Проверка параметрических гипотез
(равенство вероятностей «успеха» двух биномиальных совокупностей; проверка гипотез о параметрах гауссовских СВ)

1. В Москве было опрошено 600 человек, из которых 66 человек сказали, что они недовольны своей работой. В Московской области опросили 500 человек, и 60 из них объявили, что они недовольны своей работой. Можно ли считать, что в Московской области доля недовольных своей работой выше, чем в Москве? Принять уровень значимости равным 0.05.

2. 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. Проверьте на уровне значимости 0.05 гипотезу о том, что средний вес пакета чая соответствует номиналу.

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

4. 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. Предполагается, что наблюдения имеют гауссовское распределение. Можно ли считать, что время сборки в среднем сократилось? Уровень значимости считать равным 0.05.

5. 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	

Проверьте гипотезу о равенстве средних значений январских температур в г. Саратове и г. Алатырь, предполагая, что дисперсия температуры неизвестна, но одинакова для обоих городов. Будем считать, что наблюдения имеют гауссовское распределение. Уровень значимости считать равным 0.05.

Домашнее задание

1. Проведено исследование по выявлению факторов риска заболеваемостью туберкулезом. Один из факторов – низкий доход в семье. Было обследовано 300 семей с низким доходом, в 13 из них были выявлены больные туберкулезом. В 100 семьях с высокими доходами выявлено двое больных. Можно ли сказать, что низкие доходы являются фактором риска заболеваемости туберкулезом? Принять уровень значимости равным 0.05.

2. Два пресса штампуют детали одного наименования. Для первого пресса проверили 1000 деталей, среди них 25 оказались бракованных. Для второго – 800 деталей, среди них 18 бракованных. Согласуются ли эти результаты с предположением о равенстве доли брака для обоих прессов? Принять уровень значимости равным 0.01.

3. Измерен вес (в граммах) 6 шоколадных батончиков, которые были случайным образом выбраны из большой партии: 49,1; 50,1; 49,7; 50,5; 49,8; 50,2. Можно ли считать, что средний вес батончиков соответствует номинальному весу 50гр. Предполагается, что наблюдаемая выборка имеет гауссовское распределение. Уровень значимости считать равным 0.01.

4. Точность работы станка-автомата определяется среднеквадратическим отклонением σ размеров изделий от известного номинального размера μ. По стандарту σ не должно превышать 0.01 мм. По выборке из 25 изделий вычислена статистика $s^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \mu)^2$=0,02$мм^2$, где $X_1, \dots, X_n$ - размеры изделий. На уровне значимости 0.05 проверьте, обеспечивает ли станок необходимую точность. Предполагается, что наблюдаемая выборка имеет гауссовское распределение.

5. Задача №7 стр. 202 (учебник Кибзун А.И., Горяинова Е.Р., Наумов А.В.)

6. По выборке объёма 50 найдено выборочное среднее 20,1 мм диаметра валиков, изготовленных первым станком. По выборке объёма 60 найдено выборочное среднее 19,8 мм диаметра валиков, изготовленных вторым станком. Известно, что дисперсия диаметра валика, изготовленного первым станком, равна 1,8 мм^2. Для второго станка дисперсия равна 1.4 мм^2. Можно ли считать, опираясь на эти данные, что на втором станке изготавливаются валики, диаметр которых в среднем меньше, чем у валиков, изготовленных на первом станке? Уровень значимости принять равным 0.05. Предполагается, что наблюдения имеют гауссовское распределение.

1. В Москве было опрошено 600 человек, из которых 66 человек сказали, что они недовольны своей работой. В Московской области опросили 500 человек, и 60 из них объявили, что они недовольны своей работой. Можно ли считать, что в Московской области доля недовольных своей работой выше, чем в Москве? Принять уровень значимости равным 0.05.

1. Постановка гипотез (Но и На)

$H_0: p_1' = p_2' \Rightarrow p_1' - p_2' = 0$

$H_a: p' > p_2' \Rightarrow p_1 - p_2' > 0$

2. Выбираем α (что задана, если нет, то $\alpha = 0{,}05$)

3-4. Выбор статистики и ее распределения

$$T = \frac{\hat{p}_1 - \hat{p}_2 \boxed{- (p_1' - p_2)} = 0}{D(p_1 - p_2)} \sim N\ 0,1)$$

5. Указываем критическую область

(Д.О.)Н00 НоС

$t_{кр} = z_{1-\alpha} = 1{,}65$

Н00Но – Нет оснований отвергнуть Но

---

1. В Москве было опрошено 600 человек, из которых 66 человек сказали, что они недовольны своей работой. В Московской области опросили 500 человек, и 60 из них объявили, что они недовольны своей работой. Можно ли считать, что в Московской области доля недовольных своей работой выше, чем в Москве? Принять уровень значимости равным 0.05.

$n_1 = 600,\ k_1 = 66,\ \hat{p}_1 = \frac{66}{600} = 0{,}11$

$n_2 = 500,\ k_2 = 60,\ \hat{p}_2 = \frac{60}{500} = 0{,}12$

$H_0: p_1 = p_2$

$H_1: p_1 < p_2$

$\alpha = 0{,}05$

$$T = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{D(\hat{p}_1 - \hat{p}_2)}} \sim N(0,1)$$

$$\hat{p} = \frac{k_1 + k_2}{n_1 + n_2} = \frac{66 + 60}{600 + 500} = \frac{126}{1100} \approx 0{,}1145$$

$$T = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\hat{p}(1-\hat{p})\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} = \frac{0{,}11 - 0{,}12}{0{,}1145 \cdot 0{,}8855\left(\frac{1}{600} + \frac{1}{500}\right)} \approx -0{,}57$$

$z_{0,05} = -1{,}65$

$-0{,}57 > -1{,}65 \Rightarrow H_0$ не отвергается

Ответ: нет.

2. 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. Проверьте на уровне значимости 0.05 гипотезу о том, что средний вес пакета чая соответствует номиналу.

$\mu_0 = 125,\ \sigma = 10,\ n = 50,\ \overline{x} = 125.8$

$H_0: \mu = 125$

$H_a: \mu \neq 125$

$\alpha = 0{,}05$

$$T = \frac{(\overline{x} - \mu_0)\sqrt{n}}{\sigma} \sim N(0,1)$$

$$T=\frac{(125{,}8-125)\sqrt{50}}{10}\approx 0{,}57$$

$z_{0,975}\approx 1{,}96>|0{,}57|\Rightarrow H_0$ не отвергается

Ответ: это правда

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

$n=17$, $\bar{X}=20{,}5$, $S^2=16$, $\mu_0=20$

$H_0: \mu=20$

$H_a: \mu\neq 20$

$\alpha=0{,}05$

$$T=\frac{(\bar{X}-\mu_0)\sqrt{n}}{S}=\frac{(20.5-20)\sqrt{17}}{4}\approx 0{,}52$$

$\frac{\alpha}{2}=0{,}025$

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

$z_{0,975}\approx 1{,}96$

$|0{,}52|<1{,}96\Rightarrow H_0$ не отвергается

Ответ: это верно.

4. 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. Предполагается, что наблюдения имеют гауссовское распределение. Можно ли считать, что время сборки в среднем сократилось? Уровень значимости считать равным 0.05.

$n=10$

$x_1, x_2, \dots, x_n=\{79, 74, 112, 95, 83, 96, 77, 84, 70, 90\}$

$\mu_0=90$

$H_0: \mu=90$

$H_a: \mu<90$

$\alpha=0{,}05$

$$\bar{X}=\frac{79+74+\dots+90}{10}=86$$

$$S=\sqrt{\frac{(79-86)^2+\dots+(90-86)^2}{9}}\approx 12{,}02$$

$$T=\frac{(\bar{X}-\mu_0)\sqrt{n}}{S}=\frac{(86-90)\sqrt{10}}{12{,}02}\approx -1{,}05$$

$z_{0,5}=-1{,}65<-1{,}05\Rightarrow H_0$ не отвергается

Ответ: да, можно считать.

5. 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	

Проверьте гипотезу о равенстве средних значений январских температур в г. Саратове и г. Алатырь, предполагая, что дисперсия температуры неизвестна, но одинакова для обоих городов. Будем считать, что наблюдения имеют гауссовское распределение. Уровень значимости считать равным 0.05.

Саратов $n=13$ $\bar{X}=-11{,}88$ $S_1^2\approx 23{,}99$

Алатырь $n=13$ $\bar{X}=-13{,}75$ $S_2^2\approx 21{,}76$

$$S_p^2 = \frac{(n_1-1)S_1^2+(n_2-1)S_2^2}{n_1+n_2-2} = \frac{12\cdot 23{,}99+12\cdot 21{,}76}{24} \approx 22{,}88$$

$$SE = \sqrt{S_p^2\left(\frac{1}{n_1}+\frac{1}{n_2}\right)} = \sqrt{22{,}88\cdot\frac{2}{13}} \approx 1{,}877$$

$$t = \frac{\bar{X}_1-\bar{X}_2}{SE} = \frac{-11{,}88+13{,}75}{1{,}877} \approx 0{,}996$$

$\alpha = 0{,}05$

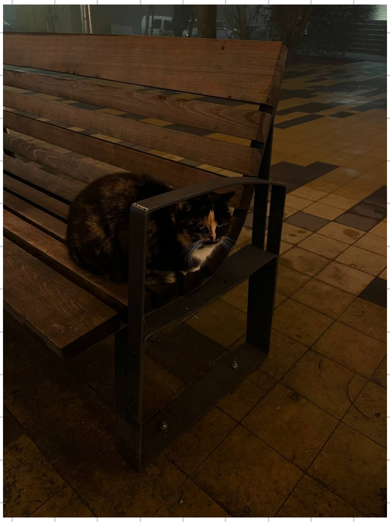
$H_0: \bar{X}_1 = \bar{X}_2$

$H_a: \bar{X}_1 \neq \bar{X}_2$

$1-\frac{\alpha}{2} = 1-0{,}25 = 0{,}975$

$z_{0,975} \approx 1{,}96 > 0{,}996 \Rightarrow H_0$ не отвергается

Ответ: это правда



Я в Калининграде, сделайте, пожалуйста, конспект с 7 семинара🙏🏻