5.3.2017

Stastical Techniques.

Q1. If $\bar{X}$ = mean of set n observations i.e. $x_1, x_2, x_3 \ldots x_n$
σ_x = standard deviation of set of n observations.
Similarly $\bar{Y}$ & σ_y are mean and standard deviation of set of m observation Show that standard deviation of pooled set
$x_1, x_2, x_3 \ldots x_n$, $y_1, y_2, \ldots y_m$ for n+m observation is

$$\sqrt{\frac{n\sigma_x^2 + m\sigma_y^2}{n+m} + \frac{nm}{(n+m)^2}(\bar{x}-\bar{y})^2}$$

Ans: The pooled set has mean $\bar{u}$ and Variance σ^2
Then

$$\bar{u} = \frac{n\bar{x} + m\bar{y}}{n+m}$$

$$\text{Also } (n+m)\sigma^2 = \sum_{j=1}^{n}(x_j - \bar{u})^2 + \sum_{j=1}^{m}(y_j - \bar{u})^2$$

$$= \sum_{j=1}^{n} x_j^2 + \sum_{j=1}^{m} y_j^2 - (n+m)\bar{u}^2$$

$$= (n\sigma_x^2 + n\bar{x}^2) + (m\sigma_y^2 + m\bar{y}^2) - (n+m)u^{-2}$$

$$= (n\sigma_x^2 + m\sigma_y^2) + \frac{nm}{n+m}(\bar{x}-\bar{y})^2$$

$$(n+m)\sigma = \sqrt{(n\sigma_x^2 + m\sigma_y^2) + \frac{nm}{n+m}(\bar{x}-\bar{y})^2}$$

Proved

Q.2. Consider a family with 2 children Assume each child is as likely to be a boy as it is to be a girl. what is the conditional probability that both children are boys given that (i) the elder child is a boy ii) at least one of the children is boy.

Ans: i) P(the elder child is a boy) = 1/2

P(both children are boys | the elder child is a boy)

$= P(b,b) / \{ P(b,b) + P(b,g) \}$

$= (1/4) / ((1/4) + (1/4))$

$= 1/2$ Ans

ii) P(at least one of the children is a boy) = 1 - P(both girls)

$= 3/4.$

P(both children are boys | at least one of the child is boy)

= P(both children are boys) / P(at least one of the child is boy)

$= (1/4)/(3/4) = 1/3$ Ans

Q.3. It has been claimed that in 60% of all android applications installations. the utility bill is reduced by at least one third Accordingly what is the probability that the utility bill we reduced by one third in
(i) Four or Five installations
(ii) at least Four or Five installations

Ans: Here the random variable follows binomial distribution with $p = 0.6$,
$r = 4$,
$n = 5$

(i) $P[X=4] = C(5,4)(0.6)^4(0.4)$
$= 0.259$

ii) We have to find the probability that X is at least 4. This probability is the Sum of the probabilities that $X = 4$ and $X = 5$, because 'at least 4 means 4 or more'
Thus we have to find $P[X=4] + P[X=5]$

$P[X=5] = (5,5)(0.6)^5$
$= 0.078$

$\therefore$ the required probability $= 0.259 + 0.078$
$= 0.337$ Ans.

Q. 4. If a hardware Support Centre receives on an average $\lambda = 6$ wrong Complaints per day then what is the probability that it will receive 4 wronge Complaints on any given day.

Ans; Since the problem deals with the receipts of wronge complaints. which is an event with rare occurence over an interval of time (a day in this case)
We Can apply Poisson distribution.
Since on an average 6 wronge complaints per day.
Substituting $\lambda = 6$ and $x = 4$ in the Poisson formula
we get $p(X=4) = \frac{6^4 e^{-6}}{4!} = \frac{1296 \times (0.0025)}{24}$

$= 0.135$ Ans.

Q5. Verify whether the following situation can be described by uniform distribution. or not.

i) The average life span of a bulb produced by manufacturing Company.

Ans: We Can model it as uniform.

(ii) The number of defective items produced by assembly process.

Ans; We cannot model it as uniform.

Q7. From a population of 200 observation a sample of $n = 50$, is selected Calculate the standard error if the population standard deviation equals 22

Ans: $S.E(\bar{x}) = \sqrt{\frac{N-n}{N-1}\left(\frac{\sigma^2}{n}\right)}$

given that

$N = 200$

$n = 50$

$\sigma = 22$

$S.E(\bar{x}) = \sqrt{\frac{200-50}{199}\left(\frac{22^2}{50}\right)}$

$= \frac{22}{\sqrt{50}} \times \sqrt{\frac{150}{199}}$

$= 3.111 \times 0.868$

$= 2.701$ Ans

Q6. An Individuals IQ score has $N(100, 15^2)$ distribution, Find the probability that the individuals IQ score is between 91 and 121.

Ans we require $P[91 < x < 121]$

Standardising gives

$P\left[\frac{91-100}{15} < \frac{x-100}{15} < \frac{121-100}{15}\right]$

The middle term is a standardised normal random Variable and so we have

$P\left[\frac{-9}{15} < z < \frac{21}{15}\right] = P[-0.6 < z < 1.4]$

$= 0.9192 - 0.2743$

$= 0.6449$ Ans.

Q8. A random sample of 100 observation is taken from a normal population having variance $\sigma^2 = 42.5$. Find the approximate probability of obtaining a sample standard deviation between 3.14 and 8.94.

Ans: We make use the Formula.

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

given that

$\sigma^2 = 42.5$

$n = 100$

$s_1^2 = 3.14$

$s_2^2 = 8.94$

$\chi^2 = ?$

$$\chi_1^2 = \frac{(n-1) \times 3.14}{42.5}$$

$$= \frac{99 \times 3.14}{42.5} = 7.314 \quad \text{—} \; (1)$$

$$\chi_2^2 = \frac{99 \times 8.94}{42.5} = 20.824 \quad (2)$$

To find the required approximate probability it is enough to find area as given

From the table we find those values of χ^2 for 99 degrees freedom which are close to

7.314 and 20.824. Ans

Q.9. If independent random samples of size $n_1 = n_2 = 8$ comes from normal populations having the same variance what is the propability that either sample variance will be at least seven times as large as the other?

Ans: given that

$n_1 = n_2 = 8$

$\therefore V_1 = V_2 = 7$

and $F = \dfrac{S_1^2}{S_2^2} = 7$

Proceeding similarly we find the value 6.99 in Table 5 which is almost equal to 7.

There for the probability is = 0.01 Ans

Q.10. A random sample of 800 computers contains 24 defective items compute 99% confidence interval for the proportion of defective computers.

Ans:

Q12:- A random sample of size 1000 from lot of computer supplied by manufacturer. 1, contains 20 defectives and a random sample of size 1500 from computers supplied by manufacturer 2 contains 40 defectives If $\alpha = 0.05$ can you say computer supplied by manufacturer 1 are better than those supplied by manufacturer 2.

Ans:- $n_1 = 1000, \quad n_2 = 1500, \quad \alpha = 0.05$

$$p_1 = \frac{20}{1000} = 0.02$$

$$p_2 = \frac{40}{1500} = 0.0267$$

$$p = \frac{60}{2500} = 0.024$$

$H_0 : \pi_1 = \pi_2$

$H_1 : \pi_1 < \pi_2$

$$|Z| = \left| \frac{0.02 - 0.0267}{\sqrt{0.024 \times 0.976\left(\frac{1}{1000} + \frac{1}{1500}\right)}} \right| = 1.0723$$

1-tailed limit for $\alpha = 0.05$ is 1.64

$\therefore$ We accept H_0.

Ans

Q.16. A sample of size 4 is to be selected from a population of 11 computer brands. list all the possible sample by i) linear systematic sampling ii) circular systematic sampling

Ans. i) $N = 11$, $n = 4$, $K = 11/4 = 3$ (approx)

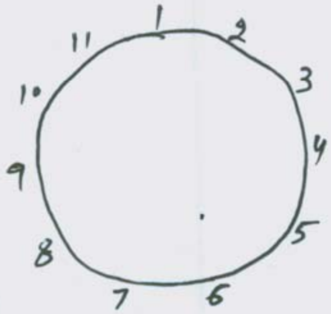
Arranging the units in 4 rows of 3 columns each we get tables as follows:

1	2	3
4	5	6
7	8	9
10	11	

Selected a number r between 1 and 3, possible samples are 1, 4, 7, 10; 2, 5, 8, 11; 3, 6, 9 of size 4 or 3

ii) Here $N = 11$, $n = 4$, $k = 3$

Consider fig:



let the random start be 2. Then sample selected is 2, 5, 7, 10. If random start be 5 then sample selected is 5, 8, 11, 3. likewise you can remaining 9 samples as the total number of possible samples is $N = 11$.

9.3.2017

BCS·040

(10)

Q.14. An economist wants to estimate relationship in a Small comunity between a family's annual income and amount that family saves.
Fallowing data of 9 families obtained.

Annual income thousand dollar.	12	13	14	15	16	17	18	19	20
Annual saving thousand dollar	0	0.1	0.2	0.2	0.5	0.5	0.6	0.7	0.8

Calculate the least regression line.

Ans: letting x_i be the income of the ith family and y_i be the saving of the ith family.
We find that
$$\sum_{i=1}^{9} x_i y_i = 63.7,$$
$$\sum_{i=1}^{9} y_i = 3.6$$
$$\bar{y} = 0.4,$$
$$\sum_{i=1}^{9} x_i^2 = 2364,$$
$$\sum_{i=1}^{9} x_i = 144.$$
$$\bar{x} = 16.$$

Thus Substitution these values in the alternate formula for $\hat{b}$ we obtain
$$\hat{b} = \frac{9(63.7) - (144)(3.6)}{9(2364) - 144^2}$$
$$= \frac{573.3 - 518.4}{21.276 - 20.736} = 0.1017$$

Consequently

$$\hat{a} = \bar{y} - b\bar{x} = 0.4 - 0.1017(16)$$
$$= -1.2272.$$

Thus the regression line is

$$y = -1.2272 + 0.1017x$$ Ans.

Q.13. A Software Company owner has 3 developer A, B and C During a particular week the owner try to evaluate the productivity of A, B and C.

Prepare the summary table and present the AnovA table

Ans:

Day	Developer		
	A	B	C
1	6	7	9
2	8	8	8
3	6	7	*
4	5	5	6
5	*	8	*

Summary	Developer			Total
	A	B	C	
n_i	4	5	3	12
Total	25	35	23	83
Average	6.25	7	7.66	21.41
Σy_{ij}^2	161	251	181	593

10.3.17.

BCS-040

$n = 12.$

ANOVA for Developers Data

	DF	SS	MS	F
Developer	2	3.50	1.75	1.5
Error	9	15.4166	1.7129	
Total	11	18.9166		

Tabulate F-Value with 2 and 9 df at 5% level is equal to 4.256. There is no substancial evidence to say that the tailors are not equally productive

ANOVA for fabric strength data

	DF	SS	MS	F
Developers.	2	81.5	40.75	29.34
Error	9	12.5	1.3888	
Total	11	94.0		

The estimate. of Variance Components are given by

$\sigma^2 = 1.3888$ and $\sigma_r^2 = \dfrac{40.75 - 1.3888}{4}$

$= 9.8403$ Ans.