Stat 330 (Spring 2020): Homework 11

Due: May 1st, 2020

Show all of your work, and *please* staple your assignment if you use more than one sheet. Write your name, the course number and the section on every sheet. Problems marked with * will be graded and one additional randomly chosen problem will be graded.

1. * Suppose you take a random sample of 30 individuals from a large population and record a numeric value for each. For this sample, the sample mean is 4.2 and sample variance is 49. You wish to estimate the unknown population mean μ.

   (a) Calculate a 90% confidence interval for μ.

   (b) Calculate a 95% confidence interval for μ.

   (c) Based on (a) and (b), comment on what happens to the width of a confidence interval (increase/decrease) when you increase your confidence level.

   (d) Suppose your sample size is 100 instead of 30. Keep the sample mean and variance at 4.2 and 49 respectively. Calculate a new 90% confidence interval for μ.

   (e) Based on (a) and (d), comment on what happens to the width of a confidence interval (increase/decrease) when you increase your sample size keeping everything else the same.

   Answer:

   (a) For a 90% confidence interval, we use $z_{\alpha/2} = z_{0.05} = 1.65$ in the calculation.

   $$\begin{aligned}&\bar{x} \pm z_{\alpha/2}\frac{s}{\sqrt{n}}\\ &= 4.2 \pm 1.65\frac{7}{\sqrt{30}}\\ &= 4.2 \pm 2.1087\\ &= (2.0913,\ 6.3087)\end{aligned}$$

   (b) For a 95% confidence interval, we use $z_{\alpha/2} = z_{0.025} = 1.96$ in the calculation.

   $$\begin{aligned}&\bar{x} \pm z_{\alpha/2}\frac{s}{\sqrt{n}}\\ &= 4.2 \pm 1.96\frac{7}{\sqrt{30}}\\ &= 4.2 \pm 2.5049\\ &= (1.6951,\ 6.7049)\end{aligned}$$

   (c) When we increase confidence (and everything else remains the same), the width of the confidence interval increases.

   (d)

   $$\begin{aligned}&\bar{x} \pm z_{\alpha/2}\frac{s}{\sqrt{n}}\\ &= 4.2 \pm 1.65\frac{7}{\sqrt{100}}\\ &= 4.2 \pm 1.155\\ &= (3.045,\ 5.355)\end{aligned}$$

   (e) When sample size increases (and everything else remains the same), the width of the confidence interval decreases.

2. In assessing the desirability of windowless schools, officials asked 144 elementary school children whether or not they like windows in their classrooms. 43 children preferred windows.

(a) Give a 95% confidence interval of the proportion of elementary school children who like windows in their classrooms.

(b) Interpret the confidence interval you obtain.

Answer:

(a) $\hat{p} = 43/144 = 0.299$ is the sample proportion of school children who prefer windows. Using the large sample C.I. for proportion p, the 95% confidence interval is

$$\begin{aligned} &\hat{p} \pm z_{\alpha/2}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}} \\ &= 0.299 \pm 1.96\frac{\sqrt{.299(.701)}}{12} \\ &= 0.299 \pm 0.075 \\ &= (0.224,\ 0.374) \end{aligned}$$

(b) We are 95% confident that the true proportion of school children that prefer windows in their classrooms is between 0.224 and 0.374.

3. In order to ensure efficient usage of a server, it is necessary to estimate the mean number of concurrent users. According to records, the average number of concurrent users at 100 randomly selected times is 37.7. The sample standard deviation is $s = 9.2$.

(a) Construct a 90% confidence interval for the expectation of the number of concurrent users.

(b) Conduct a hypothesis test to test whether the true mean number of concurrent users is *greater* than 35. Based on your hypothesis test, do you have evidence that the true mean number of concurrent users is *greater* than 35?

Answer:

(a) $\bar{x} \pm z_{.05}\frac{s}{\sqrt{n}} = 37.7 \pm 1.645\frac{9.2}{10} = 37.7 \pm 1.5$ or $(36.2, 39.2)$

(b) $H_0 : \mu = 35$
$H_A : \mu > 35$

$$Z = \frac{\bar{X} - \mu_0}{s/\sqrt{n}} = \frac{37.7 - 35}{9.2/\sqrt{100}} = 2.9348$$

The p-value is $P(Z > 2.93) = 0.0017$. Since the $p - value$ is small, we have evidence to reject the null hypothesis. We have evidence that the true mean number of concurrent users is greater than 35.

4. * A sample of 250 items from lot A contains 10 defective items, and a sample of 300 items from lot B is contains 18 defective items.

(a) Construct a 98% confidence interval for the difference of proportions of defective items.

(b) Based on your confidence interval, is there a significant difference between the quality of the two lots?

Answer:

Here $n_1 = 250$, $n_2 = 300$, $\hat{p}_1 = 10/250 = 0.04$, and $\hat{p}_2 = 18/300 = 0.06$.

(a) A 98% confidence interval for $p_1 - p_2$ is

$$\begin{aligned} & \hat{p}_1 - \hat{p}_2 \pm z_{0.02/2}\sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}} \\ = \;& 0.04 - 0.06 \pm 2.326\sqrt{\frac{(0.04)(0.96)}{250} + \frac{(0.06)(0.94)}{300}} \\ = \;& \boxed{-0.02 \pm 0.043 \text{ or } [-0.063, 0.023]} \end{aligned}$$

(b) The null hypothesis $H_0 : p_1 = p_2$ is *not rejected* against the two-sided alternative $H_A : p_1 \neq p_2$ ($p_1 - p_2 = 0$) at the 2% level because the 98% confidence interval for $p_1 - p_2$ contains 0. No, there is no significant difference between the quality of the two lots.

5. The numbers of blocked intrusion attempts on each day during the first two weeks of the month were

$$56, 47, 49, 37, 38, 60, 50, 43, 43, 59, 50, 56, 54, 58$$

After the change of firewall settings, the numbers of blocked intrusions during the next 20 days were

$$53, 21, 32, 49, 45, 38, 44, 33, 32, 43, 53, 46, 36, 48, 39, 35, 37, 36, 39, 45.$$

(a) Construct a 95% confidence interval for the difference between the average number of intrusion attempts per day before and after the change of firewall settings.

(b) Can we claim a significant reduction in the rate of intrusion attempts? The number of intrusion attempts each day has approximately Normal distribution. Conduct a hypothesis test and state your conclusion.

Answer:

6   1 = Before change, 2 = After change

Data

$\bar{x}_1 = 50$, $s_1^2 = 58$, $n_1 = 14$

$\bar{x}_2 = 40.2$, $s_2^2 = 63.33$, $n_2 = 20$

a.) 95% CI for $\mu_1 - \mu_2$

$$(\bar{x}_1 - \bar{x}_2) \pm z_{.025}\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$$

$$= (50 - 40.2) \pm 1.96\sqrt{\frac{58}{14} + \frac{63.33}{20}} = 9.8 \pm 5.299$$

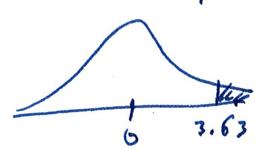
$$= (4.501, 15.099)$$

b) $H_0: \mu_1 - \mu_2 = 0$

$H_a: \mu_1 - \mu_2 > 0$

$$z = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{58}{14} + \frac{63.33}{20}}} = \frac{9.8}{2.7} = 3.63$$

Since Alternative Hypothesis is $>$, we find p-value as $P(z > 3.63)$



p-value $\simeq$ .00014

The p-value is very small $\Rightarrow$ we have evidence against H_0: and Evidence that $\mu_2 < \mu_1$ (i.e. the firewall change lowered Intrusion attempts) Note: this agrees with the CI as it contains all positive values.