آمار استنباطی

آمار : جمع آوری و سازماندهی و خلاصه کردن داده ها در قالب عدد

آمار توصیفی: مجموعه ای از روش هایی است که برای سازمان دهی، خلاصه کردن، تهیه جدول، رسم نمودار، توصیف و تفسیر داده های جمع آوری شده از نمونه آماری به کار گرفته می شود

آمار استنباطی : مشخص می کند که آیا الگوها وفرآیندهای کشف شده در نمونه آمار توصیفی ، در جامعه آماری هم کاربرد دارد

آمار استنباطی *Inferential Statistics*

آمار استنباطی

نمونه گیری و خطا **Sampling Methods Errors**

روشهای نمونه گیری **Sampling Methods**

برآوردگرها **Estimators**

خطای معیار برآوردگرها **Standard Errors** - **SE**

ویژگی‌های برآوردگرها

برآوردگر نااریب **UnBiased**

قضیه حد مرکزی

توزیع تی **T-Student Distribution**

برآورد نقطه‌ای **Point Estimation** و فاصله‌ای **Internal Estimation**

.....

برآورد واریانس و انحراف معیار

توزیع کای مربع = خی

برآوردگر باثبات (سازگار) **Consistent** (قانون قوی اعداد بزرگ)

برآوردگر کافی (بسنده) **Sufficent** (حداکثر درست‌نمایی **(Maximum Liklihood**

دقت برآوردگر (کارائی نسبی برآوردگر **(Efficiency**

آزمون فرض

آزمون فرض میانگین‌ها **Compare Means**

.....

آیا اطلاعات بدست آمده با روشهای نمونه گیری از نمونه برای جمعیت عمومیت دارد

روشهای نمونه گیری و بررسی خطاها و تخمین های نقطه ای و فاصله ای پارامترها

نمونه گیری و خطا *Sampling Methods Errors*

روشهای نمونه گیری *Sampling Methods*

۱- روش نمونه گیری تصادفی ساده

قرعه کشی = اعطای شانس مساوی برای انتخاب

این روش برای جمعیت های همگن مفید است

مثلا با استفاده از جدول اعداد تصادفی

مثال از جدول که مثلا برای جمعیت ۳۰۰ نفره (به شماره از ۰۰۱ تا ۳۰۰) ، انتخاب نمونه ۵ نفره از داخل اعداد پنج رقمی تصادفی حاصل شده در نرم افزار اکسل

و انتخاب ۵ عدد سه رقمی از وسط جدول به سمت راست و انتخاب اعداد سه رقم وسطی با حذف سه رقمهای بیشتر از ۳۰۰ که نامربوط است و حذف اعداد سه رقمی که قبلا بوجود آمده)

62413 52318 47612 71221 41111 92310 40042 ....

بنابراین ۵ شماره بعنوان نمونه انتخاب شد = نمونه تصادفی ساده

عیب این روش اینکه چون داده ها مرتب نبود برای جمعیت های همگن مفید است و امکان اریبی (**Bias**) (= اشتباهات یک طرفه) وجود دارد.

مثلا اگر از جمعیت ۳۰۰ نفری ، که شامل ۲۰۰ خانم و ۱۰۰ آقا باشد و میخواهیم ۶ نفر انتخاب کنیم وقتی بتصادف ۶ نفر انتخاب کنیم احتمال دارد همه خانم یا همه آقا یا تعداد بیشتری آقا انتخاب شود در حقیقت عمل همگن سازی انجام نشده است

۲- روش نمونه گیری خوشه ای

جمعیت به خوشه هایی تقسیم میشود و کل خوشه در نمونه ظاهر میشود

مثلا نمونه هایی از گروههای دانشجویی رشت های مختلف

مثلا در یک دانشگاه ۴ رشته و ۱۴ کلاس داریم سه کلاس دانشجویی در رشته زیست شناسی و پنج کلاس در رشته میکرو بیولوژی و دو کلاس رشته ریاضی و چهار کلاس در رشته حقوق داریم و در هر کلاس تعدادی دانشجو داریم که تعداد میتواند با هم مساوی نباشد حال میخواهیم یک نمونه خوشه ای از دانشگاه انتخاب کنیم باید از سه کلاس رشته زیست شناسی یک کلاس کل دانشجویان بطور کامل انتخاب کنیم و از پنج کلاس رشته میکروبیولوژی کل دانشجویان یک کلاس و ... انتخاب کنیم حالا این نمونه خوشه ای است در حقیقت ۴ گروه دانشجویی کامل از بین ۱۴ کلاس انتخاب شده است

==========

۳- روش نمونه گیری سیستماتیک(سامان مند)

در این روش داده های جمعیت را بر حسب یک مورد مرتب میکنیم (مثلا مرتب شدن بر حسب جنسیت یا قد یا وزن یا سن یا ....)

مثلا ۳۰۰ نفر را به پنج گروه ۶۰ تایی تقسیم میکنیم و از هر گروه مثلا نفر چهاردهم(بتصادف) را انتخاب میکنیم ، پس اولین نفر، نفر ۱۴ و سپس نفر بعدی ۷۴ و نفر بعد تری ۱۳۴ و .....

این روش از همه طیف ها حضور دارند

ولی باز هم امکان اریبی (**Bias**) (= اشتباهات یک طرفه) وجود دارد ممکن است بعضی گروهها شانس بیشتری برای اظهار نظر داشته باشند (مثلا ممکن است از ۵ نفر انتخاب شده نفر چهارمی(از گروه چهارم) پرسشنامه ما را پاسخ ندهد پس گروههای دیگر شانس‌شان افزایش می‌یابد)

۴- روش نمونه گیری طبقه بندی **Classified**

برای مواردی که میخواهیم نمونه انتخابی، معرف آن جمعیت باشد به تناسب هر جمعیت تعداد نمونه انتخاب میشود

مثلا اگر از جمعیت ۳۰۰ نفری ، که شامل ۲۰۰ خانم و ۱۰۰ آقا باشد و میخواهیم ۶ نفر انتخاب کنیم از ۲۰۰ نفر خانم ۴ نفر را به تصادف انتخاب و از ۱۰۰ آقا ۲ نفر را به تصادف انتخاب میکینم در این روش با منفک کردن خانمها و آقایان، عمل همگن سازی انجام دادیم

۵- روش نمونه گیری با هدف خاص یا در دسترس یا ...

مثلا در یک محله ، عمر خودروهایی که افراد با جقوق متوسط در جامعه دارند مد نظر ما میباشد در این روش نمونه گیری به سراغ ثروتمندان و فقرا نرفتیم و فقط به سراغ افراد با درآمد متوسط رفتیم و عدد سن خودروهایشان برای کار آماری سوال نمودیم.

فرمولها

Statistics Formulas	
Mean	$\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n}$
Variance	$s^2 = \frac{\sum (x_i - \bar{x})^2}{n-1}$
Standard Deviation	$s = \sqrt{\frac{\sum (x_i - \bar{x})^2}{n-1}}$
Mean Absolute Deviation	$MAD = \frac{\sum \lvert x_i - \bar{x} \rvert}{n}$
Standard Error	$SE = \frac{s}{\sqrt{n}}$

برآوردگرها *Estimators*

تعریف پارامتر = شاخص های آماری مرتبط با ویژگیهای جمعیت باشد

تعریف آماره = شاخص های آماری مرتبط با ویژگیهای نمونه باشد

جدول شاخص‌ها

		میانگین	واریانس	انحراف معیار	نسبت	همبستگی	میانه
جمعیت	پارامتر	μ	σ^2	σ	p	ρ	η
نمونه	آماره	$\bar{x}$	S^2	SD	$\hat{p}$	$R\ ,r$	Md

برآوردگر : اگر از آماره برای تخمین پارامتر استفاده شود برآوردگر نامند

برآورد نقطه‌ای : اگر برای تخمین برآورد پارامتر فقط از برآوردگر نمونه استفاده شود آن را برآورد نقطه‌ای گویند

برآورد فاصله‌ای : با در دست داشتن خطای معیار

خطای معیار برآوردگرها *SE - Standard Errors*

۱- در آمار ممکن است شاخص های مرکزی داده‌ها (مکان) **Central Tendency** مد نظر باشد که توجه به مرکز داده ها دارد مثل (میانگین، میانه، مد)

۲- در آمار ممکن است شاخص های پراکندگی داده‌ها **Dispersion** مد نظر باشد که توجه به پراکنده بودن داده ها دارد مثل (مثل واریانس ، انحراف معیار ، ضریب تغییرات، دامنه)

۳- در آمار ممکن است شاخص تعداد **frequency** مد نظر باشد که توجه به مقدار داده ها دارد مثل (فراوانی و فراوانی تجمعی)

۴- در آمار ممکن است شاخص‌های نسبت **Proportion** و یا درصد داده‌ها **Percent** (چندک‌ها) مد نظر باشد که توجه به نسبت داده ها دارد مثل (فراوانی نسبی و فراوانی نسبی تجمعی و چارک و دهک و صدک)

در هر یک از چهار مورد فوق اگر از طریق نمونه شاخص‌ها را بدست آوریم تقریب زدن آن به جمعیت خطا حاصل میکند که به آن خطای معیار گویند

خطای معیار : برای تخمین پارامترجمعیت از آماره نمونه استفاده میکنیم که چون داریم از شاخص نمونه برای شاخص جمعیت تخمین میزنیم پس خطا دارد این را خطای معیار گویند

اگر از جمعیت **N** تایی داده ها یک نمونه **n** تایی را بصورت زیر در نظر بگیریم

$$X_1 , X_2 , X_3 , X_4 , ...$$

و میانگین نمونه داده ها را $\bar{X}$ بنامیم و انحراف معیار نمونه **Standard Deviation** را **SD** بنامیم آنگاه خطای معیار میانگین

$$SE(\bar{X}) = \frac{SD}{\sqrt{n}}$$

و میانه نمونه داده ها را Md بنامیم و اگر آنگاه خطای معیار میانه

$$SE(Md) = 1.253\frac{SD}{\sqrt{n}} = 1.253 * SE(\bar{X})$$

با همین فرمول متوجه میشویم که خطای معیار میانه از خطای معیار میانگین ۲۵٪ بیشتر است یعنی خطا در میانه بیشتر از خطا میانگین است مگر اینکه تعداد نمونه خیلی زیاد شود

مثال : از نمرات ریاضی یک کلاس ۵۰ نفری یک نمونه ۷ تایی انتخاب میکنیم خطای معیار میانگین و خطای معیار میانه را بدست آورید و مقایسه کنید (نمرات ۱۴ و ۱۲ و ۱۷و ۶ و ۱۷ و ۱۳ و ۱۵)

==========

X_i	6	12	13	14	15	17
f_i	1	1	1	1	1	2

میانگین نمونه

$$\bar{X} = \frac{\sum x_i f_i}{\sum f_i} = \frac{(6*1)+(12*1)+(13*1)+(14*1)+(15*1)+(17*1)}{1+1+1+1+1+2} = 13.428$$ میانگین نمونه

واریانس نمونه از یکی از دو فرمول زیر میتوان بدست آورد

$$S^2 = \frac{n\sum x^2-(\sum x)^2}{n(n-1)} = \frac{7(6^2+12^2+13^2+14^2+15^2+17^2+17^2)-(6+12+13+14+15+17+17)^2}{7(7-1)} = 14.29$$

یا به این طریق

$$S^2 = \frac{\sum(x_i-\bar{x})^2 f_i}{(\sum f_i)-1}$$ واریانس نمونه

$$= \frac{(6-13.428)^2*1+(12-13.428)^2*1+(13-13.428)^2*1+(14-13.428)^2*1+(15-13.428)^2*1+(17-13.428)^2*2}{1+1+1+1+1+2-1}$$

$$= 14.29$$

حال انحراف معیار نمونه بدست میاوریم همان **SD** که به **S** نوشتیم

$$S = \sqrt{14.29} = 3.78$$ انحراف معیار نمونه

حال میخواهیم از این نمونه اگر برای جمعیت استفاده کنیم مقدار خطای میانگین چقدر است

$$SE(\bar{X}) = \frac{SD}{\sqrt{n}} = \frac{3.78}{\sqrt{7}} = 1.42$$

و اگر بخواهیم از این نمونه اگر برای جمعیت استفاده کنیم مقدار خطای میانه چقدر است

$$SE(Md) = 1.253\frac{SD}{\sqrt{n}} = 1.253 * SE(\bar{X}) = 1.253 * 1.42 = 1.78$$

اگر در یک آماره نمونه از نسبت ها استفاده کنیم مثلا هرگاه از نمونه **n** تایی $\hat{p}$ نسبت مربوطه به نمونه و p نسبت مربوط به جمعیت باشد اگر $\hat{p}$ درصد مورد نظر (موفقیت) باشد آن آنگاه درصد خلاف آن (شکست) $\hat{q} = 1 - \hat{p}$ خواهد بود و آنگاه مقدار خطا برابر است با

در صد خطا معیار نسبت نمونه $SE(\hat{p}\%) = 100 * \sqrt{\frac{\hat{p}\,\hat{q}}{n}}$ خطای معیار نسبت نمونه $SE(\hat{p}) = \sqrt{\frac{\hat{p}\,\hat{q}}{n}}$

حال اگر بخواهیم درصد این خطا برای فراوانی نمونه f را برای فراوانی تعداد کل جمعیت محاسبه کنیم

$$SE(f) = N * SE(\hat{p}) = N * \sqrt{\frac{\hat{p}\,\hat{q}}{n}}$$ خطای معیار فراوانی

مثال : بتعداد کل ۳۰۰۰ نفر به بیمارستان برای بیماری کرونا مراجعه کردند از بین آنها یک نمونه ۵۰ نفری انتخاب شد و مشخص شد تعداد ۴ نفر از این ۵۰ نفر بیماری کرونا دارند – خطای معیار نسبت و خطای معیار درصد و خطای معیار فراوانی را بدست آورید

$$\hat{p} = \frac{4}{50} = 0.08 \qquad \hat{q} = 1 - \hat{p} = 1 - 0.08 = 0.92$$

$$SE(\hat{p}) = \sqrt{\frac{\hat{p}\ \hat{q}}{n}} = \sqrt{\frac{0.08 * 0.92}{50}} = 0.038$$ خطا معیار نمونه

$$SE(\hat{p}\%) = 100 * \sqrt{\frac{0.08 * 0.92}{50}} = \%3.8$$ در صد خطای معیار نمونه

$$SE(f) = N * SE(\hat{p}) = N * \sqrt{\frac{\hat{p}\ \hat{q}}{n}} = 3000 * 0.038 = 114$$

هر گونه تقریب برای ۳۰۰۰ نفر به میزان ۱۱۴ نفر احتمال دارد خطا داشته باشد

مثلا میانگین نمونه بیماران در ۵۰ نفر

$$\bar{x} = n * p * q = 50 * 0.08 * 0.92 = 3.68$$

مثلا میانگین جمعیت ۳۰۰۰ نفر (با تقریب به میانگین نمونه)

$$\bar{x} = n * p * q = 3000 * 0.08 * 0.92 = 220$$

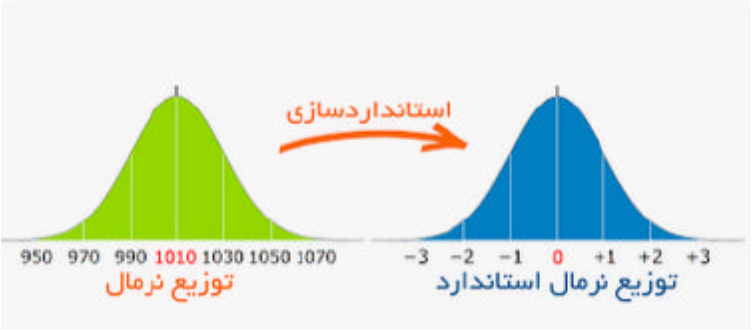
میانگین جمعیت بیماران در بین ۳۰۰۰ نفر 220 ± 114 میباشد یعنی در ۳۰۰۰ نفر بین ۶ نفر تا ۳۳۴ نفر بیمار هستند

اگر بخواهیم میزان خطا کمنر شود باید تعداد نمونه **n** را افزایش دهیم

در خصوص نمونه **n** تایی بیم دو سری عدد و ضریب همبستگی پیرسون **R** بین آن دو سری عدد (که **R** بین -۱ و ۱ است) ، آنگاه خطای معیار برآورد پیرسون برابر است با

$$SE(R) = \frac{1 - R^2}{\sqrt{n - 1}}$$

تمرین ها



ویژگی‌های برآوردگرها

در توزیع‌های متقارن میانگین ، میانه ، مود بر هم منطبق هستند

ابتدا توزیع نرمال تشریح میدهد

برآوردگر ناریب UnBiased

هرگاه امید ریاضی (مقدار انتظاری) مقدار برآوردگر به پارامتر جمعیت نزدیک باشد آنگاه میگوییم برآوردگر ناریب میباشد

قضیه حد مرکزی

اگر تعداد نمونه بیشتر از ۳۰ باشد آنگاه توزع نرمال است میگوییم توزیع نرمال است

اگر میانگین نمونه $\boldsymbol{\theta}$ باشد و $\widehat{\boldsymbol{\theta}}$ میانگین نمونه و توزیع نرمال باشد و خطای معیار نمونه $\boldsymbol{SE(\widehat{\theta})}$ باشد

اگر میانگین نمونه $\boldsymbol{\mu}$ باشد و $\bar{x}$ میانگین نمونه باشد و توزیع نرمال باشد و خطای معیار نمونه $\boldsymbol{SE(\bar{x})}$ باشد

$$\boldsymbol{z=\frac{\widehat{\theta}-\theta}{SE(\widehat{\theta})}} \qquad \boldsymbol{z=\frac{\bar{x}-\mu}{SE(\bar{x})}}$$

قبلا در مورد توزیع نرمال اشاره شد که :

مثال : در کل یک کلاس میانگین ۱۴ و واریانس ۳.۶ میباشد احتمال اینکه نمره یک دانش آموز در این کلاس بیشتز از ۱۷ باشد چقدر است؟ بیشترین نمره برای $X_{0.5}$ بدست آورید

$$X\sim N(\mu=14,\sigma^2=3.6) \qquad \sigma=\sqrt{3.6}=1.89$$

$$\boldsymbol{p(x>17)=1-p(x\le 17)=1-p\left(\frac{x-\mu}{\sigma}\le\frac{17-\mu}{\sigma}\right)=1-p\left(z\le\frac{17-14}{1.89}\right)=1-p(z\le+1.58)}$$

$$\boldsymbol{=1-0.9429=0.057}$$

به احتمال ۵,۷٪ نمره یک دانشجو بیشتر از ۱۷ میشود

$$\boldsymbol{p(x>k)=0.05}$$

$$\boldsymbol{p(x>k)+p(x\le k)=1} \qquad \boldsymbol{p(x>k)=1-p(x\le k)}$$

$$\boldsymbol{0.05=1-p\left(\frac{x-\mu}{\sigma}\le\frac{k-14}{1.89}\right)} \qquad \boldsymbol{0.05=1-p\left(z\le\frac{k-14}{1.89}\right)} \qquad \boldsymbol{p\left(z\le\frac{k-14}{1.89}\right)=0.95}$$

$$\boldsymbol{p\left(z\le\frac{k-14}{1.89}\right)=p(z\le 1.645)}$$

$$\boldsymbol{\frac{k-14}{1.89}=1.645} \qquad \boldsymbol{k-14=1.89*1.645} \qquad \boldsymbol{k=17}$$

یعنی به احتمال ۹۵٪ نمره دانشجویان زیر ۱۷ خواهد بود یا ۵٪ احتمال دارد نمره دانشجویی بیشتر از ۱۷ شود به عبارتی بیشترین نمره ۱۷ خواهد بود

مثال : در یک کلاس میانگین ۱۴ و واریانس ۳,۶ میباشد (انحراف معیار ۱,۸۹) میباشد احتمال اینکه نمره یک دانش آموز در یک کلاس ۲۵ نفری از ۱۷ بیشتر باشد

$$z = \frac{\bar{x} - \mu}{SE(\bar{x})} = \frac{\bar{x} - \mu}{\sigma/\sqrt{n}} =$$

$$p(x > 17) = 1 - p(x \le 17) = 1 - p\left(\frac{\bar{x} - \mu}{\sigma/\sqrt{n}} \le \frac{17 - \mu}{\sigma/\sqrt{n}}\right) = 1 - p\left(z \le \frac{17 - 14}{1.89/\sqrt{25}}\right)$$
$$= 1 - p(z \le +7.936) = 1 - 0.9999 = 0.0001$$

به احتمال خیلی ضعیف یعنی ۰,۰۱٪ در این کلاس نمره دانشجو بیشتر از ۱۷ میشود

توزیع تی T-Student Distribution

در قضیه حد مرکزی توزیع نرمال استاندارد اشاره شد که $z = \frac{\bar{x} - \mu}{SE(\bar{x})} = \frac{\bar{x} - \mu}{\sigma/\sqrt{n}}$ میباشد

در خصوص توزیع **T** میتوان نوشت

$$T = \frac{\bar{x} - \mu}{SE(\bar{x})} = \frac{\bar{x} - \mu}{SD/\sqrt{n}}$$

در توزیع **T** اگر حجم نمونه(**n**) زیاد باشد با توزیع نرمال استاندارد **N** مساوی است

دو نمودار نرمال و تی شبیه بهم و هر دو متقارن است اما دامنه توزیع نرمال از -۴ تا +۴ است ولی دامنه **T** بسیار بزرگ است

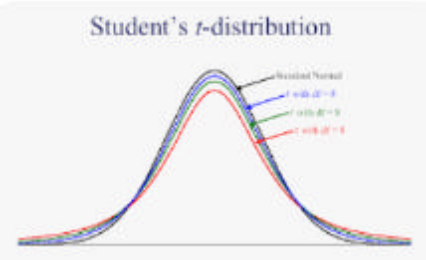
در **T** به شاخصی نیاز داریم بنام درجه آزادی **Df = Degree Of Freedom**

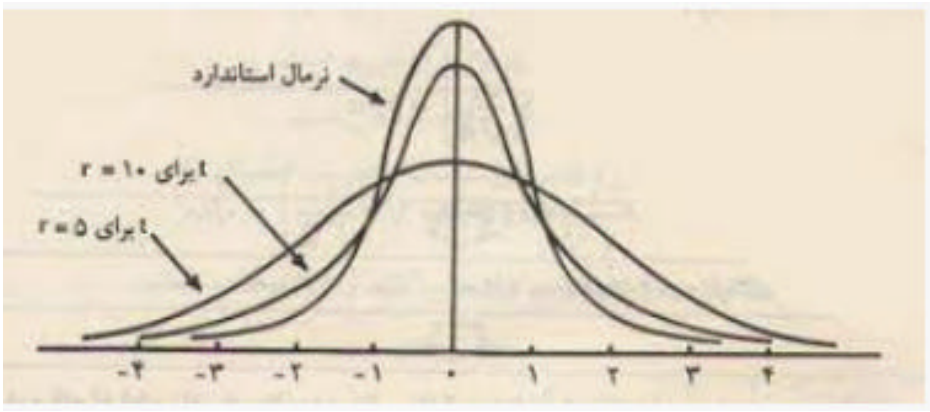
$$df = n - 1$$

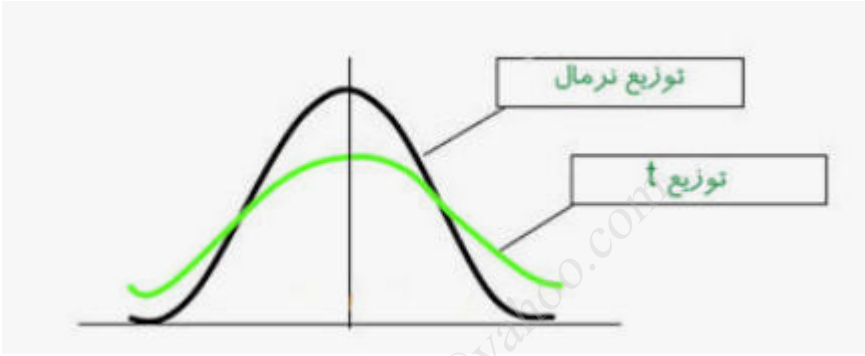
در خصوص میانگین و واریانس قبلا اشاره شد که در نمونه های با حجم کم مخرج کسر واریانس **n-1** میباشد

میانگین $\bar{x} = \frac{\sum x_i f_i}{\sum f_i} = \frac{\sum x_i}{n}$ واریانس $S^2 = \frac{\sum (x_i - \bar{x})^2 f_i}{(\sum f_i) - 1} = \frac{\sum (x_i - \bar{x})^2}{n - 1}$

انحراف معیار $SD = \sqrt{\frac{\sum (x_i - \bar{x})^2}{n - 1}}$

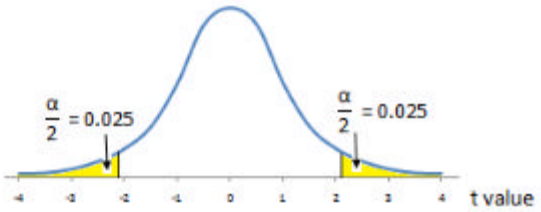






Student's t Distribution Table

For example, the t value for 18 degrees of freedom is 2.101 for 95% confidence interval (2-Tail $\alpha = 0.05$).

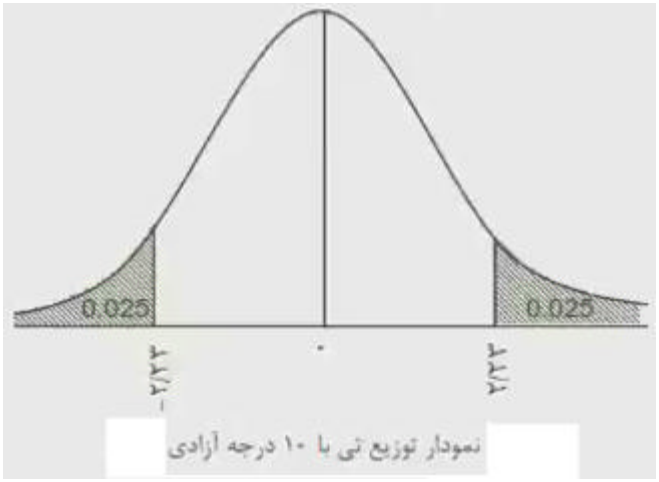


	90%	95%	97.5%	99%	99.5%	99.95%	1-Tail Confidence Level
	80%	90%	95%	98%	99%	99.9%	2-Tail Confidence Level
	0.100	0.050	0.025	0.010	0.005	0.0005	1-Tail Alpha
df	0.20	0.10	0.05	0.02	0.01	0.001	2-Tail Alpha
1	3.0777	6.3138	12.7062	31.8205	63.6567	636.6192	
2	1.8856	2.9200	4.3027	6.9646	9.9248	31.5991	
3	1.6377	2.3534	3.1824	4.5407	5.8409	12.9240	
4	1.5332	2.1318	2.7764	3.7469	4.6041	8.6103	
5	1.4759	2.0150	2.5706	3.3649	4.0321	6.8688	
6	1.4398	1.9432	2.4469	3.1427	3.7074	5.9588	
7	1.4149	1.8946	2.3646	2.9980	3.4995	5.4079	
8	1.3968	1.8595	2.3060	2.8965	3.3554	5.0413	
9	1.3830	1.8331	2.2622	2.8214	3.2498	4.7809	
10	1.3722	1.8125	2.2281	2.7638	3.1693	4.5869	
11	1.3634	1.7959	2.2010	2.7181	3.1058	4.4370	
12	1.3562	1.7823	2.1788	2.6810	3.0545	4.3178	
13	1.3502	1.7709	2.1604	2.6503	3.0123	4.2208	
14	1.3450	1.7613	2.1448	2.6245	2.9768	4.1405	
15	1.3406	1.7531	2.1314	2.6025	2.9467	4.0728	
16	1.3368	1.7459	2.1199	2.5835	2.9208	4.0150	
17	1.3334	1.7396	2.1098	2.5669	2.8982	3.9651	
18	1.3304	1.7341	2.1009	2.5524	2.8784	3.9216	
19	1.3277	1.7291	2.0930	2.5395	2.8609	3.8834	
20	1.3253	1.7247	2.0860	2.5280	2.8453	3.8495	
21	1.3232	1.7207	2.0796	2.5176	2.8314	3.8193	
22	1.3212	1.7171	2.0739	2.5083	2.8188	3.7921	
23	1.3195	1.7139	2.0687	2.4999	2.8073	3.7676	
24	1.3178	1.7109	2.0639	2.4922	2.7969	3.7454	
25	1.3163	1.7081	2.0595	2.4851	2.7874	3.7251	
26	1.3150	1.7056	2.0555	2.4786	2.7787	3.7066	
27	1.3137	1.7033	2.0518	2.4727	2.7707	3.6896	
28	1.3125	1.7011	2.0484	2.4671	2.7633	3.6739	
29	1.3114	1.6991	2.0452	2.4620	2.7564	3.6594	
30	1.3104	1.6973	2.0423	2.4573	2.7500	3.6460	

در این جدول هم نشان میدهد کل مساحت زیر منحنی **1** میباشد و نصف منحنی **0.5** میباشد

که در توزیع **T** در هر سمت **0.025** و در دو سمت با هم **0.05** مساحت زیر منحنی را نشان میدهد

مثلا توزیع **T** با **10** درجه آزادی و سطح اطمینان ۹۷,۵٪ به مقدار **2.228** میباشد



برآورد نقطه‌ای Point Estimation و فاصله‌ای Internal Estimation

آمار توصیفی جمع آوری و سازمان دهی داده ها و بدست آوردن شاخص های آماری نمونه (یا جمعیت) داده ها میباشد

آمار استباطی تحلیل و تخمین از شاخص های نمونه برای جمعیت است

برای تخمین از برآوردگرها استفاده میشود

برآورد نقطه ای : ارائه دادن گزارش دادن شاخص نمونه ، بجای جمعیت (که طبیعتا دقت کمی دارد)

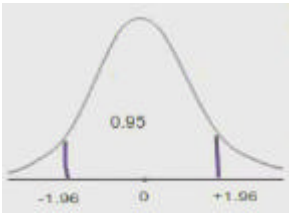
برآورد فاصله‌ای : ارائه دادن گزارش دادن شاخص نمونه ، بجای جمعیت با ارائه ضریب اطمینان

سطح اطمینان را با $(1-\alpha)\%$ نشان میدهند

قبلا در توزیع نرمال اشاره شد که سطح زیر منحنی برای ۹۵٪ بین $z = \pm 1.96$ میباشد

$$z = \frac{\hat{\theta} - \theta}{SE(\hat{\theta})} \qquad -1.96 < \frac{\hat{\theta} - \theta}{SE(\hat{\theta})} < +1.96$$

$$Low = \hat{\theta} - 1.96 * SE(\hat{\theta}) \qquad Up = \hat{\theta} + 1.96 * SE(\hat{\theta})$$



یعنی با اطمینان $(1-\alpha)\% = 95\%$ یعنی ۹۵٪ میتوان گفت که z بین ۱,۹۶- تا ۱,۹۶ + میباشد

برای داده با تعداد زیاد در جدول T هم همین بدست میاید

یعنی برای تعداد خیلی زیاد، نقطه بحرانی دو دامنه توزیع نرمال استاندارد در سطح ۵٪ از جدول توزیع نرمال با جدول توزیع T با هم مساوی میباشد

مثال: از نمونه ۲۰۰ نفری دانشجویان میانگین نمرات ۱۳ و انحراف معیار ۱,۵ و میانه نمرات ۱۲ میباشد فاصله اطمینان ۹۵٪ برای میانگین و میانه را محاسبه کنید و توضیح بنویسید

با توجه به صورت مسئله آماره ها $\bar{x} = 13$ و $SD = 1.5$ و Md=12 میباشد

برآورد نقطه ای میانگین جمعیت 13 و برآورد نقطه ای میانه جمعیت هم 12 میگردد اما این برآورد از نمونه به جمعیت مقداری خطا دارد

اگر از میانگین نمونه بعنوان میانگین جمعیت $\hat{\mu} = \bar{x} = 13$ استفاده و برآورد نقطه ای را حاصل کنیم میزان خطای معیار میانگین برابر است با

$$SE(\bar{x}) = \frac{SD}{\sqrt{n}} = \frac{1.5}{\sqrt{200}} = 0.11$$

اگر از میانه نمونه بعنوان میانه جمعیت $\hat{\eta} = \bar{x} = 12$ استفاده و برآورد نقطه‌ای شود میزان خطای معیار میانه برابر است با

$$SE(Md) = 1.253 * \frac{SD}{\sqrt{n}} = 1.253 * \frac{1.5}{\sqrt{200}} = 0.13$$

حال چون تعداد نمونه **n=200** عدد بزرگی میباشد میتوان از نرمال استاندارد برای فاصله اطمینان ۹۵٪ استفاده نمود

طبق جدول نرمال استاندارد برای فاصله اطمینان ۹۵٪ از دو طرف مقدار **Z= -1.96** تا **Z= +1.96** میگردد

$$z = \frac{\bar{x} - \mu}{SE(\bar{x})} \qquad -1.96 < \frac{\bar{x} - \mu}{SE(\bar{x})} < +1.96$$

$$Low = \hat{\mu} - 1.96 * SE(\hat{\mu}) = \bar{x} - 1.96 * SE(\bar{x}) = 13 - 1.96 * 0.11 = 12.78$$
$$Up = \hat{\mu} + 1.96 * SE(\hat{\mu}) = \bar{x} + 1.96 * SE(\bar{x}) = 13 + 1.96 * 0.11 = 13.22$$

یعنی با اطمینان ۹۵٪ میتوان گفت میانگین جمعیت بین ۱۲,۷۸ تا ۱۳,۲۲ میباشد

و برای میانه میتوان نوشت

$$Low = \hat{\eta} - 1.96 * SE(\hat{\eta}) = Md - 1.96 * SE(Md) = 12 - 1.96 * 0.13 = 11.75$$
$$Up = \hat{\eta} + 1.96 * SE(\hat{\eta}) = Md + 1.96 * SE(Md) = 12 + 1.96 * 0.13 = 12.25$$

یعنی با اطمینان ۹۵٪ میتوان گفت میانه جمعیت بین ۱۱,۷۵ تا ۱۲,۲۵ میباشد

اگر تعداد نمونه **n** کوچک بلشد باید از **T** استیودنت با درجه آزادی $df = n - 1$ و فاصله اطمینان دو دامنه **0.95** از جدول **T** استفاده کنیم و از این دو فرمول برای فاصله اطمینان استفاده میکینم

$$Low = \hat{\theta} - t * SE(\hat{\theta})$$
$$Up = \hat{\theta} + t * SE(\hat{\theta})$$

مثال: از نمونه ۲۰ نفری دانشجویان میانگین نمرات ۱۳ و انحراف معیار ۱,۵ میباشد فاصله اطمینان ۹۵٪ برای میانگین را محاسبه کنید و توضیح بنویسید

با توجه به صورت مسئله آماره ها $\bar{x} = 13$ و $SD = 1.5$ میباشد برآورد میانگین جمعیت **13** میگردد اما این برآورد از نمونه به جمعیت مقداری خطا دارد

اگر از میانگین نمونه بعنوان میانگین جمعیت $\hat{\theta} = \bar{x} = 13$ استفاده و برآورد را حاصل کنیم میزان خطای معیار میانگین برابر است با

$$SE(\hat{\theta}) = \frac{SD}{\sqrt{n}} = \frac{1.5}{\sqrt{20}} = 0.33$$

چون **n=20**کوچک میباشد از **T** با درجه آزادی **df=n-1=20-1=19** استفاده میکنیم که برای فاصله اطمینان دو طرفه **0.95** از **-2.09** تا **+2.09** میباشد

$$t = \frac{\hat{\theta} - \theta}{SE(\hat{\theta})} \qquad -2.09 < \frac{\hat{\theta} - \theta}{SE(\hat{\theta})} < +2.09$$

$$Low = \hat{\theta} - 2.09 * SE(\hat{\theta}) = 13 - (2.09 * 0.33) = 12.31$$
$$Up = \hat{\theta} + 2.09 * SE(\hat{\theta}) = 13 + (2.09 * 0.33) = 13.69$$

یعنی با اطمینان ۹۵٪ میتوان گفت میانگین جمعیت بین ۱۲,۳۱ تا ۱۳,۶۹ میباشد

از جدول **T** با درجه آزادی ۱۹ برای فاصله اطمینان دو طرفه **0.99** از **-2.86** تا **+2.86** میباشد

$$Low = \hat{\theta} - 2.86 * SE(\hat{\theta}) = 13 - (2.86 * 0.33) = 12.05$$
$$Up = \hat{\theta} + 2.86 * SE(\hat{\theta}) = 13 + (2.86 * 0.33) = 13.94$$

یعنی با اطمینان ۹۹٪ میتوان گفت میانگین جمعیت بین ۱۲,۰۵ تا ۱۳,۹۴ میباشد

مثال : در یک نمونه $n=30$ تایی دانشجویان ضریب همبستگی بین نمرات درس فیزیک و ریاضی $R=0.75$ میباشد یک فاصله اطمینان ۹۹٪ برای ضریب همبستگی بدست آورید

$$SE(R) = \frac{1 - R^2}{\sqrt{n-1}} = \frac{1 - 0.75^2}{\sqrt{30-1}} = \frac{0.437}{5.385} = 0.081$$

از جدول **T** با درجه آزادی **df=n-1=30-1=29** برای فاصله اطمینان دو طرفه **0.99** از **-2.756** تا **+2.756** میباشد

$$Low = \hat{\rho} - 2.756 * SE(\hat{\rho}) = 0.75 - (2.756 * 0.081) = 0.526$$
$$Up = \hat{\rho} + 2.256 * SE(\hat{\rho}) = 0.75 + (2.756 * 0.081) = 0.973$$

بنابراین با اطمینان ۹۹٪ میتوان گفت که ضریب همبستگی بین نمرات درس فیزیک و ریاضی بین ۰,۵۲ تا ۰,۹۷ میباشد

یعنی نمره یک دانشجو برای نمره درس قیزیک به نمره درس ریاضی هم ربط دارد

برآورد واریانس و انحراف معیار

قبلا فرمول واریانس جمعیت و نمونه ذکر شد

$$\sigma^2 = \frac{\sum(x_i - \bar{x})^2 f_i}{\sum f_i} = \frac{\sum(x_i - \bar{x})^2}{n}$$
$$S^2 = \frac{\sum(x_i - \bar{x})^2 f_i}{(\sum f_i) - 1} = \frac{\sum(x_i - \bar{x})^2}{n-1} = \frac{n\sum x^2 - (\sum x)^2}{n(n-1)}$$

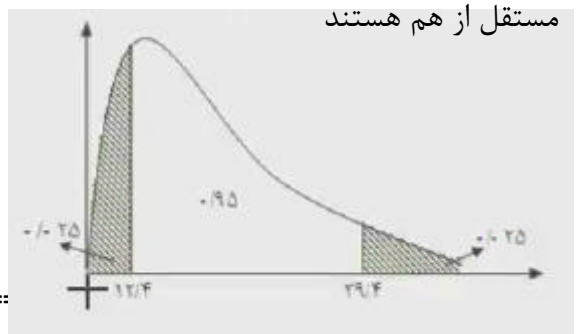
توزیع کای مربع = خی

در یک جمعیت با واریانس σ^2 ، اگر یک نمونه **n** تایی با واریانس S^2 داشته باشیم آنگاه کسر زیر دارای توزیع خاصی میباشد

$$\frac{(n-1)S^2}{\sigma^2}$$

و میتوان نوشت

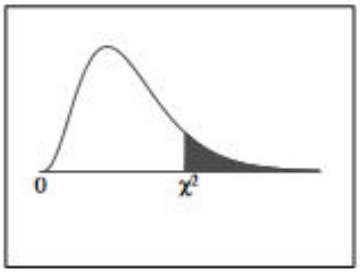
که با درجه آزادی **v = n-1** مینامند و دارای جدول خاصی میباشد نمودار این متقارن نیست و میانگین و واریانس نمونه مستقل از هم هستند



$$\chi_a^2(\nu) < \frac{(n-1)S^2}{\sigma^2} < \chi_b^2(\nu) \qquad \frac{(n-1)S^2}{\chi_b^2(\nu)} < \sigma^2 < \frac{(n-1)S^2}{\chi_a^2(\nu)}$$

بنابراین واریانس با فاصله اطمینان مورد نظر و درجه آزادی **n-1** و از جدول مربوطه ، بین دو مقدار فوق میباشد

Chi-Square Distribution Table



The shaded area is equal to α for $\chi^2 = \chi^2_\alpha$.

df	$\chi^2_{.995}$	$\chi^2_{.990}$	$\chi^2_{.975}$	$\chi^2_{.950}$	$\chi^2_{.900}$	$\chi^2_{.100}$	$\chi^2_{.050}$	$\chi^2_{.025}$	$\chi^2_{.010}$	$\chi^2_{.005}$
1	0.000	0.000	0.001	0.004	0.016	2.706	3.841	5.024	6.635	7.879
2	0.010	0.020	0.051	0.103	0.211	4.605	5.991	7.378	9.210	10.597
3	0.072	0.115	0.216	0.352	0.584	6.251	7.815	9.348	11.345	12.838
4	0.207	0.297	0.484	0.711	1.064	7.779	9.488	11.143	13.277	14.860
5	0.412	0.554	0.831	1.145	1.610	9.236	11.070	12.833	15.086	16.750
6	0.676	0.872	1.237	1.635	2.204	10.645	12.592	14.449	16.812	18.548
7	0.989	1.239	1.690	2.167	2.833	12.017	14.067	16.013	18.475	20.278
8	1.344	1.646	2.180	2.733	3.490	13.362	15.507	17.535	20.090	21.955
9	1.735	2.088	2.700	3.325	4.168	14.684	16.919	19.023	21.666	23.589
10	2.156	2.558	3.247	3.940	4.865	15.987	18.307	20.483	23.209	25.188
11	2.603	3.053	3.816	4.575	5.578	17.275	19.675	21.920	24.725	26.757
12	3.074	3.571	4.404	5.226	6.304	18.549	21.026	23.337	26.217	28.300
13	3.565	4.107	5.009	5.892	7.042	19.812	22.362	24.736	27.688	29.819
14	4.075	4.660	5.629	6.571	7.790	21.064	23.685	26.119	29.141	31.319
15	4.601	5.229	6.262	7.261	8.547	22.307	24.996	27.488	30.578	32.801
16	5.142	5.812	6.908	7.962	9.312	23.542	26.296	28.845	32.000	34.267
17	5.697	6.408	7.564	8.672	10.085	24.769	27.587	30.191	33.409	35.718
18	6.265	7.015	8.231	9.390	10.865	25.989	28.869	31.526	34.805	37.156
19	6.844	7.633	8.907	10.117	11.651	27.204	30.144	32.852	36.191	38.582
20	7.434	8.260	9.591	10.851	12.443	28.412	31.410	34.170	37.566	39.997
21	8.034	8.897	10.283	11.591	13.240	29.615	32.671	35.479	38.932	41.401
22	8.643	9.542	10.982	12.338	14.041	30.813	33.924	36.781	40.289	42.796
23	9.260	10.196	11.689	13.091	14.848	32.007	35.172	38.076	41.638	44.181
24	9.886	10.856	12.401	13.848	15.659	33.196	36.415	39.364	42.980	45.559
25	10.520	11.524	13.120	14.611	16.473	34.382	37.652	40.646	44.314	46.928
26	11.160	12.198	13.844	15.379	17.292	35.563	38.885	41.923	45.642	48.290
27	11.808	12.879	14.573	16.151	18.114	36.741	40.113	43.195	46.963	49.645
28	12.461	13.565	15.308	16.928	18.939	37.916	41.337	44.461	48.278	50.993
29	13.121	14.256	16.047	17.708	19.768	39.087	42.557	45.722	49.588	52.336
30	13.787	14.953	16.791	18.493	20.599	40.256	43.773	46.979	50.892	53.672
40	20.707	22.164	24.433	26.509	29.051	51.805	55.758	59.342	63.691	66.766
50	27.991	29.707	32.357	34.764	37.689	63.167	67.505	71.420	76.154	79.490
60	35.534	37.485	40.482	43.188	46.459	74.397	79.082	83.298	88.379	91.952
70	43.275	45.442	48.758	51.739	55.329	85.527	90.531	95.023	100.425	104.215
80	51.172	53.540	57.153	60.391	64.278	96.578	101.879	106.629	112.329	116.321
90	59.196	61.754	65.647	69.126	73.291	107.565	113.145	118.136	124.116	128.299
100	67.328	70.065	74.222	77.929	82.358	118.498	124.342	129.561	135.807	140.169

جدول کای مربع (خی۲)

مثال : در یک کلاس با ۲۵ دانشجو انحراف معیار نمرات ۱,۷۵ گردید اگر توزیع نرمال باشد ضریب اطمینان ۹۵٪ برای انحراف معیار جمعیت را بدست آورید

n = 25 , SD = 1.75

$\hat{\sigma} = SD = 1.75 \qquad \sigma^2 = S^2 = 1.75^2 = 3.06 \qquad \vartheta = n - 1 = 25 - 1 = 24$

از جدول کای مربع (خی) فاصله اطمینان ۹۵٪ (۰,۹۵ یعنی دو تا دنباله ۰,۰۲۵ در چپ و ۰,۰۲۵ در راست) و با ۲۴ درجه آزادی ، برای واریانس دو مقدار زیر حاصل میشود

$\chi^2_{0.025}(24) = 39,4 \qquad \chi^2_{0.975}(24) = 12.4.$

$$Low = \frac{(n-1)S^2}{\chi^2_a(v)} = \frac{(25-1)*3.06}{39.4} = 1.86$$

$$Up = \frac{(n-1)S^2}{\chi^2_b(v)} = \frac{(25-1)*3.06}{12.4} = 5.92$$

از این دو باید جذر گرفت تا حد بالا و پایین برای انحراف معیار حاصل شود

$$Low = \sqrt{1.86} = 1.36$$
$$Up = \sqrt{5.92} = 2.43$$

پس با اطمینان ۹۵٪ میتوان گفت انحراف معیار جمعیت بین ۱,۳۶ تا ۲,۴۳ خواهد بود

برآوردگر باثبات (سازگار) Consistent (قانون قوی اعداد بزرگ)

اگر با افزایش دادن تعداد نمونه آماره نمونه به پارامتر جمعیت نزدیک شویم برآوردگر باثبات گویند

برآوردگر کافی (بسنده) Sufficent (حداکثر درست‌نمایی Maximum Liklihood)

از تمام اطلاعات موجود در نمونه استفاده کنیم

دقت برآوردگر (کارائی نسبی برآوردگر Efficiency)

کارائی برآوردگر **T1** به **T2**

$$Ef\ (T_1 \rightarrow T_2) = (\frac{SE(T_2)}{SE(T_1)})^2$$

کارائی نسبی میانه نسبت به میانگین بدست آورید

$$Ef\ (Md \rightarrow \bar{x}) = \left(\frac{SE(\bar{x})}{SE(Md)}\right)^2 = \left(\frac{\frac{\sigma}{\sqrt{n}}}{1.253\frac{\sigma}{\sqrt{n}}}\right)^2 = 0.637 \cong 0.64$$

پس میانه کارائی کمتری نسبت به میانگین دارد

در یک نمونه تحقیقاتی ۱۰۰ نفره اگر بخواهیم از شاخص میانه بجای میانگین استفاده کنیم حجم نمونه چقدر باید باشد تا میانه کارائی ، میانگین را داشته باشد.

$$Ef\ (Md \rightarrow \bar{x}) = \left(\frac{SE(\bar{x})}{SE(Md)}\right)^2 \qquad 1 = \left(\frac{\frac{\sigma}{\sqrt{100}}}{1.253\frac{\sigma}{\sqrt{n}}}\right)^2 \qquad 1 = \frac{n}{157} \qquad n = 157$$

اگر حجم نمونه ۱۵۷ شود و میانه را بدست آوریم از این عدد میانه میتوان بجای میانگین استفاده کرد.

بقیه جزوه ادامه دارد

آزمون فرض

آزمون فرض میانگین‌ها *Compare Means*

1- **One Sample**
2- **Two Independent Samples**
3- **Two Paired Samples**

میانگین ، میانه و مود سه شاخص مرکزی (مکان مرکزی) می‌باشند

شاخص میانگین، شاخص مناسب تری برای مرکز داده ها میباشد

اگر جمعیتی دارای میانگین $\boldsymbol{\mu}$ و واریانس $\boldsymbol{\sigma^2}$ باشد

اگر از این جمعیت نمونه ای انتخاب کنیم میانگین $\bar{x}$ و واریانس S^2 باشد

اگر μ_0 یک مقدار مفروض برای میانگین جمعیت باشد سه فرضیه متصور میشود

دو دامنه $\begin{cases} H_0: \mu = \mu_0 \\ H_1: \mu \neq \mu_0 \end{cases}$ یک دامنه $\begin{cases} H_0: \mu \leq \mu_0 \\ H_1: \mu > \mu_0 \end{cases}$ یک دامنه $\boldsymbol{one\ tailed}$ $\begin{cases} H_0: \mu \geq \mu_0 \\ H_1: \mu < \mu_0 \end{cases}$

اگر واریانس جمعیت σ^2 باشد و تعداد نمونه **n** باشد (طبیعتا نمونه هم میانگین و واریانس خاص خودش را دارد) برای ای نمونه خطای معیار میانگین $SE(\bar{x}) = \sigma/\sqrt{n}$ میشود و برای هر سه حالت، آماره آزمون برابر است با $\frac{\bar{x}-\mu_0}{SE(\bar{x})}$ که آنگاه آماره آزمون نرمال استاندارد میشود و میتوان نوشت

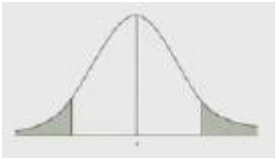
$$z = \frac{\bar{x} - \mu_0}{SE(\bar{x})} = \frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}}$$

اگر σ نداشتیم آنگاه $SE(\bar{x}) = SD/\sqrt{n}$ و دیگر توزیع نرمال نمیشود و از توزیع **t** با **n-1** درجه آزادی استفاده میکنیم

$$t = \frac{\bar{x} - \mu_0}{SE(\bar{x})} = \frac{\bar{x} - \mu_0}{SD/\sqrt{n}}$$

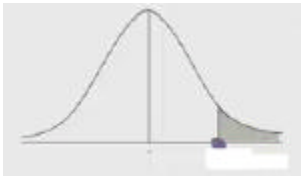
توجه شود در مشاهده نمونه ها $x_1, x_2, x_3, \dots, x_n$ میانگین $\bar{x} = \frac{\sum x_i}{n}$ میباشد که اگر این میانگین در **Z** نرمال قرار دهیم به این Z_{ob} گویند (یعنی **Z** مشاهده شده)

در خصوص آزمون دو دامنه ، ناحیه بحرانی در دو سمت محور قرار میگیرد و سطح معنا داری (**sig**) مساحت زیر منحنی با فرمول زیر حاصل میشود



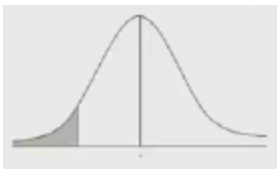
$$Si = 2\,Min\{P(z > z_{ob}), P(z < z_{ob})\}$$

در خصوص آزمون یک دامنه بعدی، که ناحیه بحرانی در سمت راست محور قرار میگیرد و سطح معنا داری (**sig**) مساحت زیر منحنی با فرمول زیر حاصل میشود



$$Si \quad = P(z > z_{ob})$$

در خصوص آزمون یک دامنه بعدی، که ناحیه بحرانی در سمت چپ محور قرار میگیرد و سطح معنا داری (**sig**) مساحت زیر منحنی با فرمول زیر حاصل میشود



$$Si \quad = P(z < z_{ob})$$

اگر **Sig** از مقدار آلفا کوچکتر باشد باید فرض صفر را رد کنیم

اگر واریانس σ^2 نامشخص باشد از t_{ob} جایگزین میشود و از توزیع **t** استفاده میشود (سطح معنا داری **t** خارج از درس میشود)

بجای محاسبه سطح معنا داری میتواند تعیین ناحیه بحرانی مد نظر باشد آنگاه مرز ناحیه بحرانی طوری انتخاب میشود که مساحت فضای هاشور خورده به میزان خطای نوع اول شود

اگر سطح معنا داری از خطای نوع اول α کمتر شود آنگاه با ضریب اطمینان $1-\alpha$ فرض صفر رد میشود

مثال : از تحقیقات قبلی مشخص میشود که در ایران هوش عاطفی زنانی که تحصیلات عالی دارند دارای توزیع نرمال با میانگین ٦٥ و انحراف معیار ١٢ میباشد این تحقیقات ادعا دارد گه در استان آذربایجان میانگین هوش عاطفی زنان بیشتر از دیگر استانها است یک نمونه ١٠٠ نفری در آذربایجان انتخاب کردیم و مشخص شد که میانگین ٦٧ است با سطح اطمینان ٩٥٪ ادعای تحقیقات را آزمون کنید

آزمونی یک دامنه است

$$\begin{matrix} \{H_0: \mu \le \mu_0\} \\ \{H_1: \mu > \mu_0\} \end{matrix} \qquad \begin{matrix} \{H_0: \mu \le 65\} \\ \{H_1: \mu > 65\} \end{matrix}$$

$$E(\bar{x}) = \frac{\sigma}{\sqrt{n}} = \frac{12}{\sqrt{100}} = 1.2 \qquad Z_{ob} \ \frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}} = \frac{68-65}{1.2} = 2.5$$

$$Si \quad = P(Z > Z_{ob}) = P(Z > 2.5) = 1 - \quad (Z \le 2.5) = 1 - \ 0.9938 = 0.0062$$

حال چون رابطه زیر برقرار است

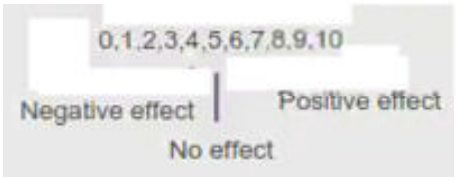
$$Si \quad = 0.0062 \le 0.05$$

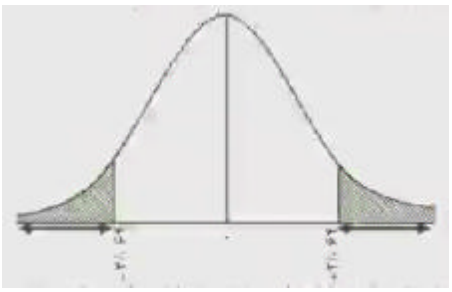
در نتیجه با اطمینان ٩٥٪ میتوان فرض صفر **H0** را رد نمود و پذیرای فزض **H1** شد

در فرض یک ذکر شده که میانگین هوش عاطفی زنان با تحصیلات دانشگاهی از ۶۵ بیشتر است پس ادعا مورد تایید است.

مثال : از ۲۵ نفر از معلمان برای رضایت از اثر بخشی نظام جدید آموزشی، نظر سنجی شد. ۱۰ سوال دو جوابی مطرح شد پاسخها بین صفر تا ۱۰ ده بود بطوریکه عدد صفر بیانگر اثر منفی و عدد پنج بی اثر و عدد ۱۰ اثر مثبت است. میانگین ٦ و انحراف معیار ۳ گردید

ادعا این است که معلمان نظام جدید آموزشی جدید را نسبت به قدیم بدون تاثیر است یعنی تغییری اساسی بوجود نیاورده است. با ضریب اطمینان ۹۵٪ ادعا را بیازمایید





$\begin{Bmatrix} H_0: \mu = \mu_0 \\ H_1: \mu \neq \mu_0 \end{Bmatrix}$, $\begin{Bmatrix} H_0: \mu = 5 \\ H_1: \mu \neq 5 \end{Bmatrix}$ دو دامنه ,

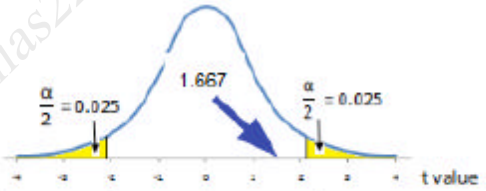
$n = 25$, $\bar{x} = 6$, $SD = 3$

$$t_{ab} = \frac{\bar{x} - \mu_0}{SE(\bar{x})} = \frac{\bar{x} - \mu_0}{SD/\sqrt{n}} = \frac{6 - 5}{3/\sqrt{25}} = 1.667$$

چون انحراف معیار جمعیت نداریم و انحراف معیار نمونه **SD** داریم آماره آزمون توزیع **t** با **24** درجه آزادی داریم که از جدول **t** برای خطای نوع اول**5%** (ضریب اطمینان **95%**) استفاده میکنیم برای دو دامنه **2.064** میگردد و چون **1.667 < 2.064** میباشد در حقیقت **1.667** در ناحیه بحرانی (هاشور زده) قرار ندارد پس دلیلی بر رد صفر نداریم ($H_0: \mu = \mu_0$) پس با اطمینان **95%** اعلام میکنیم که نظام آموزشی جدید بی تاثیر بوده است (نه منفی است نه مثبت)

Student's t Distribution Table

For example, the t value for 18 degrees of freedom is 2.101 for 95% confidence interval (2-Tail $\alpha = 0.05$).



	90%	95%	97.5%	99%	99.5%	99.95%	1-Tail Confidence Level
	80%	90%	95%	98%	99%	99.9%	2-Tail Confidence Level
	0.100	0.050	0.025	0.010	0.005	0.0005	1-Tail Alpha
df	0.20	0.10	0.05	0.02	0.01	0.001	2-Tail Alpha
22	1.3212	1.7171	2.0739	2.5083	2.8188	3.7921	
23	1.3195	1.7139	2.0687	2.4999	2.8073	3.7676	
24	1.3178	1.7109	2.0639	2.4922	2.7969	3.7454	
25	1.3163	1.7081	2.0595	2.4851	2.7874	3.7251	
26	1.3150	1.7056	2.0555	2.4786	2.7787	3.7066	

اگر محاسبه سطح معنا داری(**sig**) مد نظر نباشد اندیس **obs** در **z** و **t** نمی افزاییم یعنی اگر **z** به یک عدد برسد همان z_{ob} بعنی مشاهده شده است و اگر فرمول باشد اندیس ندارد همان آماره آزمون است ضمنا خطای نوع اول یعنی **5%** یا **1%** است.

این جزوه با تلاش این حقیر، جهت دانلود مجانی برای همه شما سروران تهیه شده است

بدیهی است خالی از ایراد نیست

هر گونه اشکالی به sedighias220@yahoo.com ارسال نمایید

با تشکر (امین صدیقی)

==========پایان==========

در هر حرفه ای که هستید نه اجازه دهید که به بدبینیهای بیحاصل آلوده شوید و نه بگذارید که بعضی لحظات تاسف بار که برای هر ملتی پیش می آید شما را به یاس و ناامیدی بکشاند. در آرامش حاکم بر آزمایشگاهها و کتابخانه هایتان زندگی کنید . نخست از خود بپرسید : " برای یادگیری و خودآموزی چه کرده ام ؟ "
سپس همچنان که پیشتر میروید بپرسید : " من برای کشورم چه کرده ام ؟ "
و این پرسش را آنقدر ادامه دهید تا به این احساس شادیبخش و هیجان انگیز برسید که شاید سهم کوچکی در پیشرفت و اعتلای بشریت داشته اید.
اما هر پاداشی که زندگی به تلاشهایمان بدهد یا ندهد هنگامی که به پایان تلاشهایمان نزدیک میشویم هر کداممان باید حق آن را داشته باشیم که با صدای بلند بگوییم
" من آنچه در توان داشته ام انجام داده‌ام " لوئی پاستور ۱۸۹۵-۱۸۲۲