Доверительные интервалы (ДИ)

1. Импортёр упаковывает чай в пакеты с номинальным весом 125 грамм. Известно, что упаковывающая машина работает с известным среднеквадратическим отклонением равным 10 г. Случайным образом для проверки было выбрано 50 пакетов чая. Выборочное среднее этих пакетов составило 125,8 г. Постройте доверительный интервал уровня надёжности 0.95 для среднего веса пакета. Найдите объём выборки n, при котором длина доверительного интервала будет равна 2г.

2. Покажите, что $s^2 = \widehat{\mu_2} - (\widehat{\mu_1})^2$.

3. Имеются следующие данные об урожайности пшеницы в России за 1998-2001 годы (в ц/га): 27.0; 31.0; 35.0; 47.0. Постройте доверительный интервал уровня надёжности 0.95 для среднего значения урожайности пшеницы.

4. Станок-автомат штампует валики. По выборке объёма n=17 вычислено выборочное среднее $\bar{X}$= 20,5 мм и выборочная дисперсия s^2= 16 (мм)^2 диаметра валиков. Постройте доверительные интервалы уровня надёжности 0,9 для среднего значения диаметра валика и для дисперсии диаметра валика. Предполагается, что диаметры валиков имеют гауссовское распределение.

5. Имеются независимые выборки $X_1, \dots, X_n$ (соответствует распределению $N(m_1, \sigma^2)$) и выборка $Y_1, \dots, Y_k$ (соответствует распределению $N(m_2, \sigma^2)$). Пользуясь утверждением теоремы 4.2 из лекции 4, постройте центральный доверительный интервал для разности средних $m_1 - m_2$ гауссовских величин с одинаковыми, но неизвестными дисперсиями σ^2.

6. В метеорологии принято характеризовать температуру месяца ее средним значением (среднее значение температуры месяца равно сумме температур всех дней данного месяца, деленной на число дней в этом месяце). В таблице ниже приведены значения средней температуры января в г. Саратове и г. Алатыре.

Год	1891	1892	1893	1894	1895	1896	1897
Саратов	−19,2	−14,8	−19,6	−11,1	−9,4	−16,9	−13,7
Алатырь	−21,8	−15,4	−20,8	−11,3	−11,6	−19,2	−13,0

Год	1899	1911	1912	1913	1914	1915	
Саратов	−4,9	−13,9	−9,4	−8,3	−7,9	−5,3	
Алатырь	−7,4	−15,1	−14,4	−11,1	−10,5	−7,2	

Требуется построить доверительный интервал уровня надёжности 0.95 для разности средних значений январских температур в городах Саратове и Алатыре предполагая, что: а) дисперсия среднеянварской температуры в Саратове равна 22, а в Алатыре - 20; б) дисперсия температуры неизвестна, но одинакова для г. Саратова и г. Алатырь. Будем считать, что наблюдения имеют гауссовское распределение.

Домашнее задание

1. Инженер предложил новый метод сборки изделий. В результате сборки 10 изделий новым способом время сборки составило 79, 74, 112, 95, 83, 96, 77, 84, 70, 90 минут. Постройте ДИ уровня надёжности 0.95 для среднего времени сборки детали. Предполагается, что наблюдения имеют гауссовское распределение.

2. Постройте ДИ уровня надёжности 0.9 для дисперсии времени сборки деталей из предыдущей задачи.

3. Имеются данные Федеральной службы государственной статистики о среднедушевых доходах населения (рублей в месяц) в 2008 году по некоторым областям Центрального и Приволжского ФО. По ЦФО - 10043; 9596;10305; 8354; 9413; 19776; 9815; 11311; 11253; 10856; 11389. По ПФО – 14253;7843; 9581;8594; 16119; 10112;10173;9756. Постройте доверительный интервал уровня надёжности 0.95 для разности средних значений среднедушевых доходов населения Центрального и Приволжского ФО.

08.02

Теорема Фишера

$Z_n = (x_1 \dots x_n) \sim N(m_x, \sigma_x^2)$

$\bar{x}$ – выборочное среднее

$\hat{D}x$ – выборочная дисперсия

1) $\frac{(\bar{x} - m_x)\sqrt{n}}{\sigma_x} \sim N(0,1)$

2) $\frac{n\hat{D}_x}{\sigma_x^2} \sim \chi^2(n-1)$ → „хи квадрат"

3) $\bar{x}$ и $\hat{D}x$ – независимы

4) $\frac{(\bar{x} - m)\sqrt{n-1}}{\sqrt{\hat{D}x}} \sim t(n-1)$

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

$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$

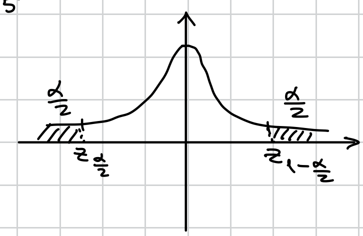
1. Импортёр упаковывает чай в пакеты с номинальным весом 125 грамм. Известно, что упаковывающая машина работает с известным среднеквадратическим отклонением равным 10 г. Случайным образом для проверки было выбрано 50 пакетов чая. Выборочное среднее этих пакетов составило 125,8 г. Постройте доверительный интервал уровня надёжности 0.95 для среднего веса пакета. Найдите объём выборки n, при котором длина доверительного интервала будет равна 2г.

$\sigma = 10$, $n = 50$, $\bar{x} = 125,8$, $\gamma = 0,95 \Rightarrow \alpha = 0,05$

$P\{T_1 \le m_x \le T_2\} = \gamma$

T_1, T_2 – ?

$T = \frac{(\bar{x} - m_x)\sqrt{n}}{\sigma_x} \sim N(0;1)$



$P\{z < \frac{(\bar{x} - m_x)\sqrt{n}}{\sigma_x} < z\} = \gamma$

$P\{z_{\frac{\alpha}{2}} < \frac{(\bar{x} - m_x)\sqrt{n}}{\sigma_x} < z_{1-\frac{\alpha}{2}}\} = \gamma$

$z_{\frac{\alpha}{2}} < \frac{(\bar{x} - m_x)\sqrt{n}}{\sigma_x} < z_{1-\frac{\alpha}{2}}$

$-z_{1-\frac{\alpha}{2}} < \frac{(\bar{x} - m_x)\sqrt{n}}{\sigma_x} < z_{1-\frac{\alpha}{2}}$

$-z_{1-\frac{\alpha}{2}}\sigma_x < (\bar{x} - m_x)\sqrt{n} < z_{1-\frac{\alpha}{2}}\sigma_x$

$\frac{-z_{1-\frac{\alpha}{2}}\sigma_x}{\sqrt{n}} - \bar{x} < -m_x < \frac{z_{1-\frac{\alpha}{2}}\sigma_x}{\sqrt{n}} - \bar{x}$

$\bar{x} - \frac{z_{1-\frac{\alpha}{2}}\sigma_x}{\sqrt{n}} < m_x < \bar{x} + \frac{z_{1-\frac{\alpha}{2}}\sigma_x}{\sqrt{n}}$

$z_{1-\frac{\alpha}{2}} = z_{0,975} = z_{0,475} = 1,96$

$125,8 - \frac{1,96 \cdot 10}{\sqrt{50}} < m_x < 125.8 + \frac{1,96 \cdot 10}{\sqrt{50}}$

$123,02 < m_x < 128,57$

$2 \cdot \frac{z_{1-\frac{\alpha}{2}}\sigma_x}{\sqrt{n}} = 2$

$z_{1-\frac{\alpha}{2}}\sigma_x = \sqrt{n}$

$n = (z_{1-\frac{\alpha}{2}})^2$

$n = 384.16 \Rightarrow n = 385$

2. Покажите, что $s^2 = \widehat{\mu_2} - (\widehat{\mu_1})^2$.

$\hat{\mu}_2 = \frac{1}{n}\sum_{i=1}^{n} x_i^2$

$\overline{x^2} \neq \overline{x}^2$

$(\hat{\mu}_2) = (\frac{1}{n}\sum_{i=1}^{n} x_i)^2$

$\hat{D}_x = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i^2 - 2x_i\bar{x} + \bar{x}^2) = \frac{1}{n}(\sum_{i=1}^{n} x_i^2 - 2\bar{x}\sum_{i=1}^{n} x_i + \sum_{i=1}^{n}\bar{x}^2) = \frac{1}{n}\sum_{i=1}^{n} x_i^2 - \frac{2\bar{x}}{n}\sum_{i=1}^{n} x_i + \frac{n\cdot\bar{x}^2}{n} = \overline{x^2} - 2\bar{x}\cdot\bar{x} +$

$+ \bar{x}^2 = \overline{x^2} - \bar{x}^2$

3. Имеются следующие данные об урожайности пшеницы в России за 1998-2001 годы (в ц/га): 27.0; 31.0; 35.0; 47.0. Постройте доверительный интервал уровня надёжности 0.95 для среднего значения урожайности пшеницы.

$x_1 = 27.0$

$x_2 = 31.0$

$x_3 = 35.0$

$x_4 = 47.0$

$n = 4$

$\gamma = 0,95$

$\alpha = 1 - 0,95 = 0,05$

$P(T_1 \le m_x \le T_2) = \gamma$

$T = \frac{(\bar{x} - m_x)\sqrt{n-1}}{\sqrt{\hat{D}_x}} \sim t(n-1)$

$P\{t < \frac{(\bar{x} - m_x)\sqrt{n-1}}{\sqrt{\hat{D}_x}} < t\} = 0,95$

$\frac{t_{\frac{\alpha}{2}}\cdot\sqrt{\hat{D}_x}}{\sqrt{n-1}} < (\bar{x} - m_x) < \frac{t_{1-\frac{\alpha}{2}}\cdot\sqrt{\hat{D}_x}}{\sqrt{n-1}}$

$\bar{x} - \frac{t_{1-\frac{\alpha}{2}}\cdot\sqrt{\hat{D}_x}}{\sqrt{n-1}} < m_x < \bar{x} - \frac{t_{\frac{\alpha}{2}}\cdot\sqrt{\hat{D}_x}}{\sqrt{n-1}}$

$\bar{x} = 35$

$n - 1 = 3$

$\hat{D}_x = \frac{1}{n}\sum_{i=1}^{n}(x_i - x)^2 = \overline{x^2} - (\bar{x})^2 = 56$

$t_{1-\frac{\alpha}{2}} = t_{0,975} = 3,182$

$t_{\frac{\alpha}{2}} = t_{0,025} = -3,182$

$35 - \frac{3,182\cdot\sqrt{56}}{\sqrt{3}} < m_x < 35 + \frac{3,182\cdot\sqrt{56}}{\sqrt{3}}$

$21,252 < m_x < 48,748$

4. Станок-автомат штампует валики. По выборке объёма n=17 вычислено выборочное среднее $\bar{X}$= 20,5 мм и выборочная дисперсия s^2= 16 (мм)^2 диаметра валиков. Постройте доверительные интервалы уровня надёжности 0,9 для среднего значения диаметра валика и для дисперсии диаметра валика. Предполагается, что диаметры валиков имеют гауссовское распределение.

$n = 17$

$\bar{X} = 20,5$

$\hat{D}_x = 16$

$\sqrt{\hat{D}_x} = 4$

$\gamma = 0,9 \Rightarrow \alpha = 0,1$

ДИ для D_x

↓

доверительный интервал

$P(T_1 \le \sigma_x^2 \le T_2) = \gamma = 1 - \alpha$

$T_1, T_2 - ?$

$T = \frac{n\hat{D}_x}{\sigma^2} \sim \chi^2(n-1)$

$$P\left\{\chi^2_{\frac{\alpha}{2}} \le \frac{n\hat{D}_x}{\sigma^2_x} \le \chi^2_{1-\frac{\alpha}{2}}\right\} = \gamma$$

$$P\left\{\frac{\chi^2_{\frac{\alpha}{2}}}{n\hat{D}_x} \le \sigma_x^2 \le \frac{\chi_{1-\frac{\alpha}{2}}}{n\hat{D}_x}\right\} = \gamma$$

$$\chi^2_{\frac{\alpha}{2}} = 7{,}96 \qquad \chi^2_{1-\frac{\alpha}{2}} = 26{,}29$$

$$\frac{7{,}96}{17\cdot 16} \le \sigma_x^2 \le \frac{26{,}29}{17\cdot 1}$$

$$0{,}03 \le \sigma_x^2 \le 0{,}097$$