Hypothesis Testing – Two Samples Tests

Jiří Neubauer

Department of Econometrics FVL UO Brno
office 69a, tel. 973 442029
email:Jiri.Neubauer@unob.cz

Testing Equality of Variances

Let $X_1, X_2, \ldots, X_{n_1}$ be a random sample from $N(\mu_1, \sigma_1^2)$ and $Y_1, Y_2, \ldots, Y_{n_2}$ be a random sample from $N(\mu_2, \sigma_2^2)$.
S_X^2 and S_Y^2 are corresponding sample variances.

Testing Equality of Variances

Let $X_1, X_2, \dots, X_{n_1}$ be a random sample from $N(\mu_1, \sigma_1^2)$ and $Y_1, Y_2, \dots, Y_{n_2}$ be a random sample from $N(\mu_2, \sigma_2^2)$.
S_X^2 and S_Y^2 are corresponding sample variances.
The statistic

$$F = \frac{S_X^2}{S_Y^2} \cdot \frac{\sigma_2^2}{\sigma_1^2}$$

has a Fisher-Snedecor distribution with $\nu_1 = n - 1$ and $\nu_2 = n_2 - 1$ degrees of freedom.

Testing Equality of Variances

Let $x_1, x_2, \ldots, x_{n_1}$ be values of a random sample from $N(\mu_1, \sigma_1^2)$,
$y_1, y_2, \ldots, y_{n_2}$ be values of a random sample from $N(\mu_2, \sigma_2^2)$,
s_x^2 and s_y^2 corresponding values of sample variances.

Testing Equality of Variances

Let $x_1, x_2, \dots, x_{n_1}$ be values of a random sample from $N(\mu_1, \sigma_1^2)$,
$y_1, y_2, \dots, y_{n_2}$ be values of a random sample from $N(\mu_2, \sigma_2^2)$,
s_x^2 and s_y^2 corresponding values of sample variances.

We test a hypothesis that the parametr σ_1^2 is equal to the parameter σ_2^2:

$$H : \sigma_1^2 = \sigma_2^2,$$

the test statistic is

$$F = \frac{s_x^2}{s_y^2},$$

which has under the null hypothesis H a Fisher-Snedecor distribution with $\nu_1 = n_1 - 1$ and $\nu_2 = n_2 - 1$ degrees of freedom.

Testing Equality of Variances

According to the alternative hypothesis we construct following regions of rejection:

alternative hypothesis	rejection region
$A: \sigma_1^2 > \sigma_2^2$	$W_\alpha = \{F, F \geq F_{1-\alpha}(\nu_1, \nu_2)\}$
$A: \sigma_1^2 < \sigma_2^2$	$W_\alpha = \{F, F \leq F_{\alpha}(\nu_1, \nu_2)\}$
$A: \sigma_1^2 \neq \sigma_2^2$	$W_\alpha = \left\{F, F \leq F_{\frac{\alpha}{2}}(\nu_1, \nu_2) \vee F \geq F_{1-\frac{\alpha}{2}}(\nu_1, \nu_2)\right\}$

where $F_\alpha(\nu_1, \nu_2)$, $F_{1-\alpha}(\nu_1, \nu_2)$, $F_{\frac{\alpha}{2}}(\nu_1, \nu_2)$, $F_{1-\frac{\alpha}{2}}(\nu_1, \nu_2)$ are quantiles of the Fisher-Snedecor distribution, $\nu_1 = n_1 - 1$, $\nu_2 = n_2 - 1$.

Testing Equality of Means ($\sigma_1^2 = \sigma_2^2$)

Let $X_1, X_2, \dots, X_{n_1}$ be a random sample from $N(\mu_1, \sigma_1^2)$ and $Y_1, Y_2, \dots, Y_{n_2}$ be a random sample from $N(\mu_2, \sigma_2^2)$. We assume that these random samples are *independent*.

Testing Equality of Means ($\sigma_1^2 = \sigma_2^2$)

Let $X_1, X_2, \ldots, X_{n_1}$ be a random sample from $N(\mu_1, \sigma_1^2)$ and $Y_1, Y_2, \ldots, Y_{n_2}$ be a random sample from $N(\mu_2, \sigma_2^2)$. We assume that these random samples are *independent*.

$\overline{X}$, $\overline{Y}$, S_X^2 a S_Y^2 are corresponding sample means and variances. If $\sigma_1^2 = \sigma_2^2$, then a statistic

$$T = \frac{\overline{X} - \overline{Y} - (\mu_1 - \mu_2)}{S} \sqrt{\frac{n_1 \cdot n_2}{n_1 + n_2}},$$

where

$$S = \left[\frac{(n_1 - 1)S_X^2 + (n_2 - 1)S_Y^2}{n_1 + n_2 - 2} \right]^{1/2}$$

(it is co called pooled estimator of the common σ) has a Student distribution with $\nu = n_1 + n_2 - 2$ degrees of freedom.

Testing Equality of Means ($\sigma_1^2 = \sigma_2^2$)

Let $x_1, x_2, \ldots, x_{n_1}$ be values of a random sample from $N(\mu_1, \sigma_1^2)$, $y_1, y_2, \ldots, y_{n_2}$ be values of a random sample from $N(\mu_2, \sigma_2^2)$
$\overline{x}$, $\overline{y}$, s_x^2 a s_y^2 are corresponding values of sample means and variances.

Testing Equality of Means ($\sigma_1^2 = \sigma_2^2$)

Let $x_1, x_2, \ldots, x_{n_1}$ be values of a random sample from $N(\mu_1, \sigma_1^2)$, $y_1, y_2, \ldots, y_{n_2}$ be values of a random sample from $N(\mu_2, \sigma_2^2)$
$\overline{x}$, $\overline{y}$, s_x^2 a s_y^2 are corresponding values of sample means and variances.

We test a hypothesis that the parameter μ_1 is equal to the parameter μ_2 ($\sigma_1^2 = \sigma_2^2$):

$$H : \mu_1 = \mu_2,$$

a test statistic is

$$t = \frac{\overline{x} - \overline{y}}{S} \sqrt{\frac{n_1 \cdot n_2}{n_1 + n_2}},$$

where

$$S = \left[\frac{(n_1 - 1)s_x^2 + (n_2 - 1)s_y^2}{n_1 + n_2 - 2} \right]^{1/2}$$

has under the null hypothesis H the Student distribution with $\nu = n_1 + n_2 - 2$ degrees of freedom.

Testing Equality of Means ($\sigma_1^2 = \sigma_2^2$)

According to the alternative hypothesis we construct following regions of rejection:

alternative hypothesis	rejection region
$A: \mu_1 > \mu_2$	$W_\alpha = \{t, t \geq t_{1-\alpha}(\nu)\}$
$A: \mu_1 < \mu_2$	$W_\alpha = \{t, t \leq -t_{1-\alpha}(\nu)\}$
$A: \mu_1 \neq \mu_2$	$W_\alpha = \left\{t, \lvert t\rvert \geq t_{1-\frac{\alpha}{2}}(\nu)\right\}$

where $t_{1-\alpha}(\nu)$, $t_{1-\frac{\alpha}{2}}(\nu)$ are quantiles of the Student distribution, $\nu = n_1 + n_2 - 2$.

Testing Equality of Means ($\sigma_1^2 \neq \sigma_2^2$)

Let $X_1, X_2, \ldots, X_{n_1}$ be a random sample from $N(\mu_1, \sigma_1^2)$ and $Y_1, Y_2, \ldots, Y_{n_2}$ be a random sample from $N(\mu_2, \sigma_2^2)$. We assume that these random samples are *independent*.

Testing Equality of Means ($\sigma_1^2 \neq \sigma_2^2$)

Let $X_1, X_2, \ldots, X_{n_1}$ be a random sample from $N(\mu_1, \sigma_1^2)$ and $Y_1, Y_2, \ldots, Y_{n_2}$ be a random sample from $N(\mu_2, \sigma_2^2)$. We assume that these random samples are *independent*.

$\overline{X}$, $\overline{Y}$, S_X^2 a S_Y^2 are corresponding sample means and variances. If $\sigma_1^2 = \sigma_2^2$, then a statistic

$$T = \frac{\overline{X} - \overline{Y} - (\mu_1 - \mu_2)}{\sqrt{\frac{S_X^2}{n_1} + \frac{S_Y^2}{n_2}}},$$

has approximately a Student distribution with ν degrees of freedom.

Testing Equality of Means ($\sigma_1^2 \neq \sigma_2^2$)

Let $x_1, x_2, \ldots, x_{n_1}$ be values of a random sample from $N(\mu_1, \sigma_1^2)$,
$y_1, y_2, \ldots, y_{n_2}$ be values of a random sample from $N(\mu_2, \sigma_2^2)$
$\overline{x}$, $\overline{y}$, s_x^2 a s_y^2 are corresponding values of sample means and variances.

Testing Equality of Means ($\sigma_1^2 \neq \sigma_2^2$)

Let $x_1, x_2, \ldots, x_{n_1}$ be values of a random sample from $N(\mu_1, \sigma_1^2)$, $y_1, y_2, \ldots, y_{n_2}$ be values of a random sample from $N(\mu_2, \sigma_2^2)$
$\overline{x}$, $\overline{y}$, s_x^2 a s_y^2 are corresponding values of sample means and variances.

We test a hypothesis that the parameter μ_1 is equal to the parameter μ_2 ($\sigma_1^2 \neq \sigma_2^2$):

$$H : \mu_1 = \mu_2,$$

a test statistic

$$t = \frac{\overline{x} - \overline{y}}{\sqrt{\frac{s_x^2}{n_1} + \frac{s_y^2}{n_2}}},$$

has under the null hypothesis H approximately a Student distribution with ν degrees of freedom.

Testing Equality of Means ($\sigma_1^2 \neq \sigma_2^2$)

Degrees of freedom are given by a formula

$$\nu \approx \frac{\left(\frac{s_x^2}{n_1} + \frac{s_y^2}{n_2}\right)^2}{\frac{1}{n_1-1}\left(\frac{s_x^2}{n_1}\right)^2 + \frac{1}{n_2-1}\left(\frac{s_y^2}{n_2}\right)^2}$$

rounded down to the nearest integer number.

Testing Equality of Means ($\sigma_1^2 \neq \sigma_2^2$)

According to the alternative hypothesis we construct following regions of rejection:

alternative hypothesis	rejection region
$A: \mu_1 > \mu_2$	$W_\alpha = \{t, t \geq t_{1-\alpha}(\nu)\}$
$A: \mu_1 < \mu_2$	$W_\alpha = \{t, t \leq -t_{1-\alpha}(\nu)\}$
$A: \mu_1 \neq \mu_2$	$W_\alpha = \left\{t, \lvert t\rvert \geq t_{1-\frac{\alpha}{2}}(\nu)\right\}$

where $t_{1-\alpha}(\nu)$, $t_{1-\frac{\alpha}{2}}(\nu)$ are quantiles of the Student distribution with ν degrees of freedom (see the previous page).

Testing Equality of Means – Large Samples

Let $X_1, X_2, \ldots, X_{n_1}$ be a random sample from a distribution with the mean μ_1 and $Y_1, Y_2, \ldots, Y_{n_2}$ be a random sample from a distribution with the mean μ_2. We assume that these random samples are *independent* and samples are large enough.

Testing Equality of Means – Large Samples

Let $X_1, X_2, \ldots, X_{n_1}$ be a random sample from a distribution with the mean μ_1 and $Y_1, Y_2, \ldots, Y_{n_2}$ be a random sample from a distribution with the mean μ_2. We assume that these random samples are *independent* and samples are large enough.

$\overline{X}$, $\overline{Y}$, S_X^2 a S_Y^2 are corresponding sample means and variances. A statistic

$$U = \frac{\overline{X} - \overline{Y} - (\mu_1 - \mu_2)}{\sqrt{\frac{S_X^2}{n_1} + \frac{S_Y^2}{n_2}}},$$

has approximately a normal distribution $N(0, 1)$.

Testing Equality of Means – Large Samples

Let $x_1, x_2, \ldots, x_{n_1}$ are values of a random sample from the first distribution, $y_1, y_2, \ldots, y_{n_2}$ are values of a random sample from the second distribution,
$\overline{x}$, $\overline{y}$, s_x^2 and s_y^2 are corresponding values of sample means and variances.

Testing Equality of Means – Large Samples

Let $x_1, x_2, \ldots, x_{n_1}$ are values of a random sample from the first distribution, $y_1, y_2, \ldots, y_{n_2}$ are values of a random sample from the second distribution,
$\overline{x}$, $\overline{y}$, s_x^2 and s_y^2 are corresponding values of sample means and variances.
We test a hypothesis that the parameter μ_1 is equal to the parameter μ_2:

$$H : \mu_1 = \mu_2,$$

a test statistic

$$u = \frac{\overline{x} - \overline{y}}{\sqrt{\frac{s_x^2}{n_1} + \frac{s_y^2}{n_2}}},$$

has under the null hypothesis H approximately a normal distribution $N(0, 1)$.

Testing Equality of Means – Large Samples

According to the alternative hypothesis we construct following regions of rejection:

alternative hypothesis	rejection region
$A: \mu_1 > \mu_2$	$W_\alpha = \{u, u \geq u_{1-\alpha}\}$
$A: \mu_1 < \mu_2$	$W_\alpha = \{u, u \leq -u_{1-\alpha}\}$
$A: \mu_1 \neq \mu_2$	$W_\alpha = \left\{u, \lvert u \rvert \geq u_{1-\frac{\alpha}{2}}\right\}$

where $u_{1-\alpha}$, $u_{1-\frac{\alpha}{2}}$ are quantiles of $N(0,1)$.

Testing Equality of Means – Paired Samples

Let us have two dependent samples, two data values – one for each sample – are collected from some source (or element). These are also called paired or matched samples. We assume two **dependent** random variables X and Y with means μ_1 and μ_2, the difference $D = X - Y$ is a random variable too. Let $D_1, D_2, \ldots, D_n$ be a random sample, where differences $D_i = X_i - Y_i$ have a normal distribution $N(\mu, \sigma^2)$, where $\mu = \mu_1 - \mu_2$ (σ^2 is not needed).

Testing Equality of Means – Paired Samples

Let us have two dependent samples, two data values – one for each sample – are collected from some source (or element). These are also called paired or matched samples. We assume two **dependent** random variables X and Y with means μ_1 and μ_2, the difference $D = X - Y$ is a random variable too. Let $D_1, D_2, \ldots, D_n$ be a random sample, where differences $D_i = X_i - Y_i$ have a normal distribution $N(\mu, \sigma^2)$, where $\mu = \mu_1 - \mu_2$ (σ^2 is not needed).
A statistic

$$T = \frac{\overline{D} - \mu}{S_D}\sqrt{n},$$

where $\overline{D}$ is a sample mean of differences and S_D is a sample standard deviation of differences, has a Student distribution with $\nu = n - 1$ degrees of freedom.

Testing Equality of Means – Paired Samples

Let $d_1 = x_1 - y_1, d_2 = x_2 - y_2, \dots, d_n = x_n - y_n$ be measure valued of differences, $\overline{d}$ is its sample mean and s_d is its sample standard deviation.

Testing Equality of Means – Paired Samples

Let $d_1 = x_1 - y_1, d_2 = x_2 - y_2, \ldots, d_n = x_n - y_n$ be measure valued of differences, $\overline{d}$ is its sample mean and s_d is its sample standard deviation.

We test a hypothesis that the parameter μ_1 is equal to the parameter μ_2:

$$H : \mu_1 = \mu_2,$$

a test statistic

$$t = \frac{\overline{d}}{s_d}\sqrt{n},$$

has under the null hypothesis H a Student distribution with $\nu = n - 1$ degrees of freedom.

Testing Equality of Means – Paired Samples

According to the alternative hypothesis we construct following regions of rejection:

alternative hypothesis	**rejection region**
$A : \mu_1 > \mu_2$	$W_\alpha = \{t, t \geq t_{1-\alpha}(\nu)\}$
$A : \mu_1 < \mu_2$	$W_\alpha = \{t, t \leq -t_{1-\alpha}(\nu)\}$
$A : \mu_1 \neq \mu_2$	$W_\alpha = \left\{t, \lvert t\rvert \geq t_{1-\frac{\alpha}{2}}(\nu)\right\}$

where $t_{1-\alpha}(\nu)$, $t_{1-\frac{\alpha}{2}}(\nu)$ are quantiles of the Student distribution, $\nu = n - 1$.