Large Sample Tests of Hypothesis About a Population Mean

We use hypothesis tests to make an inference about some population parameter of interest, for example the mean height of trees or the proportion of oak trees in a forest. We can apply hypothesis testing to answer many questions. We might want to know if one drug is better than another drug at reducing blood pressure, or whether or not the presence of air bags in a vehicle reduce the severity of injuries in an accident. The steps for conducting a hypothesis should be stated each time one is performed.

Elements of Hypothesis Testing  (Summary from lecture notes)

1) Choose the population characteristic of interest (e.g. \( \mu \) the population mean)

2) Choose a significance level for the test  (e.g. \( \alpha = .05 \), that is there is a 5% chance of incorrectly rejecting the null hypothesis - also called Type I error)

3) State the Null Hypothesis (Ho:)
   We begin by stating the null hypothesis (Ho:). For example we state that the sample will come from a population with a mean of 20.
   \( \text{Ho: } \mu = 20 \)

4) State the Alternative Hypothesis (Ha:)
   We choose an alternative hypothesis (Ha:). For example we state that the alternative hypothesis comes from a population whose mean is not equal to 20.
   \( \text{Ha: } \mu \neq 20 \)

5) Choose a test statistic

6) Choose a Rejection Region
   We choose a rejection region such that the probability of rejecting the null hypothesis incorrectly is equal to \( \alpha \).

7) Calculate the test statistic
   We then calculate the test statistic and if the test statistic to see if it falls inside or outside the rejection region.

8) Conclusion
   We state our conclusions in the context of the question we are trying to answer.
1. Go to the class webpage and open the Minitab worksheet called DOWNTIME. A manufacturer of minicomputer systems is interested in improving its customer support services. The 40 most recent customers were surveyed to determine the amount of down time (in hours) they had experienced during the previous month. The data “Customer Number” and “Down Time” are in C1 and C2 of DOWNTIME.MTW, respectively. Produce descriptive statistics for this data using the drop down menus,

Stat-Basic Statistics-Display Descriptive Statistics

The company considers that a mean downtime of 16 hours for all its customers is acceptable. You are asked to test the hypothesis that the mean downtime is more than 16 hours using $\alpha = .05$. Firstly we will do this manually and then repeat the analysis using Minitab.

**Step 1. Choose the population characteristic of interest.** $\mu$ - The population mean

**Step 2. Choose the significance level.** $\alpha = .05$ (or 5% level)

**Step 3. State null hypothesis.** $H_0: \mu = 16$

**Step 4. State alternative hypothesis.** $H_a: \mu > 16$

**Step 5. Choose a test statistic.** In this case the test statistic chosen is

$$z = \frac{\bar{x} - \mu}{\sigma_x} \approx \frac{\bar{x} - \mu}{s_x} = \frac{s}{\sqrt{n}}$$

(NB. As $\sigma$ is unknown the sample standard deviation is taken as an approximation)

**Step 6. Choose a rejection region**

Since the alternative hypothesis only includes all means greater than 16 hours, this is a one-tailed test. The rejection region will be in the upper tail of the normal distribution. Get the $z_{crit}$ value such that 5% of the standard normal is to the right and therefore 95% is to the left using the commands given below,

MTB > INVCDF .95;
SUBC> NORMAL 0 1.

What is the answer? 

Verify this answer in your Cambridge tables.
So the null hypothesis will be rejected if the sample mean is more than \( z \) standard errors above 16. This is the rejection region.

Rejection region is \( Z > \) ________

**Step 7. Calculate the test statistic**

Fill in the following blanks in this equation.

\[
Z \approx \frac{\bar{X} - \mu}{\frac{S}{\sqrt{n}}} = \frac{\bar{X} - \mu}{S} = \frac{16 - \mu}{S} = \frac{16 - \mu}{S} = \frac{16 - \mu}{S}
\]

**Step 8. State Conclusion in the context of the question**

Reject / Fail To Reject the Ho: at \( \alpha = \) ______, that the mean downtime is equal to _____ hours.

2. MINITAB has a built in function for large sample hypothesis testing about a population mean. Repeat the above test using this function. Go to the STAT - BASIC STATISTICS – 1 Sample Z. Follow these steps.
MINITAB will now print the result of the test in the session window. You will see

N - the number in the sample;
Mean - the sample mean;
Stdev - the standard deviation of the sample;
SE Mean - the standard error of the mean;
100(1-α)% CI or upper or lower bound (i.e. a two sided or one sided confidence interval depending on the H_α);
Z - the test statistic;
P - the actual probability that the sample mean came from the population specified by the H_0;

<table>
<thead>
<tr>
<th>How do you interpret the test using the P-value?</th>
</tr>
</thead>
<tbody>
<tr>
<td>P &gt; α</td>
</tr>
<tr>
<td>P &lt; α</td>
</tr>
</tbody>
</table>

What is the P value for this hypothesis test? ____________.

Conclusion: Reject / Fail To Reject the H_0: at α = ______, that the mean downtime is less than or equal to ______ hours.

Change the Ha: to a 2-sided test and redo the test. What is the two sided 95% confidence interval for the mean? ________________________________.

Has the P value changed from last time? If so explain why.

Compare the results you got from performing the test by hand and by using the inbuilt function in Minitab. Note the two different ways of interpreting a hypothesis test:

1. Comparing the z statistic and the z critical value
2. Comparing the p-value and the alpha level.
You should get the same conclusion using either method.
Assignment

Question 1

Perform suitable hypothesis tests for 1, 2 and 3 below and include the following information in your answers:

- The null and alternative hypotheses
  (The level of $\alpha$ is assumed to be 0.05 unless otherwise stated)
- The $z$ statistic
- The p-value for the test
- The $z$ critical value (use the Cambridge tables or Minitab here)
- The 2-sided or 1-sided confidence bounds (depending on which is appropriate)
- Statistical result
- The conclusion (a brief description in words what the statistical results imply)

1. Test the hypothesis that the mean downtime value is not equal to 25.

2. Suppose management in customer services decided that if the mean downtime was less than 14, some redundancies would have to occur within the department. From this data is there evidence, with $\alpha=0.1$ that redundancies will occur?

3. Suppose a consumer magazine decided to print its opinion of the worst companies who manufacture minicomputer systems. The companies with average downtime hours of more than 24 decide this criterion. Based on this data, will the company be included in the list?

Question 2

The customer numbered from 230 – 260 have business contracts with the minicomputer systems and would be more valued customers.

The company is interested in determining if the downtime for customers numbered 230 – 260 is significantly greater than 16. Is this so? Explain clearly all the steps you are taking in answering this question.
REVISION SUMMARY
After this lab you should be able to:
- Understand the 8 steps in a hypothesis test
- Perform a hypothesis test by hand and in Minitab
- Interpret a p-value
- Interpret a z-statistic
- Calculate a z-critical value by hand and in Minitab (which you have done before)