Avinash Gautam

Statistics Test

Page : 1

Date: 8/12/21

Q. 1)

Ans.

<u>Point Estimation</u> :- Point estimators are functions that are used to find an approximate value of population parameter from the random samples.

Exa. :- Average Height of the students in a university

Min, Max, Quantile etc.

(a) <u>Consistent statistics</u> :- Consistency means how close the value with the parameter. As the size of population increases, the statistic will become more consistent & accurate.

$$\lim_{n \to \infty} P(|T_n - \theta| > \varepsilon) = 0$$

(b) <u>Unbiased Statistics</u> :- An statistic is called unbiased if its expected value is equal to the value of parameter θ.

statistic $T_n = T\{x_1, \dots x_n\}$

$$E(T_n) = \theta$$

(c) <u>Sufficient Statistics</u> :- An statistic called sufficient if it holds all the ~~impor~~ information regarding the parameter of the population.

(d) Efficiency :- If two estimators T_1 & T_2 are unbiased & consistent then the efficiency E defines as

$$E = \frac{V(T_1)}{V(T_2)}$$

In simple words, An estimator with minimum variance among all other estimators ~~as~~ is called most efficient estimator.

Q.4) Given :- $n_1 = \cancel{4}\ 40$, $\bar{x}_1 = 9.1$, $s_1 = 1.9$
$n_2 = 50$, $\bar{x}_2 = 8.0$, $S_2 = 2.1$

Hypothesis :

$$\alpha = 0.05$$

$$H_0 : \mu_1 - \mu_2 = 0$$

$$H_1 : \mu_1 - \mu_2 \neq 0 \quad \text{(Two tail test)}$$

Test statistic:

$$z = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$$

$\because$ σ is unknown so $\sigma = s$

$$z = \frac{(9.1 - 8.0) - 0}{\sqrt{\frac{(1.9)^2}{40} + \frac{(2.1)^2}{50}}}$$

Page :

Date: / /

$Z = 2.604$

$Z_{0.025} = 1.96$ (From the table)

P-value : 0.0094

Conclusion :- $Z > Z_{0.025}$ and $P < \alpha$

H_0 must be rejected.

Q. 6 Hypothesis :-

$\alpha = 0.01$

$H_0 : \mu_1 - \mu_2 = 0$

$H_1 : \mu_1 - \mu_2 \neq 0$ (Two - Tail test)

Table :-

	Design 1		Design 2	
	x_1	$(x_1 - \bar{x}_1)^2$	x_2	$(x_2 - \bar{x}_2)^2$
	127	289	154	25
	168	576	135	196
	143	1	132	289
	165	441	171	484
	122	484	153	16
	139	25	149	0
Total	864	1816	894	1010

$n_1 = 6$ $\qquad$ $n_2 = 6$

$\bar{x}_1 = \frac{864}{6} = 144$ $\qquad$ $\bar{x}_2 = \frac{894}{6} = 149$

$S_1^2 = \frac{1816}{6}$ $\qquad$ $S_2^2 = \frac{1010}{6}$

$S_1^2 = 302.67$ $\qquad$ $S_2^2 = 168.33$

Test statistic :

Pooled estimator $s_p = \sqrt{\dfrac{(n_1-1)s_1^2 + (n_2-1)s_2^2}{n_1+n_2-2}}$

$$= \sqrt{\frac{(6-1)(302.67) + (6-1)(168.33)}{6+6-2}}$$

$$= \sqrt{\frac{302.67 + 168.33}{2}}$$

$$S_p = 15.346$$

Now $t_c = \dfrac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{s_p\sqrt{\dfrac{1}{n_1} + \dfrac{1}{n_2}}}$

$$= \frac{(144 - 149) - 0}{15.346\sqrt{\frac{1}{6} + \frac{1}{6}}}$$

$$t_c = \frac{-5}{8.747} = 0.57$$

$t_c = 0.57$ and $t_{.005,10} = 3.169$

Conclusion :- $t_c < t_{0.005,10}$ so H_0 can't be rejected.
We can say that difference of mean of these two designs is significant.

Q.② Problem of Interval Estimation :-

According to me, Interval estimation provides a range where parameter value may lie. but it will not give us an appropriate information about parameter and sometimes if ranges is larger then it is difficult to conclude the information regarding parameter.

Confidence Interval for ~~mean~~ Variance :-

let us consider a random variable x taken from the normal population with unknown mean μ & variance σ^2. Then we know that sample variance s^2 is an estimator for population variance σ^2. We use chi-square test to construct $100(1-\alpha)$ % confidence interval for population variance σ^2.

$$\chi^2_{n-1} = \frac{(n-1)s^2}{\sigma^2}$$

Since it is a squared estimation on $n-1$ degree of freedom so the two sided confidence interval for population variance is :

$$P\left[\chi^2_{\frac{1-\alpha}{2},n-1} \le \chi^2_{n-1} \le \chi^2_{\frac{\alpha}{2},n-1}\right] = 1-\alpha$$

$$P\left[\frac{(n-1)s^2}{\chi^2_{1-\alpha/2,n-1}} \le \sigma^2 \le \frac{(n-1)s^2}{\chi^2_{\frac{\alpha}{2},n-1}}\right] = 1-\alpha$$