Sampling variables, central limit theorem and confidences limit for unknown mean. Test of Significance for means of two small samples, students 't' distribution. Chi-square distribution as a test of goodness of fit. F-Distribution. (12 Hours)
(RBT Levels: L1, L2 and L3)

4.2 Test of significance - t test

Working rule:

- Write the null hypothesis H_0 and the alternative hypothesis H_1.
- Find the calculated value using $|t| = \left|\frac{\bar{x}-\mu}{S.E(\bar{x})}\right|$, Where $S.E(\bar{x}) = \sqrt{\frac{s^2}{n-1}}$
- Find the critical value using the table at $n-1$ degrees of freedom.
- If calculated value < critical value, Accept H_0. H_0 is the conclusion.
- If calculated value > critical value, Reject H_0. H_1 is the conclusion.

Notation:

	Mean	S.D
Sample	$\bar{x}$	s
Population	μ	σ

Problems:

1. **A Machinist making engine parts with axle diameter of 0.7 inches. A random sample of 10 parts shows mean diameter 0.742 inches with a SD of 0.04 inches. On the basis of this sample, would you say that the work is inferior at 5% level of significance?** $[t_{(.05,\ 9)} = 2.26]$

Since n = 10, apply t test.

By data, $\bar{x} = 0.742, s = 0.04, \mu = 0.7, \alpha = 0.05$

$$S.E(\bar{x}) = \sqrt{\frac{s^2}{n-1}} = \sqrt{\frac{0.04^2}{9}} = 0.0133$$

H_0: $\mu = 0.7$, The work is not inferior.

Under H_0, $|t| = \left|\frac{\bar{x}-\mu}{S.E(\bar{x})}\right| = \left|\frac{0.742-0.7}{0.0133}\right| = 3.1579$

Calculated value = 3.1579

$\alpha = 0.05, \gamma = n-1 = 9.$

Critical value of t = 2.26

Calculated value > critical value,

Reject H_0.

Therefore, the work is inferior.

2. **The nine items of the sample have the following values: 45, 47, 50, 52, 48, 47, 49, 53 and 51. Does the mean of these differ significantly from the assumed mean of 47.5?** $[t_{(.05,\ 8)} = 2.31]$

Since n = 9, apply t test.

By data, $\mu = 47.5, \alpha = 0.05$

$\bar{x} = \frac{\sum x}{n} = \frac{442}{9} = 49.11,$

$s^2 = \frac{\sum x^2}{n} - \left(\frac{\sum x}{n}\right)^2 = \frac{21762}{9} - 49.11^2 = 2418 - 2411.79 = 6.2079$

$S.E(\bar{x}) = \sqrt{\frac{s^2}{n-1}} = \sqrt{\frac{6.2079}{8}} = 0.8809$

Assume H_0: $\mu = 47.5,$

There is no significant difference from the assumed mean 47.5

$|t| = \left|\frac{\bar{x}-\mu}{S.E(\bar{x})}\right| = \left|\frac{49.11-47.5}{0.8809}\right| = 1.8276$

Calculated value = 1.8276

$\alpha = 0.05,\ \gamma = n - 1 = 8$

Critical value = 2.31

Since calculated value < critical value,

Accept H_0.

∴ There is no significant difference from the assumed mean 47.5

3. **A random sample of 10 boys had the following IQ: 70, 120, 110, 101, 88, 83, 95, 98, 107, 100. Does this data support the hypothesis that the population mean of IQ's is 100 at 5% level of significance?** $|t_{(.05,\ 9)}| = 2.26$

Since n = 10, apply t - test.

$$\bar{x} = \frac{\sum x}{n} = \frac{972}{10} = 97.2$$

$$s^2 = \frac{\sum x^2}{n} - \left(\frac{\sum x}{n}\right)^2 = \frac{96312}{10} - 97.2^2 = 183.36$$

$$SE(\bar{x}) = \sqrt{\frac{s^2}{n-1}} = \sqrt{\frac{183.36}{9}} = 4.5136$$

Assume H₀: $\mu = 100$, The population mean of IQ's is 100.

$$|t| = \left|\frac{\bar{x}-\mu}{SE(\bar{x})}\right| = \left|\frac{97.2-100}{4.5136}\right| = 0.6203$$

Calculated value = 0.6203

$\alpha = 0.05,\ \gamma = n - 1 = 9.$

Critical value = 2.26

Calculated value < Critical value

Accept H₀.

Therefore, the population mean IQ is100 at 5% level of significance.

4. A certain stimulus administered to each of 12 patients resulted in the following increases of blood pressure: 5, 2, 8, -1, 3, 0, -2, 1, 5, 0, 4, 6. Can it be concluded that the stimulus will in general be accompanied by an increase in blood pressure?

$[t_{(.05,\ 11)} = 2.2$

Since n = 12, apply t - test.

$\bar{d} = \frac{\sum d}{n} = \frac{31}{12} = 2.5833$

$s^2 = \frac{\sum d^2}{n} - \left(\frac{\sum d}{n}\right)^2 = \frac{185}{12} - \left(\frac{31}{12}\right)^2 = 15.4167 - 6.6734 = 8.7433$

$SE(\bar{d}) = \sqrt{\frac{s^2}{n-1}} = \sqrt{\frac{8.7433}{11}} = 0.8915$

$H_0: \mu = 0$, The stimulus will not increase in blood pressure.

$|t| = \left|\frac{\bar{d}-\mu}{SE(\bar{d})}\right| = \left|\frac{2.5833-0}{0.8915}\right| = 2.8977$

Calculated value = 2.8977

$\alpha = 0.05, \gamma = n - 1 = 11$

Critical value = 2.2

Calculated value > Critical value

Reject H_0.

∴The stimulus will increase blood pressure.

5. **Eleven students were given a test in statistics. They were given a month's further tuition and a second test of equal difficulty was held at the end of it. Do the marks give evidence that the students have benefitted by extra coaching?**

Boys:	1	2	3	4	5	6	7	8	9	10	11
Marks I test:	23	20	19	21	18	20	18	17	23	16	19
Marks II test:	24	19	22	18	20	22	20	20	23	20	17

$[t_{(.05,\ 10)} = 2.23]$

Since n = 11, apply t test.

$\Sigma d = \Sigma(x_2 - x_1) = 11$

$\bar{d} = \dfrac{\sum d}{n} = 1$

$s^2 = \frac{\Sigma d^2}{n} - \left(\frac{\Sigma d}{n}\right)^2 = \frac{61}{11} - 1 = \frac{50}{11}$

$SE\,(\bar{d}) = \sqrt{\frac{s^2}{n-1}} = \sqrt{\frac{50/11}{10}} = \sqrt{\frac{50}{110}} = 0.6742$

Assume H0: $\mu = 0$,

The students did not have benefitted by extra coaching.

$|t| = \left|\frac{\bar{d}-\mu}{SE(\bar{d})}\right| = \left|\frac{1-0}{0.6742}\right| = 1.4832$

Calculated value = 1.4832

$\alpha = 0.05, \gamma = n - 1 = 10.$

~~Critical value = 2.23~~

Calculated value < Critical value

Accept H0.

∴ The students did not have benefit by extra coaching.

6. **A group of boys and girls were given an intelligent test. The mean score SD's and numbers in each group are as follows:**

	Mean	S.D	n
Boys	124	12	18
Girls	121	10	14

Is the mean score of boys significantly different from that of girls?

$[t_{(.05,30)} = 2.04]$

Since $n_1 = 18, n_2 = 14$, apply t test.

By data, $\overline{x_1} = 124, \overline{x_2} = 121, s_1 = 12, s_2 = 10$

$$s^2 = \frac{n_1 s_1^2 + n_2 s_2^2}{n_1 + n_2 - 2} = \frac{18(144) + 14(100)}{18 + 14 - 2} = 133.07$$

$$SE(\overline{x_1} - \overline{x_2}) = \sqrt{s^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)} = \sqrt{133.07\left(\frac{1}{18} + \frac{1}{14}\right)} = 4.1106$$

H₀: $\mu_1 = \mu_2$, the mean score of boys does not differ significantly from that of girls.

$$|t| = \left|\frac{\overline{x_1} - \overline{x_2}}{S.E(\overline{x_1} - \overline{x_2})}\right| = \left|\frac{124 - 121}{4.1106}\right| = 0.7298$$

Therefore, calculated value of t = 0.7298

$\alpha = 0.05, \gamma = n_1 + n_2 - 2 = 30$, Therefore, critical value of t = 2.04

Since calculated value < critical value, Accept H₀.

∴The mean score of boys does not differ significantly from that of girls.

7. **The means of two random samples of sizes 9 and 7 are 196.42 and 198.82 respectively. The sum of squares of the deviations from the mean are 26.94 and 18.73 respectively. Can the sample be considered to have been drawn from the same normal population?** $[t_{(.05,\ 14)} = 2.14]$

Since $n_1 = 9$, $n_2 = 7$, apply t test.

By data, $\overline{x_1} = 196.42, \overline{x_2} = 198.82, n_1 s_1^2 = 26.94, n_2 s_2^2 = 18.73$

$$s^2 = \frac{n_1 s_1^2 + n_2 s_2^2}{n_1 + n_2 - 2} = \frac{26.94 + 18.73}{9 + 7 - 2} = 3.2621$$

$$SE(\overline{x_1} - \overline{x_2}) = \sqrt{s^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)} = \sqrt{3.2621\left(\frac{1}{9} + \frac{1}{7}\right)} = 0.9102$$

H₀: $\mu_1 = \mu_2$, sample is drawn from the same normal population.

$$|t| = \left|\frac{\overline{x_1} - \overline{x_2}}{SE(\overline{x_1} - \overline{x_2})}\right| = \left|\frac{196.42 - 198.82}{0.9102}\right| = 2.6368$$

Therefore, calculated value of t = 2.6368

$\alpha = 0.05, \gamma = n_1 + n_2 - 2 = 9 + 7 - 2 = 14$, Critical value of t = 2.14

Since calculated value > critical value, Reject H_0.

∴ Sample is not drawn from the same normal population.

8. **From a random sample of 10 pigs fed on diet A, the increases in weight in a certain period were 10, 6, 16, 17, 13, 12, 8, 14, 15, 9 lbs. For another random sample of 12 pigs fed on diet B, the increases in the same period were 7, 13, 22, 15, 12, 14, 18, 8, 21, 23, 10, 17 lbs. Test whether diets A and B differ significantly as regards their effect on increases in weight?** $[t_{(.05,\ 14)} = 2.09]$

Since $n_1 = 10$, $n_2 = 12$, apply t test.

By data, $\overline{x_1} = \frac{\sum x_1}{n_1} = \frac{120}{10} = 12$

$\overline{x_2} = \frac{\sum x_2}{n_2} = \frac{180}{12} = 15$

$s_1^2 = \frac{\sum x_1^2}{n_1} - \left(\frac{\sum x_1}{n_1}\right)^2 = \frac{1560}{10} - 144 = 12$

$s_2^2 = \frac{\sum x_2^2}{n_2} - \left(\frac{\sum x_2}{n_2}\right)^2 = \frac{3014}{12} - 225 = 26.17$

$s^2 = \frac{n_1 s_1^2 + n_2 s_2^2}{n_1 + n_2 - 2} = \frac{10(12) + 12(26.17)}{10 + 12 - 2} = 21.7020$

$SE(\overline{x_1} - \overline{x_2}) = \sqrt{s^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)} = \sqrt{21.7020\left(\frac{1}{10} + \frac{1}{12}\right)} = 1.9947$

$H_0: \mu_1 = \mu_2$, diets A and B do not differ significantly.

$|t| = \left|\frac{\overline{x_1} - \overline{x_2}}{SE(\overline{x_1} - \overline{x_2})}\right| = \left|\frac{12 - 15}{1.9947}\right| = 1.5040$

Therefore, calculated value = 1.6

$\alpha = 0.05, \gamma = n_1 + n_2 - 2 = 10 + 12 - 2 = 20$, Critical value = 2.09

Since calculated value < critical value, Accept H_0.

∴ Diets A and B do not differ significantly.

Home work:

9. A machine is expected to produce nails of length 3 inches. A random sample of 25 nails gave an average length of 31. Inches with standard deviation 0.3. Can it be said that the machine is producing nails as per specifications? Given $t_{0.05}(24) = 2.064$

10. Two horses A and B were tested according to the time (seconds) to run a particular race with the following results:

Horse A	28	30	32	33	33	29	34
Horse B	29	30	30	24	27	29	

Test whether you can discriminate between the two horses.

Given that $t_{0.05} = 2.20$ for 11 degrees of freedom

11. Two types of batteries are tested for their length of life and the following results were obtained:

	Battery A	Battery B
Mean	500	500
Variance	100	121
Sample size	10	10

Check whether there is a significant difference between two means. $[t_{0.05}(18) = 0.086]$

12. A sample of 12 measurements of the diameter of a metal ball gave the mean 7.38 mm ~~with standard deviation 1.24 mm. Find 99% confidence limits for~~ actual diameter. $[t_{0.01}(11) = 3.11]$

Note: Confidence limits for the mean are $\bar{x} \pm \frac{s}{\sqrt{n-1}} t_{\alpha}(\gamma)$

13. A group of 10 boys fed on a diet A and another group of 8 boys fed on a different diet B for a period of six months recorded the following increase in weight (lbs).

Diet A	5	6	8	1	12	4	3	9	6	10
Diet B	2	3	6	8	10	1	2	8		

Test whether diet A and B differ significantly regarding their effort on increase in weight.

5.4 Test for goodness of fit

Working rule:

Find Expected frequency using $E_i = N \times P(x)$

Assume H_0: Expected frequency distribution is a good fit to the observed frequency distribution.

Calculated value:

Under H0, $\chi^2 = \sum \frac{(O_i - E_i)^2}{E_i}$ with $n - c$ degrees of freedom.

Where $O_i -$ Observed frequency or tabulated frequency

$E_i -$ Expected frequency or theoretical frequency

$n -$ number of terms, $c -$ number of constraints

Critical value:

Level of significance $\alpha = 0.05$ *or* 0.01 (Always upper tailed)

Degrees of freedom $\gamma = n - c$. Where $c = \begin{cases} 1, & \text{In general} \\ 2, & \text{For Poisson distribution} \\ 3, & \text{For normal distribution} \end{cases}$

Use table.

~~**Conclusion:**~~

If calculated value < critical value,

Accept H_0.

The expected frequency distribution is <u>**a good fit**</u> to the observed frequency distribution.

If calculated value > critical value,

Reject H_0.

The expected frequency distribution <u>**is not a good**</u> fit to the observed frequency distribution.

---

1. **A die is thrown 60 times and the frequency distribution for the number appearing on the face x is given by the following table:**

x:	1	2	3	4	5	6
f:	15	6	4	7	11	17

Test the hypothesis that the die is unbiased. $[\chi^2_{0.05}(5) = 11.07]$

By data observed frequency $O_i: 15, 6, 4, 7, 11, 17.$

Find expected frequency $E_i = N \times P(x)$, $x: 1, 2, 3, 4, 5, 6$

$P(x): \frac{1}{6}, \frac{1}{6}, \frac{1}{6}, \frac{1}{6}, \frac{1}{6}, \frac{1}{6}$

$N \times P(x): 60 \times \left(\frac{1}{6}\right), 60 \times \left(\frac{1}{6}\right), 60 \times \left(\frac{1}{6}\right), 60 \times \left(\frac{1}{6}\right), 60 \times \left(\frac{1}{6}\right), 60 \times \left(\frac{1}{6}\right)$

$E_i: 10, 10, 10, 10, 10, 10$

Assume H_0**: The die is unbiased.**

Under H_0, $\chi^2 = \sum \frac{(O_i - E_i)^2}{E_i}$ **with** $n - 1 = 5$ **d.f**

x_i	O_i	E_i	$(O_i - E_i)^2$	$\frac{(O_i - E_i)^2}{E_i}$
1	15	10	25	2.5
2	6	10	16	1.6
3	4	10	36	3.6
4	7	10	9	0.9
5	11	10	1	0.1
6	17	10	49	4.9
				13.6

Calculated value = 13.6

Critical value= 11.07

∴ Calculated value > Critical Value.

Reject H_o.

Therefore, the die is not unbiased.

2. **The following table gives the number of road accidents that occurred in a large city during the various days of a week. Test the hypothesis that the accidents are uniformly distributed over all the days of a week.** $[\chi^2_{0.05}(6) = 12.59]$

Day:	Sun	Mon	Tue	Wed	Thu	Fri	Sat	Total
No. of accidents:	14	16	8	12	11	9	14	84

By data observed frequency O_i: 14, 16, 8, 12, 11, 9, 14.

Find expected frequency $E_i = N \times P(x)$, x: 1, 2, 3, 4, 5, 6, 7

$P(x)$: $\frac{1}{7}, \frac{1}{7}, \frac{1}{7}, \frac{1}{7}, \frac{1}{7}, \frac{1}{7}$

$N \times P(x)$: $84 \times \left(\frac{1}{7}\right), 84 \times \left(\frac{1}{7}\right), 84 \times \left(\frac{1}{7}\right), 84 \times \left(\frac{1}{7}\right), 84 \times \left(\frac{1}{7}\right), 84 \times \left(\frac{1}{7}\right)$

E_i: 12, 12, 12, 12, 12, 12, 12

Assume H_0: **The accidents are uniformly distributed over all the days of a week.**

Under H_0, $\chi^2 = \sum \frac{(O_i - E_i)^2}{E_i}$, **with** $n - 1 = 6$ **degrees of freedom.**

x	O_i	E_i	$(O_i - E_i)^2$	$\frac{(O_i - E_i)^2}{E_i}$
Sun	14	12	4	4/12
Mon	16	12	16	16/12
Tue	8	12	16	16/12
Wed	12	12	0	0
Thu	11	12	1	1/12
Fri	9	12	9	9/12
Sat	14	12	4	4/12
				50/12

Calculated value $= \frac{50}{12}$

Critical value $= 12.59$

$\therefore$ Calculated value $<$ Critical Value.

Accept H_o.

Therefore, the accidents are uniformly distributed over all the days of a week

3. A set of 5 similar coins is tossed 320 times and the result is

Number of heads: 0 1 2 3 4 5

Frequency: 6 27 72 112 71 32

Test the hypothesis that the data follows a binomial distribution. $[\chi^2_{0.05}(5) = 11.07]$

By data, observed frequency $O_i: 6, 27, 72, 112, 71, 32.$

Find expected frequency $E_i = N \times P(x),\ x: 0, 1, 2, 3, 4, 5$

$P(x) = nC_x p^x q^{n-x} = 5C_x \left(\frac{1}{2}\right)^x \left(\frac{1}{2}\right)^{5-x} = \frac{1}{32} \times 5C_x$ (n- no. of coins)

$N \times P(x) = 320 \times \frac{1}{32} \times 5C_x = 10 \times 5C_x$

$E_i: 10 \times 5C_0, 10 \times 5C_1, 10 \times 5C_2, 10 \times 5C_3, 10 \times 5C_4, 10 \times 5C_5$

Assume H_0: The data follows Binomial distribution.

Under H_0, $\chi^2 = \sum \frac{(O_i - E_i)^2}{E_i}$, with $n - 1 = 6 - 1 = 5$ degrees of freedom.

(n-number of frequencies)

$E_i x$	O_i	E_i	$(O_i - E_i)^2$	$\frac{(O_i - E_i)^2}{E_i}$
0	6	10	16	1.6
1	27	50	529	10.58
2	72	100	784	7.84
3	112	100	144	1.44
4	71	50	441	8.82
5	32	10	484	48.4
				78.68

Calculated value = 78.68

Critical Value = 11.07

∴ Calculated value > Critical Value.

Reject H_o.

Therefore, the data does not follow Binomial distribution,

4. Fit a Poisson distribution to the following data and test for its goodness of fit at level of significance 0.05.

x:	0	1	2	3	4
f:	419	352	154	56	19

$[\chi^2_{0.05}(3) = 7.82]$

By data, observed frequency $O_i: 419, 352, 154, 56, 19$

Find expected frequency $E_i = N \times P(x)$, $x: 0, 1, 2, 3, 4$

$$m = mean = \frac{\Sigma fx}{\Sigma f} = \frac{904}{1000} = 0.904$$

$$e^{-m} = e^{-0.904} = 0.4049$$

$$\therefore P(x) = \frac{e^{-m} m^x}{x!} = \frac{(0.4049)(0.904)^x}{x!}$$

$$E_i = N \times P(x) = 1000 \times P(x) = 404.9 \times \frac{(0.904)^x}{x!}$$

$$E_i = 405, 366, 165, 50, 11$$

Assume H_0: The data follows Poisson distribution.

Under H_0, $\chi^2 = \sum \frac{(O_i - E_i)^2}{E_i}$ with $n - 2 = 5 - 2 = 3$ degrees of freedom.

x (i)	O_i	E_i	$(O_i - E_i)^2$	$\frac{(O_i - E_i)^2}{E_i}$
0	419	405+1	169	0.4033
1	352	366	196	0.5355
2	154	165	121	0.7333
3	56	50	36	0.7200
4	19	11+2	36	2.7692
				5.1613

Numbers added in E_i only to preserve totality.

$\therefore$ Calculated value = 5.1613

Degrees of freedom $= n - 2 = 5 - 2 = 3$ ($\because$ It follows Poisson distribution)

Critical value = 7.82

$\therefore$ Calculated value < Critical Value.

Accept H_0. Therefore, the data follows Poisson distribution.

5. **Genetic theory states that children having one parent of blood type M and other blood type N will always be one of the three types M, MN, N and that the proportion of these types will on average 1:2:1. A report says that out of 300 children having one M parent and one N parent, 30% were found to be type M, 45% of type MN and remainder of type N. Test the hypothesis by χ^2 test. $[\chi^2_{0.05}(2) = 5.99]$**

By data, observed frequency O_i: $30\%\ of\ 300, 45\%\ of\ 300, 25\%\ of\ 300.$

Find expected frequency $E_i = N \times P(x),\ \ x: M, MN, N$

$P(x): \frac{1}{4}, \frac{2}{4}, \frac{1}{4}$

$N \times P(x): 300 \times \frac{1}{4}, 300 \times \frac{2}{4},\ 300 \times \frac{1}{4}$

E_i: $75, 150, 75$

Assume H_0: The proportion of these types is on average 1:2:1

Under H_0, $\chi^2 = \sum \frac{(O_i - E_i)^2}{E_i}$ with $n - 1 = 3 - 1 = 2$ degrees of freedom.

x	O_i	E_i	$(O_i - E_i)^2$	$\frac{(O_i - E_i)^2}{E_i}$
M	30% of 300 = 90	$300 \times \frac{1}{4} = 75$	225	3
MN	45% of 300 = 135	$300 \times \frac{2}{4} = 150$	225	1.5
M	25% of 300 = 75	$300 \times \frac{1}{4} = 75$	0	0
				4.5

Calculated value = 4.5

Critical value = 5.99

Calculated value < Critical Value.

Accept H_0.

Therefore, the Genetic theory, 'The proportion of these types is on average 1:2:1', fitted to the report.

---

6. **In experiments on Pea breading, the following frequencies of seeds were obtained:**

Round and yellow	Wrinkled and yellow	Round and green	Wrinkled and green	Total
315	101	108	32	556

Theory predicts that the frequencies should be in proportions 9: 3: 3: 1. Examine the correspondence between theory and experiment. $[\chi^2_{0.05}(3) = 7.82]$

By data, observed frequency $O_i: 315, 101, 108, 32$

Find expected frequency $E_i = N \times P(x)$, $x: RY, WY, RG, WG$

$P(x): \frac{9}{16}, \frac{3}{16}, \frac{3}{16}, \frac{1}{16}$

$N \times P(x): 556 \times \frac{9}{16},\ 556 \times \frac{3}{16},\ 556 \times \frac{3}{16},\ 556 \times \frac{1}{16}$

$E_i: 313, 104, 104, 35$

Assume H_o: The frequencies are in proportions 9: 3: 3: 1

Under H_0, $\chi^2 = \sum \frac{(O_i - E_i)^2}{E_i}$ with $n - 1 = 4 - 1 = 3$ degrees of freedom.

x (i)	O_i	E_i	$(O_i - E_i)^2$	$\frac{(O_i - E_i)^2}{E_i}$
RY	315	$556\left(\frac{9}{16}\right) = 313$	4	0.0128
WY	101	$556\left(\frac{3}{16}\right) = 104$	9	0.0865
RG	108	$556\left(\frac{3}{16}\right) = 104$	16	0.1538
WG	32	$556\left(\frac{1}{16}\right) = 35$	9	0.2571
				0.5102

Calculated value = 0.5102

Critical value = 7.82

Calculated value < Critical Value.

Accept H_o.

Therefore, the theory, 'The frequencies should be in proportions 9: 3: 3: 1, fitted to the experiment.

Tabulated values of $\chi^2_{\alpha}(\nu)$

ν	$\chi^2_{0.05}(\nu)$	$\chi^2_{0.01}(\nu)$
1	3.84	6.64
2	5.99	9.21
3	7.82	11.34
4	9.49	13.28
5	11.07	15.09
6	12.59	16.81
7	14.07	18.48
8	15.51	20.09
9	16.92	21.67
10	18.31	23.21

---

7. The theory predicates the proportion of beans in the four groups $G_1, G_2, G_3 \, and \, G_4$ should be in the ratio $9:3:3:1$. In an experiment with 1600 beans the numbers in the four groups were 882, 313, 287 and 118. Does the experimental result support the theory? $[\chi^2_{0.05}(3) = 7.815]$
8. Records taken of the number of male and female births in 800 families having four children are as follows:

Number of male births	0	1	2	3	4
Number of female births	4	3	2	1	0
Number of families	32	178	290	236	94

Test whether the data are consistent with the hypothesis that the binomial law holds and the chance of male birth is equal to that of female birth. $[\chi^2_{0.05}(4) = 9.488]$

4.3 F Distribution

Introduction:

- The F test is named in honour of the great statistician R A Fisher. The objective of the F test is to find out whether the two independent estimates of population variance differ significantly or whether the two samples may be regarded as drawn from the same population having the same variance.
- For carrying out the test significance, The ratio F is defined as

$$F = \begin{cases} \dfrac{s_1^2}{s_2^2}, & if\ s_1^2 > s_2^2 \\ \dfrac{s_2^2}{s_1^2}, & if\ s_2^2 > s_1^2 \end{cases}$$

where

$$s_1^2 = \frac{\sum(x_1 - \overline{x_1})^2}{n_1 - 1} \quad and \quad s_2^2 = \frac{\sum(x_2 - \overline{x_2})^2}{n_2 - 1}$$

- The calculated value of F is compared with the tabulated value for v_1 and v_2 at 5% or 1% level of significance. v_1 – Degrees of freedom for sample having larger variance and v_2– Degrees of freedom for sample having smaller variance.
- If the calculated value is less than the tabulated value the null hypothesis is accepted and it is inferred that both the samples come from the same population.
- Since F test is based on the ratio of two variances, it is also known as the variance ratio test.
- The ratio of two variances follows a distribution called the F distribution.
- F– Test is based on the following assumptions:
  - (i) The values in each group are normally distributed.
  - (ii) The variance within each group should equal for all groups ($\sigma_1^2 = \sigma_2^2 = -- = \sigma_r^2$)
  - (iii) The error (variation of each value around its own group mean) should be independent for each value.

1. Two random samples were drawn from two normal populations and their values are:

A: 66 67 75 76 82 84 88 90 92

B: 64 66 74 78 82 85 87 92 93 95 97

Test whether two populations have the same variance at the 5% level of significance. ($F_{10,8} = 3.36$)

Assume H_0: The two populations have the same variance.

$$\overline{x_1} = \frac{\Sigma x_1}{n_1} = \frac{720}{9} = 80$$

$$\overline{x_2} = \frac{\Sigma x_2}{n_2} = \frac{913}{11} = 83$$

To Find: s_1^2 and s_2^2

x_1	$x_1 - \overline{x_1} = x_1 - 80$	x_2	$(x_1 - \overline{x_1})^2 = (x_1 - 80)^2$	$x_2 - \overline{x_2} = x_2 - 83$	$(x_2 - \overline{x_2})^2 = (x_2 - 83)^2$
66	-14	64	196	-19	361
67	-13	66	169	-17	289
75	-5	74	25	-9	81
76	-4	78	16	-5	25
82	2	82	4	-1	1
84	4	85	16	2	4
88	8	87	64	4	16
90	10	92	100	9	81
92	12	93	144	10	100
		95		12	144
		97		14	196
720	0	913	734	0	1298
Σx_1	----	Σx_2	$\Sigma(x_1 - \overline{x_1})^2$	----	$\Sigma(x_2 - \overline{x_2})^2$

$$s_1^2 = \frac{\Sigma(x_1 - \overline{x_1})^2}{n_1 - 1} = \frac{734}{8} = 91.75$$

$$s_2^2 = \frac{\Sigma(x_2 - \overline{x_2})^2}{n_2 - 1} = \frac{1298}{10} = 129.8$$

Calculated value:

$$F = \frac{s_2^2}{s_1^2} = \frac{129.8}{91.75} = \mathbf{1.415}$$

Critical value:

Degrees of freedom $\nu_1 = n_2 - 1 = 10, \nu_2 = n_2 - 1 = 8$

At 5% level of significance, $F_{10,8} = 3.36$

Conclusion:

Since Calculated value < Critical value, accept H_0.

Therefore, the two populations have the same variance.

2. In a sample of 8 observations, the sum of square of deviations of items from the mean was 84.4. In another sample of 10 observations, the value was found to be 102.6. Test whether the difference is significant at 5% level. $[F_{7,9} = 3.29 \text{ and } F_{8,10} = 3.07]$

Assume H_0: There is no significant difference in the variances of the two samples.

By data,

$n_1 = 8$	$\sum(x_1 - \overline{x_1})^2 = 84.4$
$n_2 = 10$	$\sum(x_2 - \overline{x_2})^2 = 102.6$

Therefore,

$$s_1^2 = \frac{\sum(x_1 - \overline{x_1})^2}{n_1 - 1} = \frac{84.4}{8-1} = 12.06$$

$$s_2^2 = \frac{\sum(x_2 - \overline{x_2})^2}{n_2 - 1} = \frac{102.6}{10-1} = 11.4$$

Calculated value:

$$F = \frac{s_1^2}{s_2^2} = \frac{12.06}{11.4} = 1.06$$

Critical value:

Degrees of freedom $(\nu_1, \nu_2) = (n_1 - 1, n_2 - 1) = (7, 9)$.

At 5% level of significance, $F_{7,9} = 3.29$

Conclusion:

Calculated value < Critical value. Accept H_0.

There is no significant difference in the variances of the two samples at 5% level of significance.

3. Two samples are drawn from two normal populations. From the following data, test whether the two samples have the same variance at 5% level:

Sample 1: 60 65 71 74 76 82 85 87

Sample 2: 61 66 67 85 78 63 85 86 88 91

Assume H_0: Two samples drawn from two normal populations have the same variance.

$$\bar{x}_1 = \frac{\sum x_1}{n_1} = \frac{600}{8} = 75$$

$$\bar{x}_2 = \frac{\sum x_2}{n_2} = \frac{770}{10} = 77$$

To Find: s_1^2 and s_2^2

x_1	$x_1 - \bar{x}_1 = x_1 - 75$	$(x_1 - \bar{x}_1)^2 = (x_1 - 75)^2$	x_2	$x_2 - \bar{x}_2 = x_2 - 77$	$(x_2 - \bar{x}_2)^2 = (x_2 - 77)^2$
60	−15	225	61	−16	256
65	−10	100	66	−11	121
71	−4	16	67	−10	100
74	−1	1	85	8	64
76	1	1	78	1	1
82	7	49	63	−14	196
85	10	100	85	8	64
87	12	144	86	9	81
			88	11	121
			91	14	196
600	0	636	770	0	1200
$\sum x_1$	------	$\sum(x_1 - \bar{x}_1)^2$	$\sum x_2$	------	$\sum(x_2 - \bar{x}_2)^2$

$$s_1^2 = \frac{\sum(x_1 - \bar{x}_1)^2}{n_1 - 1} = \frac{636}{7} = 90.857$$

$$s_2^2 = \frac{\sum(x_2 - \bar{x}_2)^2}{n_2 - 1} = \frac{1200}{9} = 133.33$$

Calculated value:

$$F = \frac{s_2^2}{s_1^2} = \frac{133.33}{90.857} = 1.4675$$

Critical value:

Degrees of freedom $v_1 = n_2 - 1 = 9,\ v_2 = n_1 - 1 = 7$

At 5% level of significance, $F_{9,7} = 3.68$

Conclusion:

Since Calculated value < Critical value, accept H_0.

Therefore, the two samples taken from the two populations have the same variance.

4. **The following data present the yields in quintals of common 10 subdivisors of equal area of two agricultural plots:**

Plot 1: 6.2 5.7 6.5 6.0 6.3 5.8 5.7 6.0 6.0 5.8

Plot 2: 5.6 5.9 5.6 5.7 5.8 5.7 6.0 5.5 5.7 5.5

Test whether two samples taken from two random populations have the same variance. ($F_{9,9} = 3.18$)

Assume H_0: Two samples drawn from two normal populations have the same variance.

$$\overline{x_1} = \frac{\Sigma x_1}{n_1} = \frac{60}{10} = 6 \quad and \quad \overline{x_2} = \frac{\Sigma x_2}{n_2} = \frac{57}{10} = 5.7$$

To Find: s_1^2 and s_2^2

x_1	$x_1 - \overline{x_1} = x_1 - 6$	$(x_1 - \overline{x_1})^2 = (x_1 - 6)^2$	x_2	$x_2 - \overline{x_2} = x_2 - 5.7$	$(x_2 - \overline{x_2})^2 = (x_2 - 5.7)^2$
6.2	0.2	0.04	5.6	−0.1	0.01
5.7	−0.3	0.09	5.9	0.2	0.04
6.5	0.5	0.25	5.6	−0.1	0.01
6.0	0	0	5.7	0	0
6.3	0.3	0.09	5.8	0.1	0.01
5.8	−0.2	0.04	5.7	0	0
5.7	−0.3	0.09	6.0	0.3	0.09
6.0	0	0	5.5	−0.2	0.04
6.0	0	0	5.7	0	0
5.8	−0.2	0.04	5.5	−0.2	0.04
60	----	0.64	57	----	0.24
Σx_1	----	$\Sigma(x_1 - \overline{x_1})^2$	Σx_2	----	$\Sigma(x_2 - \overline{x_2})^2$

$$s_1^2 = \frac{\Sigma(x_1 - \overline{x_1})^2}{n_1 - 1} = \frac{0.64}{9} = 0.071$$

$$s_2^2 = \frac{\Sigma(x_2 - \overline{x_2})^2}{n_2 - 1} = \frac{0.24}{9} = 0.027$$

Calculated value:

$$F = \frac{s_1^2}{s_2^2} = \frac{0.071}{0.027} = 2.63$$

Critical value:

Degrees of freedom $\nu_1 = n_1 - 1 = 9, \nu_2 = n_2 - 1 = 9$

At 5% level of significance, $F_{9,9} = 3.18$

Conclusion:

Since Calculated value < Critical value, accept H_0.

Therefore, the two samples taken from the two populations have the same variance.

5. Two samples of sizes 9 and 8 give the sum of squares of deviations from their respective means equal to 160 square inches and 91 square inches respectively. Can these be regarded as drawn from the same population?

Assume H_0: Two samples are drawn from the same population.
By data,

$n_1 = 9$	$\sum(x_1 - \overline{x_1})^2 = 160$
$n_2 = 8$	$\sum(x_2 - \overline{x_2})^2 = 91$

Therefore,

$$s_1^2 = \frac{\sum(x_1 - \overline{x_1})^2}{n_1 - 1} = \frac{160}{9-1} = 20$$

$$s_2^2 = \frac{\sum(x_2 - \overline{x_2})^2}{n_2 - 1} = \frac{91}{8-1} = 13$$

Calculated value:

$$F = \frac{s_1^2}{s_2^2} = \frac{20}{13} = 1.54$$

Critical value:

Degrees of freedom $(\nu_1, \nu_2) = (n_1 - 1, n_2 - 1) = (8, 7)$.
At 5% level of significance, $F_{8,7} = \mathbf{3.73}$

Conclusion:

Calculated value < Critical value. Accept H_0.
Two samples are drawn from the same population.

6. **Measurements on the length of a copper wire were taken in two experiments A and B as under:**

A's Measurements (mm): 12.29 12.25 11.86 12.13 12.44 12.78 12.77 11.90 12.47

B's Measurements (mm): 12.39 12.46 12.34 12.22 11.98 12.46 12.23 12.06

Test whether B's measurements are more accurate than A's. (The readings taken in both cases being unbiased.)

Assume H_0: Both measurements have the same variance.

$$\overline{x_1} = \frac{\Sigma x_1}{n_1} = \frac{110.89}{9} = 12.32 \quad and \quad \overline{x_2} = \frac{\Sigma x_2}{n_2} = \frac{98.14}{8} = 12.27$$

To Find: s_1^2 and s_2^2

x_1	$x_1 - \overline{x_1} = x_1 - 12.32$	$(x_1 - \overline{x_1})^2 = (x_1 - 6)^2$	x_2	$x_2 - \overline{x_2} = x_2 - 12.27$	$(x_2 - \overline{x_2})^2 = (x_2 - 5.7)^2$
12.29	−0.03	0.0009	12.39	0.12	0.0144
12.25	−0.07	0.0049	12.46	0.19	0.0361
11.86	−0.46	0.2116	12.34	0.07	0.0049
12.13	0.19	0.0361	12.22	−0.05	0.0025
12.44	0.12	0.0144	11.98	−0.29	0.0841
12.78	0.46	0.2116	12.46	0.19	0.0361
12.77	0.45	0.2025	12.23	−0.04	0.0016
11.90	−0.42	0.1764	12.06	−0.21	0.0441
~~12.47~~	~~0.15~~	~~0.0225~~			
110.89	----	0.8809	98.14	----	0.2238
Σx_1	----	$\Sigma(x_1 - \overline{x_1})^2$	Σx_2	----	$\Sigma(x_2 - \overline{x_2})^2$

$$s_1^2 = \frac{\Sigma(x_1 - \overline{x_1})^2}{n_1 - 1} = \frac{0.8809}{8} = 0.1101$$

$$s_2^2 = \frac{\Sigma(x_2 - \overline{x_2})^2}{n_2 - 1} = \frac{0.2238}{7} = 0.032$$

Calculated value:

$$F = \frac{s_1^2}{s_2^2} = \frac{0.1101}{0.032} = \mathbf{3.441}$$

Critical value:

Degrees of freedom $\nu_1 = n_1 - 1 = 8, \nu_2 = n_2 - 1 = 7$

At 5% level of significance, $\boldsymbol{F_{8,7}} = \mathbf{3.73}$

Conclusion:

Since Calculated value < Critical value, accept H_0.

Therefore, Both measurements have the same variance.