برآورد :

منظور از برآورد، تخمین پارامتر مجهول جامعه به کمک آماره به دست آمده از نمونه است.

دو نوع برآورد داریم : برآورد نقطه‌ای و برآورد فاصله‌ای .

برآورد نقطه‌ای : در برآورد نقطه‌ای از روی آماره به دست آمده از نمونه تنها یک عدد یا یک نقطه به عنوان تخمین پارامتر جامعه معرفی می‌شود. مثلاً اگر میانگین نمونه‌ای برابر 20 شود ($\bar{x}=20$) در این صورت 20 برآوردی از میانگین جامعه خواهد بود.

آماره استفاده شده جهت برآورد پارامتر θ را برآوردگر θ گفته و با $\hat{\theta}$ نمایش می‌دهیم.

مثلاً برآوردگر μ را با $\hat{\mu}$ ، برآوردگر σ^2 را با $\hat{\sigma}^2$ ، برآوردگر نسبت P در جامعه را با $\hat{P}$ نمایش می‌دهیم. به طور مثال $\bar{X}$ یک برآوردگر نقطه‌ای برای پارامتر μ ، S^2 (واریانس نمونه‌ای) یک برآوردگر نقطه‌ای از σ^2 و $\hat{P}$ (نسبت نمونه‌ای) یک برآوردگر نسبت جامعه است.

برآوردگر نااریب : $\hat{\theta}$ را یک برآوردگر نااریب پارامتر مجهول θ گوییم هرگاه

$$E[\hat{\theta}] = \theta$$

مثلاً در بخش قبل دیدیم که $E[\bar{X}] = \mu$ پس $\bar{X}$ یک برآوردگر نااریب از μ می‌باشد. می‌توان نشان داد که $E[S^2] = \sigma^2$ لذا S^2 نیز یک برآوردگر نااریب σ^2 است.

برآورد فاصله‌ای (فاصله اطمینان) : فاصله اطمینان دامنه‌ای از مقادیری است با یک حد پایین و بالا که با اطمینان خاصی می‌توان گفت که پارامتر مجهول μ در این دامنه قرار دارد. در ادامه به تعریف دقیق فاصله اطمینان می‌پردازیم:

تعریف: فاصله اطمینان، فاصله‌ای (بازه‌ای) است با یک کران پایین L و یک کران بالای U که می‌توان گفت پارامتر مجهول θ از جامعه (مثلاً میانگین μ یا واریانس σ^2) با اطمینان خاصی، احتمال $1-\alpha$، در این بازه قرار دارد، یعنی:

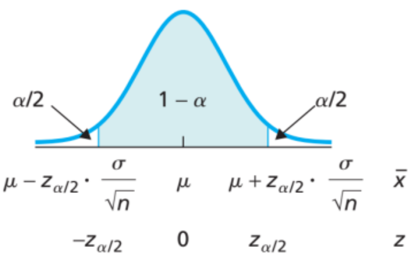
$$P(L < \theta < U) = 1-\alpha$$

در این صورت L و U را کران‌های اطمینان پایینی و بالایی و فاصله اطمینان را یک فاصله اطمینان $100(1-\alpha)$ درصدی و همچنین $1-\alpha$ را درجه اطمینان می‌نامند. معمولاً $\alpha=0.05$ یا $\alpha=0.01$ در نظر گرفته می‌شود. مثلاً برای $\alpha=0.05$ درجه اطمینان 0.95 و فاصله اطمینان را یک فاصله اطمینان 95 درصدی می‌گوییم.

فاصله اطمینان برای میانگین جامعه با واریانس معلوم:

اگر X_1، X_2، ...، X_n یک نمونه تصادفی جامعه‌ای نرمال با میانگین μ و واریانس σ^2 باشد و یا n به اندازه کافی بزرگ باشد ($n>30$) در این صورت با توجه به قضیه حد مرکزی داریم:

$$Z = \frac{\bar{X}-\mu}{\frac{\sigma}{\sqrt{n}}} \sim N(0,1)$$



با توجه به اینکه:

$$P(|Z| \leq z_{\alpha/2}) = 1-\alpha$$

پس:

$$P\left(\left|\frac{\bar{X}-\mu}{\frac{\sigma}{\sqrt{n}}}\right| \leq z_{\alpha/2}\right) = P\left(-z_{\alpha/2} \leq \frac{\bar{X}-\mu}{\frac{\sigma}{\sqrt{n}}} \leq z_{\alpha/2}\right)$$

$$= P\left(\bar{X} - \frac{\sigma}{\sqrt{n}} z_{\frac{\alpha}{2}} \leq \mu \leq \bar{X} + \frac{\sigma}{\sqrt{n}} z_{\alpha/2}\right) = 1-\alpha$$

لذا فاصله زیر یک فاصله اطمینان $100(1-\alpha)$ درصدی برای میانگین جامعه است:

$$\left(\bar{x} - \frac{\sigma}{\sqrt{n}} z_{\alpha/2} , \bar{x} + \frac{\sigma}{\sqrt{n}} z_{\alpha/2}\right)$$

$$\bar{x} \pm \frac{\sigma}{\sqrt{n}} z_{\alpha/2}$$

مثال : فرض کنید یک پژوهشگری سطح متوسط آنزیمی را در جهت انسانی خاصی برآورد می کند. بدین منظور یک نمونه ۱۰ تایی انتخاب کرده و سطح آنزیم را در هر یک اندازه گیری کرده است و میانگین 22 به دست آورده است. علاوه بر آن این تصور وجود دارد که متغیر مورد نظر از توزیع نرمال با واریانس 45 تبعیت می کند. فاصله اطمینان 95 درصدی برای میانگین سطح آنزیم به دست آورید.

$$n=10,\quad \bar{x}=22,\quad X_i \sim N(\mu,45) \qquad \sigma^2=45 \Rightarrow \sigma=\sqrt{45}=6.7$$

$$\alpha=0.05 \Rightarrow \frac{\alpha}{2}=0.025 \qquad Z_{\alpha/2}=Z_{0.025}=1.96$$

$$\bar{x} \pm Z_{\alpha/2}\frac{\sigma}{\sqrt{n}} \Rightarrow 22 \pm 1.96\,\frac{6.7}{\sqrt{10}} \Rightarrow (17.84\,,\,26.15)$$

مثال : میزان حداکثر قوت یک ماهیچه خاص از توزیع نرمال با واریانس 144 پیروی می کند. یک نمونه تصادفی 15 تایی گرفته شده است و میانگین 84.3 به دست آمده است. یک فاصله اطمینان ۹۹ درصدی برای میانگین حداکثر قوت این ماهیچه به دست آورید.

$$\sigma^2=144,\quad n=15,\quad \bar{x}=84.3,\quad \alpha=0.01 \Rightarrow \frac{\alpha}{2}=0.005$$

$$\bar{x} \pm Z_{\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}} \qquad Z_{\frac{\alpha}{2}}=Z_{0.005}=2.576$$

$$\Rightarrow 84.3 \pm 2.576\,\frac{12}{\sqrt{15}} \Rightarrow (76.318\,,\,92.281)$$

مثال : نمونه ای ۱۰۰ تایی از پسران بالغ 25 ساله گرفته شده است. فشار خون سیستولیک این افراد اندازه گیری شده و میانگین آن 125 می باشد. اگر انحراف معیار جامعه 15 باشد، یک فاصله اطمینان ۹۰ درصدی برای میانگین فشار خون سیستولیک افراد 25 ساله بیابید.

$$n=100,\quad \bar{x}=125,\quad \sigma=15,\quad \alpha=0.1 \Rightarrow \frac{\alpha}{2}=0.05 \qquad Z_{\frac{\alpha}{2}}=Z_{0.05}=1.645$$

$$\bar{x} \pm Z_{\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}} \Rightarrow 125 \pm 1.645\,\frac{15}{\sqrt{100}} \Rightarrow (122.53\,,\,127.46)$$

تذکر: اگر حجم نمونه بیش از ۳۰ و واریانس جامعه نامعلوم باشد کافیست از انحراف معیار نمونه‌ای s به جای σ در فرمول فاصله اطمینان استفاده کنیم:

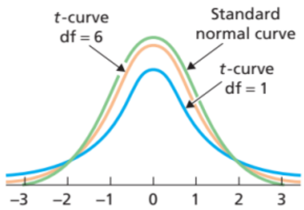
$$\left(\bar{x} - Z_{\frac{\alpha}{2}} \frac{s}{\sqrt{n}} , \bar{x} + Z_{\frac{\alpha}{2}} \frac{s}{\sqrt{n}}\right)$$

فاصله اطمینان میانگین یک جامعه نرمال با واریانس مجهول:

در عمل در بیشتر مسائل وقتی میانگین جامعه مجهول است، واریانس نیز مجهول می‌باشد. در اینگونه موارد بخصوص وقتی حجم نمونه کمتر از ۳۰ است باید از توزیع t استیودنت کمک بگیریم. اگر $X_1, \dots, X_n$ یک نمونه تصادفی به حجم n از جامعه‌ای نرمال با میانگین μ و واریانس σ^2 باشد، در این صورت متغیر تصادفی

$$T = \frac{\bar{X} - \mu}{\frac{s}{\sqrt{n}}}$$

دارای توزیع t استیودنت با $n-1$ درجه آزادی است.



همانطور که مشاهده می‌شود توزیع t استیودنت بسیار شبیه توزیع نرمال استاندارد است. در واقع اگر درجه آزادی این توزیع به اندازه کافی بزرگ (بزرگتر از ۳۰) باشد می‌توان این توزیع را به کمک توزیع نرمال استاندارد تقریب زد. همانطور که از نمودار بالا مشخص است توزیع t یک توزیع با میانگین صفر و متقارن نسبت به میانگین است.

نمادگذاری : $t_{\alpha,n}$ را مقداری تعریف می‌کنیم که سطح زیر منحنی چگالی توزیع t با n درجه آزادی در سمت راست $t_{\alpha,n}$ برابر با α باشد.

$$P(t_n \geq t_{\alpha,n}) = \alpha$$

$$P(t_n \leq t_{\alpha,n}) = 1-\alpha$$

حال می‌توانیم مشابه قبل فاصله اطمینان میانگین را بیابیم :

$$\begin{aligned}1-\alpha &= P\left(|t| < t_{\frac{\alpha}{2},n-1}\right)\\ &= P\left(\left|\frac{\bar{x}-\mu}{\frac{s}{\sqrt{n}}}\right| < t_{\frac{\alpha}{2},n-1}\right)\\ &= P\left(\bar{x} - t_{\frac{\alpha}{2},n-1}\frac{s}{\sqrt{n}} \leq \mu \leq \bar{x} + t_{\frac{\alpha}{2},n-1}\frac{s}{\sqrt{n}}\right)\end{aligned}$$

لذا فاصله زیر یک فاصله اطمینان $100(1-\alpha)$ درصدی برای میانگین جامعه است :

$$\left(\bar{x} - t_{\frac{\alpha}{2},n-1}\frac{s}{\sqrt{n}} \, , \, \bar{x} + t_{\frac{\alpha}{2},n-1}\frac{s}{\sqrt{n}}\right)$$

$$\bar{x} \pm t_{\frac{\alpha}{2},n-1}\frac{s}{\sqrt{n}}$$

مثال : تعیین سرم آمیلاز بر روی 15 نفر انسان سالم انجام شده است. نمونه انتخابی دارای میانگین 96 واحد در 100 میلی لیتر و انحراف معیار 35 واحد در 100 میلی لیتر است. یک فاصله اطمینان 95 درصدی برای میانگین میزان سرم آمیلاز در 100 میلی لیتر به دست آورید.

$$n=15, \quad \bar{x}=96, \quad s=35 \quad \alpha=0.05 \Rightarrow \frac{\alpha}{2}=0.025$$

$$t_{\frac{\alpha}{2},n-1} = t_{0.025,14} = 2.145$$

$$\bar{x} \pm t_{\frac{\alpha}{2},n-1}\frac{s}{\sqrt{n}} \Rightarrow 96 \pm 2.145\,\frac{35}{\sqrt{15}}$$

$$(76.615 \, , \, 115.384)$$

مثال: در زیر وزن 15 نوزاد دختر که به صورت تصادفی انتخاب شده اند برحسب کیلوگرم داده شده است:

$28, 31, 33, 28, 33, 37, 31, 32, 31, 28, 34, 28, 33, 26, 30$

یک فاصله اطمینان 95 درصدی برای میانگین وزن نوزادان دختر هنگام تولد بیابید.

$$n = 15, \quad \bar{x} = \frac{\sum x_i}{n} = \frac{28+31+33+\cdots+30}{15} = 30.9$$

$$s^2 = \frac{\sum x_i^2 - n\bar{x}^2}{n-1} = 8.41 \Rightarrow s = \sqrt{8.41} = 2.9$$

$$\alpha = 0.05 \Rightarrow t_{\frac{\alpha}{2}, n-1} = t_{0.025, 14} = 2.145$$

$$\bar{x} \pm \frac{s}{\sqrt{n}} t_{\frac{\alpha}{2}, n-1} \Rightarrow 30.9 \pm \frac{2.9}{\sqrt{15}} \times 2.145$$

$$(29.293, \ 32.506)$$

فاصله اطمینان برای نسبت p: فرض کنید $X_1, X_2, \ldots, X_n$ یک نمونه تصادفی به حجم n، که در آن $n > 30$ از یک جامعه برنولی با نسبت p باشد. در این صورت اگر $\hat{p}$ نسبت نمونه‌ای باشد یعنی،

$$\hat{p} = \frac{\text{تعداد موفقیت‌ها در نمونه}}{n}$$

با تعریف $\hat{q} = 1 - \hat{p}$، فاصله اطمینان $100(1-\alpha)$ درصدی برای نسبت p به صورت زیر است

$$\left(\hat{p} - z_{\frac{\alpha}{2}} \sqrt{\frac{\hat{p}\hat{q}}{n}}, \ \hat{p} + z_{\frac{\alpha}{2}} \sqrt{\frac{\hat{p}\hat{q}}{n}}\right)$$

$$\hat{p} \pm z_{\frac{\alpha}{2}} \sqrt{\frac{\hat{p}\hat{q}}{n}}$$

مثال: برای بررسی شیوع بیماری فشار خون در افراد با بیش از 30 سال سن، یک نمونه تصادفی 100 تایی در نظر گرفته ایم و مشاهده کردیم که 23 نفر آنها دارای این بیماری هستند. یک فاصله اطمینان 99 درصدی برای نسبت p شیوع این بیماری در افراد با بیش از 30 سال سن بدست آورید.

$\hat{p} = 0.23, \; n = 100, \; \hat{q} = 1 - 0.23 = 0.77$

$\alpha = 0.01 \Rightarrow \frac{\alpha}{2} = 0.005 \Rightarrow Z_{\frac{\alpha}{2}} = Z_{0.005} = 2.57$

$$\hat{p} \pm Z_{\frac{\alpha}{2}} \sqrt{\frac{\hat{p}\hat{q}}{n}} \Rightarrow 0.23 \pm 2.57 \sqrt{\frac{0.23 \times 0.77}{100}}$$

تعیین اندازه نمونه برای برآورد نسبت :

با توجه به فرمول فاصله اطمینان برای نسبت p :

$$\hat{p} \pm Z_{\frac{\alpha}{2}} \sqrt{\frac{\hat{p}\hat{q}}{n}}$$

اگر بخواهیم با احتمال $1-\alpha$ خطای برآورد p با $\hat{p}$ یعنی $d = |p - \hat{p}|$ کمتر از e شود

با توجه به فاصله اطمینان برای تعیین n داریم :

$$e = Z_{\frac{\alpha}{2}} \sqrt{\frac{\hat{p}\hat{q}}{n}} \Rightarrow e^2 = Z^2_{\alpha/2} \frac{\hat{p}\hat{q}}{n} \Rightarrow n = Z^2_{\alpha/2} \frac{\hat{p}\hat{q}}{e^2}$$

که $\hat{p}$ و $\hat{q}$ خود برآوردی از p و q هستند. اگر چنین برآوردی وجود نداشته باشد با توجه

به اینکه $\hat{p}$ و $\hat{q}$ اعدادی بین صفر و یک هستند لذا $\max(\hat{p}\hat{q}) = \frac{1}{4}$ پس مقدار n را

به کمک فرمول زیر تعیین می کنیم :

$$n = \frac{Z^2_{\alpha/2}}{4e^2}$$

خلاصه بحث : اگر برآوردی از p و q موجود باشد : $\left[n = Z^2_{\frac{\alpha}{2}} \frac{\hat{p}\hat{q}}{e^2} \right]$

اگر $\hat{p}$ و $\hat{q}$ موجود نباشند : $\left[n = \frac{Z^2_{\alpha/2}}{4e^2} \right]$

مثال : در مثال قبل حجم نمونه چقدر باشد که خطای $\hat{p}$ با احتمال 95 درصد کمتر از 0.02 باشد.

$e = 0.02, \; \alpha = 0.05 \Rightarrow \frac{\alpha}{2} = 0.025 \quad Z_{\frac{\alpha}{2}} = Z_{0.025} = 1.96$

با توجه به فرمول بالا : $n = \frac{(1.96)^2}{4 \times 4 \times 10^{-4}} = 2401$

تعیین اندازه نمونه برای برآورد میانگین:

با توجه به فرمول فاصله اطمینان برای میانگین یک نمونه

$$\bar{x} \pm Z_{\alpha/2} \frac{\sigma}{\sqrt{n}}$$

اگر بخواهیم خطای برآورد میانگین با $\bar{x}$ کمتر از e باشد، با توجه به فرمول فوق

$$e = Z_{\alpha/2} \frac{\sigma}{\sqrt{n}} \Rightarrow e^2 = Z^2_{\alpha/2} \frac{\sigma^2}{n} \Rightarrow n = \left(\frac{Z_{\alpha/2}\,\sigma}{e}\right)^2$$

$$\left[\, n = \left(\frac{Z_{\alpha/2}\,\sigma}{e}\right)^2 \,\right]$$

اگر σ مجهول باشد می‌توان از s استفاده کرد و یا $\frac{R}{6}$ را به عنوان برآوردی از σ قرار داد. (R دامنه متغیر)

مثال: طول قد نوزادان در یک جامعه از توزیع نرمال با انحراف معیار 3 سانتیمتر پیروی می‌کند. اگر محققی بخواهد میانگین قد نوزادان را با خطای کمتر از یک سانتیمتر با احتمال 95 درصد برآورد کند، حجم نمونه چقدر باشد؟

$\sigma = 3, \quad \alpha = 0.05 \quad \frac{\alpha}{2} = 0.025 \quad Z_{\alpha/2} = Z_{0.025} = 1.96$

$e = 1$

$$n = \left(\frac{Z_{\alpha/2}\,\sigma}{e}\right)^2 = \left(\frac{1.96 \times 3}{1}\right)^2 = 35$$

فاصله اطمینان برای اختلاف میانگین دو جامعه با واریانس‌های معلوم:

فرض کنید X_1 ، X_2 ، ... ، X_{n_1} نمونه‌ای به حجم n_1 از جامعه‌ای نرمال با میانگین μ_1 و واریانس σ_1^2 و Y_1 ، ... ، Y_{n_2} نمونه‌ای به حجم n_2 از جامعه‌ای نرمال با میانگین μ_2 و واریانس σ_2^2 باشد و این دو جامعه از هم مستقل باشند. در اینصورت

$$X_i \sim N(\mu_1, \sigma_1^2) \Rightarrow \bar{X} \sim N\left(\mu_1, \frac{\sigma_1^2}{n_1}\right)$$

$$Y_i \sim N(\mu_2, \sigma_2^2) \Rightarrow \bar{Y} \sim N\left(\mu_2, \frac{\sigma_2^2}{n_2}\right)$$

لذا

$$\bar{X} - \bar{Y} \sim N\left(\mu_1 - \mu_2, \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}\right)$$

پس:

$$Z = \frac{(\bar{X} - \bar{Y}) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$$

پس فاصله اطمینان $100(1-\alpha)$ درصدی برای اختلاف میانگین دو جامعه به صورت زیر است:

$$(\bar{x} - \bar{y}) \pm Z_{\alpha/2} \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$$

نکته 1: اگر حجم دو نمونه از 30 بیشتر باشد، فرمول فوق برای جامعه غیرنرمال نیز درست است.

نکته 2: اگر حجم دو نمونه بیش از 30 باشد و واریانس‌های دو جامعه مجهول باشد کافیست در فرمول فوق از واریانس‌های نمونه‌ای به جای واریانس جامعه استفاده کنیم:

$$(\bar{x} - \bar{y}) \pm Z_{\frac{\alpha}{2}} \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$$

مثال: در بیمارستان بزرگ روانی، نمونه‌ای 12 نفره از بیماران مبتلا به بیماری خاصی را انتخاب کرده و میانگین مقدار اسید اوریک سرم 4.5 میلی‌گرم در صد میلی‌لیتر به دست آمد.

در بیمارستان دیگری نمونه‌ای 15 نفری از افراد طبیعی گرفته شده و میانگین اسید اوریک

3.4 بوده است. اگر واریانس هر دو جمعیت 1 و توزیع نرمال باشند. یک فاصله اطمینان 95 درصدی برای اختلاف میانگین اسیداوریک افراد سالم و بیمار به دست آورید.

$\bar{x} = 4.5, \quad \sigma_1 = 1 \quad , \quad n_1 = 12$

$\bar{y} = 3.4 \quad \sigma_2 = 1 \quad , \quad n_2 = 15 \quad , \quad \alpha = 0.05 \Rightarrow \frac{\alpha}{2} = 0.025 \quad Z_{0.025} = 1.96$

$$(\bar{x} - \bar{y}) \pm Z_{\alpha/2} \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} \Rightarrow (4.5 - 3.4) \pm 1.96 \sqrt{\frac{1}{12} + \frac{1}{15}}$$

$$\Rightarrow 1.1 \pm 1.96\,(0.39) \Rightarrow (0.3, 1.9)$$

یعنی با احتمال 95 درصد اطمینان داریم $\mu_1 - \mu_2$ مابین 0.3 و 1.9 است.

فاصله اطمینان برای اختلاف میانگین دو جامعه نرمال با واریانس های مجهول:

دو نمونه به حجم n_1 و n_2 از دو جامعه نرمال مستقل از هم با میانگین های μ_1 و μ_2 و واریانس های مجهول σ_1^2 و σ_2^2 در نظر بگیرید. در اینجا باید حتماً واریانس دو جامعه با هم برابر باشند، یعنی $\sigma_1^2 = \sigma_2^2 = \sigma^2$. در این صورت فاصله زیر یک فاصله اطمینان $100(1-\alpha)$ درصدی برای اختلاف میانگین دو جامعه یعنی $\mu_1 - \mu_2$ است:

$$(\bar{x}_1 - \bar{x}_2) \pm t_{\alpha/2,\, n_1+n_2-2}\; s_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}$$

که در آن s_p^2 واریانس آمیخته نام دارد و برابر است با:

$$s_p^2 = \frac{(n_1 - 1)\, s_1^2 + (n_2 - 1)\, s_2^2}{n_1 + n_2 - 2}$$

مثال: سرم آمیلاز در 15 فرد سالم اندازه گیری شده و میانگین آن 96 و انحراف معیار آن 35 بوده است. همچنین نمونه ای 25 نفره از افراد بستری در نظر گرفته ایم و میانگین سرم

آمیلاز در این گروه برابر ۱۲۰ و انحراف معیار ۴۰ است. همچنین فرض کنید هر دو جمعیت دارای توزیع نرمال با واریانس های برابر باشند. یک فاصله اطمینان ۹۵ درصدی برای $\mu_2 - \mu_1$ به دست آورید.

$n_1 = 15, \; \bar{x}_1 = 96, \; s_1 = 35$

$n_2 = 22, \; \bar{x}_2 = 120, \; s_2 = 40$

$$s_p^2 = \frac{(15-1)(35)^2 + (22-1)(40)^2}{15+22-2} = 1450$$

$$s_p = 38.07$$

$$(\bar{x}_2 - \bar{x}_1) \pm t_{\alpha/2} \, s_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}$$

$$\alpha = 0.05 \Rightarrow \frac{\alpha}{2} = 0.025 \Rightarrow t_{0.025,35} = 2.03$$

$$(120-96) \pm (2.03)(38.07)\sqrt{\frac{1}{15} + \frac{1}{22}}$$

$$(-1.877, \; 49.877)$$

فاصله اطمینان اختلاف نسبت:

اگر $X_1, \ldots, X_{n_1}$ یک نمونه تصادفی به حجم n_1 از یک جامعه برنولی با نسبت P_1 و $Y_1, \ldots, Y_{n_2}$ یک نمونه تصادفی به حجم n_2 از یک جامعه برنولی با نسبت P_2 باشد. همچنین این دو جامعه مستقل از هم باشند و $n_1, n_2 \geq 30$ در این صورت

$$(\hat{P}_1 - \hat{P}_2) \pm z_{\frac{\alpha}{2}} \sqrt{\frac{\hat{P}_1 \hat{q}_1}{n_1} + \frac{\hat{P}_2 \hat{q}_2}{n_2}}$$

یک فاصله اطمینان $100(1-\alpha)$ درصدی برای $P_1 - P_2$ است.

مثال: برای بررسی نسبت شیوع بیماری فشار خون در بین زنان و مردان یک نمونه ۱۰۰ نفره از هر کدام در نظر گرفته‌ایم. مشاهده شد که 32 مرد و 25 زن مبتلا به فشار خون بالا هستند. یک فاصله اطمینان ۸۰ درصدی برای اختلاف نسبت شیوع این بیماری بین مردان و زنان به دست آورید.

$$\hat{P}_1 = 0.32 , \quad \hat{P}_2 = 0.25 , \quad n_1 = n_2 = 100 \quad \alpha = 0.2 \Rightarrow \frac{\alpha}{2} = 0.1$$

$$Z_{0.1} = 1.282$$

$$\hat{q}_1 = 1 - \hat{P}_1 = 0.68$$

$$\hat{q}_2 = 1 - \hat{P}_2 = 0.75$$

$$(\hat{P}_1 - \hat{P}_2) \pm Z_{0.1} \sqrt{\frac{\hat{P}_1 \hat{q}_1}{n_1} + \frac{\hat{P}_2 \hat{q}_2}{n_2}}$$

$$(0.32 - 0.25) \pm 1.282 \sqrt{\frac{(0.68)(0.32)}{100} + \frac{(0.75)(0.25)}{100}}$$

فاصله اطمینان برای مشاهدات زوجی: فرض کنید می‌خواهیم میزان اثربخشی یک رژیم غذایی خاص را در کاهش وزن افراد بررسی کنیم. برای این منظور یک نمونه تصادفی به حجم n در نظر گرفته و وزن آنها را ثبت می‌کنیم و سپس پس از یک دوره شش ماهه از مصرف این رژیم غذایی وزن آنها را دوبار ثبت می‌نماییم. واضح است که وزن هر فرد قبل و بعد از مصرف رژیم غذایی به هم وابسته است. بنابراین در اینجا مشاهدات زوجی مانند (X_i, Y_i) داریم. برای رفع مشکل استقلال متغیر D_i را به صورت $D_i = X_i - Y_i$ تعریف می‌کنیم. اگر D_i دارای توزیع نرمال با میانگین μ_D باشد، در اینصورت

$$T = \frac{\bar{D} - \mu_D}{\frac{S_D}{\sqrt{n}}} \sim t_{(n-1)}$$

پس فاصله زیر یک فاصله اطمینان $100(1-\alpha)$ درصدی برای $\mu_D = \mu_X - \mu_Y$ است:

$$\bar{d} \pm t_{\frac{\alpha}{2}, n-1} \frac{s_D}{\sqrt{n}}$$

مثال: می‌خواهیم تأثیر یک رژیم غذایی خاص را در کاهش وزن افراد بررسی کنیم. برای بررسی ۱۰ نفر به طور تصادفی انتخاب شده، وزن آنها قبل و شش ماه بعد از مصرف رژیم غذایی اندازه‌گیری شده است و اطلاعات آن در جدول زیر آمده است. بر اساس اطلاعات این جدول یک فاصله اطمینان ۹۸ درصدی برای اختلاف میانگین وزن این افراد قبل و بعد از مصرف رژیم غذایی به دست آورید.

	وزن بعد	وزن قبل	d_i
1	76	81	-5
2	60	52	8
3	85	87	-2
4	58	70	-12
5	91	86	5
6	75	77	-2
7	82	90	-8
8	64	63	1
9	79	85	-6
10	88	83	5

$$\bar{d} = \frac{\sum d_i}{n} = -1.6$$

$$s_D = 6.38$$

$$n = 10, \quad \alpha = 0.02 \Rightarrow \frac{\alpha}{2} = 0.01$$

$$t_{0.01, 9} = 2.82$$

$$\bar{d} \pm t_{\frac{\alpha}{2}, n-1} \frac{s_d}{\sqrt{n}}$$

$$-1.6 \pm 2.82 \frac{6.38}{\sqrt{10}}$$

$$\mu_D \in (-7.29, 4.09)$$