Subject	Statistics
Semester	04
Paper no	10
Paper Name	Testing of Hypothesis
Topic no	13
Topic name	LRTP for Testing Equality of Two Variances of Two Univariate Normal Distributions
SME	Ms Shubharekha
ID	Ms Varsha Shetty

E-Learning Module on LRTP for Testing Equality of Two Variances of Two Univariate Normal Distributions

Learning Objectives

By the end of this session, you will be able to:

- Explain LRTP to test the difference between the variances of two Normal populations when the means are known and unknown
- Describe the procedure to test for the equality of two population variances

Introduction

Application:1

To test the null hypothesis

$$H_0 : \sigma_1 = \sigma_2 = \sigma$$

against the alternative hypothesis

$$H_1 : \sigma_1 \neq \sigma_2$$

where σ_1 and σ_2 are the unknown standard deviations of the two Normal populations

Respective means μ_1 and μ_2 are unknown

Let $\mathbf{x_1, x_2 \ldots x_n}$ be a random sample of size $\mathbf{n}$ from a Normal population with mean $\boldsymbol{\mu_1}$, standard deviation $\boldsymbol{\sigma_1}$ and

Let $\mathbf{x_1, x_2 \ldots x_m}$ be a random sample of size m from a Normal population with mean $\boldsymbol{\mu_2}$ and standard deviation $\boldsymbol{\sigma_2}$

Then the likelihood ratio **λ** is given by

$$\lambda = \frac{\underset{\sigma_1=\sigma_2=\sigma,\mu_1,\mu_2}{Sup} L(x_1, x_2 \ldots . x_n)\ L(y_1, y_2 \ldots y_m)}{\underset{\sigma_1,\sigma_2,\mu_1,\mu_2,}{Sup} L(x_1, x_2 \ldots x_n)\ L(y_1, y_2 \ldots y_m)}$$

Since all the observations are independent of each other

,

$$\lambda=\frac{\sup\limits_{\sigma_1=\sigma_2=\sigma,\mu_1,\mu_2}\left(\frac{1}{\sigma^2}\right)^{\frac{n}{2}}\left(\frac{1}{\sqrt{2\pi}}\right)^n e^{\frac{-1}{2\sigma^2}\sum_i(x_i-\mu_1)^2}\left(\frac{1}{\sigma^2}\right)^{\frac{m}{2}}\left(\frac{1}{\sqrt{2\pi}}\right)^m e^{\frac{-1}{2\sigma^2}\sum_i(y_i-\mu_2)^2}}{\sup\limits_{\sigma_1,\sigma_2,\mu_1,\mu_2}\left(\frac{1}{\sigma^2}\right)^{\frac{n}{2}}\left(\frac{1}{\sqrt{2\pi}}\right)^n e^{\frac{-1}{2\sigma^2}\sum_i(x_i-\mu_1)^2}\left(\frac{1}{\sigma^2}\right)^{\frac{m}{2}}\left(\frac{1}{\sqrt{2\pi}}\right)^m e^{\frac{-1}{2\sigma^2}\sum_i(y_i-\mu_2)^2}}$$

μ_1,μ_2,σ^2

Under the null hypothesis the m.l.e's are obtained as

$$\hat{\mu}_1 = \bar{x} \quad , \hat{\mu}_2 = \bar{y} \quad \text{and} \quad \hat{\sigma}^2 = \frac{\sum (x_i - \bar{x})^2 + \sum (y_i - \bar{y})^2}{m+n} = \frac{ns_1^2 + ms_2^2}{m+n}$$

Otherwise the m.l.e's are $\hat{\mu}_1 = \bar{x} \quad , \hat{\mu}_2 = \bar{y}$ and

$$\hat{\sigma}_1^2 = s_1^2 \qquad , \hat{\sigma}_2^2 = s_2^2$$

where $s_1^2 = \dfrac{\sum (xi - \bar{x})^2}{n}$, and $s_2^2 = \dfrac{\sum (yi - \bar{y})^2}{m}$,

$$\lambda = \frac{\left(\frac{1}{\hat{\sigma}^2}\right)^{\frac{m+n}{2}} \left(\frac{1}{\sqrt{2\pi}}\right)^{n+m} e^{\frac{-1}{2\hat{\sigma}^2}\left[\sum_i (xi-\bar{x})^2 + \sum_i (yi-\bar{y})^2\right]}}{\left(\frac{1}{\hat{\sigma}_1^2}\right)^{\frac{n}{2}} \left(\frac{1}{\sqrt{2\pi}}\right)^{n+m} e^{\frac{-1}{2\hat{\sigma}_1^2}\sum_i (xi-\bar{x})^2} \left(\frac{1}{\hat{\sigma}_1^2}\right)^{\frac{m}{2}} e^{\frac{-1}{2\hat{\sigma}_2^2}\sum_i (yi-\bar{y})^2}}$$

$$\lambda = \frac{\left[\dfrac{1}{\dfrac{ns_1^2 + ms_2^2}{m+n}}\right]^{\frac{m+n}{2}} e^{\frac{-1}{2\frac{ns_1^2+ms_2^2}{m+n}}[\sum_i (xi-\bar{x})^2 + \sum_i (yi-\bar{y})^2]}}{(\frac{1}{s_1^{\ 2}})^{\frac{n}{2}} e^{\frac{-1}{2s_1^{\ 2}}\sum_i (xi-\bar{x})^2} \quad (\frac{1}{s_1^{\ 2}})^{\frac{m}{2}} e^{\frac{-1}{2s_2^{\ 2}}\sum_i (yi-\bar{y})^2}}$$

$$\lambda = \left[\frac{1}{ns_1^2 + ms_2^2}\right]^{\frac{m+n}{2}} (s_1{}^2)^{\frac{n}{2}} (s_2{}^2)^{\frac{m}{2}} (m+n)^{\frac{m+n}{2}}$$

$$= \left[\frac{1}{ns_1^2 + ms_2^2}\right]^{\frac{m+n}{2}} (ns_1{}^2)^{\frac{n}{2}} (ms_2{}^2)^{\frac{m}{2}} \frac{(m+n)^{\frac{m+n}{2}}}{n^{\frac{n}{2}} m^{\frac{m}{2}}}$$

Dividing both numerator and denominator by

$(ms_2{}^2)^{\frac{n+m}{2}}$ and with $k = \dfrac{(m+n)^{\frac{m+n}{2}}}{n^{\frac{n}{2}} m^{\frac{m}{2}}}$

$$\lambda = \frac{\left(\frac{ns_1^2}{ms_2^2}\right)^{\frac{n}{2}}}{\left(1+\frac{ns_1^2}{ms_2^2}\right)^{\frac{m+n}{2}}} k = F_1^{\frac{n}{2}}(1+F_1)^{-(\frac{m+n}{2})} k$$

Where

$$F_1 = \frac{ns_1^2}{ms_2^2}$$

Now $\lambda \leq \lambda_\alpha \Rightarrow F_1^{\frac{n}{2}}(1+F_1)^{-(\frac{m+n}{2})} k \leq \lambda_\alpha$

Which implies $F_1 \leq C_1$ or $F_1 \geq C_2$ where

$$C_1^{\frac{n}{2}}(1+C_1)^{-(\frac{m+n}{2})} k = C_2^{\frac{n}{2}}(1+C_2)^{-(\frac{m+n}{2})} = \lambda_\alpha$$

$$\Leftrightarrow \; F_1 \frac{m-1}{n-1} \le C_1 \frac{m-1}{n-1} \quad \text{or} \quad F_1 \frac{m-1}{n-1} \ge C_2 \frac{m-1}{n-1}$$

Which implies $F \le C_1'$ or $F \ge C_2'$ where

$$F = F_1 \frac{m-1}{n-1} = \frac{(ns_1^{\,2})/n-1}{(ms_2^{\,2})/m-1} = \frac{\chi^2(n-1)}{\chi^2(m-1)}$$

is distributed as their F variable with (n-1) and (m-1) degrees of freedom

Now size of the test =α

$$\Rightarrow P\ [reject\ \ H_0 / H_0\ \ true] = \alpha$$

$$\Rightarrow P\,[F \le C_1^{'}\ or\ \ F \ge C'_2 / \sigma_1 = \sigma_2 = \sigma] = \alpha$$

$$\Rightarrow P\,[F \le F_{1-\alpha}(n-1, m-1) + P\,[F \ge F_{\alpha}(n-1, m-1)] = \alpha$$

From the table of probabilities of Snedecor's F distribution and by knowing the relation among C_1', C_1 and λ_α (or C_2', C_2 and λ_α) the test can be determined.

One Sided Tests:

Incase we have to test the null hypothesis

$$H_0: \sigma_1=\sigma_2=\sigma$$

against the alternative hypothesis

$$H_1: \sigma_1>\sigma_2$$

When the means are unknown, then we obtain the following test procedure:

To reject H_0 if the computed value of F as defined above , exceeds the table value

$F_\alpha(n-1,m-1)$

Similarly in case of testing

$$\mathbf{H_0: \sigma_1=\sigma_2=\sigma}$$

against the alternative hypothesis

$$\mathbf{H_1: \sigma_1< \sigma_2}$$

When the means are unknown, then we obtain the following test procedure:

To reject $\mathbf{H_0}$ if the computed value of F is less than the table value $F_{1-\alpha}(n-1,m-1)$

In practical situation the following procedure is followed

$$s_x^{\,2} = \frac{\sum (xi - \bar{x})^2}{n-1}, \qquad \text{and} \qquad s_y^{\,2} = \frac{\sum (yi - \bar{y})^2}{m-1},$$

are firstly computed.

Then F is defined by taking the larger of s_x^2 or s_y^2 in the numerator.

In case $s_x^2 > s_y^2$, F is defined by $F = s_x^2 / s_y^2$ and compared with the table value $F_{\alpha/2}(n-1, m-1)$

However if $s_x^2 < s_y^2$, F is defined by $F = s_y^2 / s_x^2$ and compared with the table value $F_{\alpha/2}(m-1, n-1)$

Application 2:

To test the null hypothesis

$$\mathbf{H_0: \sigma_1=\sigma_2=\sigma}$$

against the alternative hypothesis

$$\mathbf{H_1: \sigma_1\neq\sigma_2}$$

where $\boldsymbol{\sigma_1}$ and $\boldsymbol{\sigma_2}$ are the unknown standard deviations of two Normal populations when respective means $\boldsymbol{\mu_1}$ and $\boldsymbol{\mu_2}$ are known

$$\lambda=\frac{\underset{\sigma_1=\sigma_2=\sigma,\mu_1,\mu_2}{Sup}\left(\frac{1}{\sigma^2}\right)^{\frac{n}{2}}\left(\frac{1}{\sqrt{2\pi}}\right)^n e^{\frac{-1}{2\sigma^2}\sum_i (xi-\mu_1)^2}\left(\frac{1}{\sigma^2}\right)^{\frac{m}{2}}\left(\frac{1}{\sqrt{2\pi}}\right)^m e^{\frac{-1}{2\sigma^2}\sum_i (yi-\mu_2)^2}}{\underset{\sigma_1,\sigma_2,\mu_1,\mu_2}{Sup}\left(\frac{1}{\sigma^2}\right)^{\frac{n}{2}}\left(\frac{1}{\sqrt{2\pi}}\right)^n e^{\frac{-1}{2\sigma^2}\sum_i (xi-\mu_1)^2}\left(\frac{1}{\sigma^2}\right)^{\frac{m}{2}}\left(\frac{1}{\sqrt{2\pi}}\right)^m e^{\frac{-1}{2\sigma^2}\sum_i (yi-\mu_2)^2}}$$

Under the null hypothesis the m.l.e's are obtained as

$$\hat{\sigma}^2 = \frac{\sum(xi-\mu_1)^2 + \sum(yi-\mu_2)^2}{m+n}$$

Otherwise the m.l.e's are

$$\hat{\sigma}_1^{\,2} = \frac{\sum(xi-\mu_1)^2}{n}, \quad \text{and} \quad \hat{\sigma}_2^{\,2} = \frac{\sum(yi-\mu_2)^2}{m},$$

By simplifying we get

$$\lambda=\left[\frac{1}{\sum(xi-\mu_1)^2+\sum(yi-\mu_2)^2}\right]^{\frac{m+n}{2}}\left(\sum(xi-\mu_1)^2\right)^{\frac{n}{2}}\left(\sum(yi-\mu_2)^2\right)^{\frac{m}{2}}\frac{(m+n)^{\frac{m+n}{2}}}{n^{\frac{n}{2}}m^{\frac{m}{2}}}$$

Dividing both numerator and denominator by

$$\left(\sum(yi-\mu_2)^2\right)^{\frac{n+m}{2}}$$

and with

$$k=\frac{(m+n)^{\frac{m+n}{2}}}{n^{\frac{n}{2}}m^{\frac{m}{2}}}$$

$$\lambda = \frac{\left(\frac{\sum (x_i - \mu_1)^2}{\sum (y_i - \mu_2)^2}\right)^{\frac{n}{2}}}{\left(1 + \frac{\sum (x_i - \mu_1)^2}{\sum (y_i - \mu_2)^2}\right)^{\frac{m+n}{2}}} k = F_1^{\frac{n}{2}} (1 + F_1)^{-(\frac{m+n}{2})} k$$

Where $$F_1 = \frac{\sum (xi - \mu_1)^2}{\sum (yi - \mu_2)^2}$$

Now $$\lambda \le \lambda_\alpha \Rightarrow F_1^{\frac{n}{2}} (1 + F_1)^{-(\frac{m+n}{2})} k \le \lambda_\alpha$$

$$\Leftrightarrow F_1 \frac{m}{n} \le C_1 \frac{m}{n}$$ or $$F_1 \frac{m}{n} \ge C_2 \frac{m}{n}$$

Which implies $F \le C_1'$ or $F \ge C_2'$ where

$$F = F_1 \frac{m}{n} = \frac{\sum (x_i - \mu_1)^2 / n}{\sum (y_i - \mu_2)^2 / m}$$

is distributed as F variable with (n) and (m) degrees of freedom

Now size of the test $=\alpha$

$$\Rightarrow P[F \leq F_{1-\alpha}(n,m) + P[F \geq F_{\alpha}(n,m)] = \alpha$$

From the table of probabilities of Snedecor's F- distribution and by knowing the relation among C_1', C_1 and λ_α (or C_2', C_2 and λ_α), the test can be determined.

One Sided Tests:

Incase we have to test the null hypothesis

$$H_0: \sigma_1 = \sigma_2$$

against the alternative hypothesis

$$H_1: \sigma_1 > \sigma_2$$

When the means are known then we obtain the following test procedure:

To reject $\mathbf{H_0}$ if the computed value of **F** as defined above, exceeds the table value $\mathbf{F_\alpha(n,m)}$

Similarly in case of testing

$$\mathbf{H_0: \sigma_1 = \sigma_2}$$

against the alternative hypothesis

$$\mathbf{H_1: \sigma_1 < \sigma_2}$$

when the means are known then we obtain the following test procedure:

To reject $\mathbf{H_0}$ if the computed value of **F** is less than the table value $\mathbf{F_{1-\alpha}(n, m)}$