Stat 201: Statistics I
Week 9

Metropolitan State University

Week 9
Inference for Numerical Data

Section 9.1
One sample hypothesis tests for means

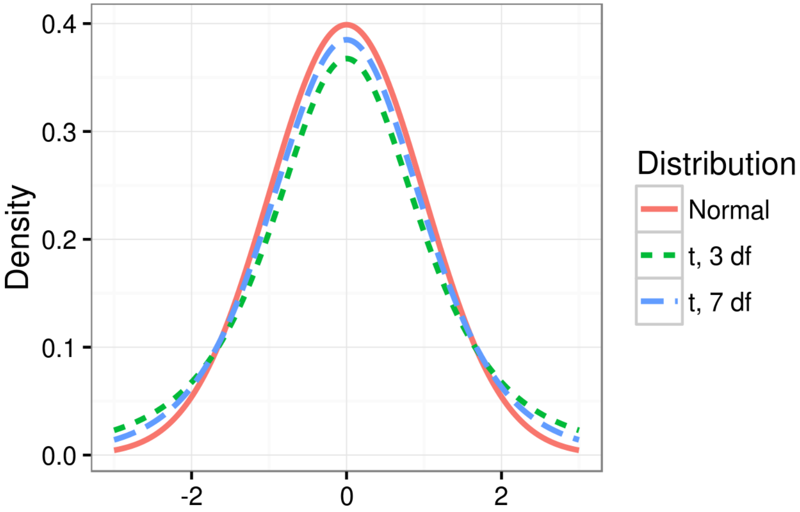
Student's t distribution

Recall, the Central Limit Theorem states that, under the right conditions, a sampling distribution is normally distributed with a standard error defined by the population standard deviation. However, when the population standard deviation is unknown, it must be estimated by the sample standard deviation.

Student's t distribution is similar to a normal distribution, except with an adjusted shape to account for an estimated standard deviation.

- The t distribution has an added parameter known as the degrees of freedom (df).
- The degrees of freedom for a sampling distribution is defined as sample size minus one ($df = n - 1$).
- As degrees of freedom increases, the t distribution approaches a normal distribution.

Student's t distribution



Steps for a mean hypothesis test

1. Identify null and alternative hypotheses from research question
   $H_0 : \mu = \mu_0$
   $H_a : \mu \neq \mu_0,\ \mu < \mu_0,\ \mu > \mu_0$
2. Use a t distribution
3. Calculate t test statistic
4. Calculate p-value
5. Compare p-value to significance level α and report decision
6. State conclusion in terms of original research question

Test statistic for mean tests

Tests for population means use the t sampling distribution. The test statistic is a t-score.

$$t = \frac{\bar{x} - \mu_0}{s/\sqrt{n}}$$

where $\bar{x}$ is the sample mean, μ_0 is the population mean under the null hypothesis, s is the sample standard deviation and n is the sample size.

p-values for mean tests

Given a t test statistic, the p-value is the probability of getting t-scores more extreme on a t distribution.

- For two-sided test:
  $H_a : \mu \neq \mu_0$, p-value $= P(T < -t) + P(T > t) = \begin{cases} 2 \times P(T > t), & t > 0 \\ 2 \times P(T < t), & t < 0 \end{cases}$
- For one-sided test:
  $H_a : \mu < \mu_0$, p-value $= P(T < t)$
  $H_a : \mu > \mu_0$, p-value $= P(T > t)$

Hypothesis test for a mean, example

Example

A study is conducted of the heights of Metro State students. The heights of a sample of 35 male students are measured. The mean height from the sample is $\bar{x} = 68.1$ with a standard deviation of $s = 3.5$. Conduct a test at $\alpha = 0.05$ level of significance of the claim that male Metro State students are shorter than the general population of 69.2 inches.

1. Identify null and alternative hypotheses from research question
   $H_0 : \mu = 69.2$
   $H_a : \mu < 69.2$
   Population: Male Metro State students

Hypothesis test for a mean, example

Example

2. Use a t distribution
3. Calculate t test statistic
   $$t = \frac{\bar{x} - \mu_0}{s/\sqrt{n}} = \frac{68.1 - 69.2}{3.5/\sqrt{35}} = -1.859339$$
4. Calculate p-value
   $$p = P(T < t) = P(T < -1.859) = 0.0358$$
5. Compare p-value to significance level α and report decision
   $p = 0.0358 < \alpha = 0.05$. Reject the null hypothesis.
6. State conclusion in terms of original research question
   There is evidence that male Metro State students are shorter than the general population.

Hypothesis test for a mean, example

Example

A clinical trial of an experimental drug for a rare disease is conducted. The survival times in months of 10 patients given the new drug are below. Test the claim that the drug changes the survival time for the expected survival time of 36 months at a 0.01 level of significance.

35.5, 40.5, 44.7, 37.2, 34.7, 41.3, 37.8, 34.5, 39.8, 37.1

1. Identify null and alternative hypotheses from research question
   $H_0 : \mu = 36$
   $H_a : \mu \neq 36$
   Population: Patients with disease taking experiment drug

Hypothesis test for a mean, example

Example

2. Use a t distribution
3. Calculate t test statistic
   $t = 2.2462$
4. Calculate p-value
   $p = 0.05132$
5. Compare p-value to significance level α and report decision
   $p = 0.05132 > \alpha = 0.01$. Do not reject the null hypothesis.
6. State conclusion in terms of original research question
   There is not evidence that the experimental drug changes survival time.

Section 9.2
Two sample hypothesis tests for means

Two sample hypothesis testing

Previously, a single sample was used to test whether the population from which the sample was drawn was the same or similar to a known population. More precisely, an hypothesis test compares an unknown population parameter to a known value.

Hypothesis testing can be used to compare two unknown populations using samples drawn from each population. This is unsurprisingly known as **two sample hypothesis testing**. These tests compare two unknown population parameters.

Null hypotheses

The null hypothesis is the claim that nothing interesting has occurred. For one sample tests, the null hypothesis states that the unknown population parameter is the same as the known parameter or some other known value.

For two sample tests, the null hypothesis states that the two unknown parameters are the same.

For example, if comparing means, $H_0 : \mu_1 = \mu_2$, or the mean of population 1 is the same as the mean of population 2.

Alternative hypotheses

The alternative hypothesis is then the claim that something interesting has occurred. For one sample tests, the alternative hypothesis states that the unknown population parameter is somehow different from the known parameter. The unknown parameter might be less than, greater than or simple not equal to the known value.

For two sample tests, the alternative hypothesis states that the two unknown parameters are different. Similar to a one sample test, the alternative hypothesis could state that the first population parameter is less than, greater than or not equal to the second population parameter.

For example, if comparing means, $H_a : \mu_1 < \mu_2$ or $\mu_1 > \mu_2$ or $\mu_1 \neq \mu_2$.

Conducting two sample hypothesis test

Once the null and alternative hypotheses are determined, two sample hypothesis tests are carried out almost identically to one sample hypothesis tests.

Assuming the null hypothesis is true, a test statistic representing the place of the samples within the appropriate sampling distribution is calculated. A p-value representing the probability of obtaining the test statistic or one more extreme is calculated. If the p-value is below a pre-determined threshold (the significance level), then the null hypothesis is rejected and it can be said that there is evidence for the alternative hypothesis.

Independent vs. dependent samples

Before conducting a test, it must be determined whether the samples are independent or dependent.

Independent samples are samples that come from distinct populations where the values from one sample do not affect the values from the other samples.

Conversely, **dependent samples** are samples that often involve the same subjects with measurements taken at different times.

Independent vs dependent samples, example

Example

Independent samples:

- Scores on statistic final in one class vs another class
- Heights of men vs heights of women
- In a clinical trial, outcomes for patients given experiment treatment vs patients given placebo (control)

Dependent samples:

- Scores on midterm exam vs final exam
- Heights of husbands vs heights of wives
- In a case control study, risk factors for subjects with disease vs matched subjects without disease

Null and alternative hypotheses, independent samples

The null hypothesis for two sample mean tests is always $H_0 : \mu_1 = \mu_2$ or $H_0 : \mu_1 - \mu_2 = 0$.

Alternative hypotheses can be any of

$$H_a : \mu_1 < \mu_2,\ \mu_1 > \mu_2,\ \mu_1 \neq \mu_2$$

or the equivalent forms of

$$H_a : \mu_1 - \mu_2 < 0,\ \mu_1 - \mu_2 > 0,\ \mu_1 - \mu_2 \neq 0$$

Requirements

Both samples must be independent and simple random samples drawn from normal populations or have a sample size of at least 30.

Test statistic

As with one sample mean tests, two sample mean tests use a t distribution. Similar to two sample proportion tests, a t statistic is calculated by

$$t = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

Recall, the degrees of freedom of a t distribution is defined by sample size minus one $(n-1)$. For a two sample test, a conservative approach is to use the smaller of the degrees of freedom from the two samples.

$$df = \min\left[(n_1 - 1), (n_2 - 1)\right]$$

When using technology, a more complicated, but accurate value will be used for the degrees of freedom.

Hypothesis test for two means, example

Example

A study is conducted of the heights of Metro State students. The heights of a sample of 32 male students taking statistics (population 1) have a mean of $\bar{x} = 70.1$ with a standard deviation of $s = 3.5$. The heights of a sample of 38 male students taking accounting (population 2) have a mean of $\bar{x} = 67.9$ with a standard deviation of $s = 3.1$.

Conduct a test at $\alpha = 0.05$ level of significance of the claim that statistics students are taller than accounting students.

1. Identify null and alternative hypotheses from research question
   $H_0 : \mu_1 = \mu_2$ of $H_0 : \mu_1 - \mu_2 = 0$
   $H_a : \mu_1 > \mu_2$ or $H_a : \mu_1 - \mu_2 > 0$

Hypothesis test for two means, example

Example

2. Use a t distribution
3. Calculate t test statistic
$$t = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} = \frac{70.1 - 67.9}{\sqrt{\frac{3.5^2}{32} + \frac{3.1^2}{38}}} = 2.759269$$
4. Calculate p-value
$p = P(T > t) = P(T > 2.759) = 0.0048$
5. Compare p-value to significance level α and report decision
$p = 0.0048 < \alpha = 0.05$. Reject the null hypothesis.
6. State conclusion in terms of original research question
There is evidence that male statistics students are taller than male accounting students.

Hypothesis tests for dependent samples

Conceptually, when working with two dependent samples, also known as matched pairs, a new value $d = x_1 - x_2$ is calculated for each pair. Then, the null hypothesis $H_0 : \mu_1 = \mu_2$ becomes $H_0 : \mu_D = 0$ and the test is essentially a one sample hypothesis test for the mean of d.

Matched pairs tests, when appropriate, are preferable to independent samples tests because the variability between subjects is almost eliminated. Thus, the total variability is reduced resulting in a more accurate test.

Conducting a dependent samples test

Since a dependent samples test is essentially a one sample test of the sample of difference values d, null and alternative hypotheses should be stated in terms of μ_D the population mean of the differences.

For example, $H_0 : \mu_D = 0$ and $H_a : \mu_D < 0$.

The test statistic is calculated as a one sample statistic from the sample of differences

$$t = \frac{\bar{d}}{\frac{s_d}{\sqrt{n}}}$$

Hypothesis test for matched pairs, example

Example

A study is conducted to see if statistics students change their scores from the midterm exam to the final exam. 42 students who took both exams are randomly selected. Their scores are found in the file "scores.csv".

Test at significance level of 0.01 whether scores changed between the midterm and the final.

1. Identify null and alternative hypotheses from research question
   $H_0 : \mu_D = 0$
   $H_a : \mu_D \neq 0$

Hypothesis test for matched pairs, example

Example

2. Use a t distribution
3. Calculate t test statistic
   $t = -3.8176772$
4. Calculate p-value
   $p = 0.0004$
5. Compare p-value to significance level α and report decision
   $p = 0.0004 < \alpha = 0.01$. Reject the null hypothesis.
6. State conclusion in terms of original research question
   There is evidence that statistics exam scores change from the midterm to the final.

Teaching the null hypothesis