Kunthavai Naachiyaar Govt. Arts College (W) Autonomous, Thanjavur

M.Stat

CORE COURSE - X 18KP3S10

Testing of Statistical Hypothesis Hours : 6

Credit: 4

UNIT – I

Statistical Hypothesis - Simple and Composite, Null and Alternative Hypothesis. Concept on Critical Region, Types of errors, Level of Significance, Power of a test. Optimum tests – Most Powerful Test (MPT), Uniformly Most Powerful Test (UMPT) and Neyman – Pearson Lemma – Simple Problems.

UNIT – II

Likelihood Ratio Test – Definition and Properties – Likelihood ratio test for a mean of a Normal Population, equality of means of two Normal Population, Variance of Normal Population, equality of Variances of two Normal Populations.

UNIT – III

Hypothesis testing – Prior and Posterior odds, Base factor for Simple VS Simple Hypothesis, Base factor for Composite VS Composite Hypothesis. Lindley's Procedure for test of Significance, Lindley's Paradox and Decision Theoretic Approach to testing Problems.

UNIT – IV

Sequential Analysis – Wald's Sequential Probability Ratio Test, Properties, efficiency and Fundamental Identity of Sequential Analysis.

UNIT – V

Non – Parametric tests – Advantages and Disadvantages – Sign test, Median test, test for randomness, Wald – Wolfowitz run test, Kolmogrow – Smirnov (one and two samples) tests and Mann Whitney Wilcoxon U-test. Kruskal wally Test

Test Books:

1.Gupta S.C and Kapoor V.K (1993), Fundamental of Mathematical Statistics, Sultan Chand & Sons, New Delhi (Unit I, Unit II and Unit V).

2.Radhakrishna Rao C., Linear Statistical Inference and its Applications – Second Edition, Wiley Eastern Limited (Unit IV).

3. Leonard T. and Hsu.JSJ, Bayesian Methods, Cambridge University Press (Unit III).

Books for Reference:

1. Rohatgi. V and Saleh (2002), Statistical Inference, Asia Publications.

2. Lehmann.E.L, Testing of Statistical Hypothesis, John Wiley.

3. Gibbons.J.D, Non – Parametric Statistical Inference, Duxbury.

4. Berger J.O, Statistical Decision Theory and Bayesian Analysis, Sriges Verlog.

UNIT-I

Statistical hypothesis - simple and composite hypothesis:

simple definition

A statistical hypothesis is some statement or assection about a population (or) equivalently about the probability distribution characterising a population, which we want to verify on the basis of information available from a sample.

definition

If the statistical hypothesis specifies the population completely then it is termed as a simple statistical hypothesis otherwise It is called a composite statistical hypothesis.

For example:

If $x_1, x_2 \ldots x_n$ is a random sample of size n. from a normal population with mean μ and variance σ^2, then the hypothesis $H_0: \mu = \mu_0, \sigma^2 = \sigma_0^2$ is a sample hypothesis where as each of the following hypothesis is a composite hypothesis:

(i) $\mu = \mu_0$ (ii) $\sigma^2 = \sigma_0^2$ (iii) $\mu < \mu_0, \sigma^2 = \sigma_0^2$

(iv) $\mu > \mu_0, \sigma^2 = \sigma_0^2$ (v) $\mu = \mu_0, \sigma^2 < \sigma_0^2$

(vi) $\mu = \mu_0, \sigma^2 > \sigma_0^2$ (vii) $\mu < \mu_0, \sigma^2 > \sigma_0^2$

A hypothesis which does not specify completely "r" parameter of a population is termed as a composite hypothesis with r degrees of freedom.

Test of a statistical hypothesis :-

A test of a statistical hypothesis is a two-action decision problem after the experimental sample values have been obtained. the two-actions being the acceptance or rejection of the hypothesis.

Null hypothesis :-

In hypothesis testing, a statistician or decision-maker should not be motivated by prospects of profits or loss resulting from the acceptance or rejection of the hypothesis. He should be completely impartial and should have no brief for any party or company nor should he allow his personal views to influence the decision.

For example: let us consider the 'light - bulbs' problem. let us suppose that the bulbs manufactured under some standard manufacturing process have an average life of μ hours and it is proposed to test a new procedure for manufacturing light bulbs. Thus, we have two populations of bulbs, those manufactured by standard process and those manufactured by the new process. In this problem the following three hypotheses may be setup:

(i) New process is better than standard process.

(ii) New process is inferior to standard process

(iii) There is no difference between the two processes.

The first statements appear to be biased since they reflect a preferential attitude to one or the other of the two processes. Hence the best course is to adopt the hypothesis of no difference, as stated in (iii) This suggests that the statistician should take up the neutral or null hypothesis attitute regarding the outcome of the test. His attitute should be on the null or zero line in which the experimental data has the new importance and complete sale in the matter. This neutral or non-committal ~~before~~ attitude of the statistician or decision-maker. before the sample observations are taken as the keynote of the null hypothesis.

Thus in the above examples of light bulb if μ_0 is the mean life (in hours) of the bulb manufactured by the new process then the null hypothesis which is usualy denoted by H_0, can be stated as follows;

$$H_0 ; \mu = \mu_0$$

An another example let us suppose that two different concerns manufacture drugs for inclusing sleep, drug A manufactured by first concern and B manufactured by second concern each company claims that it's drug is supperier to that of the other and it is desied to test which is a supperior drug A or B?

To formulate the statistical hypothesis let x be a r.v which denotes the additional hours of sleep gined by an individual then drug A is given and let the hours of sleep gined when drug B is used. let us suppose that x and y follow the probability distribution with meane μ_x and μ_y respectively, here our null hypothesis would be that there is no difference between the effects of two drugs simpolicaly, $H_0; \mu_x = \mu_y$.

Alternative hypothesis: It is desirable to state what is called an alternative hypothesis is respective of every statistical hypothesis being tested because the acceptance or rejection are null hypothesis meaning full only when it is being tested against a rivdhypotheris which should rather be explicity mendionod.

alternative hypothesis is usually denoted by H_1. For example. In the example of light bulbs, alternative hypothesis could be $H_1 : \mu > \mu_0$ or $\mu < \mu_0$ or $\mu \neq \mu_0$. In the Example of light bulbs, alternative hypothesis could be $H_1 : \mu_x > \mu_y$ or $\mu_x < \mu_y$ or $\mu \neq \mu_y$

In both the cases the first two of the alternative hypotheses due rise to what are called 'one tailed' test and the Third alternative hypothesis results in "two tailed" test

Critical region :

Let $x_1, x_2 \dots x_n$ be the sample observations denoted by o all the values of o will be agrecate of a sample and they constitute or space called the sample space. Which is denoted by S.

Since the sample values $x_1, x_2 \dots x_n$ can be taken as a Point in N - dimensional space, We specify some region of the N - dimensional space and see whether this point lies with in this region or outside this region. We divided the whole sample space S into two disjoint parts W and S - W or $\bar{W}$ or w', The null hypothersis H_0 is rejected if the observed sample Point fake in W and if it false in $\bar{W}$ We reject H_1 and accept H_0 The region of rejection of H_0 when H_0 is true is that region of the outcome set where H_0 is rejected if the sample Point false in that region and his called (critical region ardantity the size of the critical region is α the Probability of committing type 1 error (discussed below) suppose with the test is based on of size two. Then the outcome set or the sample space is the first quatrant in a two. diamentional space and a test criterion will enable us to seprate own out come set into two complementary subsets w and $\bar{w}$

Two Types of errors :-

The decision to accept or reject the null hypothesis H_0 is made on the basis of the information supplied by the observed sample observations. The conclusion drawn on the basis of a particular sample may not always be true in respect of the population. The four possible situations that arise in any test procedure are given in the following table.

Double Dichotomy Relating to decision and hypothesis.

		Decision from sample.	
		reject H_0	Accept H_0
True State	H_0 True	wrong (Type I Error)	Correct
	H_0 false (H_1 True)	Correct	wrong (Type II Error)

from the above table it is obvious that in any testing problem we are liable to commit two types of errors.

Errors of Type I and Type II in thesis

The error of rejecting H_0 (accepting H_1) when H_0 is true is called Type I error and the error of accepting H_0 when H_0 is false (H_1 is true) is called Type II error. The probabilities of type I and type-II errors are denoted by α and β respectively. Thus

α = probability of type I error

= probability of rejecting H_0 when H_0 is true.

β = probability of type-II error

= probability of accepting H_0 when H_0 is false.

Symbolically:

$$P(x \in w \mid H_0) = \alpha, \text{ where } X = (x_1, x_2 \cdots x_n)$$

$$\Rightarrow \int_W L_0 \, dx = \alpha \longrightarrow (1)$$

Where L_0 is the likelihood function of the sample observation under H_0 and $\int dx$ represents the n-fold integral $\int\int \cdots \int dx_1 \, dx_2 \cdots dx_n$.

again

$P(x \in w \mid H_1) = \beta \Rightarrow \int_{\bar{w}} L_1 \, dx = \beta \rightarrow ①$

Where L_1 is the likelihood function of the sample observations under H_1. Since

$$\int_{w} L_1 \, dx + \int_{\bar{w}} L_1 \, dx = 1,$$

we get

$$\int_{w} L_1 \, dx = 1 - \int_{\bar{w}} L_1 \, dx = 1 - \beta \rightarrow ②a$$

$$\Rightarrow P(x \in w \mid H_1) = 1 - \beta \longrightarrow ②b.$$

level of significance:

α, the probability of type I error, is known as the level of significance of the test. It is also called the size of the critical region.

power of the Test: $1-\beta$, define in is called the power function of the test hypothesis H_0 against the alternative hypothesis H_1. the value of the power function at a parameter point is called the power of the test at that point.

Steps in solving testing of hypothesis problem:

The major steps involved in the solution of a testing of hypothesis problem may be outlined as follows:

1. Explicit knowledge of the nature of the population distribution and the parameter(s) of interest, i.e., the parameter(s) about which the hypothesis are set up.

2. Setting up of the null hypothesis H_0 and the alternative hypothesis H_1 in terms of the range of the parameter values each one embodies.

3. The choice of a suitable statistic $t = t(x_1, x_2, \ldots, x_n)$ called the test statistic which will best reflect upon the probability of H_0 and H_1.

4. Partitioning the set of possible values of the test statistic t into two disjoint sets w (called the rejection region or critical region) and $\bar{w}$ (called the acceptance region) and framing the following test:

(i) reject H_0 (i.e., accept H_1) if the value of t falls in w.

(ii) Accept H_0 if the value of t falls $\bar{W}$.

5. After framing the above test. obtain experimental sample observations, compute the appropriate test statistic and take action accordingly.

Optimum test under different situations

The discussion enables us to obtain the so called best test under different situations. In any testing problem the first two steps, viz., the from of the population distribution, the parameters(s) of interest and the framing of H_0 and H_1 should be obvious from the description of the problem. The most crucial step is the choice of the best test, i.e., the best statistic 't' and the critical region W where by best test, i.e., The best statistic 't' and the critical region W where by best we mean one which in addition to controlling α at any desired low level has the minimum type-II error β or maximum power $1-\beta$, compared to β of all other tests having this α This leads to the following difinition.

most powerful test (mp test) :- let us consider the problem of testing a simple hypothesis : $H_0 : \theta = \theta_0$ against a simple alternative hypothesis : $H_1 : \theta = \theta_1$

Definition: The critical region w is the most powerful (MP) critical region of size α (and the corresponding test a most powerful test of level α) for testing $H_0 ; \theta = \theta_0$ against $H_1 ; \theta = \theta_1$ if

$$P(x \in w \mid H_0) = \int_w L_0 \, dx = \alpha \longrightarrow ①$$

and

$$P(x \in w \mid H_1) \geq P(x \in w_1 \mid H_1) \longrightarrow ①a$$

Uniformly most powerful test (UMP test)

Let us now take up the case of testing a simple null hypothesis against a composite alternative hypothesis, e.g., of testing

$H_0 ; \theta = \theta_0$

against the alternative

$H_1 ; \theta \neq \theta_0$

In such a case, for a predetermined α, the best test for H_0 is called the uniformly most powerful test of level α.

Definition: The region w is called uniformly most powerful (UMP) critical region of size α [and the corresponding test as uniformly most powerful (UMP) test of level α] for testing $H_0 ; \theta = \theta_0$ against $H_1 : \theta \neq \theta_0$ i.e. $H_1 ; \theta = \theta_1 \neq \theta_0$

if $P(x \in w \mid H_0) = \int_w L_0 \, dx = \alpha \longrightarrow ②$

and $P(x \in w \mid H_1) \geq P(x \in w_1 \mid H_1)$ for all $\theta \neq \theta_0$

whatever the region w_1 satisfying ② ~~may be~~

critical values or significant values:

The value of test statistic which separates the critical (or rejection) region and the acceptance region is called the critical value or significant value. It depends upon:

(i) The level of significance used and

(ii) The alternative hypothesis, whether it is two taild or single-tailed.

As has been pointed out earlier, for large samples, the standardised variable corresponding to the statistic t, viz.,

$$Z = \frac{t - E(t)}{S.E(t)} \sim N(0,1), \text{ asymptotically} \quad (*)$$

as $n \to \infty$. The value of Z given by $(*)$ under the null hypothesis is known as test statistic. The critical value of the test statistic at level of significance α for a two-tailed test is given by Z_α, where Z_α is determined by the equation:

$$P(|Z| > Z_\alpha) = \alpha \longrightarrow ②c$$

i.e., Z_α is the value so that the total area of the critical region on both ~~sides~~ tails is α. Since normal probability curve is a symmetrical curve, from ②c we get

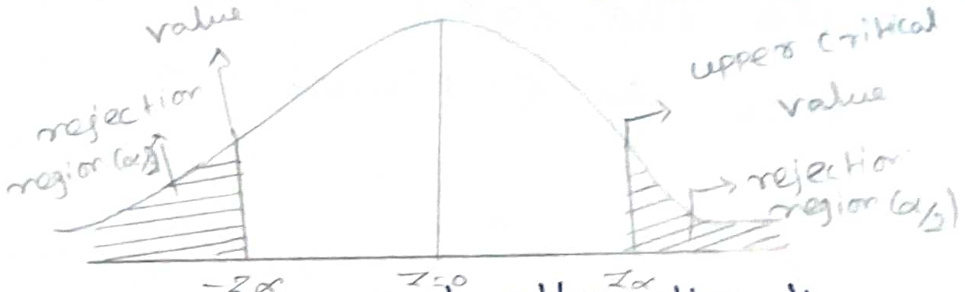
$P(Z > Z_\alpha) + P(Z < -Z_\alpha) = \alpha \Rightarrow$

$P(Z > Z_\alpha) + P(Z > Z_\alpha) = \alpha$ (By symmetry)

$2P(Z > Z_\alpha) = \alpha, \Rightarrow P(Z > Z_\alpha) = \alpha/2$

In other words, the area of each tail is $\alpha/2$. Thus Z_α is the value such that area to the right of Z_α is $\alpha/2$ and to the left of $(-Z_\alpha)$ is $\alpha/2$, as shown in the following diagram:

Two-Tailed test
(level of significance)

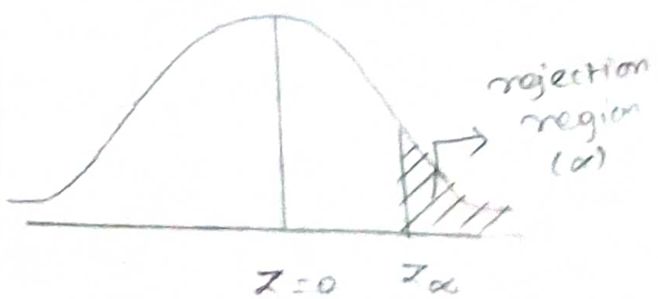


In case of single tail alternative, the critical value Z_α is determined so that total area to the right of it (For right-tailed test) is α and for left tailed test the total area to the left of $(-Z_\alpha)$ is α (see diagrams below), i.e.,

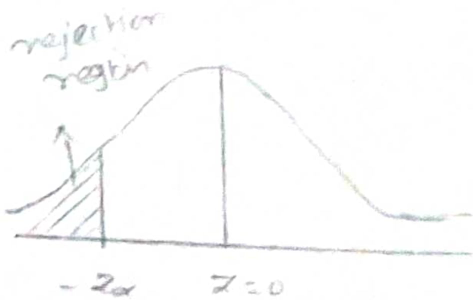
for Right tailed test : $P(Z > Z_\alpha) = \alpha$

for left tailed test : $P(Z < -Z_\alpha) = \alpha$

Right-tailed test
(level of significance α)



left-tailed test
(level of significance)



Thus the significant or critical value of Z for a single-tailed test (left or right) at level of significance 'α' is same as the critical value of Z for a two-tailed test at level of significance '2α'

We given below, the critical values of Z at commonly used levels of significance for both two taild and single tailed tests. These values have been obtained from equation (2c) (2d) (2e), on using the normal probability Tables as explained

Critical value (Z_α)	Level of Significance		
	1%	5%	10%
Two tailed test	$\lvert Z_\alpha \rvert = 2.58$	$\lvert Z_\alpha \rvert = 1.96$	$\lvert Z_\alpha \rvert = 1.645$
Right-tailed test	$Z_\alpha = 2.33$	$Z_\alpha = 1.645$	$Z_\alpha = 1.28$
left-tailed test	$Z_\alpha = -2.33$	$Z_\alpha = -1.645$	$Z_\alpha = -1.28$

likelihood Ratio test :-

Neyman-pearson lemma based on the magnitude of the ratio of two probability density functions provides the best test for testing simple hypothesis against simple alternative hypothesis. the best test in any given situation depends on the nature of the population distribution and the form of the alternative hypothesis being Considered. In this section we shall discuss a general method of test Construction Called the likelihood Ratio (L.R). Test Introduced by Neyman and pearson for testing a hypothesis, Simple or Composite, against a Simple or composite alternative hypothesis. This test is related to the maximum likelihood estimates.

properties of likelihood Ratio Test :-

likelihood ratio (L.R) test principle is an Intuitive one. If we are testing a simple hypothesis H_0 against a Simple alternative hypothesis H_1 then the LR principle leads to the Same test as given by the Neyman-pearson lemma. This suggests that L.R test has some desirable properties, specialy large sample properties.

In L.R Test probability of type I error is Controlled by suitably Choosing the cut off point λ_0. LR Test is generally. UMP if an ump test at all exists. we state below, the two asymptotic properties of LR Tests.

(i) under certain Conditions, $-2\log \lambda$ has an asymptotic chi-square distribution.

2. under certain assumptions. LR Test is Consistent.

Test for the mean of a Normal population:-

let us take the problem of testing if the mean of a normal population has a specified value. let $(x_1, x_2 \dots x_n)$ be a random sampe of size n from the normal population with mean μ and variance σ^2, where μ and σ^2 are unknown, suppose we want to test the (composite) null hypothesis.

$$H_0 ; \mu = \mu_0 \text{ (specified)}, \; 0 < \sigma^2 < \infty$$

against the Composite alternative hypothesis

$$H_1 ; \mu \neq \mu_0 ; \; 0 < \sigma^2 < \infty$$

In this case the Parameter Space Θ is given by

$$\Theta = [(\mu, \sigma^2) : -\infty < \mu < \infty, 0 < \sigma^2 < \infty]$$

and the supspace Θ determined by the null hypothesis H_0 is given by

$$\Theta_0 = [(\mu, \sigma^2) : \mu = \mu_0, 0 < \sigma^2 < \infty]$$

The likelihood function of the sample observations $x_1, x_2 \dots x_n$ is given by

$$L = \left(\frac{1}{2\pi\sigma^2}\right)^{n/2} \cdot \exp\left[-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2\right]$$

$\hookrightarrow$ ①

The maximum likelihood estimates of μ and σ^2 are given by: (3)

$$\hat{\mu} = \frac{1}{n}\sum_{i=1}^{n} x_i = \bar{x}$$

$$\hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n} (x_i - \bar{x})^2 = s^2 \quad \Bigg\} \rightarrow (2)$$

Hence substituting in (1), the maximum of L in the parameter space Θ is given by

$$L(\hat{\Theta}) = \left[\frac{1}{2\pi s^2}\right]^{n/2} \cdot \exp\left(-n/2\right) \rightarrow (3)$$

In Θ_0, the only variable parameter is σ^2 and MLE of σ^2 for given $\mu = \mu_0$ is given by

$$\hat{\sigma}^2 = \frac{1}{n}\Sigma(x_i - \mu_0)^2 = S_0^2, \text{(say)} \longrightarrow (4)$$

$$= \frac{1}{n}\Sigma(x_i - \bar{x} + \bar{x} - \mu_0)^2$$

$$= \frac{1}{n}\Sigma(x_i - \bar{x})^2 + (\bar{x} - \mu_0)^2,$$

The product term vanishes, since

$$\Sigma(x_i - \bar{x})(\bar{x} - \mu_0) = (\bar{x} - \mu_0)\Sigma(x_i - \bar{x}) = 0$$

$$\therefore \hat{\sigma}^2 = s^2 + (\bar{x} - \mu_0)^2 = S_0^2, \text{(say)} \longrightarrow (4)$$

Hence substituting in (1), we get

$$L(\hat{\Theta}_0) = \left[\frac{1}{2\pi S_0^2}\right]^{n/2} \exp\left(-n/2\right) \rightarrow (5)$$

The ratio of (4) and (3) gives the likelihood Ratio criterion

$$\lambda = \frac{L(\hat{\Theta}_0)}{L(\hat{\Theta})} = \left[\frac{s^2}{S_0^2}\right]^{n/2} \longrightarrow (7)$$

(7)

$$= \left[\frac{s^2}{s^2 + (\bar{x} - \mu_0)^2} \right]^{n/2} = \left\{ \frac{1}{1 + \left[\frac{(\bar{x} - \mu_0)^2}{s^2} \right.} \right\}^{n/2} \rightarrow (8)$$

that under H_0, the statistic

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

where

$$S^2 = \frac{1}{n-1} \sum (x_i - \bar{x})^2 = \frac{ns^2}{n-1},$$

follows Student's t distribution with $(n-1)$ d.f.

Thus

$$t = \frac{\bar{x} - \mu_0}{S/\sqrt{n}} = \frac{\bar{x} - \mu_0}{s/\sqrt{n-1}} \sim t_{n-1} \rightarrow (9)$$

Substituting in (8) we get

$$\lambda = \frac{1}{\left(1 + \frac{t^2}{n-1}\right)^{n/2}} = \phi(t^2), \text{ say} \rightarrow (10)$$

The likelihood Ratio test for testing H_0 against H_1 Consists in finding a critical region of the type $0 < \lambda < \lambda_0$, where λ_0 is given by, (which requires the distribution of λ under H_0. In this Case, it is not neccessary to obtain the distribution of λ since $\lambda = \phi(t^2)$ is a monotonic function of t^2 and the test can well be carried on with t^2 as a criterion as with λ. Now $t^2 = 0$ when $\lambda = 1$ and t^2 becomes infinite when $\lambda = 0$. The critical region of the LR Test viz., $0 < \lambda < \lambda_0$ on using (10) is equivalent to

⑤

$$\left(1+\frac{t^2}{n-1}\right)^{-n/2} \le \lambda_0$$

$$\Rightarrow \left(1+\frac{t^2}{n-1}\right)^{n/2} \ge \lambda_0^{-1} \Rightarrow \frac{t^2}{n-1} \ge (\lambda_0)^{-2/n} - 1$$

$$\Rightarrow t^2 \ge (n-1)\left[\lambda_0^{-2/n} - 1\right] = A^2, \text{ (say)}$$

Thus the critical region may well be defined by

$$|t| = \left|\frac{\sqrt{n}(\bar{x}-\mu_0)}{s}\right| \ge A \longrightarrow (11)$$

Where the constant A is determined such that

$$P\left[|t| \ge A \mid H_0\right] = \alpha \longrightarrow (12)$$

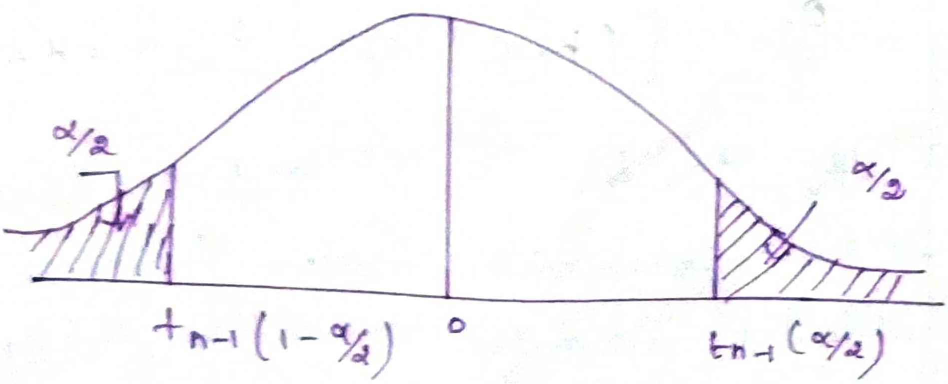
Since under H_0, the statistic t follows Student's t distribution with $(n-1)$ d.f.,

$$A = t_{n-1}(\alpha/2)$$

where the symbol $t_n(\alpha)$ stands for the right tail $100\alpha\%$ point of the t-distribution with n d.f given by

$$P\left[t > t_n(\alpha)\right] = \int_{t_n(\alpha)}^{\infty} f(t)\,dt = \alpha \longrightarrow (13)$$

where $f(\cdot)$ is the p.d.f of Student's with n d.f. The critical region is shown in the following diagram.



Thus for testing $H_0 : \mu = \mu_0$ against $\mu \neq \mu_0$ (σ^2, unknown), we have the two-tailed t-test defined as follows:

If $|t| = \left| \dfrac{\sqrt{n}(\bar{x} - \mu_0)}{S} \right| > t_{n-1}(\alpha/2)$, reject H_0 and if $|t| < t_{n-1}(\alpha/2)$, H_0 may be accepted.

Test for the Equality of means of Two Normal populations :-

Let us consider two independent random variables X_1 and X_2 following normal distribution $N(\mu_1, \sigma_1^2)$ and $N(\mu_2, \sigma_2^2)$ respectively where the mean μ_1, μ_2 and the variances σ_1^2, σ_2^2 are unspecified. Suppose we want to test the hypothesis

$$H_0 : \mu_1 = \mu_2 = \mu, \text{ (say) (unspecified)};$$

$$0 < \sigma_1^2 < \infty, \; 0 < \sigma_2^2 < \infty.$$

against the alternative hypothesis

$$H_1 ; \; \mu_1 \neq \mu_2 \; \sigma_1^2 > 0, \; \sigma_2^2 > 0.$$

Case 1. Population variance are unequal :-

$$\Theta = \{(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2) : -\infty < \mu_i < \infty, \; \sigma_i^2 > 0, \; i = 1, 2\}$$

and

$$\Theta_0 = \{(\mu, \sigma_1^2, \sigma_2^2) : -\infty < \mu < \infty, \; \sigma_i^2 > 0, \; i = 1, 2\}$$

Let x_{1i} $(i = 1, 2 \ldots m)$ and x_{2j} $(j = 1, 2 \ldots n)$ be two independent random samples of size m and n from the populations $N(\mu_1, \sigma_1^2)$ and $N(\mu_2, \sigma_2^2)$ respectively. Then the likelihood

(7) function is given by

$$L = \left(\frac{1}{2\pi\sigma_1^2}\right)^{m/2} \cdot \exp\left[-\frac{1}{2\sigma_1^2}\sum_{i=1}^{m}(x_{1i}-\mu_1)^2\right]$$
$$\times \left(\frac{1}{2\pi\sigma_2^2}\right)^{n/2} \cdot \exp\left[-\frac{1}{2\sigma_2^2}\sum_{j=1}^{n}(x_{2j}-\mu_2)^2\right] \quad \to (1)$$

The maximum likelihood estimates for μ_1, μ_2, σ_1^2 and σ_2^2 are given by the equations:

$$\frac{\partial}{\partial \mu_1}\log L = 0 \Rightarrow \hat{\mu}_1 = \frac{1}{M}\sum_{i=1}^{m} x_{1i} = \bar{x}_1$$

$$\frac{\partial}{\partial \mu_2}\log L = 0 \Rightarrow \hat{\mu}_2 = \frac{1}{n}\sum_{j=1}^{n} x_{2j} = \bar{x}_2$$

$$\frac{\partial}{\partial \sigma_1^2}\log L = 0 \Rightarrow \hat{\sigma}_1^2 = \frac{1}{m}\sum_{i=1}^{m}(x_{1i}-\bar{x}_1)^2 = s_1^2 \text{ (say)}$$

and
$$\frac{\partial}{\partial \sigma_2^2}\log L = 0 \Rightarrow \sigma_2^2 = \frac{1}{n}\sum_{j=1}^{n}(x_{2j}-\bar{x}_2)^2 = s_2^2 \text{ (say)}$$

$\to$ (1a)

Substituting in (1) we get

$$L(\hat{\Theta}) = \left(\frac{1}{2\pi s_1^2}\right)^{m/2} \cdot \left(\frac{1}{2\pi s_2^2}\right)^{n/2} \cdot e^{-\frac{(m+n)}{2}} \quad \to (2)$$

In Θ_0, we have $\mu_1 = \mu_2 = \mu$ and the likelihood function is given by:

$$L(\Theta_0) = \left(\frac{1}{2\pi\sigma_1^2}\right)^{m/2} \exp\left[-\frac{1}{2\sigma_1^2}\sum_{i=1}^{m}(x_{1i}-\mu)^2\right] \times$$
$$\left(\frac{1}{2\pi\sigma_2^2}\right)^{n/2} \exp\left[-\frac{1}{2\sigma_2^2}\sum_{j=1}^{n}(x_{2j}-\mu)^2\right]$$

To obtain the maximum value of $L(\Theta_0)$ for variations in μ, σ_1^2 and σ_2^2, it will be seen that estimate of μ is obtained as the root of a cubic equation

(8)

$$\frac{m^2(\bar{x}_1-\mu)}{\sum_{i=1}^{m}(x_{1i}-\hat{\mu})^2} + \frac{n^2(\bar{x}_2-\mu)}{\sum_{j=1}^{n}(x_{2j}-\hat{\mu})^2} \longrightarrow (3)$$

and is thus a Complicated function of the Sample observations. Consequently the likelihood ratio Criterion λ will be a Complex function of the observations and its distribution is quite tedious since it involves the ratio of two variances. Consequently, it is Impossible to obtain the Critical region $0<\lambda<\lambda_0$, for given α, since the distribution of the population variances is ordinary unknown. However, in any given Instance the Cubic equation (3) can be solved for μ by numerical analysis technique and thus λ can be computed. finally, as an approximate test, $-2\log_e\lambda$ can be regarded as a χ^2-variate with 1 d.f.

Case 2: Population variances are equal:-

$\sigma_1^2=\sigma_2^2=\sigma^2$ (say). In this Case

$$\Theta = \{(\mu_1,\mu_2,\sigma^2) : -\infty<\mu_i<\infty,\ \sigma^2>0,\ (i=1,2)\}$$

$$\Theta_0 = \{(\mu,\sigma^2) : -\infty<\mu<\infty,\ \sigma^2>0\}$$

The likelihood function is then given by

$$L = \left(\frac{1}{2\pi\sigma^2}\right)^{\frac{(m+n)}{2}} \cdot \exp\left[-\frac{1}{2\sigma^2}\left\{\sum_{i=1}^{m}(x_{1i}-\mu_1)^2+\sum_{j=1}^{n}(x_{2j}-\mu_2)^2\right\}\right] \rightarrow (1)$$

(9) for $\mu_1, \mu_2, \sigma^2 \in \Theta$. The maximum likelihood equations are given by

$$\frac{\partial}{\partial \mu_1} \log L = 0 \Rightarrow \hat{\mu}_1 = \bar{x}_1$$

$$\frac{\partial}{\partial \mu_2} \log L = 0 \Rightarrow \hat{\mu}_2 = \bar{x}_2 \quad \Bigg\} \rightarrow (2) \text{ and}$$

$$\frac{\partial}{\partial \sigma^2} \log L = 0 \Rightarrow \hat{\sigma}^2 = \frac{1}{m+n}\left[\Sigma (x_{1i} - \hat{\mu}_1)^2 + \Sigma (x_{2j} - \hat{\mu}_2)^2\right]$$

$$\Rightarrow \hat{\sigma}^2 = \frac{1}{m+n}\left[\Sigma (x_{1i} - \bar{x}_1)^2 + \Sigma (x_{2j} - \bar{x}_2)^2\right]$$

$$\hat{\sigma}^2 = \frac{1}{m+n}\left[m s_1^2 + n s_2^2\right] \longrightarrow (2a)$$

Substituting the values from (2) and (2a) in (1), we get

$$L(\hat{\Theta}) = \left[\frac{(m+n)}{2\pi (m s_1^2 + n s_2^2)}\right]^{\frac{(m+n)}{2}} \cdot \exp\left[-\frac{1}{2} \cdot (m+n)\right] \rightarrow (3)$$

In Θ_0, $\mu_1 = \mu_2 = \mu$ (say), and we get

$$L(\Theta_0) = \left(\frac{1}{2\pi\sigma^2}\right)^{\frac{(m+n)}{2}} \cdot \exp\left[-\frac{1}{2\sigma^2}\left\{\sum_{i=1}^{m} (x_{1i} - \mu)^2 + \sum_{j=1}^{n} (x_{2j} - \mu)^2\right\}\right] \rightarrow (4)$$

$$\Rightarrow \log L(\Theta_0) = C - \frac{m+n}{2} \log \sigma^2 - \frac{1}{2\sigma^2}\left[\sum_i (x_{1i} - \mu)^2 + \sum_j (x_{2j} - \mu)^2\right].$$

Where C is a constant independent of μ and σ^2.

The likelihood equation for estimating μ gives

10

(10) $$\frac{\partial}{\partial \mu} \log L = \frac{1}{\sigma^2}\left[\sum_{i=1}^{m}(x_{1i}-\mu) + \sum_{j=1}^{n}(x_{2j}-\mu)\right] = 0$$ (11)

$$\Rightarrow \quad (m\bar{x}_1 + n\bar{x}_2) - (m+n)\mu = 0$$

$$\Rightarrow \quad \hat{\mu} = \frac{1}{m+n}\left[m\bar{x}_1 + n\bar{x}_2\right] \longrightarrow (5)$$

Also

$$\frac{\partial}{\partial \sigma^2} \log L = 0$$

$$\Rightarrow \quad -\frac{(m+n)}{2\sigma^2} + \frac{1}{2\sigma^4}\left[\Sigma(x_{1i}-\mu)^2 + \Sigma(x_{2j}-\mu)^2\right] = 0$$

$$\Rightarrow \quad \hat{\sigma}^2 = \frac{1}{m+n}\left[\Sigma(x_{1i}-\hat{\mu})^2 + \Sigma(x_{2j}-\hat{\mu})^2\right] \rightarrow (6)$$

But $$\sum_{i=1}^{m}(x_i-\hat{\mu})^2 = \sum_{i=1}^{m}(x_{1i}-\bar{x}_1+\bar{x}_1-\hat{\mu})^2.$$

$$= \Sigma(x_{1i}-\bar{x}_1)^2 + m(\bar{x}_1-\hat{\mu})^2,$$

The product term vanishes since

$$\sum_{i}(x_{1i}-\bar{x}_1) = 0$$

$$\therefore \quad \sum_{i=1}^{m}(x_{1i}-\hat{\mu})^2 = m s_1^2 + m\left[\bar{x}_1 - \frac{m\bar{x}_1+n\bar{x}_2}{m+n}\right]^2$$

$$= m s_1^2 + \frac{mn^2(\bar{x}_1-\bar{x}_2)^2}{(m+n)^2}$$

Similarly, we shall get

$$\sum_{j=1}^{n}(x_{2j}-\hat{\mu})^2 = n s_2^2 + \frac{nm^2(\bar{x}_2-\bar{x}_1)^2}{(m+n)^2}$$

Substituting in (6), we get

$$\hat{\sigma}^2 = \frac{1}{m+n}\left[m s_1^2 + n s_2^2 + \frac{mn}{m+n}(\bar{x}_1-\bar{x}_2)^2\right] \rightarrow (6a)$$

Substituting from (5) and (6a) in (4), we get

$$L(\hat{\Theta}_0) = \left\{ \frac{(m+n)}{2\pi\left(ms_1^2 + ns_2^2 + \frac{mn}{m+n}(\bar{x}_1 - \bar{x}_2)^2\right)} \right\}^{\frac{(m+n)}{2}}$$

$$\times \exp\left(-\frac{m+n}{2}\right) \rightarrow (7)$$

$$\therefore \lambda = \frac{L(\hat{\Theta}_0)}{L(\hat{\Theta})} = \left\{ \frac{ms_1^2 + ns_2^2}{ms_1^2 + ns_2^2 + \frac{mn}{m+n}(\bar{x}_1 - \bar{x}_2)^2} \right\}^{\frac{m+n}{2}}$$

$$\lambda = \left\{ \frac{1}{\left(1 + \frac{mn\,(\bar{x}_1 - \bar{x}_2)^2}{(m+n)(ms_1^2 + ns_2^2)}\right)} \right\}^{\frac{m+n}{2}} \rightarrow (8)$$

We know that, under the null hypothesis $H_0 : \mu_1 = \mu_2$, the statistic

$$t = \frac{\bar{x}_1 - \bar{x}_2}{S\sqrt{\frac{1}{m} + \frac{1}{n}}} \rightarrow (9)$$

where

$$S^2 = \frac{1}{m+n-2}(ms_1^2 + ns_2^2) \rightarrow (9a)$$

follows Student's t distribution with $(m+n-2)$ d.f. Thus in terms of t, we get

$$\lambda = \left[1 + \frac{t^2}{m+n-2}\right]^{-\frac{(m+n)}{2}} \rightarrow (10)$$

The test can as well be carried with t rather than with λ. The critical region $0 < \lambda < \lambda_0$, transforms to the critical region of the type.

$$t^2 > (m+n-2)\left[\frac{1}{\lambda_0^{2}/(m+n)} - 1\right] = A^2 \text{ (say)}$$

i.e., by

$$|t| > A.$$

Where A is determined so that

12

$$P\left[\,|t| > A \mid H_0\,\right] = \alpha \rightarrow (11)$$

13

since under H_0, the statistic t follows Student's t distribution with $(m+n-2)$ d.f., we get (11)

$$A = t_{m+n-2}(\alpha/2) \rightarrow (12)$$

where, $t_n(\alpha)$ is the right $100\,\alpha\%$ point of the t-distribution with n d.f.

Thus for testing the null hypothesis

$$H_0 : \mu_1 = \mu_2 \; ; \quad \sigma_1^2 = \sigma_2^2 = \sigma^2 > 0$$

against the alternative

$$H_1 ; \quad \mu_1 \neq \mu_2 , \; \sigma_1^2 = \sigma_2^2 = \sigma^2 > 0 .$$

we have the two-tailed t test defined as follows: if

$$|t| = \left| \frac{\bar{x}_1 - \bar{x}_2}{s\sqrt{\frac{1}{m} + \frac{1}{n}}} \right| > t_{m+n-2}(\alpha/2)$$

reject H_0, otherwise H_0 may be accepted.

Test for the variance of a Normal population :-

Let us now consider the problem of testing if the variance of a normal population has a specified value σ_0^2, on the basis of a random sample $x_1, x_2 \ldots x_n$ of size n from normal population $N(\mu, \sigma^2)$. We want to test the hypothesis

$$H_0 : \sigma^2 = \sigma_0^2 \text{ (specified)},$$

(12) against the alternative hypothesis

$H_1 ; \sigma^2 \neq \sigma_0^2$

Here we have

$$\Theta = \{ (\mu, \sigma^2) : -\infty < \mu < \infty, \sigma^2 > 0 \}$$

and

$$\Theta_0 = \{ (\mu, \sigma^2) : -\infty < \mu < \infty, \sigma^2 = \sigma_0^2 \}$$

The likelihood function of the sample observations is given by

$$L = \left(\frac{1}{2\pi\sigma^2} \right)^{n/2} \cdot \exp\left\{ -\frac{1}{2\sigma^2} \sum_{i=1}^{n} (x_i - \mu)^2 \right\} \to ①$$

we shall get

$$L(\hat{\Theta}) = \left(\frac{1}{2\pi s^2} \right)^{n/2} \exp\left(-n/2\right) \to ②$$

In Θ_0, we have only one variable parameter, viz., μ and

$$L(\Theta_0) = \left(\frac{1}{2\pi\sigma_0^2} \right)^{n/2} \exp\left[-\frac{1}{2\sigma_0^2} \sum_{i=1}^{n} (x_i - \mu)^2 \right] \to ③$$

The MLE for μ is given by

$$\frac{\partial}{\partial \mu} \log L = 0 \Rightarrow \hat{\mu} = \bar{x}$$

$$\therefore L(\hat{\Theta}_0) = \left(\frac{1}{2\pi\sigma_0^2} \right)^{n/2} \exp\left[-\frac{1}{2\sigma_0^2} \sum_{i=1}^{n} (x_i - \bar{x})^2 \right]$$

$$= \left(\frac{1}{2\pi\sigma_0^2} \right)^{n/2} \exp\left[-\frac{ns^2}{2\sigma_0^2} \right] \to ④$$

The likelihood ratio criterion is given by

$$\lambda = \frac{L(\hat{\Theta}_0)}{L(\hat{\Theta})} = \left[\frac{s^2}{\sigma_0^2} \right]^{n/2} \exp\left[-\frac{1}{2} \left(\frac{ns^2}{\sigma_0^2} - n \right) \right]$$

We know that under H_0, the statistic

$$\chi^2 = \frac{ns^2}{\sigma_0^2} \rightarrow (5)$$

follows Chi-square distribution with $(n-1)$ d.f.

In terms of χ^2, we have

$$\lambda = \left[\frac{\chi^2}{n}\right]^{n/2} \cdot \exp\left[-\frac{1}{2}(\chi^2 - n)\right] \rightarrow (6)$$

Since λ is a monotonic function of χ^2, the test may be done using χ^2 as a criterion. The critical region $0 < \lambda < \lambda_0$ is now equivalent to

$$(\chi^2/n)^{n/2} \exp\left[-\frac{1}{2}(\chi^2 - n)\right] < \lambda_0$$

or

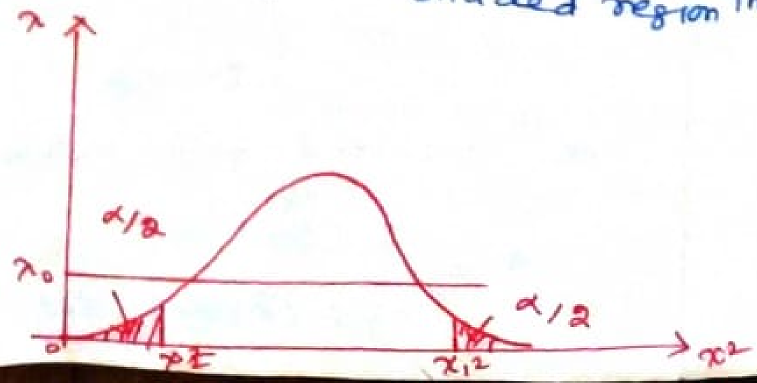
$$\exp\left(-\tfrac{1}{2}\chi^2\right)(\chi^2)^{n/2} < \lambda_0 \cdot (ne^{-1})^{n/2} = B \text{ (say)}$$

$\rightarrow (7)$

Since χ^2 has Chi-square distribution with $(n-1)$ d.f., The critical region (7) is determined by a pair of Intervals $0 < \chi^2 < \chi_2^2$ and $\chi_1^2 < \chi^2 < \infty$, where χ_1^2 and χ_2^2 are to be determined such that the ordinates of (6) are equal i.e.,

$$(\chi_1^2)^{n/2} \exp\left(-\tfrac{1}{2}\chi_1^2\right) \Rightarrow (\chi_2^2)^{n/2} \exp\left(\tfrac{-1}{2}\chi_2^2\right)$$

Critical region is shown as shaded region in the above diagram



In other words, x_1^2 and x_2^2 are defined by the equations

$$P(x^2 > x_1^2) = \alpha/2$$

and $$\left.P(x^2 > x_2^2) = 1 - \alpha/2\right\} \rightarrow (7)$$

and In other words,

$$x_1^2 = x^2_{n-1}(\alpha/2) \text{ and } x_2^2 = x^2_{n-1}(1-\alpha/2),$$

where $x^2_{(n-1)}(\alpha)$ is the upper α-point of the chi-square distribution with $(n-1)$ d.f. thus the critical region for testing H_0; $\sigma^2 = \sigma_0^2$ against H_1; $\sigma^2 \neq \sigma_0^2$, is a two-tailed region given by

$$x^2 > x^2_{n-1}(\alpha/2) \text{ and } x^2 < x^2_{n-1}(1-\alpha/2) \rightarrow (8)$$

Thus, in this case we have a two-tailed test.

Test for Equality of variance of two normal populations :- Consider two normal population $N(\mu_1, \sigma_1^2)$ and $N(\mu_2, \sigma_2^2)$ where the mean μ_1 and μ_2 and variances σ_1^2, σ_2^2 are unspecified. we want to test the hypothesis.

H_0 ; $\sigma_1^2 = \sigma_2^2 = \sigma^2$ (unspecified), with μ_1 and μ_2 (unspecified) against the alternative hypothesis.

H_1 ; $\sigma_1^2 \neq \sigma_2^2$; $\cancel{\mu_1 \text{ and } \mu_2}$ μ_1 and μ_2 (unspecified).

If x_{1i}, $(i = 1, 2 \ldots m)$ and x_{2j} $(j = 1, 2 \ldots n)$ be Independent random samples of sizes m and n from $N(\mu_1, \sigma_1^2)$ and $N(\mu_2, \sigma_2^2)$

respectively then

$$L = \left(\frac{1}{2\pi\sigma_1^2}\right)^{m/2} \exp\left[-\frac{1}{2\sigma_1^2}\sum_{i=1}^{m}(x_{1i}-\mu_1)^2\right] \times$$ (17)

$$\left(\frac{1}{2\pi\sigma_2^2}\right)^{n/2} \exp\left[-\frac{1}{2\sigma_2^2}\sum_{j=1}^{n}(x_{2j}-\mu_2)^2\right]^2 \rightarrow (1)$$

In this case

$$\Theta = [(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2) : -\infty < \mu_i < \infty;\ \sigma_i^2 > 0, (i=1,2)]$$ and

$$\Theta_0 = \{(\mu_1, \mu_2, \sigma^2) : -\infty < \mu_i < \infty;\ (i=1,2), \sigma^2 > 0$$

As

$$L(\hat{\Theta}) = \left(\frac{1}{2\pi s_1^2}\right)^{m/2} \cdot \left(\frac{1}{2\pi s_2^2}\right)^{n/2} \cdot \exp\left[-\frac{1}{2}(m+n)\right] \longrightarrow (2)$$

where s_1^2 and s_2^2 are as defined

In Θ_0, The likelihood function (1) is given by

$$L(\Theta_0) = \left[\frac{1}{2\pi\sigma^2}\right]^{\frac{(m+n)}{2}} \cdot \exp\left[-\frac{1}{2\sigma^2}\left\{\sum_i (x_{1i}-\mu_1)^2 + \sum_j (x_{2j}-\mu_2)^2\right\}\right] \rightarrow (3)$$

and the MLE's for μ_1 and μ_2 and σ^2 are now given by

$$\hat{\mu_1} = \bar{x}_1, \quad \hat{\mu_2} = \bar{x}_2 \rightarrow (4)$$ and

$$\hat{\sigma}^2 = \frac{1}{(m+n)}\left[\sum_i (x_{1i}-\hat{\mu_1})^2 + \sum_j (x_{2j}-\hat{\mu_2})^2\right]$$

$$\hat{\sigma}^2 = \frac{1}{m+n}\left[\sum_i (x_{1i}-\bar{x}_1)^2 + \sum_j (x_{2j}-\bar{x}_2)^2\right]$$

7

$$\hat{\sigma}^2 = \frac{ms_1^2 + ns_2^2}{m+n} \longrightarrow (4a)$$

Substituting from (4) and (4a) in (3), we get

$$L(\hat{\theta}_0) = \left[\frac{m+n}{2\pi(ms_1^2+ns_2^2)}\right]^{\frac{(m+n)}{2}} \cdot \exp\left[-\frac{1}{2}(m+n)\right]$$

↳ (5)

$$\therefore \lambda = \frac{L(\hat{\theta}_0)}{L(\hat{\theta})}$$

$$= (m+n)^{\frac{(m+n)}{2}} \left\{\frac{(s_1^2)^{m/2}(s_2^2)^{n/2}}{\{ms_1^2+ns_2^2\}^{m+n/2}}\right\}$$

$$= \frac{(m+n)^{\frac{m+n}{2}}}{m^{m/2}\cdot n^{n/2}} \left\{\frac{(ms_1^2)^{m/2}(ns_2^2)^{n/2}}{[ms_1^2+ns_2^2]^{\frac{m+n}{2}}}\right\}$$

↳ (6)

We know that under H_0 the statistic

$$F = \frac{\Sigma(x_{1i}-\bar{x}_1)^2/m-1}{\Sigma(x_{2j}-\bar{x}_2)^2/n-1} = \frac{S_1^2}{S_2^2} \longrightarrow (7)$$

follows F-distribution with $(m-1, n-1)$ d.f also

Implies

$$F = \frac{m(n-1)\,s_1^2}{n(m-1)\,s_2^2}$$

$$\Rightarrow \left(\frac{m-1}{n-1}\right) F = \frac{ms_1^2}{ns_2^2} \longrightarrow (7a)$$

Substituting in (6) and simplifying, we get

$$\lambda = \frac{(m+n)^{\frac{(m+n)}{2}}}{m^{m/2}\, n^{n/2}} \left\{\frac{\left(\frac{m-1}{n-1}F\right)^{m/2}}{\left[1+\frac{m-1}{n-1}F\right]^{\frac{(m+n)}{2}}}\right\}$$

↳ (8)

(18) Thus λ is a monotonic function of F and hence the test can be carried on with F, defined (7) as test statistic. The critical region $0 < \lambda < \lambda_0$ can be equivalently seen to be given by pairs intervals $F \leq F_1$ and $F \geq F_2$, where F_1 and F_2 are determined so that under H_0.

$$P(F \geq F_2) = \alpha/2 \text{ and}$$

$$P(F \geq F_1) = 1 - \alpha/2$$

Since, under H_0, F follows Snedecor's F-distribution with $(m-1, n-1)$ d.f., we have

$$F_2 = F_{m-1, n-1}(\alpha/2) \text{ and}$$

$$F_1 = F_{m-1, n-1}(1-\alpha/2).$$

where $F_{m,n}(\alpha)$ is the upper α-point of F-distribution with (m, n) d.f. Consequently for testing $H_0: \sigma_1^2 = \sigma_2^2$ against the alternative hypothesis $H_1: \sigma_1^2 \neq \sigma_2^2$, we have a two-tailed F-test, the critical region being given by

$$F > F_{m-1, n-1}(\alpha/2) \text{ and } F < F_{m-1, n-1}(1-\alpha/2) \rightarrow (9)$$

Where F is defined (7) or (7a)