کاربرد آمار در علوم اجتماعی

آمار توصیفی

علم آمار به ۳ قسمت تقسیم می‌شود:

۱) آمار توصیفی ۲) آمار تحلیلی ۳) آمار ناپارامتریک

مؤلفه‌های موجود در آمار:

جامعه: تمام مؤلفه‌های مورد نظر را در غالب جامعه بیان می‌کنیم.

مثال: میانگین سنی دانشجویان دانشگاه‌های علمی ـ کاربردی

نمونه: قسمتی از جامعه را به عنوان نمونه انتخاب می‌کنیم و مورد مطالعه قرار می‌دهیم.

در آمار توصیفی «جامعه» مورد مطالعه قرار می‌گیرد و در آمار تحلیلی «نمونه». درجه اطمینان نتیجه، در آمار توصیفی ۱۰۰٪ است اما در آمار تحلیلی نتیجه، به صورت احتمالی درنظر گرفته می‌شود.

ادبیات گفتگویی در هر دو آمار متفاوت است. به عنوان مثال: N، در آمار توصیفی، جامعه مدنظر است و n، در آمار تحلیلی مورد مطالعه قرار می‌گیرد به معنی نمونه.

مثال: نمونه $\leftarrow$ $n = ۴۰۰$ جامعه $\leftarrow$ $N = ۴۰۰$

انواع شاخص‌ها:

۱) شاخص‌های مرکزی ۲) شاخص‌های پراکندگی

شاخص‌های مرکزی:

الف) *میانگین :* ۱) میانگین حسابی ۲) میانگین هندسی ۳) میانگین هارمونیک

۱) میانگین حسابی: *فرمول میانگین حسابی:*

$$\boxed{\mu = \frac{\sum x_i}{N}}$$

$$\frac{۲ + ۴ + ۶ + ۸ + ۱۰}{۵} = ۶$$

چهار خاصیت میانگین حسابی:

خاصیت اول: اگر میانگین را از هر x_iها کم کنیم و جواب‌ها را با هم جمع بـزنیم نتیجـه صفر می‌شود.

فرمول خاصیت اول: $\boxed{\sum (x_i - \mu) = ۰}$

مثال:

$$\begin{array}{ccccc} ۲ & ۴ & ۶ & ۸ & ۱۰ \\ -۶ & -۶ & -۶ & -۶ & -۶ \\ \hline -۴+ & -۲+ & ۰+ & ۲+ & ۴=۰ \end{array}$$

خاصیت دوم: اگر x_iها را با یک عدد ثابت جمع کنیم، به میانگین هم همان عدد اضافه می‌شود. به بیان دیگر میانگین اعداد جدید می‌شود: میانگین اعداد قبلی به اضافه عدد ثابت.

فرمول خاصیت دوم: $\boxed{y_i = x_i + a \quad \Rightarrow \quad \mu_y = \mu_x + a}$

مثال :

$$\begin{array}{cccccc} x & ۲ & ۴ & ۶ & ۸ & ۱۰ \\ a & +۳ & +۳ & +۳ & +۳ & +۳ \\ \hline y & ۵ & ۷ & ۹ & ۱۱ & ۱۳ \end{array}$$

خاصیت سوم: اگر x_i ها را در یک عدد ثابت ضرب کنیم. میانگین اعداد جدید می‌شود: میانگین اعداد قبلی در همان عدد ثابت.

فرمول خاصیت سوم: $\boxed{y_i = ax_i \;\Rightarrow\; \mu_y = a\mu_x}$

مثال:

$$\begin{array}{cccccc} x & ۲ & ۴ & ۶ & ۸ & ۱۰ \\ a & \times ۲ & \times ۲ & \times ۲ & \times ۲ & \times ۲ \\ \hline y & ۴ & ۸ & ۱۲ & ۱۶ & ۲۰ \end{array}$$

خاصیت چهارم: اگر xi ها را با دسته‌ی yi ها جمع کنیم. میانگین جواب آن‌ها (zi ها) برابر است با جمع میانگین x و y.

فرمول خاصیت چهارم: $\boxed{z_i = x_i + y_i \;\Rightarrow\; \mu_z = \mu_x + \mu_y}$

مثال:

x	۲	۴	۶	۸	۱۰	$\mu = ۶$
y	۱	۳	۵	۷	۹	$\mu = ۵$
z	۳	۷	۱۱	۱۵	۱۹	$\mu = ۱۱$

***تمرین:**

$x_i = ۱\ \ ۳\ \ ۵\ \ ۷ \qquad \mu = ۴$

$y_i = ۱\ \ ۲\ \ ۳\ \ ۶ \qquad \mu = ۳$

$a = ۲$

خاصیت ۱ : $\sum(x_i - \mu) = ۰ \;\Rightarrow\; (-۳ + -۱ + ۱ + ۳) = ۰$

خاصیت ۲ : $y_i = xi + a \;\Rightarrow\; \mu_y = \mu_x + a$

۱	۲	۳	۶	$\mu = ۳$
+۲	۲	۲	۲	
۳	۴	۵	۸	$\mu = ۵$

خاصیت ۳ : $y_i = ax_i \;\Rightarrow\; \mu_y = a\,\mu_x$

۱	۲	۳	۶	$\mu = ۳$
×۲	۲	۲	۲	
۲	۴	۶	۱۲	$\mu = ۶$

خاصیت ۴: $z_i = x_i + y_i \;\Rightarrow\; \mu_z = \mu_x + \mu_y$

x	۱	۳	۵	۷	$\mu = ۴$
y	۱	۲	۳	۶	$\mu = ۳$
z	۲	۵	۸	۱۳	$\mu = ۷$

۲) میانگین هندسی: برای محاسبه‌ی اندازه‌های نسبی، درصدها، نرخ‌های رشد و شاخص‌ها از میانگین هندسی استفاده می‌کنیم. مثال:

نسبت سود شرکت زمزم در سال ۶۷ نسبت به سال ۶۶ مساوی ۳، سال ۶۸ نسبت به ۶۷ مساوی۲ و سال ۶۹ نسبت به ۶۸ مساوی ۴/۵ می‌باشد. میانگین سود شرکت زمزم چقدر است؟

❋کلمه‌ی نسبت در این گونه مسایل نشان می‌دهد که باید از میانگین هندسی استفاده کنیم.

فرمول:

$$\boxed{\mu_G = (x_۱ \times x_۲ \times x_۳ \times \ldots)^{\frac{۱}{n}}}$$

$$(۳ \times ۲ \times ۴/۵)^{\frac{۱}{۳}} = \sqrt[۳]{(۳ \times ۲ \times ۴/۵)} = ۳$$

۳) میانگین هارمونیک: زمانی که مقیاس ترکیبی‌ست، مانند کیلومتر در ساعت یا دور در ثانیه یا نفر ساعت، از مقیاس‌های ترکیبی باشند، از میانگینی به نام میانگین هارمونیک استفاده می‌کنیم.

فرمول:

$$\boxed{\mu_H = \frac{N}{\frac{۱}{x_۱} + \frac{۱}{x_۲} + \frac{۱}{x_۳} + \ldots}}$$

مثال: راننده‌ای مسافت یزد به تهران را با سرعت ۸۰ کیلومتر در ساعت و همین مسافت را با سرعت ۱۰۰ کیلومتر در ساعت برمی‌گردد. متوسط سرعتی که این راننده داشته است، چند کیلومتر است؟

$$\mu_H = \frac{N}{\frac{۱}{x_۱} + \frac{۱}{x_۲} + \frac{۱}{x_۳} + \ldots} = \frac{۲}{\frac{۱}{۸۰} + \frac{۱}{۱۰۰}} = ۸۸/۸۹$$

ب) مد: داده‌ای است که بیش از همه تکرار شده باشد.

$۲\ ۳\ ۲\ ۴\ ۵ \quad : \mathrm{mo} = ۲$

$۲\ ۳\ ۲\ ۴\ ۳\ ۵ : \mathrm{mo} = ۲,۳$

$۱\ ۲\ ۴\ ۵\ ۶ \quad : \mathrm{mo} = \phi$

$۲\ ۲\ ۴\ ۴\ ۵\ ۵ : \mathrm{mo} = \phi$

مثال:

$۲\ \ ۲\ \ ۵\ \ ۱۵\ \ ۱۰\ \ ۸\ \ ۵\ \ ۲\ : \ \mathrm{mo} = ۲$

$۸\ ۵\ ۲\ ۸\ ۵\ ۲\ \ ۸\ ۵\ ۲\ : \ \mathrm{mo} = \phi$

ج) چارک‌ها

فرمول چارک:

$$\boxed{\frac{a\,N}{۴}+\frac{۱}{۲}}$$

در فرمول بالا: a : چارک و N تعداد اعداد است.

مثال: اعداد ۸۵، ۱۴۰، ۱۲۰، ۸۰، ۹۰ و ۱۰۰. چارکهای اول و دوم و سوم را به دست آورید.

راه حل: ۱- ابتدا داده‌ها را مرتب می‌کنیم. ۲- سپس داده‌ها را از سمت چپ به ترتیب کدگذاری می‌کنیم. (از کد ۱ به بعد). ۳- محل چارک را طبق فرمول مورد محاسبه قرار می‌دهیم. ۴- مقدار چارک‌ها را تعیین می‌نماییم.

۸۰ – ۸۵ – ۹۰ – ۱۰۰ – ۱۲۰ – ۱۴۰
۱ ۲ ۳ ۴ ۵ ۶

$$Q_۱=\frac{۱(۶)}{۴}+\frac{۱}{۲}=۲ \Rightarrow ۸۵$$: چارک اول

$$Q_۲=\frac{۲\,(۶)}{۴}+\frac{۱}{۲}=۳/۵ \Rightarrow \frac{۹۰+۱۰۰}{۲}=۹۵$$: چارک دوم

$$Q_۳=\frac{۳(۶)}{۴}+\frac{۱}{۲}=۵ \Rightarrow ۱۲۰$$: چارک سوم

* نکته: چارک اول $\frac{۱}{۳}$ داده‌ها، چارک دوم، نصف داده‌ها و چارک سوم $\frac{۳}{۴}$ داده‌ها است.

مثال: با اضافه کردن عدد ۱۶۰ به مجموعه‌یس اعداد مذکور، سه چارک را محاسبه کنید.

۸۰ – ۸۵ – ۹۰ – ۱۰۰ – ۱۲۰ – ۱۴۰ – ۱۶۰
۱ ۲ ۳ ۴ ۵ ۶ ۷

$$Q_۱=\frac{۱(۷)}{۴}+\frac{۱}{۲}=۲/۲۵ \Rightarrow ۸۵+\frac{۱}{۴}(۹۰-۸۵)=۸۶/۲۵$$: چارک اول

$$Q_۲=\frac{۱(۷)}{۴}+\frac{۱}{۲}=۴ \Rightarrow ۱۰۰$$: چارک دوم

$$Q_۳=\frac{۳(۷)}{۴}+\frac{۱}{۲}=\frac{۲۱+۲}{۴}=\frac{۲۳}{۴}=۵/۷۵ \Rightarrow ۱۲۰+\frac{۳}{۴}(۱۴۰-۱۲۰)=۱۳۵$$: چارک سوم

شاخص‌های پراکندگی

۱) دامنه تغییرات (R) :

دامنه تغییرات = $\max - \min$

مثال: ۱ ۵ ۸ ۴ ۳

$R = ۸ - ۱ = ۷$

۲) واریانس و ۳) انحراف معیار

توضیح: چنانچه بخواهیم از خاصیت اول میانگین حسابی استفاده کنیم، به عدد صفر خواهیم رسید. برای این که از بنبست خلاص شویم، xi ها را منهای میانگین می‌کنیم و سپس جواب ها را به توان ۲ می‌رسانیم. پاسخ‌ها را با هم جمع کرده و تقسیم بر تعداد می‌کنیم. با این کار *واریانس* به دست می‌آید. حال اگر واریانس را زیر رادیکال قرار داده و از آن جذب بگیریم، *انحراف معیار* به دست می‌آید.

فرمول واریانس :
$$\boxed{\delta^{۲} = \frac{\sum (xi - \mu)^{۲}}{N}}$$

فرمول انحراف معیار:
$$\boxed{\delta = \sqrt{\frac{\sum (xi - \mu)^{۲}}{N}}}$$

مثال:

۱۵	۲۰	۲۵	۲۷	۱۳	۲۰
−۲۰	−۲۰	−۲۰	−۲۰	−۲۰	−۲۰
$(-۵)^{۲}$	$(۰)^{۲}$	$(۵)^{۲}$	$(۷)^{۲}$	$(-۷)^{۲}$	$(۰)^{۲}$

واریانس $(\delta^{۲})$: $۲۵ + ۰ + ۲۵ + ۴۹ + ۴۹ + ۰ = \frac{۱۴۸}{۶} = ۲۴/۶۶$

انحراف معیار (δ) : $\sqrt{۲۴/۶۶} = ۴/۹۶$

مثال: فرض کنید متوسط درآمد ۷ شهروندان اصفهانی و ۷ شهروند یزدی به شرح زیر باشد، مشخص نمایید که فاصله طبقاتی در شهر یزد بیشتر است یا اصفهان؟

متوسط درآمد شهروندان اصفهانی	۱	۲	۳	۴	۵	۶	۷	$\mu = ۴$
	-۴							
	-۳	-۲	-۱	۰	۱	۲	۳	$\mu = ۰$
واریانس	۹	۴	۱	۰	۱	۴	۹	$\mu = ۴$

$$\delta^{۲} = \frac{\sum (x_i - \mu)^{۲}}{N} = \frac{۲۸}{۷} = ۴$$ واریانس

$$\delta = \sqrt{\frac{\sum (xi - \mu)^{۲}}{N}} = \sqrt{۴} = ۲$$ انحراف معیار

متوسط درآمد شهروندان یزدی	۲	۳	۴	۵	۶	۷	۸	$\mu = ۵$
	-۵							
	-۳	-۲	-۱	۰	۱	۲	۳	$\mu = ۰$
واریانس	۹	۴	۱	۰	۱	۴	۹	$\mu = ۴$

$$\delta^{۲} = \frac{\sum (x_i - \mu)^{۲}}{N} = \frac{۲۸}{۷} = ۴$$ واریانس

$$\delta = \sqrt{\frac{\sum (xi - \mu)^{۲}}{N}} = \sqrt{۴} = ۲$$ انحراف معیار

نتیجه : *فاصله‌ی طبقاتی در اصفهان و یزد مساوی است.*

طبقه‌بندی داده‌ها:

۶ ۳ ۲ ۹ ۹ ۸ ۷ ۳ ۵ ۷ ۳ ۵ ۶ ۸ ۲ ۴ ۳ ۴ ۲ ۲

فراوانی تجمعی ↑	فراوانی تجمعی ↓	فراوانی نسبی fi	حد میانی xi	فراوانی مطلق Fi	C-L حدود طبقات
۲۰	۸	$\frac{۸}{۲۰}$	۳	۸	۲-۴
۱۲	۱۲	$\frac{۴}{۲۰}$	۵	۴	۴-۶
۸	۱۶	$\frac{۴}{۲۰}$	۷	۴	۶-۸
۴	۲۰	$\frac{۴}{۲۰}$	۹	۴	۸-۱۰

حد میانی = (حد بالا + حدپایین) ÷ ۲

فراوانی نسبی $(fi) = \frac{Fi}{\sum Fi}$

چه نسبتی از افراد نمره ۴-۲ گرفته‌اند؟

چندنفر از دانشجویان، نمره زیر ۸ گرفته‌اند؟

چندنفر از دانشجویان، نمره بالاتر از ۸ گرفته‌اند؟

چندنفر از دانشجویان، بالاتر از ۶ گرفته‌اند؟

متوسط درآمد شهروندان یزدی چه مبلغ است؟

* برای گرفتن میانگین شهروندان یزدی، به صورت میانگین حسابی، مشکلات زیر وجود دارد:

ـ نیاز به هزینه‌ی زیادی دارد. (نداریم)

ـ فرصت زیادی می‌طلبد. (نداریم)

ـ نیروی انسانی زیادی می‌خواهد. (نداریم)

ـ مشکلات فرهنگی بسیاری نیز سر راهمان قرار دارد.

بنابراین یک نمونه‌ی آماری در نظر می‌گیریم (از آن جامعه) و بر روی آن‌ها تحقیق انجام می‌دهیم و به وسیله‌ی *علم آمار* نتیجه را به کل جامعه تعمیم می‌دهیم.

در اینگونه مسائل یا جامعه نرمال است و یا غیرنرمال. که هر کدام باز ممکن است شرایط متفاوتی داشته باشند. یعنی در جامعه‌ی نرمال، یا انحراف معیار جامعه معلوم است و یا انحراف معیار جامعه معلوم نیست. زمانی که انحراف معیار جامعه معلوم نباشد ممکن است تعداد نمونه کمتر یا بیشتر از ۳۰ باشد، که هر کدام از این شرایط روش‌های متفاوتی برای حل مسأله دارند.

۱) میانگین جامعه نرمال، با انحراف معیار معلوم:

فرمول:

$$\boxed{P(\bar{x} - z\frac{\alpha}{2}.\delta\bar{x} \le \mu \le \bar{x} + z\frac{\alpha}{2}.\delta\bar{x}) = 1-\alpha}$$

$$\boxed{\delta\bar{x} = \frac{\delta x}{\sqrt{n}}}$$

مثال: محققی علاقه دارد، متوسط (میانگین) درآمد شهروندان یکی از مناطق تهران را برآورد نماید. با توجه به این که جامعه نرمال است و انحراف معیار آن ۳/۸ است، پس از محاسبات لازم، ۱۰ نفر از شهروندان را انتخاب نموده که میانگین درآمد آن‌ها ۱۰ می‌باشد. در سطح اطمینان ۹۹٪ میانگین درآمد شهروندان آن منطقه را برآورد نمایید.

میانگین جامعه $\mu = ?$ ، انحراف معیار جامعه $\delta x = ۳/۸$ ، میانگین نمونه $\bar{x} = ۱۰$ ، $n = ۱۰$

$$P(۱۰ - z\frac{۰/۰۱}{۲}.\frac{۳/۸}{\sqrt{۱۰}} \le \mu \le ۱۰ + z\frac{۰/۰۱}{۲}.\frac{۳/۸}{\sqrt{۱۰}}) = ۰/۹۹$$

$$P(۱۰ - ۲/۵۷۵ \times \frac{۳/۸}{\sqrt{۱۰}} \le \mu \le ۱۰ + ۲/۵۷۵ \times \frac{۳/۸}{\sqrt{۱۰}}) = ۰/۹۹$$

نتیجه : *به احتمال ۹۹٪ درآمد شهروندان ساکن در آن منطقه بین ۱۳/۱ تا ۶/۹ می‌باشد.*

مثال: دانشجویی می‌خواهد بداند میانگین (متوسط) نمرات دانشجویان دانشگاه علمی-کاربردی فرهنگ و هنر یزد چه قدر است. بدین منظور ۳۶ دانشجو انتخاب گردیده که میانگین نمرات آن‌ها ۱۲ می‌باشد. با توجه به این که انحراف معیار جامعه ۴ می‌باشد و همچنین جامعه نرمال است در سطح اطمینان ۹۵٪ نمرات دانشجویان را برآورد کنید.

$n = ۳۶$

$\overline{x} = ۱۲$

$\delta x = ۴$

$z\frac{\alpha}{۲} = z_{۰/۰۲۵} = ۱/۹۶$

$$P(\overline{x} - z\frac{\alpha}{۲}.\frac{\delta x}{\sqrt{n}} \le \mu \le \overline{x} + z\frac{\alpha}{۲}.\frac{\delta x}{\sqrt{n}}) = ۱ - \alpha$$

$$P(۱۲ - ۱/۹۶ \times ۰/۶۶ \le \mu \le ۱۲ + ۱/۹۶ \times ۰/۶۶) = ۰/۹۵$$

$$P(۱۰/۷۱ \le \mu \le ۱۳/۲۹) = ۰/۹۵$$

نتیجه : *به احتمال ۹۵٪ میانگین نمرات دانشجویان بین دو عدد ۱۰/۷۱ و ۱۳/۲۹ می‌باشد.*

مثال: دانشجویی می‌خواهد میانگین سن دانشجویان علمی-کاربردی را برآورد نماید. بدین منظور یک نمونه‌ی ۲۰ نفره از بین دانشجویان انتخاب نموده که میانگین سن آنها ۲۱ می‌باشد. با توجه به این که جامعه نرمال است و انحراف معیار آن ۳ می‌باشد، با ۹۰٪ اطمینان، سن دانشجویان را برآورد نمایید.

$n = ۲۰$

$\overline{x} = ۲۱$

$\delta x = ۳$

$\alpha = ۰/۱$

$z\frac{\alpha}{۲} = z_{۰/۰۵} = ۱/۶۴۵$

$\delta \overline{x} = \frac{\delta x}{\sqrt{n}} = \frac{۳}{\sqrt{۲۰}} = ۰/۶۷۱$

$$P(\overline{x} - z\frac{\alpha}{۲}.\frac{\delta x}{\sqrt{n}} \le \mu \le \overline{x} + z\frac{\alpha}{۲}.\frac{\delta x}{\sqrt{n}}) = ۱ - \alpha$$

$$P(۲۱ - z\frac{۰/۱}{۲}.\frac{۳}{\sqrt{۲۰}} \le \mu \le ۲۱ + z\frac{۰/۱}{۲}.\frac{۳}{\sqrt{۲۰}}) = ۰/۹۰$$

$$P(۲۱ - ۱/۶۴۵ \times ۰/۶۷۱ \le \mu \le ۲۱ + ۱/۶۴۵ \times ۰/۶۷۱) = ۰/۹$$

$$P(۱۹/۹ \le \mu \le ۲۲/۱) = ۰/۹$$

نتیجه : *به احتمال ۹۰٪ میانگین سن دانشجویان علمی کابردی بین دو عدد ۱۹/۹ و ۲۲/۱ می‌باشد.*

۲) میانگین جامعه نرمال، با انحراف معیار نامعلوم:

فرمول:

$$\boxed{P(\overline{x}-t_{\frac{\alpha}{2}^{n-1}}.S\overline{x}\le\mu\le\overline{x}+t_{\frac{\alpha}{2}^{n-1}}.S\overline{x})=1-\alpha}$$

$$\boxed{S\overline{x}=\frac{Sx}{\sqrt{n}}}$$

مثال: دانشجویی می‌خواهد میانگین سن دانشجوی علمی-کاربردی را برآورد نماید. بدین منظور یک نمونه‌ی ۲۰ نفره از بین دانشجویان انتخاب نموده که میانگین سن آن‌ها ۲۱ و انحراف معیار آن‌ها ۳ می‌باشد. با توجه به این که جامعه نرمال است، با اطمینان ۹۰٪، سن دانشجویان را برآورد نمایید.

$n=20$

$\overline{x}=21$

$Sx=3$

$t\frac{\alpha}{2}=t0/05=1/729$

$$P(\overline{x}-t\frac{\alpha}{2}.\frac{Sx}{\sqrt{n}}\le\mu\le\overline{x}+t\frac{\alpha}{2}.\frac{Sx}{\sqrt{n}})=1-\alpha$$

$$P(21-t\frac{0/1}{2}\times\frac{3}{\sqrt{20}}\le\mu\le21+t\frac{0/1}{2}\times\frac{3}{\sqrt{20}})=0/90$$

$$P(21-1/729\times0/671\le\mu\le21+1/729\times0/671)=0/90$$

$$P(19/84\le\mu\le22/16)=0/90$$

نتیجه : *به احتمال ۹۰٪ میانگین سن دانشجویان علمی کابردی بین دو عدد ۱۹/۸۴ و ۲۲/۱۶ می‌باشد.*

مثال: دانشجویی می‌خواهد بداند، میانگین درآمد شهروندان یزدی چه مبلغ است؟ بدین منظور از بین شهروندان نمونه‌ای به تعداد ۲۵ نفر انتخاب می‌نماید که میانگین درآمدی آن‌ها ۲۰ و انحراف معیار آن‌ها ۴ می‌باشد. با اطمینان ۹۹٪ و با توجه به اینکه جامعه نرمال است. میانگین درآمدی شهروندان یزدی را برآورد کنید.

$n=25$

$\overline{x}=20$

$Sx=4$

$t\frac{\alpha}{2}=t0/005=2/797$

$$P(\overline{x}-t\frac{\alpha}{2}.\frac{Sx}{\sqrt{n}}\le\mu\le\overline{x}+t\frac{\alpha}{2}.\frac{Sx}{\sqrt{n}})=1-\alpha$$

$$P(20-t\frac{0/01}{2}\times\frac{4}{\sqrt{25}}\le\mu\le20+t\frac{0/01}{2}\times\frac{4}{\sqrt{25}})=0/99$$

$$P(20-2/797\times0/8\le\mu\le20+2/797\times0/8)=0/99$$

$$P(17/763\le\mu\le22/237)=0/99$$

نتیجه : *با ۹۹٪ اطمینان، درآمد شهروندان یزدی بین ۱۷/۷۶۳ و ۲۲/۲۳۷ می‌باشد.*

*** نکته:** *اگر تعداد نمونه در مسأله بیشتر از ۳۰ نفر باشد می‌توان از فرمول زیر نیز برای حل آن استفاده کرد:*

$$\boxed{P(\bar{x} - z\frac{\alpha}{۲} \times \frac{Sx}{\sqrt{n}} \leq \mu \leq \bar{x} + z\frac{\alpha}{۲} \times \frac{Sx}{\sqrt{n}}) = ۱ - \alpha}$$

مثال: دانشجویی می‌خواهد بداند، میانگین درآمد شهروندان یزدی چه مبلغ است؟ بدین منظور از بین شهروندان نمونه‌ای به تعداد ۳۵ نفر انتخاب می‌نماید که میانگین درآمدی آن‌ها ۲۰ و انحراف معیار آن‌ها ۴ می‌باشد. با اطمینان ۹۹٪ و با توجه به اینکه جامعه نرمال است. میانگین درآمدی شهروندان یزدی را برآورد کنید.

$n = ۳۵$

$\bar{x} = ۲۰$

$Sx = ۴$

$z\frac{\alpha}{۲} = z_{۰/۰۰۵} = ۲/۵۷۵$

$$P(\bar{x} - z\frac{\alpha}{۲}.\frac{Sx}{\sqrt{n}} \leq \mu \leq \bar{x} + z\frac{\alpha}{۲}.\frac{Sx}{\sqrt{n}}) = ۱ - \alpha$$

$$P(۲۰ - z\frac{۰/۰۱}{۲} \times \frac{۴}{\sqrt{۳۵}} \leq \mu \leq ۲۰ + z\frac{۰/۰۱}{۲} \times \frac{۴}{\sqrt{۳۵}}) = ۰/۹۹$$

$$P(۲۰ - ۲/۵۷۵ \times ۰/۶۷۶ \leq \mu \leq ۲۰ + ۲/۵۷۵ \times ۰/۶۷۶) = ۰/۹۹$$

$$P(۱۸/۲۷۵ \leq \mu \leq ۲۱/۷۲۵) = ۰/۹۹$$

نتیجه : *با ۹۹٪ اطمینان، درآمد شهروندان یزدی بین ۱۸/۲۷۵ و ۲۱/۷۲۵ می‌باشد.*

۳) میانگین جامعه غیر نرمال، با انحراف معیار نامعلوم:

فرمول :

$$\boxed{P(\overline{x}-KS\overline{x}\leq\mu\leq\overline{x}+KS\overline{x})\geq ۱-\alpha}$$

$$\boxed{K=\sqrt{\frac{۱}{\alpha}}} \quad \text{و} \quad \boxed{S\overline{x}=\frac{Sx}{\sqrt{n}}}$$

مثال: دانشجویی می‌خواهد بداند μ درآمد شهروندان یزدی چه مبلغ است. بدین منظـور از بین شهروندان، نمونه‌ای به تعداد ۲۵ نفر انتخاب می‌نمایـد کـه میـانگین درآمـدی آن‌هـا ۲۰ و انحراف معیار آن‌ها ۴ می‌باشد با اطمینان ۹۹٪ و با توجه به اینکه دلیلی بر نرمال بـودن جامعـه نیست. میانگین درآمدی شهروندان یزدی را برآورد نمایید.

$n=۲۵$

$\overline{x}=۲۰$

$Sx=۴$

$\alpha=۰/۰۱$

$$P(\overline{x}-\sqrt{\frac{۱}{\alpha}}\times\frac{Sx}{\sqrt{n}}\leq\mu\leq\overline{x}+\sqrt{\frac{۱}{\alpha}}\times\frac{Sx}{\sqrt{n}})\geq ۱-\alpha$$

$$P(۲۰-\sqrt{\frac{۱}{۰/۰۱}}\times\frac{۴}{\sqrt{۲۵}}\leq\mu\leq ۲۰+\sqrt{\frac{۱}{۰/۰۱}}\times\frac{۴}{\sqrt{۲۵}})\geq ۰/۹۹$$

$$P(۲۰-۱۰\times ۰/۸\leq\mu\leq ۲۰+۱۰\times ۰/۸)\geq ۰/۹۹$$

$$P(۱۲\leq\mu\leq ۲۸)\geq ۰/۹۹$$

نتیجه : *با ۹۹٪ اطمینان، میانگین درآمد شهروندان یزدی بین ۱۲ و ۲۸ می‌باشد.*

۴) میانگین جامعه <u>غیر نرمال</u>، با انحراف معیار <u>معلوم</u>:

فرمول :

$$\boxed{P(\bar{x} - K\delta\bar{x} \le \mu \le \bar{x} + K\delta\bar{x}) \ge ۱ - \alpha}$$

$$\boxed{K = \sqrt{\frac{۱}{\alpha}}} \quad \text{و} \quad \boxed{\delta\bar{x} = \frac{\delta x}{\sqrt{n}}}$$

مثال: دانشجویی می‌خواهد بداند میانگین درآمد شهروندان یـزدی چـه مبلـغ اسـت. بـدین منظور از بین شهروندان نمونه‌ای به تعداد ۲۵ نفر انتخاب می‌نماید که میانگین درآمـدی آن‌هـا ۲۰ می‌باشد. همچنین با توجه به این که انحراف معیار جامعه ۴ می‌باشد با اطمینـان ۹۹٪ و بـا توجه به این که دلیلی بر نرمال بودن جامعه نیست، میانگین درآمدی شهروندان یزدی را برآورد نمایید.

$n = ۲۵$
$\bar{x} = ۲۰$
$\delta x = ۴$
$\alpha = ۰/۰۱$

$$P(\bar{x} - \sqrt{\frac{۱}{\alpha}} \times \frac{\delta x}{\sqrt{n}} \le \mu \le \bar{x} + \sqrt{\frac{۱}{\alpha}} \times \frac{\delta x}{\sqrt{n}}) \ge ۱ - \alpha$$

$$P(۲۰ - \sqrt{\frac{۱}{۰/۰۱}} \times \frac{۴}{\sqrt{۲۵}} \le \mu \le ۲۰ + \sqrt{\frac{۱}{۰/۰۱}} \times \frac{۴}{\sqrt{۲۵}}) \ge ۰/۹۹$$

$$P(۲۰ - ۱۰ \times ۰/۸ \le \mu \le ۲۰ + ۱۰ \times ۰/۸) \ge ۰/۹۹$$

$$P(۱۲ \le \mu \le ۲۸) \ge ۰/۹۹$$

نتیجه : *با ۹۹٪ اطمینان، میانگین درآمد شهروندان یزدی بین ۱۲ و ۲۸ می‌باشد.*

مقایسه میانگین دو جامعه آماری در شرایطی که انحراف معیار معلوم است:

فرمول:

$$\boxed{P[(\bar{x}_۱-\bar{x}_۲)-z\frac{\alpha}{۲}\times\delta\,\bar{x}-\bar{x}_۱\le\mu_۱-\mu_۲\le(\bar{x}_۱-\bar{x}_۲)+z\frac{\alpha}{۲}\times\delta\,\bar{x}-\bar{x}_۱)=۱-\alpha}$$

$$\boxed{\delta\,\bar{x}-\bar{x}=\sqrt{\frac{\delta_۱^۲}{n_۱}+\frac{\delta_۲^۲}{n_۲}}}$$

نکات:

۱- اگر هر دو دامنه مثبت باشد، در سطح اطمینان مورد نظر، $\mu_۱ > \mu_۲$ است.

۲- اگر هر دو دامنه منفی باشد، در سطح اطمینان مورد نظر، $\mu_۱ < \mu_۲$ است.

۳- در غیر از موارد فوق، بین $\mu_۱$ و $\mu_۲$ اختلاف معناداری مشاهده نمی‌شود.

مثال: دانشجویی می‌خواهد بداند دانشجویان علمی- کاربردی یزد جوان‌ترند و یا دانشجویان علمی کاربردی تهران. بدین منظور از بین یزدی‌ها ۲۵ نفر و از تهران‌ها ۱۶ نفر را انتخاب کرد که میانگین آن‌ها به ترتیب ۲۴ و ۲۱ می‌باشد. با فرض نرمال بودن جامعه و با توجه به اینکه انحراف معیار دو جامعه به ترتیب ۲ و ۴ می‌باشد و با اطمینان ۹۰٪ مشخص کنید. کدام دانشگاه دانشجویانش جوان‌ترند؟

$$P[(\bar{x}_۱-\bar{x}_۲)-z\frac{\alpha}{۲}\times\sqrt{\frac{\delta_۱^۲}{n_۱}+\frac{\delta_۲^۲}{n_۲}}\le\mu_۱-\mu_۲\le(\bar{x}_۱-\bar{x}_۲)+z\frac{\alpha}{۲}\times\sqrt{\frac{\delta_۱^۲}{n_۱}+\frac{\delta_۲^۲}{n_۲}})=۱-\alpha$$

یزدی	تهرانی
$n_۱=۲۵$	$n_۲=۱۶$
$\bar{x}_۱=۲۴$	$\bar{x}_۲=۲۱$
$\delta x_۱=۲$	$\delta x_۲=۴$
$\alpha=۰/۱۰$	$\alpha=۰/۱۰$

$$P[(۲۴-۲۱)-z\frac{۰/۱}{۲}\sqrt{\frac{۴}{۲۵}+\frac{۱۶}{۱۶}}\le\mu_۱-\mu_۲\le(۲۴-۲۱)+z\frac{۰/۱}{۲}\sqrt{\frac{۴}{۲۵}+\frac{۱۶}{۱۶}}]=۰/۹۰$$

$$P[۳-۱/۶۴۵\times۱/۰۷\le\mu_۱-\mu_۲\le۳+۱/۶۴۵\times۱/۰۷]=۰/۹۰$$

$$P[۱/۲۳\le\mu_۱-\mu_۲\le۴/۷۶]=۰/۹۰$$

نتیجه : *به لحاظ این که دامنه اول در مثال قبل مثبت می‌باشد (۱/۲۴) و همچنین دامنه دوم نیز مثبت می‌باشد (۴/۷۶)، لذا نتیجه می‌گیریم که به احتمال ۹۰٪ سن دانشجویان علمی-کاربردی یزد بیشتر از دانشگاه علمی-کاربردی تهران است.*

مثال: هدف از این تحقیق، مقایسه کارمندان دو سازمان الف و ب می‌باشد. بـدین منظـور یـک نمونه تصادفی ۲۵ نفره از الف و ۲۰ نفره از ب انتخاب نموده‌ایم که میانگین آن‌ها بـه ترتیـب ۶۰ و ۵۵ می‌باشد. با توجه به این که انحراف معیار آن‌ها (جامعه) به ترتیب ۱۰ و ۱۲ مـی‌باشـد، در سـطح اطمینان ۹۹٪ عملکرد کارمندان آن دو سازمان را با یکدیگر مقایسه کنید. (جامعه نرمال می‌باشد.)

$$P[(\bar{x}_۱-\bar{x}_۲)-z\frac{\alpha}{۲}\times\sqrt{\frac{\delta_۱^۲}{n_۱}+\frac{\delta_۲^۲}{n_۲}}\le\mu_۱-\mu_۲\le(\bar{x}_۱-\bar{x}_۲)+z\frac{\alpha}{۲}\times\sqrt{\frac{\delta_۱^۲}{n_۱}+\frac{\delta_۲^۲}{n_۲}})=۱-\alpha$$

$$P[(۶۰-۵۵)-۲/۵۷۵\times\sqrt{\frac{۱۰^۲}{۲۵}+\frac{۱۲^۲}{۲۰}}\le\mu_۱-\mu_۲\le(۶۰-۵۵)+۲/۵۷۵\times\sqrt{\frac{۱۰^۲}{۲۵}+\frac{۱۲^۲}{۲۰}}]=۰/۹۹$$

الف	ب
$n_۱=۲۵$	$n_۲=۲۰$
$\bar{x}_۱=۶۰$	$\bar{x}_۲=۵۵$
$\delta x_۱=۱۰$	$\delta x_۲=۱۲$
$\alpha=۰/۰۱$	$\alpha=۰/۰۱$

$$P[۵-۲/۵۷۵\times۳/۳۴۶\le\mu_۱-\mu_۲\le۵+۲/۵۷۵\times۳/۳۴۶]=۰/۹۹$$

$$P[۵-۸/۶۱۵\le\mu_۱-\mu_۲\le۵+۸/۶۱۵]=۰/۹۹$$

$$P[-۳/۶۱۵\le\mu_۱-\mu_۲\le۱۳/۶۱۵]=۰/۹۹$$

نتیجه : *به لحاظ اینکه یک دامنه منفی شده است (۳/۶۱۵-) و دامنه دیگر مثبت (۱۳/۶۱۵). بنابراین می‌توانیم چنین نتیجه بگیریم که به احتمال ۹۹٪ تفاوت معنی‌داری بین عملکرد کارمندان این دو سازمان وجود ندارد.*

مثال: فرض کنید از بین ساکنین (جامعه) یزد ۱۶ نفر انتخاب، که درآمد ماهیانه آنان، ۴۴ می‌باشد. و همچنین از بین شهروندان اصفهانی ۲۴ نفر انتخاب نموده‌ایم که میانگین درآمدی آنها ۳۶ می‌باشد. با توجه به نرمال بودن هر دو جامعه و انحراف معیار آنها که به ترتیب ۲ و ۳ می‌باشد به احتمال ۹۵٪ مشخص کنید شهروندان یزدی، پولدارترند یا اصفهانی؟

$$P[(۴۴-۳۶)-۱/۹۶\times\sqrt{\frac{۲^۲}{۱۶}+\frac{۳^۲}{۲۴}}\le\mu_۱-\mu_۲\le(۴۴-۳۴)+۱/۹۶\times\sqrt{\frac{۲^۲}{۱۶}+\frac{۳^۲}{۲۴}}]=۰/۹۵$$

یزدی	اصفهانی
$n_۱=۱۶$	$n_۲=۲۴$
$\bar{x}_۱=۴۴$	$\bar{x}_۲=۳۶$
$\delta x_۱=۲$	$\delta x_۲=۳$
$\alpha=۰/۰۵$	$\alpha=۰/۰۵$

$$P[۸-۱/۲۲۵\le\mu_۱-\mu_۲\le۸+۱/۲۲۵]=۰/۹۵$$

$$P[۶/۷۷۵\le\mu_۱-\mu_۲\le۹/۲۲۵]=۰/۹۵$$

نتیجه: *به لحاظ اینکه هر دو دامنه مثبت است، به احتمال ۹۵٪ شهروندان یزدی پولدارتر از شهروندان اصفهانی هستند.*

مقایسه میانگین دو جامعه آماری در شرایطی که انحراف معیار نامعلوم است:

فرمول:

$$\boxed{P[(\bar{x}_۱-\bar{x}_۲)-t\frac{\alpha}{۲}df.SP\sqrt{\frac{۱}{n_۱}+\frac{۱}{n_۲}}\le\mu_۱-\mu_۲\le(\bar{x}_۱-\bar{x}_۲)+t\frac{\alpha}{۲}df.SP\sqrt{\frac{۱}{n_۱}+\frac{۱}{n_۲}}]=۱-\alpha}$$

$$\boxed{SP=\sqrt{\frac{(n_۱-۱)S_۱^۲+(n_۲-۱)S_۲^۲}{n_۱+n_۲-۲}}} \quad ، \quad \boxed{df=n_۱+n_۲-۲}$$

مثال: هدف از این تحقیق، مقایسه سطح کارمندان در سازمان الف با سازمان ب است. در این تحقیق از سازمان الف، یک نمونه ۹ نفره انتخاب شده که میانگین آن ۴۵ و انحراف معیارش ۱۲ است. در حالیکه میانگین و انحراف معیار سطح آمادگی کارمندان سازمان ب در یک نمونه ۱۵ نفره به ترتیب ۵۵ و ۱۴ می‌باشد. فرض کنید توزیع نمره‌ها در دو سازمان نرمال باشد، در سطح اطمینان ۹۰٪ تخمین لازم برای مقایسه میانگین دو جامعه به عمل آورید. (با فرض: $\delta_۱^۲=\delta_۲^۲$)

یزدی	اصفهانی
$n_۱=۹$	$n_۲=۱۵$
$\bar{x}_۱=۴۵$	$\bar{x}_۲=۵۵$
$Sx_۱=۱۲$	$Sx_۲=۱۴$
$\alpha=۰/۱$	$\alpha=۰/۱$

$$SP=\sqrt{\frac{(۹-۱)۱۲^۲+(۱۵-۱)۱۴^۲}{۹+۱۵-۲}}=\sqrt{\frac{۱۱۵۲+۲۷۴۴}{۲۲}}=۱۳/۳$$

$$P[(۴۵-۵۵)-t\frac{۰/۰۱}{۲}\times۱۳/۳\times\sqrt{\frac{۱}{۹}+\frac{۱}{۱۵}}\le\mu_۱-\mu_۲\le(۴۵-۵۵)+t\frac{۰/۰۱}{۲}\times۱۳/۳\times\sqrt{\frac{۱}{۹}+\frac{۱}{۱۵}}]=۰/۹۰$$

$$P[-۱۰-۱/۷۱۷\times۳/۱۳\times۰/۴۲\le\mu_۱-\mu_۲\le-۱۰+۱/۷۱۷\times۳/۱۳\times۰/۴۲]=۰/۹۰$$

$$P[-۱۰-۹/۵۹۱\le\mu_۱-\mu_۲\le-۱۰+۹/۵۹۱]=۰/۹۰$$

$$P[-۱۹/۵۹۱\le\mu_۱-\mu_۲\le-۰/۴۰۹]=۰/۹۰$$

نتیجه: *چون هر دو دامنه منفی است، پس در خطای ۱۰٪ (یا سطح اطمینان ۹۰٪) می‌توان نتیجه گرفت، که $\mu_۱<\mu_۲$ است. یعنی میانگین عملکرد کارمندان در سازمان الف کمتر از میانگین عملکرد کارمندن سازمان ب است.*

مثال: هدف از این تحقیق، مقایسه سن دانشجویان دو دانشگاه الف و ب است. از دانشگاه الف یک نمونه ۸ نفره انتخاب که میانگین آنها ۲۱ و انحراف معیارش ۴ است. در حالی که میانگین و انحراف معیار سن دانشجویان دانشگاه ب در یک نمونه ۹ نفره به ترتیب ۲۴ و ۳ می‌باشد. فرض کنید دو جامعه نرمال و واریانس هر جامعه با هم برابر است. در سطح اطمینان ۹۵٪ میانگین دو جامعه را مقایسه کنید.

$$P[(\bar{x}_۱-\bar{x}_۲)-t\frac{\alpha}{۲}df.SP\sqrt{\frac{۱}{n_۱}+\frac{۱}{n_۲}}\le\mu_۱-\mu_۲\le(\bar{x}_۱-\bar{x}_۲)+t\frac{\alpha}{۲}df.SP\sqrt{\frac{۱}{n_۱}+\frac{۱}{n_۲}}]=۱-\alpha$$

$$df=n_۱+n_۲-۲=۸+۹-۲=۱۵\Rightarrow t\frac{\alpha}{۲},df=۲/۱۳۱$$

$$P[(۲۱-۲۴)-۲/۱۳\times۳/۵\times۰/۴۸۵\le\mu_۱-\mu_۲\le(۲۱-۲۴)+۲/۱۳۱\times۳/۵\times۰/۴۸۵]=۰/۹۵$$

$$P[-۳-۳/۶۱۷\le\mu_۱-\mu_۲\le-۳+۳/۶۱۷]=۰/۹۵$$

$$P[-۶/۶۱۷\le\mu_۱-\mu_۲\le+۰/۶۱۷]=۰/۹۵$$

نتیجه: *چون یک دامنه منفی شده است (۶/۶۱۷-) و دامنه دیگر مثبت شده است (۰/۶۱۷+) لذا نتیجه می‌گیریم که تفاوت معنی‌داری بین سن دانشجویان در دو دانشگاه الف و ب وجود ندارد.*

*** نکته: *اگر*** $n_۱+n_۲-۲\ge۳۰$ ***شد، می‌توان از فرمول زیر استفاده کرد:***

$$\boxed{P[(\bar{x}_۱-\bar{x}_۲)-z\frac{\alpha}{۲}.\sqrt{\frac{S_۱^۲}{n_۱}+\frac{S_۲^۲}{n_۲}}\le\mu_۱-\mu_۲\le(\bar{x}_۱-\bar{x}_۲)+z\frac{\alpha}{۲}.\sqrt{\frac{S_۱^۲}{n_۱}+\frac{S_۲^۲}{n_۲}}]=۱-\alpha}$$

مثال: مسأله ص ۱۷:

یزدی	اصفهانی
$n_۱=۹$	$n_۲=۱۵$
$\bar{x}_۱=۴۵$	$\bar{x}_۲=۵۵$
$Sx_۱=۱۲$	$Sx_۲=۱۴$
$\alpha=۰/۱$	$\alpha=۰/۱$

$$z\frac{\alpha}{۲}=\frac{۰/۱}{۲}=۰/۰۵\Rightarrow۱/۶۴۵$$

$$P[(۴۵-۵۵)-۱/۶۴۵\times\sqrt{\frac{۱۴۴}{۹}+\frac{۱۹۶}{۱۵}}\le\mu_۱-\mu_۲\le(۴۵-۵۵)-۱/۶۴۵\times\sqrt{\frac{۱۴۴}{۹}+\frac{۱۹۶}{۱۵}}]=۰/۹۰$$

$$P[(-۱۰-۱/۶۴۵\times۴/۵۴\le\mu_۱-\mu_۲\le-۱۰+۱/۶۴۵\times۴/۵۴]=۰/۹۰$$

$$P[-۱۷/۴۶\le\mu_۱-\mu_۲\le-۲/۵۳]=۰/۹۰$$

نتیجه: *به لحاظ اینکه هر دو دامنه منفی شده است، نتیجه می‌گیریم که عملکرد سازمان الف ضعیفتر از عملکرد سازمان ب است.*

تعیین حجم نمونه

فرمول:

$$\boxed{n = \frac{(z\frac{\alpha}{۲})^{۲}.\delta x^{۲}}{\sum^{۲}}}$$

مثال: فرض کنید یکی از محققان مایل است، مطالعه‌ای به منظور تعیین میـانگین رشـد کـاری کارمندان سازمانی را انجام دهد، در این مطالعه او دقت بـرآورد را ۵ نمـره در نظـر گرفتـه و تـصور می‌کند که انحراف معیار نمره‌های رشد کاری کارمندان ۲۰ نمره باشد، اندازه‌ی نمونه‌ی لازم را برای بررسی نهایی در سطح خطای ۵٪ محاسبه کنید.

$z\frac{۰/۰۵}{۲} = ۱/۹۶$

$\delta x = ۲۰$

$\Sigma = ۵$

$$n = \frac{۱/۹۶^{۲} \times ۲۰^{۲}}{۵^{۲}} = ۶۱/۴۴ \approx ۶۲$$ به سمت عدد بالا گرد می‌کنیم

** اگر حجم نمونه‌ی محاسبه شده، بزرگتر یا مساوی ۳۰ بود، فرایند حجم نمونه را متوقف می‌نماییم و به همین عدد به دست آمده اکتفا می‌کنیم.*

مثال: بررسی مقدماتی حاصل از ۴ مشتری، نشان می‌دهد که انحراف معیار زمان اشتغال یک فروشنده در مغازه‌ای ۷۲٪ است. هدف از این تحقیق تخمین درصد زمان اشتغال با دقت برآورد ۰/۰۴+ و با احتمال ۹۵ درصد است. حجم نمونه موردنظر را برآورد نمایید.

$z\frac{\alpha}{۲} = ۱/۹۶$

$$n = \frac{۱/۹۶^{۲} \times ۰/۰۷۲}{۰/۰۴} = \frac{۰/۰۱۹۵}{۰/۰۰۱۶} = ۱۲/۴۴ \approx ۱۳$$

	۱۳	۱۴	۱۵
$t = ۰/۰۲۵,۱۲$ $t = ۰/۰۲۵,۱۳$ $t = ۰/۰۲۵,۱۴$	۱/۱۷۹	۲/۱۶۰	۲/۱۴۵
$(\frac{t\frac{\alpha}{۲}df.\delta x}{\sum})^{۲}$	۱۵/۳۶	۱۵/۰۵	۱۴/۹

** عددی که در پایین جدول به دست می‌آید (عدد واریانس) باید از عددی که در بالای جدول وجود دارد، کوچکتر باشد. اگر اینگونه نبود باید هر بار یک عدد به عدد بالای جدول اضافه کنیم و طبق فرمول واریانس را به دست آوریم تا جایی که عدد واریانس از عدد بالای جدول کوچکتر شود. (در این مسأله ۱۴/۹ را به سمت عدد بالا گرد می‌کنیم، یعنی حجم نمونه ۱۵ به دست می‌آید.)*

مثال: هدف از یک تحقیق، برآورد میانگین تجربه دبیران شهر تهران است. بدین منظور یک نمونه تصادفی ۱۰ نفره از بین دبیران انتخاب شده که پراکندگی تجربه‌شان ۲/۵ سال است. اگر دقت برآورد را ۱ سال در نظر بگیریم در سطح اطمینان ۹۵٪ حجم نمونه را تعیین کنید.

$$z\frac{\alpha}{۲} = ۱/۹۶ \qquad n = \frac{(z\frac{\alpha}{۲})^{۲}.\delta\, x^{۲}}{\sum^{۲}} = \frac{۱/۹۶ \times ۲/۵}{۱} = ۲۴/۰۱ \approx ۲۵$$

	۲۵	۲۶	۲۷
$t = ۰/۰۲۵, ۲۵$ $t = ۰/۰۲۵, ۲۶$ $t = ۰/۰۲۵, ۲۷$	۲/۰۶۴	۲/۰۶	۲/۰۵۶
$(\frac{t\frac{\alpha}{۲} df.\delta\, x}{\sum})^{۲}$	۲۶/۶۲	۲۶/۵۲	۲۶/۴۱

حجم نمونه را ۲۷ می‌گیریم.

مثال: مدت زمان مشتریان در یک فروشگاه بزرگ از توزیع نرمال برخوردار است. از مشتریان ۱۰ نفر، به طور تصادفی انتخاب شده که پس از محاسبات لازم پراکندگی زمان ماندن مشتریان ۶ دقیقه محاسبه گردیده است. اگر دقت برآورد را ۳ دقیقه در نظر بگیریم با سطح اطمینان ۹۵ درصد حجم نمونه را محاسبه کنید.

$$z\frac{\alpha}{۲} = ۱/۹۶ \qquad n = (\frac{z\frac{\alpha}{۲} \times \delta x}{\Sigma})^{۲} = (\frac{۱/۹۶ \times ۶}{۳})^{۲} = ۱۵/۳۶ \approx ۱۶$$

	۱۶	۱۷	۱۸
$t = ۰/۰۲۵, ۱۶$ $t = ۰/۰۲۵, ۱۷$ $t = ۰/۰۲۵, ۱۸$	۲/۱۳۱	۲/۱۲۰	۲/۱۱۰
$(\frac{t\frac{\alpha}{۲} df.\delta\, x}{\sum})^{۲}$	۱۸/۱۶	۱۷/۹۷	۱۷/۸۰

حجم نمونه را ۱۸ می‌گیریم.

* *در مسأله، ریسک، پراکندگی، تفاضل و فاصله طبقاتی هم‌معنی واریانس است.*

از روی واریانس نمونه

$$\boxed{P\left[\frac{(n-۱)S^۲}{K^۲\,\frac{\alpha}{۲}\,df}\le \delta^۲ \le \frac{(n-۱)S^۲}{K^۲ ۱-\frac{\alpha}{۲}\,df}\right]=۱-\alpha}$$

مثال: یکی از کارگزاران بازار بورس تهران، در صدد تعیین ریسک شرکتهایی است که سهام خود را عرضه می‌کنند. در این زمینه یک نمونه‌ی تصادفی ۱۵ تایی از بین شرکت‌ها انتخاب کرده که میانگین و واریانس سود سالانه آن‌ها به ترتیب ۲۵,۰۰۰ و ۱۲۲۵ است. توزیع سود سالانه‌ی شرکت‌ها از تقریب نرمال برخوردار است. در سطح خطای ۵٪ حدود اطمینان ریسک شرکت‌ها را به دست آورید.

$$n=۱۵ \qquad \bar{x}=۲۵۰۰۰ \qquad S^۲x=۱۲۲۵ \qquad \alpha=۰/۰۵$$

$$P\left[\frac{(n-۱)S^۲}{K^۲\,\frac{\alpha}{۲}\,df}\le \delta^۲ \le \frac{(n-۱)S^۲}{K^۲ ۱-\frac{\alpha}{۲}\,df}\right]=۱-\alpha$$

$$P\left[\frac{(۱۵-۱)۱۲۲۵}{K^۲\,\frac{۰/۰۵}{۲}\,۱۴}\le \delta^۲ \le \frac{(۱۵-۱)۱۲۲۵}{K^۲ ۰/۹۷۵\times ۱۴}\right]=۰/۹۵$$

$$P\left[\frac{۱۴(۱۲۲۵)}{۲۶/۱۱۹۰}\le \delta^۲ \le \frac{۱۴(۱۲۲۵)}{۵/۶۲۸۷۲}\right]=۰/۹۵$$

$$P[۶۵۶/۶۱۵ \le \delta^۲ \le ۳۰۵۱/۶۰]=۰/۹۵$$

نتیجه: *به احتمال ۹۵٪ ریسک یا واریانس داخلی بورس ما به این دو ختم می‌شود.*

نکته:

* *برای به دست آوردن میانگین جامعه از روی نمونه در شرایطی که انحراف معیار معلوم است: از جدول z استفاده می‌کنیم.*
* *برای به دست آوردن میانگین جامعه از روی نمونه در شرایطی که انحراف معیار معلوم نیست: از جدول t استفاده می‌کنیم.*
* *برای به دست آوردن واریانس جامعه از روی واریانس نمونه:* $x^۲=k^۲$.

مثال: محققی درصدد تعیین واریانس جامعه‌ای مورد بررسی تحقیق خود می‌باشند. بدین منظور یک نمونه تصادفی ۱۰ تایی انتخاب نموده که واریانس و میانگین آن به ترتیب ۹ و ۱۴ می‌باشد. با اطمینان ۹۰٪ حدود اطمینان واریانس جامعه مورد نظر را مشخص کنید.

$n=۱۰$
$\bar{x}=۱۴$
$S^۲x=۱۲۲۵$
$\alpha=۰/۱$

$$P\left[\frac{(۱۰-۱)۹}{۱۶/۹۱۹۰} \le \delta^۲ \le \frac{(۱۰-۱)۹}{۳/۳۲۵۱۱}\right]=۰/۹۰$$

$$P\left[۴/۷۹ \le \delta^۲ \le ۲۴/۳۹\right]=۰/۹۰$$

مثال: هدف از این تحقیق مقایسه پراکندگی نمره‌های دانشجویان در دو دانشگاه الف و ب است. در این تحقیق از دانشکده الف یک نمونه تصادفی ۱۶ تایی انتخاب شده که میانگین آن ۱۴ و واریانس آن ۱۶ است در حالی که میانگین یک نمونه تصادفی ۱۰ تایی از دانشجویان دانشکده ب به ترتیب ۱۵ و ۱۲ بوده است در سطح اطمینان ۹۰٪ پراکندگی نمره‌ها را در دو دانشکده مقایسه کنید.

ب	الف
$n_۲=۱۰$	$n_۱=۱۶$
$\bar{x}_۲=۱۵$	$\bar{x}_۱=۱۴$
$S_۲^۲=۱۲$	$S_۱^۲=۱۶$

$$\boxed{P\left[\frac{\frac{S_۱^۲}{S_۲^۲}}{F\frac{\alpha}{۲},df,df_۲} \le \frac{\delta_۱^۲}{\delta_۲^۲} \le \frac{S_۱^۲}{S_۲^۲}\times F\frac{\alpha}{۲},df_۲,df_۱\right]=۱-\alpha}$$

$$C=f\,\frac{۰/۱}{۲}.۱۵/۹=۳/۰۱ \qquad , \qquad D=f\,\frac{۰/۰۱}{۲}.۹/۱۵=۲/۵۹$$

$$P\left[\frac{\frac{۱۶}{۱۲}}{۳/۰۱} \le \frac{\delta_۱^۲}{\delta_۲^۲} \le \frac{۱۶}{۱۲}\times ۲/۵۹\right]=۰/۹۰ \quad \Rightarrow \quad P\left[۰/۴۴ \le \frac{\delta_۱^۲}{\delta_۲^۲} \le ۳/۴۵\right]=۰/۹۰$$

۱- اگر هر دو دامنه بزرگتر از یک باشد در سطح اطمینان مورد نظر می‌توان گفت که واریانس جامعه اول بزرگتر از واریانس جامعه دوم است.

۲- اگر هر دو دامنه کوچکتر از یک باشد در سطح اطمینان مورد نظر می‌توان گفت که واریانس جامعه اول کوچکتر از واریانس جامعه دوم است.

۳- در حالتی غیر از موارد فوق یعنی این که یک دامنه کوچکتر از یک و دامنه دوم بزرگتر از یک باشد می‌توان بگوییم که بین واریانس دو جامعه تفاوت معنی‌داری وجود ندارد.

آزمون فرض

ادعا: نمره‌ی آمار من بیشتر از ۱۵ می‌شود.

نقیض ادعا: نمره‌ی آمار من ۱۵ و کمتر از ۱۵ می‌شود.

مثال: میانگین دستمزد کارگران یک کارخانه کمتر از ۲۰۰,۰۰۰ تومان است. (ادعا)

میانگین دستمزد کارگران یک کارخانه مساوی و بیشتر از ۲۰۰,۰۰۰ تومان است. (نقیض ادعا)

$H_1: \mu < ۲۰۰.۰۰۰$ ادعا

$H_0: \mu \geq ۲۰۰.۰۰۰$ نقیض ادعا

مثال: میانگین نمره مسئولیت‌پذیری سازمان، کمتر از ۵۰ است. (ادعا)

میانگین نمره مسئولیت‌پذیری سازمان، مساوی و بیشتر از ۵۰ است. (نقض ادعا)

$H_1: \mu < ۵۰$ ادعا

$H_0: \mu \geq ۵۰$ نقیض ادعا

مثال: میانگین نمره مسئولیت‌پذیری مدیران روابط عمومی سازمان مساوی ۵۰ است. (ادعا)

میانگین نمره مسئولیت‌پذیری مدیران روابط عمومی سازمان مساوی ۵۰ نیست. (نقیض ادعا)

$H_0: \mu = ۵۰$ ادعا

$H_1 \neq ۵۰$ نقیض ادعا

مثال: میانگین نمره مسئولیت‌پذیری مدیران روابط عمومی بیشتر از ۵۰ است. (ادعا)

میانگین نمره مسئولیت‌پذیری مدیران روابط عمومی مساوی و کمتر از ۵۰ است. (نقیض ادعا)

$H_1: \mu > ۵۰$ ادعا

$H_0: \mu \leq ۵۰$ نقیض ادعا

* در هر کدام از موارد بالا یک مساوی است. هر کدام از آن‌ها که مساوی دارند H_0 و آن که مساوی ندارد H_1 نامیده می‌شود. (می‌توان ابتدا را آورد و بعد H_0 را).

هیچ رابطه‌ی بین مساوی بودن و ادعا و نقیض ادعا نیست.

مثال: میانگین نمره مسئولیت‌پذیری سازمان دست کم (حداقل) ۵۰ است.

میانگین نمره مسئولیت‌پذیری سازمان بیشتر از ۵۰ است.

$H_0: \mu \geq ۵۰$ ادعا

$H_1: \mu < ۵۰$ نقیض ادعا

مثال: میانگین بلوغ روانی کارمندان سازمان «الف» حداکثر ۶۰ است.

میانگین بلوغ روانی کارمندان سازمان «الف» بیشتر از ۶۰ است.

$H_0 : \mu \leq ۶۰$ ادعا

$H_1 : \mu > ۶۰$ نقیض ادعا

مثال: بیش از ۶۰ درصد تولیدات کارخانه دست کم ۲۰ سال عمر است.

۶۰ درصد و کمتر از ۶۰ درصد تولیدات کارخانه دست کم ۲۰ سال عمر است.

$H_1 : \mu > ۶۰$ ادعا

$H_0 : \mu \leq ۶۰$ نقیض ادعا

مثال: نمره‌ی مسئولیت‌پذیری سازمان «الف» بهتر از سازمان «ب» است.

نمره‌ی مسئولیت‌پذیری سازمان «الف» بدتر از سازمان «ب» است.

$H_1 : \mu_1 > \mu_2$ ادعا

$H_0 : \mu_1 \leq \mu_2$ نقیض ادعا

مثال: نمره‌ی سازمان «الف» مساوی سازمان «ب» است.

نمره‌ی سازمان «الف» مساوی سازمان «ب» نیست.

$H_0 : \mu_1 = \mu_2$ ادعا

$H_1 : \mu_1 \neq \mu_2$ نقیض ادعا

مثال: نمره‌ی سازمان «الف» کمتر از سازمان «ب» است.

نمره‌ی سازمان «الف» بیشتر از سازمان «ب» است.

$H_1 : \mu_1 < \mu_2$ ادعا

$H_0 : \mu_1 \geq \mu_2$ نقیض ادعا

مثال: فرضیه‌ای به این صورت بیان شده است «میانگین نمره مسئولیت‌پذیری مدیران روابط عمومی کشور، حداقل ۵۰ است» محقق برای بررسی فرضیه فوق، یک نمونه ۶۴ تایی از بین مدیران کشور، به طور تصادفی انتخاب نموده که میانگین و انحراف معیار آن، به ترتیب، ۴۵ و ۱۶ است، در سطح خطای ۵٪ صحت قضیه فوق را بررسی کنید.

* *چهار مرحله برای حل این گونه مسایل وجود دارد:*

مرحله ۱) ادعا و نقیض ادعا را بنویسیم. ($H_۱$ و $H_\circ$)

$$H_\circ : \mu \geq ۵۰$$
$$H_۱ : \mu < ۵۰$$

مرحله ۲) آماره آزمون را طبق فرمول زیر به دست می‌آوریم:

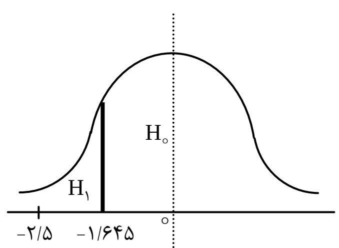
آماره آزمون (در صورتی که $n \geq ۳۰$) $z = \dfrac{\overline{x} - \mu}{\delta \overline{x}} \quad \Rightarrow \quad z = \dfrac{۴۵ - ۵۰}{\frac{۱۶}{\sqrt{۱۶}}} = -۲/۵$

آماره آزمون (در صورتی که $n < ۳۰$) $z = \dfrac{\overline{x} - \mu}{S \overline{x}}$

مرحله ۳) مقدار بحران (خط قرمز یا سطح خطا) را به دست می‌آوریم. (در صورتی که تعداد نمونه ۳۰ یا بیشتر از ۳۰ باشد باید $z\alpha$ و در غیر این صورت $t\alpha$ را به دست آوریم. اگر آماره‌ی آزمون عدد منفی به دست آید مقدار بحران (سطح خطا) را منفی می‌کنیم و اگر مثبت بود، سطح خطا نیز مثبت می‌شود.

$z_{۰/۰۵} = -۱/۶۴۵$

مرحله ۴) یک منحنی به صورت زیر رسم می‌کنیم. (منحنی نرمال) خط تقارن منحنی نرمال را مشخص می‌نماییم. خط تقارن رسم شده نقطه‌ی صفر نمودار در نظر گرفته می‌شود.



بنابراین اگر آماره‌ی آزمون منفی شده باشد (یا علامت $H_۱$ نشانه‌ی کمتر از عدد مورد نظر باشد)، سمت چپ منحنی نرمال، که قسمت منفی نمودار است، خطی به نشانه‌ی خط قرمز (مرز خطا) رسم می‌نماییم. در این صورت نمودار منحنی به دو بخش تقسیم می‌شود؛ بخش کوچکتر محدوده‌ی $H_۱$ و بخش بزرگتر محدوده‌ی $H_\circ$ می‌باشد.

سپس آماره‌ی آزمون را در نمودار مشخص می‌نماییم. در صورتی که آماره‌ی آزمون در محدوده‌ی $H_۱$ قرار بگیرد، ادعا رد شده و نقیض ادعا تایید می‌شود. و بالعکس، اگر آماره‌ی آزمون در محدوده‌ی $H_\circ$ واقع شود، به معنی این است که ادعا تأیید شده و نقیض ادعا رد شده می‌گردد.

نتیجه مثال: *پس از مقایسه‌ی آماره آزمون با مقدار بحرانی، مشخص می‌شود که، آماره آزمون در ناحیه H_1 قرار می‌گیرد. بنابراین در سطح اطمینان ۹۵٪ می‌توان گفت که مشاهدات دلالت کافی بر تأیید H_0 ندارد. از آنجایی که فرض H_0 بیان‌کننده‌ی فرضیه‌ی پژوهشی است (ادعا) پس در سطح خطای ۵٪ می‌توان گفت فرضیه پژوهشی رد می‌شود و نقیض آن یعنی میانگین نمره‌ی مسئولیت‌پذیری مدیران در کشور کمتر از ۵۰ است، پذیرفته می‌شود.*

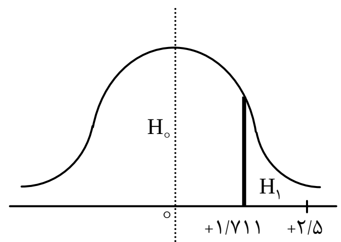
مثال: فرضیه‌ای به این صورت بیان شده است که، «میانگین نمره‌ی درس آمار دانشجویان، حداکثر ۱۴ می‌باشد. بدین منظور از بین دانشجویان یک نمونه ۲۵ نفره انتخاب، که میانگین نمره آنها ۱۶ می‌باشد همچنین انحراف معیار آن ۴ می‌باشد. در سطح خطای ۵٪ فرضیه‌ی فوق را بررسی نمایید.

$$H_0 : \mu \leq ۱۴$$

$$H_1 : \mu > ۱۴$$

$$t = \frac{\bar{x} - \mu}{S\bar{x}} = \frac{۱۶ - ۱۴}{\frac{۴}{\sqrt{۲۵}}} = \frac{۲}{۰/۸} = +۲/۵$$

$$t_{۰/۰۵} = +۱/۷۱۱$$



نتیجه: *چون مقدار آزمون (۲/۵) در ناحیه‌ی H_1 قرار گرفته است، لذا با توجه به این که H_1 در مرحله‌ی اول مشخص گردید، که نقیض ادعای ما را بیان می‌نماید، لذا فرضیه‌ی آزمون مورد تأیید قرار نمی‌گیرد.*

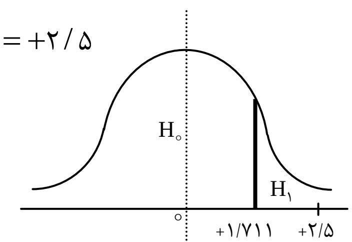
مثال: فرضیه‌ای به این صورت بیان شده است که، «میانگین نمره‌ی درس آمار دانشجویان، حداکثر ۱۵ می‌باشد. بدین منظور از بین دانشجویان یک نمونه ۲۵ نفره انتخاب، که میانگین نمره آنها ۱۶ می‌باشد همچنین انحراف معیار آن ۴ می‌باشد. در سطح خطای ۵٪ فرضیه‌ی فوق را بررسی نمایید.

$$H_0 : \mu \leq ۱۴$$

$$H_1 : \mu > ۱۴$$

$$t = \frac{\bar{x} - \mu}{S\bar{x}} = \frac{۱۶ - ۱۴}{\frac{۴}{\sqrt{۲۵}}} = \frac{۲}{۰/۸} = +۲/۵$$

$$t_{۰/۰۵} = +۱/۷۱۱$$



نتیجه: *به لحاظ این که آماده آزمون در ناحیه‌ی H_0 قرار گرفته و همچنین با توجه به صورت مسأله H_0 ادعا را مشخص می‌نماید لذا فرضیه‌ی ما تأیید می‌شود.*

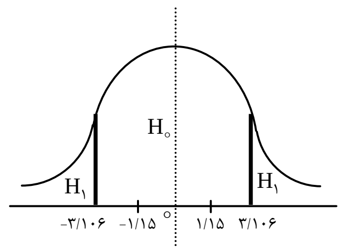
مثال: فرضیه‌ای به این صورت توسط یک دانشجوی روابط عمومی طراحی شده است «میانگین مهارتهای انسانی مدیران روابط عمومی سازمان الف ۵۵ است» به منظور بررسی فرضیه‌ی فوق دانشجو از بین مدیران سازمان الف یک نمونه تصادفی ۱۲ نفره انتخاب که میانگین و انحراف معیار آن به ترتیب ۶۰ و ۱۵ است، در سطح اطمینان ۹۹٪ صحت فرضیه فوق را بررسی کنید.

$$H_{۰}: \mu = ۵۵$$

$$H_{۱}: \mu \neq ۵۵$$

$$t_{\frac{\alpha}{۲}, n-۱} = t_{\frac{۰/۰۱}{۲}} = \pm ۳/۱۰۶$$

$$t = \frac{\bar{x} - \mu}{\frac{Sx}{\sqrt{n}}} = \frac{۶۰ - ۵۵}{\frac{۱۵}{\sqrt{۱۲}}} = ۱/۱۵$$



* چون علامت بیشتر و کمتر نداریم، بنابراین هر دو طرف (مثبت و منفی) را مشخص می‌کنیم و به عنوان مقدار بحران در نظر می‌گیریم و آماره‌ی آزمون نیز یک بار مثبت و یک بار منفی مشخص می‌کنیم.

نتیجه: *به لحاظ این که مقدار آماره آزمون در ناحیه $H_{۰}$ قرار می‌گیرد و همچنین با توجه به مسأله که مشخص می‌نماید $H_{۰}$ ادعا می‌باشد، لذا نتیجه می‌گیریم که ادعای ما تأیید می‌شود.*