Доверительные интервалы

Определение. Доверительным интервалом для одномерного параметра θ с доверительной вероятностью $1-\alpha$ называется любой интервал (θ_1, θ_2), содержащий истинное значение параметра с вероятностью $1-\alpha$.

Для построения доверительного интервала нужно построить статистику (случайную величину), которая удовлетворяла бы трем условиям:

1. Эта статистика должна иметь известный закон распределения,
2. Она должна содержать искомый параметр,
3. Она должна быть вычисляема, т.е. не должна содержать неизвестные параметры.

Принцип построения доверительных интервалов:

Пусть $\overrightarrow{X_n} = (X_1, X_2, \ldots X_n)$ – выборка из некоторого закона распределения, зависящего от параметра θ.

1. Находим статистику $\eta = \varphi\left(\overrightarrow{X_n}, \theta\right)$, закон распределения которой $F(x)$ не зависит от неизвестного параметра θ.

2. Находим квантили $t_{\alpha/2}$ и $t_{1-\alpha/2}$ функции распределения $F(x)$ уровней $\frac{\alpha}{2}$ и $1-\frac{\alpha}{2}$, то есть такие значения, что монотонно возрастающая функция $F(x)$ достигает указанных уровней в этих точках: $F(t_{\alpha/2}) = \alpha/2$, $F(t_{1-\alpha/2}) = 1-\alpha/2$. При этом

$$P\left(t_{\alpha/2} < \varphi\left(\overrightarrow{X_n}, \theta\right) < t_{1-\alpha/2}\right) = F(t_{1-\alpha/2}) - F(t_{\alpha/2}) = 1-\alpha.$$

3. Неравенство $t_{\alpha/2} < \varphi\left(\overrightarrow{X_n}, \theta\right) < t_{1-\alpha/2}$ разрешается относительно параметра θ:

$$t_{\alpha/2} < \varphi\left(\overrightarrow{X_n}, \theta\right) < t_{1-\alpha/2} \leftrightarrow \underline{\varphi}\left(\overrightarrow{X_n}, \alpha\right) < \theta < \overline{\varphi}\left(\overrightarrow{X_n}, \alpha\right)$$

Обозначив $\theta_1 = \underline{\varphi}\left(\overrightarrow{X_n}, \alpha\right)$, $\theta_2 = \overline{\varphi}\left(\overrightarrow{X_n}, \alpha\right)$, получаем доверительный интервал уровня $1-\alpha$.

Построение доверительных интервалов для параметров нормального распределения

1. Доверительный интервал для (неизвестного) математического ожидания a при известном с.к.о. σ.

Так как доверительный интервал строится для параметра a, то используем статистику, которая используется для его оценки это $\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$ и она распределена по нормальному закону с параметрами

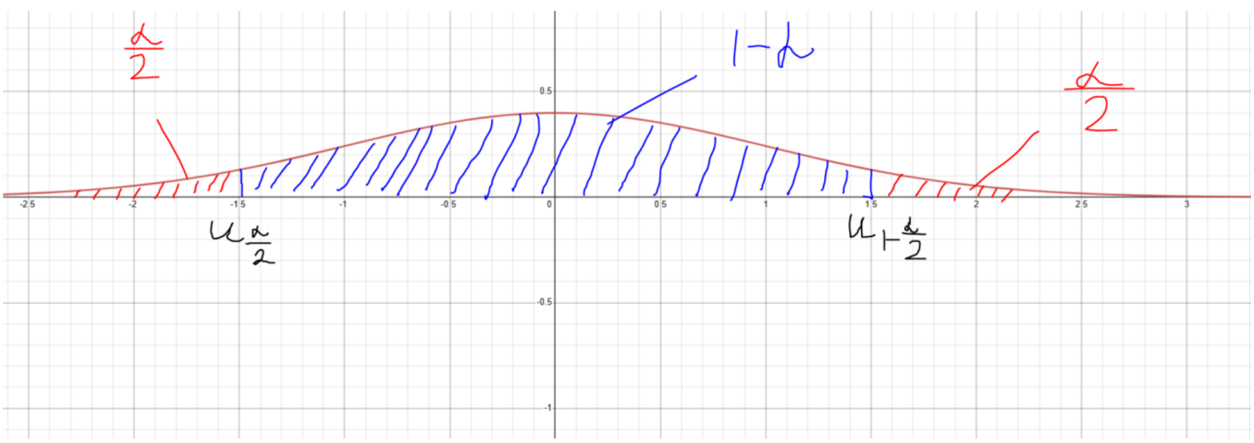
$$M\bar{X} = M\frac{1}{n}\sum_{i=1}^{n} X_i = \frac{1}{n}\sum_{i=1}^{n} MX_i = \frac{1}{n} na = a$$

$$D\bar{X} = D\frac{1}{n}\sum_{i=1}^{n} X_i = \frac{1}{n^2}\sum_{i=1}^{n} DX_i = \frac{1}{n^2} n\sigma^2 = \frac{\sigma^2}{n}$$

Следовательно, $\bar{X} \sim N\left(a, \frac{\sigma}{\sqrt{n}}\right)$, но нам нужно, чтобы закон распределения не зависел от параметра a, поэтому приведём к стандартному нормальному закону.

$$\bar{X} - a \sim N\left(0, \frac{\sigma}{\sqrt{n}}\right) \to \frac{\bar{X} - a}{\sigma} \cdot \sqrt{n} \sim N(0,1)$$

Теперь составим интервал, в который заданная статистика попадает с вероятностью $1-\alpha$, тогда с вероятность α она в него не попадает, и пусть она выходит за правую и левую граница интервала с одинаковыми вероятностями $\frac{\alpha}{2}$, так как это нормальный закон, то на графике это будет выглядеть так



$$P\left(u_{\alpha/2} \le \frac{\bar{X}-a}{\sigma}\cdot\sqrt{n} \le u_{1-\alpha/2}\right) = 1-\alpha$$

Далее, для квантилей $u_{\frac{\alpha}{2}}$ и $u_{1-\frac{\alpha}{2}}$ в силу симметрии имеет место соотношение $u_{\frac{\alpha}{2}} = -u_{1-\frac{\alpha}{2}}$.

Разрешаем относительно параметра a неравенство $-u_{1-\frac{\alpha}{2}} \le \frac{\bar{X}-a}{\sigma}\cdot\sqrt{n} \le u_{1-\frac{\alpha}{2}}$:

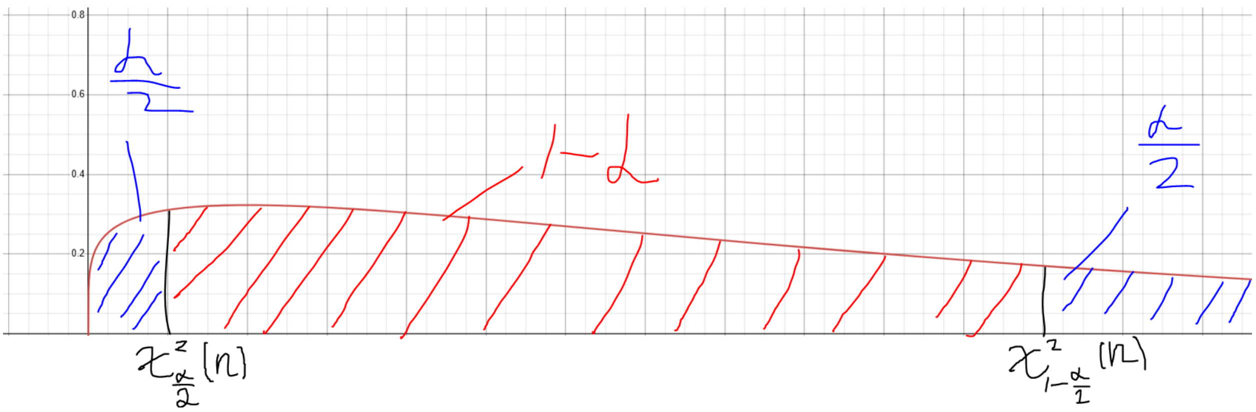
$$-u_{1-\alpha/2}\frac{\sigma}{\sqrt{n}} \le \bar{X} - a \le \frac{\sigma}{\sqrt{n}}u_{1-\alpha/2} \leftrightarrow \bar{X} - u_{1-\alpha/2}\frac{\sigma}{\sqrt{n}} \le a \le \bar{X} + \frac{\sigma}{\sqrt{n}}u_{1-\alpha/2}$$

Таким образом, $\left(\bar{X} - u_{1-\alpha/2}\frac{\sigma}{\sqrt{n}}; \bar{X} + \frac{\sigma}{\sqrt{n}}u_{1-\alpha/2}\right)$ –доверительный интервал для параметра a с доверительной вероятностью $1-\alpha$.

2. *Доверительный интервал для (неизвестного) с.к.о. σ при известном математическом ожидании* $\mathbf{a}$.

Рассмотрим статистику $S_0^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - a)^2$. Поскольку $X_i \sim N(a,\sigma)$, то $X_i - a \sim N(0,\sigma)$, $\frac{X_i - a}{\sigma} \sim N(0,1)$, следовательно, $\frac{S_0^2 \cdot n}{\sigma^2}$ распределена по закону $\chi^2(n)$, как сумма квадратов n стандартный нормальных велитчин

Теперь составим интервал, в который заданная статистика попадает с вероятностью $1-\alpha$, тогда с вероятность α она в него не попадает, и пусть она выходит за правую и левую граница интервала с одинаковыми вероятностями $\frac{\alpha}{2}$, так как это распределение хи квадрат с n степенями свободы, то на графике это будет выглядеть так



$$P\left(\chi^2_{\frac{\alpha}{2}}(n) \le \frac{S_0^2 \cdot n}{\sigma^2} \le \chi^2_{1-\alpha/2}(n)\right) = 1-\alpha$$

разрешаем неравенство $\chi^2_{\frac{\alpha}{2}}(n) \le \frac{S_0^2 \cdot n}{\sigma^2} \le \chi^2_{1-\alpha/2}(n)$ относительно параметра σ

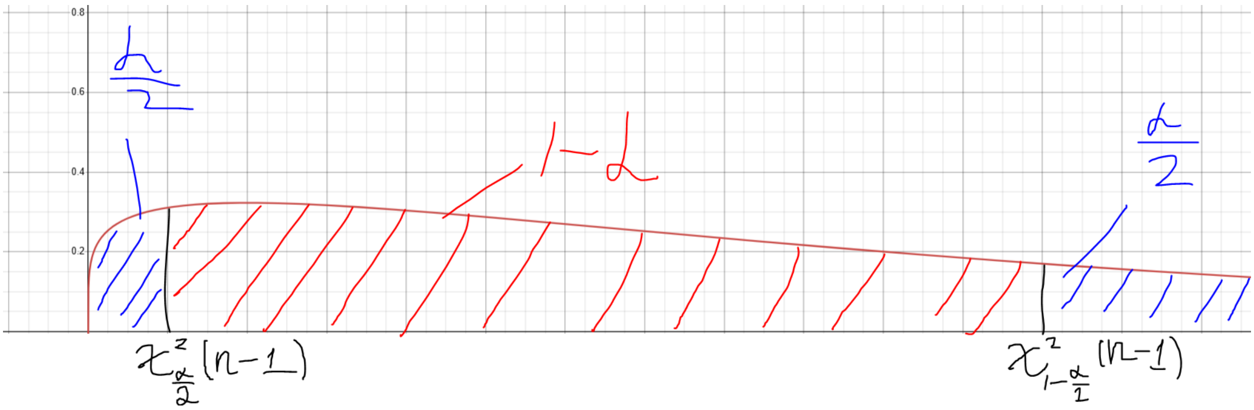
$\frac{S_0^2 \cdot n}{\chi^2_{1-\frac{\alpha}{2}}(n)} \le \sigma^2 \le \frac{S_0^2 \cdot n}{\chi^2_{\frac{\alpha}{2}}(n)}$, $\sqrt{\frac{S_0^2 \cdot n}{\chi^2_{1-\frac{\alpha}{2}}(n)}} \le \sigma \le \sqrt{\frac{S_0^2 \cdot n}{\chi^2_{\frac{\alpha}{2}}(n)}}$ – доверительный интервал для σ при известном a.

3. Доверительный интервал для (неизвестного) с.к.о. σ при неизвестном математическом ожидании а .

Рассмотрим статистику $S^2 = \frac{1}{n-1}\sum_{i=1}^{n}\left(X_i - \bar{X}\right)^2$.

По теореме статистика $\frac{S^2(n-1)}{\sigma^2}$ распределена по закону $\chi^2(n-1)$. Тогда, по аналогии с предыдущем случаем

$$P\left(\chi^2_{\frac{\alpha}{2}}(n-1) \le \frac{S_0^2 \cdot n}{\sigma^2} \le \chi^2_{1-\alpha/2}(n-1)\right) = 1-\alpha$$



Находим квантили $\chi^2_{\frac{\alpha}{2}}(n-1)$ и $\chi^2_{1-\frac{\alpha}{2}}(n-1)$, разрешаем неравенство

$\chi^2_{\frac{\alpha}{2}}(n-1) \le \frac{S^2(n-1)}{\sigma^2} \le \chi^2_{1-\frac{\alpha}{2}}(n-1)$ относительно σ:

$$\frac{S^2(n-1)}{\chi^2_{1-\frac{\alpha}{2}}(n-1)} \le \sigma^2 \le \frac{S^2(n-1)}{\chi^2_{\frac{\alpha}{2}}(n-1)} \ ,$$

$$\sqrt{\frac{S^2(n-1)}{\chi^2_{1-\frac{\alpha}{2}}(n-1)}} \le \sigma \le \sqrt{\frac{S^2(n-1)}{\chi^2_{\frac{\alpha}{2}}(n-1)}}$$

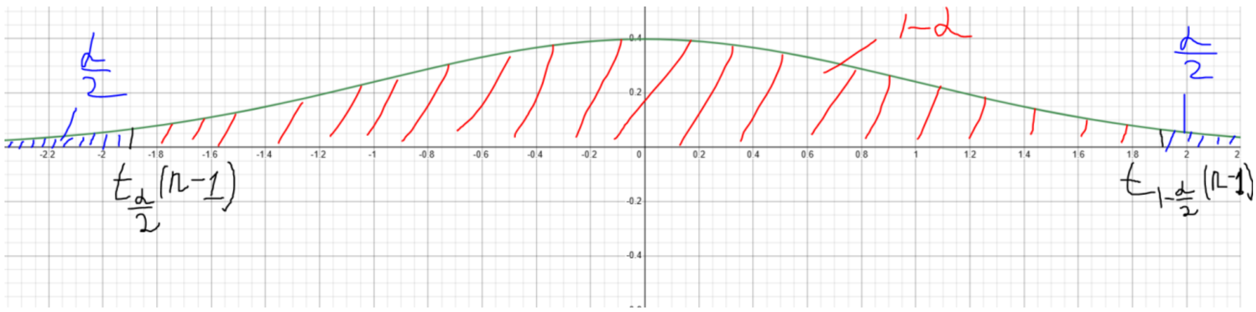
- доверительный интервал для σ при неизвестном a.

Примечание. Доверительный интервал для σ при неизвестном математическом ожидании a шире, чем при известном, что объяснятся меньшим количеством информации.

4. *Доверительный интервал для (неизвестного) математического ожидания* **a** *при неизвестном с.к.о.* **σ**.

По теореме, статистика $\frac{\bar{X}-a}{\sigma}\sqrt{n} \sim N(0,1)$ и независима от статистики $\frac{S^2(n-1)}{\sigma^2}$, распределенной по закону $\chi^2(n-1)$. Таким образом, отношение $\frac{\frac{\bar{X}-a}{\sigma}\sqrt{n}}{\sqrt{\frac{S^2}{\sigma^2}}}$ распределено по закону Стьюдента с $t(n-1)$.

Строим доверительный интервал:



$$1-\alpha = P\left(t_{\frac{\alpha}{2}}(n-1) \le \frac{\frac{\bar{X}-a}{\sigma}\sqrt{n}}{\sqrt{\frac{S^2}{\sigma^2}}} \le t_{1-\frac{\alpha}{2}}(n-1) \right) =$$

$$= P\left(t_{\frac{\alpha}{2}}(n-1) \le \frac{(\bar{X}-a)\sqrt{n}}{\sqrt{S^2}} \le t_{1-\frac{\alpha}{2}}(n-1) \right) =$$

$$= P\left(\bar{X} - \frac{S}{\sqrt{n}} t_{1-\frac{\alpha}{2}}(n-1) \le a \le \bar{X} + \frac{S}{\sqrt{n}} t_{1-\frac{\alpha}{2}}(n-1) \right)$$

$$\bar{X} - \frac{S}{\sqrt{n}} t_{1-\frac{\alpha}{2}}(n-1) \le a \le \bar{X} + \frac{S}{\sqrt{n}} t_{1-\frac{\alpha}{2}}(n-1)$$

Пример 1.

Пусть $\bar{X}=9,3; n=36,\ \alpha=0,05\ 1)\sigma=2\,кОм;\ 2)S^2=6.25\,кОм^2$

Построим доверительные интервалы

$$\bar{X}=9,3; n=36,\ \alpha=0,05\ ;\ \sigma=2\,кОм$$

$$\alpha=0,05\to u_{1-\alpha/2}=u_{0,975}=1,96$$

$$\left(\bar{X}-u_{1-\alpha/2}\frac{\sigma}{\sqrt{n}};\bar{X}+\frac{\sigma}{\sqrt{n}}u_{1-\alpha/2}\right)=\left(9,3-\frac{2}{6}\cdot 1,96;9,3+\frac{2}{6}\cdot 1,96\right)=\left(8,65;9,95\right)$$

$$\bar{X}=9,3; n=36,\ \alpha=0,05,\ S^2=6.25\,кОм^2$$

$$\alpha=0,05\to t_{1-\frac{\alpha}{2}}\left(n-1\right)=t_{0,975}\left(35\right)=2,03$$

$$\left(\bar{X}-\frac{S}{\sqrt{n}}t_{1-\frac{\alpha}{2}}\left(n-1\right)\le a\le \bar{X}+\frac{S}{\sqrt{n}}t_{1-\frac{\alpha}{2}}\left(n-1\right)\right)=$$

$$=\left(9,3-\frac{2,5}{6}\cdot 2,03;9,3+\frac{2,5}{6}\cdot 2,03\right)=\left(8,45;10,15\right)$$

5. Построение доверительного интервала для разности математических ожиданий двух независимых выборок из нормального распределения при известных значениях СКО

Пусть $X_1, X_2,\ldots.X_n$ –выборка из нормального распределения с параметрами a и σ_X ,

$Y_1, Y_2,\ldots.Y_m$ –выборка из нормального распределения с параметрами b и σ_Y .

Рассмотрим статистику $\bar{X}-\bar{Y}$, распределение которой нормально с параметрами

$M\left(\bar{X}-\bar{Y}\right)=a-b$, $D\left(\bar{X}-\bar{Y}\right)=D\bar{X}+D\left(\bar{Y}\right)=\frac{\sigma_X^2}{n}+\frac{\sigma_Y^2}{m}$.

Следовательно, статистика $\dfrac{\bar{X}-\bar{Y}-\left(a-b\right)}{\sqrt{\dfrac{\sigma_X^2}{n}+\dfrac{\sigma_Y^2}{m}}}$ распределена по стандартному нормальному закону $N\left(0,1\right)$.

Строим доверительный интервал:

$$1-\alpha = P\left(u_{\frac{\alpha}{2}} \le \frac{\bar{X}-\bar{Y}-(a-b)}{\sqrt{\frac{\sigma_X^2}{n}+\frac{\sigma_Y^2}{m}}} \le u_{1-\frac{\alpha}{2}} \right) =$$

$$= P\left(-u_{1-\frac{\alpha}{2}}\sqrt{\frac{\sigma_X^2}{n}+\frac{\sigma_Y^2}{m}} \le \bar{X}-\bar{Y}-(a-b) \le u_{1-\frac{\alpha}{2}}\sqrt{\frac{\sigma_X^2}{n}+\frac{\sigma_Y^2}{m}} \right) =$$

$$= P\left(\bar{X}-\bar{Y}-u_{1-\frac{\alpha}{2}}\sqrt{\frac{\sigma_X^2}{n}+\frac{\sigma_Y^2}{m}} \le a-b \le \bar{X}-\bar{Y}+u_{1-\frac{\alpha}{2}}\sqrt{\frac{\sigma_X^2}{n}+\frac{\sigma_Y^2}{m}} \right).$$

6. Построение доверительного интервала для разности математических ожиданий двух независимых выборок из нормального распределения при неизвестном СКО

Пусть $X_1, X_2, \dots X_n$ –выборка из нормального распределения с параметрами a и σ,

$Y_1, Y_2, \dots Y_m$ –выборка из нормального распределения с параметрами b и σ.

Параметр σ неизвестен, но известно, что $\dfrac{\bar{X}-\bar{Y}-(a-b)}{\sigma\sqrt{\frac{1}{n}+\frac{1}{m}}}$ распределено по

стандартному закону и не зависит от статистики $\dfrac{(n-1)S_X^2}{\sigma^2}+\dfrac{(m-1)S_Y^2}{\sigma^2}$,

подчиняющейся распределению $\chi^2(m+n-2)$.

После очевидных преобразований

$$\frac{\bar{X}-\bar{Y}-(a-b)}{\sqrt{(n-1)S_X^2+(m-1)S_Y^2}}\sqrt{nm}\sqrt{\frac{n+m-2}{n+m}} = \frac{\dfrac{\bar{X}-\bar{Y}-(a-b)}{\sigma\sqrt{\frac{1}{n}+\frac{1}{m}}}}{\sqrt{\dfrac{1}{n+m-2}\left(\dfrac{(n-1)S_X^2}{\sigma^2}+\dfrac{(m-1)S_Y^2}{\sigma^2}\right)}}$$, получаем

распределение

Стьюдента $t(n+m-2)$.

Найдем квантили $t_{1-\alpha/2}(n+m-2)$ и $t_{\alpha/2}(n+m-2) = -t_{1-\alpha/2}(n+m-2)$ и построим

доверительный интервал:

$$1-\alpha = P\left(t_{\frac{\alpha}{2}}\left(n+m-2\right) \le \frac{\overline{X}-\overline{Y}-\left(a-b\right)}{\sqrt{\frac{1}{n}+\frac{1}{m}}\sqrt{\frac{1}{n+m-2}\left(\sum_{i=1}^{n}\left(X_i-\overline{X}\right)^2+\sum_{i=1}^{m}\left(Y_i-\overline{Y}\right)^2\right)}} \le t_{1-\frac{\alpha}{2}}\left(n+m-2\right) \right).$$

Следовательно, с вероятностью $1-\alpha$

$$t_{\frac{\alpha}{2}}\left(n+m-2\right) \le \frac{\overline{X}-\overline{Y}-\left(a-b\right)}{\sqrt{\frac{1}{n}+\frac{1}{m}}\sqrt{\frac{1}{n+m-2}\left(\sum_{i=1}^{n}\left(X_i-\overline{X}\right)^2+\sum_{i=1}^{m}\left(Y_i-\overline{Y}\right)^2\right)}} \le t_{1-\frac{\alpha}{2}}\left(n+m-2\right) \Rightarrow$$

$$\begin{cases} \overline{X}-\overline{Y}-\left(a-b\right) \ge -t_{1-\frac{\alpha}{2}}\left(n+m-2\right)\sqrt{\frac{n+m}{nm\left(n+m-2\right)}}\sqrt{\left(n-1\right)S_X^2+\left(m-1\right)S_Y^2} \\ \overline{X}-\overline{Y}-\left(a-b\right) \le t_{1-\frac{\alpha}{2}}\left(n+m-2\right)\sqrt{\frac{n+m}{nm\left(n+m-2\right)}}\sqrt{\left(n-1\right)S_X^2+\left(m-1\right)S_Y^2} \end{cases} \Rightarrow$$

$a-b \in \left(\overline{X}-\overline{Y}-\delta; \overline{X}-\overline{Y}+\delta\right)$, где $\delta = t_{1-\frac{\alpha}{2}}\left(n+m-2\right)\sqrt{\frac{n+m}{nm\left(n+m-2\right)}}\sqrt{\left(n-1\right)S_X^2+\left(m-1\right)S_Y^2}$.

Распределение Фишера-Снедекора $\mathbf{F}\left(\boldsymbol{m},\boldsymbol{n}\right)$.

$$p_{\mathrm{F}(m,n)}\left(x\right) = \begin{cases} \dfrac{m^{\frac{m}{2}} n^{\frac{n}{2}} x^{\frac{m}{2}-1}}{\left(n+mx\right)^{\frac{m+n}{2}} \mathrm{B}\left(\frac{m}{2};\frac{n}{2}\right)}, & x>0 \\ 0, & x \le 0 \end{cases}$$

Теорема. Если случайные величины ξ_i, $i=1,\dots,m$, η_i, $j=1,\dots,n$, независимы и распределены по стандартному нормальному закону $\mathrm{N}\left(0,1\right)$, то плотность

распределения случайной величины $\zeta = \dfrac{\frac{1}{m}\left(\xi_1^2+\xi_2^2+\dots+\xi_m^2\right)}{\frac{1}{n}\left(\eta_1^2+\eta_2^2+\dots+\eta_n^2\right)}$

7. Построение доверительного интервала для отношения дисперсий двух независимых выборок из нормального распределения

Пусть $X_1, X_2, \dots X_n$ –выборка из нормального распределения с параметрами a и σ_X, $Y_1, Y_2, \dots Y_m$ –выборка из нормального распределения с параметрами b и σ_Y.

Рассмотрим статистику: $\dfrac{S_x^2}{S_y^2}\cdot\dfrac{\sigma_Y^2}{\sigma_X^2}=\dfrac{\dfrac{(n-1)S_x^2}{\sigma_X^2}}{\dfrac{(m-1)S_y^2}{\sigma_Y^2}}\cdot\dfrac{m-1}{n-1}$.

Отношение, стоящее справа имеет распределение Фишера $F(n-1,m-1)$.

Находим квантили $F_{\alpha/2}(n-1,m-1)$ и $F_{1-\alpha/2}(n-1,m-1)$ распределения Фишера и строим доверительный интервал для отношения $\dfrac{\sigma_Y^2}{\sigma_X^2}$:

$$1-\alpha=P\left(F_{\alpha/2}(n-1,m-1)\le\frac{S_x^2}{S_y^2}\cdot\frac{\sigma_Y^2}{\sigma_X^2}\le F_{1-\alpha/2}(n-1,m-1)\right),$$

$$1-\alpha=P\left(F_{\alpha/2}(n-1,m-1)\frac{S_y^2}{S_x^2}\le\frac{\sigma_Y^2}{\sigma_X^2}\le\frac{S_y^2}{S_x^2}F_{1-\alpha/2}(n-1,m-1)\right).$$

8. Построение доверительного интервала для отношения дисперсий двух независимых выборок из нормального распределения при известных средних

В случае известных средних значений, используя статистики $S_{0X}^2=\dfrac{1}{n}\sum_{i=1}^{n}(X_i-a)^2$ и $S_{0Y}^2=\dfrac{1}{m}\sum_{i=1}^{m}(Y_i-b)^2$, получаем аналогично

$$1-\alpha=P\left(F_{\alpha/2}(n,m)\frac{S_{0y}^2}{S_{0x}^2}\le\frac{\sigma_Y^2}{\sigma_X^2}\le\frac{S_{0y}^2}{S_{0x}^2}F_{1-\alpha/2}(n,m)\right).$$

Пример 2

Постройте доверительный интервал для оценки параметра λ, если выборка $X_1,\ldots,X_n$ объема n получена из показательного распределения

$$p(x,\theta)=\begin{cases}\lambda e^{-\lambda} & x\le 0\\ 0 & x>0\end{cases}.$$

Для распределений, которые не устойчивы относительно суммы (в результате сумме одинаково распределённых независимых величин получается другой закон распределения) можно воспользоваться центральной предельной теоремой.

Вспомним числовые характеристики показательного распределения

$$MX_i=\frac{1}{\lambda};DX_i=\frac{1}{\lambda^2}$$

В качестве исходной статистики для построения доверительно интервала используем выборочное среднее (так как оно является достаточной статистикой для оценки параметра λ). По центральной предельной теореме, выборочное среднее будет иметь нормальное распределение

$$\bar{X}=\frac{1}{n}\sum_{i=1}^{n}X_i\sim N\left(MX_i,\sqrt{\frac{DX_i}{n}}\right)\sim N\left(\frac{1}{\lambda},\frac{1}{\lambda\sqrt{n}}\right)$$

$$\frac{\bar{X}-\frac{1}{\lambda}}{\frac{1}{\lambda\sqrt{n}}}=\left(\lambda\bar{X}-1\right)\sqrt{n}\sim N(0,1)$$

Тогда можем записать

$$P\left(u_{\alpha/2}\le\left(\lambda\bar{X}-1\right)\sqrt{n}\le u_{1-\alpha/2}\right)=1-\alpha$$

Далее, для квантилей $u_{\frac{\alpha}{2}}$ и $u_{1-\frac{\alpha}{2}}$ в силу симметрии имеет место соотношение $u_{\frac{\alpha}{2}}=-u_{1-\frac{\alpha}{2}}$

Разрешаем относительно параметра λ неравенство:

$$u_{\alpha/2}\le\left(\lambda\bar{X}-1\right)\sqrt{n}\le u_{1-\alpha/2}$$

$$\frac{\frac{-u_{1-\alpha/2}}{\sqrt{n}}+1}{\bar{X}}\le\lambda\le\frac{\frac{u_{1-\alpha/2}}{\sqrt{n}}+1}{\bar{X}}$$

Таким образом, $\left(\dfrac{\dfrac{-u_{1-\alpha/2}}{\sqrt{n}}+1}{\bar{X}};\dfrac{\dfrac{u_{1-\alpha/2}}{\sqrt{n}}+1}{\bar{X}}\right)$ –доверительный интервал для параметра λ с доверительной вероятностью $1-\alpha$.

2-й способ

Можно использовать статистику $2\lambda\sum_{i=1}^{n}X_i$, которая будет распределена, по закону $\chi^2(n)$ (для доказательства рассмотрите закон распределения $2\lambda X_i$ и воспользуйтесь свойствами распределения хи-квадрат)

$$2\lambda\sum_{i=1}^{n}X_i \sim \chi^2(2n)$$

Тогда можем составить доверительный интервал

$$P\left(\chi^2_{\frac{\alpha}{2}}(2n)\le 2\lambda\sum_{i=1}^{n}X_i \le \chi^2_{1-\alpha/2}(2n)\right)=1-\alpha$$

$$P\left(\frac{\chi^2_{\frac{\alpha}{2}}(2n)}{2\sum_{i=1}^{n}X_i}\le\lambda\le\frac{\chi^2_{1-\alpha/2}(2n)}{2\sum_{i=1}^{n}X_i}\right)=1-\alpha$$

$$\frac{\chi^2_{\frac{\alpha}{2}}(2n)}{2\sum_{i=1}^{n}X_i}\le\lambda\le\frac{\chi^2_{1-\alpha/2}(2n)}{2\sum_{i=1}^{n}X_i}$$

Этот интервал будет точнее, чем интервал построенный первым способом, так как используется статистика распределённая по известному закону, а не приближенная, но при больших значениях n разница будет незначительной

Пример 3

Постройте доверительный интервал для оценки параметра θ, если выборка $X_1,\ldots,X_n$ объема n получена из распределения с плотностью

$$p(x,\theta)=\begin{cases} e^{x-\theta} & x\le\theta \\ 0 & x>\theta \end{cases}.$$

Указание. Используйте статистику $X_{(n)}$.

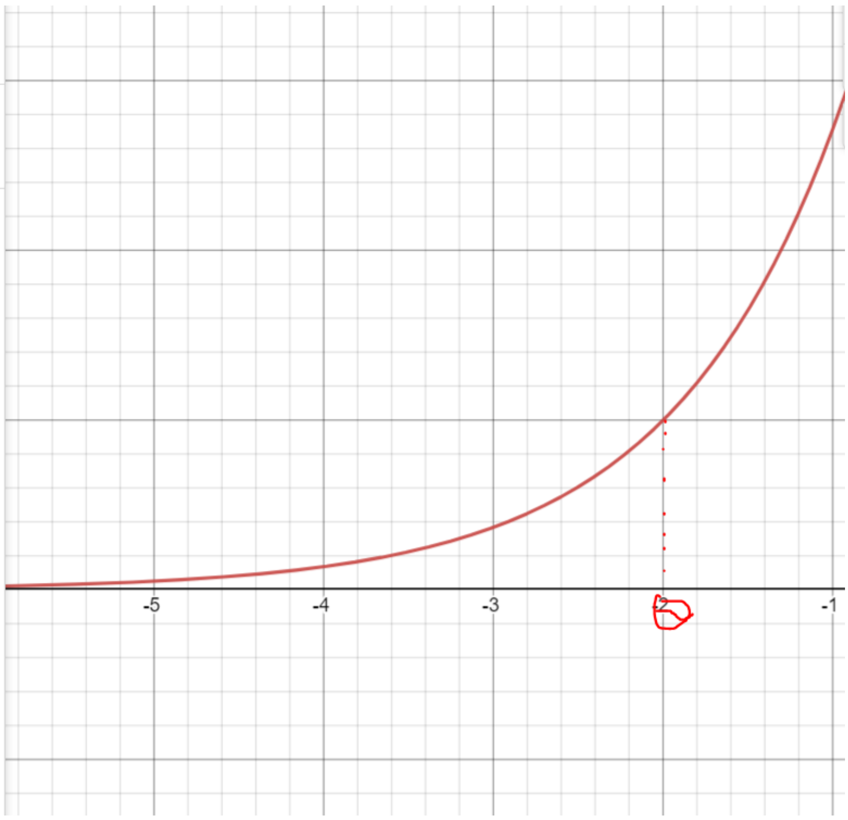
Найдём функцию распределения каждого элемента выборки

$$F(x,\theta)=\begin{cases} \int\limits_{-\infty}^{x} e^{x-\theta}, & x\le\theta \\ 0, & x>\theta \end{cases}=\begin{cases} e^{x-\theta}, & x\le\theta \\ 0, & x>\theta \end{cases}$$

Тогда функция распределения порядковой статистики $X_{(n)}$ имеет вид

$$F_{X_{(n)}}(x,\theta)=\left(F(x,\theta)\right)^n=\begin{cases} \left(e^{x-\theta}\right)^n, & x\le\theta \\ 0, & x>\theta \end{cases}=\begin{cases} e^{n(x-\theta)}, & x\le\theta \\ 0, & x>\theta \end{cases}$$

Изобразим график плотности распределения выборки



Параметр θ отвечает за смещение графика от оси y, и θ это максимальное значение, которое принимает случайная величина

Так как $X_{(n)}$ это максимальный элемент выборки, и он не может быть больше параметра θ. Тогда можем записать доверительный интервал в таком виде

$$P\left(\theta - d < X_{(n)} < \theta\right) = 1 - \alpha$$

$$P\left(-d < X_{(n)} - \theta < 0\right) = 1 - \alpha$$

Отметим, что статистика $X_{(n)} - \theta$ не зависит от параметра θ,

Но для начала найдём значение d, потом разреши неравенство относительно θ

Так как функция распределения $X_{(n)}$ известна, то можем записать

$$P\left(\theta - d < X_{(n)} < \theta\right) = F_{X_{(n)}}\left(\theta\right) - F_{X_{(n)}}\left(\theta - d\right) = e^{n(\theta-\theta)} - e^{n(\theta-d-\theta)} = 1 - e^{-nd}$$

$$1 - e^{-nd} = 1 - \alpha \Rightarrow e^{-nd} = \alpha \Rightarrow \ln\alpha = -nd \Rightarrow d = \frac{1}{n}\ln\frac{1}{\alpha}$$

Теперь разрешим неравенство относительно параметра θ

$$\theta - \frac{1}{n}\ln\frac{1}{\alpha} < X_{(n)} < \theta$$

$$-X_{(n)} - \frac{1}{n}\ln\frac{1}{\alpha} < -\theta < -X_{(n)}$$

$$X_{(n)} < \theta < X_{(n)} + \frac{1}{n}\ln\frac{1}{\alpha}$$

Пример 4

Постройте доверительный интервал для оценки параметра λ, если выборка $X_1,\ldots,X_n$ объема n получена из биномиального распределения

$$P(\xi = j) = C_k^j p^j (1-p)^{k-j}.$$

Можно воспользоваться центральной предельной теоремой.

Вспомним числовые характеристики биномиального распределения

$MX_i = kp; DX_i = kp(1-p)$

В качестве исходной статистики для построения доверительно интервала используем выборочное среднее (так как оно является достаточной статистикой для оценки параметра p). По центральной предельной теореме, выборочное среднее будет иметь нормальное распределение

$$\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i \sim N\left(MX_i, \sqrt{\frac{DX_i}{n}}\right) \sim N\left(kp, \frac{\sqrt{kp(1-p)}}{\sqrt{n}}\right)$$

$$\frac{\bar{X} - kp}{\sqrt{kp(1-p)}}\sqrt{n} \sim N(0,1)$$

Тогда можем записать

$$P\left(u_{\alpha/2} \le \frac{\bar{X} - kp}{\sqrt{kp(1-p)}}\sqrt{n} \le u_{1-\alpha/2}\right) = 1-\alpha$$

Полученное неравенство не просто преобразовать относительно p, так как этот параметр содержится и в числителе, и в знаменателе. Для упрощения задачи будет в знаменателе вместо истинного значения параметра p использовать его оценку $\hat{p} = \frac{\bar{X}}{k}$

$$P\left(u_{\alpha/2} \le \frac{\bar{X} - kp}{\sqrt{k\frac{\bar{X}}{k}\left(1 - \frac{\bar{X}}{k}\right)}}\sqrt{n} \le u_{1-\alpha/2}\right) = 1-\alpha$$

Далее, для квантилей $u_{\frac{\alpha}{2}}$ и $u_{1-\frac{\alpha}{2}}$ в силу симметрии имеет место соотношение $u_{\frac{\alpha}{2}} = -u_{1-\frac{\alpha}{2}}$

Разрешаем относительно параметра p неравенство:

$$-u_{\alpha/2} \le \frac{\bar{X} - kp}{\sqrt{\bar{X}\left(1-\dfrac{\bar{X}}{k}\right)}}\sqrt{n} \le u_{1-\alpha/2}$$

$$\frac{\bar{X} - \dfrac{u_{1-\alpha/2}\cdot\sqrt{\bar{X}\left(1-\dfrac{\bar{X}}{k}\right)}}{\sqrt{n}}}{k} \le p \le \frac{\bar{X} + \dfrac{u_{1-\alpha/2}\cdot\sqrt{\bar{X}\left(1-\dfrac{\bar{X}}{k}\right)}}{\sqrt{n}}}{k}$$

$$\frac{\bar{X}}{k} - \frac{u_{1-\alpha/2}\cdot\sqrt{\bar{X}\left(1-\dfrac{\bar{X}}{k}\right)}}{k\sqrt{n}} \le p \le \frac{\bar{X}}{k} + \frac{u_{1-\alpha/2}\cdot\sqrt{\bar{X}\left(1-\dfrac{\bar{X}}{k}\right)}}{k\sqrt{n}}$$