده هاي بدست آمده نسبت ، درصد ، شاخص نرخ رشد باشند ، براي بدست آوردن مقدار متوسط از ميانگين هندسي استفاده مي كنيم.

$$G = \sqrt[n]{x_1, x_2, \ldots, x_n} = \sqrt[n]{\prod_{i=1}^{n} x_i}$$

مثال:

فرض کنید میزان تولید کارخانه اي در چهار سال متوالي ۲ ، ۴ ، ۶ و ۲۷ برابر نسبت به سال قبل باشد. مطلوبست میزان افزایش متوسط تولید کارخانه.

پاسخ:

اگر از میانگین حسابي استفاده کنیم داریم:

$$\bar{x} = \frac{1}{4}(2+4+6+27) = 9.75$$

مقدار بدست آمده درست نیست زیرا در این صورت باید در سال چهارم ، برابر سال اول تولید کرده باشد. در صورتي که در سال چهارم برابر سال اول تولید شده باشد:

$$(9.75)^4 = 9036$$

$$2 \times 4 \times 6 \times 27 = 1296$$

$$G = \sqrt[4]{2 \times 4 \times 6 \times 27} = \sqrt[4]{1296} = 6$$

مثال:

ميزان سود شركت بتا در ۵ سال گذشته بر حسب درصد فروش به ترتيب ۲ ، ۳ ، ۴ ، ۴ ، ۳ مي باشد. شاخص مركزي وضع سودآوري شركت را محاسبه كنيد.

پاسخ:

$$G = \sqrt[5]{2 \times 3 \times 4 \times 4 \times 3} = \sqrt[5]{288} = 3.104$$

ج- ميانگين هارمونيك (Harmonic Mean)

اگر داده ها بصورت تركيبي باشند و يا اينكه داده ها داراي واحدهايي بصورت كسري (واحد سرعت متوسط (زمان/ مسافت)) باشند از ميانگين هارمونيك استفاده مي شود.

$$H = \frac{n}{\sum_{i=1}^{n} \frac{1}{x_i}}$$

مثال:

اتومبيلي فاصله بين دو شهر A و B را كه ۹۰ كيلومتر است با سرعت متوسط ۶۰ كيلومتر در ساعت پيموده است و با سرعت متوسط ۳۰ كيلومتر در ساعت همين مسير را برگشته است. سرعت متوسط اين اتومبيل را در رفت و برگشت محاسبه كنيد.

پاسخ:

ميانگين حسابي

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i = \frac{1}{2}(60+30) = 45$$

ميانگين هارمونيك

$$H = \frac{n}{\sum_{i=1}^{n}\frac{1}{x_i}} = \frac{2}{\frac{1}{60}+\frac{1}{30}} = 40$$

ادامه:

حل فيزيكي مسئله:

$$t_1 = 1.5 \qquad t_2 = 3 \qquad t = t_1 + t_2 = 1.5 + 3 = 4.5$$

$$x_1 = 90 \qquad x_2 = 90 \qquad x = x_1 + x_2 = 90 + 90 = 180$$

$$\bar{v} = \frac{x}{t} = \frac{180}{4.5} = 40$$

نتيجه:

براي حل اين مثال بايد از ميانگين هارمونيك استفاده شود.

مثال:

اگر در مثال قبل اتومبیل نصف مدت حرکت را با سرعت متوسط ۶۰ کیلومتر در ساعت و نصف مدت دیگر را با سرعت ۳۰ کیلومتر در ساعت پیموده بود ، سرعت متوسط در رفت و برگشت را حساب کنید.

پاسخ:

حل فیزیکي مسئله:

$$180 = 60 \times \frac{t}{2} + 30 \times \frac{t}{2} = 45t \qquad t = 4 \qquad \bar{v} = 45$$

میانگین حسابي

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i = \frac{1}{2}(60+30) = 45$$

میانگین هارمونیك

$$H = \frac{n}{\sum_{i=1}^{n}\frac{1}{x_i}} = \frac{2}{\frac{1}{60}+\frac{1}{30}} = 40$$

سوال:

با توجه به دو مثال حل شده قبلي مشخص كنيد كه براي محاسبه ميانگين داده هاي با واحد كسري چه زماني مي توان از ميانگين حسابي استفاده كرد و چه زماني از ميانگين هارمونيك.

مثال:

سه ماشين به توليد يك كالا مشغول اند ، اولي يك كالا را در ۲ ، دومي در ۳ و سومي در ۶ دقيقه توليد مي كنند. اگر اين سه ماشين با هم كار كنند بطور متوسط يك كالا در چند دقيقه توليد مي شود ؟

پاسخ:

$$\overline{X} = \frac{n}{\sum_{i=1}^{n} \frac{1}{x_i}} = \frac{3}{\frac{1}{2} + \frac{1}{3} + \frac{1}{6}} = 3$$

د- میانگین پیراسته (Truncated Mean)

اگر از میان داده هاي بدست آمده تعداد **k** داده که با بقیه اختلاف بیشتري دارند حذف گردند ، میانگین پیراسته بدست خواهد آمد.

$$\bar{x}_p = \frac{1}{n-k}\sum_{i=1}^{n-k} x_i$$

مساله:

اتومبيلي مسيري را با سرعت ۱۰۰ کيلومتر بر ساعت رفته ، و ۱/۳ مسير را با سرعت ۸۰ کيلومتر بر ساعت و باقيمانده را با سرعت ۱۲۰ کيلومتر بر ساعت برگشته ، سرعت متوسط اين اتومبيل چقدر بوده است ؟

۱) ۹۰ ۲) ۱۰۰ ۳) ۱۰۱/۴ ۴) ۱۰۲/۸

۲- ميانه (Median)

ميانه مقداري است كه ۵۰% داده ها قبل و ۵۰% داده ها بعد از آن قرار دارند. و يا به عبارتي احتمال وقوع آن ۵۰% مي باشد.

ويژگی ها:

الف- ميانه مشاهدات را به دو بخش مساوي تقسيم مي كند.
ب- منحصر به فرد است.
ج- تحت تأثير داده هاي پرت قرار نمي گيرد.
د- محاسبه آن ساده است.

اگر داده ها را بصورت غير نزولي مرتب كنيم ، ميانه عدد وسطي مي باشد به شرط اينكه تعداد داده ها فرد باشد و اگر تعداد داده ها زوج باشد ميانه برابر با ميانگين دو عدد وسطي است.

مثال:

میانه اعداد زیر را بدست آورید.

A: ۱۴، ۱۱، ۳، ۷، ۹، ۱۱، ۷، ۵، ۴

B: ۱۲، ۷، ۳، ۴، ۳، ۵

پاسخ:

مرتب سازي

A: ۳، ۴، ۵، ۷، ۷، ۹، ۱۱، ۱۱، ۱۴

B: ۳، ۳، ۴، ۵، ۷، ۱۲

$$Median = \frac{4+5}{2} = 4.5$$

مثال:

فرض كنيد در يك جشن از ۱۰ نفر ميهمان مي خواهيد كه هر كدام مبلغي را براي كمك به يك موسسه خيريه پرداخت نمايند. اعداد زير مربوط به مبالغ پول پرداختي توسط اين ۱۰ نفر برحسب هزار تومان است. از چه پارامتري براي نشان دادن ميزان شاخص پرداختي اين ده نفر استفاده مي كنيد؟

۳۵ ، ۳۷ ، ۳۹ ، ۴۰ ، ۴۳ ، ۴۴ ، ۴۴ ، ۹۹۰۰

پاسخ:

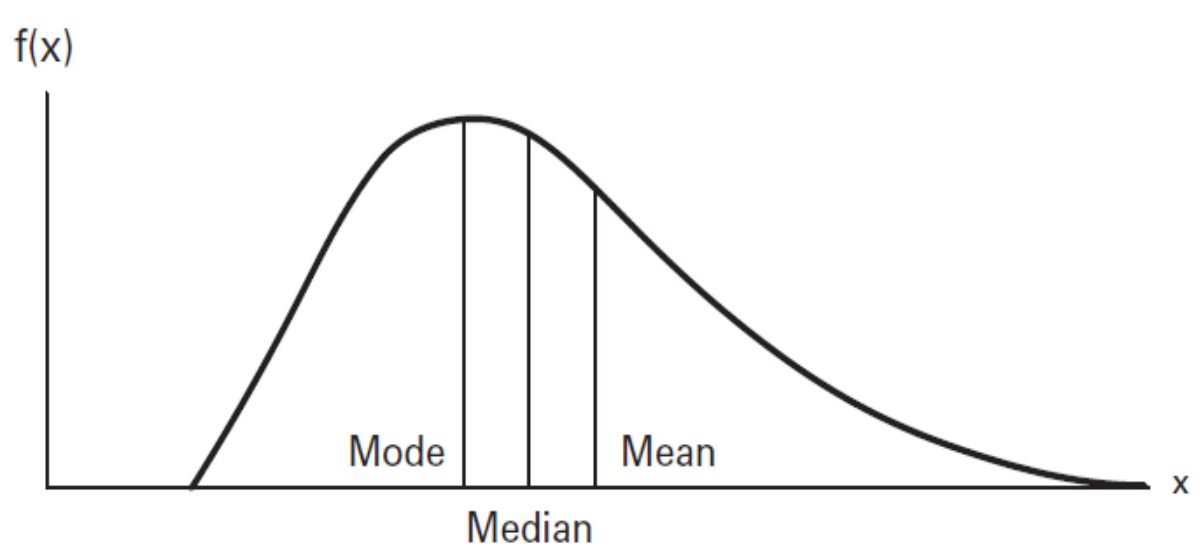
$$\bar{x} = 4630$$

$$Median = \frac{40+43}{2} = 41.5$$

گرايش افراد به سمت عدد ۴۱/۵ مي باشد. يعني مهمانان با احتمال ۵۰ درصد مبلغ ۴۱/۵ هزار تومان كمك مي كنند.

۳- نما(Mode)

متغيري كه داراي بيشترين فراواني است نما ناميده مي شود.



شاخص هاي پراكندگي

مثال:

فرض کنید طول عمر باتري هاي ساخته شده توسط دو کارخانه بصورت زیر باشد. به نظر شما باتري هاي توليدي کدام کارخانه بهتر است ؟

A: ۱۰۰ ، ۱۰۲ ، ۱۰۳ ، ۱۰۵ ، ۱۰۷ ، ۱۰۷ ، ۱۱۱

B: ۹۶ ، ۱۰۰ ، ۱۰۴ ، ۱۰۵ ، ۱۰۷ ، ۱۰۷ ، ۱۱۶

پاسخ:

میانگین: $\overline{X_A} = 105$ $\overline{X_B} = 105$

میانه: $Median_A = 105$ $Median_B = 105$

نما: $Mode_A = 107$ $Mode_B = 107$

نتیجه: پارامترهاي مرکزي دقیقاً برابر هم هستند. بنابراین باید از پارامترهاي پراکندگي کمك گرفت.

۱- دامنه

دامنه عبارتست از اختلاف بزرگترین و کوچکترین داده.

$$R = X_{Max} - X_{Min}$$

۲- واریانس

واریانس عبارتست از میانگین مربع انحراف داده ها از میانگین.

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

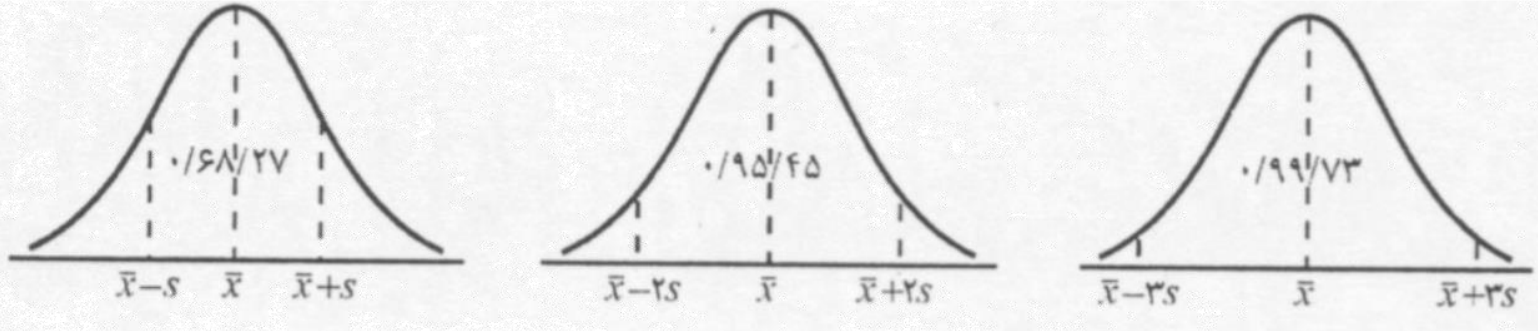
۳-انحراف معیار

جذر واریانس را انحراف معیار یا انحراف استاندارد مي گویند.

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

در توزيع هاي متقارن روابط زير بطور تقريبي برقرار مي باشد.

- در حدود ۶۸/۲۷ درصد داده ها در فاصله ($\bar{x} - s$, $\bar{x} + s$) قرار دارند.
- در حدود ۹۵/۴۵ درصد داده ها در فاصله ($\bar{x} - 2s$, $\bar{x} + 2s$) قرار دارند.
- در حدود ۹۹/۷۳ درصد داده ها در فاصله ($\bar{x} - 3s$, $\bar{x} + 3s$) قرار دارند.



مثال:

دامنه ، واریانس و انحراف معیار داده هاي زیر را محاسبه کنید.

A: ۳ ، ۴ ، ۶ ، ۷ ، ۱۰ B: -۲۰ ، ۵ ، ۱۵ ، ۲۴

پاسخ:

دامنه: $R_A = 10 - 3 = 7$ $R_B = 24 - (-20) = 44$

میانگین:

$$\overline{X_A} = (3+4+6+7+10)/5 = 6$$

$$\overline{X_B} = (-20+5+15+24)/4 = 6$$

واریانس:

$$S_A{}^2 = \left((-3)^2 + (-2)^2 + 0^2 + 1^2 + 4^2\right)/4 = 7.5$$

$$S_B{}^2 = \left((-20)^2 + 5^2 + 15^2 + 24^2\right)/3 = 360.67$$

انحراف معیار:

$$S_A = \sqrt{S_A{}^2} = 2.74 \qquad S_B = \sqrt{S_B{}^2} = 18.99$$

مثال:

حقوق پرداختي به كارمندان شركت آلفا بطور متوسط ۱۵ هزار تومان با انحراف معيار ۳ هزار تومان است. اگر ۲۰% ميانگين به حقوق كارمندان اضافه شود ، به ترتيب ميانگين و انحراف معيار حقوق پرداختي چقدر خواهد شد؟

پاسخ:

ميانگين:

$$\overline{X} = 15000 + 15000 \times \frac{20}{100} = 18000$$

انحراف معيار:

$$S = 3000$$

www.engclubs.net

پایان