Lecture 4 – Tests concerning μ and σ^2

Conducting a one-sample test on independent data

One sample tests are of the type “is the average I am observing from the sample different from some sort of value?”

The test statistic is calculated as the difference between the sample average and the null hypothesised value divided by an index of the repeatability of sample – the standard error of the mean. Large differences would suggest departure from the null hypothesis.

The test statistic T is calculated as

$$T = \frac{\bar{y} - \mu}{s/\sqrt{n}}$$

where T is t-distributed for n < 30 and z-distributed for larger samples. I suggest sticking with the t-distribution to avoid mental clutter.

The assumptions of this test is that:

- Data are normally distributed
- Data are independently sampled of one another and
- Data were randomly sampled

Example:
You have sampled a population of fish from a dam that is supposed to have fish averaging 110 mm. The sample statistics are $\bar{y} = 130$, $s^2 = 400$ and n = 20. Is the mean length of fish different from 110 mm? Assume $\alpha = 0.05$.

Step 1: Hypothesis: $H_0 : \mu = 110mm$
$H_a : \mu \neq 110mm$

Step 2: Test statistic: Calculate a T test statistic

Step 3: Rejection region: Reject H_0 if $|T| > t_{(0.025,19)} = 2.093$. We are looking for any difference therefore it is a two-tailed test. The sample could be bigger or smaller than the null hypothesized mean.

Step 4: Calculating the test statistic

$$T = \frac{130 - 110}{20/\sqrt{20}} = 4.47$$

Step 5: Making a decision

As the observed *T* value is greater than the rejection value, then we reject the null hypothesis and accept the alternative hypothesis that mean fish length in the population is different to 110mm.

Alternatively, if the Null hypothesis were $H_a : \mu > 110mm$ then we have a two-tailed test and would reject the null hypothesis at $T > t_{(0.05,19)} = 1.729$.

Conducting a hypothesis test on two variances

A test for variance simply compares the ratio of variances and assesses whether this ratio is significantly greater than unity. This test is most often used in conjunction with a two-sample *t*-test to assess validity of the *t*-test's assumptions.

The test statistic is calculated from the ratio of the two sample variances is $F = \frac{s_1^2}{s_2^2}$ (note that $s_1^2 > s_2^2$) and is *F*-distributed with $n_1 - 1$ numerator and $n_2 - 1$ denominator degrees of freedom.

An example is provided in the next section.

Conducting a two-sample test on independent data

Two sample tests are of the type "are the sample averages that I am observing from two samples different similar to each other?" Alternatively, are the sample averages that I am observing from two samples come from the same population with a common population mean?

The test statistic, as with the one-sample test, is calculated as the difference between the sample averages and the null hypothesised value divided by an index of the repeatability of the samples – a common the standard error of the mean. Large differences would suggest departure from the null hypothesis.

The test statistic for independent samples is calculated as

$$T = \frac{\bar{y}_1 - \bar{y}_2 - D_0}{s_{pooled}\left(\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}\right)}$$

where the pooled estimate variance is $s^2_{pooled} = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}$ and D_0 is the difference between sample means, and *T* is *t*-distributed for n < 30 and *z*-distributed for larger samples.
The assumptions of this test is that:

- Data are normally distributed
- Data are independently sampled of one another and
- Data were randomly sampled
- Samples have equal variance

Note that the last assumption is crucial. Conduct a test for equality of variance <u>before</u> embarking on a two-sample test.

Example:
You have sampled two populations of a species fish and weighed both the fish and their stomach contents. If one assumes that the ratio of stomach mass to body mass is an indicator of feeding intensity, are the population feeding at the same intensity? The sample statistics relating to the stomach mass : body mass ratio are
$\bar{y}_1 = 10.04$, $s_1^2 = 4.70$ $n_1 = 25$, and $\bar{y}_1 = 12.66$, $s_1^2 = 4.74$ $n_1 = 30$. Assume $\alpha = 0.05$.

Step 1: Hypothesis: $\begin{array}{l} H_0 : \mu_1 - \mu_2 = 0 \\ H_a : \mu_1 - \mu_2 \neq 0 \end{array}$

Step 2: Test statistic: Calculate a *T* test statistic

Step 3: Rejection region: Reject H_0 if $|T| > t_{(0.025,53)} = 1.96$.

Step 4: Calculating the test statistic

- First check for equality of variances

  $F = \frac{4.73}{4.70} = 1.01$ This is not larger than $F_{0.05,29,24} = 1.67$ so the variances are equal.

- If variances are equal then calculate

$$s^2_{pooled} = \frac{(25-1)4.70 + (30-1)4.73}{25+30-2} = 4.72$$

$$T = \frac{10.04 - 12.66 - 0}{2.17\left(\sqrt{\frac{1}{25} + \frac{1}{30}}\right)} = -4.45$$

Step 5: Making a decision

As the absolute observed $|T|$ value is greater than the rejection value, then we reject the null hypothesis and accept the alternative hypothesis that the ratios between populations are different.

Conducting a two-sample test on dependent data

Data that are dependent, the measurements are from the same items, require special treatment. These data are correlated as what happens to an individual in the second sample is dependent on its value in the first sample.

Dependent, or paired, two-sample tests are of the type – is the average of the values BEFORE and AFTER an experiment conducted on the same individuals different from some value? In most instances there will be some measurement conducted before a treatment and then another one after the treatment. Does the treatment have an effect?

The test statistic is calculated from the differences between measurements, denoted as d_i where i is the ith individual.

The test statistic is, therefore, the difference between the average of the sample differences and the null hypothesised value divided by the standard error of the mean from the differences. Large differences would suggest departure from the null hypothesis.

The test statistic T for dependent (paired) samples is

$T = \dfrac{\bar{d}}{s_d / \sqrt{n}}$ where d are the deviations.

where T is t-distributed for $n < 30$ and z-distributed for larger samples. If there were two sample each with 20 data points, then because we are dealing with differences then $n = 20$ and not the summation of the two samples.

Note: This test is simply a one-sample t-test that uses the differences between the results instead of the actual data.

The assumptions of this test is that:

- Differences are normally distributed
- Differences are independently sampled of one another and
- Differences were randomly sampled

Example:

Two people measure a sample of fish to assess whether or not both researchers have similar measurements. Each fish was measured twice and the difference between measurement 1 and measurement 2 noted. The sample statistics on these differences are $\bar{d} = -0.32mm$, $s^2 = 0.99$ and $n = 20$. Do the researchers measure similarly? Assume $\alpha = 0.05$.

Step 1: Hypothesis: $$H_0 : \mu_d = 0mm$$ $$H_a : \mu \neq 0mm$$

Step 2: Test statistic: Calculate a *T* test statistic

Step 3: Rejection region: Reject H_0 if $|T| > t_{(0.025,9)} = 2.262$. We are looking for any difference therefore it is a two-tailed test. The sample could be bigger or smaller than the null hypothesized mean.

Step 4: Calculating the test statistic

$$T = \frac{-0.32 - 0}{\sqrt{0.99/20}} = -1.43$$

Step 5: Making a decision

As the observed $|T|$ value is smaller than the rejection value, we fail to reject the null hypothesis that the measurements are equal.