Теория вероятностей и математическая статистика.

Домашнее задание Сурова София

Проверка гипотез

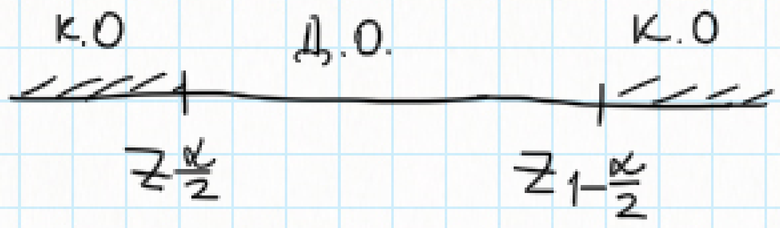
(1) Импортёр упаковывает чай в пакеты с номинальным весом 125 граммов. Известно, что упаковывающая машина работает с известным среднеквадратическим отклонением равным 10 г. Случайным образом для проверки было выбрано 50 пакетов чая. Выборочное среднее этих пакетов составило 125,8 г. Проверьте на уровне значимости 0.05 гипотезу о том, что средний вес пакета чая соответствует номиналу.

1) $H_0: m = 125$

$H_A: m \neq 125$

2) $T = \frac{(\overline{X} - 125)\sqrt{n}}{\sigma}$

3) $T|_{H_0} \sim N(0,1)$

4) 

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

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

5) $T_{набл.} = \frac{(125{,}8 - 125)\sqrt{50}}{10} \approx 0{,}5657$

6) $T_{набл.} = 0{,}5657 \in [-1{,}96; 1{,}96]$ ← доверительная область

=> нет оснований отвергнуть нулевую гипотезу в пользу альтернативной

Ответ: можно считать, что средний вес пакета чая соответствует номиналу

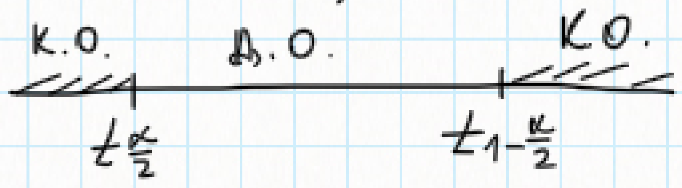
(2) Станок-автомат штампует валики. По выборке объёма n=17 вычислено выборочное среднее $\bar{X}$= 20,5 мм и выборочная дисперсия s^2= 16 (мм)^2 диаметра валиков. Проверьте на уровне значимости 0.05 гипотезу о том, что среднее значение диаметра валика равно 20.

1) $H_0: m = 20$

$H_A: m \neq 20$

2) $T = \frac{(\overline{X} - 20)\sqrt{n-1}}{\sqrt{S^2}}$

3) $T|_{H_0} \sim t(n-1)$

4) 

$t_{16,\,0.975} = 2{,}120$

$t_{16,\,0.025} = -2{,}120$

5) $T_{набл.} = \frac{(20{,}5 - 20)\sqrt{16}}{\sqrt{16}} = 0{,}5$

6) $T_{набл.} = 0{,}5 \in [-2{,}120; 2{,}120]$ ← д.о.

=> нет оснований отвергнуть нулевую гипотезу в пользу альтернативной

Ответ: можно считать, что среднее значение диаметра валика равно 20

③ Среднее время сборки изделия составляет 90 минут. 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. Можно ли считать, что время сборки в среднем сократилось. Уровень значимости считать равным 0.05.

1) $H_0: m = 90$

$H_A: m < 90$

2) $T = \frac{(\overline{x} - 90)\sqrt{n-1}}{\sqrt{S^2}}$

3) $T|_{H_0} \sim t(n-1)$

4) к.о. | д.о. ; t_α $\qquad t_{9,0.05} = -t_{9,0.95} = -1,833$

5) $\overline{x} = \frac{1}{n}\sum^{n} x_i = 860/10 = 86$

$S^2 = \frac{1}{n}\sum^{n}(x_i - \overline{x})^2 = 1416/10 = 141.6 \qquad \sqrt{S^2} \approx 11.9$

$T_{набл.} = \frac{(86-90)\sqrt{9}}{11.9} \approx -1,0084$

6) $T_{набл.} = -1,0084 \in [-1,833; +\infty)$ ← д.о.

=> нет оснований отвергнуть нулевую гипотезу в пользу альтернативной

Ответ: нет оснований считать, что время сборки сократилось

④ Используя данные задачи 16.1 на стр. 163–164, проверьте гипотезу о равенстве средних значений январских температур в г. Саратове и Алатырь. Уровень значимости считать равным 0.05. Будем считать, что дисперсия температуры неизвестна, но одинакова для г. Саратова и г. Алатырь

Задача 16.1. 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 табл. 16.2 приведены значения средней температуры января в г. Саратове и г. Алатыре.

Таблица 16.2

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	

1) $H_0: m_1 - m_2 = 0$

$H_A: m_1 - m_2 \neq 0$

2) $T = \frac{\overline{x} - \overline{y}}{\sqrt{S^2_{xy}}\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$

3) $T|_{H_0} \sim t(n_1 + n_2 - 2)$

4) к.о. | д.о. | к.о. ; $t_{\frac{\alpha}{2}}$, $t_{1-\frac{\alpha}{2}}$

$t_{24,0.975} = 2,064$

$t_{24,0.025} = -2,064$

5) $\overline{x} = \frac{1}{n_1}\sum^{n_1} x_i \approx -11,69$

$S^2_x = \frac{1}{n_1}\sum^{n_1}(x_i - \overline{x})^2 \approx 22,1449 \qquad \sqrt{\frac{1}{13} + \frac{1}{13}} \approx 0,3922$

$\overline{y} = \frac{1}{n_2}\sum^{n_2} y_i \approx -13,7538$

$S^2_y = \frac{1}{n_2}\sum^{n_2}(y_i - \overline{y})^2 \approx 20,0902$

$S^2_{xy} = \frac{n_1 S^2_x + n_2 S^2_y}{n_1 + n_2 - 2} \approx 22,8773 \qquad \sqrt{S^2_{xy}} \approx 4,7830$

$T_{набл.} = \frac{-11,69 + 13,7538}{4,7830 \cdot 0,3922} \approx 1,1002$

6) $T_{набл.} = 1,1002 \in [-2,064; 2,064]$ ← д.о.

=> нет оснований отвергнуть нулевую гипотезу в пользу альтернативной

Ответ: можно считать, что средние значения температур в январе в Саратове и Алтыре равны

⑤ Изучается зависимость между показателями X (оценка студента за контрольную работу по теории вероятностей) и Y (оценка студента за экзамен по программированию). Используя данные о показателях X и Y для группы 182, в которой обучается 30 человек, студент Кристиан Бенуа вычислил выборочный коэффициент корреляции и получил значение 0.45. Можно ли, опираясь на эти данные, считать, что показатели X и Y являются положительно коррелированными? Уровень значимости принять равным 0.05

1) $H_0: \rho_{XY} = 0$

$H_A: \rho_{XY} > 0$

2) $T = \dfrac{\sqrt{n-2}\,\hat{\rho}_{XY}}{\sqrt{1-\hat{\rho}_{XY}^2}}$

3) $T|_{H_0} \sim t(n-2)$

4) Д.О. К.О. $t_{1-\alpha}$ $\quad t_{28,\,0.95} = 1{,}701$

5) $T_{\text{набл.}} = \dfrac{\sqrt{28}\cdot 0{,}45}{\sqrt{1-0{,}45^2}} \approx 2{,}6664$

6) $T_{\text{набл.}} = 2{,}6664 \notin (-\infty;\, 1{,}701]$ ← Д.О.

$\Rightarrow$ есть основания отвергнуть нулевую гипотезу в пользу альтернативной

Ответ: есть основания считать, что оценки положительно коррелированы

⑥ Влияние прививки на холерную инфекцию (данные Macmillan Publishing Company). Среди 1630 человек, сделавших прививку против холеры, оказалось 5 заболевших. В контрольной группе из 1033 человек, не сделавших прививку, заболевших оказалось 11. Имеется ли зависимость между наличием прививки и заболеваемостью?

1) H_0: величины независимы

H_A: величины зависимы

2) $T = n\left(\sum\limits_{i}^{s}\sum\limits_{j}^{r} \dfrac{n_{ij}^2}{n_{i\cdot}\, n_{\cdot j}} - 1\right)$

	B_1	B_2	
A_1	n_{11}	n_{12}	$n_{1\cdot}$
A_2	n_{21}	n_{22}	$n_{2\cdot}$
	$n_{\cdot 1}$	$n_{\cdot 2}$	n

3) $T|_{H_0} \sim \chi^2((2-1)(2-1))$

4) Д.О. К.О. $\chi^2_{1-\alpha}$ $\quad \chi^2_{1,\,0.95} = 3{,}84$

5)

	заболел	не заболел	
привился	5	1625	1630
не привился	11	1022	1033
	16	2647	2663

$T_{\text{набл.}} = 2663\left(\dfrac{5^2}{1630\cdot 16} + \dfrac{1625^2}{1630\cdot 2647} + \dfrac{11^2}{1033\cdot 16} + \dfrac{1022^2}{1033\cdot 2647} - 1\right) \approx 6{,}0849$

6) $T_{\text{набл.}} = 6{,}0849 \notin (-\infty;\, 3{,}84]$ ← Д.О.

$\Rightarrow$ есть основания опровергнуть нулевую гипотезу в пользу альтернативной

Ответ: можно считать, что есть зависимость между наличием прививки и заболеваемостью