(III) Inferential Statistics

Standardized test statistic: \( \frac{\text{statistic} - \text{parameter}}{\text{standard deviation of statistic}} \)

Confidence interval: statistic \( \pm \) (critical value) \( \times \) (standard deviation of statistic)

Single-Sample

<table>
<thead>
<tr>
<th>Statistic</th>
<th>Standard Deviation of Statistic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sample Mean</td>
<td>( \frac{\sigma}{\sqrt{n}} )</td>
</tr>
<tr>
<td>Sample Proportion</td>
<td>( \sqrt{\frac{p(1-p)}{n}} )</td>
</tr>
</tbody>
</table>

Two-Sample

<table>
<thead>
<tr>
<th>Statistic</th>
<th>Standard Deviation of Statistic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Difference of sample means</td>
<td>( \frac{\sigma_1^2 + \sigma_2^2}{n_1 + n_2} )</td>
</tr>
<tr>
<td>Special case when ( \sigma_1 = \sigma_2 )</td>
<td>( \sigma \sqrt{\frac{1}{n_1} + \frac{1}{n_2}} )</td>
</tr>
<tr>
<td>Difference of sample proportions</td>
<td>( \sqrt{\frac{p_1(1-p_1)}{n_1} + \frac{p_2(1-p_2)}{n_2}} )</td>
</tr>
<tr>
<td>Special case when ( p_1 = p_2 )</td>
<td>( \sqrt{\frac{1}{n_1} + \frac{1}{n_2}} )</td>
</tr>
</tbody>
</table>

Chi-square test statistic = \( \sum \frac{(\text{observed} - \text{expected})^2}{\text{expected}} \)

STATISTICS
SECTION II
Part A
Questions 1-5
Spend about 65 minutes on this part of the exam.
Percent of Section II score: 75

Directions: Show all your work. Indicate clearly the methods you use, because you will be graded on the correctness of your methods as well as on the accuracy and completeness of your results and explanations.

1. A hospital administrator noticed that the first nonemergency surgery scheduled each day often started late. If the first scheduled surgery got delayed, then all of the other surgeries scheduled for that day also got delayed. For three weeks (a total of 15 days) the administrator recorded how many minutes past the scheduled time the first surgery began each weekday. The data are shown in the table below.

<table>
<thead>
<tr>
<th>Minutes Past Scheduled Starting Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>25 45 75 50 81 40 56 40 45 14 39 60 50 39 45</td>
</tr>
</tbody>
</table>

The administrator sent a memo to the hospital's entire surgical staff to ask that everyone work to reduce the delay in the starting time for the first nonemergency surgery each day. The administrator recorded how many minutes past the scheduled starting time the first scheduled surgery began each weekday for the three weeks after the memo was sent out. The data are shown in the table below. A negative number in the table indicates that the surgery started earlier than the scheduled time.

<table>
<thead>
<tr>
<th>Minutes Past Scheduled Starting Time After the Memo Was Sent Out</th>
</tr>
</thead>
<tbody>
<tr>
<td>64 43 20 -28 15 45 72 82 25 50 62 89 -10 45 50</td>
</tr>
</tbody>
</table>

The dot plots below display the distributions of minutes past the scheduled starting time before the memo went out and after the memo went out.
2. A certain company makes three grades (A, B, and C) of a particular electrical component. Historically, grade A components have a 2 percent defective rate, grade B components have a 5 percent defective rate, and grade C components have a 10 percent defective rate. Since grade A components are less likely to be defective, the company can charge more money for those components than it can charge for the grade B or C components. Similarly, the company can charge more money for grade B components than it can charge for grade C components.

Recently, the company found a batch of components in a warehouse that were known to be of the same grade, but the grade was not labeled on the components. To determine the grade (A, B, or C), the company selected from that batch a random sample of 200 components, which contained 16 defective components.

(a) Construct and interpret a 95 percent confidence interval for the proportion of defective components in the batch.

Construct a 95% CI for the proportion of defective components in the batch.

\[
\text{Random} - \text{the data were not randomly selected because the administrator selected 15 days before & after the memo to collect data.}
\]

\[
\text{Independent} - \text{the starting time for surgery one day should have no impact on surgery another day.}
\]

\[
\text{Normal} - \text{the problem does not state that the population is normal, and the sample sizes are not large. The data gathered before the memo appears symmetric and mound shape with no outliers or significant skew. However, after the memo, there are outliers or at least significant skew in the lower direction.}
\]

\[
\text{If conditions are met, we will perform a 1-proportion z-interval.}
\]

\[
\text{Randomly given in the problem.}
\]

\[
\text{Independent - assuming the batch had } n \geq 20, \text{ no components, } n \leq N \text{ Normal: np (0.004 \times 0.25) } \geq 10
\]

\[
\text{We are 95% confident that the interval 0.0424 - 0.1176 captures the true proportion of defective parts.}
\]

\[
\hat{p} = 0.08 \\
\text{n} = 200
\]
If you need more room for your work in part (a), use the space below.

(b) Does the interval calculated in part (a) allow the company to clearly determine the grade of component that was produced in the batch? Explain.

We are unable to determine clearly from which batch these parts originated. Since both batches B (5% defective rate) and C (10% defective rate) are captured by the confidence interval, we are only confident that the components were not from batch A.

3. An environmental research agency conducted a study of a certain state's roadsides to estimate the mean number of discarded cans and bottles per mile of public road. The state's public roads were grouped into three types:

- Major highways: major paved roads designed for high traffic volume
- Minor highways: smaller paved roads designed for low traffic volume
- Unpaved roads: gravel and dirt roads

There are about 100,000 miles of public roads in the state. The environmental research agency defined a sampling unit to be a one-mile segment of public road. Using a database supplied by the state's department of transportation, the agency randomly selected 30 one-mile road segments for each of the three types of roads. Researchers from the agency searched the roadsides along each of the selected one-mile road segments and recorded the number of discarded cans and bottles. Results are shown in the table below.

<table>
<thead>
<tr>
<th>Type of Road</th>
<th>Sample Size</th>
<th>Sample Mean</th>
<th>Sample Standard Deviation</th>
<th>Total Number of Miles</th>
</tr>
</thead>
<tbody>
<tr>
<td>Major highways</td>
<td>30</td>
<td>11.2</td>
<td>6.4</td>
<td>24,000</td>
</tr>
<tr>
<td>Minor highways</td>
<td>30</td>
<td>32.6</td>
<td>9.6</td>
<td>58,000</td>
</tr>
<tr>
<td>Unpaved roads</td>
<td>30</td>
<td>21.7</td>
<td>8.3</td>
<td>18,000</td>
</tr>
</tbody>
</table>

(a) What is the variable of interest in the study?

Number of discarded cans and bottles in a one-mile segment of public road

What is the parameter of interest?

The mean number of discarded cans and bottles per mile of public road.
(b) Were the data in the study obtained by a simple random sample, a stratified random sample, or a cluster sample? Explain.

These data were obtained through a stratified random sample. The research agency intentionally divided public roads into 3 categories and took random samples from each category.

(c) Two methods for estimating the mean number of discarded cans and bottles per mile along all public roads in the state are given below.

Method 1
\[0.24x_{\text{inter}} + 0.38x_{\text{inter}} + 0.18x_{\text{inter}} =
\]
\[(0.24)(11.2) + (0.38)(32.6) + (0.18)(21.7) = 25.5\]

Method 2
\[
\frac{x_{\text{inter}} + x_{\text{inter}} + x_{\text{inter}}}{3} = \frac{11.2 + 32.6 + 21.7}{3} = 21.83
\]

Which of these methods gives a better estimate of this mean? Explain.

Method 1 provides a better estimate of the mean because it takes into account the relative proportions of the number of miles for each type of roadway, whereas method 2 treats all types of roadway equally, even though there are clear differences in number of miles for each road type.

4. A taxi company in a large city charges passengers a flat fee to enter a cab plus an additional fee per mile. There is also a charge for time spent stopped in traffic. The company wants to develop a new method for determining fares based on mileage and a flat fee only, not on time spent stopped in traffic. A random sample of 10 recent cab fares was selected, and the distance, in miles, and the fare, in dollars, were recorded. A regression model was fit to the data, and the output, scatter plot, and residual plot are given below.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Estimate</th>
<th>Std. Err.</th>
<th>T-Stat</th>
<th>P-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>4.296</td>
<td>0.298</td>
<td>14.40</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Mileage</td>
<td>1.229</td>
<td>0.166</td>
<td>7.418</td>
<td>&lt;0.0001</td>
</tr>
</tbody>
</table>

\[R^2 = 0.81209003\]
\[s = 0.4429646\]

(a) State the equation of the least squares regression line for these data. Define any variables used in the equation.

\[\text{Fare} = 4.296 + 1.229 \times \text{(Distance)}\]

Fare is the predicted total fare in $ for the cab ride.

GO ON TO THE NEXT PAGE.
(b) A 95 percent confidence interval for the intercept of the least squares regression line is (3.64, 4.58). Construct and interpret a 95 percent confidence interval for the slope of the least squares regression line. Assume the conditions for inference are met.

Construct a 95% confidence interval for the slope of the LSRL

<table>
<thead>
<tr>
<th>Conditions assumed to be met (per the problem)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Linear regression t-interval</td>
</tr>
<tr>
<td>We are 95% confident that the interval 0.846 - 1.612 captures the slope of the true regression line</td>
</tr>
<tr>
<td>Stat ± se × std. dev. stat 1.229 ± 7.306 × 0.166</td>
</tr>
<tr>
<td>df = 10 - 2 = 8</td>
</tr>
<tr>
<td>(0.846 - 1.612)</td>
</tr>
</tbody>
</table>

(c) The company wants to know if charging a flat fee of $5.00 and a per-mile charge of $1.50 will maintain its current revenue. Based on the information in part (b), is a flat fee of $3.00 a reasonable value? Explain.

A $3 flat fee will be insufficient to maintain revenue. While the $1.50 per mile fee is within the slope interval, the $3 flat fee is not within the intercept interval but is, in fact, below that interval. Therefore, the company will not make up for the amounts currently charged.

5. In a report to the department of transportation of a western state, a large trucking firm stated that the distribution of weights of its fully loaded tractor-trailer trucks is approximately normal with a mean of 19,016 pounds and a standard deviation of 2,324 pounds. The state police decided to check a sample of 40 of the company's trucks to test the company's claim concerning the mean weight and standard deviation of the weights of its trucks.

(b) Assume that the company's claim is true. Describe the distribution of the sample mean weight for random samples, each consisting of 40 trucks.

Since the sample size is large and the population is normal, the sample distribution will be normal with a mean the same as the population (19,016 pounds) but will have a standard deviation of 367.46 pounds ($\sigma/\sqrt{n} = 2,324/\sqrt{40}$)

(b) At the company's large terminal, a state police crew selects a random sample of 40 fully loaded trucks and finds that the mean weight of those trucks is 19,168 pounds. What is the probability that a random sample of 40 of the company's fully loaded trucks would have a mean weight of 19,168 pounds or more if the company's claim is true?

\[
\begin{align*}
\text{Given } & \mu = 19,016, \\
\text{what is the probability that a random sample of 40 trucks will have a weight } \bar{x} \geq 19,168 \\
\text{P}(\bar{x} \geq 19,168) &= \frac{z}{\sqrt{\frac{\sigma^2}{n}}} \\
&= 19168 - 19016 \\
&= 367.46 \\
\text{P} &= .340
\end{align*}
\]

The probability that a random sample of 40 trucks would have an average weight of 19,168 pounds is .340.
6. A drug company currently sells a prescription pain reliever that has been shown to be effective at lowering arthritis pain. However, since the drug also causes stomach irritation in some patients, the company has created a new formulation that it hopes will reduce that side effect.

To see if the new formulation reduces the occurrence of stomach irritation for users of the pain reliever, the company conducted a small preliminary study to compare the new formulation with the current pain reliever. In the preliminary study of 100 subjects with arthritis, 50 were randomly assigned to take the current pain reliever and 50 were randomly assigned to take the new formulation.

Patient responses at the end of the study are summarized in the table below.

<table>
<thead>
<tr>
<th>Patient Response</th>
<th>Current Pain Reliever</th>
<th>New Formulation</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Had stomach irritation</td>
<td>24</td>
<td>17</td>
<td>31</td>
</tr>
<tr>
<td>Had no stomach irritation</td>
<td>26</td>
<td>33</td>
<td>60</td>
</tr>
<tr>
<td>Total</td>
<td>50</td>
<td>50</td>
<td>100</td>
</tr>
</tbody>
</table>

(a) Do the data from the preliminary study indicate, at the 10 percent level of significance, that the new formulation helps to reduce the proportion of patients with stomach irritation compared to the current pain reliever? (The conditions for inference have been checked and verified.)

There are multiple tests you could use here. Since there are two populations, you could do a chi-square test of homogeneity. I will do a 2-proportion z-test.
If you need more room for your work for part (a), use the space below.

\[ H_0: P = P_0 \]
\[ H_a: P > P_0 \]
Where \( P \) is the proportion of patients with stomach pain using the new medicine and \( P_0 \) is the true proportion of patients with stomach pain using the old medicine.

Told in problem that conditions met 2 proportion z-test

\[
\begin{align*}
P_{\text{new}} &= .24 \\
P_{\text{old}} &= .205 \\
P_{\text{diff}} &= .34 \\
P_{\text{crit}} &= .38
\end{align*}
\]

Since \( p \) (.205) > \( \alpha \) (.10)

We fail to reject the null hypothesis that there is no sufficient evidence to show that the new medicine reduces stomach irritation.

(b) Based on your conclusion in part (a), which type of error, Type I or Type II, is possible? Describe the consequences of each error in the context of this study.

A Type II error is possible because we failed to reject the null hypothesis. If the medicine does produce less stomach irritation than the old formulation, we would have made a Type II error. In that case, the company would not manufacture the new formulation, even though it is better, and patients would have more stomach irritation than they would have if the new formulation were used.

(c) After the preliminary study, one of the researchers of the company suggested that the procedure described above would have more statistical power if the sample size was increased and recommended that a simulation study be conducted to investigate that more completely.

For the purposes of the simulation, the researcher assumed that 40 percent of the patients on the current pain reliever will have stomach irritation and that only 30 percent of the patients on the new formulation will have stomach irritation. Further, the total sample size will now be 200, with 100 subjects randomly assigned to each treatment.

One hundred trials of the simulation were conducted, and a p-value was calculated for each trial based on the testing procedure, significance level, and hypotheses used in part (a). A histogram of the p-values is given below.

![Histogram of p-values](image)

Use the information in the histogram to estimate the power of the test if the test is performed at the 0.10 level of significance. Indicate how you obtained your estimate.

Since power is the ability to correctly reject a false null hypothesis at a given alternative value, the power in this case would be around 55%. 55% of the simulations produced p-values low enough to reject the null.

(Note: this is from an old test - you would not be required to estimate the power now.)