B.Tech II Semester (2020 Batch)
PROBABILITY AND STATISTICS (20BM1104)
(For CSE-3 & CSE-4)
Unit – 4: Estimation and Test of Hypothesis of Variances and Proportions

(*Estimation of variance, hypothesis concerning one variance, hypothesis concerning two variances, estimation of proportion, hypothesis concerning one proportion, hypothesis concerning several proportions*)

Proportion:

Let p, q be the success and failure probabilities of an event of a trial. Let the trial be conducted in any number of times. Then the collection of all successes and failures of the event is a population known as Binomial population. For this population, p is called **Proportion or True proportion**. If x is the number of successes in n trials, then $\frac{x}{n}$ is called the **Sample Proportion** and it is denoted by P; that is, $P = \frac{x}{n}$

Example:

Let a coin be tossed 10 times and 'getting head H' be the event. Suppose that the outcomes in these 10 tosses are H,H,H,T,T,H,T,H,H,T respectively. Now the collection of all the successes and failures of the event is a population; that is, $Population = \{S,S,S,F,F,S,F,S,S,F\}$.

For this population, **Proportion** $p = \frac{1}{2}$

(i) If we collect 1st five outcomes $\{H,H,H,F,F\}$, then it is a sample of size $n = 5$

For this sample, sample proportion $P = \frac{x}{n} = \frac{3}{5}$

(ii) If we collect 1st two outcomes $\{H,H\}$, then it is a sample of size $n = 2$

For this sample, sample proportion $P = \frac{x}{n} = \frac{2}{2} = 1$

(iii) If we collect 1st four outcomes $\{H,H,H,F\}$, then it is a sample of size $n = 4$

For this sample, sample proportion $P = \frac{x}{n} = \frac{3}{4}$

(iv) If we collect last two outcomes $\{H,F\}$, then it is a sample of size $n = 2$

For this sample, sample proportion $P = \frac{x}{n} = \frac{1}{2}$

Maximum Error of the Proportion:

(i) $E = z_{\frac{\alpha}{2}} \times \sqrt{\frac{\frac{x}{n}\left(1-\frac{x}{n}\right)}{n}}$ is the maximum error of the Proportion with the probability $1-\alpha$

(ii) If we know the value of p, then the sample size is given by $n = p\,(1-p)\left[\frac{z_{\frac{\alpha}{2}}}{E}\right]^2$

(iii) If we do not know the value of p, then the sample size is given by $n = \frac{1}{4}\left[\frac{z_{\frac{\alpha}{2}}}{E}\right]^2$

Confidence Interval – One Proportion:

With the probability $1-\alpha$ or $(1-\alpha)100\%$ confidence,

Upper and Lower confidence limits: $\dfrac{x}{n} \pm E$

Confidence Interval: $\left(\dfrac{x}{n}-E,\ \dfrac{x}{n}+E\right)$ where, $E = z_{\frac{\alpha}{2}} \times \sqrt{\dfrac{\frac{x}{n}\left(1-\frac{x}{n}\right)}{n}}$

Examples:

(1) Among 900 people in a state 90 are found to be rice eaters. Construct 99% confidence interval for the true proportion p

Here $n = 900$, $x = 90$ and $\dfrac{x}{n} = \dfrac{90}{900} = 0.1$

Also, confidence $= 99\%$, probability $1-\alpha = 0.99$ and so $\alpha = 0.01$

Now, $\dfrac{\alpha}{2} = 0.005$ and $z_{\frac{\alpha}{2}} = z_{0.005} = 2.575$

The Maximum error of the proportion, $E = z_{\frac{\alpha}{2}} \times \sqrt{\dfrac{\frac{x}{n}\left(1-\frac{x}{n}\right)}{n}}$

$$= (2.575)\sqrt{\frac{(0.1)(1-0.1)}{900}} = 0.02575$$

Lower confidence limit, $\dfrac{x}{n} - E = 0.1 - 0.02575 = 0.07425$

Upper confidence limit, $\dfrac{x}{n} + E = 0.1 + 0.02575 = 0.12575$

Therefore, confidence interval $\left(\dfrac{x}{n}-E,\ \dfrac{x}{n}+E\right) = (0.07425,\ 0.12575)$

(2) In a random sample of 400 industrial accidents, it was found that 231 were due at least partially to unsafe working conditions. Construct 99% confidence interval for the true proportion using the large sample confidence interval formula

Here $n = 400$, $x = 231$ and $\dfrac{x}{n} = \dfrac{231}{400} = 0.5775$

Also, confidence $= 99\%$, probability $1-\alpha = 0.99$ and so $\alpha = 0.01$

Now, $\dfrac{\alpha}{2} = 0.005$ and $z_{\frac{\alpha}{2}} = z_{0.005} = 2.575$

The Maximum error of the proportion, $E = z_{\frac{\alpha}{2}} \times \sqrt{\dfrac{\frac{x}{n}\left(1-\frac{x}{n}\right)}{n}}$

$$= (2.575)\sqrt{\frac{(0.5775)(1-0.5775)}{400}} = 0.0636$$

Lower confidence limit, $\frac{x}{n} - E = 0.5775 - 0.0636 = 0.5139$

Upper confidence limit, $\frac{x}{n} + E = 0.5775 + 0.0636 = 0.6411$

Therefore, confidence interval $\left(\frac{x}{n} - E,\ \frac{x}{n} + E\right) = (0.5139,\ 0.6411)$

(3) If $x = 36$ of $n = 100$ persons interviewed are familiar with the tax incentives for installing certain energy saving devices construct a 95% confidence interval for the corresponding true proportion

Here $n = 100$, $x = 36$ and $\frac{x}{n} = \frac{36}{100} = 0.36$

Also, confidence $= 95\%$, probability $1 - \alpha = 0.95$ and so $\alpha = 0.05$

Now, $\frac{\alpha}{2} = 0.025$ and $z_{\frac{\alpha}{2}} = z_{0.025} = 1.96$

The Maximum error of the proportion, $E = z_{\frac{\alpha}{2}} \times \sqrt{\frac{\frac{x}{n}\left(1 - \frac{x}{n}\right)}{n}}$

$$= (1.96)\sqrt{\frac{(0.36)(1-0.36)}{100}} = 0.0941$$

Lower confidence limit, $\frac{x}{n} - E = 0.36 - 0.0941 = 0.2659$

Upper confidence limit, $\frac{x}{n} + E = 0.36 + 0.0941 = 0.4541$

Therefore, confidence interval $\left(\frac{x}{n} - E,\ \frac{x}{n} + E\right) = (0.2659,\ 0.4541)$

(4) Find the sample size if the true proportion does not exceed 0.12 to estimate the true Proportion of defective items with at least 95% confidence with error 0.04.

Here the maximum error of the proportion, $E = 0.04$ and $p = 0.12$

Also, confidence $= 95\%$, probability $1 - \alpha = 0.95$ and so $\alpha = 0.05$

Now, $\frac{\alpha}{2} = 0.025$ and $z_{\frac{\alpha}{2}} = z_{0.025} = 1.96$

For known proportion, the sample size $n = p\,(1-p)\left[\frac{z_{\frac{\alpha}{2}}}{E}\right]^2$

$$= (0.12)(1-0.12)\left[\frac{1.96}{0.04}\right]^2 = 253.55 \approx 254$$

(5) What is the size of the smallest sample required to estimate an unknown proportion to within a maximum error of 0.06 with at least 95% confidence

Here the maximum error of the proportion, $E = 0.06$
Also, confidence $= 95\%$, probability $1-\alpha = 0.95$ and so $\alpha = 0.05$

Now, $\frac{\alpha}{2} = 0.025$ and $z_{\frac{\alpha}{2}} = z_{0.025} = 1.96$

For unknown proportion, the sample size $n = \frac{1}{4}\left[\frac{z_{\frac{\alpha}{2}}}{E}\right]^2 = \frac{1}{4}\left[\frac{1.96}{0.06}\right]^2 = 266.7 \approx 267$

Exercise:

(1) In a sample survey conducted in a large city, 136 of 400 persons answered 'yes' to the question of whether their cities public transportation is adequate. With 99% confidence, what can be say about the maximum error, if $\frac{x}{n} = \frac{136}{400} = 0.34$ is used as an estimate of the corresponding true proportion

(2) If sample size $n = 400$ and proportion to success $p = 0.578$ then construct 98% confidence interval for the proportion p

(3) Among 900 people in a state 90 are found to be blind. Construct 98% confidence interval for the true proportion

(4) A random sample of 500 apples was taken from, a large consignment and 60 were found to be bad. Obtain the 98% confidence limits for the percentage number of bad apples in the consignment

(5) What is the size of the smallest sample required to estimate an unknown proportion to within a maximum error of 0.08 with at least 98% confidence

(6) Find the sample size if the true proportion does not exceed 0.2 to estimate the true Proportion of defective items with at least 95% confidence with error 0.05

Test of Hypothesis – One Proportion:

(1) **Null Hypothesis** H_0 **:** $p = p_0$

(2) **Alternative Hypothesis** H_1 **:** Any one of the following $p < p_0$, $p > p_0$, $p \neq p_0$

(3) **Level of Significance** **:** α

(4) **Test statistic** **:** $Z = \frac{x - np}{\sqrt{np(1-p)}}$ or $Z = \frac{\frac{x}{n} - p}{\sqrt{\frac{p(1-p)}{n}}}$

(5) **Criterion** **:**

H_1	Reject H_0 if
$p < p_0$	$Z < -z_\alpha$
$p > p_0$	$Z > z_\alpha$
$p \neq p_0$	$\lvert Z \rvert > z_{\frac{\alpha}{2}}$

Examples:

(1) An airline claims that only 6% of all lost luggage is never found. If, in a random sample, 17 of 200 pieces of lost luggage are not found, test the null hypothesis $p = 0.06$ against the alternate hypothesis $p > 0.06$ at 0.05 level of significance

Here $n = 200,\ x = 17,\ p = 0.06$ and $\alpha = 0.05$

(i) Null Hypothesis H_0 : $p = 0.06$

(ii) Alternative Hypothesis H_1 : $p > 0.06$

(iii) Level of Significance : $\alpha = 0.05$

(iv) Test statistic : $Z = \dfrac{x - n\,p}{\sqrt{n\,p\,(1-p)}}$

(v) Criterion : Reject H_0 if $Z > z_\alpha$ and $z_\alpha = z_{0.05} = 1.645$

(vi) Calculation :

$$Z = \frac{x - n\,p}{\sqrt{n\,p\,(1-p)}} = \frac{17 - (200)(0.06)}{\sqrt{(200)(0.06)(1-0.06)}} = 1.4887$$

And $z_\alpha = z_{0.05} = 1.645$

(vii) Decision: Since $Z < z_\alpha$, accept H_0 based on the sample data at $\alpha = 0.05$

(2) In a study designed to investigate whether certain detonators used with explosives in coal mining meet the requirement that at least 90% will ignite the explosive when charged, it is found that 174 of 200 detonators function properly. Test the null hypothesis $p = 0.9$ against the alternative hypothesis $p < 0.9$ at the 0.05 level of significance

Here $n = 200,\ x = 174,\ p = 0.9$ and $\alpha = 0.05$

(i) Null Hypothesis H_0 : $p = 0.9$

(ii) Alternative Hypothesis H_1 : $p < 0.9$

(iii) Level of Significance : $\alpha = 0.05$

(iv) Test statistic : $Z = \dfrac{x - n\,p}{\sqrt{n\,p\,(1-p)}}$

(v) Criterion : Reject H_0 if $Z < -z_\alpha$ and $-z_\alpha = -z_{0.05} = -1.645$

(vi) Calculation :

$$Z = \frac{x - n\,p}{\sqrt{n\,p\,(1-p)}} = \frac{174 - (200)(0.9)}{\sqrt{(200)(0.9)(1-0.9)}} = -1.4142$$

And $-z_\alpha = -z_{0.05} = -1.645$

(vii) Decision: Since $Z > -z_\alpha$, accept H_0 based on the sample data at $\alpha = 0.05$

(3) In a big city 325 men out of 600 men were found to be smokers. Does this information support the conclusion that the majority of men in this city are smokers

Here $n = 600,\ x = 325$ and $\alpha = 0.05$ (assumed)
We have the check the Hypothesis $p > 0.5$

(i) Null Hypothesis H_0 : $p = 0.5$
(ii) Alternative Hypothesis H_1 : $p > 0.5$
(iii) Level of Significance : $\alpha = 0.05$
(iv) Test statistic : $Z = \dfrac{x - n\,p}{\sqrt{n\,p\,(1-p)}}$
(v) Criterion : Reject H_0 if $Z > z_\alpha$ and $z_\alpha = z_{0.05} = 1.645$
(vi) Calculation :

$$Z = \frac{x - n\,p}{\sqrt{n\,p\,(1-p)}} = \frac{325 - (600)(0.5)}{\sqrt{(600)(0.5)(1-0.5)}} = 2.0412$$

And $z_\alpha = z_{0.05} = 1.645$

(vii) Decision: Since $Z > z_\alpha$, reject H_0 based on the sample data at $\alpha = 0.05$
That is, the majority of men in this city are smokers

(4) In a large consignment of oranges, a random sample of 64 oranges revealed that 14 oranges were bad. Is it reasonable to assume that 20% of oranges were bad at 5% level of significance?

Here $n = 64,\ x = 14$ and $\alpha = 0.05$

We have to check the hypothesis $p = \dfrac{20}{100} = 0.2$

(i) Null Hypothesis H_0 : $p = 0.2$
(ii) Alternative Hypothesis H_1 : $p \neq 0.2$
(iii) Level of Significance : $\alpha = 0.05$
(iv) Test statistic : $Z = \dfrac{x - n\,p}{\sqrt{n\,p\,(1-p)}}$
(v) Criterion : Reject H_0 if $|Z| > z_{\frac{\alpha}{2}}$ and $z_{\frac{\alpha}{2}} = z_{0.025} = 1.96$
(vi) Calculation :

$$Z = \frac{x - n\,p}{\sqrt{n\,p\,(1-p)}} = \frac{14 - (64)(0.2)}{\sqrt{(64)(0.2)(1-0.2)}} = 0.375, \quad |Z| = 0.375$$

And $z_{\frac{\alpha}{2}} = z_{0.025} = 1.96$

(vii) Decision: Since $|Z| < z_{\frac{\alpha}{2}}$, accept H_0 based on the sample data at $\alpha = 0.05$

(5) A coin is tossed 960 times and head turns up 183 times. Is the coin biased? Use the 0.05 level of significance

Here $n = 960,\ x = 183$ and $\alpha = 0.05$
If a coin is unbiased, then the probability of getting a Head in a single toss, $p = 0.5$
So we have to check the hypothesis $p = 0.5$

(i) Null Hypothesis H_0 : $p = 0.5$

(ii) Alternative Hypothesis H_1 : $p \neq 0.5$

(iii) Level of Significance : $\alpha = 0.05$

(iv) Test statistic : $Z = \dfrac{x - n\,p}{\sqrt{n\,p\,(1-p)}}$

(v) Criterion : Reject H_0 if $|Z| > z_{\frac{\alpha}{2}}$ and $z_{\frac{\alpha}{2}} = z_{0.025} = 1.96$

(vi) Calculation :

$$Z = \frac{x - n\,p}{\sqrt{n\,p\,(1-p)}} = \frac{183-(960)(0.5)}{\sqrt{(960)(0.5)(1-0.5)}} = -19.1713, \quad |Z| = 19.1713$$

And $z_{\frac{\alpha}{2}} = z_{0.025} = 1.96$

(vii) Decision: Since $|Z| > z_{\frac{\alpha}{2}}$, reject H_0 based on the sample data at $\alpha = 0.05$

That is, the coin is biased

Exercise:

(1) Write the test statistic and criteria for the testing a hypothesis concerning one proportion and two proportions

(2) A manufacturer of submersible pumps claims that at most 30% of the pumps require repairs within the first five years of operation. If a random sample of 120 of these pumps includes 47 which required repairs within the first five years of operation, test the null hypothesis $p = 0.3$ against the alternative hypothesis $p > 0.3$ at the 0.05 level of significance

(3) An ambulance service's claim that at least 40% of its calls are life-threatening emergencies, a random sample was taken from its files, and it was found that only 49 of 150 calls were life-threatening emergencies. Can the null hypothesis $p = 0.4$ be rejected against the alternative hypothesis $p < 0.4$ at the 0.01 level of significance

(4) In a random sample of 600 cars making a right turn at a certain intersection, 157 pulled into the wrong lane. Test the null hypothesis that actually 30% of all drivers make this mistake at the given intersection, using the alternative hypothesis $p \neq 0.3$ and the level of significance 0.01

(5) In a random sample of 160 workers exposed to a certain amount of radiation, 24 experienced some ill effects. Test the null hypothesis $p = 0.18$ versus the alternative hypothesis $p \neq 0.18$ at the 0.01 level

(6) A coin is tossed 960 times and head turns up 450 times. Is the coin biased? Use the 0.05 level of significance $\left(\text{Hint } n = 960,\ x = 450,\ H_0 : p = 0.5,\ H_1 : p \neq 0.5\right)$

(7) A coin is tossed 400 times and head turns up 216 times. Is the coin biased? Use the 0.05 level of significance $\left(\text{Hint } n = 400,\ x = 216,\ H_0 : p = 0.5,\ H_1 : p \neq 0.5\right)$

(8) A die is thrown 900 times and it falls with 5 upwards 185 times. Is the die biased? Use the 0.01 level of significance $\left(\text{Hint } n = 900,\ x = 185,\ H_0 : p = \dfrac{1}{6},\ H_1 : p \neq \dfrac{1}{6}\right)$

(9) A marketing expert for a pasta-making company believes that 40% of pasta lovers prefer lasagna. If 9 out of 20 pasta lovers choose lasagna over other pastas, what can be concluded about the expert's claim? Use a 0.05 level of significance.

Test of Hypothesis – Two Proportions:

(1) Null Hypothesis H_0 **:** $p_1 - p_2 = 0$

(2) Alternative Hypothesis H_1 **:** Any one of the following $p_1 - p_2 < 0$, $p_1 - p_2 > 0$, $p_1 - p_2 \neq 0$

(3) Level of Significance **:** α

(4) Test statistic **:** $Z = \dfrac{\frac{x_1}{n_1} - \frac{x_2}{n_2}}{\sqrt{\frac{\hat{p}(1-\hat{p})}{n_1} + \frac{\hat{p}(1-\hat{p})}{n_2}}}$ where, $\hat{p} = \dfrac{x_1 + x_2}{n_1 + n_2}$

(5) Criterion **:**

H_1	Reject H_0 if
$p_1 - p_2 < 0$	$Z < -z_\alpha$
$p_1 - p_2 > 0$	$Z > z_\alpha$
$p_1 - p_2 \neq 0$	$\lvert Z \rvert > z_{\frac{\alpha}{2}}$

Note: $\hat{p} = \dfrac{x_1 + x_2}{n_1 + n_2}$ is known as the **proportion by pooling**

Confidence Interval - Two Proportions:

With the probability $1-\alpha$ or $(1-\alpha)100\%$ confidence,

Upper and Lower confidence limits: $\left(\dfrac{x_1}{n_1} - \dfrac{x_2}{n_2}\right) \pm E$

Confidence Interval: $\left(\left(\dfrac{x_1}{n_1} - \dfrac{x_2}{n_2}\right) - E,\ \left(\dfrac{x_1}{n_1} - \dfrac{x_2}{n_2}\right) + E\right)$

Where, $E = z_{\frac{\alpha}{2}} \times \sqrt{\dfrac{\hat{p}(1-\hat{p})}{n_1} + \dfrac{\hat{p}(1-\hat{p})}{n_2}}$ and $\hat{p} = \dfrac{x_1 + x_2}{n_1 + n_2}$

Examples:

(1) A manufacturer of electronic equipment subject's samples of two competing brands of transistors to an accelerated performance test. If 45 of 180 transistors of the first kind and 34 of 120 transistors of the second kind fail the test, what can he conclude at the level of significance $\alpha = 0.05$ about the difference between the corresponding sample proportions

Here $n_1 = 180$, $x_1 = 45, n_2 = 120$, $x_2 = 34$ and $\alpha = 0.05$

We have to check the hypothesis $p_1 - p_2 \neq 0$

(i) Null Hypothesis H_0 : $p_1 - p_2 = 0$

(ii) Alternative Hypothesis H_1 : $p_1 - p_2 \neq 0$

(iii) Level of Significance : $\alpha = 0.05$

(iv) Test statistic : $Z = \dfrac{\dfrac{x_1}{n_1} - \dfrac{x_2}{n_2}}{\sqrt{\dfrac{\hat{p}\,(1-\hat{p})}{n_1} + \dfrac{\hat{p}\,(1-\hat{p})}{n_2}}}$ where, $\hat{p} = \dfrac{x_1 + x_2}{n_1 + n_2}$

(v) Criterion : Reject H_0 if $|Z| > z_{\frac{\alpha}{2}}$ and $z_{\frac{\alpha}{2}} = z_{0.025} = 1.96$

(vi) Calculation :

$$\hat{p} = \frac{x_1 + x_2}{n_1 + n_2} = \frac{45 + 34}{180 + 120} = \frac{79}{300} = 0.2633$$

$$Z = \frac{\dfrac{x_1}{n_1} - \dfrac{x_2}{n_2}}{\sqrt{\dfrac{\hat{p}\,(1-\hat{p})}{n_1} + \dfrac{\hat{p}\,(1-\hat{p})}{n_2}}} = \frac{\dfrac{45}{180} - \dfrac{34}{120}}{\sqrt{\dfrac{0.2633\,(1-0.2633)}{180} + \dfrac{0.2633\,(1-0.2633)}{120}}}$$

$$= \frac{0.25 - 0.2833}{\sqrt{\dfrac{0.194}{180} + \dfrac{0.194}{120}}} = \frac{-0.0333}{\sqrt{0.0011 + 0.0016}} = -0.641$$

$|Z| = 0.641$

And $z_{\frac{\alpha}{2}} = z_{0.025} = 1.96$

(vii) Decision: Since $|Z| < z_{\frac{\alpha}{2}}$, accept H_0 based on the sample data at $\alpha = 0.05$

(2) A study shows that 16 of 200 tractors produced one assembly line required extensive adjustments before they could be shipped, while the same was true for 14 of 400 tractors produced another assembly line. At the 0.01 level of significance, does this support the claim that the second line does superior work?

Here $n_1 = 200,\ x_1 = 16, n_2 = 400,\ x_2 = 14$ and $\alpha = 0.01$

Second line does superior work means that tractors produced by 2nd line required less adjustments; that is, $p_1 > p_2$ or $p_1 - p_2 > 0$

So we have to check the hypothesis $p_1 - p_2 > 0$

(i) Null Hypothesis H_0 : $p_1 - p_2 = 0$

(ii) Alternative Hypothesis H_1 : $p_1 - p_2 > 0$

(iii) Level of Significance : $\alpha = 0.01$

(iv) Test statistic : $Z = \dfrac{\dfrac{x_1}{n_1} - \dfrac{x_2}{n_2}}{\sqrt{\dfrac{\hat{p}\,(1-\hat{p})}{n_1} + \dfrac{\hat{p}\,(1-\hat{p})}{n_2}}}$ where, $\hat{p} = \dfrac{x_1 + x_2}{n_1 + n_2}$

(v) Criterion : Reject H_0 if $Z > z_{\alpha}$ and $z_{\alpha} = z_{0.01} = 2.33$

(vi) Calculation :

$$\hat{p} = \frac{x_1 + x_2}{n_1 + n_2} = \frac{16+14}{200+400} = \frac{30}{600} = 0.05$$

$$Z = \frac{\frac{x_1}{n_1} - \frac{x_2}{n_2}}{\sqrt{\frac{\hat{p}\,(1-\hat{p})}{n_1} + \frac{\hat{p}\,(1-\hat{p})}{n_2}}} = \frac{\frac{16}{200} - \frac{14}{400}}{\sqrt{\frac{0.05\,(1-0.05)}{200} + \frac{0.05\,(1-0.05)}{400}}} = 2.3841$$

And $z_\alpha = z_{0.01} = 2.33$

(vii) Decision: Since $Z > z_\alpha$, reject H_0 based on the sample data at $\alpha = 0.01$

(3) A machine puts out 9 imperfect articles in a sample of 200 articles. After the machine is overhauled it puts out 5 imperfect articles in a sample of 700 articles. Test at 5% level of significance, whether the machine is improved.

Here $n_1 = 200,\ x_1 = 9, n_2 = 700,\ x_2 = 5$ and $\alpha = 0.05$

The machine is improved means that the machine puts out less number of imperfect articles; that is, $p_1 > p_2$ or $p_1 - p_2 > 0$

So we have to check the hypothesis $p_1 - p_2 > 0$

(i) Null Hypothesis H_0 : $p_1 - p_2 = 0$

(ii) Alternative Hypothesis H_1 : $p_1 - p_2 > 0$

(iii) Level of Significance : $\alpha = 0.05$

(iv) Test statistic : $Z = \dfrac{\frac{x_1}{n_1} - \frac{x_2}{n_2}}{\sqrt{\frac{\hat{p}\,(1-\hat{p})}{n_1} + \frac{\hat{p}\,(1-\hat{p})}{n_2}}}$ where, $\hat{p} = \dfrac{x_1 + x_2}{n_1 + n_2}$

(v) Criterion : Reject H_0 if $Z > z_\alpha$ and $z_\alpha = z_{0.05} = 1.645$

(vi) Calculation :

$$\hat{p} = \frac{x_1 + x_2}{n_1 + n_2} = \frac{9+5}{200+700} = \frac{14}{900} = 0.0156$$

$$Z = \frac{\frac{x_1}{n_1} - \frac{x_2}{n_2}}{\sqrt{\frac{\hat{p}\,(1-\hat{p})}{n_1} + \frac{\hat{p}\,(1-\hat{p})}{n_2}}} = \frac{\frac{9}{200} - \frac{5}{700}}{\sqrt{\frac{0.0156\,(1-0.0156)}{200} + \frac{0.0156\,(1-0.0156)}{700}}}$$

$$= \frac{0.045 - 0.0071}{\sqrt{\frac{0.0154}{200} + \frac{0.0154}{700}}} = \frac{0.0379}{\sqrt{\frac{0.0154}{200} + \frac{0.0154}{700}}} = 3.8091$$

And $z_\alpha = z_{0.05} = 1.645$

(vii) Decision: Since $Z > z_\alpha$, reject H_0 based on the sample data at $\alpha = 0.05$
That is, the machine is improved

(4) Random samples of 400 men and 600 women were asked whether they would like to have a flyover near their residence. 200 men and 325 women were in favor of the proposal. Test the hypothesis that proportions of men and women in favor of the proposal are same, at 5% level of significance.

Here $n_1 = 400,\ x_1 = 200, n_2 = 600,\ x_2 = 325$ and $\alpha = 0.05$

Here we have to test *the proportions of men and women in favor of the proposal are same;* that is, $p_1 = p_2$ or $p_1 - p_2 = 0$

(i) Null Hypothesis H_0 : $p_1 - p_2 = 0$

(ii) Alternative Hypothesis H_1 : $p_1 - p_2 \neq 0$

(iii) Level of Significance : $\alpha = 0.05$

(iv) Test statistic : $Z = \dfrac{\dfrac{x_1}{n_1} - \dfrac{x_2}{n_2}}{\sqrt{\dfrac{\hat{p}\,(1-\hat{p})}{n_1} + \dfrac{\hat{p}\,(1-\hat{p})}{n_2}}}$ where, $\hat{p} = \dfrac{x_1 + x_2}{n_1 + n_2}$

(v) Criterion : Reject H_0 if $|Z| > z_{\frac{\alpha}{2}}$ and $z_{\frac{\alpha}{2}} = z_{0.025} = 1.96$

(vi) Calculation :

$$\hat{p} = \frac{x_1 + x_2}{n_1 + n_2} = \frac{200 + 325}{400 + 600} = \frac{525}{1000} = 0.525$$

$$Z = \frac{\dfrac{x_1}{n_1} - \dfrac{x_2}{n_2}}{\sqrt{\dfrac{\hat{p}\,(1-\hat{p})}{n_1} + \dfrac{\hat{p}\,(1-\hat{p})}{n_2}}} = \frac{\dfrac{200}{400} - \dfrac{325}{600}}{\sqrt{\dfrac{0.525\,(1-0.525)}{400} + \dfrac{0.525\,(1-0.525)}{600}}}$$

$$= \frac{0.5 - 0.5417}{\sqrt{\dfrac{0.2494}{400} + \dfrac{0.2494}{600}}} = \frac{-0.0417}{\sqrt{0.0006235 + 0.0004157}} = -1.2936$$

So $|Z| = 1.2936$ and $z_{\frac{\alpha}{2}} = z_{0.025} = 1.96$

(vii) Decision: Since $|Z| < z_{\frac{\alpha}{2}}$, accept H_0 based on the sample data at $\alpha = 0.05$

Exercise:

(1) One method of seeding clouds was successful in 57 of 150 attempts while another method was successful in 33 of 100 attempts. At the 0.05 level of significance, can we conclude that the first method is better than second?

(Hint $n_1 = 150,\ x_1 = 57,\ n_2 = 100,\ x_2 = 33,\ H_0 : p_1 - p_2 = 0,\ H_1 : p_1 - p_2 > 0$)

(2) Before an increase in excise duty on tea, 800 persons out of a sample of 1000 persons were found to be tea drinkers. After an increase in duty, 800 people were tea drinkers in a sample of 1200 people. Test whether there is a significant decrease in the consumption of tea after the increase in excise duty. (Use 5% LOS).

(3) In a large city *A*, 25% of a random sample of 900 school boys had defective eye-sight. In another large city *B*, 15·5% of a random sample of 1,600 school boys had the same defect Is this difference between the two proportions significant? (Ans. Not significant.)

(4) A machine puts out 16 imperfect articles in a sample of 500 articles. After the machine is overhauled it puts out 3 imperfect articles in a sample of 100 articles. Test at 1% level of significance, whether the machine is improved.

(5) In a random sample of 500 men from a particular district of U.P., 300 are found to be smokers. In one of 1,000 men from another district, 550 are smokers. Do the data indicate that the two districts are significantly different with respect to the prevalence of smoking among men? Ans. Z = 1·85, (not significant).

(6) In a referendum submitted to the students' body at a university, 850 men and 566 women voted. 530 of the men and 304 of the women voted, yes. Does this indicate a significant difference of opinion on the matter at 1 % level, between men and women students? [Ans. Z = 3·2, (significant)]

(7) In a random sample of 800 adults from the population of a certain large city, 600 are found to have dark hair. In a random sample of 1,000 adults from the habitants of another large city, 700 are dark haired. Show that the difference of the proportion of dark haired people is nearly 2·4 times, the standard error of the difference for samples of above sizes.

(8) An urban community would like to show that the incidence of breast cancer is higher in their area than in a nearby rural area. If it is found that 20 of 200 adult women in the urban community have breast cancer and 10 of 150 adult women in the rural community have breast cancer, can we conclude at the 0.05 level of significance that breast cancer is more prevalent in the urban community?

Test of Hypothesis – Several Proportions:

Let $n_1, n_2, n_3, \cdots n_k$ be the sizes of k number of samples taken from k number of populations respectively.

Let $x_1, x_2, x_3, \cdots x_k$ be the numbers of successes of these k samples respectively.

Let $n = n_1 + n_2 + n_3 + \cdots + n_k$ and $x = x_1 + x_2 + x_3 + \cdots + x_k$.

Now all these values can be tabulated as follows

	Sample 1	Sample 2	…	Sample k	Total
No. of successes	x_1	x_2	…	x_k	x
No. of failures	$n_1 - x_1$	$n_2 - x_2$	…	$n_k - x_k$	$n - x$
Total	n_1	n_2	…	n_k	n

In the above table, each entry in (i, j)th cell is called Observed frequency and it is denoted by O_{ij} for $i = 1, 2$ and $j = 1, 2, 3, \cdots k$; that is,

$$O_{11} = x_1,\ O_{12} = x_2,\ \cdots\ O_{1k} = x_k,$$

$$O_{21} = n_1 - x_1,\ O_{22} = n_2 - x_2,\ \cdots\ O_{2k} = n_k - x_k$$

And the Expected frequency of each (i, j)th cell is given by

$$e_{ij} = \frac{\left(i^{\text{th}} \text{ row total}\right) \times \left(j^{\text{th}} \text{ column total}\right)}{n} \quad \text{for } i = 1, 2 \text{ and } j = 1, 2, 3, \cdots k;$$ that is,

$$e_{11} = \frac{x\,n_1}{n},\ e_{12} = \frac{x\,n_2}{n},\ \cdots\ e_{1k} = \frac{x\,n_k}{n},$$

$$e_{21} = \frac{(n - x)n_1}{n},\ e_{22} = \frac{(n - x)n_2}{n},\ \cdots\ e_{2k} = \frac{(n - x)n_k}{n}$$

(1) **Null Hypothesis** H_0 : $p_1 = p_2 = p_3 = \cdots = p_k$

(2) **Alternative Hypothesis** H_1 : Not all $p_1, p_2, p_3, \cdots p_k$ are equal

(3) **Level of Significance** : α

(4) **Test statistic** : $\chi^2 = \sum_{i=1}^{2}\sum_{j=1}^{k} \frac{(O_{ij} - e_{ij})^2}{e_{ij}}$ with $\nu = k - 1$

(5) **Criterion** : Reject H_0 if $\chi^2 > \chi^2_\alpha$

Examples:

(1) Samples of three kinds of materials, subjected to extreme temperature changes, produced the results shown in the following table

	Material A	Material B	Material C	Total
Crumbled	41	27	22	90
Remained intact	79	53	78	210
Total	120	80	100	300

Use the 0.05 level of significance to test whether, under the stated conditions, the probability of crumbling is the same for the three kinds of materials

Here the data is in 2 rows and 3 columns; that is the number of samples, $k = 3$

Therefore, $\nu = k - 1 = 2$

(i) Null Hypothesis H_0 : p_1, p_2, p_3 are all equal **or** $p_1 = p_2 = p_3$

(ii) Alternative Hypothesis H_1 : p_1, p_2, p_3 are not all equal

(iii) Level of Significance : $\alpha = 0.05$

(iv) Test statistic : $\chi^2 = \sum_{i=1}^{2}\sum_{j=1}^{k} \frac{(O_{ij} - e_{ij})^2}{e_{ij}}$ with $\nu = k - 1$

(v) Criterion : Reject H_0 if $\chi^2 > \chi^2_\alpha$ and $\chi^2_\alpha = \chi^2_{0.05} = 5.991$

(vi) Calculation :

Observed frequency, O_{ij} = entry in (i, j)$^{\text{th}}$ cell, for $i = 1,2$ and $j = 1,2,3$

$O_{11} = 41, O_{12} = 27, O_{13} = 22$

$O_{21} = 79, O_{22} = 53, O_{23} = 78$

Expected frequency, $e_{ij} = \frac{(i^{\text{th}} \text{ row total}) \times (j^{\text{th}} \text{ column total})}{n}$ for $i = 1,2$ and $j = 1,2,3$

$$e_{11} = \frac{(90)(120)}{300} = 36,\ e_{12} = \frac{(90)(80)}{300} = 24,\ e_{13} = \frac{(90)(100)}{300} = 30$$

$$e_{21} = \frac{(210)(120)}{300} = 84,\ e_{22} = \frac{(210)(80)}{300} = 56,\ e_{23} = \frac{(210)(100)}{300} = 70$$

Now $\chi^2 = \sum_{i=1}^{2}\sum_{j=1}^{k} \frac{(O_{ij} - e_{ij})^2}{e_{ij}}$

$$= \frac{(O_{11}-e_{11})^2}{e_{11}} + \frac{(O_{12}-e_{12})^2}{e_{12}} + \frac{(O_{13}-e_{13})^2}{e_{13}}$$

$$+ \frac{(O_{21}-e_{21})^2}{e_{21}} + \frac{(O_{22}-e_{22})^2}{e_{22}} + \frac{(O_{23}-e_{23})^2}{e_{23}}$$

$$= \frac{(41-36)^2}{36} + \frac{(27-24)^2}{24} + \frac{(22-30)^2}{30} + \frac{(79-84)^2}{84} + \frac{(53-56)^2}{56} + \frac{(78-70)^2}{70}$$

$$= 0.6944 + 0.375 + 2.1333 + 0.2976 + 0.1607 + 0.9143$$

$$= 4.5754$$

And for $\nu = k-1 = 2$, $\chi_\alpha{}^2 = \chi_{0.05}{}^2 = 5.991$

(vii) Decision: Since $\chi^2 < \chi_\alpha{}^2$, accept H_0 based on the sample data at $\alpha = 0.05$

(2) Four methods are under development for making discs of a superconducting material. Fifty discs are made by each method and they are checked for superconductivity when cooled with liquid nitrogen

	Method I	Method II	Method III	Method IV	Total
Super conductors	31	42	22	25	120
Failures	19	8	28	25	80
Total	50	50	50	50	200

Test whether there is any significant difference between the proportions of super conductors produced at the 0.05 level of significance

Here the data is in 2 rows and 4 columns; that is the number of samples, $k = 4$

Therefore, $\nu = k-1 = 3$

(i) Null Hypothesis H_0 : p_1, p_2, p_3, p_4 are all equal **or** $p_1 = p_2 = p_3 = p_4$

(ii) Alternative Hypothesis H_1 : p_1, p_2, p_3, p_4 are not all equal

(iii) Level of Significance : $\alpha = 0.05$

(iv) Test statistic : $\chi^2 = \sum_{i=1}^{2}\sum_{j=1}^{k} \frac{(O_{ij}-e_{ij})^2}{e_{ij}}$ with $\nu = k-1$

(v) Criterion : Reject H_0 if $\chi^2 > \chi_\alpha^2$ and $\chi_\alpha^2 = \chi_{0.05}^2 = 7.815$

(vi) Calculation :

Observed frequency, O_{ij} = entry in (i, j)$^{\text{th}}$ cell, for $i = 1,2$ and $j = 1,2,3,4$

$O_{11} = 31, O_{12} = 42, O_{13} = 22, O_{14} = 25$

$O_{21} = 19, O_{22} = 8, O_{23} = 28, O_{24} = 25$

Expected frequency, $e_{ij} = \frac{(i^{\text{th}} \text{ row total}) \times (j^{\text{th}} \text{ column total})}{n}$ for $i = 1,2$ and $j = 1,2,3,4$

$$e_{11} = \frac{(120)(50)}{200} = 30,\ e_{12} = \frac{(120)(50)}{200} = 30,\ e_{13} = \frac{(120)(50)}{200} = 30,\ e_{14} = \frac{(120)(50)}{200} = 30$$

$$e_{21} = \frac{(80)(50)}{200} = 20,\ e_{22} = \frac{(80)(50)}{200} = 20,\ e_{23} = \frac{(80)(50)}{200} = 20,\ e_{24} = \frac{(80)(50)}{200} = 20$$

Now $\chi^2 = \sum_{i=1}^{2}\sum_{j=1}^{k}\frac{(O_{ij}-e_{ij})^2}{e_{ij}}$

$$= \frac{(O_{11}-e_{11})^2}{e_{11}} + \frac{(O_{12}-e_{12})^2}{e_{12}} + \frac{(O_{13}-e_{13})^2}{e_{13}} + \frac{(O_{14}-e_{14})^2}{e_{14}}$$

$$+ \frac{(O_{21}-e_{21})^2}{e_{21}} + \frac{(O_{22}-e_{22})^2}{e_{22}} + \frac{(O_{23}-e_{23})^2}{e_{23}} + \frac{(O_{24}-e_{24})^2}{e_{24}}$$

$$= \frac{(31-30)^2}{30} + \frac{(42-30)^2}{30} + \frac{(22-30)^2}{30} + \frac{(25-30)^2}{30}$$

$$+ \frac{(19-20)^2}{20} + \frac{(8-20)^2}{20} + \frac{(28-20)^2}{20} + \frac{(25-20)^2}{20}$$

$$= 0.0333 + 4.8 + 2.1333 + 0.8333 + 0.05 + 7.2 + 3.2 + 1.25$$

$$= 19.4999$$

And for $\nu = k-1 = 3$, $\chi^2_\alpha = \chi^2_{0.05} = 7.815$

(vii) Decision: Since $\chi^2 > \chi^2_\alpha$, reject H_0 based on the sample data at $\alpha = 0.05$

(3) Tests are made on the proportion of defective castings produced by 5 different molds. If there were 14 defectives among 100 castings made with Mold I, 33 defectives among 200 castings made with Mold II, 21 defectives among 180 castings made with Mold III, 17 defectives among 120 castings made with Mold IV, and 25 defectives among 150 castings made with Mold V, use the 0.01 level of significance to test whether the true proportion of defectives is the same for each mold.

The give data can be tabulated as follows

	Mold I	Mold II	Mold III	Mold IV	Mold V	Total
Defectives	14	33	21	17	25	110
Non defectives	86	167	159	103	125	640
Total	100	200	180	120	150	750

Here the data is in 2 rows and 5 columns; that is the number of samples, $k = 5$

Therefore, $\nu = k-1 = 4$

(i) Null Hypothesis H_0 : p_1, p_2, p_3, p_4, p_5 are all equal **or** $p_1 = p_2 = p_3 = p_4 = p_5$

(ii) Alternative Hypothesis H_1 : p_1, p_2, p_3, p_4, p_5 are not all equal

(iii) Level of Significance : $\alpha = 0.01$

(iv) Test statistic : $\chi^2 = \sum_{i=1}^{2}\sum_{j=1}^{k}\frac{(O_{ij}-e_{ij})^2}{e_{ij}}$ with $\nu = k-1$

(v) Criterion : Reject H_0 if $\chi^2 > \chi_\alpha{}^2$ and $\chi^2_\alpha = \chi^2_{0.01} = 13.277$

(vi) Calculation :

Observed frequency, O_{ij} = entry in (i, j)$^{\text{th}}$ cell, for $i = 1,2$ and $j = 1,2,3,4,5$

$O_{11} = 14, O_{12} = 33, O_{13} = 21, O_{14} = 17, O_{15} = 25$

$O_{21} = 86, O_{22} = 167, O_{23} = 159, O_{24} = 103, O_{25} = 125$

Expected frequency, $e_{ij} = \frac{(i^{\text{th}} \text{ row total}) \times (j^{\text{th}} \text{ column total})}{n}$ for $i = 1,2$ and $j = 1,2,3,4,5$

$$e_{11} = \frac{(110)(100)}{750} = 14.67,\ e_{12} = \frac{(110)(200)}{750} = 29.33,\ e_{13} = \frac{(110)(180)}{750} = 26.4,$$

$$e_{14} = \frac{(110)(120)}{750} = 17.6,\ e_{15} = \frac{(110)(150)}{750} = 22$$

$$e_{21} = \frac{(640)(100)}{750} = 85.33,\ e_{22} = \frac{(640)(200)}{750} = 170.66,\ e_{23} = \frac{(640)(180)}{750} = 153.6,$$

$$e_{24} = \frac{(640)(120)}{750} = 102.4,\ e_{25} = \frac{(640)(150)}{750} = 128$$

Now $\chi^2 = \sum_{i=1}^{2}\sum_{j=1}^{k} \frac{(O_{ij} - e_{ij})^2}{e_{ij}}$

$$= \frac{(O_{11}-e_{11})^2}{e_{11}} + \frac{(O_{12}-e_{12})^2}{e_{12}} + \frac{(O_{13}-e_{13})^2}{e_{13}} + \frac{(O_{14}-e_{14})^2}{e_{14}} + \frac{(O_{15}-e_{15})^2}{e_{15}}$$

$$+ \frac{(O_{21}-e_{21})^2}{e_{21}} + \frac{(O_{22}-e_{22})^2}{e_{22}} + \frac{(O_{23}-e_{23})^2}{e_{23}} + \frac{(O_{24}-e_{24})^2}{e_{24}} + \frac{(O_{25}-e_{25})^2}{e_{25}}$$

$$= \frac{(14-14.67)^2}{14.67} + \frac{(33-29.33)^2}{29.33} + \frac{(21-26.4)^2}{26.4} + \frac{(17-17.6)^2}{17.6} + \frac{(25-22)^2}{22}$$

$$+ \frac{(86-85.33)^2}{85.33} + \frac{(167-170.66)^2}{170.66} + \frac{(159-153.6)^2}{153.6_{23}} + \frac{(103-102.4)^2}{102.4} + \frac{(125-128)^2}{128}$$

$$= 0.0306 + 0.4592 + 1.1046 + 0.0205 + 0.4091$$

$$+ 0.0053 + 0.0785 + 0.1898 + 0.0035 + 0.0703$$

$$= 2.3714$$

And for $\nu = k - 1 = 4$, $\chi^2_\alpha = \chi^2_{0.01} = 13.277$

(vii) Decision: Since $\chi^2 < \chi^2_\alpha$, accept H_0 based on the sample data at $\alpha = 0.01$

Exercise:

(1) The following data come from a study in which random samples of the employees of three government agencies were asked questions about their pension plan:

	Agency 1	Agency 2	Agency 3
For the pension plan	67	84	109
Against the pension plan	33	66	41

Use the 0.01 level of significance to test the null hypothesis that the actual proportions of employees favoring the pension plan are the same

(2) The following table gives the classification of 100 workers according to gender and nature of work. Test whether the nature of work is independent of the gender of the worker at the 0.05 level of significance

	Stable	Unstable	Total
Males	40	20	60
Females	10	30	40
Total	50	50	100

ESTIMATION AND TEST OF HYPOTHESIS OF VARIANCES

Confidence Interval – One Variance:

With the probability $1-\alpha$ or $(1-\alpha)100\%$ confidence,

Lower and Upper confidence limits are respectively given by $\dfrac{(n-1)s^2}{\chi^2_{\frac{\alpha}{2}}}$ and $\dfrac{(n-1)s^2}{\chi^2_{1-\frac{\alpha}{2}}}$

Confidence Interval is given by $\left(\dfrac{(n-1)s^2}{\chi^2_{\frac{\alpha}{2}}}, \dfrac{(n-1)s^2}{\chi^2_{1-\frac{\alpha}{2}}}\right)$

Examples:

(1) Suppose that the refractive indices of 20 pieces of glass (randomly selected from a large shipment purchased by the optical firm) have a variance of 1.20×10^{-4}. Construct a 95% confidence interval for σ, the standard deviation of the population.

Here $n=20, \nu=n-1=19$ and $s^2=1.20\times10^{-4}$

Also, confidence $=95\%$, probability $1-\alpha=0.95$ and so $\alpha=0.05$

Now, $\dfrac{\alpha}{2}=0.025$ and $1-\dfrac{\alpha}{2}=0.975$

From the tables, $\chi^2_{\frac{\alpha}{2}}=\chi^2_{0.025}=32.852$ and $\chi^2_{1-\frac{\alpha}{2}}=\chi^2_{0.975}=8.907$

Lower confidence limit for σ^2 is given by $\dfrac{(n-1)s^2}{\chi^2_{\frac{\alpha}{2}}}=\dfrac{19\times1.20\times10^{-4}}{32.852}=0.000069$

Upper confidence limit for σ^2 is given by $\dfrac{(n-1)s^2}{\chi^2_{1-\frac{\alpha}{2}}}=\dfrac{19\times1.20\times10^{-4}}{8.907}=0.000256$

Therefore, the lower confidence limit for σ is $\sqrt{0.000069}=0.0083$

And the upper confidence limit for σ is $\sqrt{0.000256}=0.0160$

Hence the confidence interval for σ is $(0.0083, 0.0160)$

(2) If five pieces of a certain kind of ribbon have a standard deviation of 5.7 pounds, construct a 98% confidence interval for the true standard deviation of the breaking strength of the given kind of ribbon.

Here $n=5, \nu=n-1=4$ and $s=5.7$

Also, confidence $=98\%$, probability $1-\alpha=0.98$ and so $\alpha=0.02$

Now, $\dfrac{\alpha}{2}=0.01$ and $1-\dfrac{\alpha}{2}=0.99$

From the tables, $\chi^2_{\frac{\alpha}{2}}=\chi^2_{0.01}=13.277$ and $\chi^2_{1-\frac{\alpha}{2}}=\chi^2_{0.99}=0.297$

Lower confidence limit for σ^2 is given by $\dfrac{(n-1)s^2}{\chi^2_{\frac{\alpha}{2}}}=\dfrac{4\times(5.7)^2}{13.277}=9.7884$

Upper confidence limit for σ^2 is given by $\dfrac{(n-1)s^2}{\chi^2_{1-\frac{\alpha}{2}}} = \dfrac{4\times(5.7)^2}{0.297} = 437.5758$

Therefore, the lower confidence limit for σ is $\sqrt{9.7884} = 3.1286$

And the upper confidence limit for σ is $\sqrt{437.5758} = 20.9183$

Hence the confidence interval for σ is $(3.1286,\ 20.9183)$

(3) The heights of 10 males of given locality are found to be 70, 67, 62, 68, 61, 68, 70, 64, 64, 66 inches. Construct 99% confidence interval for population variance

Here $n=10, \nu = n-1 = 9$

Also, confidence $=99\%$, probability $1-\alpha = 0.99$ and so $\alpha = 0.01$

Now, $\dfrac{\alpha}{2} = 0.005$ and $1-\dfrac{\alpha}{2} = 0.995$

From the tables, $\chi^2_{\frac{\alpha}{2}} = \chi^2_{0.005} = 23.589$ and $\chi^2_{1-\frac{\alpha}{2}} = \chi^2_{0.995} = 1.735$

Sample mean, $\bar{x} = \dfrac{1}{n}\sum_{i=1}^{n} x_i = \dfrac{70+67+62+68+61+68+70+64+64+66}{10} = \dfrac{660}{10} = 66$

Sample variance, $s^2 = \dfrac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2$

$$= \frac{(70-66)^2 + (67-66)^2 + (62-66)^2 + (68-66)^2 + \cdots + (64-66)^2 + (66-66)^2}{9}$$

$$= \frac{90}{9} = 10$$

(*First clear the data*: on shift mode 3 ==, *enter the data*: SD mode, 70M+67M+…66M+, *For mean*: shift2, 1=….)

Lower confidence limit for σ^2 is given by $\dfrac{(n-1)s^2}{\chi^2_{\frac{\alpha}{2}}} = \dfrac{9\times 10}{23.589} = 3.8153$

Upper confidence limit for σ^2 is given by $\dfrac{(n-1)s^2}{\chi^2_{1-\frac{\alpha}{2}}} = \dfrac{9\times 10}{1.735} = 51.8731$

Hence the confidence interval for σ^2 is $(3.8153,\ 51.8731)$

Exercise:

(1) While performing a particular task, the pulse rate of 25 workers increased on the average by 18.4 beats per minute with a standard deviation of 4.9 beats per minute. Construct 95% confidence interval for the corresponding population variance.

(2) A random sample {10, 112, 19, 14, 15, 18, 11, 13} are drawn from a normal population. Construct 95% confidence interval for population variance

(3) Suppose that the refractive indices of 20 pieces of glass (randomly selected from a large shipment purchased by the optical firm) have a variance of 1.20×10^{-4}. Construct a 99% confidence interval for σ^2, the variance of the population.

(4) If 31 measurements of the boiling point of sulfur have a standard deviation of 0.83 degree Celsius, construct a 98% confidence interval for the true standard deviation of such measurements.

(5) If five pieces of a certain kind of ribbon have a standard deviation of 5.7 pounds, construct a 95% confidence interval for the true standard deviation of the breaking strength of the given kind of ribbon.

Test of Hypothesis – One Variance:

(1) Null Hypothesis H_0 : $\sigma^2 = \sigma_0^2$

(2) Alternative Hypothesis H_1 : Any one of the following $\sigma^2 < \sigma_0^2$, $\sigma^2 > \sigma_0^2$, $\sigma^2 \neq \sigma_0^2$

(3) Level of Significance : α

(4) Test statistic : $\chi^2 = \dfrac{(n-1)S^2}{\sigma^2}$, $\nu = n-1$

(5) Criterion :

H_1	Reject H_0 if
$\sigma^2 < \sigma_0^2$	$\chi^2 < \chi^2_{1-\alpha}$
$\sigma^2 > \sigma_0^2$	$\chi^2 > \chi^2_{\alpha}$
$\sigma^2 \neq \sigma_0^2$	$\chi^2 < \chi^2_{1-\frac{\alpha}{2}}$ or $\chi^2 > \chi^2_{\frac{\alpha}{2}}$ that is, $\chi^2 \notin \left(\chi^2_{1-\frac{\alpha}{2}}, \chi^2_{\frac{\alpha}{2}}\right)$

Problems:

(1) If 12 determinations of the specific heat of iron have a standard deviation of 0.0086, test the null hypothesis $\sigma = 0.01$ for such determinations. Use the alternative hypothesis $\sigma \neq 0.01$ and the level of significance $\alpha = 0.01$

Here $n = 12, \nu = n-1 = 11, s = 0.0086,\ \sigma = 0.01$ and $\alpha = 0.01$

(i) Null Hypothesis H_0 : $\sigma = 0.01$

(ii) Alternative Hypothesis H_1 : $\sigma \neq 0.01$

(iii) Level of Significance : $\alpha = 0.01$

(iv) Test statistic : $\chi^2 = \dfrac{(n-1)S^2}{\sigma^2}$, $\nu = n-1$

(v) Criterion : Reject H_0 if $\chi^2 < \chi^2_{1-\frac{\alpha}{2}}$ or $\chi^2 > \chi^2_{\frac{\alpha}{2}}$

That is, $\chi^2 \notin \left(\chi^2_{1-\frac{\alpha}{2}}, \chi^2_{\frac{\alpha}{2}}\right)$

(vi) Calculation :

$$\chi^2 = \frac{(n-1)S^2}{\sigma^2} = \frac{(11)(0.0086)^2}{(0.01)^2} = 8.1356$$

For $\nu = n-1 = 11$, $\chi^2_{1-\frac{\alpha}{2}} = \chi^2_{1-0.005} = \chi^2_{0.995} = 2.603$ and $\chi^2_{\frac{\alpha}{2}} = \chi^2_{0.005} = 26.757$

(vii) Decision: Since $\chi^2 \in \left(\chi^2_{1-\frac{\alpha}{2}}, \chi^2_{\frac{\alpha}{2}}\right)$, accept H_0 based on the sample data at $\alpha = 0.01$

(2) The security department of a large office building wants to test the null hypothesis $\sigma = 2.0$ minutes for the time it takes a guard to walk his round against the alternative hypothesis that $\sigma \neq 2.0$ minutes. What can it conclude at the 0.01 level of significance if a random sample of size $n = 30$ yields $s = 1.8$ minutes?

Here $n=30, \nu=n-1=29, s=1.8,\ \sigma=2.0$ and $\alpha=0.01$

(i) Null Hypothesis H_0 : $\sigma=2.0$

(ii) Alternative Hypothesis H_1 : $\sigma\neq 2.0$

(iii) Level of Significance : $\alpha=0.01$

(iv) Test statistic : $\chi^2=\dfrac{(n-1)S^2}{\sigma^2},\ \nu=n-1$

(v) Criterion : Reject H_0 if $\chi^2<\chi^2_{1-\frac{\alpha}{2}}$ or $\chi^2>\chi^2_{\frac{\alpha}{2}}$

That is, $\chi^2\notin\left(\chi^2_{1-\frac{\alpha}{2}},\ \chi^2_{\frac{\alpha}{2}}\right)$

(vi) Calculation :

$$\chi^2=\frac{(n-1)S^2}{\sigma^2}=\frac{(29)(1.8)^2}{(2)^2}=23.49$$

For $\nu=n-1=29$, $\chi^2_{1-\frac{\alpha}{2}}=\chi^2_{1-0.005}=\chi^2_{0.995}=13.121$ and $\chi^2_{\frac{\alpha}{2}}=\chi^2_{0.005}=52.336$

(vii) Decision: Since $\chi^2\in\left(\chi^2_{1-\frac{\alpha}{2}},\ \chi^2_{\frac{\alpha}{2}}\right)$, accept H_0 based on the sample data at $\alpha=0.01$

Note: If $s=2.8$, then $\chi^2=\dfrac{(n-1)S^2}{\sigma^2}=\dfrac{(29)(2.8)^2}{(2)^2}=56.84$ and $\chi^2\notin\left(\chi^2_{1-\frac{\alpha}{2}},\ \chi^2_{\frac{\alpha}{2}}\right)=(13.121, 52.336)$

Therefore, reject H_0

(3) A random sample of 6 steel beams has a mean compressive strength of 58,392 psi with standard deviation 648 psi. Use this information and the level of significance 0.05, test the null hypothesis $\sigma=600$ psi against the alternative hypothesis $\sigma>600$

Here $n=6, \nu=n-1=5, s=648,\ \sigma=600$ and $\alpha=0.05$

(i) Null Hypothesis H_0 : $\sigma=600$

(ii) Alternative Hypothesis H_1 : $\sigma>600$

(iii) Level of Significance : $\alpha=0.05$

(iv) Test statistic : $\chi^2=\dfrac{(n-1)S^2}{\sigma^2},\ \nu=n-1$

(v) Criterion : Reject H_0 if $\chi^2>\chi^2_{\alpha}$

(vi) Calculation :

$$\chi^2=\frac{(n-1)S^2}{\sigma^2}=\frac{(5)(648)^2}{(600)^2}=5.832$$

For $\nu=n-1=5$, $\chi^2_{\alpha}=\chi^2_{0.05}=11.070$

(vii) Decision: Since $\chi^2<\chi^2_{\alpha}$, accept H_0 based on the sample data at $\alpha=0.05$

(4) While performing a particular task, the pulse rate of 25 workers increased on the average by 18.4 beats per minute with a standard deviation of 4.9 beats per minute. Use the 0.05 level of significance to test the null hypothesis $\sigma^2=30.0$ against the alternative hypothesis $\sigma^2<30.0$

Here $n=25, \nu=n-1=24, s=4.9,\ \sigma^2=30$ and $\alpha=0.05$

(i) Null Hypothesis H_0 : $\sigma^2 = 30$

(ii) Alternative Hypothesis H_1 : $\sigma^2 < 30$

(iii) Level of Significance : $\alpha = 0.05$

(iv) Test statistic : $\chi^2 = \dfrac{(n-1)S^2}{\sigma^2}$, $\nu = n-1$

(v) Criterion : Reject H_0 if $\chi^2 < \chi^2_{1-\alpha}$

(vi) Calculation :

$$\chi^2 = \frac{(n-1)S^2}{\sigma^2} = \frac{(24)(4.9)^2}{30} = 19.208$$

For $\nu = n-1 = 24$, $\chi^2_{1-\alpha} = \chi^2_{1-0.05} = \chi^2_{0.95} = 13.848$

(vii) Decision: Since $\chi^2 > \chi^2_{\alpha}$, accept H_0 based on the sample data at $\alpha = 0.05$

Exercise:

(1) Use 1% level of significance to test the null hypothesis , $\sigma = 0.015$ inch for the diameters of certain bolts against the alternative hypothesis , $\sigma \neq 0.015$ inch, given that a random sample of size 15 yielded $S^2 = 0.00011$.

(2) Playing 10 rounds of golf on his home course, a golf professional scored with a standard deviation of 1.32. Test the null hypothesis that the consistency of his game on his home course is actually measured by $\sigma =$ 1.20, against the alternative hypothesis that he is less consistent. Use the level of significance 0.05.

(3) The lapping process which is used to grind certain silicon wafers to the proper thickness is acceptable only if σ, the population standard deviation of the thickness of dice cut from the wafers, is at most 0.50 mil. Use the 0.05 level of significance to test the null hypothesis $\sigma = 0.50$ against the alternative hypothesis $\sigma > 0.50$, if the thickness of 15 dice cut from such wafers have a standard deviation of 0.64 mil.

(4) The security department of a large office building wants to test the null hypothesis that $\sigma = 2.0$ minutes for the time it takes a guard to walk his round against the alternative hypothesis that $\sigma \neq 2.0$ minutes. What can it conclude at the 0.01 level of significance if a random sample of size $n = 30$ yields $s = 2.75$ minutes?

(5) Past experience indicates that the time required for high school seniors to complete a standardized test is a normal random variable with a standard deviation of 6 minutes. Test the hypothesis that σ = 6 against the alternative that σ < 6 if a random sample of the test times of 20 high school seniors has a standard deviation s = 4.51. Use a 0.05 level of significance

Test of Hypothesis – Two Variances:

(1) **Null Hypothesis** H_0 **:** $\sigma_1^2 = \sigma_2^2$

(2) **Alternative Hypothesis** H_1 **:** Any one of the following $\sigma_1^2 < \sigma_2^2$, $\sigma_1^2 > \sigma_2^2$, $\sigma_1^2 \neq \sigma_2^2$

(3) **Level of Significance** **:** α

(4) **Test statistic & Criterion** **:**

H_1	Test Statistic	Reject H_0 if
$\sigma_1^2 < \sigma_2^2$	$F = \dfrac{S_2^2}{S_1^2}$	$F > F_\alpha(n_2 - 1,\ n_1 - 1)$
$\sigma_1^2 > \sigma_2^2$	$F = \dfrac{S_1^2}{S_2^2}$	$F > F_\alpha(n_1 - 1,\ n_2 - 1)$
$\sigma_1^2 \neq \sigma_2^2$	If $s_1^2 < s_2^2$ then $F = \dfrac{S_2^2}{S_1^2}$ If $s_1^2 > s_2^2$ then $F = \dfrac{S_1^2}{S_2^2}$	$F > F_{\frac{\alpha}{2}}(n_2 - 1,\ n_1 - 1)$ $F > F_{\frac{\alpha}{2}}(n_1 - 1,\ n_2 - 1)$

Problems:

(1) It is desired to determine whether there is less variability in the silver plating done by company 1 than in that done by company 2. If independent random samples of size 12 of the two companies work yield S_1 = 0.035 mil and S_2 = 0.062 mil, test the null hypothesis $\sigma_1^2 = \sigma_2^2$ against the alternative hypothesis $\sigma_1^2 < \sigma_2^2$ at the 0.05 level of significance

Here $n_1 = 12, n_2 = 12,\ s_1 = 0.035,\ s_2 = 0.062$ and $\alpha = 0.05$

Observe that $s_1^2 < s_2^2$

(i) Null Hypothesis H_0 : $\sigma_1^2 = \sigma_2^2$

(ii) Alternative Hypothesis H_1 : $\sigma_1^2 < \sigma_2^2$

(iii) Level of Significance : $\alpha = 0.05$

(iv) Test statistic : $F = \dfrac{S_2^2}{S_1^2}$

(v) Criterion : Reject H_0 if $F > F_\alpha(n_2 - 1, n_1 - 1) = F_{0.05}(11,11)$

(vi) Calculation : $F = \dfrac{S_2^2}{S_1^2} = \dfrac{(0.062)^2}{(0.035)^2} = 3.138$

And $F_\alpha(n_2 - 1, n_1 - 1) = F_{0.05}(11,11) = 2.82$

(vii) Decision: Since $F > F_\alpha(n_2 - 1, n_1 - 1)$, reject H_0 based on the sample data at $\alpha = 0.05$

(2) The average marks scored by 36 boys, is 72 with a standard deviation of 8. While that for 25 girls is 70 with a standard deviation of 6. Use 0.02 level of significance, to test whether the variances of boys and girls are equal

Here $n_1 = 36, n_2 = 25, s_1 = 8,\ s_2 = 6$ and $\alpha = 0.02$

We have to test $\sigma_1^2 = \sigma_2^2$

Observe that $s_1^2 > s_2^2$

(i) Null Hypothesis H_0 : $\sigma_1^2 = \sigma_2^2$

(ii) Alternative Hypothesis H_1 : $\sigma_1^2 \neq \sigma_2^2$

(iii) Level of Significance : $\alpha = 0.02$

(iv) Test statistic : $F = \dfrac{S_1^2}{S_2^2}$ (since $s_1^2 > s_2^2$)

(v) Criterion : Reject H_0 if $F > F_{\frac{\alpha}{2}}(n_1 - 1, n_2 - 1) = F_{0.01}(35, 24) = 2.535$

(vi) Calculation : $F = \dfrac{S_1^2}{S_2^2} = \dfrac{(8)^2}{(6)^2} = 1.7778$

And $F_{\frac{\alpha}{2}}(n_1 - 1, n_2 - 1) = F_{0.01}(35, 24) = 2.535$

(vii) Decision: Since $F < F_{\frac{\alpha}{2}}(n_1 - 1, n_2 - 1)$, accept H_0 based on the sample data at $\alpha = 0.02$

(3) The following random samples are measurements of the heat-producing capacity (in millions of calories per ton) of specimens of coal from two mines

Mine 1	8,260	8,130	8,350	8,070	8,340	
Mine 2	7,950	7,890	7,900	8,140	7,920	7,840

Use 0.02 level of significance, to test whether it is reasonable to assume that the variances of the two populations are equal

Here $n_1 = 5, n_2 = 6, s_1 = 125.499,\ s_2 = 104.499$ and $\alpha = 0.02$

We have to test $\sigma_1^2 = \sigma_2^2$

Observe that $s_1^2 > s_2^2$

(i) Null Hypothesis H_0 : $\sigma_1^2 = \sigma_2^2$

(ii) Alternative Hypothesis H_1 : $\sigma_1^2 \neq \sigma_2^2$

(iii) Level of Significance : $\alpha = 0.02$

(iv) Test statistic : $F = \dfrac{S_1^2}{S_2^2}$ (since $s_1^2 > s_2^2$)

(v) Criterion : Reject H_0 if $F > F_{\frac{\alpha}{2}}(n_1 - 1, n_2 - 1) = F_{0.01}(4, 5) = 11.39$

(vi) Calculation : $F = \dfrac{S_1^2}{S_2^2} = \dfrac{(125.499)^2}{(104.499)^2} = 1.4423$

And $F_{\frac{\alpha}{2}}(n_1 - 1, n_2 - 1) = F_{0.01}(4, 5) = 11.39$

(vii) Decision: Since $F < F_{\frac{\alpha}{2}}(n_1 - 1, n_2 - 1)$, accept H_0 based on the sample data at $\alpha = 0.02$

Exercise:

(1) The measurements of the output of two units have given the following results. Assuming that both samples have been obtained from the normal populations at 1% significant level, test whether the two populations have the same variance.

Unit A	14.1	10.1	14.7	13.7	14.0
Unit B	14.0	14.5	13.7	12.7	14.1

(2) A study is conducted to compare the lengths of time required by men and women to assemble a certain product. Past experience indicates that the distribution of times for both men and women is approximately normal but the variance of the times for women is less than that for men. A random sample of times for 11 men and 14 women produced the following data:

Men	Women
$n_1 = 11$	$n_2 = 14$
$s_1 = 6.1$	$s_2 = 5.3$

Test the hypothesis that $\sigma_1^2 = \sigma_2^2$ against the alternative that $\sigma_1^2 > \sigma_2^2$ at the 0.05 level of significance.

Maximum Error & Confidence Interval with probability $1-\alpha$

S.No.		Maximum Error	Confidence Interval
1	One Mean Large Samples	$E = z_{\frac{\alpha}{2}}\left(\frac{\sigma}{\sqrt{n}}\right)$	$\left(\bar{x}-E, \bar{x}+E\right)$
2	One Mean Small Samples	$E = t_{\frac{\alpha}{2}}\left(\frac{s}{\sqrt{n}}\right)$	$\left(\bar{x}-E, \bar{x}+E\right)$
3	Two Means Large Samples	$E = z_{\frac{\alpha}{2}}\sqrt{\left(\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}\right)}$	$\left(\left(\bar{x}_1-\bar{x}_2\right)-E,\ \left(\bar{x}_1-\bar{x}_2\right)+E\right)$
4	Two Means Small Samples	$E = t_{\frac{\alpha}{2}}\sqrt{\left(\frac{\sigma_p^2}{n_1}+\frac{\sigma_p^2}{n_2}\right)}$ where $\sigma_p^2 = \frac{(n_1-1)\,s_1^{\,2}+(n_2-1)\,s_2^{\,2}}{n_1+n_2-2}$ or $\sigma_p^2 = \frac{\sum\left[(x_1-\bar{x}_1)^2+(x_2-\bar{x}_2)^2\right]}{n_1+n_2-2}$	$\left(\left(\bar{x}_1-\bar{x}_2\right)-E,\ \left(\bar{x}_1-\bar{x}_2\right)+E\right)$
5	One Proportion	$E = z_{\frac{\alpha}{2}} \times \sqrt{\frac{\frac{x}{n}\left(1-\frac{x}{n}\right)}{n}}$	$\left(\frac{x}{n}-E,\ \frac{x}{n}+E\right)$
6	Two Proportions	$E = z_{\frac{\alpha}{2}} \times \sqrt{\frac{\hat{p}\,(1-\hat{p})}{n_1}+\frac{\hat{p}\,(1-\hat{p})}{n_2}}$	$\left(\left(\frac{x_1}{n_1}-\frac{x_2}{n_2}\right)-E,\ \left(\frac{x_1}{n_1}-\frac{x_2}{n_2}\right)+E\right)$
7	One Variance		$\left(\frac{(n-1)s^2}{\chi^2_{\frac{\alpha}{2}}},\ \frac{(n-1)s^2}{\chi^2_{1-\frac{\alpha}{2}}}\right)$

Test statistics and Critical regions for tests of Hypotheses

S.No	Test of Hypothesis	Test Statistic	H_1	Reject H_0 if
1	One Mean Large Samples	$Z = \dfrac{\bar{X} - \mu}{\left(\dfrac{\sigma}{\sqrt{n}}\right)}$	$\mu < \mu_0$ $\mu > \mu_0$ $\mu \neq \mu_0$	$Z < -z_\alpha$ $Z > z_\alpha$ $\lvert Z \rvert > z_{\frac{\alpha}{2}}$
2	One Mean Small Samples	$t = \dfrac{\bar{X} - \mu}{\left(\dfrac{s}{\sqrt{n}}\right)}, \qquad \nu = n-1$	$\mu < \mu_0$ $\mu > \mu_0$ $\mu \neq \mu_0$	$t < -t_\alpha$ $t > t_\alpha$ $\lvert t \rvert > t_{\frac{\alpha}{2}}$
3	Two Means Large Samples	$Z = \dfrac{\left(\bar{X}_1 - \bar{X}_2\right) - \delta}{\sqrt{\left(\dfrac{\sigma_1^2}{n_1} + \dfrac{\sigma_2^2}{n_2}\right)}}$	$\mu_1 - \mu_2 < \delta$ $\mu_1 - \mu_2 > \delta$ $\mu_1 - \mu_2 \neq \delta$	$Z < -z_\alpha$ $Z > z_\alpha$ $\lvert Z \rvert > z_{\frac{\alpha}{2}}$
4	Two Means Small Samples	$t = \dfrac{\left(\bar{X}_1 - \bar{X}_2\right) - \delta}{\sqrt{\left(\dfrac{\sigma_p^2}{n_1} + \dfrac{\sigma_p^2}{n_2}\right)}}, \qquad \nu = n_1 + n_2 - 2$ where, $\sigma_p^2 = \dfrac{(n_1 - 1)\, s_1{}^2 + (n_2 - 1)\, s_2{}^2}{n_1 + n_2 - 2}$ or $\sigma_p^2 = \dfrac{\sum\left[(x_1 - \bar{x}_1)^2 + (x_2 - \bar{x}_2)^2\right]}{n_1 + n_2 - 2}$	$\mu_1 - \mu_2 < \delta$ $\mu_1 - \mu_2 > \delta$ $\mu_1 - \mu_2 \neq \delta$	$t < -t_\alpha$ $t > t_\alpha$ $\lvert t \rvert > t_{\frac{\alpha}{2}}$
5	One Proportion	$Z = \dfrac{x - n\,p}{\sqrt{n\,p\,(1-p)}}$ or $Z = \dfrac{\dfrac{x}{n} - p}{\sqrt{\dfrac{p\,(1-p)}{n}}}$	$p < p_0$ $p > p_0$ $p \neq p_0$	$Z < -z_\alpha$ $Z > z_\alpha$ $\lvert Z \rvert > z_{\frac{\alpha}{2}}$
6	Two Proportions	$Z = \dfrac{\dfrac{x_1}{n_1} - \dfrac{x_2}{n_2}}{\sqrt{\dfrac{\hat{p}\,(1-\hat{p})}{n_1} + \dfrac{\hat{p}\,(1-\hat{p})}{n_2}}}$ where, $\hat{p} = \dfrac{x_1 + x_2}{n_1 + n_2}$	$p_1 - p_2 < \delta$ $p_1 - p_2 > \delta$ $p_1 - p_2 \neq \delta$	$Z < -z_\alpha$ $Z > z_\alpha$ $\lvert Z \rvert > z_{\frac{\alpha}{2}}$
7	Several Proportions	$\chi^2 = \sum_{i=1}^{2} \sum_{j=1}^{k} \dfrac{\left(O_{ij} - e_{ij}\right)^2}{e_{ij}}$ with $\nu = k - 1$	$H_0 : p_1, p_2, p_3, \cdots p_k$ are equal $H_1 : p_1, p_2, p_3, \cdots p_k$ are not equal	$\chi^2 > \chi^2_\alpha$
8	One Variance	$\chi^2 = \dfrac{(n-1)\, S^2}{\sigma^2}$ with $\nu = n - 1$	$\sigma^2 < \sigma_0^2$ $\sigma^2 > \sigma_0^2$ $\sigma^2 \neq \sigma_0^2$	$\chi^2 < \chi^2_{1-\alpha}$ $\chi^2 > \chi^2_\alpha$ $\chi^2 < \chi^2_{1-\frac{\alpha}{2}}$ or $\chi^2 > \chi^2_{\frac{\alpha}{2}}$
9	Two Variances	$F = \dfrac{s_2^2}{s_1^2}$ (provided $s_2^2 > s_1^2$), with (ν_2, ν_1)	$\sigma_1^2 < \sigma_2^2$	$F > F_\alpha(\nu_2, \nu_1)$

		$F=\frac{s_1^2}{s_2^2}$ (provided $s_1^2 > s_2^2$), with (ν_1,ν_2) $F=\frac{s_M^2}{s_m^2}$ (provided $s_M^2 > s_m^2$), with (ν_M,ν_m)	$\sigma_1^2 > \sigma_2^2$ $\sigma^2 \neq \sigma_0^2$	$F > F_\alpha(\nu_1,\nu_2)$ $F > F_{\frac{\alpha}{2}}(\nu_M,\nu_m)$