HW5

Сметанин Александр

16 апреля 2020 г.

1. Пусть $X_1, \dots, X_n$ — выборка, о плотности распределения которой высказаны две гипотезы: гипотеза H_0 о том, что X_i имеют распределение с плотностью $f_1(y) = e^{-(y-6)}$ при $y > 6$, и альтернатива H_1, состоящая в том, что X_i имеют плотность $f_2(y) = 2e^{-2(y-3)}$ при $y > 3$. Найти пределы при $n \to \infty$ вероятностей ошибок первого и второго рода следующего критерия: гипотеза H_0 принимается тогда и только тогда, когда

a. $\bar{X} > 3.5 + \frac{1}{\sqrt{n}}$

$$\alpha_I = P(\bar{X} \leq 3.5 + \frac{1}{\sqrt{n}}) \overset{n \geq 1}{\leq} P(\bar{X} \leq 4.5) \overset{X_i \geq 6}{=\!=} 0$$

Заметим, что $EX_i = 3.5$ и $DX_i = 0.25$, если верна H_1.

$$\alpha_{II} = P(\bar{X} > 3.5 + \frac{1}{\sqrt{n}}) = P(\sqrt{n}\frac{\bar{X} - EX_i}{\sqrt{DX_i}} > \frac{1}{\sqrt{DX_i}}) \xrightarrow[n\to\infty]{CLT} 1 - \Phi\left(\frac{1}{0.5}\right) = 0.023$$

b. $\bar{X} > 3.5 + \frac{1}{n}$
Аналогично,

$$\alpha_I = P(\bar{X} \leq 3.5 + \frac{1}{n}) \leq P(\bar{X} \leq 4.5) = 0$$

$$\alpha_{II} = P(\bar{X} > 3.5 + \frac{1}{n}) = P(\sqrt{n}\frac{\bar{X} - EX_i}{\sqrt{DX_i}} > \frac{1}{\sqrt{n \cdot DX_i}}) \xrightarrow[n\to\infty]{CLT} 1 - \Phi(0) = 0.5$$

c. $\bar{X} > 3.5$

$$\alpha_I = P(\bar{X} \leq 3.5) \overset{X_i \geq 6}{=\!=} 0$$

$$\alpha_{II} = P(\bar{X} > 3.5) = P(\sqrt{n}\frac{\bar{X} - EX_i}{\sqrt{DX_i}} > 0) \xrightarrow[n\to\infty]{CLT} 1 - \Phi(0) = 0.5$$

2. The manager of the Danvers-Hilton Resort Hotel stated that the mean guest bill for a weekend is \$600 or less. A member of the hotel's accounting staff noticed that the total charges for guest bills have been increasing in recent months. The accountant will use a sample of weekend guest bills to test the manager's claim.

a. Which form of the hypotheses about the mean μ should be used to test the manager's claim? Explain.

I. $H_0 : \mu = 600; H_1 : \mu < 600$

II. $H_0 : \mu = 600; H_1 : \mu > 600$

III. $H_0 : \mu = 600; H_1 : \mu \neq 600$

Стоит выбрать вторую (II) форму гипотезы, так как нужно проверить, что средний чек не увеличился.

b. What conclusion is appropriate when H_0 cannot be rejected?
Можно сделать вывод, что нет оснований не доверять позиции менеджера о том, что чек меншье \$600.

c. What conclusion is appropriate when H_0 can be rejected?
Можно сделать вывод, что скорее всего средний чек увеличился.

3. Wall Street securities firms paid out record year-end bonuses of \$125,500 per employee for 2005 (Fortune, February 6, 2006). Suppose we would like to take a sample of employees at the Jones & Ryan securities firm to see whether the mean year-end bonus is different from the reported mean of \$125,500 for the population.

a. State the null and alternative hypotheses you would use to test whether the year-end bonuses paid by Jones & Ryan were different from the population mean.
$H_0 : \mu = 125500; H_1 : \mu \neq 125500$

b. Suppose a sample of 40 Jones & Ryan employees showed a sample mean year-end bonus of \$118,000. Assume a population standard deviation of \$30,000 and with $\alpha = 0.05$ as the level of significance, what is your conclusion?
Воспользуемся Z-критерием:

$$\bar{Z} = \sqrt{n}\frac{\bar{X} - \mu}{\sigma} = \sqrt{40}\frac{118000 - 125500}{30000} = -1.58$$

Для уровня значимости $\alpha = 0.05$. P(-1.96 < Z < 1.96) = 0.95. Так как $\bar{Z} \in (-1.96, 1.96)$, у нас нет опровержения того, что средний размер бонуса составил \$125,500.

4. Many investors and financial analysts believe the Dow Jones Industrial Average (DJIA) provides a good barometer of the overall stock market. On January 31, 2006, 9 of the 30 stocks making up the DJIA increased in price (The Wall Street Journal, February 1, 2006). On the basis of this fact, a financial analyst claims we can assume that 30% of the stocks traded on the New York Stock Exchange (NYSE) went up the same day.

a. Formulate null and alternative hypotheses to test the analyst's claim.
Исходим из предположения, что в обоих гипотезах количество ценных бумаг, показывающих рост определяется биномиальным распределением с вероятностью p.
$H_0 : p = 0.3; H_1 : p \neq 0.3$

b. A sample of 50 stocks traded on the NYSE that day showed that 24 went up. What is your point estimate of the population proportion of stocks that went up?

$$\hat{p} = \frac{24}{50} = 0.48$$

c. Conduct your hypothesis test using $\alpha = 0.01$ as the level of significance. What is your conclusion?

$$\bar{Z} = \sqrt{n}\frac{\hat{p} - p}{\sqrt{p(1-p)}} = 2.78$$

Для уровня значимости $\alpha = 0.01$. P(-2.58 < Z < 2.58) = 0.99. Так как $\bar{Z} \notin (-2.58, 2.58)$, у нас есть основания для того, чтобы отвергнуть гипотезу H_0.

5. $H_0 : \mu = 120$ and $H_1 : \mu \neq 120$ are used to test whether a bath soap production process is meeting the standard output of 120 bars per batch. Use a 0.05 level of significance for the test and a planning value of 5 for the standard deviation.

a. If the mean output drops to 117 bars per batch, the firm wants to have a 98% chance of concluding that the standard production output is not being met. How large a sample should be selected?
Так как мы хотим определять снижения производства до 117, нам нужен критерий вида

$$\delta(X_1, \dots X_n) = \begin{cases} H_0, & \bar{X} > c, \\ H_1, & \bar{X} \leq c. \end{cases}$$

Выберем такую c, чтобы $a_I = 0.05$, $a_{II} = 0.02$:

$$\begin{cases} c &= \mu_0 - t_{\alpha_I} \cdot \frac{\sigma}{\sqrt{n}} \\ c &= \mu_1 + t_{\alpha_{II}} \cdot \frac{\sigma}{\sqrt{n}} \end{cases}$$

$$\frac{\sigma}{\sqrt{n}} = \frac{\mu_0 - \mu_1}{t_{\alpha_I} + t_{\alpha_{II}}}$$

$$n = \sigma^2 \frac{(t_{\alpha_I} + t_{\alpha_{II}})^2}{(\mu_0 - \mu_1)^2}$$

$$n = \left(5 \cdot \frac{1.96 + 2.05}{3}\right)^2 = 44.7$$

Стоит брать выборку размером 45.

b. With your sample size from part (a), what is the probability of concluding that the process is operating satisfactorily for each of the following actual mean outputs: 117, 118, 119, 121, 122, and 123 bars per batch? That is, what is the probability of a Type II error in each case?
Найдем - границы для нашего критерия $\delta(X_1, \ldots X_n)$:

$$c_- = \mu_0 - t_{\alpha_I} \cdot \frac{\sigma}{\sqrt{n}} = 120 - 1.96 \cdot \frac{5}{\sqrt{45}} = 118.54$$

$$c_+ = 121.46$$

Посчитаем вероятность ошибки II рода при настоящем среднем:

b1. $\mu_1 = 117$

$$\alpha_{II} = P(\bar{X} > c_-) = P(\sqrt{n}\frac{\bar{X} - \mu_1}{\sigma} > \sqrt{n}\frac{c_- - \mu_1}{\sigma}) = 1 - \Phi(\sqrt{n}\frac{c_- - \mu_1}{\sigma}) = 0.019$$

b2. $\mu_2 = 118$

$$\alpha_{II} = P(\bar{X} > c_-) = P(\sqrt{n}\frac{\bar{X} - \mu_2}{\sigma} > \sqrt{n}\frac{c_- - \mu_2}{\sigma}) = 1 - \Phi(\sqrt{n}\frac{c_- - \mu_2}{\sigma}) = 0.23$$

b3. $\mu_3 = 119$

$$\alpha_{II} = P(\bar{X} > c_-) = P(\sqrt{n}\frac{\bar{X} - \mu_3}{\sigma} > \sqrt{n}\frac{c_- - \mu_3}{\sigma}) = 1 - \Phi(\sqrt{n}\frac{c_- - \mu_3}{\sigma}) = 0.73$$

b4. $\mu_4 = 121$

$$\alpha_{II} = P(\bar{X} < c_+) = P(\sqrt{n}\frac{\bar{X} - \mu_4}{\sigma} < \sqrt{n}\frac{c_+ - \mu_4}{\sigma}) = \Phi(\sqrt{n}\frac{c_+ - \mu_4}{\sigma}) = 0.73$$

b5. $\mu_5 = 122$

$$\alpha_{II} = P(\bar{X} < c_+) = P(\sqrt{n}\frac{\bar{X} - \mu_5}{\sigma} < \sqrt{n}\frac{c_+ - \mu_5}{\sigma}) = \Phi(\sqrt{n}\frac{c_+ - \mu_5}{\sigma}) = 0.23$$

b6. $\mu_6 = 123$

$$\alpha_{II} = P(\bar{X} < c_+) = P(\sqrt{n}\frac{\bar{X} - \mu_6}{\sigma} < \sqrt{n}\frac{c_+ - \mu_6}{\sigma}) = \Phi(\sqrt{n}\frac{c_+ - \mu_6}{\sigma}) = 0.019$$