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

$\widehat{D}x$ – выборочная дисперсия

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

2) $\frac{n\widehat{D}x}{\sigma_x^2} \sim \underbrace{\chi^2}_{\text{„хи квадрат"}}(n-1)$

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

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

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

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

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

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

$\frac{(\bar{X} - m)\sqrt{n-1}}{\sqrt{\widehat{D}x}}$

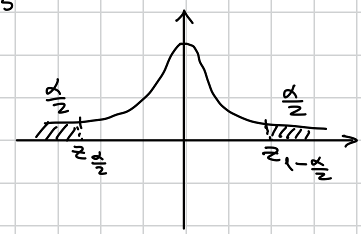
1. Импортёр упаковывает чай в пакеты с номинальным весом 125 грамм. Известно, что упаковывающая машина работает с известным среднеквадратическим отклонением равным 10 г. Случайным образом для проверки было выбрано 50 пакетов чая. Выборочное среднее этих пакетов составило 125,8 г. Постройте доверительный 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 $s^2 = \widehat{\mu_2} - (\widehat{\mu_1})^2$.

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

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

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

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

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

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.

$x_1 = 27.0$
$x_2 = 31.0$
$x_3 = 35.0$
$x_4 = 47.0$
$n = 4$
$\gamma = 0{,}95$
$\alpha = 1 - 0{,}95 = 0{,}05$
$P(T_1 \le m_x \le T_2) = \gamma$

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

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

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

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

$x_1 = 27$
$x_2 = 31$
$x_3 = 35$
$x_4 = 47$
$n = 4$
$\gamma = 0{,}95$
$\alpha = 1 - 0{,}95 = 0{,}05$
$P(T_1 \le m_x \le T_2) = \gamma$

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

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

$\overline{x} = 35$
$n - 1 = 3$
$\hat{D}x = \frac{1}{n}\sum_{i=1}^{n}(x_i - x)^2 = \overline{x^2} - (\overline{x})^2 = 56$

$\overline{x^2} = \frac{1}{n}\sum_{i=1}^{n} x_i^2 =$
$= \frac{1}{4}(27^2 + 31^2 + 35^2 + 47^2) = 1281$
$\overline{x} = \frac{1}{4}(27 + 31 + 35 + 47) = 35$
$1281 - 35^2 = 56$

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

улыбок тебе,
без Мокар!

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

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

$21{,}252 < m_x < 48{,}748$

$\hat{D}_x = \overline{x^2} - (\overline{x})^2$

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.

$n = 17$
$\overline{x} = 20{,}5$
$\hat{D}x = 16$
$\sqrt{\hat{D}x} = 4$

$\gamma = 0{,}9 \Rightarrow \alpha = 0{,}1$
ДИ для Dx
↓
доверительный интервал

$P(T_1 \le \sigma_x^2 \le T_2) = \gamma = 1 - \alpha$
$T_1, T_2 - ?$
$T = \frac{n\hat{D}x}{\sigma^2} \sim \chi^2(n-1)$

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

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

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

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

$$0{,}03 \le \sigma^2_x \le 0{,}097$$

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.

$$x_1, \dots, x_n \sim N(m_1, \sigma^2);$$

$$x_1, \dots, x_n \sim N(m_2, \sigma^2)$$

$$\overline{x} = \frac{\sum_{i=1}^{n} x_i}{n} \qquad S_x^2 = \frac{1}{n-1} \sum_{i=1}^{n} (x_i - \overline{x})^2$$

$$\overline{y} = \frac{\sum_{i=1}^{k} y_i}{k} \qquad S_y^2 = \frac{1}{k-1} \sum_{i=1}^{k} (y_i - \overline{y})^2$$

$$S_{xy} = \sqrt{\frac{(n-1)S_x^2 + (k-1)S_y^2}{n+k-2}}$$

св-ва:

1) $\frac{S_x^2}{S_y^2} \sim F(n-1, k-1)$

2) $\frac{(\overline{x} - \overline{y}) - (m_1 - m_2)}{\sqrt{\frac{(n-1)S_x^2 + (k-1)S_y^2}{n+k-2}}} \sim t(n+k-2)$

$$P(T_1 < m_1 - m_2 < T_2) = 1 - \alpha$$

$$P(t_{\frac{\alpha}{2}} < T < t_{1-\frac{\alpha}{2}}) = 1 - \alpha$$

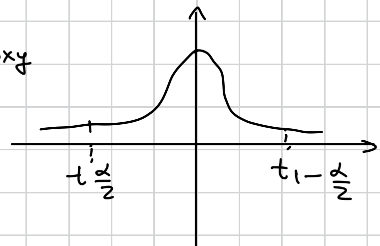
$$t_{\frac{\alpha}{2}} < \frac{(\overline{x} - \overline{y}) - (m_1 - m_2)}{\underbrace{\sqrt{\frac{(n-1)S_x^2 + (k-1)S_y^2}{n+k-2}}}_{*}} < t_{1-\frac{\alpha}{2}}$$

$$t_{\frac{\alpha}{2}} \cdot * < (\overline{x} - \overline{y}) - (m_1 - m_2) < t_{1-\frac{\alpha}{2}} \cdot *$$

$$t_{\frac{\alpha}{2}} \cdot S_{xy} - (\overline{x} - \overline{y}) < -(m_1 - m_2) < t_{1-\frac{\alpha}{2}} - (\overline{x} - \overline{y})$$

$$(\overline{x} - \overline{y}) - t_{1-\frac{\alpha}{2}} \cdot S_{xy} < m_1 - m_2 < (\overline{x} - \overline{y}) - t_{\frac{\alpha}{2}} S_{xy}$$

$$(\overline{x} - \overline{y}) - t_{1-\frac{\alpha}{2}} \cdot S_{xy} < (m_1 - m_2) < (\overline{x} - \overline{y}) + t_{1-\frac{\alpha}{2}} S_{xy}$$



$$t_{\frac{\alpha}{2}} = -t_{1-\frac{\alpha}{2}}$$

Вариант на 8

Фамилия Имя **СР-02**

1 (5 баллов). Найти оценку максимального правдоподобия параметра α > 0 распределения с плотностью

$$f(y;\alpha) = \begin{cases} 3y^2\alpha^{-3}e^{-(y/\alpha)^3} & \text{при } y \geq 0, \\ 0 \text{ при } y < 0 \end{cases}$$

2 (3 балла). Пусть $X = (X_1, \dots, X_n)$ – случайная выборка из равномерного распределения на отрезке $[-\theta;\theta]$, $\theta > 0$. При помощи метода моментов найдите оценку для неизвестного параметра θ. Для решения используйте начальные моменты второго порядка.

1 (5 баллов). Найти оценку максимального правдоподобия параметра $\alpha > 0$ распределения с плотностью

$$f(y;\alpha) = \begin{cases} 3y^2\alpha^{-3}e^{-(y/\alpha)^3} & \text{при } y \ge 0, \\ 0 \text{ при } y < 0 \end{cases}$$

$$f(y;\alpha) = \begin{cases} 3y^2\alpha^{-3}e^{-\left(\frac{y}{\alpha}\right)^3}, & y>0 \\ 0, & y\le 0 \end{cases}$$

$$\mathcal{L}(\alpha) = \prod_{i=1}^{n} 3y_i^2\alpha^{-3}e^{-\left(\frac{y_i}{\alpha}\right)^3} = 3^n\alpha^{-3n}\prod_{i=1}^{n} y_i^2 \cdot e^{-\sum_{i=1}^{n}\left(\frac{y_i}{\alpha}\right)^3}$$

$$\ln(\mathcal{L}(\alpha)) = n\ln 3 - 3n\ln\alpha + \sum_{i=1}^{n}\ln y_i^2 - \sum_{i=1}^{n}\left(\frac{y_i}{\alpha}\right)^3$$

$$\frac{\partial \ln(\mathcal{L}(\alpha))}{\partial\alpha} = -\frac{3n}{\alpha} + 3\sum_{i=1}^{n}\frac{y_i^3}{\alpha^4}$$

$$-\frac{3n}{\alpha} + 3\sum_{i=1}^{n}\frac{y_i^3}{\alpha^4} = 0$$

$$\frac{n}{\alpha} = \sum_{i=1}^{n}\frac{y_i^3}{\alpha^4}$$

$$n\alpha^3 = \sum_{i=1}^{n} y_i^3$$

$$\hat{\alpha} = \sqrt[3]{\frac{\sum_{i=1}^{n} y_i^3}{n}}$$

2 (3 балла). Пусть $X = (X_1, \dots, X_n)$ – случайная выборка из равномерного распределения на отрезке $[-\theta;\theta]$, $\theta > 0$. При помощи метода моментов найдите оценку для неизвестного параметра θ. Для решения используйте начальные моменты второго порядка.

$X \sim R(-\theta, \theta)$ $\qquad E x = \frac{-\theta+\theta}{2} = 0$ $\qquad f(x) = \frac{1}{2\theta}$

$$E[X] = \frac{-\theta+\theta}{2} = 0$$

$$E[X^2] = \int_{-\theta}^{\theta} x^2 f(x)\,dx = \int_{-\theta}^{\theta} x^2\cdot\frac{1}{2\theta}\,dx = \frac{1}{2\theta}\int_{-\theta}^{\theta} 2x\,dx = \frac{1}{2\theta}\cdot\frac{2\theta^3}{3} = \frac{\theta^2}{3}$$

$$\int_{-\theta}^{\theta} x^2\,dx = \frac{x^3}{3}\Big|_{-\theta}^{\theta} = \frac{\theta^3}{3} + \frac{\theta^3}{3} = \frac{2\theta^3}{3}$$

$$\frac{\theta^2}{3} = \frac{1}{n}\sum_{i=1}^{n} x_i^2$$

$$\hat{\theta} = \sqrt{\frac{3}{n}\sum_{i=1}^{n} x_i^2}$$

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.

$n = 13$, $\gamma = 0{,}95$, $\alpha = 0{,}05$ $\qquad E\bar{X} = EX_1$ $\qquad D\bar{X} = \frac{DX}{n}$

1) $\sigma_x^2 = 22$ $\qquad \sigma_y^2 = 20$ $\qquad D[X-Y] = D(X) + D(Y) - 2\operatorname{cov}(X,Y)$

$\bar{x} = -11{,}88$ $\qquad \bar{y} = -13{,}75$ $\qquad \bar{x} - \bar{y} = 1{,}87$

$\bar{x} - \bar{y} \sim N\left(m_1 - m_2; \frac{\sigma_1^2}{n} + \frac{\sigma_2^2}{n}\right)$ $\qquad D(\bar{x} - \bar{y}) = \frac{Dx}{n} + \frac{Dy}{n}$

$$\frac{\bar{x} - \bar{y} - (m_1 - m_2)}{\sqrt{\frac{\sigma_1}{n} + \frac{\sigma_2}{n}}} \sim N(0;1)$$

$$P(T_1 \le m_1 - m_2 \le T_2) = 0{,}95$$

$$z_{\frac{\alpha}{2}} \le \frac{\bar{x} - \bar{y} - (m_1 - m_2)}{\sqrt{\frac{\sigma_1^2}{n} + \frac{\sigma_2^2}{n}}} \le z_{1-\frac{\alpha}{2}}$$

$$(m_1 - m_2) \in \left[\bar{x} - \bar{y} \pm z_{1-\frac{\alpha}{2}}\sqrt{\frac{1}{n}(\sigma_1^2 + \sigma_2^2)}\right]$$

$$(m_1 - m_2) \in \left[1{,}87 \pm 1{,}96\cdot\sqrt{\frac{42}{13}}\right]$$

строим доверит интервал $\qquad X \sim N(m_1; \sigma^2)$

$P(T_1 \le m_1 - m_2 \le T_2)$ $\qquad Y \sim N(m_2; \sigma^2)$

$$\frac{\overline{x}-\overline{y}-(m_1-m_2)}{\sqrt{\frac{(n-1)s_x^2+(n-1)s_y^2}{2n-2}}} \sim t(2n-2)$$

посмотрели
по табличке
что x и y
одинаков. число

$$m_1-m_2 \in [1{,}87 \pm 2{,}064 \cdot 4{,}78]$$

Асимптотические доверительные интервалы (ДИ)

1. По выборке $X_1, \dots, X_n$,соответствующей распределению Пуассона $\Pi(\theta)$, постройте асимптотический доверительный интервал для параметра θ.

2. Для проверки качества деталей из большой партии выбрали 200 деталей. Среди выбранных деталей оказалось 12 бракованных. Постройте доверительный интервал уровня надёжности 0.95 для доли бракованных деталей.

3. Проводится две серии независимых испытаний по схеме Бернулли. В первой серии проведено n_1 испытаний, во второй - n_2 испытаний. Предполагается, что n_1 и n_2 достаточно велики. Построить асимптотический доверительный интервал для разности вероятностей "успеха" в двух сериях. (Показать, что статистика $\frac{\hat{p}_1-\hat{p}_2-(p_1-p_2)}{\sqrt{\widehat{D}(\hat{p}_1-\hat{p}_2)}}$, где $\widehat{D}(\hat{p}_1-\hat{p}_2)=\frac{\hat{p}_1(1-\hat{p}_1)}{n_1}+\frac{\hat{p}_2(1-\hat{p}_2)}{n_2}$, имеет асимптотически стандартное нормальное распределение).

4. Два года назад на втором курсе ОПИ ФКН обучалось 252 студента. На экзамене по ТВиМС 29 человек получили неудовлетворительные оценки. В прошлом учебном году на втором курсе ОПИ ФКН обучалось 286 студентов. На экзамене по ТВиМС 42 человека получили неудовлетворительные оценки. Постройте асимптотический доверительный интервал уровня надёжности 0.95 для разности вероятностей «успеха» по выборкам прошлого и позапрошлого года.

Проверка гипотез в биномиальной модели

5. Для проверки качества деталей из большой партии выбрали 200 деталей. Среди выбранных деталей оказалось 12 бракованных. Считается, что партия качественная, если она содержит не более 5% брака. Можно ли считать (на уровне значимости 0.05), что данная партия качественная? (! H_a односторонняя)

6. Известно, что вероятность рождения мальчика равна 0.52. В случайной выборке из 5000 человек в возрасте от 30 до 60 лет оказалось 2500 мужчин и 2500 женщин. Можно ли считать, основываясь на этих данных, что смертность среди мужчин и женщин одинакова? Уровень значимости считать равным 0.05. (! H_a односторонняя)

Домашнее задание

1. Из 500 опрошенных покупателей сети магазинов «Фиалка» 100 человек ответили, что они полностью довольны обслуживанием в данном магазине. Построить доверительный интервал уровня надёжности 0.95 для доли покупателей, довольных обслуживанием в сети «Фиалка».

2. Из 400 опрошенных покупателей сети «Жасмин» 70 человек ответили, что они довольны обслуживанием в данной сети. Постройте асимптотический доверительный интервал уровня надёжности 0.98 для разности долей покупателей, довольных обслуживание в сетевых магазинах «Фиалка» (см. предыдущую задачу) и «Жасмин».

3. Служба мониторинга опросила 1000 избирателей, среди которых оказалось 68 избирателей, собирающихся голосовать за партию А. Постройте 99% доверительный интервал для доли избирателей, собирающихся голосовать за партию А.

4. Для того, чтобы партия прошла в парламент, ей требуется набрать не менее 7% голосов избирателей. Служба мониторинга опросила 1000 избирателей, среди которых оказалось 68 избирателей, собирающихся голосовать за партию А. Можно ли считать (на уровне значимости 0.05), что партия А пройдёт в парламент? (! H_a односторонняя)

5. На экзамене по английскому языку студент должен ответить на 100 вопросов. На каждый вопрос предлагается 4 варианта ответа, один из которых правильный. Студент NN указал верные ответы на 30 вопросов. Можно ли считать (на уровне значимости 0.05), что студент NN не знает английский язык (т.е., выбор ответа на вопросы является случайным)? (! H_a односторонняя)

6. Известно, что женщины- водители составляют 30% от общего числа водителей. Из 635 зафиксированных ДТП 132 ДТП произошли по вине женщин-водителей. Можно ли считать (на уровне значимости 0.01), что женщины водят машину аккуратнее мужчин? (! H_a односторонняя)

$\hat{\theta}$ – сост. оценка θ

$\sqrt{n}(\hat{\theta} - \theta) \sim N(0, \sigma^2(\theta))$

по теор о наслед сход.

$\hat{\theta} \xrightarrow{P} \theta \Rightarrow \sigma(\hat{\theta}) \to \sigma(\theta)$

АДИ: $\left(\hat{\theta} \pm z_{1-\frac{\alpha}{2}} \cdot \frac{\sigma(\hat{\theta})}{\sqrt{n}}\right)$

1) $X_1 \dots X_n \sim \Pi(\theta) \Rightarrow$ $EX = \theta$, $DX = \theta \Rightarrow \sigma = \sqrt{\theta}$, $\sigma_X^2 = DX$

$\hat{\theta} = \overline{X}$ состоят. и эффективная оценка θ

$\underbrace{\sigma(\hat{\theta})}_{f(\hat{\theta})} = \sqrt{\hat{\theta}} = \sqrt{\overline{X}}$

ДЗ: N2, N3 и N4 (нужно использовать N3)