Chapter 8

hypothesis Testing: Two-Sample

Inference

**8.1 Introduction**

In this chapter, we cover inference using two samples.

**Objectives**:

Perform two-sample

* Paired T-test for **two dependent** samples
* T-tests for means with **equal and unequal** variance,
* F-test for **equality of two variances**
* Z-tests for two means

**Problem specifications in brief:**

We wish to test hypothesis in relation to the unknown mean (), variance () and proportion (). For example, we wish to perform the test of null hypotheses of the form:

The three possible alternative hypotheses for are

Similar alternative hypotheses follows for proportions () and variances (). The three alternatives refer to right-tailed (upper-tailed), left-tailed (lower-tailed) and two-tailed tests, respectively.

**Sections 8.2/8.3 Paired T-test/ CI for paired samples**

The paired t-test applies to two correlated or dependent samples due to measurements from the paired or before-after situations. For example, if we let and be population means that represent paired or before-after measurements, then we wish to test or equivalently, by letting, test against . This test is called a ***paired T-test***.

Given a sample of size with paired measurements, let ; ; and . Then, the test statistic to test against is given by

under hypothesis. The computation of p-value for T-test is as follows:

if (left-tailed test)

if (right-tailed test)

if when (i.e., two-tailed test)

where the test statistic follows a -distribution with and is the observed value of the test statistic .

Reject at significance level if -value, or accept otherwise.

Often, we are interested for the confidence interval (CI) for . A 100(1-)% CI for is given by:

For example, for a 95% CI for , we use .

**Assumptions for paired T-test/ CI for paired mean-difference**

The paired T-test or CI for paired mean-differencerequires the following assumptions:

1. Correlated paired samples,
2. Either the sample is large () or the population of paired differences is normally distributed,
3. Standard deviation of the population of paired differences is unknown.

**Sections 8.4/8.5 2-Sample T test/ CI with Equal Variances**

The test of , or equivalently , when and are assumed unknown but equal, is called the ***two-sample pooled T-test***. The test statistic for the test is given by

where

and are the mean of two independent samples,

is the pooled estimate of the common variance

The statistic follows a -distribution with (.

An observed value of the test statistic is denoted by , and as such, the p-value for the test is computed as in section 8.3 with degrees of freedom df=.

A 100(1-)% CI for is given by:

, where

For example, for a 95% CI for we use .

**Examples 8.2/8.3: See Activity 12.2.**

**Section 8.7 2-Sample T test with Unequal Variances**

The test of , or equivalently , when and are assumed unknown and unequal, is called the ***two-sample non-pooled T-test***. The test statistic for the test is given by

where

and are the mean of two independent samples,

and are sample variances of two independent samples,

follows an approximate -distribution with an approximate degrees of freedom ()

rounded to the nearest integer.

The computation of the p-value is same for all T-tests with the proper specification of the df.

A 100(1-)% CI for is given by:

For example, for a 95% CI for , we use

**Section 8.6 F Test for Equality of Two Variances**

The test of , or equivalently , is called an F test for equality of two population variances. The test statistic for the test is given by

where

and are sample variances of two independent samples

and are degrees of freedom associated with and , respectively. The above test statistic follows from the facts that .

The p-value for the test is computed as follows:

1. if
2. if
3. if

where is observed value of the test statistic , and F follows an distribution with numerator and denominator .

Note that in SAS, .

Reject the null hypothesis at , if the -value , and accept the null hypothesis otherwise.

An F test requires two samples be independent coming from two normal distribution.

**Section 8.8 2-Sample Z tests**

The test of , when are known, is called the ***two-sample Z-test***. The test statistic for the test is given by

When are known and equal (i.e., ) the test statistic is

The p-value for the test is computed as follows:

Note that in SAS,

A 100(1-)% CI for is given by:

Or

where . For example, for a 95% CI for , we use

**See Activities Uploaded in Canvas**