Assignment 4

Introduction to Statistics (Athabasca University)

Assignment 4

Overview

Total marks: / 92

This assignment covers content from Unit 4 of the course. It assesses your knowledge of sampling distributions that refer to probability distributions of sample statistics, such as the sample mean and sample proportion, and your ability to use sampling distributions in estimation and in hypothesis testing about population means and population proportions.

Instructions

- Show all your work and justify all of your answers and conclusions, except for the TRUE/FALSE questions.
- Keep your work to 4 decimals, unless otherwise stated.
- **Note:** Finishing a test of hypotheses with a statement like "reject H_0" or "do not reject H_0" will be insufficient for full marks. You must also provide a written concluding statement in the context of the problem itself. For example, if you are testing hypotheses about the effectiveness of a medical treatment, you must conclude with a statement like, "we can conclude that the treatment is effective" or "we cannot conclude that the treatment is effective."

(9 total marks)

1. The duration of long-distance telephone calls is normally distributed with a mean of $\mu=18$ minutes and a standard deviation of $\sigma=12$ minutes. If a random sample of 64 telephone calls is used to reflect on the population of all long-distance calls, what is the probability that the sample mean call duration: $\mu=18$ $\sigma=12$ $n=64$ ~~$\bar{x}=14$~~

   (3 marks)

   a. will be more than 14 minutes? $P(x \geq 14) = P\left(z \geq \frac{14-18}{12/\sqrt{64}}\right) = \frac{-4}{1.5} = -2.67$

   $P(z \geq -2.67) = \boxed{.9962}$

   (6 marks)

   b. will be either less than 15 minutes or more than 20.5 minutes?

   $P(\bar{x} < 15 \text{ or } \bar{x} > 20.5) = 1 - P(15 < \bar{x} < 20.5)$

   $= 1 - P\left(\frac{15-18}{12/\sqrt{64}} < z < \frac{20.5-18}{12/\sqrt{64}}\right) = 1 - P(-2 < z < 1.67) = 1 - .9297 \boxed{= .0703}$

(7 total marks)

2. An insurance company states that 8% of all house insurance claims are fraudulent. If this estimate is correct, what is the probability that in a random sample of 184 house insurance claims, the proportion of claims that are fraudulent is $n=184$ $p=.08$

   (3 marks)

   a. less than 0.05?

   $P(\hat{p} < 0.05) = P\left(z < \frac{0.05-0.08}{\sqrt{\frac{0.08 \times .92}{184}}}\right)$

   $= P(z < -1.5)$

   $\boxed{= .0668}$

   (4 marks)

   b. more than 0.10? $P(\hat{p} > .10) = P\left(z > \frac{.10 - .08}{\sqrt{\frac{.08 \times .92}{184}}}\right)$

   $= P(z > 1) \boxed{= .1587}$

(11 total marks)

3. Researchers were interested in the number of monitors and screens (televisions and computer monitors) owned within households in Canada. They collected data from a random sample of ten households. The number of monitors/screens for the ten households was as follows:

   5 8 1 3 3 4 2 7 6 4

   $\bar{x} = \frac{\Sigma x}{n} = \frac{43}{10} = 4.3$ $s = \sqrt{\frac{\Sigma(x-\bar{x})^2}{n-1}} = \sqrt{\frac{44.1}{9}} = 2.2136$

   (8 marks)

   a. Assuming that this variable is normally distributed, construct a 90% confidence interval for the population mean.

   $\alpha = 1 - .90 = .10$

   $t_{\alpha/2} = t(\alpha, n-1) = t(.10, 9) = 1.833$

   $CI = \bar{x} \pm t_{\alpha/2} \times \frac{s}{\sqrt{n}}$
   $= 4.3 \pm 1.833 \times \frac{2.2136}{\sqrt{10}}$
   $= 4.3 \pm 1.203$
   $\boxed{= 3.017,\ 5.503}$

   (2 marks)

   b. In a sentence or two, describe what this confidence interval represents.

   We are 90% confident that the population mean lies between 3.017 and 5.503.

   (1 marks)

   c. Which of the following would produce a confidence interval with a larger margin of error than the 90% confidence interval? Clearly circle only one response.

   A. Sampling only 5 households instead of 10, because 5 are easier to manage.

   B. Sampling 5 households rather than 10, because a smaller sample size will result in a larger margin of error.

   C. Sampling 20 households rather than 10, because a larger sample size will result in a larger margin of error.

   D. Computing an 85% confidence interval rather than a 90% confidence interval, because a larger confidence interval will result in a larger margin of error.

(10 total marks)

4. The manager of a large production plant would like to estimate the mean amount of time a worker takes to assemble a new component. A random sample of 64 workers indicates a mean time of 16.2 minutes. Assume that the population standard deviation of this assembly time is known to be 3.6 minutes.

   $n = 64$ $\bar{x} = 16.2$ $\sigma = 3.6$

   $Z_{.05/2} = 1.96$

   $\left(16.2 \pm 1.96 \times \frac{3.6}{\sqrt{64}}\right) = (16.2 \pm .882)$
   $= 16.2 - .882,\ 16.2 + .882$
   $\boxed{= 15.318,\ 17.082}$

   (5 marks)

   a. Construct a 95% confidence interval for the mean assembly time.

   The 95% CI for μ assembly time is 15.318 - 17.082 minutes.

   (4 marks)

   b. How many workers should be involved in this study in order to have the mean assembly time estimated within 0.5 minutes of the population mean with 97.5% confidence?

   $n = \left(\frac{Z_{\left(\frac{.025}{2}\right)} \cdot \sigma}{E}\right)^2$ $E = .05$ $\sigma = 3.6$ $Z = 2.2414$

   $= \left(\frac{2.2414 \times 3.6}{.05}\right)^2$

   $= (161.38)^2$

   $n = 260$

   260 workers should be involved in the study

(8 total marks)

6. Suppose a consumer advocacy group would like to conduct a survey to find the proportion of consumers who bought the newest model of a particular vehicle who were happy with their purchase.

(4 marks)

a. How large a sample should they take so that a 90% confidence interval for the population proportion has a margin of error of 0.02? Assume that a preliminary sample suggests that 85% of consumers are satisfied.

Here z value for 90% CI is 1.645 as $P(-1.645<z<1.645)=.90$

(4 marks) $n=\left(\frac{z}{E}\right)^2 \times p \times q=\left(\frac{1.645}{.02}\right)^2 \times .85 \times .15=862.55$ $\boxed{863}$

b. Assuming that the results of the preliminary sample are not available, what is the most conservative estimate of the sample size that would limit the margin of error to within 0.05 of the population proportion for a 98% confidence interval? Z value for 98% CI is 2.326 as $P(-2.326<z<2.326)=.98$

(11 total marks) $n=\left(\frac{z}{E}\right)^2 \times p \times q=\left(\frac{2.326}{.05}\right)^2 \times .5 \times .5=541.63$ $\boxed{=542}$

7. For the past several years, the mean literacy-achievement score for a population of third-grade students has been 45 with a population standard deviation of 15. A researcher is interested in whether an experimental teaching program is more or less effective than current teaching methods for literacy. After participating in the experimental teaching program, a random sample of 100 students had a mean score of 48.75.

(8 marks)

a. Formulate and test the appropriate hypotheses using the p-value approach. Would you reject the null hypothesis $at = 0.05$? What would be your conclusion, explained within the context of the test? $\bar{x}=45$ $n=100$

(2 marks) $\sigma=15$

b. Would you reject the null hypothesis $at = 0.01$? What would be your conclusion, explained within the context of the test?

(1 mark)

c. How strong is the evidence against H_0? (Hint: Refer to Unit 4, Section 7, in the *Study Guide*).

① $H_0: \mu=48.75$
$H_A: \mu \neq 48.75$

② population Standard deviation

The significance level is $\alpha=.01$, and the critical value for a two tailed test is $z=2.58$

③ $z=\frac{\bar{x}-\mu}{\sigma/\sqrt{n}}=\frac{45-48.75}{15/\sqrt{100}}=-2.5$

Rejection region $= Z>2.58$

$Z=2.5 \leq z=2.58$, the null hypotheses is not rejected.

Using the P-value approach: the p-value is $p=.0124$, and since $p=.0124 \geq .01$, the null hypothesis is not rejected.

It is concluded that the null hypothesis H_0 is not rejected. Therefore, there is not enough evidence to claim that the population μ is different than 48.75 at the .01

 SL.

(14 total marks)

8. A researcher thinks that if hip-surgery patients go to physical therapy three times a week, instead of the usual twice a week, their recovery period will be shorter. In the past, population mean recovery time for hip-surgery patients who attended twice a week was 8.2 weeks. A random sample of 81 hip-surgery patients is selected. Each patient is asked to attend physical therapy three times a week. The sample results show a mean recovery time of 7.7 weeks and a standard deviation of 3.15 weeks.

   Using a 5% level of significance, can the researcher conclude that the mean recovery time has decreased with more physical therapy? Formulate and test the appropriate hypotheses, using both the critical value approach and the p-value approach. Provide a concluding statement, interpreted within the context of the question.

   Note: When you consult the appropriate statistical table to answer this question, use the degrees of freedom in the row following $df = 75$ (where infinity shows as the df).

   $H_0: \mu = 8.2$ $\quad n = 81$

   $H_A: \mu < 8.2$ $\quad \bar{x} = 7.7$

   $s = 3.15$

   $t = \frac{\bar{x} - \mu}{s/\sqrt{n}} = \frac{7.7 - 8.2}{3.15/\sqrt{81}} = -1.429$

   

   -1.664

   (8 marks)

   a. Critical value approach

   Reject H_0 if the test stat < -1.664

   $-1.429 > -1.664$, so we would fail to reject H_0

   (6 marks)

   b. p-value approach

   $P(t < -1.429)$ $\boxed{= .0784}$

   P-value $> \alpha = .05$, fail to reject H_0.

   So, we can not conclude that mean recovery time is decreased.

(14 total marks)

9. A teacher believes that students with natural music abilities are more likely to be left-handed than typical students are. To test this, a random sample of 80 students in a performing arts school were selected. The results showed that 25 of these students were left-handed. Within the general population, it has been determined that 20% of students are left-handed.

   Can it be concluded that the percentage of students in the performing arts school who are left-handed is greater than 20%? Formulate and test the appropriate hypotheses at a 1% significance level, using both the critical value approach and the p-value approach. Provide a concluding statement, interpreted within the context of the question.

   $H_0: p = .2$

   $H_A: p > .2$

   $n = 80$

   $z = \frac{\hat{p} - p}{\sqrt{\frac{pq}{n}}}$

   $\hat{p} = \frac{25}{80} = .3125$

   $z = \frac{.3125 - .2}{\sqrt{\frac{.2 \times .8}{80}}} = 2.5156$

   (8 marks)

   a. Critical value approach

   At a 1% z critical is 2.326. Since $2.5156 > 2.326$, so we reject H_0.

   (6 marks)

   b. p-value approach

   P-value $= .0059$

   Since p-value $.0059 < .01$ we reject H_0.

   There is sufficient evidence to conclude that the % of students in the performing art school who are left handed is greater than 20%.