UNIVERSITY OF THE PUNJAB

Part-I : Supplementary Examination 2018

Examination:- M.A./M.Sc.

Roll No. .....

Subject: Statistics

PAPER: I (Statistical Methods)

MAX. TIME: 3 Hrs.

MAX. MARKS: 100

NOTE: Attempt any FIVE questions. All questions carry equal marks.

Q.1.a) i) In airport luggage screening it is known that 3% of people screwed have questionable objects in their luggage. What is the probability that a string of 15 people pass through successfully before an individual is caught with a questionable object? What is the expected number in a row that pass through before an individual is stopped. (4)

b) A soft-drink machine is regulated so that it discharges an average of 200 millimeters per cup. If the amount of drink is normally distributed with a standard deviation equal to 15 millimeters. (12)

i) What fraction if cups will contain more than 224 millimeters.

ii) What is the probability that a cup contains between 191 and 209 millimeters.

iii) How many cups will probably overflow if 230 millimeters cups are used for next 1000 drinks?

iv) Below what value do we get the smallest 25% of the drinks?

c) An absent-minded professor does not remember which of his twelve keys will open his office door. If he tries at random and with replacement. (4)

i) On average how many keys should he try before unlocking his office.

ii) What is the probability that he open his office door after only three tries?

Q.2.a) A marketing expert for a pasta-making company believes that 40% of pasta lovers prefer lasagna. If 9 out of 20 pasta lovers chose lasagna over other pastas What can be concluded about the experts claims. Use a 0.05 level of significance. (8)

b) The following data represents the running times of films produced by 2 motion-picture companies: (12)

Company	Time (minutes)						
1	102	86	98	109	92		
2	81	165	97	134	92	87	114

Test the hypothesis that average running time of films produced by Company 2 exceeds the average running time of films produced by company 1 by 10 minutes, against the one-sided alternative that difference is less than 10 minutes. Use a 0.01 level of significance and assume the distributions of times to be approximately normal with universal variances.

Q.3.a) Explain the relationship between Poisson process and exponential distribution. (3)

b) A fair die is tossed 180 times with the following results: (7)

x	1	2	3	4	5	6
f	28	36	36	30	27	23

Is this a balanced die? Use a 0.01 level of significance.

c) For the data given below: (10)

Sample I	4	7	6	6		
Sample II	5	1	3	5	3	4
Sample III	3	8	6	8	9	5

Use a Bartlett test to test the hypothesis that variances of three populations are equal.

Q.4.a) Use fisher exact test to test the hypothesis that inoculation is independent of immunity from attack among a population exposed to a certain disease from the following data: (8)

	Inoculated	Not – Inoculated
Attacked	9	2
Not-Attacked	7	6

b) A new cure has been developed for a certain type of cement that results in a compressive strength of 5000 kilograms per square centimeters and a standard deviation of 120. To test the hypothesis that $\mu = 5000$ against the alternative that $\mu < 5000$, a random sample of 50 pieces of content is tested. The critical region is defined to be $\bar{X} \leq 4970$. (12)

i) Find the probability of commuting a type I error when H_0 is true.

ii) Evaluate β for alternatives $\mu = 4960$ and $\mu = 4970$.

Q.5.a) For the following data: (15)

Y	110	135	120	120	140	130	135	150
X_1	30	60	60	62	62	62	64	64
X_2	40	20	30	20	30	40	20	30

i) Calculate a, b_1, and b_2.

ii) Estimate $\mu_{Y.3,1,2,4}$ with 99% confidence interval.

iii) Estimate $\sigma^2_{y.x_1x_2}$ with 95% confidence interval.

b) State the assumptions of linear regression line. (5)

Q.6.a) Explain the sequential testing of hypothesis for proportion. A null hypothesis $H_0: p_1 - p_2 \leq 0.1$ is to be tested at $\alpha = 5\%$ and it is decided to have $\beta = 0.10$ if $H_0: p = 0.30$ is true. (8)

Sketch a sequential probability ratio test.

b) Prove Brandt – snedecor formula for chi-square. (6)

c) In an experiment to study the dependence of hypotension on smoking habits, the following data were taken on 180 individuals. (6)

Hypertension	Non Smoker	Moderate Smokers	Heavy Smokers
	21	36	30
No Hypertension	48	26	19

Test the hypothesis that pressure or absence of hypertension is independent of smoking habits. Use 0.05 level of significance.

Q.7.a) Write a note on Parametric and Non – Parametric tests. (6)

b) Discuss the procedure of Kolmogorov's smirnov test for two samples. (4)

c) The following data show the Kidney weights in grows of 30 dogs prior to their use in experiments. (10)

58	78	84	90	97	70	90	86	82
59	90	70	74	83	90	76	88	84
68	93	70	74	70	110	67	68	75
80	68	82	104	92	112	84	98	80

Do the data provide enough evidence that they have come from normally distributed population with a mean of 85 grows and a standard deviation of 15 grows. Use Kolmogorov – smirnov test.

Q.3 a) Use the Mann-Whitney U-test to test the hypothesis that the difference between mean scores of students in authentic computations in two types of schools are equal, using $\alpha = 0.05$ (12)

Given the data

	0 – 9	10 – 19	20 – 29	30 – 39	40 – 49	50 – 59
Residential	1	3	10	63	38	5
Non-Residential	4	7	25	37	13	4

b) A true-false examination was constructed with answers running in the following sequence. (8)

T F F T F T F T T F F T F F T F T F T T

Does this sequence indicate a departure from randomness in the arrangements of T and F answers.

Q.5. a) Obtain $\Delta^n\ (x$ when (4)

$(x = (1 - \alpha x)\ (1 - \beta x^2)(1 - \alpha x^3)(1 - \beta x^4)$

b) Calculate the values of $f(x)$ corresponding to $x = 7$ (12)

Given Values

x	16	17	34	35
$f(x)$	68.7	64	44	39.1

c) Define by means of a table. (4)

i) Forward differences

ii) Divided differences

UNIVERSITY OF THE PUNJAB

Part-I : Supplementary Examination 2018

Examination:- M.A./M.Sc.

Roll No. .....

Subject: Statistics

PAPER: II (Probability and Probability Distributions)

MAX. TIME: 3 Hrs.

MAX. MARKS: 100

NOTE: Attempt any FOUR questions. All questions carry equal marks.

Q.1.a) Write short notes on (10)

i) Independence of event
ii) Indicator function
iii) Permutation and Combination
iv) Cumulative distribution function and its properties
v) Mutually exclusive events

b) A father, mother, 2 boys, and 3 girls are asked to line up for a photograph. Determine the number of ways they can line up if (5)
i) there are no restrictions
ii) the parents stand together
iii) the parents do not stand together

c) A father buys 9 different toys for his four children. In how many ways can he give one child 3 toys and the remaining 3 children two toys each? (5)

d) Mr. John has 10 books that he is going to put on his bookshelf. Of these 4 are mathematics, 3 are chemistry, 2 are history books and 1 is language book. John wants to arrange his books so that all the books dealing with the same subjects are together on the shelve. How many different arrangements are possible? (5)

Q.2.a) State and prove Baye's theorem. (8)

b) There are "n" Socks, 3 of which are red in a drawer. What is the value of "n". When 2 of the socks are chosen randomly. Where the probability that they both are red is 1 / 2 . (5)

c) From a faculty of 6 Professors, 6 Associate Professors,6 Assistant Professors and 12 instructors , a committee of size 6 is formed randomly. What is the probability that there is at least one person from each rank in the committee? (5)

d) If 4 married couples are seated at random in a row. Compute the Probability that (7)
i. At least one couple sits together
ii. No wife sit next to her husband

Q.3.a) Derive Poisson distribution as limiting form of the binomial distribution (9)

P.T.O.

b) Find the probability generating function and moment generating function of the Poisson distribution and show that its cumulants are equal. (8)

c) Find mean and variance of Hypergeometric distribution. (8)

Q.4.a) If the p.d.f. of X is given by (7)

$$f(x)=\begin{cases}630x^4(1-x^4) & 0<x<1\\ 0 & \text{otherwise}\end{cases}$$

Find the probability that it will take a value within two standard deviations of the mean and compare this probability with the lower bound provided by Chebychev's inequality.

b) Find the mean and variance of Weibull Distribution (9)

c) Define lognormal distribution and obtain its rth moment about origin. If $\log_e X$ is normally distributed with $\mu=1, \sigma^2=4$ find $P(1/2<x<2)$. (9)

Q.5.a) Assume that $X_1, X_2 \ldots\ldots, X_n$ are independent r.v's and $X_i \sim N(\mu_i, \sigma_i^2)$. Find the p.d.f of $Y=\sum a_i X_i$ where a_i's are constants (5)

b) Derive χ^2-distribution. (20)

Q.6.a) Define Bivariate Normal distribution and prove that its probability density function is a proper probability function. (7)

b) Let $\underset{\sim}{X}=(X_1, X_2, X_3)'$ have a multivariate normal distribution with mean vector (9)

$$E(Y)=\begin{bmatrix}2 & 2 & 3\end{bmatrix}' \text{ and covariance matrix } \sum=\begin{bmatrix}1 & 1 & 1\\ 1 & 2 & 2\\ 1 & 2 & 3\end{bmatrix}$$

Find the conditional distribution of $(X_1, X_2)'$ given (X_3).

c) Three fair coins are tossed let X denote the number of heads and Y denote the number of runs of heads (9)
i) Find the joint distribution of X and Y
ii) Find E(Y / X = 2)

Q.7.a) A sample of odd size $n=2r+1$ is taken from the rectangular distribution $f(x)=1 \quad 0\le x\le 1$ (10)

The median is $(r+1)th$ member of the sample arranged in ascending order. Find the variance of the median.

b) If the joint p.d.f. of X_1 and X_2 is given by (10)

$$f(x_1, x_2)=\begin{cases}e^{-(x_1+x_2)} & x_1>0,\ x_2>0\\ 0 & \text{otherwise}\end{cases}$$

Find the p.d.f. of $Y=\dfrac{X_1}{X_1+X_2}$.

c) State and prove Law of Large numbers. (5)

UNIVERSITY OF THE PUNJAB

Part-I : Supplementary Examination 2018

Examination:- M.A./M.Sc.

Roll No. .....

Subject: Statistics

PAPER: III (Design and Analysis of Experiments)

MAX. TIME: 3 Hrs.

MAX. MARKS: 100

NOTE: Attempt any FOUR questions. All questions carry equal marks.

Q.1 a) Discuss the importance of Randomization, Replication and Local Control in designing an experiment. (15)

b) Explain what is meant by the terms fixed-effect model and random-effect model in the analysis of variance. (10)

Q.2 a) Derive formula for estimating 2 missing values in RCB design when values are missing in different blocks but for same treatments. What changes will occur in ANOVA table after estimating missing observations. (15)

b) Why a Latin square design is sometimes referred to as a double blocking? Also explain what is a reduced Latin square design? (10)

Q.3 a) What are the uses and assumptions of co-variance analysis? What factors are considered in interpreting the results of the analysis? (10)

b) If A, B and C are the three methods of teaching, and X and Y stand for original spelling performance and later spelling performance of each of the four students allocated to systems A, B and C as tabulated below, set up the the table of analysis of covariance. What conclusions could be drawn from the table of analysis of covariance set up by you? (15)

A		B		C	
X	Y	X	Y	X	Y
3	10	4	12	1	6
2	8	3	12	2	5
1	8	3	10	3	8
2	11	5	13	1	7

Q.4 a) Write the difference between Yates technique and sign table method for computing contrasts by giving examples. (10)

b) The following data is the result of an experiment in a RCB design involving five rates of nitrogen fertilizer, three verities and four replications. The total of four replications for the 15 treatment combination is shown in the table. The total SS and Blocks SS are 53.53 and 2.599 respectively. Analyzed the data and draw conclusions. (15)

	N_0	N_1	N_2	N_3	N_4
V_1	12.496	18.894	17.838	22.868	23.292
V_2	14.192	19.224	21.744	22.958	22.722
V_3	15.112	19.014	21.280	22.210	23.546

Q.5 a) What do you know by the term confounding? Also explain in detail what is partial confounding? (10)

P.T.O.

b) Complete ANOVA table for the following factorial experiment. ABC is completely confounded (15)

Replicate I		Replicate II	
BI	BII	BI	BII
(1)= 10	a=6	ab= 8	a= 6
ab= 7	b= 4	(1)= 12	abc= 8
ac= 6	c= 9	ac= 7	b= 6
bc= 8	abc=5	bc= 6	c= 7

Q.6 a) Explain how a split-plot experiment differs from a factorial experiment. (10)

b) In an experiment on rice, there are 4 methods of irrigation to be compared, on main plot, with 3 fertilizer mixtures on subplots. Two complete blocks were used. The total SS, Block SS and Whole plot Error Ss are 3.0200, 0.0004 and 0.6509 respectively. The following table shows the sum of two blocks. Complete the ANOVA Table And draw your conclusion. (15)

	I_1	I_2	I_3	I_4
F_1	4.68	4.34	3.78	4.67
F_2	5.02	4,91	4.01	4.67
F_3	6.00	5.16	4.10	5.45

i. Find the Standard Error of the difference between two Fertilizer means.

ii. Find the Standard Error of the difference between two Irrigation means

Q.7 a) Define partially balanced incomplete block design. Describe the parameters specifying PBIBD with 2 associate classes. (10)

b) The effect on current flow of four treatment applied to the cols of TV tube filaments are studied. The days were taken as blocks. Due to some constraints it is not possible to run all the treatments on the same day, therefore incomplete blocks design is used. Analysis the data and draw conclusions. (15)

Blocks	**Treatment**			
(days)	A	B	C	D
1	12	-	30	17
2	-	42	42	13
3	14	23	41	-
4	10	33	-	21

UNIVERSITY OF THE PUNJAB

Part-I : Supplementary Examination 2018

Examination:- M.A./M.Sc.

Roll No. .....

Subject: Statistics

PAPER: IV (Sampling Techniques)

MAX. TIME: 3 Hrs.

MAX. MARKS: 100

NOTE: Attempt any FIVE questions. All questions carry equal marks.

Q#1 (a)	Why it is important to reduce bias in statistical investigation? Discuss the sources of bias.	10
(b)	What is simple random sampling? Explain the procedure of simple random sampling by the help of an example.	10
Q#2(a)	If the loss function due to an error in $\bar{y}$ is $\lambda\|\bar{y} - \bar{Y}\|$ and the cost function is $C = C_0 + C_1 n$, then show that the most economical value of 'n' in simple random sampling, ignoring finite population correction is $(\frac{\lambda S}{C_1\sqrt{2\pi}})^{2/3}$	08
(b)	If the terms in $\frac{1}{N_h}$ are ignored relative to unity, show that for estimated mean from stratified random sample of size n_h, $V_{opt} \leq V_{prop} \leq V_{ran}$ Where the optimum allocation is for fixed 'n'.	12
Q#3 (a)	Describe the different ways of allocation of sample size in strata.	10
(b)	A sampler proposes to take a stratified random sample. He expects that his field costs will be of the form $\sum c_h n_h$. His advance estimates of relevant quantities for the two strata are as follows.	10

Stratum	W_h	S_h	C_h
1	0.4	10	\$4
2	0.6	20	\$9

a. Find the values of $\frac{n_1}{n}$ and $\frac{n_2}{n}$ that minimize the total field cost for a given value of $V(\bar{y}_{st})$.

b. Find the sample size required, under the optimum allocation, to make $V(\bar{y}_{st}) = 1$. Ignore the fpc.

c. How much will the total field cost be?

Q#4(a)	Define systematic sampling. Describe the limitations of systematic sampling procedure.	10
(b)	Show that the mean of a systematic sample is more precise than the mean of a simple random sample if and only if $S^2_{wsy} > S^2$	10

P.T.O.

Q#5(a)	What is a ratio type estimator? Also describe if $\hat{R}$ is a biased estimator of R then why we prefer it?	12
(b)	Obtain the condition upon ρ for which $V(\hat{Y}_R)$ is smaller than $V(\hat{Y})$.	08
Q#6 (a)	For simple random sampling in which b_0 is a pre assigned constant, show that the linear regression estimate $\bar{y}_{lr} = \bar{y} + b_0(\bar{X} - \bar{x})$ is an unbiased estimate of $\bar{Y}$ with variance $$V(\bar{y}_{lr}) = \frac{1-f}{n}(S_y^2 - 2b_0 S_{yx} + b_0^2 S_x^2)$$	10
(b)	Compare combined and separate regression estimates. Under what condition they are equally good.	10
Q#7(a)	What is non-response error? Also describe the sources of non-response.	08
(b)	If S_b^2 is the variance between units in the population and $S_w^2 = AM^g \quad (g > 0)$ is the variance between elements that lie in the same unit, find the optimum value of M, the size of the unit.	12
Q#8 (a)	What is two-phase sampling? Also describe a practical example in which two-phase sampling can be used.	10
(b)	If $\pi_i > 0\ (i = 1,2,\ldots\ldots,N)$ then show that $\hat{y}_{HT} = \sum_{i=1}^{n} \frac{y_i}{\pi_i}$ is an unbiased estimator of population total Y with variance $$V(\hat{Y}_{HT}) = \sum_{i=1}^{N}\sum_{j>i}^{N}(\pi_i\pi_j - \pi_{ij})[\frac{y_i}{\pi_i} - \frac{y_j}{\pi_j}]^2$$ where π_{ij} is the probability that units i and j both are in the sample.	10
Q#9	Write a short note on the following: i. Two way stratification ii. Inverse sampling iii. Design Effect iv. Unequal Probability Sampling	5 each