UNIVERSITY OF THE PUNJAB

Part-I A/2018
Examination:- M.A./M.Sc.

Roll No. .....

Subject: Statistics
PAPER: I (Statistical Methods)

TIME ALLOWED: 3 hrs.
MAX. MARKS: 100

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

Q.1.a) Suppose that airplane engines operate independently and fail with probability equal to 0.4. Assuming that a plane makes a safe flight if at least one-half of its engines run, determine whether a 4-engine plane or a 2- engine plane has the higher probability for a successful flight.

b) If the probability is 0.4 that a child exposed to a certain contagious disease will catch it, what is the probability that tenth child exposed to the disease will be third to catch it.

c) The average number of trucks arriving on any one day at a truck depot in a certain city is known to be 12. What is the probability that on a given day fewer than 9 trucks arrive at this depot.

(8+6+6)

Q.2.a) Suppose that the life of a certain type of electronic component has an exponential distribution with a mean life of 500 hours. If x denotes the life of this component. Suppose that this component has been in operation for 300 hours. Find the probability that it will last for another 600 hours.

b) The average life of a certain type of small motors is 10 years with a standard deviation of 2 years. The manufacturer replaces free all motor that fail while under guarantee. If he is willing to replace only 3% of the motors that fail, how long a guarantee should be offer?

c) Assume that no of deaths by suicide in a city is a Poisson process with parameter 2 per day. Let X is the no of deaths by suicide that would occur in a week. Find the probability that there would be at least 10 deaths by suicide.

(8+8+4)

Q.3.a) If a (person) population has σ = \$500, how many observations would be needed to estimate is within \$100 with 95% confidence.

b) A geneticist is interested in the proportion of African males that have a certain minor blood disorder. In a random sample of 100 African males, 24 are found to be afflicted.

i) Compute a 99% confidence interval for the proportion of African males who have this blood disorder.

ii) What can we assert with 99% confidence about the possible size of our error if we estimate the proportion of African males with this blood disorder to be 0.24?

c) Explain

i) Confidence interval and testing of hypothesis
ii) Simple and composite hypothesis
iii) Chebyshev's inequality
iv) Power of test

(4+8+8)

Q.4.a) State the assumptions of regression line.

b) Consider the following set of data:

X	320	330	335	338	342	349	355	358	360	369
Y	24	27	29	32	31	35	30	42	40	52

i) Fit estimated regression line. Also calculate variance of estimate.

ii) Obtain ANOVA table and test the hypothesis that there is no association between regression and response variable.

iii) Test β=0 with t-test at 1% level of significance.

(5+15)

Q.5.a) Describe how would you test the equality of k (k>2) variances.

b) Each of the sets of observations is a random sample drawn from normal populations. (A, B, C, D):

PTO

A	48	40	47	75	58			
B	50	40	50	40	70			
C	44	50	30	40	50			
D	60	55	63	64	52	42	58	54

Use Bartlett test to test the hypothesis of equal variances.

(5+15)

Q.6.a) How parametric test differ from nonparametric test. Discuss advantages and disadvantages of non-parametric test.

b) Two varieties of tomato were experimented with concerning their fruit producing abilities, measured in pounds, the following data were obtained.

Variety A	3.03	3.10	2.35	3.86	3.92	1.71	2.60	2.3	2.70
Variety B	2.08	3.63	2.17	3.56	3.73	1.80	1.40	1.80	2.76

Apply. (i) Sign test, (ii) Wilcoxon signed rank test at 0.05 test of significance to test the hypothesis that there is no difference in fruit-producing abilities of two varieties.

(6+14)

Q.7.a) The following information reflect the results of tensile strength of the modeled parts produced by different methods (A,B,C,D):

A	80	88	85	86	90	85	88			
B	99	91	98	96	99	96	92	98		
C	89	82	81	80	85	86	86	84		
D	75	77	78	76	73	71	80	75	80	90

Use Kruskall-Wallis test to determine whether the mean tensile strengths are same for all methods.

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

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

Is there departure from arrangements of T and F answers.

(12+8)

Q.8.a) Prove that

$$\chi^2 = \frac{(ad - bc)^2 (a + b + c + d)}{(a + b)(c + d)(a + c)(b + d)}$$

b) A random sample of 30 adults is classified according to sex and the number of hours they watch television during a week.

Classes	Male	Female
Over 25 hours	5	9
Under 25 hours	9	7

Using $\alpha = 0.01$ level of significance, test the hypothesis that a Person's sex and time watching television are independence.

(10+10)

Q.9.a) Using Method of successive approximation find the x when f(x) = 14.00 given:

x	0	5	10	15
f(x)	16.35	14.88	13.59	12.46

b) By given $u_0 = 0$, $u_{10} = 15$, $u_{20} = 50$. Estimate u_{15}, if you were given in addition $u_5 = 35$. How would your estimate be revised?

c) Apply Lagrange's formula to find f (5), given f (1) =2, f (3) =5, f (4) =10, f (7) =128.

(8+6+6)

UNIVERSITY OF THE PUNJAB

Part-I A/2018
Examination:- M.A./M.Sc.

Roll No. .....

Subject: Statistics **TIME ALLOWED: 3 hrs.**
PAPER: II (Probability and Probability Distributions) **MAX. MARKS: 100**

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

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

i) Properties of Binomial Distribution
ii) Characteristic function and its properties
iii) Chebyshev's inequality
iv) Uniqueness Theorem
v) Moment Generating Function

b) Find the probability that at a game of bridge one hand will get ace and king of exactly 2 suits. **(5)**

c) When two dice are rolled, find the probability of getting a greater number on the first die than the one on the second, given that the sum should equal 8. **(5)**

d) In an examination 30% students failed in Statistics, 25% students failed in English and 10% students failed in both subjects. Find the probability that a randomly selected student **(5)**

i) Failed in statistics if it is known that he failed in English

ii) Failed in Statistics but passed in English.

Q.2.a) Poker dice is played by simultaneously rolling 5 dices. Find the probability that **(6)**

i. No two alike
ii. One pair
iii. Two pair

b) In a thirteen cards bridge game find the probability that **(6)**

i) All the 13 cards in a bridge hand have different face value

ii) A hand contains 2 spades, 7 hearts, 3 diamonds and 1 club.

c) In a certain day care class, 30% of the children have grey eyes, 50% of them have blue and the other 20%'s eyes are in other colors. One day they play a game together. In the first run, 65% of the grey eye ones, 82% of the blue eyed ones and 50% of the children with other eye color were selected. Now, if a child is selected randomly from the class, and we know that he/she was not in the first game, what is the probability that the child has blue eyes? **(8)**

d) Find the c.d.f. of the random variable from the p.d.f. **(5)**

$$f(x) = \begin{cases} 3e^{-3x} & 0 < x < \alpha \\ 0 & \text{otherwise} \end{cases}$$

and use it tc evaluate $P(0.5 \leq X \leq 1)$.

PTO

Q.3.a) If three independent variables X_1, X_2 *and* X_3 have Poisson distribution with means λ_1, λ_2 *and* λ_3 respectively, then show that their sum $X_1 + X_2 + X_3$ is a Poisson variate with mean $\lambda_1 + \lambda_2 + \lambda_3$. **(5)**

b) Derive the m.g.f. of the discrete uniform distribution. Use it to find mean and variance of the discrete uniform distribution. **(8)**

c) Obtain first four factorial cumulants and first four mean moments of the Binomial distribution and also find coefficient of skewness and kurtosis. **(12)**

Q.4.a) Find expected value, standard deviation, coefficient of skewness and kurtosis of the Negative binomial distribution. **(9)**

b) Define Pareto distribution and derive its first four mean moments. **(8)**

c) Derive the m.g.f of normal distribution and thus find its mean and variance. **(8)**

Q.5.a) Find the mean deviation of exponential distribution. **(5)**

b) Derive the m.g.f of bivariate normal distribution. **(20)**

Q.6.a) If X follow normal distribution with mean μ and variance σ^2. Obtain the distribution of $Y = |X|$ and show that the area under the probability curve is unity. **(10)**

b) Let X_1 *and* X_2 be two independent standard normal variates **(8)**

Let $Y = (X_2 - X_1)^2 / 2$

Find the distribution of Y.

c) Derive the mean and variance of F-distribution. **(7)**

Q.7.a) State and prove the Central limit theorem. **(10)**

b) Define order statistics. Derive the distribution of *r*th order statistic (y_r). **(10)**

c) Show that the square of a t-variate with *n* d.f is distributed as *F* with 1 and n d.f. **(5)**

UNIVERSITY OF THE PUNJAB

Part-I A/2048
Examination:- M.A./M.Sc.

Roll No. .....

Subject: Statistics **TIME ALLOWED: 3 hrs.**
PAPER: III (Design and Analysis of Experiments) **MAX. MARKS: 100**

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

Q.1 a) Drive the partition of the total sum of squares for a two-way analysis of variance without interaction. Find the number of degrees of freedom associated with each of the sources of variation. (15)

b) What is the purpose of experimental design? Write down the steps which must be accomplished before any experiment is performed. (10)

Q.2 a) Compare the efficiency of Latin Square Design with RCB and CR design. (10)

b) Derive formula for estimating 2 missing values in Latin Square Design when values are missing in same columns, different rows and treatments. What changes will occur in ANOVA table after estimating missing observations. (15)

Q.3 a) Given below are the sums of cross products for RCB design with one covariate. Compute the analysis and draw conclusions (10)

Source of Variation	d.f	$\sum X^2$	$\sum Y^2$	$\sum XY$
Blocks	2	1.2667	0.933	1.9111
Treatments	14	19.4000	68.1333	10.3111
Error	28	7.4000	13.7333	10.7556
Total	44	28.0667	82.7996	22.9778

b) In a RCB design with p treatments and r blocks is assumed to be represented by the model (15)

$$Y_{ij} = \mu + \alpha_i + \gamma_j + \beta(X_{ij} - \bar{X}_{..}) + \varepsilon_{ij}$$

$$i = 1, 2, \ldots, p; \qquad j = 1, 2, \ldots, r$$

Develop the procedure for testing the null hypothesis that adjusted treatment means are equal.

Q.4 a) What are factorial experiments? If we use a single replicate in a 2^k factorial design then in this situation how can we estimate the Error SS? (10)

b) The following data were obtained from a 2^3-factorial experiment; use Yate's method to analyze the data. (15)

Blocks	Treatment combinations							
	(1)	a	b	c	ab	ac	bc	abc
I	12	14	15	14	14	10	11	11
II	21	19	16	17	18	16	18	15
III	10	12	10	14	11	11	10	12

Q.5 a) What is fractional replication? Write the advantages and disadvantages of fractional replication as compare to simple factorial experiment. (10)

b) A 2^3 experiment with 2 replication measured the yield of Brussels sprouts (kg/plot) as follows, where the factors were A,B, and C; Obtain the ANOVA for the below data and draw conclusions. (15)

PTO

Replicate I		Replicate II	
ABC Confounded		AC Confounded	
BI	BII	BI	BII
(1)= 10	a=12	a= 7	(1)= 11
ab= 18	b= 11	c= 14	b= 14
ac= 17	c= 16	ab= 16	ac= 19
bc= 16	abc=25	bc= 16	abc= 22

Q.6 a) Define Split-plot Design? In what situation this design is used? (10)

b) Complete the following ANOVA table and test the significance of treatment means. Also compute standard errors for (i) whole plot treatment mean, (ii) the difference between 2 sub-plot means. (15)

S.O.V	d.f	SS	MS	F-ratio
Blocks	4	18.44	-	-
A	-	4.92	-	-
Error (a)	-	-	-	
Subtotal	-	50.85		
B	3	29.30	-	-
AB	12	-	-	-
Error(b)	-	86.65	-	

Q.7 a) Compare the RCB design with Incomplete block design. Also give the merits and demerits of Incomplete block design over RCB design. (10)

b) Complete the analysis for the following BIB design (15)

Blocks	Treatments				
	1	2	3	4	5
1	257	230	279	287	-
2	245	283	-	280	260
3	231	352	280	-	250
4	203	-	227	193	259
5	-	271	286	334	328

UNIVERSITY OF THE PUNJAB

Part-I A/2018
Examination:- M.A./M.Sc.

Roll No. .....

Subject: Statistics **TIME ALLOWED: 3 hrs.**
PAPER: IV (Sampling Techniques) **MAX. MARKS: 100**

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

Q#1 (a)	Describe the advantages of a sample survey in comparison with a census survey.	10
(b)	If the loss function due to an error in $\bar{y}$ is $d(\bar{y}-\bar{Y})^2$ 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 $\sqrt{\frac{dS^2}{C_1}}$.	10
Q#2(a)	Explain the difference between stratification and post stratification. Explain briefly.	10
(b)	If the terms in $\frac{1}{N_h}$ are not ignored then prove that $$V_{ran} = V_{prop} + \frac{1-f}{n(N-1)}\left[\sum_{i=1}^{L} N_h(\bar{Y}_h - \bar{Y})^2 - \frac{1}{N}\sum_{i=1}^{L}(N-N_h)S_h^2\right]$$	10
Q#3 (a)	Discuss the important factors to be considered while determining the size of sample.	10
(b)	The farms of a small nation were divided into seven strata based on their areas reported in the last census. The yields of the chief cash crops were estimated from the results of a sample survey and rounded to convenient numbers.	10

h	N_h	$\bar{Y}_h$	S_h^2
1	50,000	0.13	0.25
2	23,000	0.72	2.89
3	20,000	3.34	72.25
4	5,300	18.03	1225
5	1,500	68.85	9025
6	120	786.00	40,000
7	80	434.00	28,900

Design a sample of size n=3000 using proportional allocation and find its variance.

Q#4(a)	Define systematic sampling. Describe the procedure to draw a sampling using linear systematic sampling.	10
(b)	In a directory of 6 households on a street, the persons are listed as below: M= male adult, F= female adult, m= male child, f=female child **1 2 3 4 5 6**	10

PTO

	M M M M M M F F F F F F f f m m f m m f m m f f f Compare the variances given by a systematic sample of 1 in 5 persons and a 20% simple random sample for estimating the proportion of female adults.	
Q#5(a)	Derive the following expression to compare the separate and combined ratio estimates: $V(\hat{Y}_{RC}) - V(\hat{Y}_{RS}) = \sum_{h}^{L} \frac{N_h^2(1-f_h)}{n_h}[(R-R_h)^2 S_{x_h}^2 + 2(R_h - R)(\rho_h S_{x_h} S_{y_h} - R_h S_{x_h}^2)]$	12
(b)	Obtain the variance of $\hat{Y}_R$ which is minimum, under the model $y_i = \beta x_i + \varepsilon_i$	08
Q#6 (a)	Show that the value of b_0 that minimizes $V(\bar{y}_{lr})$ is $B = \frac{S_{yx}}{S_x^2}$.	10
(b)	If b is the least square estimate of B and $\bar{y}_{lr} = \bar{y} + b(\bar{X} - \bar{x})$, then prove that for simple random sampling of size n (assuming large), the resulting minimum variance is $V(\bar{y}_{lr}) = \frac{1-f}{n} S_y^2 (1-\rho^2)$ Where ρ is the population correlation coefficient between y and x.	10
Q#7 (a)	Differentiate cluster and stratum by help of a practical example.	08
(b)	Calculate an expression for the optimum value of M, the size of cluster, by using the cost function $C = c_1 nM + c_2\sqrt{n}$ Where C is the total cost of the survey.	12
Q#8 (a)	Explain briefly how you would select a PPS sample using Hansen-Hurwitz scheme.	08
(b)	If n units and m subunits from each chosen unit are selected by simple random sampling, prove that $v(\bar{\bar{y}}) = \frac{1-f_1}{n} s_1^2 + \frac{f_1(1-f_2)}{mn} s_2^2$ is an unbiased estimate of $V(\bar{\bar{y}})$.	12
Q#9	Write a short note on the following: i. Non-response error ii. Post stratification iii. Double sampling iv. Unequal Probability Sampling	5 each