

Задача 1.

$$n = 61$$

$$s^2 = 14$$

$$\alpha = 1 - \beta = 1 - 0.9 = 0.1$$

$$\chi^2(0.95, df = 60) :$$

$$\frac{x-60}{\sqrt{120}} = Z_{0.95}$$

$$\frac{x-60}{\sqrt{120}} = 1.64$$

$$x = 1.64 \cdot \sqrt{120} + 60 = 77.97$$

$$\chi^2(0.05, df = 60) :$$

$$\frac{x-60}{\sqrt{120}} = Z_{1-0.05}$$

$$\frac{x-60}{\sqrt{120}} = -1.64$$

$$x = -1.64 \cdot \sqrt{120} + 60 = 42.03$$

$$\frac{s^2(n-1)}{\chi^2(1 - \frac{\alpha}{2}, df = n-1)} \leq \text{Var}(X) \leq \frac{s^2(n-1)}{\chi^2(\frac{\alpha}{2}, df = n-1)}$$

$$10.77 \leq \text{Var}(X) \leq 19.99$$

Вывод: С 90%-ной уверенностью мы можем утверждать, что генерального параметра дисперсии X лежит в интервале от 10.77 до 19.99. Если мы будем проводить аналогичное исследование на выборках одного и того же размера много раз, независимо друг от друга, 90% доверительных интервалов будут включать истинное значение генерального параметра дисперсии X .

Задача 2.

$$n = 10$$

$$H_0 : p = 0.5$$

$H_1 : p > 0.5$ (если выбрать решки, то мы работаем с правосторонней H_1 , так как 8 - это больше чем половина бросков)

$$P(x \geq 8) = C_{10}^8 \cdot 0,5^8 \cdot 0,5^2 + C_{10}^9 \cdot 0,5^9 \cdot 0,5^1 + C_{10}^{10} \cdot 0,5^{10} \cdot 0,5^0 \approx 0.05468$$

Вывод: Так как в условии задачи не был задан уровень значимости, мы не можем однозначно сделать вывод о принятии или об отвержении нулевой гипотезы, ведь значение p -value не является однозначно маленьким, чтобы, например отвергать нулевую гипотезу без уровня значимости. Если мы предполагаем, что уровень значимости 5%, то H_0 не отвергается, так как $p\text{-value} > \alpha$. Если уровень значимости больше 5%, то H_0 отвергается в пользу альтернативной.

Задача 3.

$$n = 120$$

$$p_0 = 0.5$$

$$\hat{p} = \frac{75}{120} = 0.625$$

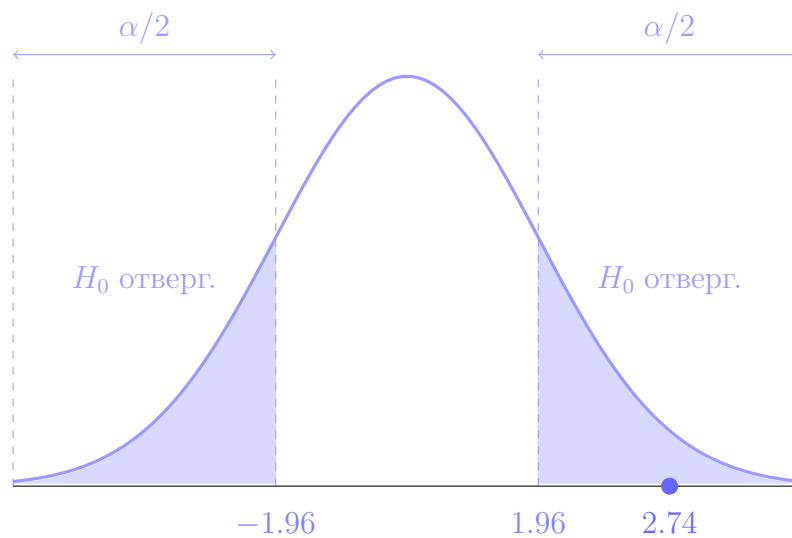
$$1. \alpha = 0.05$$

$$H_0 : p = 0.5$$

$$H_1 : p \neq 0.5$$

$$Z_{\text{крит.}} = 1.96$$

$$Z_{\text{набл.}} = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0 \cdot q_0}{n}}} = \frac{0.625 - 0.5}{\sqrt{\frac{0.5 \cdot 0.5}{120}}} = \frac{0.625 - 0.5}{0.0456} = 2.74$$



Вывод : $Z_{\text{набл.}}$ попадает в критическую область, поэтому мы отвергаем $H_0 : p = 0.5$ в пользу альтернативной гипотезы $H_1 : p \neq 0.5$

$$2. H_0 : p = 0.5$$

$$H_1 : p > 0.5$$

$$\text{p-value} = 1 - \Phi(2.74) = 1 - 0.9969 = 0.0031$$

Вывод : Так как значение p-value мало, у нас есть основание отвергнуть нулевую гипотезу в пользу альтернативной

Задача 4.

$$L(X_1 \dots X_n | \theta) = \frac{1}{\theta} e^{\frac{-2}{\theta}} \cdot \frac{1}{\theta} e^{\frac{-5}{\theta}} \cdot \frac{1}{\theta} e^{\frac{-6}{\theta}} \cdot \frac{1}{\theta} e^{\frac{-9}{\theta}} \cdot \frac{1}{\theta} e^{\frac{-8}{\theta}} \cdot (e^{\frac{-15}{\theta}})^5 = \frac{1}{\theta^5} e^{\frac{-105}{\theta}}$$

$$\ln(L) = -5 \cdot \ln(\theta) - \frac{105}{\theta}$$

$$\frac{\partial \ln(L(\theta))}{\partial \theta} = \frac{-5}{\theta} + \frac{105}{\theta^2}$$

$$\frac{-5}{\theta} + \frac{105}{\theta^2} = 0$$

$$105 - 5\theta = 0$$

$$5\theta = 105$$

$$\theta = 21$$

Задача 5.

$$f(X) = \frac{1}{\theta}$$

$$L(X_1 \dots X_n | \theta) = \frac{1}{\theta^n}$$

$$\ln(L(\theta)) = -n \cdot \ln(\theta)$$

$$\frac{\partial \ln(L(\theta))}{\partial \theta} = \frac{-n}{\theta}$$

$$\frac{-n}{\theta} = 0 \quad \begin{cases} n = 0 \\ \theta \neq 0 \end{cases}$$

Вывод: θ должно быть бесконечно большим значением, чтобы максимизировать функцию правдоподобия, однако в рамках задачи θ ограничена максимальным значением по выборке (X_{max}), поэтому θ должна быть равна X_{max}

$$\hat{\theta}_{MLE} = \max(X_1, X_2 \dots X_n)$$

Задача 6.

1. Используем правостороннюю альтернативную гипотезу, так как наиболее высокие значения статистики маловероятны при верной H_0 , и мы отвергнем H_0 в пользу альтернативной в данной области значения статистики.

Для определения значений, которые входят в критическую область, считаем сумму вероятностей ровно до того момента, пока кумулятивное значение не превысит $\alpha = 0.1$.

$$0.01 + 0.04 = 0.05 \Rightarrow \text{в критическую область статистики входят значения 7 и 8.}$$

$$\text{Type I Error: } 0.01 + 0.04 = 0.05$$

2. **Type II Error:** - сумма вероятностей значений статистики S меньше 7:

$$0.01 + 0.02 + 0.09 + 0.21 + 0.31 + 0.14 = 0.78$$

$$\text{Power: } 1 - 0.78 = 0.22$$

3.1 Мощность критерия можно увеличить путем увеличения уровня значимости

3.2 Мощность критерия можно увеличить путем увеличения effect size, то есть необходимо поставить альтернативу дальше от H_0

Задача 7.

$$n = 20$$

$$H_0 : p = 0.5$$

$$H_1 : p = 0.1$$

RR : значение статистики не превышает 2 (значение статистики критерия ≤ 2)

$$1. \text{ Type I Error: } C_{20}^0 \cdot 0,5^0 \cdot 0,5^{20} + C_{20}^1 \cdot 0,5^1 \cdot 0,5^{19} + C_{20}^2 \cdot 0,5^2 \cdot 0,5^{18} \approx 0.0002$$

Вывод: p-value очень мало, поэтому мы отвергаем нулевую гипотезу о том, что вероятность того, что человек знаком с понятием криптовалюты равна 0.5 в пользу альтернативной гипотезы о равенстве такой вероятности 0.1

$$2. \text{ Type II Error: } 1 - (P(k=0) + P(k=1) + P(k=2)) \Bigg\} \Rightarrow$$

$$\text{Power} = 1 - \text{Type II Error}$$

$$\Rightarrow \text{Power: } C_{20}^0 \cdot 0,1^0 \cdot 0,9^{20} + C_{20}^1 \cdot 0,1^1 \cdot 0,9^{19} + C_{20}^2 \cdot 0,1^2 \cdot 0,9^{18} \approx 0.68$$