Imp questions from large samples

1) 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. Has the machine improved?

Sol: Let P_1 and P_2 be the proportions of imperfect articles in the population of articles manufactured by the machine before and after overhauling respectively

1) Let $H_0 : P_1 = P_2$

2) then $H_1 : P_1 > P_2$

3) let $\alpha = 0.05$

4) Test statistic is $z = \dfrac{P_1 - P_2}{\sqrt{pq\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$

where $P_1 = \dfrac{x_1}{n_1} = \dfrac{16}{500} = 0.032$, $P_2 = \dfrac{x_2}{n_2} = \dfrac{3}{100} = 0.03$

Here $P_1 > P_2$ then $p = \dfrac{n_1 P_1 + n_2 P_2}{n_1 + n_2} = \dfrac{x_1 + x_2}{n_1 + n_2} = \dfrac{16+3}{500+100}$

$= \dfrac{19}{600} = 0.032$

$q = 1 - p = 1 - 0.032 = 0.968$

Since H_1 is right-tailed, we apply ~~right~~ one tailed test and $z_\alpha = 1.645$

Now $z = \dfrac{0.032 - 0.03}{\sqrt{0.032 \times 0.968 \left(\frac{1}{500} + \frac{1}{100}\right)}} = \dfrac{0.002}{0.019} = 0.104$

As $z = 0.104 < 1.645$

We accept null hypothesis H_0 at 5% level of significance and ~~machine~~ conclude that the machine has improved.

(2) A researcher wants to know the intelligence of students of students in a school. He selected two groups of students. In the first group there 150 students having mean IQ of 75 with a S.D. of 15 in the second group there are 250 students

having mean IQ of 70 with S.D. of 20, Is there a significant difference between the means of two group?

Sol: Given $n_1 = 150$, $\bar{x}_1 = 75$, $\sigma_1 = 15$

and $n_2 = 250$, $\bar{x}_2 = 70$, $\sigma_2 = 20$.

1. $H_0 : \mu_1 = \mu_2$
2) $H_1 : \mu_1 \neq \mu_2$
3) Let $\alpha = 1\% = 0.01$ then z_α from the table is 2.33
4) Test statistic, $z = \dfrac{\bar{x}_1 - \bar{x}_2}{\sqrt{\dfrac{\sigma_1^2}{n_1} + \dfrac{\sigma_2^2}{n_2}}} = \dfrac{75-70}{\sqrt{\dfrac{225}{150} + \dfrac{400}{250}}} = \dfrac{5}{\sqrt{\dfrac{9}{5} + \dfrac{8}{5}}} = \dfrac{5\sqrt{5}}{\sqrt{17}} = 2.7116$
5) Calculated z > tabulated z

Hence we reject H_0 at 1% level of significance and Conclude that the groups have not taken from the same population

3) 20 people were attacked by a disease and only 18 survived. Will you reject the hypothesis that the survival rate if attacked by this disease is 85% in favour of the hypothesis that is more at 5% level.

Sol: $n = 20$, ~~x=18~~, p = proportion of survived people

x = no. of survived people $= \dfrac{x}{n} = \dfrac{18}{20} = 0.9$

$\therefore p = 0.9$

Given $P = 85\% = 0.85$

$\therefore Q = 1 - P = 1 - 0.85 = 0.15$

1. $H_0 : P = 0.85$
2. $H_1 : P > 0.85$ (right tailed test)
3. $\alpha = 0.05$
4. Test statistic, $z = \dfrac{p - P}{\sqrt{\dfrac{PQ}{n}}} = \dfrac{0.9 - 0.85}{\sqrt{\dfrac{0.85 \times 0.15}{20}}} = \dfrac{0.05}{0.08} = 0.625$
5. Tabulated z at 5% level of significance, $z_\alpha = 1.645$

tab(z) > cal(z), we accept H_0.

④ The average breaking strength of the steel rods is specified to be 18.5 thousand pounds. To test this, sample of 14 rods were tested. The mean and standard deviations obtained were 17.85 and 1.955 respectively. Is the result of the experiment significant?

Sol :- We are given $n=14$, sample mean $\bar{x} = 17.85$,

Sample S.D. (s) $= 1.955$

population mean, $\mu = 18.5$

Degree of freedom $= n-1 = 13$

1. Null hypothesis $H_0 : \mu = 18.5$
2. Alternative hypothesis $H_1 : \mu \neq 18.5$
3. $\alpha = 0.05$
4. Test Statistic : $t = \dfrac{\bar{x}-\mu}{S/\sqrt{n-1}} = \dfrac{17.85-18.5}{1.955\sqrt{13}}$

$= \dfrac{-0.65}{0.542} = -1.199$

$\therefore |t| = 1.199$

ie Calculated $t = 1.199$

5. Tabulated t at 5% level of significance for 13 degree of freedom for two tailed test $= 2.16$

Since Calculated $t <$ tabulated t, we accept H_0 at 5% level

$\therefore$ Result of the experiment is not significant

5) In two indepedent samples of sizes 8 and 10, the sum of squares of the sample values from the respective sample means were 84.4 and 102.6. Test whether the difference of variances of the population is significant or not. Use 5% level of significance

1: let σ_1^2 and σ_2^2 be the variances of the two populations from which the samples are drawn

1. let $H_0 : \sigma_1^2 = \sigma_2^2$
2. let $H_1 : \sigma_1^2 \neq \sigma_2^2$
3. $\alpha = 0.05 = 5\%$, $\Sigma(x_i-\bar{x})^2 = 84.4$, $\Sigma(y_i-\bar{y})^2 = 102.6$
4. $n_1 = 8$, $n_2 = 10$

Let s_1^2 and s_2^2 be the estimates of σ_1^2 and σ_2^2

then $s_1^2 = \dfrac{\Sigma(x_i-\bar{x})^2}{n_1-1}$, $s_2^2 = \dfrac{\Sigma(y_i-\bar{y})^2}{n_2-1} = \dfrac{102.6}{9} = 11.4$

$= \dfrac{84.4}{7} = 12.057$ $\quad = \cancel{\Sigma(y_i-\bar{y})^2}$

Since $s_1^2 > s_2^2$, test-statistic is

$$F = \frac{s_1^2}{s_2^2} = \frac{12.057}{11.4} = 1.057$$

ie Calculated $F = 1.057$

Degree of freedom is given by $\nu_1 = n_1 - 1 = 8 - 1 = 7$

and $\nu_2 = n_2 - 1 = 10 - 1 = 9$.

Tabulated value of F at 5% level for (7, 9) degree of freedom is 3.29

ie $F_{0.05}(7, 9) = 3.29$

Since Calculated F < tabulated F

we accept H_0 and

conclude that the populations have the same

Variance.