LECTURE 7-2: Hypothesis tests for a mean

Chris Robertson

One sample z test for a mean

$H_0: \ \mu = \ \mu_0$

Null Hypothesis

Specify Significance Level α

$H_1: \ \mu < \mu_0$

$H_1: \ \mu < \mu_0$

$H_1: \ \mu \neq \mu_0$

Pick one appropriate alternative Hypothesis

$$Z = \frac{\bar{X} - \mu_0}{\sigma / \sqrt{n}}$$

Calculate the test statistic from the data and null hypothesis

Work out the p-value of the test

Compare the p value to the significance level

Example 2

- Suppose we are interested the average number of hours students work in part time employment while studying
- Assume this is normally distributed with a standard deviation of $\sigma = 6$ hours
- We believe that its mean value used to be $\mu_0 = 15$ hours
- We want to see if this has reduced
- We then take a sample of 36 observations from the population and the mean is 12 hours

$$\bar{x} = 12; \sigma = 6; n = 36$$

Example 2

- We are now interested in finding out whether or not there is evidence to suggest that the mean value is really less than $\mu_0 = 15$

- Null hypothesis: population mean (μ) = $\mu_0 = 15$

- Alternative hypothesis: population mean (μ) < μ_0
- We now calculate a z statistic using the sample values, i.e. $\bar{x}$, etc.

$$z = \frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}} = \frac{12 - 15}{6/\sqrt{36}} = -3$$

  - This will be normally distributed with mean 0, standard deviation 1 assuming null hypothesis is true

Example 2

- The value of the z statistic will depend on the difference between the hypothesised mean (μ_0) and the mean of the sample (which estimates the population mean μ)
  - Positive value will mean that the mean is bigger that μ_0
  - A small or 0 value will mean that they are very close to each other
  - A negative value will mean that the mean is smaller
  - The larger the z statistic the greater the difference between the sample mean the μ_0

- We go the the tables of the normal distribution and calculate $P(Z < -3)$

Example 2

- Since the alternative hypothesis contains the less than (<) symbol we use the Low option in the Normal tables

- We find that $P(Z < -3) = 0.00135$

- This means that there is a 0.1% that we would see a value of the sample mean smaller that that observed if the population mean was really μ_0, or

- There is a 0.1% chance that the difference between μ_0 and our sample mean $\bar{x}$ could have occurred by chance assuming the null hypothesis is true

Example 2

- Does this mean that there is evidence to suggest that we need to reject the null hypothesis?

- That depends on the cut-off value for evidence
  - Based on the α significance level

- In this example we have not been given a cut-off value to work with but let us assume that we are working to a 1% significance level

- Since p-value was $0.001 < 0.01$ we **reject H_0**
  - **Conclude:** There is evidence at the 1% significance level to suggest the population mean μ value is not 15 and is less than 15 hours

Testing the mean

- There are many different hypothesis tests a statistician can use

- The first test we have introduced is one which allows us to test a sample mean value
  - One-sample Z-test

- The test compares the sample mean to the hypothesised mean, looks to see if there is some kind of difference that could have occurred by chance

- Hypothesis test questions are wordy. You have to read them carefully to find the information you need but it is all there

Testing the mean – Example 3

- A random sample of 47 young adult men were asked how many minutes of sport they watched daily on TV. The sample mean was 75 minutes. Previous studies have suggested that the mean number of minutes watching sport on TV is 67 minutes. If the standard deviation of such watching is known to be 23 minutes, test at the 5% significance level to see if there is evidence to suggest that the mean time spent watching sport is greater than 67 minutes.
- A test has to be applied at the 5% significance level so the cut-off for evidence is $\alpha = 0.05$

Testing the mean – Example 3

- A random sample of 47 young adult men were asked how many minutes of sport he watched daily on TV. The sample mean was 75 minutes. Previous studies have suggested that the mean number of minutes watching sport on TV is 67 minutes. If the standard deviation of such watching is known to be 23 minutes, test at the 5% significance level to see if there is evidence to suggest that the mean time spent watching sport is greater than 67 minutes.
- Null hypothesis (H_0) : $\mu = 67$
- Alternative hypothesis (H_1) : $\mu > 67$

Testing the mean – Example 3

- Null hypothesis: $\mu = 67$
- Alternative hypothesis: $\mu > 67$

$$\mu_0 = 67$$

$$\alpha = 0.05$$

$$\bar{x} = 75$$

$$n = 47$$

$$\sigma = 23$$

- Next we calculate the test statistic

$$z = \frac{(\bar{x} - \mu)}{(\sigma / \sqrt{(n)})} = 2.385$$

- Next we calculate the p-value for the test statistic

$$p = 0.008549$$

Testing the mean – Example 3

- We now compare this p-value to our cut-off value of 0.05

- Since 0.008549 < 0.05 we make the following conclusions
  - The result is significant at the 5% level
  - We **reject H_0**
  - There is evidence to reject the hypothesis that the mean time spent watching TV sport is 67 minutes at the 5% significance level and to suggest that the mean time spent watching sport on TV is greater than 67 minutes

Testing the mean – Example 4

- A random sample of 47 university students who take a first year statistics class was assembled. At the end of the semester the students were asked how many hours they had spent working on the class outside of lectures. The sample mean was 58 hours. The standard deviation was known (from previous studies) to be 8hrs. Staff teaching the course recommend that each students spends 6 hours per week for 10 weeks working on the course material. Test at the 5% significance level to see if there is evidence to suggest the students are spending less than the recommended time on their studies

Testing the mean – Example 4

$$\mu_0 = 60$$
$$\bar{x} = 58$$
$$n = 47$$
$$\sigma = 8$$

- What is the null hypothesis?

$$\mu = \mu_0 = 6 \times 10 = 60$$

- What is the alternative hypothesis?

$$\mu < 60$$

- What is the standardised test statistic?

$$z = \frac{(\bar{x} - \mu)}{(\sigma/\sqrt{n})} = \frac{(58 - 60)}{(8/\sqrt{(47})} = -1.7139$$

Testing the mean – Example 4

- What is p-value for the test?

$$p = 0.04327$$

- What are your conclusions?
  - $0.04327 < 0.05$
  - The test is significant at the 5% significance level
  - **There is evidence to reject H_0 at the 5% significance level**
  - We conclude that there is evidence to suggest that students are spending less than the recommended time on self study

Summary

- Null hypothesis (H_0) represents belief before collecting the data.
    - Very specific hypothesis so that the distribution of the test statistic is known if H_0 is true
- Alternative hypothesis (H_1) represents the case if H_0 is not true
    - General hypothesis
- Work out the test statistic using the sample data
- Calculate the p value
    - probability of getting a more extreme value of the test statistic than the one calculated from the data

Summary

- If the p-value is small
  - Less that the specified significance level
  - There is evidence to reject the null hypothesis
- If the p-value is large
  - Bigger than the specified significance level
  - There is no evidence from the data to reject the null hypothesis