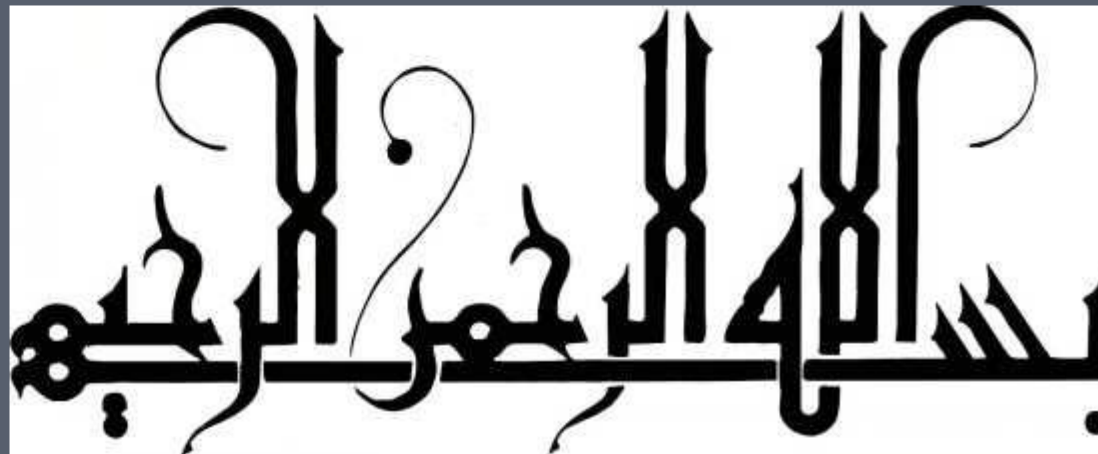


به نام آنکه به شماره موجودات آگاه است

مفاهیم آمار و استفاده از نرم افزار SPSS

امین قنبرنژاد
عضو هیات علمی دانشکده بهداشت

اهداف کارگاه

- مفاهیم اولیه آمار
- تعریف انواع متغیرها
- نقش متغیرها در پژوهش
- مقیاس اندازه گیری متغیرها
- نحوه ارائه آمار توصیفی
- آمار استنباطی

مفاهیم اولیه آمار

- تعریف آمار و انواع آن از نظر تجزیه و تحلیل
- جامعه آماری و انواع آن
- سرشماری و نمونه گیری
- متغیر و انواع آن
- مقیاسهای اندازه گیری

مفاهیم اولیه آمار

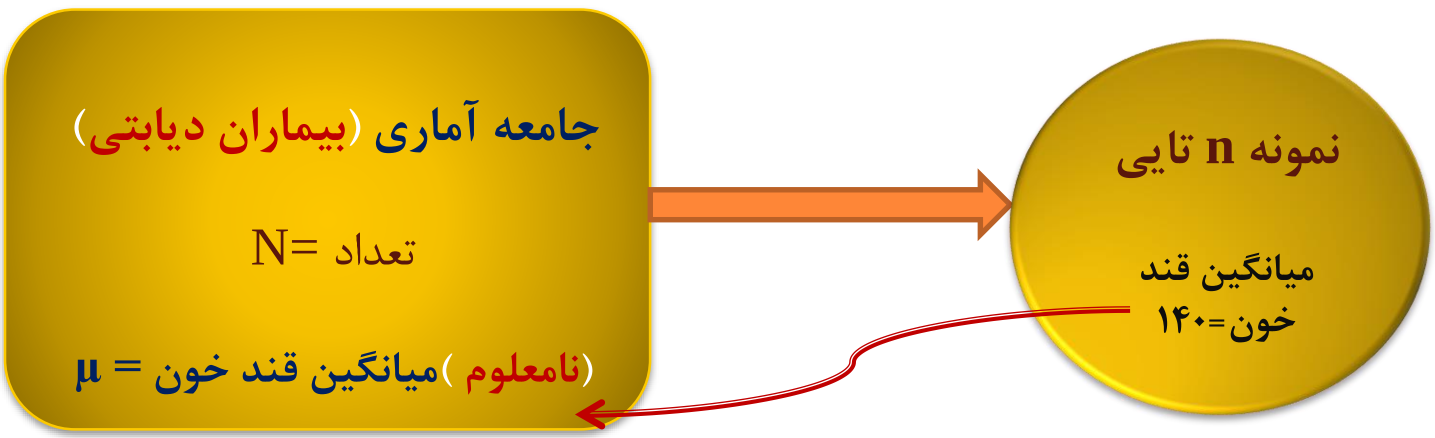
معادل کلمه آمار در زبان انگلیسی «statistics» است.
علم روش‌های **جمع‌آوری** (Collection)، **سازماندهی** (Organization)، **تلخیص** (Summarization)، **نمایش** (Representation) و **تجزیه و تحلیل** (Analysis) داده‌ها دانست.

مفاهیم اولیه آمار

- آمار از نظر تجزیه و تحلیل اطلاعات
  - **آمار توصیفی**

    جداول توزیع فروانی و درصد، نمودار و آماره‌های توصیفی دیگری چون میانگین، میانه ، مد و..
  - **آمار استنباطی**

    تعمیم از جزء به کل (برآورد نمونه به پارامتر جامعه)



مفاهیم اولیه آمار

- **مجموعه‌ای از افراد، اشیاء، مکانها و...که در یک یا چند صفت با هم وجه اشترک دارند را جامعه آماری گویند.**

- مطالعه « بررسی وضعیت بهداشتی اغذیه فروش های استان هرمزگان در سال ۱۳۹۷ »
- مطالعه « تعیین میزان شیوع فشار خون در مراجعه کنندگان به مرکز سلامت خلیج فارس شهر بندرعباس »
- مطالعه « طول مدت بستری در بخش جراحی بیمارستانهای تابعه علوم پزشکی هرمزگان »

- **جامعه آماری محدود:** بیماران بستری شده در بیمارستان شهید محمدی بندرعباس - افراد مبتلا به بیماری ایدز در کشور

- **جامعه آماری نامحدود:** مانند تعداد ستارگان آسمان، تعداد گلبولهای سفید خون یا تعداد ماهیهای یک دریا

- **در مطالعات عموماً جامعه هایی که بیش از ۱۰۰۰۰ می باشد را به عنوان جامعه های نا محدود در نظر می گیرند.**

مفاهیم اولیه آمار: روش جمع آوری اطلاعات

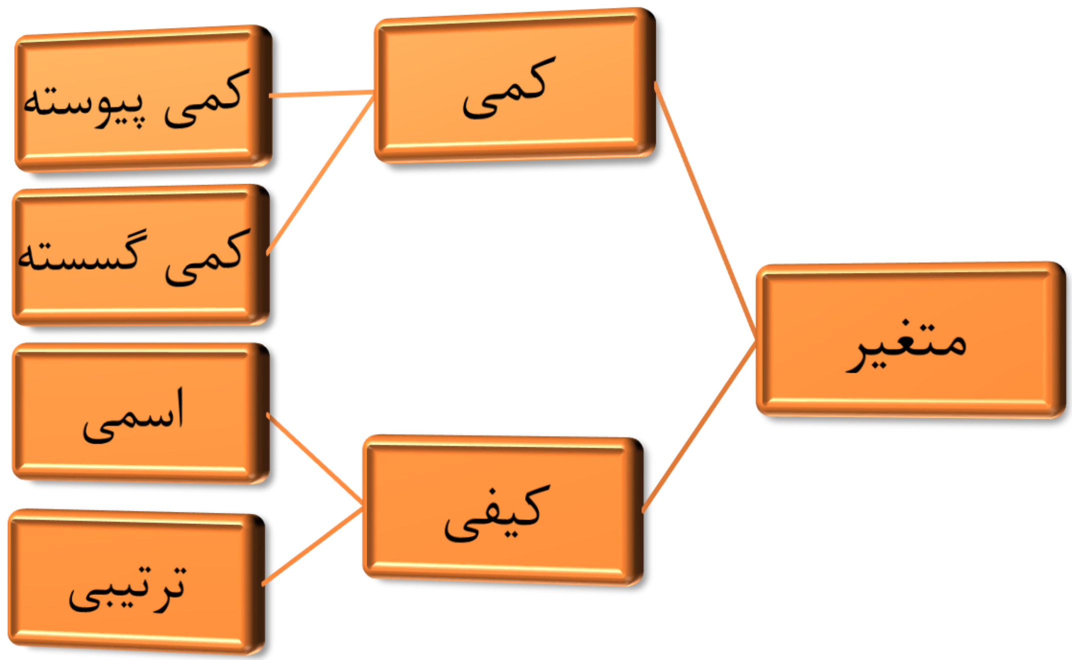
روش سرشماری: جمع آوری اطلاعات کلیه افراد جامعه آماری

«سرشماری عمومی نفوس و مسکن»

روش نمونه گیری: نمونه آماری جزئی از جامعه آماری می‌باشد که اگر به درستی انتخاب شده باشد معرف تمامی خصوصیات جامعه آماری می باشد.

متغیر و انواع آن

- **تعریف: هر ویژگی از اعضای جامعه آماری می‌باشد که از عضوی به عضو دیگر تغییر پذیر بوده و اندازه‌های مختلفی را به خود می‌گیرد.**
- **مثال: وزن، جنسیت، قد، گروه خون، طول عمر انسان، سابقه کار، تعداد فرزندان و ...**



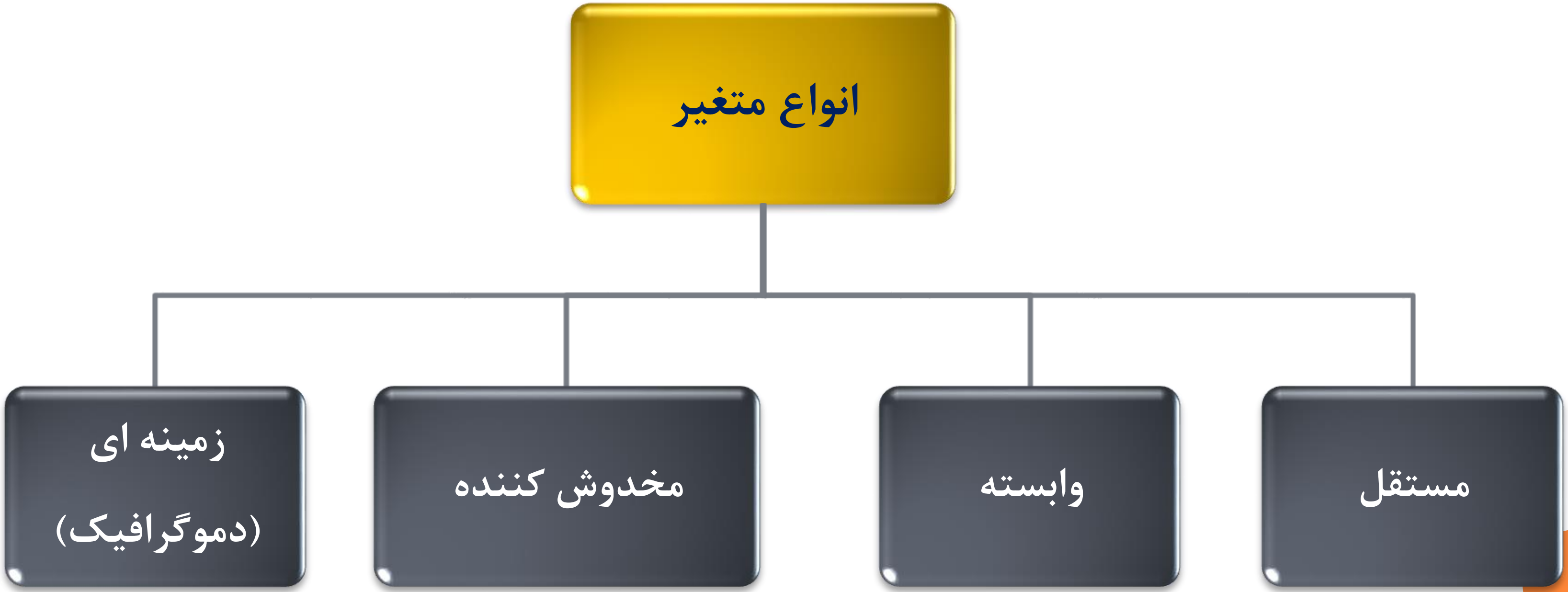
متغیر و انواع آن

- **متغیر کمی:** متغیری که با ابزاری که جامعه بشری در اختیار دارد مقدار آن را با دقتی معین می توان اندازه گیری نمود و میتوان آنرا با یک عدد واحد دار نمایش داد. مانند: سن، وزن، قد، هزینه بستری بیمار در بیمارستان و...

- **متغیر کیفی:** صفات یا ویژگیهایی که واحد اندازه‌گیری ندارند و قابل اندازه‌گیری نمی‌باشند، مانند: جنسیت، رنگ چشم، شغل، گروه خون و...

نقش متغیرها در پژوهش

نقش متغیر بطور کامل مربوط به اهداف مطالعه است. به عبارت دیگر این هدف مطالعه است که تعیین می کند که یک متغیر چه نقش و جایگاهی دارد.



نقش متغیرها در پژوهش

- **متغیر مستقل: متغیری است که دستکاری می شود یا رفتار آن مطالعه می شود تا اثر یا اثرات آن مشخص شود. متغیر مستقل را با حرف x نمایش می دهند.**

- **متغیر وابسته: متغیری که تغییرات آن ناشی از تغییرپذیری متغیر مستقل می باشد. آنرا با حرف Yنشان می دهند.**

مثال: تاثیر میزان BMI بر چربی خون (LDL) در بین بیماران مبتلا به فشار خون در بیمارستان شهید محمدی بندرعباس در سال ۱۳۹۷

تمرین : در مطالعه «بررسی تاثیر الکل بر ابتلا به بیماری قلبی در بین مردان ۶۰-۳۵ سال شهر بندرعباس»

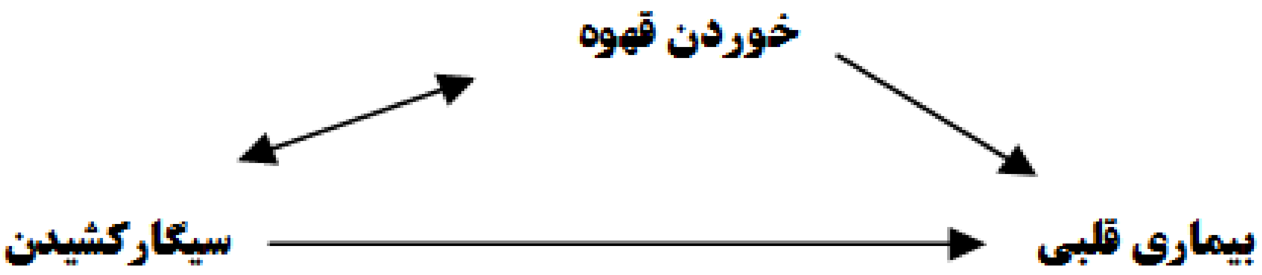
جامعه آماری- متغیر مستقل – متغیر وابسته

؟

نقش متغیرها در پژوهش

- **متغیر مداخله گر (مخدوش کننده):**

آن متغیری است که بر روی رابطه علت معلولی بین دو یا چند متغیر تاثیر می گذارد و باعث قویتر یا ضعیفتر شدن رابطه بین متغیرها نسبت به حد واقعی آنها می شود.



نقش متغیرها در پژوهش

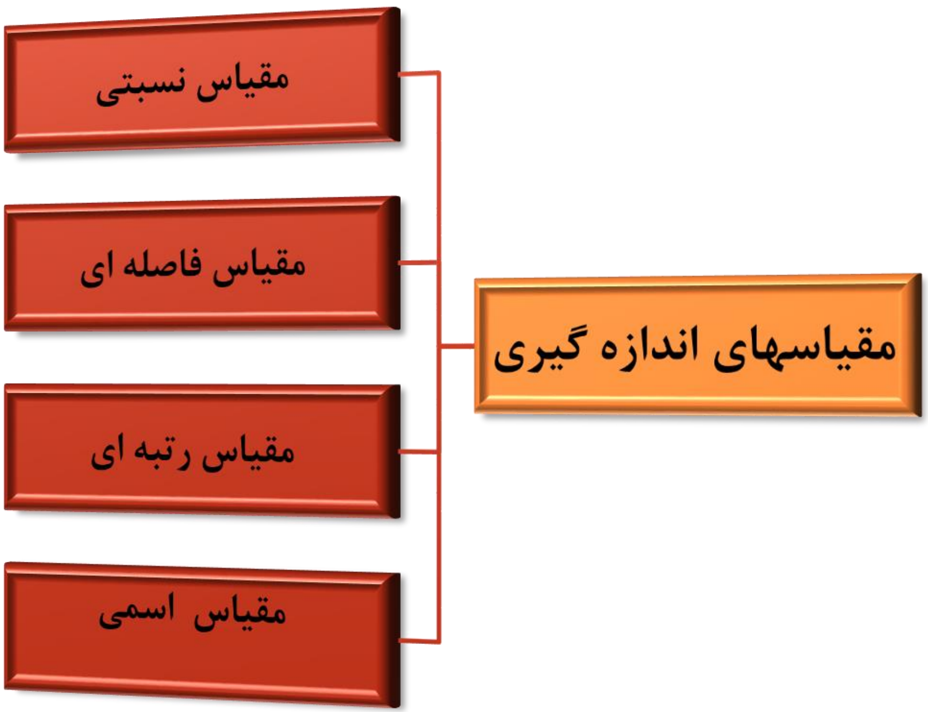
- **متغیر زمینه ای (جمعیت شناسی یا دموگرافیک):**

در مطالعاتی که بر روی انسانها صورت می گیرد برخی ویژگیهای افراد جامعه که خصوصیات جامعه آماری را بیشتر و بهتر توصیف می کنند و به شناخت کامل از آن جامعه کمک می کنند.

تحصیلات، محل سکونت، شغل، سابقه کار و....

مقیاس اندازه گیری متغیرها

اندازه گیری : یعنی اختصاص اعداد به حالتها یا وضعیتهای مختلف یک متغیر، بمنظور کمی ساختن آن



مقیاس اندازه گیری متغیرها: مقیاس اسمی

- مقادیر عددی صرفاً نام مقوله‌ها هستند
- اعداد فقط برای متمایز کردن وضعیتها استفاده می‌شوند
- این مقادیر هیچ اندازه‌ای را نشان نمی‌هند
- این مقادیر هیچ برتری یا رابطه ای را مشخص نمی‌کنند
- از عملهای حسابی نمی‌توان در مورد اعداد استفاده کرد

- مانند: گروه خون

O=4   AB=3   B=2   A=1

مقیاس اندازه گیری متغیرها: مقیاس رتبه ای

- مقادیر عددی نام مقوله‌ها هستند
- اعداد برای متمایز کردن وضعیتها استفاده می‌شوند
- این مقادیر هیچ اندازه‌ای را نشان نمی‌هند
- این مقادیر مشخص می کند که فردی نسبت به فرد دیگر از نظر صفت مورد نظر اندازه‌گیری شده برتر و یا پایین تر می‌باشد

مثال: کلاس تحصیلی دانش آموزان اول تا پنجم ابتدایی چنین مقیاسی را دارد.

کلاس اول(۱) کلاس دوم (۲) کلاس سوم (۳) کلاس چهارم(۴) کلاس پنجم(۵)

اعداد ۱ تا ۵ را به کلاسهای اول تا پنچم نسبت داده‌ایم. این اعداد گویای اندازه واقعی نمی‌باشند.

مقیاس اندازه گیری متغیرها: مقیاس فاصله ای

- مقادیر عددی هم ترتیب و هم اندازه را نشان می‌دهند.
- فاصله میان مقادیر یکسان است.
- مقدار صفر نشان دهنده صفر مطلق نیست.
- این مقادیر روابط را برحسب ترتیب واندازه نشان می دهند.
- از این مقادیر نمی‌توان برای ایجاد نسبت‌ها استفاده کرد.
- دستکاری آماری خیلی محدود نیست.

- مثال : درجه حرارات که صفر آن صفر واقعی نیست.

مقیاس اندازه گیری متغیرها: مقیاس فاصله ای

$$F = \frac{9}{5}C + 32$$

$$\left| \begin{array}{l} C = 10^{o} \Rightarrow F = (\frac{9}{5})(10^{o}) + 32 = 18 + 32 = 50 \\ C = 20^{o} \Rightarrow F = (\frac{9}{5})(20^{o}) + 32 = 36 + 32 = 68 \\ C = 30^{o} \Rightarrow F = (\frac{9}{5})(30^{o}) + 32 = 54 + 32 = 86 \end{array} \right\rangle \begin{array}{l} 68 - 50 = 18 \\ \\ 86 - 68 = 18 \end{array}$$

مقیاس اندازه گیری متغیرها: مقیاس نسبتی

- مقادیر عددی هم ترتیب و هم اندازه را نشان می‌دهند.
- فاصله میان مقادیر عددی یکسان است.
- مقدار صفر نشان دهنده فقدان موضوع اندازه‌گیری شده‌است.
- این مقادیر روابط را برحسب ترتیب واندازه نشان می دهند.
- بهترین نوع مقیاس می باشد .
- دستکاری آماری محدود نیست.

موجودی حساب - وزن افراد – فشار خون – سن - غلظت

مقیاس اندازه گیری متغیرها

نوع مقیاس	وجود ترتیب در طبقات	فواصل مساوی طبقات	وجود صفر مطلق	عملیات مجاز آماری	عملیات مجاز ریاضی
اسمی	–	–	–	نما- فراوانی	هیچکدام
ترتیبی	+	–	–	نما- فراوانی- میانه- همبستگی اسپیرمن	هیچکدام
فاصله ای	+	+	–	تمام عملیات آماری	جمع و تفریق
نسبتی	+	+	+	تمام عملیات آماری	تمام عملیاتهای ریاضی

تمرین

بررسی شیوع کم شنوایی مادرزادی در نوزادان استان هرمزگان طی سالهای ۱۳۹۲ -۱۳۹۷

مشخصات متغیر	نقش متغیر در مطالعه				کمی		کیفی		مقیاس
	مستقل	زمینه ای	مخدوش کننده	وابسته	پیوسته	گسسته	اسمی	رتبه ای	
جنس									
سال تولد									
شدت کم شنوایی									
سابقه در خانواده									
سن مادر									
سن حاملگی (gestational age)									

تمرین: پاسخ

بررسی شیوع کم شنوایی مادرزادی در نوزادان استان هرمزگان طی سالهای ۱۳۹۲ -۱۳۹۷

مشخصات متغیر	نقش متغیر در مطالعه				کمی		کیفی		مقیاس
	مستقل	زمینه ای	مخدوش کننده	وابسته	پیوسته	گسسته	اسمی	رتبه ای	
جنس	*						*		اسمی
سال تولد		*				*			نسبتی
شدت کم شنوایی				*				*	رتبه ای
سابقه در خانواده	*						*		اسمی
سن مادر	*				*				نسبتی
سن حاملگی (gestational age)	*				*				نسبتی

روشهای تجزیه و تحلیل داده ها

- تقسیم بندی روشهای تجزیه و تحلیل
  - توصیفی
    - مانند: جدول توزیع فراوانی-نمودار-میانگین انحراف معیار و...
  - مقایسه ای
    - مانند: آزمون t-test آزمون F
  - ارتباط یا همبستگی
    - مانند : همبستگی پیرسون-همبستگی اسپیرمن و...

ارائه آمار توصیفی

- **جداول توزیع فراوانی:**
  - **(کیفی–کمی گسسته–کمی پیوسته)**

- **ترسیم نمودار مناسب داده ها :**
  - **(میله ای –ستونی– هیستوگرام–دایره ای–ساقه و برگ–جعبه ای)**

- **شاخصهای تمرکز و پراکندگی**
  - **(میانگین – میانه – نما – واریانس – انحراف معیار)**

آمار توصیفی: جدول توزیع فراوانی

جدول توزیع فراوانی

- جدول توزیع فراوانی برای داده های کیفی
- جدول توزیع فراوانی برای داده های کمی گسسته
- جدول توزیع فراوانی برای داده های کمی پیوسته

ارائه آمار توصیفی: جدول فراوانی

فرض کنید که X یک متغیر است که می‌تواند حالات A_1 و A_2 و ...و A_s را اختیار کند و فراوانیهای هر یک از این حالات به ترتیب n_1، n_2، ...و n_s باشد.

سطوح متغیر	n_i (فراوانی مطلق)	r_i (فراوانی نسبی)	$100 \times r_i$ (درصد فراوانی نسبی)	cr_i فراوانی نسبی تجمعی	cf_i (فراوانی تجمعی)
A_1	n_1	r_1	$r_1 \times 100$	r_1	$n_1 = cf_1$
A_2	n_2	r_2	$r_2 \times 100$	$r_1 + r_2$	$n_1 + n_2 = cf_2$
.	.	.	.	.	.
.	.	.	.	.	.
.	.	0	.	.	.
A_s	n_s	r_s	$r_s \times 100$	1	$n_1 + n_{2+} \ldots + n_s = n = cf_s$
	n	۱	۱۰۰		

آمار توصیفی: جدول فراوانی- کیفی اسمی

مثال: به منظور بررسی گروه خون افراد یک منطقه نمونه ای به اندازه تعداد ۳۰ نفر از افراد آن منطقه را انتخاب و گروه خونشان را پرسیده ایم اطلاعات آنها بصورت زیر نمایش داده شده:

O AB B B AB A AB O AB AB B AB B O O
A B O O AB O AB B A B AB A O A O

گروه خون	f_i	r_i	$r_i \times 100$	cr_i	cf_i
A	۵	۰/۱۷	۱۷		
B	۷	۰/۲۳	۲۳		
AB	۹	۰/۳	۳۰		
O	۹	۰/۳	۳۰		
جمع	۳۰	۱	۱۰۰		

آمار توصیفی: جدول فراوانی- کمی گسسته

مثال: تعداد دندانهای افتاده در بین ۵۰ نفر دانش آموز زیر ۱۰ سال

تعداد دندان	f_i	r_i	$r_i \times 100$	cr_i	cf_i
۱	۲	۰/۰۴	۴	۰/۰۴	۲
۲	۹	۰/۱۸	۱۸	۰/۲۲	۱۱
۳	۱۳	۰/۲۶	۲۶	۰/۴۸	۲۴
۴	۱۳	۰/۲۶	۲۶	۰/۷۴	۳۷
۵	۹	۰/۱۸	۱۸	۰/۹۲	۴۶
۶	۳	۰/۰۶	۶	۰/۹۸	۴۹
۷ و بیشتر	۱	۰/۰۲	۲	۱	۱۰۰
جمع	۵۰	۱	۱۰۰		

آمار توصیفی: جدول فراوانی- کمی پیوسته

مثال: در یک نمونه ۲۰ نفری از بین دانش آموزان سال چهارم دبیرستان میزان قد آنها بصورت زیر حاصل شده است. جدول توزیع فراوانی این مشاهدات بصورت زیر بدست می آید.

۱۸۲–۱۷۵–۱۶۱–۱۶۰–۱۴۹–۱۵۰–۱۷۳–۱۸۵–۱۶۳–۱۵۳

۱۷۰–۱۷۶–۱۷۹––۱۶۸–۱۷۰–۱۷۸-۱۴۵–۱۶۶–۱۵۹–۱۴۶

$$k = 1 + 3/3 \log n = 1 + 3/3 \log 20 = 1 + 3/3(1/3) = 1 + 4/2 \cong 5$$

فواصل طبقات را با W نمایش داده و از رابطه زیر بدست می آید:

$$W = \frac{X_{\max} - X_{\min}}{k} = \frac{185 - 145}{5} = \frac{40}{5} = 8$$

آمار توصیفی: جدول فراوانی- کمی پیوسته

حدود طبقات	نماینده طبقه X_i	f_i	r_i	$r_i \times 100$	cr_i	cf_i
[145-153)	149	4	0.2	20	20	4
[153-161)	157	3	0.15	15	35	4+3=7
[161-169)	165	3	0.15	15	50	4+3+3=10
[169-177)	173	5	0.25	25	75	4+3+3+5=15
[177-185]	181	5	0.25	25	100	4+3+3+5+5=20
جمع		20	1	100		

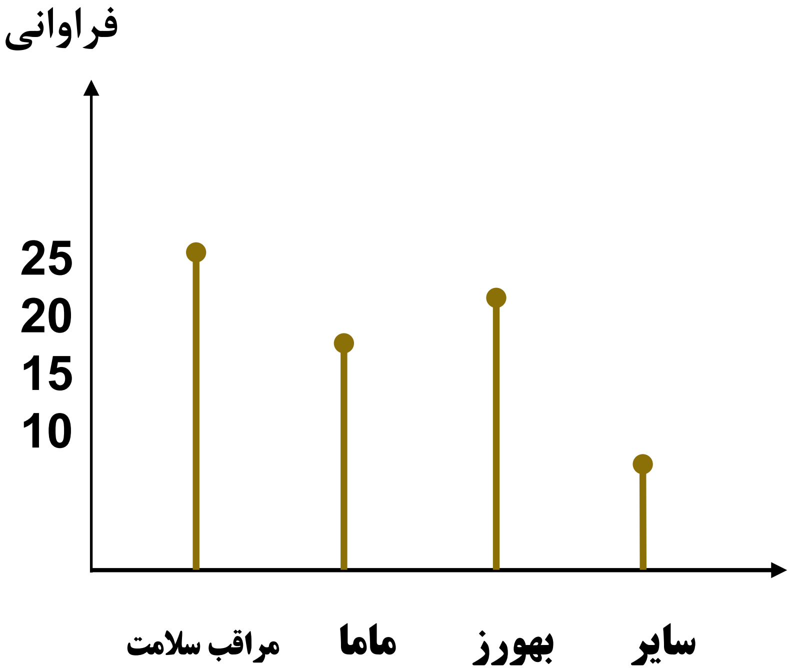
آمار توصیفی: نمودار ها

- **انواع نمودار**
  - نمودار میله ای
  - نمودار ستونی
  - هیستوگرام
  - نمودار چند بر فراوانی
  - نمودار دایره ای
  - نمودار شاخه و برگ
  - نمودار جعبه ای

آمار توصیفی: نمودار میله ای

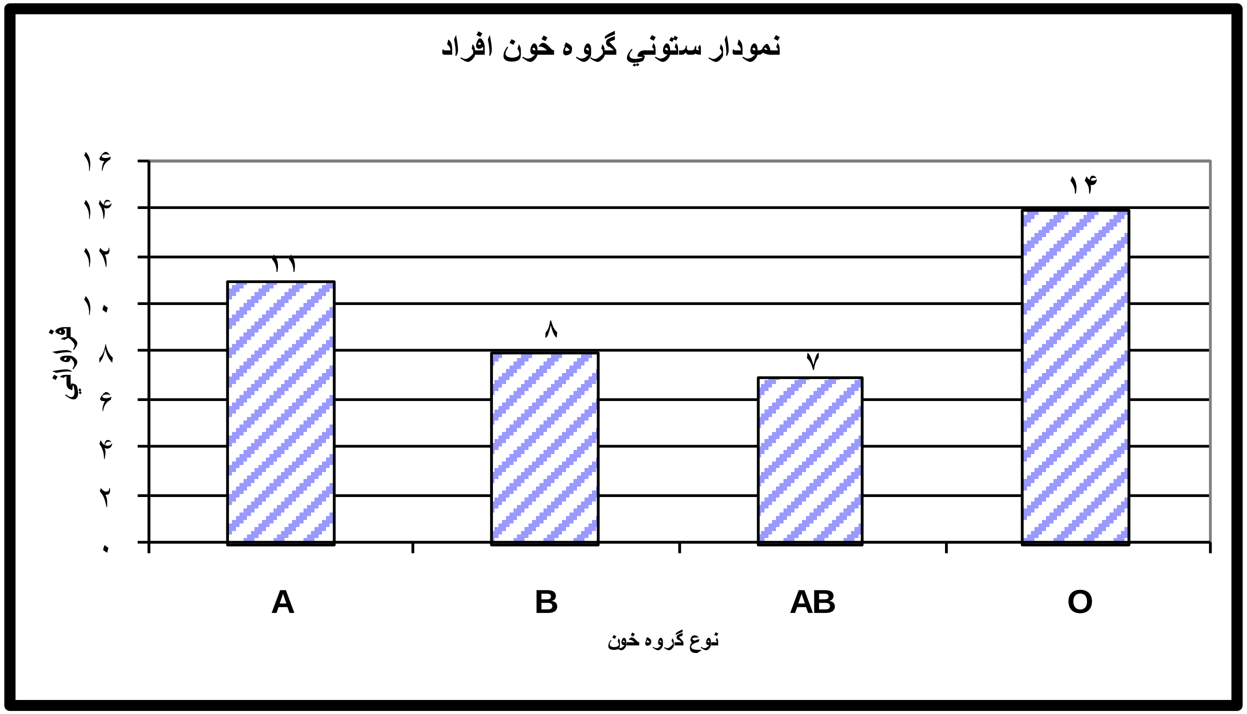
مثال: توزیع فراوانی تعداد ۷۰ نفر از پرسنل یک مرکز جامع سلامت شهری - روستایی بصورت زیر است

تخصیص	فراوانی
مراقب سلامت	۲۵
ماما	۱۶
بهورز	۲۰
سایر	۹



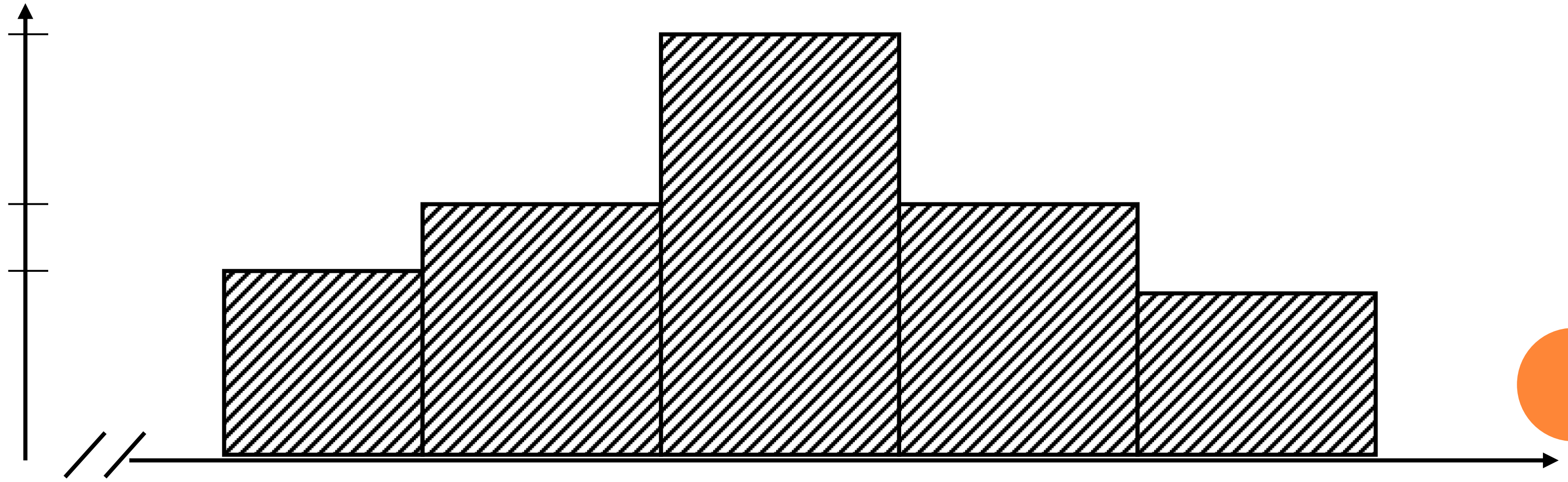
آمار توصیفی: نمودار ستونی (BAR CHART)

- این نمودار برای داده های کیفی و کمی گسسته بکار می آید.
- روی محور عمودی تقسیم بندی را بر اساس فراوانی مطلق یا فراوانی نسبی و یا درصد فراوانی انجام می دهیم.



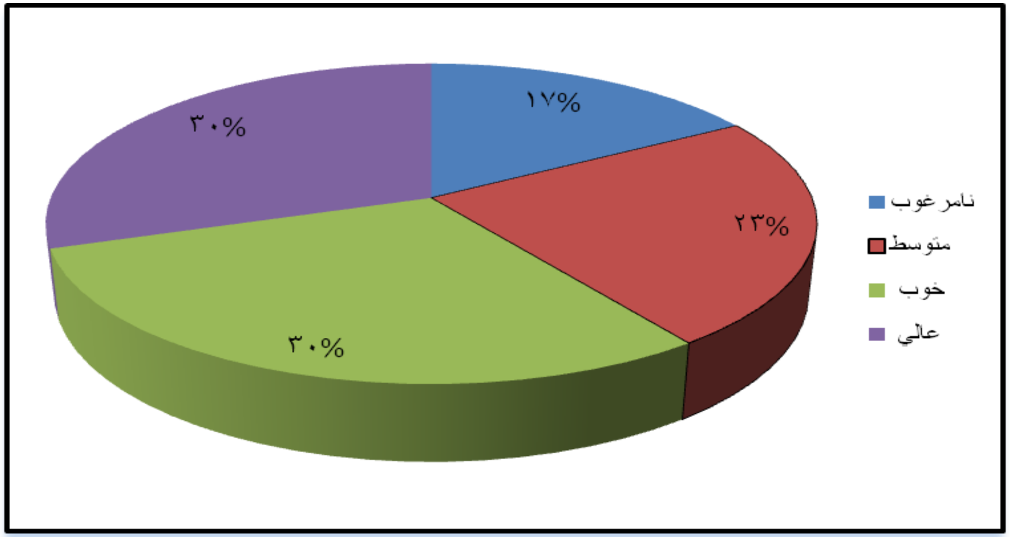
آمار توصیفی: نمودار هیستوگرام (بافت نگار)

- برای متغیرهای کمی پیوسته مناسب است.
- رسم آن شبیه نمودار ستونی است.
- از ستونها یا مستطیلهای به هم چسبیده تشکیل شده است.
- بر روی محور افقی سطوح متغیر و بر محور عمودی فراوانی مطلق یا نسبی.
- کل مساحت مستطیلهای به هم چسبیده برابر با فراوانی کل .
- مستطیلهای این نمودار به هم چسبیده می‌باشند.



آمار توصیفی: نمودار دایره ای (PIE CHART)

کیفیت	f_i	r_i	$D_i=r_i\times$۳۶۰
نامرغوب(۱)	۵	۰/۱۷	۰/۱۷*۳۶۰=۶۰
متوسط(۲)	۷	۰/۲۳	۸۴
خوب (۳)	۹	۰/۳۰	۱۰۸
عالی (۴)	۹	۰/۳۰	۱۰۸
جمع	۳۰	۱	۱۰۰



$$D_i = \frac{f_i}{n} \times 360^o = r_i \times 360^o$$

آمار توصیفی: نمودار ساقه و برگ (Stem & Leaf)

- شبیه نمودار هیستوگرام است.
- اطلاعات بیشتری در مورد داده ها به ما می دهد.
- هر مشاهده به دو قسمت ساقه و قسمت برگ تقسیم می شود. بطور مثال عد ۵۶ به دو قسمت ساقه(۵) و قسمت برگ(۶) تقسیم می شود.

آمار توصیفی: نمودار ساقه و برگ (Stem & Leaf)

مثال: در زیر نمودار ساقه و برگ مربوط به تعداد مراجعات ۲۰ نفر از مادران کودکان زیر ۵ سال به مرکز بهداشتی محل سکونت در طول یکسال می باشد.

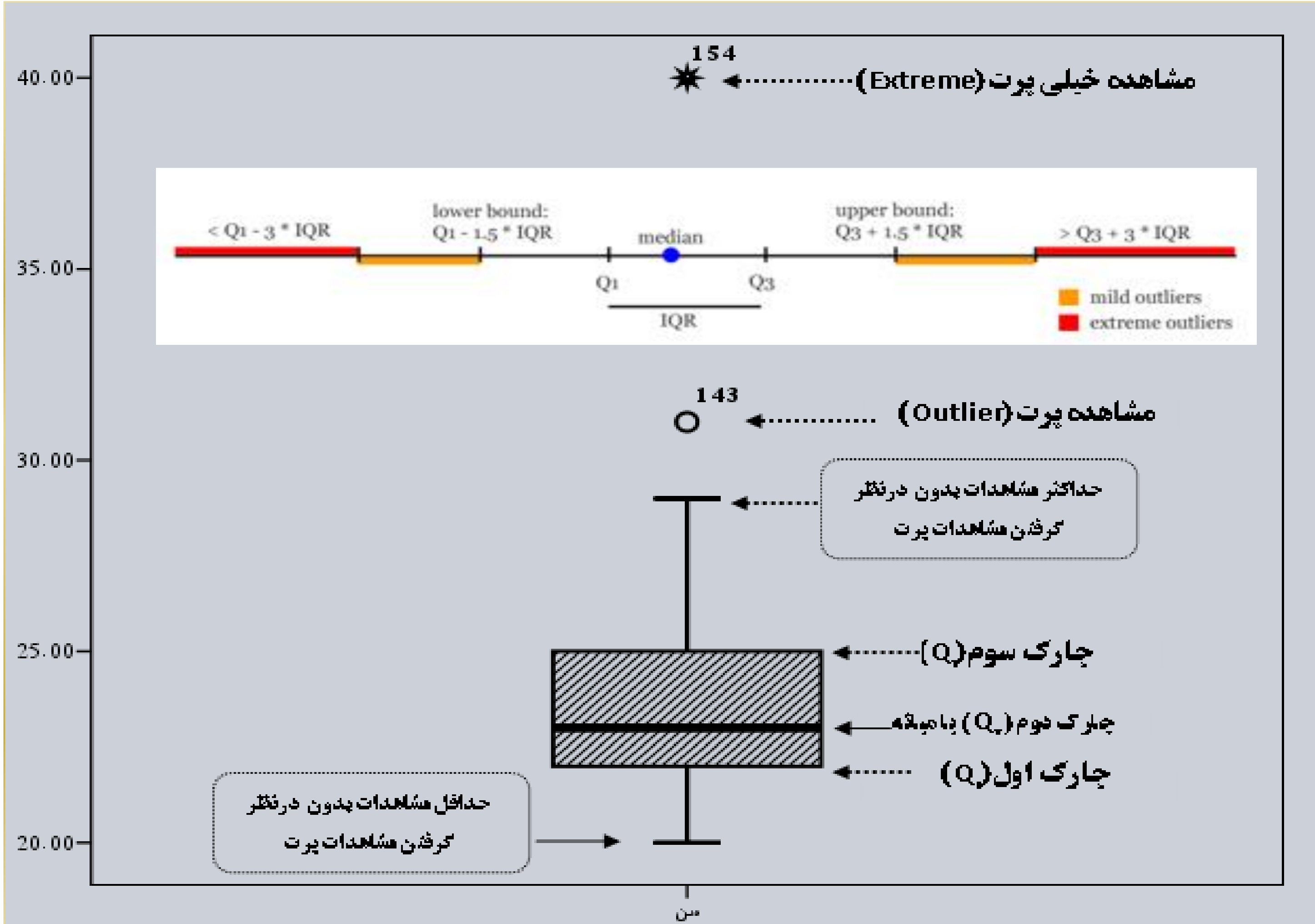
-۳۸-۵ - ۱۵ - ۱۳ - ۱۳- ۱۹- ۲ - ۴ – ۲۲ – ۹ – ۰
۳۳–۴– ۲۵ – ۳۲ – ۱۴– ۱۶–۱۰– ۲ –۷ –۲۶ –۲۰

برگ	ساقه	فراوانی
0, 2, 2, 4, 4 , 5 , 7 , 9	0	8
0 , 3 , 3 , 4 , 5 , 6 ,9	1	7
0 , 2 , 5 , 6	2	4
2,3,8	3	1

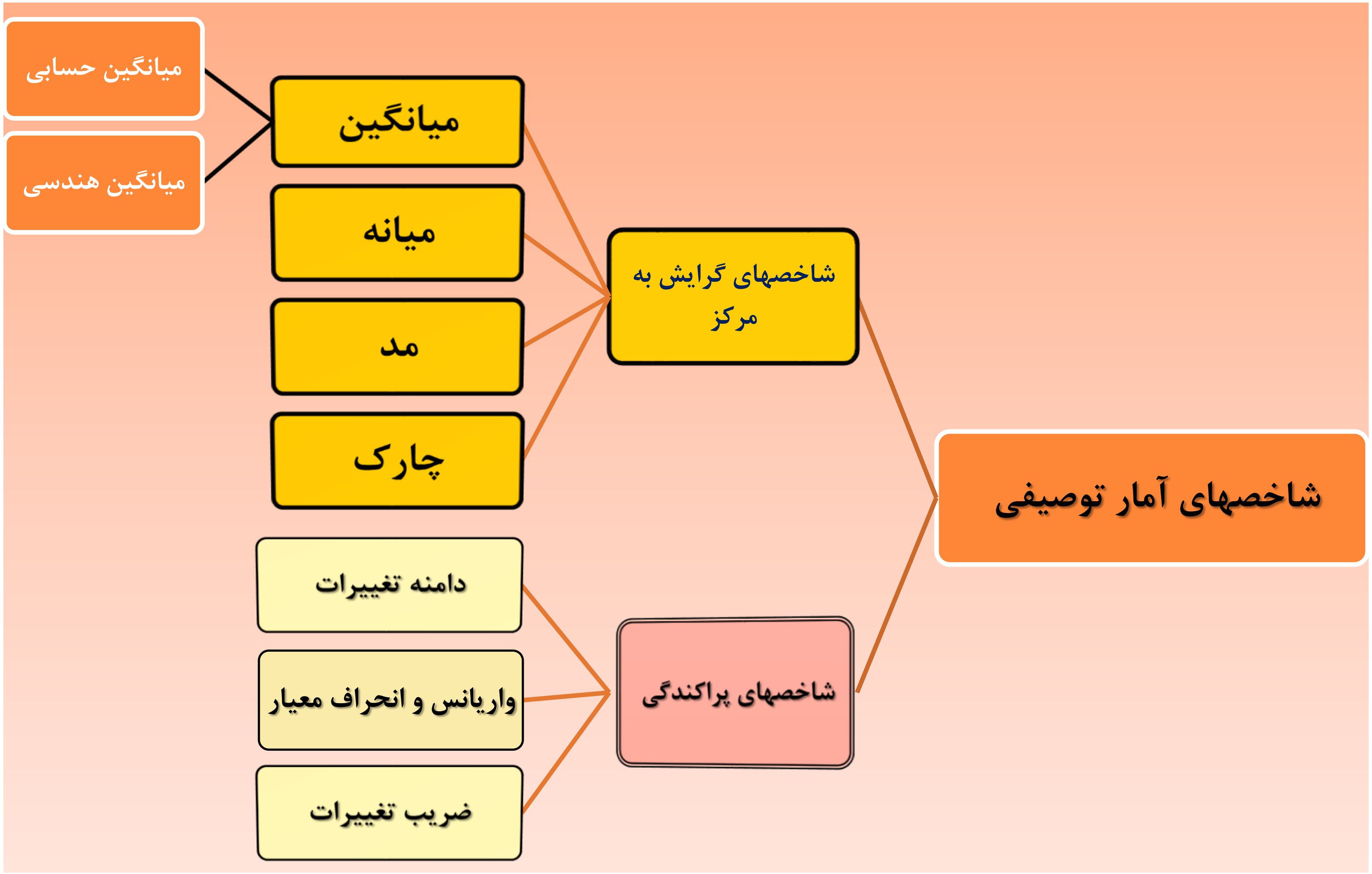
آمار توصیفی: نمودار جعبه ای (Box-Plot)

- روشی مناسب برای تشخیص نوع توزیع مشاهدات.
- برای مقایسه توزیع مشاهدات در دو یا چند گروه.(تمرکز وپراکندگی)
- عموماً قبل از بکارگیری آزمون آنالیز واریانس (**ANOVA**) ترسیم آن مناسب است.
- شاخصهای مناسبی مانند کمترین مشاهده، بیشترین مشاهده، مشاهدات پرت و خیلی پرت، میانه، چارک اول، دوم و سوم و پراکندگی مشاهدات را به ما می دهد.

آمار توصیفی: نمودار جعبه ای (Box-Plot)



آمار توصیفی: شاخصهای گرایش مرکزی و پراکندگی مشاهدات



آمار توصیفی: شاخص های گرایش مرکزی – میانگین حسابی

آنرا با $\bar{X}$ نشان می‌دهند و به آن معدل نیز می‌گویند و روش محاسبه آن بدین صورت می‌باشد که اگر **n** متغیر $x_1, x_2, .., x_n$ داشته باشیم، میانگین حسابی این **n** مشاهدات بصورت زیر محاسبه می شود.

جامعه:

$$\mu = \frac{\sum_{i=1}^{n} X_i}{N}$$

نمونه:

$$\bar{X} = \frac{x_1 + x_2 + \ldots + x_n}{n} = \frac{\sum_{i=1}^{n} X_i}{n}$$

مثال: نمونه ای به اندازه ۱۰ نفر از افراد یک جامعه را انتخاب و قند خون ناشتای آنها را اندازه گیری نموده ایم، اطلاعات بدست آمده بصورت زیر است.

۷۸-۶۶-۹۴-۹۰-۱۰۰-۸۷-۵۹-۱۱۰-۶۷-۸۰

$$\bar{X} = \frac{\sum_{i=1}^{n} x_i}{n} \frac{x_1 + x_2 + \ldots + x_{10}}{10} = \frac{78 + 66 + \ldots + 79}{10} = \frac{830}{10} = 83$$

آمار توصیفی: شاخص های گرایش مرکزی – میانگین حسابی

حدود طبقات	f_i	نماینده طبقه x_i	$f_i \times x_i$
۲/۱۹–۴/۱۹	۱	۳/۱۹	۱×۳/۱۹=۳/۱۹
۵/۱۹–۷/۱۹	۲	۶/۱۹	۲×۶/۱۹=۲/۳۹
۸/۱۹–۰/۲۰	۸	۹/۱۹	۸×۹/۱۹=۲/۱۵۹
۱/۲۰–۳/۲۰	۴	۲/۲۰	۴×۲/۲۰=۸/۸۰
۴/۲۰–۶/۲۰	۳	۵/۲۰	۳×۵/۲۰=۵/۶۱
۷/۲۰–۹/۲۰	۲	۸/۲۰	۲×۸/۲۰=۶/۴۱
جمع	۲۰		۶/۴۰۱

آمار توصیفی: شاخص های گرایش مرکزی – میانگین هندسی

معیار مرکزی مناسب برای داده هایی از نوع درصد، نسبت، نرخ، شاخص ها و غیره است.
محاسبه میانگین هندسی، برای n مشاهده $x_1,x_2,..,x_n$ که همگی مثبت باشند، از رابطه زیر استفاده می شود.

$$\bar{X}_g = \sqrt[n]{X_1.X_2.\dots.X_n}$$

مثال: نرخ رشد جمعیت در سه استان به ترتیب ۲، ۳ و ۱/۵می باشد.
میانگین میزان رشد جمعیت را برای این سه استان بدست آورید.

$$\bar{X}_g = \sqrt[3]{X_1.X_2.X_3} = \sqrt[3]{2\times 3\times 1/5} = \sqrt[3]{9} = 2.8$$

آمار توصیفی: شاخص های گرایش مرکزی – میانه

شاخصی است که ۵۰٪ از مشاهدات از آن کوچکتر و ۵۰٪ دیگر از آن بزرگتر می‌باشند. میانه را با **Md** نمایش می‌دهند و برای بدست آوردن میانه داده‌ها مراحل زیر را باید رعایت نمود.

۱– ابتدا داده‌ها را بترتیب صعودی (از کوچک به بزرگ) مرتب کرد.

۲– با در نظر داشتن اینکه تعداد داده‌ها زوج یا فرد هستند، روشهای زیر را درنظر می‌گیریم.

$$Md = X_{\frac{n+1}{2}} \qquad , when \ \mathbf{n} \ is \ odd$$

$$Md = \frac{X_{n/2} + X_{n/2+1}}{2} \qquad , when \ \mathbf{n} \ is \ even$$

آمار توصیفی: شاخص های گرایش مرکزی - میانه

مثال: وزن ۵ نوزاد به شرح زیر است، مطلوب است میانه مشاهدات.

۳/۲ ,۳/۰۷, ,۳/۸۴ ۲/۷۵ , ۳/۸۲

-در ابتدا داده‌ها را به ترتیب صعودی مرتب می‌کنیم.

۲/۷۵ , ۳/۰۷ , ۳/۲, ۳/۸۲, ۳/۸۴

$$Md = X_3 = 3.2$$

مشاهده شماره سوم یعنی **۳/۲** میانه مشاهدات است.

مثال: سن ۶ نفر مبتلا به آسم به شرح زیر است، مطلوب است میانه مشاهدات.

۱ ,۴, ۸, ۹ , ۱۱, ۱۵

-در ابتدا داده‌ها را به ترتیب صعودی مرتب می‌کنیم.

۱ , ۴, ۸, ۹ , ۱۱, ۱۵

$$Md = \frac{X_3 + X_4}{2} = \frac{8+9}{2} = \frac{17}{2} = 8.5$$

آمار توصیفی: شاخص های گرایش مرکزی - مد

- نما ترجمه کلمه مد است، که یک لغت فرانسوی و به معنای " متداول ترین" است.
- نما، برای مجموعه ای از داده‌ها عبارتست از اندازه ای که بیشترین فراوانی را دارا می باشد.

مثال(۱): نمره آپگار ۱۰ نوزاد متولد شده به ترتیب زیر است، مد را برای این مشاهدات بدست آورید.

۹ , ۷ ,۸, ۹, ۸ , ۹ , ۹, ۹ , ۹, ۹,

mo=9

مثال(۲): برای داده های ۵،۳،۳،۴،۱،۲،۳،۱،۱

mo=1, 3

آمار توصیفی: شاخص های گرایش مرکزی – مزایا و معایب

شاخص	مزایا	معایب
میانگین	-در محاسبه آن از همه داده ها استفاده می شود. -براوردی نااریب برای میانگین جامعه است. -از میانگین برای محاسبه معیارهای پراکندگی استفاده می شود. -میانگین برای مجموعه ای از داده ها یکتا است. -مجموع انحرافات از میانگین برابر با صفر است. $\sum(x_i - \bar{x}) = 0$ -مجموع مجذورات انحرافات از آن از مجموع مجذورات انحرافات از هرعددی کوچکتر است.	-اندازه میانگین تحت تأثیر مقادیر بسیار بزرگ و یا مقادیر بسیار کوچک قرار می گیرد و به همین دلیل برای داده های خیلی چوله مناسب نیست. -میانگین را نمی توان برای یک جدول توزیع فراوانی که طبقه اول و طبقه آخر آن محدود نمی باشند محاسبه نمود. -مقدار میانگین نسبت به دو معیار دیگر در نمونه گیری های متفاوت از یک جمعیت کمتر تغییر می‌یابد.
میانه	-میانه نیز مانند میانگین منحصر به فرد است(یکتاست) -میانه معیار مرکزی مناسبی برای داده های جدول توزیع فراوانی است که طبقه اول و طبقه آخر آن محدود نیست، می باشد. -میانه زمانی محاسبه می گردد که نیاز به شناخت ارزش میانی داده ها باشد. -میانه کمتر تحت تأثیر مقادیر بسیار کوچک و بسیار بزرگ قرار می گیرد. مجموع قدرمطلق انحرافات از میانه حداقل است.	خیلی از مشاهدات را نادیده می گیرد. تعریف جبری ندارد. توزیع نمونه گیری آن پیچیده است.
مد	نما را می توان به عنوان یک معیار مرکزی برای داده های کیفی نیز به کار برد. برای تعیین متداول ترین اندازه داده ها از معیار نما استفاده می شود. محاسبه معیار مرکزی نما از سایر معیارها ساده تر است.	اکثر داده ها را نادیده می گیرد. تعریف جبری ندارد. ممکن است برای مجموعه ای از داده ها نما وجود نداشته باشد و یا بیش از یک نما موجود باشد.

آمار توصیفی: شاخص های گرایش مرکزی - چارکها

چارکها مقادیری از صفات هستند که کل مشاهدات را به چهار قسمت مساوی تقسیم می‌کنند.

چارک اول : Q_1

چارک اول : Q_2

چارک اول : Q_3

آمار توصیفی: شاخص های پراکندگی- دامنه تغییرات

دامنه تغییرات یا برد (Range)

ساده ترین اندازه پراکندگی برای یک مجموعه داده، برد آنها است که آن را با R نشان داده و به صورت زیر تعریف می شود:

$$\boxed{R = X_{\max} - X_{\min}}$$

مثال: دامنه تغییرات داده‌های 63 ,97 ,9 ,16 ,104 ,16 ,15 ,7 برابر است با:

$$\boxed{R = X_{\max} - X_{\min} = 104 - 7 = 97}$$

آمار توصیفی: شاخص های پراکندگی- واریانس، انحراف معیار

واریانس(Variance)

میانگین توان دوم انحرافات از میانگین، واریانس نام دارد وآنرا V(x) یا σ^2 نشان می دهند و برای داده‌های که بصورت سری داده‌ها هستند و داده‌های که بصورت جدول توزیع فراوانی هستند روشهای محاسباتی متفاوت دارد.

جامعه :

$$\sigma^2 = \frac{\sum_{i=1}^{N} {x_i}^2}{N} - \mu$$

$$\sigma^2 = \frac{\sum_{i=1}^{k} f_i (x_i - \mu)^2}{N}$$

نمونه:

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

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

انحراف معیار (Standard Deviation)

جذر واریانس

آمار توصیفی: شاخص های پراکندگی- ضریب تغییرات

اگربخواهیم تغییرات دو صفت با واحد های اندازه گیری متفاوت یا تغییرات یک صفت با دو واحد اندازه‌گیری را مورد بررسی قرار دهیم.

مثال: اگر در یک بیمارستان زنان و زایمان قصد بررسی توزیع پراکندگی میزان هماتوکریت و میزان هموگلوبین خون زنان باردار را داشته باشیم.

$$C.V = \frac{\sigma}{\overline{\overline{X}}} \times 100$$

آمار توصیفی: شاخص های پراکندگی- ضریب تغییرات

مثال: اگر میانگین وانحراف معیار سن اولین زایمان زنان یک منطقه به ترتیب ۲۵و ۵ سال باشد و میانگین وانحراف معیار قند خون همین افراد به ترتیب برابر با ۱۰۰ و ۱۰ باشد، میزان پراکندگی کدام صفت این افراد بیشتر است؟

- ضریب تغییرات برای سن اولین زایمان:

$$C.V_1 = \frac{\sigma_1}{\overline{X}_1} \times 100 = \frac{5}{25} \times 100 = \%20$$

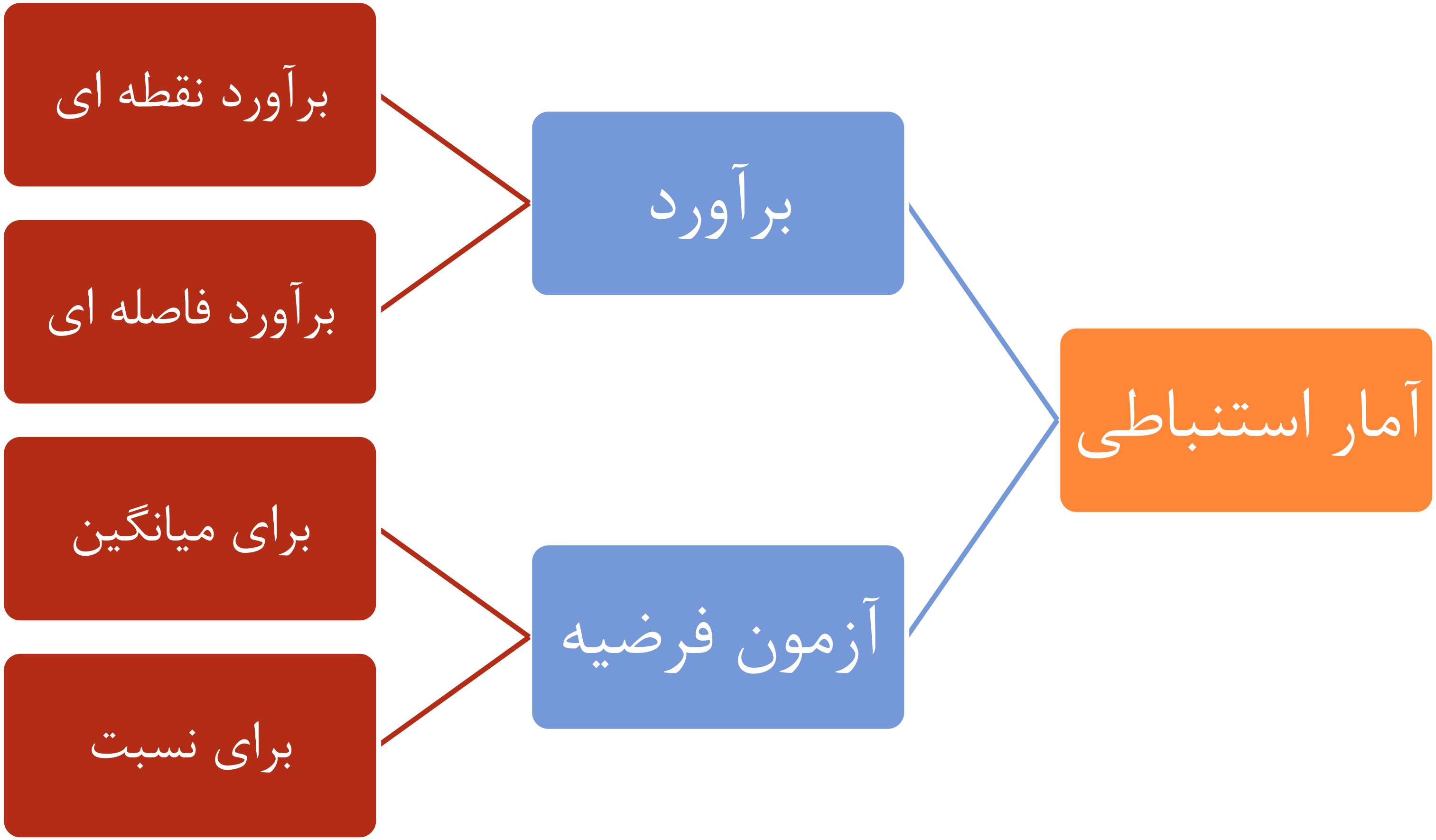
- ضریب تغییرات برای قند خون:

$$C.V_2 = \frac{\sigma_2}{\overline{X}_2} \times 100 = \frac{10}{100} \times 100 = \%10$$

آمار توصیفی: مقیاس داده ها و تحلیل توصیفی مناسب آنها

مقیاس سنجش	رده بندی داده ها	نمایش گرافیکی داده ها	تلخیص داده ها
اسمی	جدول توزیع فراوانی	نمودار میله ای و دایره ای	نما
رتبه ای	جدول توزیع فراوانی	نمودار میله ای، دایره ای و ستونی	نما، میانه
فاصله ای یا نسبتی	اول طبقه بندی و بعد جدول توزیع فراوانی	نمودار هیستوگرام، جعبه ای، ساقه و برگ و...	میانگین، میانه، واریانس، انحراف معیار، و...

آمار استنباطی



آمار استنباطی: برآورد نقطه ای

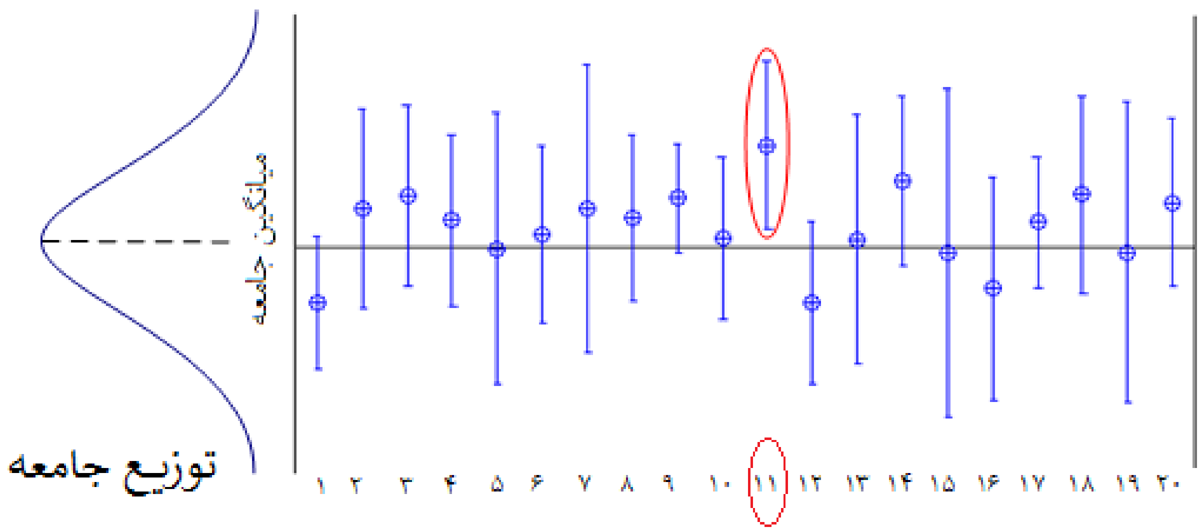
- برآورد میانگین در جامعه : میانگین در نمونه
- برآورد واریانس در جامعه : واریانس در نمونه
- برآورد نسبت (شیوع) در جامعه : نسبت در نمونه

آمار استنباطی: برآورد فاصله ای

فرض کنید قد ۴۰ نفر که به طور تصادفی انتخاب شده‌اند، اندازه‌گیری شده است. میانگین قد این افراد برابر است با ۱۷۵ سانتی‌متر است. از طرفی می‌دانیم که انحراف معیار قد در جامعه آماری ما برابر است با ۲۰ سانتی‌متر. با توجه به این موضوع، فاصله اطمینان برای میانگین قد افراد جامعه با سطح اطمینان ۹۵ درصدی برابر است با

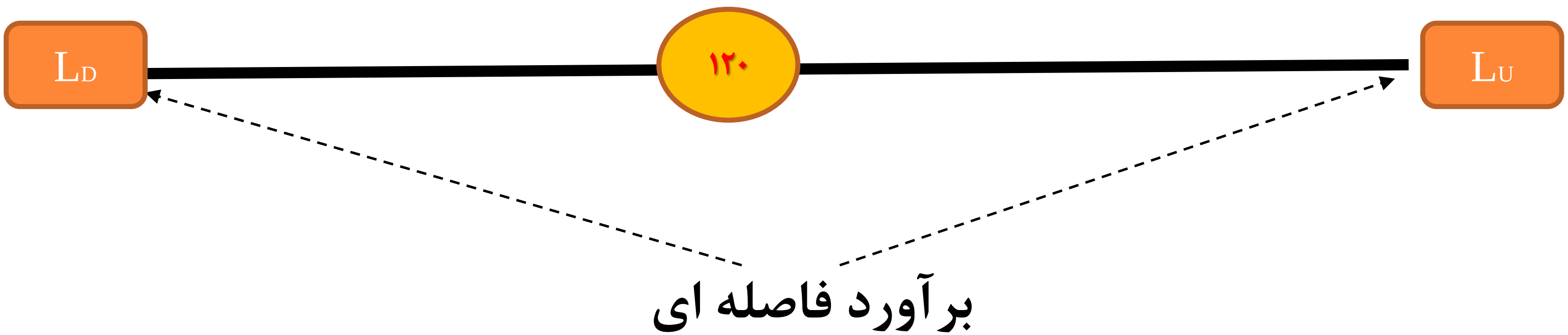
175cm±6.2cm

فاصله اطمینان بیان می‌کند که اگر ۱۰۰ بار نمونه‌گیری تکرار شود و ۱۰۰ فاصله اطمینان ۹۵٪ تولید شود، ۹۵ فاصله شامل پارامتر جامعه خواهند بود و فقط ۵ تا از این فاصله‌ها شامل میانگین جامعه نمی‌شوند.



آمار استنباطی: برآورد فاصله ای برای میانگین

- **فاصله اطمینان برای میانگین:**



$$\overline{X} \mp Z_{1-\alpha/2} SE_{(x)}$$

$$SE(x)=S/\sqrt{n}$$

آمار استنباطی: برآورد فاصله ای برای میانگین

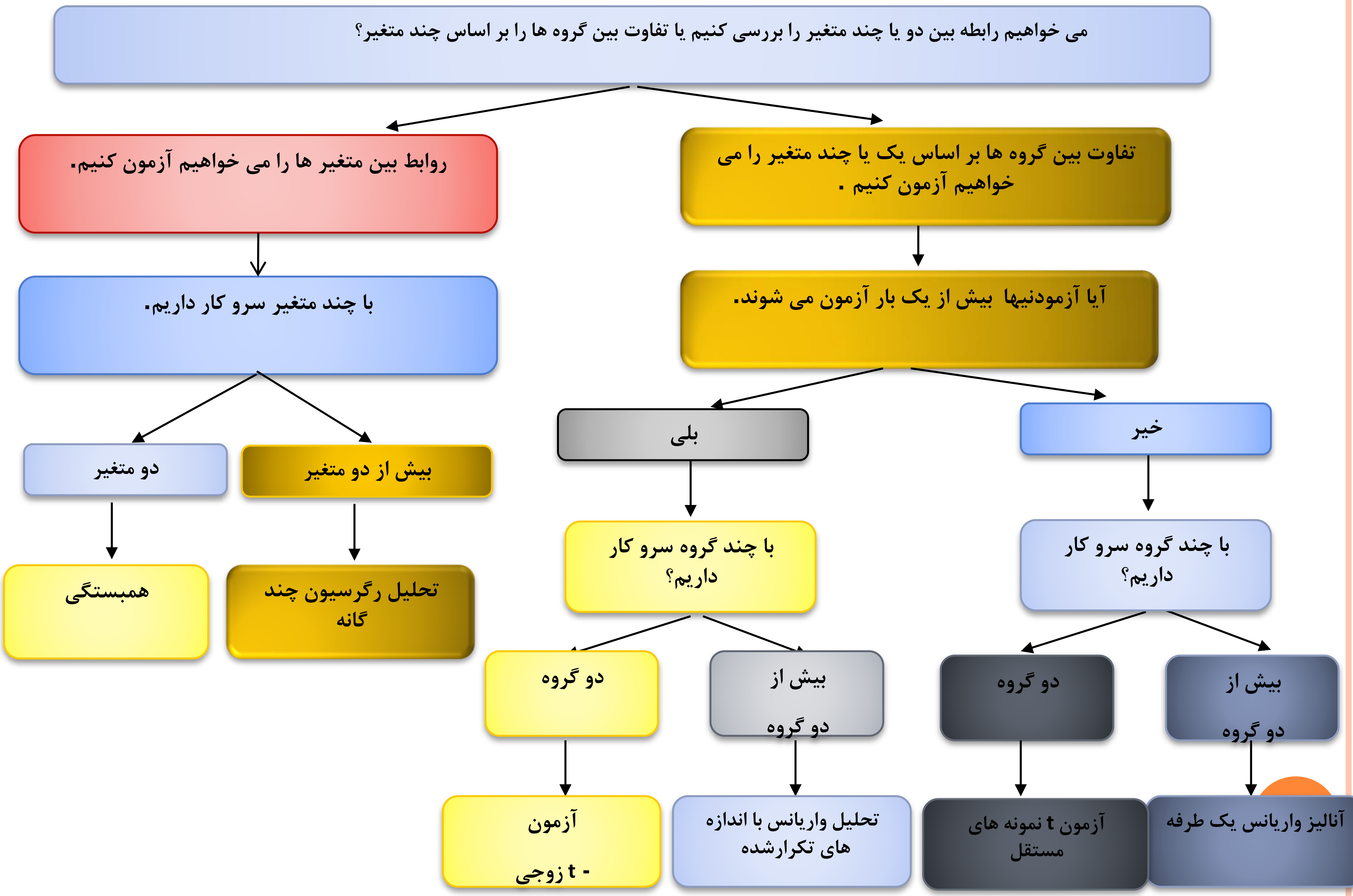
مثال: بمنظور برآورد وزن نوزادان در حین تولد نمونه ای به اندازه ۹۰۰ پرونده از بین نوزادان متولد شده در یک زایشگاه انتخاب کرده ایم. میانگین و انحراف معیار این نمونه به ترتیب ۳۰۸۷ گرم و ۳۹۷ گرم بدست آمده است. درسطح ۹۵٪ اطمینان یک برآورد فاصله ای برای میانگین وزن نوزادان در حین تولد جامعه بدست آورید.

$$S_{\bar{x}} = \frac{s_x}{\sqrt{n}} = \frac{397}{\sqrt{900}} = \frac{397}{30} = 13.2$$

$$C = (1-\alpha) = 0.95 \Rightarrow \alpha = 1-0.95 \Rightarrow \frac{\alpha}{2} = \frac{1-0.95}{2} = 0.025 \Rightarrow Z_{\frac{\alpha}{2}} = 1.96$$

$$\mu : \bar{X} \pm Z_{\frac{\alpha}{2}} \times \frac{s_x}{\sqrt{n}} \Rightarrow \begin{cases} \bar{L} = \bar{X} + Z_{\frac{\alpha}{2}} \times \frac{s}{\sqrt{n}} = 3087 + (1.96)(13.2) = 3087 + 25/87 = 3112.87 \\ \underline{L} = \bar{X} - Z_{\frac{\alpha}{2}} \times \frac{s}{\sqrt{n}} = 3087 - (1.96)(13.2) = 3087 - 25.87 = 3061.30 \end{cases}$$

آمار استنباطی: چه آزمونی ؟



آزمون فرضیه: مقایسه میانگین یک جامعه با عدد ثابت:
آزمون T تک نمونه ای

موارد کاربرد: وقتی بخواهیم بررسی کنیم که آیا میانگین یک نمونه برابر با یک مقدار خاص(مثلا میانگین فشار خون برابر با ۱۲۰) می باشد یا نه از آن استفاده می شود.

مفروضات: نرمال بودن متغیر مورد مطالعه

مثال: محققی ادعا می کند، میانگین هموگلوبین خون مردان ساکن درمناطق کوهستانی بیشتر از ۱۵ است. به منظور بررسی این گفته نمونه ای ۱۳ نفری از بین افراد این مناطق انتخاب نموده ایم. ادعای محقق را ازمون کنید.

فرضیه صفر: میانگین هموگلوبین خون مردم این منطقه ۱۵ است.
فرضیه تحقیق(ادعا): میانگین هموگلوبین خون مردم این منطقه بیشتر از ۱۵ است.

شماره	۱	۲	۳	۴	۵	۶	۷	۸	۹	۱۰	۱۱	۱۲	۱۳
نمره	۱۵	۱۱.۷۵	۱۴	۱۸	۱۸.۲۵	۱۳	۱۰	۸	۱۳	۱۶	۱۴.۵	۱۶.۲۵	۱۶.۷۵

آزمون فرضیه: مقایسه میانگین دو جامعه:
آزمون T دو نمونه مستقل

- **هدف**: **این آزمون برای موقعی بکار می آید که بخواهیم میانگین یک صفت یا ویژگی کمی (مثلاً میزان قند خون) را برای دو گروه مستقل از هم (مانند دو گروه دریافت کنند داروی A و B)مقایسه کنیم.**
- **شرایط**:

۱- متغیر مورد مطالعه کمی باشد (مقیاس نسبتی یا فاصله ای)

۲- توزیع متغیر کمی نرمال باشد.

۳- اگر واریانس گروه ها مساوی نباشد باز می توان آزمون را اجرا نمود. (اگر واریانس برابر باشد بهتر است)

$$\begin{cases} H_o : \mu_1 = \mu_2 \\ H_1 : \mu_1 \neq \mu_2 \end{cases}$$

میانگین پلاکت در دو گروه بیمار یکسان است

میانگین پلاکت در دو گروه بیمار متفاوت است

$$tc = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{S^2{}_p}{n_1} + \frac{S^2_p}{n_2}}}$$

$$S^2_p = \frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1 + n_2 - 2}$$

آزمون فرضیه: مقایسه میانگین قبل و بعد:
آزمون T زوجی

فرضیه تحقیق: مصرف دارو منجر به کاهش تعداد نبض بیمار می شود.(دارو موثر است)

شماره افراد	تعداد نبض (قبل از مصرف دارو)	تعداد نبض (بعد از مصرف دارو)	Di=Xi-Yi	$(Di)^2$
1	66	58	-8	64
2	69	65	-4	16
3	75	68	-7	49
4	68	70	2	4
5	73	66	-7	49
6	75	75	0	0
7	68	62	-6	36
8	69	72	3	9
جمع	563	536	-27	227
میانگین	70.375	67		

$$\overline{D} = \frac{\sum_{i=1}^{8} D_i}{8} = \frac{-27}{8} = -3/375$$

$$S_D = \sqrt{\frac{\sum_{i=1}^{n} D_i^2 - \frac{(\sum_{i=1}^{n} D_i)^2}{n}}{n-1}} = \sqrt{\frac{227 - \frac{(-27)^2}{8}}{8-1}} \cong 4/4$$

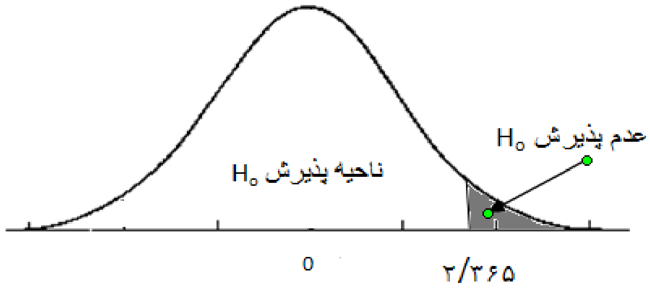
آزمون فرضیه: مقایسه میانگین قبل و بعد:
آزمون T زوجی

$$\begin{cases} H_o : \mu_B = \mu_A \\ H_1 : \mu_B < \mu_A \end{cases}$$

میانگین نبض بیماران قبل از و بعد از مصرف قرص یکسان است

بعد از مصرف دارو میانگین نبض بیماران کاهش می یابد

$$t_c = \frac{|\overline{D}|}{\frac{S_D}{\sqrt{n}}} = \frac{|-3.375|}{\frac{4.4}{\sqrt{8}}} = \frac{3.375}{1.55} = 2.17$$



$$\left| \begin{array}{l} 1-\alpha = 0.95 \Rightarrow \alpha = 0.05 \\ \\ \quad df = n-1 = 8-1 = 7 \end{array} \right\rangle t_{\alpha,df} = t_{0/05,7} = 2.365$$

آزمون فرضیه: مقایسه میانگین در چند نوبت:
آزمون Repeated measures

هدف: این آزمون برای موقعی بکار می آید که بخواهیم میانگین یک صفت (متغیر کمی) را که برای یک گروه افراد در سه بار یا بیشتر اندازه گیری شده، با هم مقایسه کنیم.

شرایط:
۱-متغیر مورد مطالعه کمی باشد(مقیاس فاصله ای یا نسبتی) و دارای توزیع نرمال باشد.
۲-افرادی که برای روشها (درمانی) برروی آنها اعمال می شود یک گروه افراد ثابت باشند(گروه ها به هم وابسته باشند) .
۳-اگر آزمون ماخلی(**Mauchly's Test**) معنادار نشود یعنی بین کوواریانس داده ها تفاوتی وجود ندارد و می توان از F استفاده کرد در غیر این صورت باید از یکی از آزمونهای دیگر استفاده کرد.

مثال: محققی معتقد است که فشار خون افراد در سه زمان هنگام بیدار شدن از خواب(صبح)-در طول روز(ساعت ۱۱ صبح) و هنگام خواب متفاوت است.
به منظور بررسی گفته این محقق، نمونه ای به اندازه ۲۰ نفر از افراد جامعه انتخاب و فشار خون آنها را در سه زمان مذکور اندازه گیری می کنیم.

آزمون فرضیه: مقایسه میانگین در چند گروه مستقل:
آزمون آنالیز واریانس یکطرفه

هدف: وقتی بخواهیم میانگین سه گروه مستقل(نمرات درس آمار زیستی سه گروه دانشجویان رشته های مامایی- پزشکی- بهداشت) را با هم مقایسه کنیم از این آزمون استفاده می کنیم.

این آزمون تعمیم یافته آزمون t-test می باشد.

شرایط:

۱-متغیر مورد مطالعه که می خواهیم آن را در گروه ها مقایسه کنیم کمی (مقیاس فاصله ای یا نسبتی) باشد.

۲-مشاهدات گروه ها مستقل از هم باشد.

۳-توزیع متغیر کمی نرمال باشد.

۴- واریانس گروه ها یکسان باشد. (خیلی ضروری نیست)

آزمون فرضیه: مقایسه میانگین در چند گروه مستقل: آزمون آنالیز واریانس یکطرفه

- **مثال**: ادعا شده است که میانگین مدت زمان بقای بیماران لوسمی حاد در گروه های خونی مختلف با هم متفاوت است. برای این منظور نمونه ای به اندازه ۲۰۰ نفر بیمار را انتخاب و این موضوع را مورد مطالعه قرار داده ایم.

فرضیه صفر: میانگین زمان بقای بیماران در ۴ گروه خونی یکسان است.
فرضیه تحقیق(ادعا) : میانگین زمان بقای حداقل دو گروه خونی با هم متفاوت است.

آزمون فرضیه: رابطه دو متغیر کمی:
همبستگی پیرسون

هدف: وقتی که بخواهیم شدت و جهت ارتباط بین دو متغیرکمی را مورد بررسی قرار دهیم.
یک متغیر، متغیر وابسته و متغیر دیگر، متغیر مستقل.

شرایط و ویژگی ها:

۱- متغیر مستقل و متغیر وابسته کمی(مقیاس نسبتی یا فاصله ای) باشند.

۲-ضریب همبستگی واحد اندازه گیری ندارد.

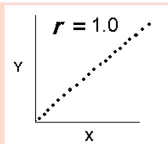
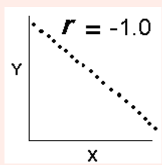
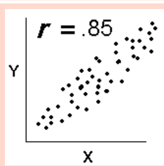
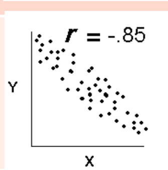
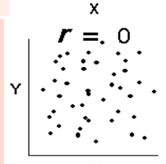
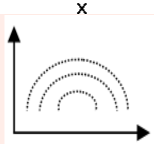
فرضیه صفر: **بین دو متغیر رابطه وجود ندارد.(r=0)**

فرضیه مقابل(تحقیق): **بین دو متغیر رابطه وجود دارد(r≠0)**

$$r=\frac{n\sum_{i=1}^{n}x_iy_i-(\sum_{i=1}^{n}x_i)(\sum_{i=1}^{n}y_i)}{\sqrt{[n\sum_{i=1}^{n}x_i^2-(\sum_{i=1}^{n}x_i)^2][n\sum_{i=1}^{n}y_i^2-(\sum_{i=1}^{n}y_i)^2]}}$$

آزمون فرضیه: رابطه دو متغیر کمی:

همبستگی پیرسون

وضعیت همبستگی	تفسیر	منحنی
$r=+1$	همبستگی کامل ومستقیم	
$r=-1$	همبستگی کامل ومعکوس	
$0<r<+1$	همبستگی ناقص ,و مستقیم	
$-1<r<0$	همبستگی ناقص ومعکوس	
$r=0$	همبستگی وجود ندارد	
$r=0$	همبستگی خطی وجود ندارد (همبستگی درجه دوم)	

آزمون فرضیه: تشخیص استفاده از نوع همبستگی ها

متغیر مستقل	متغیر وابسته	نوع همبستگی	مثال
کمی (فاصله ای یا نسبتی)	کمی (فاصله ای یا نسبتی)	پیرسون(r)	همبستگی وزن و قد
(رتبه ای)	رتبه ای	اسپیرمن(رو)	رتبه قبولی و معدل دیپلم
اسمی یا رتبه ای (با سطوح کم)	اسمی یا رتبه ای (با سطوح کم)	تاوی کندال-فی-کرامر-سامرز-خطر نسبی – odds ratio	سیگاری بودن و بیماری قلبی

رگرسیون خطی ساده: رابطه خطی بین دو متغیر کمی

هدف: پیش بینی مقادیر متغیر وابسته ($\mathbf{Y}$)به ازای مقادیر معلوم متغیر مستقل ($\mathbf{x}$).

شرایط:

۱- متغیر مستقل و متغیر وابسته کمی باشند.

۲- متغیر مستقل و وابسته دارای مقیاسهای فاصله ای و نسبتی باشند.

۳- هر دو متغیر یا حداقل متغیر وابسته دارای توزیع نرمال باشد.

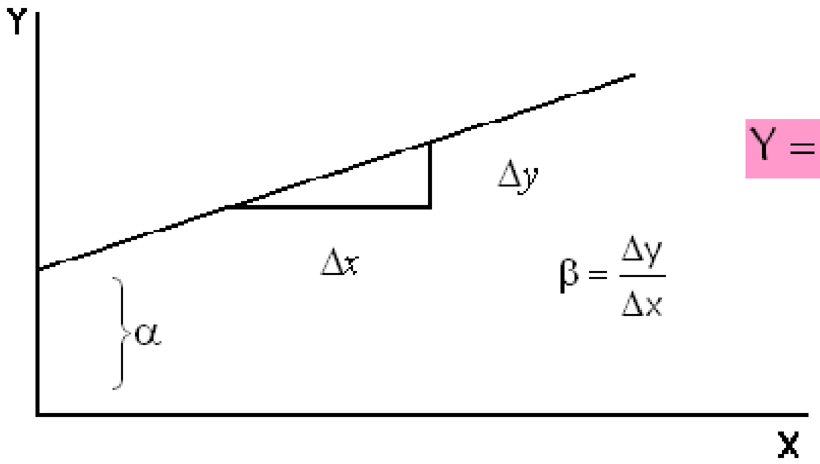
۴- مشاهدات پرت (بخصوص خیلی پرت) وجود نداشته باشد.

۵- قبل از اجرای رگرسیون خطی بهتر است نمودار پراکنش ترسیم شود.

مدل رگرسیون خطی ساده:

$$\boxed{Y = \alpha + \beta\, x + e_i}$$

رگرسیون خطی ساده: رابطه خطی بین دو متغیر کمی



$$Y = \alpha + \beta X$$

$$\hat{\beta} = \frac{\sum XY - n\bar{X}\bar{Y}}{\sum X^2 - n\bar{X}^2}$$

$$\hat{\alpha} = \bar{Y} - \hat{\beta}\bar{X}$$

آزمون فرضیه: رابطه دو متغیر کیفی:
آزمون کای دو

هدف: بررسی رابطه بین دو متغیر کیفی (اسمی یا رتبه ای)

شرایط:

۱-تعداد نمونه مناسب باشد.(بالای ۴۰)

۲-خانه ای از جدول فراوانی صفر نداشته باشد.

۳-فراوانی مورد انتظار بیش از ۲۰٪ خانه های جدول زیر ۵ نباشد.

- **حالات خاص:**

۱-اگر نمونه بین ۲۰ تا ۴۰ باشد و فراوانی مورد انتظار بیش از ۲۰٪ کمتر از ۵ باشد، تصحیح یتس یا آزمون دقیق فیشر (جداول ۲*۲)را بکار می بریم.

۲-اگرتعداد نمونه کمتر از ۲۰ باشد از آزمون فیشر استفاده می شود.

فرمول:

O_{ij}= فراوانی مشاهده شده

e_{ij}= فراوانی مورد انتظار (تئوریک)

$$\boxed{\chi_c^{\,2} = \sum \sum \frac{(o_{ij} - e_{ij})^2}{e_{ij}}}$$

آزمون فرضیه: رابطه دو متغیر کیفی:
آزمون کای دو- مثال

مثال: در یک مطالعه کارآزمایی بالینی ۴۶۰ نفر را انتخاب و به تصادف به تعداد ۲۴۰ نفر آنها واکسن موثر و به ۲۲۰ نفر آنها واکسن پلاسیبو تزریق نموده ایم پس از مدتی پیگیری، اطلاعات زیر در رابطه با بیماری آنها بدست آمده با درنظر گرفتن خطای ۰/۰۵ بررسی کنید که آیا واکسن بر پیشگیری از بیماری موثر است یا خیر؟

واکسناسیون / آنفولانزا	واکسن موثر	پلاسیبو	جمع
مبتلا شده	20(52.2)	80(47.8)	100
مبتلا نشده	220(187.8)	140(172.2)	360
جمع	240	220	460

فرضیه صفر(H_O): نسبت ابتلاء دربین کودکان واکسن دریافت کرده و پلاسیبو دریافت کرده یکسان است.(**واکسن موثر نیست**)

فرضیه تحقیق(H_1) : نسبت ابتلاء دربین کودکان واکسن دریافت کرده کمتر از پلاسیبو دریافت کرده است. **(واکسن موثر است)**

$$e_{11} = \frac{n_{1.} \times n_{.1}}{n} = \frac{100 \times 240}{460} = 52.2$$

$$e_{12} = \frac{n_{2.} \times n_{.2}}{n} = \frac{100 \times 220}{460} = 47.8$$

$$e_{21} = \frac{n_{20} \times n_{01}}{n} = \frac{240 \times 360}{460} = 187.8$$

$$e_{22} = \frac{n_{20} \times n_{02}}{n} = \frac{220 \times 360}{460} = 172.2$$

آزمون فرضیه: رابطه دو متغیر کیفی:

آزمون کای دو– ادامه مثال

$$\chi_c^{\ 2}=\sum\sum\frac{(o_{ij}-e_{ij})^2}{e_{ij}}=\frac{(20-52.2)^2}{52.2}+\frac{(80-47.8)^2}{47.8}+\frac{(220-187.8)^2}{187.8}+\frac{(140-172.8)^2}{172.8}=53$$

$$\left.\begin{array}{r}\alpha=0.05\\ df=(r-1)(c-1)=(2-1)(2-1)=1\end{array}\right\rangle\Rightarrow\chi_{0/05,1}=3.84$$

نتیجه گیری: با توجه به اینکه مقدار محاسبه شده (53) بزرگتر از مقدار کای دو جدول(3.84) می باشد (نتیجه می گیریم که فرضیه صفر رد می شود و فرضیه مقابل یعنی (نسبت ابتلاء دربین کودکان واکسن دریافت کرده کمتر از پلاسیبو دریافت کرده، می باشد) پذیرفته می شود. به عبارتی واکسن در پیشگیری موثر است.

آزمون فرضیه: رابطه دو متغیر کیفی وابسته:
آزمون مک نمار

- **هدف:** تعیین رابطه بین دو متغیر کیفی (دو وضعیتی) که به هم وابسته هستند.
- **مثال: به منظور بررسی تاثیر سخنرانی یک آموزش دهنده حوزه سلامت، نمونه ای تصادفی به اندازه ۵۰۰ نفر از افراد یک جامعه را انتخاب و از آنها در خصوص اینکه آیا یک رفتار را انجام می دهند یا خیر سوال می شود. پس از سخنرانی این فرد (با فرض اینکه همه افراد نمونه به سخنرانی گوش داده اند) مجدداً درخصوص اینکه آیا یا یک رفتار را انجام می دهند یا خیر سئوال می شود.**

فرضیه صفر: سخنرانی بر تغییر رفتار افراد تاثیر نداشته.

فرضیه تحقیق: سخنرانی بر تغییر رفتار افراد موثر بوده.

قبل \ بعد	بلی	خیر	جمع
بلی	۱۵۰	۵۰	۲۰۰
خیر	۲۱۰	۹۰	۳۰۰
جمع	۳۶۰	۱۴۰	۵۰۰

- آماره آزمون:

$$x^2 = \frac{(b-c)2}{(b+c)}$$

آزمون فرضیه: آزمونهای ناپارامتری

تاریخ آمار ناپارامتری هم مانند آمار پارامتری به اوایل قرن هیجدهم میلادی بر می گردد.

در سال ۱۹۷۰ میلادی مقاله ای منتشر شد و طی آن براساس آمار نوزادان شهر لندن در فاصله ۱۶۲۹ تا ۱۷۱۰ میلادی ادعا شد که مشیت الهی بر این است که تعداد نوزادان پسر بیشتر از تعداد نوزادان دختر باشد.در حقیقت این ادعا یک آزمون ناپارامتری معروف می باشد که امروزه به نام آزمون علامت(**sing test**) شهرت دارد.

با اینحال آمار ناپارامتری (**non parametric statistics**) بیش از دو قرن ناشناخته بود، تا اینکه پیشرفت آن با انتشار دو مقاله پژوهشی ، یکی توسط آماردانی به نام ویلکاکسون (**Wilcoxon**) در سال ۱۹۴۵و دیگری توسط دو آماردان به نام های ((من))**Mann** و ((ویتنی)) **Whitney** آغاز شد.

آزمون فرضیه: چه وقت آزمون ناپارامتری؟

- مشاهدات از توزیع نرمال تبعیت نکند(مفروضات ازمون پارامتری برقرار نباشد).
- اندازه نمونه کوچک باشد (عموماً کمتر از ۱۵).
- مشاهدات بصورت رتبه ای یا اسمی باشند.
- دربرخی موارد که آزمون پارامتریک معنادار نشده برای بررسی معناداری به سراغ این آزمونها هم می آییم.

آزمون فرضیه: آزمون های پارامتری و معادل ناپارامتری آنها

آزمون ناپارامتریک	آزمون پارامتریک	هدف
آزمون دو جمله ای	One sample t-test	مقایسه میانگین یک جامعه با یک مقدار خاص
آزمون من-ویتنی	t-testاستقلال دو گروه	مقایسه دو نمونه مستقل از هم
آزمون ویلکاکسن	آزمون t-زوجی	مقایسه دو نمونه وابسته به هم
آزمون کروسکال- والیس	آزمون آنالیز واریانس یک طرفه	مقایسه بیش از دو نمونه مستقل از هم
آزمون فریدمن	آزمون Repeted Measurment	مقایسه بیش از دو نمونه وابسته به هم
همبستگی اسپیرمن	همبستگی پیرسون	بررسی ارتباط بین دو متغیر


با سپاس از توجه
شما