Assignment 6

Monday, June 13, 2022 4:48 PM

Problem #1: **(a)** The following data shows the number of hours that 8 hospital patients slept following the administration of a certain anesthetic.

6, 6, 9, 11, 8, 11, 5, 10,

Find a 98% confidence interval for the average hours slept following the administration of the anesthetic for the sampled population.

(b) Which of the following statements is true regarding part **(a)**?

(A) The population must be normal. **(B)** The population does not need to be normal.

(C) The population must follow a *t*-distribution. **(D)** The population mean must be inside the confidence interval.

(E) The population standard deviation σ must be known.

5 6 6 8 9 10 11 11 $\Rightarrow$ $n = 8$

a.) $\alpha = 0.02$

$$\bar{X} = \frac{5+6+6+8+9+10+11+11}{8} = \frac{33}{4}$$

$$S^2 = \frac{\Sigma(X_i - \bar{X})^2}{n-1}$$

$$= \frac{\left(-\frac{13}{4}\right)^2 + \left(-\frac{9}{4}\right)^2 + \left(-\frac{9}{4}\right)^2 + \left(-\frac{1}{4}\right)^2 + \left(\frac{3}{4}\right)^2 + \left(\frac{7}{4}\right)^2 + \left(\frac{11}{4}\right)^2 + \left(\frac{11}{4}\right)^2}{8-1}$$

$$= \frac{\frac{79}{2}}{7}$$

$$= \frac{79}{14}$$

$$S = \sqrt{\frac{79}{14}} = 2.37547$$

$$\text{Upper limit} = \bar{X} + t_{\frac{\alpha}{2}, n-1} \cdot \frac{S}{\sqrt{n}}$$

$$= \frac{33}{4} + t_{\frac{0.02}{2}, 8-1} \cdot \frac{\sqrt{\frac{79}{14}}}{\sqrt{8}}$$

$$= \frac{33}{4} + 2.99795 \cdot 0.8399$$

$$= 10.767978$$

lower limit $= \bar{X} - t_{\frac{\alpha}{2}, n-1} \cdot \frac{S}{\sqrt{n}}$

$$= \frac{33}{4} - 2.99795 \cdot 0.8399$$

$$= 5.732022$$

$\therefore$ the 98% confidence interval is $[5.732022, 10.767978]$

b.) **(A)** The population must be normal. **(B)** The population does not need to be normal.

(C) The population must follow a *t*-distribution. **(D)** The population mean must be inside the confidence interval.

(E) The population standard deviation σ must be known.

Problem #3: In a study of perception, 127 men are tested and 27 are found to have red/green color blindness.

(a) Find a 90% confidence interval for the true proportion of men from the sampled population that have this type of color blindness.

(b) Using the results from the above mentioned survey, how many men should be sampled to estimate the true proportion of men with this type of color blindness to within 3% with 99% confidence?

(c) If no previous estimate of the sample proportion is available, how large of a sample should be used in **(b)**?

Binomial Distribution

$n = 127$

$X = 27$

a.) $\alpha = 0.1$

$Z_{\frac{\alpha}{2}} = Z_{\frac{0.1}{2}} = Z_{0.05} = 1.645$

$\hat{p} = \frac{X}{n} = \frac{27}{127} = 0.212598$

upper limit $= \hat{p} + Z_{\frac{\alpha}{2}} \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$

$$= \frac{27}{127} + 1.645 \cdot \sqrt{\frac{\left(\frac{27}{127}\right)\left(1-\frac{27}{127}\right)}{127}}$$

$$= 0.272322$$

$$\text{lower limit} = \hat{p} - z_{\frac{\alpha}{2}} \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$

$$= \frac{27}{127} - 1.645 \cdot \sqrt{\frac{\left(\frac{27}{127}\right)\left(1-\frac{27}{127}\right)}{127}}$$

$$= 0.152875$$

$\therefore$ the confidence interval is $[0.153, 0.272]$

b.) $\pm 3\%$ of 99% confidence

$E = 0.03$

$\alpha = 0.01$

$z_{\frac{\alpha}{2}} = z_{\frac{0.01}{2}} = 2.576$

$$n = \left(\frac{z_{\frac{\alpha}{2}} \cdot \sqrt{\hat{p}(1-\hat{p})}}{E}\right)^2$$

$$= \left(\frac{2.576 \cdot \sqrt{\left(\frac{27}{127}\right)\left(\frac{100}{127}\right)}}{0.03}\right)^2$$

$$= 1234.2568$$

$\therefore$ 1234 men should be sampled

c.) no previous estimate, so use $\hat{p} = 0.5$

$$n = \left(\frac{2.576 \cdot \sqrt{(0.5)(0.5)}}{0.03}\right)^2$$

$$= 1843.27$$

$\therefore$ 1843 men should be sampled

Problem #5: Suppose that we want to test the hypothesis that mothers with low socioeconomic status (SES) deliver babies whose birthweights are different than "normal". To test this hypothesis, a list of birthweights from 84 consecutive, full-term, live-born deliveries from the maternity ward of a hospital in a low-SES area is obtained. The mean birghweight is found to be 116 oz with a sample standard deviation of 23 oz. Suppose that we know from nationwide surveys based on millions of deliveries that the mean birthweight in the United States is 120 oz.

At α = .03, can it be concluded that the average birthweight from this hospital is different from the national average?

(a) Find the value of the test statistic for the above hypothesis.

(b) Find the critical value.

(c) Find the *p*-value.

(d) What is the correct way to draw a conclusion regarding the above hypothesis test?

$\bar{x} = 116$; $\sigma = 23$; $\mu = 120$; $\alpha = 0.03$; $n = 84$

sample mean; standard deviation; population mean; level of significance

$H_0: \mu = 120$ vs. $H_1: \mu \neq 120$

a.)

$$\text{Test statistic} = \frac{\bar{x} - \mu}{\frac{\sigma}{\sqrt{n}}} = \frac{116 - 120}{\frac{23}{\sqrt{84}}} = -1.59394$$

b.) critical value

$$Z_{\frac{\alpha}{2}} = Z_{\frac{0.03}{2}} = Z_{0.015} = -2.17 \Rightarrow |Z_{0.015}| = 2.17$$

c.) p-value

$$\begin{aligned} \text{p-value} &= 2[1 - \Phi(|Z_0|)] \\ &= 2[P(Z < -1.59394)] \\ &= 2[0.05592] \\ &= 0.11184 \end{aligned}$$

d.)

(A) If the answer in (c) is greater than 0.03 then we conclude at the 3% significance level that the average birthweight from this hospital is different from the national average.

(B) If the answer in (a) is greater than the answer in (c) then we **cannot** conclude at the 3% significance level that the average birthweight from this hospital is different from the national average.

(C) If the answer in (c) is less than 0.03 then we conclude at the 3% significance level that the average birthweight from this hospital is different from the national average.

(D) If the answer in (c) is less than 0.03 then we **cannot** conclude at the 3% significance level that the average birthweight from this hospital is different from the national average.

d.)

(A) If the answer in (c) is greater than 0.03 then we conclude at the 3% significance level that the average birthweight from this hospital is different from the national average.

(B) If the answer in (a) is greater than the answer in (c) then we **cannot** conclude at the 3% significance level that the average birthweight from this hospital is different from the national average.

(C) If the answer in (c) is less than 0.03 then we conclude at the 3% significance level that the average birthweight from this hospital is different from the national average.

(D) If the answer in (c) is less than 0.03 then we **cannot** conclude at the 3% significance level that the average birthweight from this hospital is different from the national average.

(E) If the answer in (b) is greater than the answer in (c) then we conclude at the 3% significance level that the average birthweight from this hospital is different from the national average.

(F) If the answer in (a) is greater than the answer in (c) then we conclude at the 3% significance level that the average birthweight from this hospital is different from the national average.

(G) If the answer in (a) is greater than the answer in (b) then we **cannot** conclude at the 3% significance level that the average birthweight from this hospital is different from the national average.

(H) If the answer in (b) is greater than the answer in (c) then we **cannot** conclude at the 3% significance level that the average birthweight from this hospital is different from the national average.

if p-value $< \alpha$ $\Rightarrow$ reject Ho
p-value $> \alpha$ $\rightarrow$ don't reject Ho

Problem #6: **(a)** Erythromycin is a drug that has been proposed to possibly lower the risk of premature delivery. A related area of interest is its association with the incidence of side effects during pregnancy. Assume that 30% of all pregnant women complain of nausea between the 24th and 28th week of pregnancy. Furthermore, suppose that of 189 women who are taking erythromycin regularly during this period, 67 complain of nausea. Find the *p*-value for testing the hypothesis that incidence rate of nausea for the erythromycin group is greater than for a typical pregnant woman.

(b) At the 1% significance level, what is the conclusion of the above hypothesis test?

$\bar{X} = 67$; $n = 189$; $P_0 = 30\% = 0.3$

$H_0: p = p_0$ vs. $H_1: p > p_0$

a.) test statistic $= z_0 = \dfrac{\bar{X} - np_0}{\sqrt{np_0(1-p_0)}}$

$= \dfrac{67 - 189(0.3)}{\sqrt{189(0.3)(1-0.3)}}$

$= \dfrac{103}{63}$

$= 1.634921$

$p > p_0 \Rightarrow$ p-value $= 1 - P(z_0 > z_\alpha)$

$= 1 - P(z < 1.63)$

$= 1 - 0.94845$

$= 0.05155$

b.) $\alpha = 0.01$

p-value $= 0.05155 > \alpha \Rightarrow$ do not reject H_0

(A) We conclude that the incidence rate of nausea for the erythromycin group is greater than for a typical pregnant woman since the *p*-value is greater than or equal to .02

(B) We **cannot** conclude that the incidence rate of nausea for the erythromycin group is greater than for a typical pregnant woman since the *p*-value is greater or equal to 0.01

(C) We conclude that the incidence rate of nausea for the erythromycin group is greater than for a typical pregnant woman since the *p*-value is greater than or equal to 0.01

(D) We conclude that the incidence rate of nausea for the erythromycin group is greater than for a typical pregnant woman since the *p*-value is less than .02

(E) We **cannot** conclude that the incidence rate of nausea for the erythromycin group is greater than for a typical pregnant woman since the *p*-value is greater or equal to .02

(F) We conclude that the incidence rate of nausea for the erythromycin group is greater than for a typical pregnant woman since the *p*-value is less than 0.01

(G) We **cannot** conclude that the incidence rate of nausea for the erythromycin group is greater than for a typical pregnant woman since the *p*-value is less than 0.01

(H) We **cannot** conclude that the incidence rate of nausea for the erythromycin group is greater than for a typical pregnant woman since the *p*-value is less than .02

Problem #7: In a 1868 paper, German physician Carl Wunderlich reported based on over a million body temperature readings that the mean body temperature for healthy adults is 98.6° F. However, it is now commonly believed that the mean body temperature of a healthy adult is less than what was reported in that paper. To test this hypothesis a researcher measures the following body temperatures from a random sample of healthy adults.

98.2, 98.5, 98.5, 98.6, 97.5, 98.4, 98.0

(a) Find the value of the test statistic.

(b) Find the 5% critical value.

(c) Work through this example on R, and then use R to find the *p*-value for the relevant hypothesis test.

(d) Since the sample size is less than 30, the above analysis requires that the population follows a normal distribuiton. What method could be used to check this assumption?

$\mu = 98.6$

97.5 98.0 98.2 98.4 98.5 98.5 98.6

$n = 7$

$H_0: \mu = 98.6$ vs. $H_1: \mu < 98.6$

a) $$\bar{x} = \frac{97.5 + 98.0 + 98.2 + 98.4 + 98.5 + 98.5 + 98.6}{7}$$

$$= 98.242857$$

$$s^2 = \frac{\Sigma (x_i - \bar{x})^2}{n-1} = 0.149524$$

$$s = \sqrt{0.149524} = 0.386683$$

$$\text{test statistic} = T_0 = \frac{\bar{x} - \mu}{\frac{s}{\sqrt{n}}} = \frac{98.242857 - 98.6}{\frac{0.386683}{\sqrt{7}}} = -2.443634$$

b) 5% critical value $\Rightarrow \alpha = 0.05$

negative because $\mu < \mu_0$

$$t_{\alpha, n-1} = t_{0.05, 7-1} = t_{0.05, 6} = -1.943$$

c) p-value = 0.02511

d) **(A)** Check to see if the *p*-value in **(c)** above is less than the α value. **(B)** One-sample *t*-test for a population mean.

(C) One-sample *z*-test for a population mean.

(D) check to see if the answer in **(a)** is greater than or equal to the answer in **(b)**.

(E) check to see if the answer in **(a)** is less than the answer in **(b)**.

d.)

(A) Check to see if the p-value in **(c)** above is less than the α value. **(B)** One-sample t-test for a population mean.

(C) One-sample z-test for a population mean.

(D) check to see if the answer in **(a)** is greater than or equal to the answer in **(b)**.

(E) check to see if the answer in **(a)** is less than the answer in **(b)**.

(F) Check to see if the p-value in **(c)** above is greater than .10.

(G) Draw a normal probability plot, and see if the points are bell-shaped (like the normal curve).

(H) Draw a normal probability plot, and see if the points fall close to a straight line.