

PSG COLLEGE OF ARTS & SCIENCE
(AUTONOMOUS)
MSc DEGREE EXAMINATION DECEMEBR 2024
(First Semester)

Branch - STATISTICS

SAMPLING THEORY

Time: Three Hours

Maximum: 75 Marks

SECTION-A (10 Marks)

Answer ALL questions

ALL questions carry EQUAL marks

(10 × 1 = 10)

Module No.	Question No.	Question	K Level	CO
1	1	When an investigator wants a sample containing 'm' units which possess a rare attribute, the appropriate sampling procedure is (a) Simple random sampling (b) Systematic sampling (c) Cluster sampling (d) Stratified random sampling	K1	CO1
	2	Stratification means division into _____. (a) Groups (b) Layers (c) classes (d) sub groups	K2	CO2
2	3	Which of the following is the most efficient random sampling technique? (a) Simple random sampling (b) systematic sampling (c) cluster sampling (d) stratified random sampling.	K1	CO1
	4	Let y_{ij} denote the observation on j^{th} unit of the i^{th} sample, ($i=1,2,...,k; j=1,2,...,n$) then the mean of the i^{th} systematic sample is _____. (a) $\frac{1}{n} \sum_{j=1}^n y_{ij}, (i=1,2,...,k)$ (b) $\frac{1}{n} \sum_{j=1}^n x_{ij}, (i=1,2,...,k)$ (c) $\frac{N}{n} \sum_{j=1}^n y_{ij}, (i=1,2,...,k)$ (d) $\sum_{j=1}^n x_{ij}, (i=1,2,...,k)$	K2	CO2
3	5	Cluster sampling is most useful when: (a) The population is small and concentrated (b) The population is large and geographically dispersed (c) All members of the population have the same characteristics (d) The cost of sampling is not a concern	K1	CO1
	6	The optimum cluster size in cluster sampling is primarily determined by: (a) The total population size (b) The homogeneity within clusters and the cost of sampling (c) The number of clusters available (d) The sampling frame used	K2	CO2
4	7	Two stage sampling design is more efficient than single stage sampling if the correlation between units in the first stage is (a) Positive (b) Negative (c) No symbol (d) zero	K1	CO1
	8	In two-stage sampling, the second stage involves: (a) Selecting a sample from the entire population directly (b) Dividing the population into strata and sampling from each (c) Taking a sample from selected primary units (d) Randomly selecting individuals from the population.	K2	CO2
5	9	The ratio estimator is typically used when: (a) The variable of interest is categorical (b) There is a strong linear relationship between the study Variable and an auxiliary variable (c) The sample size is too large to manage (d) The population is evenly distributed.	K1	CO1

Cont....

5	10	A regression estimator is most effective when: (a) There is no relationship between the auxiliary variable and the study variable (b) There is a linear relationship between the auxiliary variable and the study variable (c) The auxiliary variable is randomly distributed (d) The sample size is small.	K2	CO2
---	----	---	----	-----

SECTION - B (35 Marks)

Answer ALL questions

ALL questions carry EQUAL Marks

(5 × 7 = 35)

Module No.	Question No.	Question	K Level	CO
1	11.a.	For a stratified random sampling, show that $Var_{opt} \leq Var_{prop} \leq Var_{SR}$.	K1	CO1
		(OR)		
	11.b.	Describe the construction of number of strata.		
2	12.a.	Prove that the systematic sample mean $\bar{y}_{sys}$ is more efficient than simple random sample mean $\bar{y}$ if $S_w^2 > S^2$	K2	CO2
		(OR)		
	12.b.	Describe the auto correlated population.		
3	13.a.	Show that the simple arithmetic mean given by the relation $\bar{y}_n = \sum_i \frac{\bar{y}_i}{n}$ is not an unbiased estimator. The bias and sampling variance of the estimator are given by $B(\bar{y}_n) = -\frac{cov(\bar{y}_i, M_i)}{M}$ and $V(\bar{y}_n) = \frac{(1-f)}{n} S_b^2$.	K3	CO3
		(OR)		
	13.b.	Explain the relative efficiency of unequal cluster sampling.		
4	14.a.	Prove that if 'n' first stage units and 'm' second stage units from each chosen first stage units are selected by simple random sampling, without replacement, $\bar{y}$ with sampling variance $V(\bar{y}) = \frac{(N-n) S_b^2}{N} + \frac{(M-m) S_w^2}{M mn}$	K3	CO3
		(OR)		
	14.b.	Show that the estimator given by the relation $\bar{y}_1 = \sum_i \frac{\bar{y}_i}{n}$ is biased and its bias is given by $B = -\sum_i (M_i - \bar{M}) \frac{\bar{y}_i}{NM}$ and sampling variance by $V(\bar{y}_1) = (1-f_1) \frac{S_b^2}{n} + \frac{1}{nN} \sum_i (1-f_{2i}) \frac{S_{w_i}^2}{m_i}$ Where $S_b^2 = \sum_i \frac{(\bar{y}_i - \bar{Y})^2}{N-1}$ and $S_{w_i}^2$ as usual.		
5	15.a.	Show that to the first order approximation the variance of $\hat{R}$ can be expressed as $V(\hat{R}) = \left\{ \frac{(1-f)}{nX^2} [S_y^2 + R^2 S_x^2 - 2\rho R S_x S_y] \right. \\ \left. \frac{(1-f)R^2}{n} [C_y^2 + C_x^2 - 2\rho C_x C_y] \right\}$	K4	CO5
		(OR)		
	15.b.	Discuss the variance of difference of regression estimators.		

Cont....

SECTION -C (30 Marks)

Answer ANY THREE questions

ALL questions carry EQUAL Marks

(3 × 10 = 30)

Module No.	Question No.	Question	K Level	CO																																																											
1	16	200 cultivators' holding in Uttar Pradesh were stratified according to their sizes. The number of holdings (N_i), mean area under wheat per holdings ($\bar{Y}_i$) and standard deviation of under wheat per holding (S_i) are given below each stratum.	K1	CO1																																																											
		Stratum number			1	2	3	4	5	6	7																																																				
		Number of holdings(N_i)			394	461	381	334	169	113	148																																																				
		mean area under wheat per holdings ($\bar{Y}_i$)			5.4	16.3	24.3	34.5	42.1	50.1	63.8																																																				
		standard deviation of under wheat per holding (S_i)			8.3	13.3	15.1	19.8	24.5	26.0	35.2																																																				
		For a sample of 200 farms, compute the sample size in each stratum under proportional and optimum allocation. Calculate the sampling variance of the estimated area under wheat from the sample.																																																													
2	17	Compare systematic sampling with stratified random sampling.	K2	CO2																																																											
3	18	Illustrate the optimum cluster size sampling for proportions.	K3	CO3																																																											
4	19	At an experimental station, there were 100 fields sown with wheat. Each field was divided in to 16 plots of equal size ($1/16^{\text{th}}$ hectare). Out of 100 fields, 10 were selected by simple random sampling wor. From each selected field, 4 plots were chosen by random sampling wor. The yields in kg/plot are given below.	K3	CO4																																																											
		<table><tr><th rowspan="2">Selected field</th><th colspan="4">Plots</th></tr><tr><th>1</th><th>2</th><th>3</th><th>4</th></tr><tr><td>1</td><td>4.32</td><td>4.84</td><td>3.96</td><td>4.04</td></tr><tr><td>2</td><td>4.16</td><td>4.36</td><td>3.50</td><td>5.00</td></tr><tr><td>3</td><td>3.06</td><td>4.24</td><td>4.76</td><td>3.12</td></tr><tr><td>4</td><td>4.00</td><td>4.84</td><td>4.32</td><td>3.72</td></tr><tr><td>5</td><td>4.12</td><td>4.68</td><td>3.46</td><td>4.02</td></tr><tr><td>6</td><td>4.08</td><td>3.96</td><td>3.42</td><td>3.08</td></tr><tr><td>7</td><td>5.16</td><td>4.24</td><td>4.96</td><td>3.84</td></tr><tr><td>8</td><td>4.40</td><td>4.72</td><td>4.04</td><td>3.98</td></tr><tr><td>9</td><td>4.20</td><td>4.66</td><td>3.64</td><td>5.00</td></tr><tr><td>10</td><td>4.28</td><td>4.36</td><td>3.00</td><td>3.52</td></tr></table>			Selected field	Plots				1	2	3	4	1	4.32	4.84	3.96	4.04	2	4.16	4.36	3.50	5.00	3	3.06	4.24	4.76	3.12	4	4.00	4.84	4.32	3.72	5	4.12	4.68	3.46	4.02	6	4.08	3.96	3.42	3.08	7	5.16	4.24	4.96	3.84	8	4.40	4.72	4.04	3.98	9	4.20	4.66	3.64	5.00	10	4.28	4.36	3.00	3.52
		Selected field				Plots																																																									
					1	2	3	4																																																							
		1			4.32	4.84	3.96	4.04																																																							
		2			4.16	4.36	3.50	5.00																																																							
		3			3.06	4.24	4.76	3.12																																																							
		4			4.00	4.84	4.32	3.72																																																							
		5			4.12	4.68	3.46	4.02																																																							
		6			4.08	3.96	3.42	3.08																																																							
		7			5.16	4.24	4.96	3.84																																																							
		8			4.40	4.72	4.04	3.98																																																							
		9			4.20	4.66	3.64	5.00																																																							
10	4.28	4.36	3.00	3.52																																																											
Estimate the wheat yield per hectare for the experimental station along with its standard error.																																																															
5	20	Discuss the comparison of mean per unit and ratio estimators.	K4	CO5																																																											

Z-Z-Z

END