INDIAN STATISTICAL INSTITUTE
Research and Training School

Mid-year Examination

B.Stat. IV Year

PROBABILITY

Duration : 3 hours | Maximum Marks: 100 | Date: 7 December 1964

1. Define a) the Borel σ field and b) a measurable function on the real line R. Show that the space of all measurable functions on R is closed under addition and multiplication. Show that if $f_1(x)$, $f_2(x)$, ... is a sequence of measurable functions on R then $\overline{\lim}\, f_n(x)$ and $\underline{\lim}\, f_n(x)$ are measurable functions. (35)

2. Define a) the Lebesgue measure and b) the integral of a function with respect to Lebesgue measure over an interval.

Evaluate $\int_0^1 f(x)\,dx$ where

$$f(x) = \begin{cases} \text{Sin } x & \text{if } x \text{ is irrational} \\ 0 & \text{otherwise.} \end{cases}$$

State the Lebesgue dominated convergence theorem. Deduce the following :

If $f(x)$ is a non-negative function defined on the real line R and is integrable over every bounded interval then prove that the set function

$$P(S) = \begin{cases} \int_S f(x)\,dx & \text{if } f(x) \text{ is integrable over } S \\ +\infty & \text{otherwise} \end{cases}$$

is a measure on the Borel σ-field. (35)

3. Let $F(x)$ be the distribution function associated with a probability measure on the real line. Show the following :

1) $F(x)$ is monotonic increasing, non-negative and right continuous.

2) $\lim_{x \to -\infty} F(x) = 0$, $\lim_{x \to +\infty} F(x) = 1$

3) $F(x) = F_1(x) + F_2(x)$ where $F_1(x)$ is a continuous monotonic increasing function and $F_2(x)$ is a pure jump function. (30)

ooooooo

ad/1.12.64

INDIAN STATISTICAL INSTITUTE
Research and Training School

Mid-year Examination

B. Stat. IV Year

STATISTICAL INFERENCE

Duration : 3 hours | Maximum Marks: 100 | Date: 8 December 1964

Answer any four questions

1(a) In the context of a family of probability distributions on a countable sample space, define sufficient statistics, state the factorisation theorem for sufficient statistics and prove it. (10)

(b) How is the definition of sufficient statistics modified when the sample space is the n-dimensional Euclidean space on which there is a continuous density function? State the factorisation theorem in this set up. (Mention the regularity conditions that are involved. No proof is needed). (8)

(c) (X, Y) follows bivariate normal distribution with mean vector $(\theta, 2\theta)$ and covariance matrix $\begin{pmatrix} 1 & 1 \\ 1 & 4 \end{pmatrix}$, where θ is unknown.

Show that $2X + Y$ is a sufficient statistic for θ. (7)

2(a) $X_1, \ldots, X_n$ are independent random variables whose common distribution involves a single unknown parameter θ. Obtain a lower bound for the variance of an estimator of θ. When is this bound attained? (Mention all the regularity conditions that are involved). (20)

(b) How will the above bound be modified if the frequency function of X_i is $f_i(x;\theta)$, $i=1,\ldots,n$ where the functions f_i are not identical? (5)

3. In a certain production process, the proportion θ of defective items is unknown. Independent samples are taken from the production process till a defective item is observed. The above experiment is carried out twice and let X_i be the number of samples needed in the i-th experiment, $i = 1,2$. Let

$$T = 1 \text{ if } X_1 = 1, \quad = 0 \text{ otherwise} \qquad \text{and} \qquad S = X_1 + X_2$$

(a) Show that T is an unbiased estimator of θ. (2)

(b) Show that S is sufficient for θ. (5)

(c) Obtain the minimum variance unbiased estimator of θ. (10)

(d) Find the variance of the estimator obtained in (c). (8)

4(a) State and prove the Neyman-Pearson lemma for the most powerful test of a simple hypothesis against a simple alternative at a given level of significance. (18)

(Please turn over)

4(b) $X_1, \ldots, X_n$ are independent Poisson variates, X_r having mean $r\theta$, $r=1,\ldots, n$. Derive the most powerful test at level α of $H_0 : \theta = \theta_0$ against $H_1 : \theta = \theta_1 > \theta$ on the basis of $X_1, \ldots, X_n$. (7)

5(a) What is a monotone likelihood family of probability distributions? (5)

(b) Show that if the frequency function f_θ of X belongs to a monotone likelihood ratio family, then the most powerful test of $H_0 : \theta = \theta_0$ against $H_1 : \theta = \theta_1 > \theta_0$ at level α is also the uniformly most powerful test of $H_0^* : \theta \leq \theta_0$ against $H_1^* : \theta > \theta_0$ at level α. (12)

(c) $X_1, \ldots, X_n$ are independent normal variates with mean 5 and unknown variance σ^2. Find the uniformly most powerful test of the hypothesis $\sigma^2 \leq 2$ against alternative $\sigma^2 > 2$ at level α. Write down the expression for the power of this test at $\sigma^2 = 4$. (8)

ooooooooo

ad/2.12.64

INDIAN STATISTICAL INSTITUTE
Research and Training School

Mid-year Examination

B. STAT. IV YEAR

Statistical Methods (Theory)

Duration 3 hours | Maximum marks : 100 | Date : 4 December 1964

Attempt any four questions
All questions carry equal marks.
[Some marks are reserved for home assignments]

1. Obtain the likelihood ratio criterion for testing on the basis of samples, the hypothesis that k different Normal Populations have the same variance. How does Bartlett modify the criterion? Work out the t-th moment (about origin) of Bartlett's statistic in the null case. How is this test related with the usual variance-ratio(F) test when $k = 2$.

2. Denoting by r the correlation coefficient in a sample of size n from a bivariate Normal Population, show that the statistic

$$t = \frac{r}{\sqrt{1-r^2}} \sqrt{n-2}$$

follows the t-distribution with (n-2) degrees of freedom, if the two variates are independent.

Show that two correlated variables X_1 and X_2 can have equal variances if and only if $S = X_1 + X_2$ and $D = X_1 - X_2$ are uncorrelated. Hence suggest a test for the hypothesis that the two variables have a common variance on the basis of a sample of size n from a bivariate Normal Population.

3. How will you test on the basis of a sample of size n from a p variate Normal Population $(p = p_1 + p_2)$ the hypothesis that the first p_1 variates are independent of the last p_2 variates? Work out the t-th moment about origin of the appropriate likelihood-ratio statistic in the null case. Discuss in particular the case $p_1 = 1$, $p_2 = p-1$.

4. How will you test on the basis of samples from two p-variate Normal Population the hypothesis that the mean vectors are identical when the dispersion matrices (a) are known (b) are known to be equal, but the common value is unknown?

Assuming the hypothesis to be true, work out in each case the sampling distribution of the statistic you would use.

5. Obtain the likelihood-ratio statistic for testing on the basis of sample of size n from a bivariate Normal Population the hypothesis that the two variates are independent and have equal means and equal variances. Work out the t-th moment (about origin) of the statistic in the null case.

INDIAN STATISTICAL INSTITUTE
Research and Training School
Mid-year Examination

B. STAT. IV YEAR

Statistical Methods (Practical)

Duration: 3 hours. Maximum marks. 100 Date. 4 December 1968
Ten marks are reserved for class records.

1. The correlations obtained from 6 samples of sizes 10, 14, 16, 20, 25, 28 are 0.318, 0.106, 0.253, 0.340, 0.116, 0.112 respectively. Can these be considered homogeneous? (State your assumptions clearly). [20]

2. The following statistics were computed to study the length (L)-weight (W) relationship in three common species of fish in an experimental tank in a fishery. It has been suggested that weight varies as $(L)^n$. Examine if the length - weight relationship is identical for the three species. Comment on the suggestion that n does not depend on the species and is nearly 3.

Mean and corrected sums of squares and products.

species	number of fish	mean: log length (x)	mean: log weight (y)	corrected sums of squares and products: (x^2)	(y^2)	(xy)
Catla	394	1.5723	0.7805	108.443	1340.003	371.861
Rohita	121	1.4571	0.4846	37.000	431.748	124.122
Mrigal	124	1.7493	1.3468	40.923	467.813	135.927

[30+10]

3 The following table gives the values of the variances of the stature in centimetres for Muslims in eight different districts of Bengal as obtained in the Bengal Anthropometric Survey, 1945.
Examine if the variance of stature differs from district to district (use Bartlett's correction). Obtain a pooled estimate for the common variance if you accept the hypothesis of equal variances.

Variances of stature of Muslims in eight districts of Bengal.

Districts	Sample size	Variances of stature
Barisal	131	39.5068
Burdwan	59	24.4685
Dacca	337	39.2262
Faridpur	77	24.8924
Murshidabad	124	24.3490
Mymensingh	299	30.3306
Nadia	170	32.4325
Rangpur	139	33.4436

[30]

INDIAN STATISTICAL INSTITUTE
Research and Training School

Mid-year Examination

B. Stat. IV Year

Design of experiments (theory and practical)

Duration : 4 Hours | Maximum Marks: 100 | Date: 10 December 1964

1. Explain the importance and role of the following in design and analysis of experiments :

   a) randomisation, b) replication, c) local control (4+4+4=12)

2. In a simple randomised block experiment, one of the observations was found to be missing. Explain, with brief mathematical arguments, how you will analyse the data. (10)

3(a) Explain the purpose of confounding in factorial experiments. (6)

(b) Prepare a layout for a single replicate 2^5 factorial experiment in blocks of size 8, confounding only interactions involving 3 or more factors. (12)

4. Analyse the following data obtained from an experiment arranged in balanced incomplete blocks. Which of the four treatments A, B, C, D do you recommend as the one giving the maximum yield? (27)

Wheels	Blocks	Treatments and yields		
Right Front	1	(A) 42	(D) 40	(B) 38
Left Front	2	(A) 41	(B) 36	(C) 36
Right Rear	3	(D) 30	(C) 33	(A) 33
Left Rear	4	(B) 32	(D) 26	(C) 29

5. The Table below gives the gains in weight (gms.) and the initial weights (given within brackets) in an experiment to compare four diets A, B, C, D with 5 litters of rats.

litter \ diet	A	B	C	D
1	57 (93)	65 (97)	71 (94)	68 (90)
2	55 (93)	67 (95)	59 (89)	67 (89)
3	62 (95)	70 (98)	64 (92)	69 (90)
4	74 (98)	61 (96)	74 (96)	73 (93)
5	87 (102)	92 (104)	78 (103)	91 (107)

(a) Examine whether the gain in weight is affected by the initial weight. (8)

(b) Assuming that the answer in (a) is "yes", analyse the data utilising all the information given. (25)

ooooooooooooo

ad/3.12.64

INDIAN STATISTICAL INSTITUTE
Research and Training School

Mid-year Examination

B.Stat. IV Year

SQC – Theory & Practical

Duration: 4 hours | Maximum Marks: 100 | Date: 9 December 1964

Separate Answer-book should be used for each Group

(Attempt any TWO questions from Group A and the WHOLE of Group B)

GROUP A

1. The production from a certain process is at a continuous rate. It is required to set up an $\bar{x}$-chart to control the level of production. The standards μ and σ are known. Two schemes of sampling are suggested as follows :

Scheme 1 : Samples of size 4 taken every 15 minutes.

Scheme 2 : Samples of size 16 taken every hour.

(a) Suppose that you start taking samples from 10 a.m. on a certain day. If a shift in the level of production from μ to $\mu + 1.5\sigma$ occurs at 11-10 a.m., what is the chance that the shift will go undetected on taking the next sample under each scheme? (4)

(b) Assuming that defectives are produced at a constant rate after the shift has occurred, compare the average number of defectives produced under the two schemes before the shift is detected. (6)

(c) Have you got any comments to make in the light of your answers to (a) and (b)? (4)

(d) The problem above is evidently one of choosing between large samples at infrequent intervals and small samples at frequent intervals. What are the other factors which will influence our decision? Explain clearly. (6)

2. To control the level of production of a process, $\bar{x}$-chart is set up. x is normal with mean m' and standard deviation σ. The control limits are set at

$$m \pm B\frac{\sigma}{\sqrt{n}}$$

(a) If $p = \text{Prob.}\ \bar{x} > m + B\frac{\sigma}{\sqrt{n}}$, show that the average number of samples required to obtain λ successive points above the upper control limit is given by

$$S = \frac{1 - p^{\lambda}}{p^{\lambda}(1-p)}$$ (10)

(b) If S = 1000 and $\lambda = 2$ in (a), find the value of B. (5)

(c) Briefly indicate how the above problem of choosing the control limits such that a specified number of samples, say 1000, will be required to obtain λ successive points above (or below) the upper (or lower) control limit could be used to modify and improve the conventional control chart. (5)

Please turn over

3(a) Write a short note on accelerated life testing. (5)

(b) The length of life X of a certain population of electronic tubes is known to have an exponential distribution characterized by the probability density function $f(x, \theta)$ given by

$$f(x,\theta) = \begin{cases} \frac{1}{\theta} e^{-\frac{x}{\theta}}, & x \geq 0, \ \theta > 0 \\ 0, & \text{elsewhere.} \end{cases}$$

A sample of n tubes was put on test. The tubes, which failed, were not replaced. The test was terminated at the r-th $(r<n)$ failure. The observed completed lives in the order of their failure are $x_1, \ldots, x_r$.

(i) What is the physical meaning of θ ? (1)

(ii) Find $\hat{\theta}$, the maximum likelihood estimate (m.l.e.) of θ. (5)

(iii) If is found that $\frac{2r\hat{\theta}}{\theta}$ is distributed as a χ^2 with 2r degrees of freedom. Show that $\hat{\theta}$ is an unbiased estimate of θ and obtain the variance of $\hat{\theta}$. (3)

(iv) 20 electronic tubes were placed on test and the experiment was terminated at the fifth failure. The completed lives were 107, 112, 115, 124 and 125 hours. Obtain the m.l.e. of θ and also the one-sided and two-sided 95% confidence intervals for θ. (6)

4. Write a detailed note on any ONE of the following :

(a) Economics of control charts.

(b) Military Standard 105. (20)

GROUP -B

5(a) Obtain the values at which the control charts lines will be drawn in the following cases.

(i) p-chart, given that in 10 samples of size 200 each the numbers of defectives found are 6,6,6,5,0,0,6,1,4 and 0 respectively. (3)

(ii) c-chart, given that the numbers of surface defects observed in the inspection of 15 galvanised sheets of given area are 15,9,13,20,11,15,7,11,22,12,24,16,.., 8,24,16. (3)

(b) Subgroups of five items each are taken from a manufacturing process at regular intervals. A certain quality characteristic (x) is measured. After 25 subgroups, $\Sigma \bar{x} = 357.50$ and $\Sigma R = 9.90$.

(i) Compute the control limits for $\bar{x}$ - R charts. (3)

(ii) Both the charts exhibit a perfect state of control. Find the natural tolerance limits of the process. (2)

(iii) If the specification limits are 14.40 ± 0.45, what conclusions can you draw regarding the ability of the process to produce items within these specifications? (2)

(iv) Can you suggest possible ways, if any, of improving the situation? (2)

5(c) Obtain 95% confidence interval for the following :

(i) process mean μ , given that from a sample of 15 items the mean and standard deviation (with divisor $(n-1)$) obtained are 15.6 and 5.6 respectively. (3)

(ii) process standard deviation, given that from a sample of six items the range obtained was 2.8. (3)

6. A one-sided specification for a measurable quality x [which follows $N(\mu,\sigma)$] of a product is as follows :

If $x > 5.6$ cms, the item is a defective

$x \leq 5.6$ cms, it is a non-defective.

A lot of the product is defined to be 'good', if it contains not more than 4% defectives and 'bad' if it contains not less than 6% defectives.

(a) Find the interval in which μ can wander in order to produce lots of indifference quality. (3)

(b) If μ = 3.8 cms, what will be your verdict on the lot? (1)

(c) Derive a suitable single sampling plan with producer's risk = 0.05 and consumer's risk = 0.10. (8)

(d) Suppose that a sample inspected as per the plan obtained in (c) gives $\bar{x}$ = 5.2 cms.

(i) Can you accept the lot? (1)

(ii) If not, how will you revise the specification so that it is accepted? (2)

(e) If only a division into defectives and non-defectives were made, explain how you can derive a single sampling plan with the same stipulations as above. (5)

7(a) Set up a graphical procedure of a sequential sampling plan to descriminate between lots of proportion defectives p_0 and p_1, if the risks are taken to be α each. (10)

(b) Draw the 5-point OC and ASN curves, given that $p_1 = 2p_0 = \alpha = 0.10$. (8)

(c) Indicate clearly how you will find the AOQL of the plan with stipulations given in (b). (2)

....

ad/4.12.64

INDIAN STATISTICAL INSTITUTE
Research and Training School

Mid-year Examination

B.Stat. IV Year

ECONOMICS I

Duration: 2 Hours | Maximum Marks: 100 | Date: 5 December 1964

Separate Answer-book should be used for each Group

Group A

1. The Keynesian consumption function may be regarded as an aggregate demand function. Can you show any formal relationship between the consumption function and demand function of macroeconomic theory? Should you make any simplifying assumptions they must be clearly stated. (25)

Or,

Why are single cross-section samples not suitable for the estimation of price elasticities of demand? What a time-series of cross-section samples be more suitable for this purpose? (25)

2. What do you understand by 'quality' elasticity. Indicate how you would proceed to estimate it from grouped survey data. Is the method of concentration curves more appropriate than the conventional method of least squares? (25)

Or,

Formulate a demand function for the exports of Indian goods to the United Kingdom. State how you would proceed to estimate its parameters? What type of statistics would you need in order to do this? (25)

Group B

(Answer any two questions)

3. Consider a one-commodity economy and assume that input requirement per unit of output is 0.4. Given the initial stock (i.e. in period 0) of the commodity as 50 units, analyse the choices open to the economy for a planning horizon extending from period 0 to period 3. How much will the economy gain in consumption in period 3, if it sacrifices one unit of consumption in period 0, given the terminal stock, and consumption of periods 1 and 2? (25)

4. The technology of an economy consisting of four commodities is described by the matrix where the columns denote the processes and rows the commodities.

$$A = \begin{bmatrix} 1 & 1 & 1 \\ -1 & 2 & 1 \\ 0 & -1 & 2 \\ -1 & -3 & -4 \end{bmatrix}$$

(Please turn over)

(i) Interpret and show that the first commodity may be produced alone, jointly with the second commodity or jointly with the second and third commodities.

(ii) Show the different ways of producing one unit of the first commodity and examine under what conditions the other three commodities are outputs and inputs. (25)

3. In a two-sector economy, assume that the flow inputs are supplied from current production, but the stock inputs are available from the last period. Set down the balance relations for such an economy and show how a determinate solution of the planned outputs of the two sectors can be obtained. (25)

ooooooooo

ad/1.12.64

INDIAN STATISTICAL INSTITUTE
Research and Training School

Mid-year Examination

B.Stat. IV Year

ECONOMICS II

Duration : 2 Hours · Maximum Marks: 100 · Date: 5 December 1964

Expenditures by urban consumer units from a cross-section sample in Japan, April 1930, show the mean expenditures by income class in Table 1.

Income class (yen per month)	Average income	Average expenditure	No. of households
Under 5,000	1,414	17,653	39
5,000 – 9,999	7,279	11,958	85
10,000 – 14,999	12,456	14,103	129
15,000 – 19,999	17,725	18,529	242
20,000 – 24,999	22,205	22,618	381
25,000 – 29,999	27,316	25,786	397
30,000 – 34,999	32,219	30,092	336
35,000 – 39,999	37,279	34,500	285
40,000 – 44,999	42,268	39,816	192
45,000 – 49,999	47,226	42,048	136
50,000 – 59,999	54,274	49,585	193
60,000 – 69,999	64,251	57,022	83
70,000 – 79,999	73,887	68,117	52
80,000 – 89,999	83,323	66,500	19
90,000 – 99,999	93,381	90,130	21
100,000 and over	131,884	113,714	27

1. Compute the average income and expenditure in the whole sample. (10)

2. Compute the cumulative proportions and draw the concentration curves for the distributions of income and expenditure. (25)

3. Obtain the respective concentration ratios, and see whether or not the concentration of income is more than the concentration of expenditure. (20)

4. From the concentration curves, calculate the relative shares of income and expenditure accruing to the bottom 10 per cent and top 10 per cent of households in the sample. (15)

5. Finally obtain the estimate of average and marginal propensities of expenditure by assuming (i) a linear relation between income and expenditure, and (ii) a log-linear relationship between income and expenditure. (30)

ooooooooo

ad/1.12.64

INDIAN STATISTICAL INSTITUTE
Research and Training School

Mid-year Examination

B. Stat. IV Year

Demography (Theory and Practical)

Duration: 3 hours | Maximum Marks: 100 | Date: 11 December 1964

1. Explain the terms (a) Gross reproduction rate (b) Net reproduction rate. How do they help in the study of population growth? (15)

2. Obtain the differential equation of the growth of population and derive the logistic stating clearly the conditions under which it is applicable.

   Discuss the law of growth of population with reference to India from the following data.

Year	1901	1911	1921	1931	1941	1951	1961
Population (in millions)	235	249	248	276	313	357	438

(20)

3. The number of survivors l_x and the complete expectation of life e_x^o of a life table for females are given in table 1. If these figures are applicable to a female population group 'P' having at present an age distribution given in table 2, estimate the age composition of the population 'P' 10 years hence. (35)

Table 1 : Life table for females

Age x	No. of survivors at age x l_x	Complete expectation of life at age x e_x^o
0	1000	66.5
1	967	67.8
5	961	64.2
10	958	59.4
15	955	54.6
20	950	49.8
25	943	45.2
30	936	40.5
35	928	35.8
40	917	31.2
45	901	26.7
50	872	22.5
55	828	18.6
60	759	15.1
65	659	12.0
70	530	9.3
75	379	7.0
80	222	5.2
85	94	3.9

(Please turn over)

Table 2 : Present age composition of the female population 'F' (in 000's)

Age group	Female population (in 000's)
25-29	891
30-34	1207
35-39	1108
40-44	1301
45-49	941
50-54	611
55-59	504

4(a) Define 'underlying cause of death'.

(b) Enter the details given in the following example in the International Form of Medical Certificate of Cause of Death (The approximate interval between onset of the disease and death is given within brackets).

Example: An adult died of septicemia (6 days) resulting from diabetes (2 years) complicated by gangrene (1 month). (30)

ooooooo

ad/3.12.64

INDIAN STATISTICAL INSTITUTE
Research and Training School
Periodical Examination
B. Stat. IV class
PROBABILITY

Date: 8 April 1968 | Maximum marks 100 | Time: 3 hours

1. What is a probability space ?

Define a random variable. If (X, S, P) is a probability space and $f(x)$ is a random variable show that $f(x)$ possesses a distribution on the real line.

When are two random variables $f(x)$ and $g(x)$ (defined on the probability space (X, S, P)) said to be independent ?

If $f(x)$ and $g(x)$ are two independent random variables on (X, S, P) and $F(t)$ and $G(t)$ are real valued Borel measurable functions defined on the real line show that $F(f(x))$ and $G(g(x))$ are also independent random variables on (X, S, P).

If $f_1(x)$ and $f_2(x)$ are two independent random variables on (X, S, P) with distribution functions F_1 and F_2 what is the distribution of the random variable $f_1 + f_2$? Show that, if either F_1 or F_2 is continuous then the distribution of $f_1 + f_2$ is also continuous.

3+3+12+4+12+4+12 = 50

2. Define the characteristic function of a distribution function.

Suppose (X, Y) is a random variable in the plane with density function

$$f(x, y) = \frac{1}{4}\left[1 + xy\,(x^2 - y^2)\right] \quad \text{if } |x| < 1,\ |y| < 1,$$
$$= 0 \qquad \text{otherwise.}$$

Evaluate the characteristic functions of the distributions of X, Y and $X + Y$ respectively. What do you infer from this example ?

Let $\phi(t)$ be the characteristic function of a distribution $F(x)$. Show that

$$F(x + 0) - F(x - 0) = \lim_{T \to \infty} \frac{1}{2T} \int_{-T}^{T} e^{-itx}\, \phi(t)\, dt.$$

If $p_1, p_2, \ldots$ are the magnitudes of the jumps of the distribution $F(x)$ show that

$$\Sigma\, p_i^2 = \lim_{T \to \infty} \frac{1}{2T} \int_{-T}^{T} |\phi(t)|^2\, dt.$$

2+12+18+18

INDIAN STATISTICAL INSTITUTE
Research and Training School
Periodical Examination
B. Stat. IV year
STATISTICAL INFERENCE

Date: 29 March 1965 | Maximum marks: 100 | Time: 2½ hrs.

All questions carry equal marks.

1. $f_1, \ldots, f_n, f_{n+1}$ are integrable functions on the k-dimensional Euclidean space E_k and $c_1, \ldots, c_n$ are given real numbers. Let $\mathcal{C}$ denote the class of all critical functions ϕ (i.e. $0 \leq \phi(x) \leq 1$ for all $x \in E_k$) such that

$$\int \phi(x)\, f_i(x)\, dx = c_i, \quad i = 1, \ldots, n.$$

Show that if there exist $k_1, \ldots, k_n$ and a function $\phi_0 \in \mathcal{C}$ such that

$$\phi_0(x) = 1 \text{ if } f_{n+1}(x) > \sum_1^n k_i f_i(x)$$
$$= 0 \text{ if } f_{n+1}(x) < \sum_1^n k_i f_i(x),$$

then $\int \phi_0(x)\, f_{n+1}(x)\, dx \geq \int \phi(x)\, f_{n+1}(x)\, dx$ for all $\phi \in \mathcal{C}$

State (without proof) a condition under which such $k_1, \ldots, k_n$ and $\phi_0 \in \mathcal{C}$ exist.

2. The frequency function of a random variable X belongs to the exponential family

$$f(x, \theta) = c(\theta)\, e^{\theta t(x)}\, h(x).$$

Show that the power function of an arbitrary test for $H_0: \theta = \theta_0$ is differentiable and that the differentiation can be carried out under the sign of integration.

3. $X_1, \ldots, X_n$ are independent random variables with common frequency function

$$f(x, \theta, \gamma_1, \ldots, \gamma_r) = c(\theta, \gamma_1, \ldots, \gamma_r)\, e^{\theta t(x) + \sum_1^r \gamma_i u_i(x)}\, h(x)$$

a) What are the similar region tests of level α for $H_0: \theta = \theta_0$ when $\gamma_1, \ldots, \gamma_r$ are treated as nuisance parameters ?

b) What are the tests of Neyman - structure for the same problem ?

c) Prove the equivalence of the class of similar region tests and the tests of Neyman - structure, stating the conditions involved.

4. $Y_1, \ldots, Y_n$ are independent random variables, Y_i being normally distributed with mean $\alpha + \beta x_i$ and variance σ^2, $x_1, \ldots, x_n$ are known constants but α, β and σ^2 are unknown. Derive the uniformly most powerful unbiased test for $H_0: \beta = 0$ against $H_1: \beta \neq 0$ on a given level of significance.

INDIAN STATISTICAL INSTITUTE
Research and Training School
Periodical Examination
B.Stat. IV.
Operation Research

Date: 13.4.65. Maximum marks : 100 Time: 3 hrs.

Use separate answer book for each group.

Group A

1. a) Define (1) a basic feasible solution (2) extreme point of the set of feasible solutions of a linear programming problem and establish the relationship between them.

b) State and prove simplex optimality test to find out whether a given basic feasible solution is optimal.

c) State and formulate the diet problem of linear programming.

(8+6+6)

2. A company makes 3 types of products which require the use of two different machines whose capacities are limited.

The times on each machine required for manufacturing unit amount of each 3 products, the profit per unit amount of each product and the capacities of each of two machines are given below:

Machines	Products I	II	III	Capacities of machines (unit time)
I	2	2	4	3
II	3	2	1	1
profit per unit	1	2	3	

Find the optimal amounts of each of the 3 products to be manufactured so as to maximise the total profit. (20)

3. Formulate the following nut-mix problem as a problem in linear programming:-

A manufacturer wishes to determine an optimal program for mixing three grades of nuts consisting of Cashew-nuts, hazel nuts and pea-nuts according to specifications of 3 different types of mixtures and their selling prices given in Table 1 and capacity limit on the inputs and buying prices given in Table 2.

Table 1

Mixture	Specifications	selling price Rs. per lb.
1	at least 50 per cent cashews, at most 25 per cent p. nuts	5
2	at least 25 per cent cashews, at most 50 per cent p. nuts.	3.5
3	nil	2.5

Table 2

Input	Capacity lb / day	Price Rs./lb.
Cashew-nut	100	6.5
Pea-nut	100	2.5
Hazel	60	3.5

Hint:- Let x_{ij} be the amount of jth type of nut in ith type of mixture i,j = 1,2,3. Profit = Sales price - buying price
There are 4 restrictions on specifications and 3 on capacities.
(20)

Please Turn Over

Group B.

1. A manufacturer produces two products A and B. The following information is given

	A	B
Rate of demand (units per unit time)	200	200
Set-up cost (in Rs.)	1000	100
Rate of production (units per unit time)	1000	4000
Inventory carrying cost (Rs. per unit per unit time)	1	1
Shortage cost (Rs. per unit per unit time)	1	1

Find the optimal order quantities for the two products

a) without any restriction,

b) if more than 1000 units cannot be stored due to the limitations on the warehouse facilities. (25)

2. The demand for a certain product has a continuous distribution with probability density function

$$f(x) = \frac{x}{50} \quad (0 \leq x \leq 10)$$

The storage cost is Rs.0.50 per unit closing inventory and shortage cost is Rs.2.00 per unit closing shortage. Find the optimal order quantity. (15)

INDIAN STATISTICAL INSTITUTE
Research and Training School
Annual Examination, 1965
B. Stat. IV Year

PROBABILITY

Duration: 3 hours | Maximum marks: 100 | Date: 24.6.1965

Even though the total score for the paper is specified as 100 questions have been set for 120. Choose questions for answering in such a way that the total of marks does not exceed 100.

---

1. Let $\phi(t)$ be the characteristic function of a probability distribution function $F(x)$ on the real line. Show that

   a) $|\phi(t)| \leq 1$;

   b) $\phi(t)$ is uniformly continuous in t ;

   c) if F has moments upto order k, then

   $$i^r \int x^r e^{itx} \, dF(x) = \frac{d^r \cdot \phi}{dt^r}.$$

   d) If X and Y are two real valued independent random variables defined on a probability space) such that the distribution of X+Y is degenerate show that the distributions of X and Y are also degenerate. (Hint: Show that, for every t, the variances of cos tX and cos tY are zero).

   (2 + 8 + 15 + 15) = (40

2. a) Define the notion of weak convergence for a sequence of probability distribution functions on the real line.

   b) Show that, if a sequence of probability distribution functions $F_n(x)$ converges weakly to a continuous probability distribution function $F(x)$, then

   $$\lim_{n \to \infty} \sup_{x} |F_n(x) - F(x)| = 0,$$

   where the supremum is taken over the entire real line.

   c) Suppose $F_n(x)$ and $F(x)$ are discrete probability distribution functions with jumps exactly at the points 1, 2, 3, ... of magnitude $P_{n1}, P_{n2}, P_{n3}, \ldots$ and $P_1, P_2, P_3 \ldots$ respectively. If $F_n(x)$ converges weakly to $F(x)$, show that

   $$\lim_{n \to \infty} \sum_{j=1}^{\infty} |P_{nj} - P_j| = 0$$

   (5 + 20 + 15) (40)

3. a) State precisely Lyapunov's central limit theorem for a sequence of (not necessarily identically distributed) independent (real valued) random variables.

   b) If $X_1, X_2, \ldots$ are independent binomially distributed random variables such that the probability for success of X_r is $P_r (r = 1, 2, \ldots)$ and $\sum_{r=1}^{\infty} P_r(1-P_r)$ diverges, show that the central limit theorem holds.

   c) Let $F_n(x)$ denote the distribution function of the random variable $\frac{X_n - n}{\sqrt{2n}}$ where X_n is distributed as χ^2_n (χ^2 with n degrees of freedom). Show that

   $$\lim_{n \to \infty} \sup_{x} \left| F_n(x) - \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-u^2/2} \, du \right| = 0$$

   (5 + 15 + 20) (40)

INDIAN STATISTICAL INSTITUTE
Research and Training School
Annual Examination, 1965
B. Stat. IV Year

STATISTICAL INFERENCE

Duration: 3 hours   Maximum marks: 100   Date: 22.5.1965

Answer any four questions

1.a) Discuss the role of complete sufficient statistics in

i) unbiased estimation of real parameters

and ii) construction of similar tests for composite hypotheses involving nuisance parameters.

(State the relevant theorems and prove them). (16)

b) $X_1, \ldots, X_n$ are independent observations from a normal population with mean θ and standard deviation 1. Find the minimum variance unbiased estimator of $\Phi(a - \theta)$ where a is a given number and Φ is the standard normal distribution function. (9)

2.a) n independent observations are available from a population with frequency function $f(X, \theta)$, it is known a priori that θ is contained in a finite set $\{\theta_1, \ldots, \theta_k\}$. Define the maximum likelihood estimator T_n of θ based on the sample. Show that there is an $h > 0$ such that

$$P_\theta [T_n \neq \theta] < k e^{-nh}$$ (15)

b) State (without proofs) the asymptotic properties of the maximum likelihood estimator of a real parameter when the parameter space contains a non-degenerate interval mentioning the regularity conditions that are involved. (10)

3. Write notes on :-

a) Sufficient statistics and exponential family of distributions. (9)

b) Estimation of parameters by the method of moments. (8)

c) Estimation of parameters by the method of minimum χ^2 (8)

4. $X_1, \ldots, X_n$ are independent observations from a normal population with mean ξ and variance σ^2, $Y_1, \ldots, Y_n$ are independent observations from a normal population with mean η and variance τ^2. Derive the uniformly most powerful similar test of size α for the null hypothesis $\sigma^2 = \tau^2$ against the alternative $\sigma^2 > \tau^2$. (State all the general propositions that you will use in course of the derivation). (25)

5.a) The probability distribution on a sample space involves an unknown real parameter θ. What is meant by the uniformly 'most accurate' lower confidence bound for θ with a given confidence coefficient? (7)

b) X is normally distributed with unknown mean θ and variance 1. (i) Express c in terms of α such that $X - c$ is a lower confidence bound for θ with confidence coefficient $1 - \alpha$. (ii) Show that when c is determined in this way, $X - c$ is the uniformly most accurate lower confidence bound for θ with confidence coefficient $1 - \alpha$. (18)

6.a) Obtain an approximate formula for the ASN function of the sequential probability ratio test (SPRT), for a simple hypothesis against a simple alternative with given error probabilities. (13)

b) $X_1, X_2, \ldots$ are independent observations from a normal population with mean θ and variance 1. To test $H_0: \theta = 0$ against $H_1: \theta = 2$, keeping the probabilities of the first and second kind of error at .05 and .10 respectively. Compare the ASN of the SPRT at $\theta = 0$ and $\theta = 2$ with the number of observations needed by the best fixed sample size test. (12)

INDIAN STATISTICAL INSTITUTE
Research and Training School
Annual Examination, 1965
B. Stat. IV Year
STATISTICS – Theory and Practical

Duration: 4 hours Maximum marks: 100 Date: 19.5.1965.

Attempt all questions.
Figures in the margin indicate marks allotted to different questions.

1.a) Describe the sign test for examining if the median of a continuous population is zero, on the basis of a random sample of size n drawn from the population. (5)

b) Compute the power function of the test when $n = 100$ and the alternative hypothesis is that the population is Normal with median μ and standard deviation unity for $\mu = 0.05\ (0.05)\ 0.25$ (8)

c) What is the most powerful test for the hypothesis that for a Normal population with standard deviation unity, the median is 0 against the alternative that it is $\mu > 0$?

Compute the power function of this test for samples of size $n=100$ and $\mu = 0.05\ (0.05)\ 0.25$. (7)

d) Determine the size of sample n for which the power of the test in (c) is approximately the same as the power of the sign test in (b) with $n = 100$ for $\mu = 0.05$. (5)

2. What are tolerance limits ? (5)

Work out the probability $\alpha = \alpha(n, \beta)$ that the range in a sample of size n from a population with a continuous density function covers a fraction β of the population. (10)

Find the value of n for which approximately $\alpha = 0.9$, $\beta = 0.8$. (15)

3. Define (Hoeffding's) U – statistic. (5)

Work out the mean and the variance in a general case of (a) U-statistic, describing the basic assumptions underlying your computations. (10)

Hence, or otherwise, work out the exact mean and variance of the variance in a sample of size n. (10)

Show that under certain conditions (to be stated by you) the limiting distribution of a normalised U-statistic is Normal. (10)

4. The result of a radio listener's sample survey as regards preference for types of music and age-group are tabulated below. Is preference for type of music influenced by age ?

type of music preferred	age-group 19–25	26–35	above 36
national	80	60	9
foreign	210	325	44
indifferent	16	45	132

(10)

INDIAN STATISTICAL INSTITUTE
Research and Training School
Annual Examination, 1965
B. Stat. IV Year

SAMPLE SURVEYS – THEORY

Duration: 3 hours Maximum marks: 100 Date : 20.5.1965

Answer any four questions

1.a) If a sample of size n is drawn from a population of size N, by simple random sampling without replacement, give the conventional estimator of the population total. Also deduce an unbiased estimator of the variance of this estimator. (3 + 9) = (12)

b) If from a simple random sample of n units drawn without replacement, a random sub-sample of n^* units is selected without replacement, duplicated, and added to the original sample, prove that the mean (t^*) based on the $n + n^*$ units is an unbiased estimate of the population mean. Also show that the variance of t^* is greater than the variance of the mean based on the original sample of n units by an approximate factor $(1 + 3n^* n^{-1})(1 + n^* n^{-1})^{-2}$. For what value of n^*/n does the relative loss in efficiency attain its maximum value ? (3 + 7 + 3) = (13)

2.a) Explain what you understand by the terms 'inclusion probability of a unit u_i' and 'inclusion probability of the pair (u_i, u_j)' in a sampling design.

b) Give an unbiased estimator of the population total of a characteristic Y as proposed by Horvitz and Thompson.

c) Derive an expression for the variance of the above estimator.

d) Show that, under the super-population set up, when auxiliary information on a correlated character X is available for all the units, there exists an optimum class of designs with given π_i's for which the conditional expectation of the above variance is uniformly minimised. (2 + 3 + 8 + 12) = (25)

3.a) For stratified simple random sampling (without replacement), to estimate the population mean, derive optimum allocation of a fixed total sample size n to the various strata.

b) If some of the optimum values n_i's obtained in (a) exceed the corresponding stratum sizes N_i's, how do you modify them, keeping n fixed ? What happens to the variance of the estimator by this modification ?

c) If we deviate from the optimum allocation by using a sample of size $\hat{n}_i$ in the ith stratum, show that the proportional increase in variance cannot exceed g^2, where g is the maximum deviation $|\hat{n}_i - n_i|$, expressed as a fraction of $\hat{n}_i$.

d) A population is divided into two strata of sizes N_1 and N_2 and a simple random sample of size n_i is taken from the ith stratum, $i = 1,2$. Suppose that optimum allocation of the total sample size n_1+n_2, results in the sample size n_i' for the ith stratum, $i = 1,2$, Show that the ratio of the optimum variance of the mean of a characteristic y to the actual variance is never less than

$\frac{4t}{(1+t)^2}$, where $t = (\frac{n_1}{n_2})(\frac{n_1'}{n_2'})^{-1}$. (6 + 5 + 7 + 7) = (25)

Please Turn Over

4. a) Explain how to take a sample of size n from a population of size N, by linear systematic sampling.

b) Give an unbiased estimator of population mean and prove its unbiasedness.

c) Suppose you have a sample of size n selected as in (a). Can you estimate the variance of the estimator of population mean? Give reasons. (3 + 5 + 6) = (14)

d) A first sample of size n_1 is drawn from a population of size N and the values of X are observed. From this sample a sub-sample of size n_2 is drawn and the values of Y are also observed. If the cost of enumerating a unit for X is c_1 and for Y is c_2, find the optimum values of n_1 and n_2 which minimize the variance of the estimate of the total of Y, for a given fixed total cost C. (11)

5. Write briefly on:

i) unbiased ratio estimators

ii) Self-weighting designs

iii) Sources of non-sampling errors

(10 + 9 + 6) = (25)

INDIAN STATISTICAL INSTITUTE
Research and Training School
Annual Examination, 1965
B. Stat. IV Year

SAMPLE SURVEYS PRACTICAL

Duration: 3 hours   Maximum marks: 100   Date: 20.5.1965

Answer all questions.

1. A survey is conducted to estimate the industrial output in India where the distribution of the factories according to the number of workers is given in the table below. The figures in the last two columns give the values as estimated in a previous survey.

sl. no.	size class (average no. of workers)	no. of factories	estimated output per factory (in '000 Rs.)	estimated standard deviation of output (in '000 s)
1	1 – 49	18260	100	80
2	50 – 99	4315	250	200
3	100 – 249	2233	500	600
4	250 – 999	1057	1760	1900
5	1000 and above	517	2250	2500

i) Taking these size classes as strata, compare the efficiencies of the following allocations of a sample of 3000 factories when the selection within each stratum is made with equal probability and with replacement.

a) allocation proportional to number of factories
b) allocation proportional to estimated output
c) optimum allocation. (7 + 7 + 7 + 3) = (24)

ii) Obtain the allocation of the sample size in (a), (b) and (c) above. (3 + 3 + 3) = (9)

iii) Compare the efficiency of stratified simple random sampling with optimum allocation with that of unstratified simple random sampling in the above case. (12)

2. A sample survey was conducted to estimate the total yield of paddy in a district. The design adopted was a stratified two-stage one with villages as first-stage units and plots within them as second-stage units. From each stratum 4 villages were selected with probabilities proportional to area and with replacement and 4 plots were selected from each selected village with equal probability and without replacement. The data on yield for the sample plots together with information on selection probabilities are given in the table below. (next page).

i) Obtain an unbiased estimate of the total yield of paddy in the district. (15)

ii) Obtain an unbiased estimate of the sampling variance of the estimate. (20)

iii) EITHER

Compare the efficiency of the above design with that of unistage simple random sampling with replacement of plots in each stratum. (20)

OR

If it were proposed to make the above design self-weighting so as to have a sample of 3 plots per village on an average:

a) find the common multiplier for this group of 3 strata,

b) obtain an estimate of the total yield of paddy in the district. (12 + 8) = (20)

Please Turn Over

stratum	sample village	inverse of probability of selection	total no. of plots	yield of sample plots 1	2	3	4
I	1	440.21	28	104	182	148	87
	2	660.43	14	108	64	132	156
	3	31.50	240	100	115	60	172
	4	110.38	76	348	350	157	119
II	1	21.00	256	124	111	135	216
	2	16.80	288	123	177	106	138
	3	24.76	222	264	78	144	55
	4	49.99	69	300	114	68	111
III	1	67.68	189	110	281	120	114
	2	339.14	42	80	61	118	124
	3	100.00	134	121	212	174	106
	4	68.07	161	243	116	314	129

Total number of plots in stratum I : 8423

" " " " II : 6353

" " " " III : 12853

INDIAN STATISTICAL INSTITUTE
Research and Training School
Annual Examination, 1965
B. Stat. I' Year

OPERATIONAL RESEARCH (THEORY AND PRACTICAL)

Duration: 4 hours. Maximum marks: 100 Date: 25.5.1965

Answer different groups in separate booklets

GROUP A

Maximum marks : 60

1.a) Explain the concept of duality in linear programming and the relationship between a pair of dual linear programming problems.

b) Write down the dual of the transportation problem of linear programming.

c) State the simplex optimality test to examine whether a given basic feasible solution of a linear programming problem
[Max $c'x$, $x \geq 0$ $Ax = b$] is optimal. (4 + 4 + 4) = (12)

2. State and formulate the following classical problems in linear programming

i) the travelling salesman's problem
ii) the production scheduling problem
iii) the assignment problem (3 + 3 + 3) = (9)

3. A potato merchant has a go-down capable of stocking 400 tons. He can order in the middle of a season, at prices given below for delivery in the beginning of the following season. During any season he can sell any amount up to his total stock at the beginning of the season at market prices below. If he starts the year with 200 tons at the beginning of the winter, how much should he plan to purchase and sell each season to maximise his profits. Solve this problem by Bellman's method of dynamic programming).

season	winter	spring	summer	autumn
cost prices	75	55	65	60
sales prices	65	15	60	70

(18)

4. Obtain an initial basic feasible solution to the following transportation problem. Examine whether it is optimal. Otherwise obtain another adjacent basic feasible solution having a cheaper value.

Table giving data regarding costs, supplies and demands.

origins	destinations 1	2	3	4	supplies
1	10	5	6	7	25
2	8	2	7	6	25
3	9	3	4	8	50
demands	15	20	30	35	

(15)

5. A machine costs Rs.9,000. Annual operating costs are Rs.400 for the first year and increase by Rs.800 every year. The machine has no resale or salvage value. Determine the age at which to replace the machine. (6)

Please Turn Over

GROUP B

Maximum marks: 40

1. A fan manufacturing company sells ceiling fans and table fans through retail outlets. The costs of procuring a ceiling fan and a table fan are Rs.150 and Rs.50 respectively. Holding and shortage costs are 5 per cent and 30 per cent of the cost respectively. The demand for each type is as follows:

demand/month	probability	
	ceiling fan	table fan
0	0.30	0.35
500	0.40	0.35
1000	0.20	0.25
1500	0.10	0.05

Find the optimal order quantities for the two products

i) if there is no restriction on total order quantity
ii) if the total order quantity is restricted to 1500
Formulae used are to be derived) (15)

2. a) For the case of one service station, Poisson arrivals with mean arrival rate λ and exponential service times with mean service rate μ, obtain the following:

i) steady-state probabilities
ii) distribution of the waiting time in queue and the expected waiting time in queue
iii) distribution of the waiting time in the system and the expected waiting time in the system. (6 + 6 + 6) = (18)

b) At a public telephone booth in a post office, arrivals are considered to be Poisson with an average inter-arrival time of 12 minutes. The length of a phone call may be assumed to be distributed exponentially with an average of 4 minutes.

i) Estimate the fraction of a day that the phone will be in use.
ii) What is the probability that an arrival will have to wait for more than 10 minutes before the phone is free ?
iii) What is the probability that it will not take more than 15 minutes for an arrival to wait for the phone and complete his call ? (1 + 3 + 3) = (7)

INDIAN STATISTICAL INSTITUTE
Research and Training School
Annual Examination, 1965
B. Stat. IV Year
ECONOMICS I.

Duration: 3 hours. Maximum marks: 100 Date 18.5.1965.

Answer each Group in separate answerscripts.
(Each question carries 20 marks)

GROUP A

Maximum marks: 60

1. EITHER

Linear arithmetic and linear logarithmic demand functions have been suggested as possible parametric forms to be used in statistical demand analysis. Can you suggest other possible forms of statistical demand functions. For what commodities would you use such functions ? Do they satisfy the homogeneity conditions ? (8 + 8 + 4) (20)

OR

Under what conditions is it advantageous to pool cross-section and time series samples in demand analysis? Do estimates from the two samples have the same meaning? (10 + 10) (20)

2. EITHER

Does a log normal distribution usually fit the distribution of personal incomes better than a normal distribution? If so, why? Do you think a normal distribution may give a better fir to the distribution of income in a group which is comparatively homogeneous with respect to occupation or some other relevant characteristic than to the distribution of income in the total population? Derive the estimates of the parameters of the log normal distribution assuming that you are supplied with grouped income distribution data.

(4 + 4 + 4 + 8) (20)

OR

Define the concept of specific concentration curve for consumption of a given commodity. Does it differ from the concept of the Lorenz curve for income or total expenditure? Work out analytical expressions for such curves assuming that the distributions of income is log logistic and the Engel curve has a constant elasticity. Show further that an estimate of the Engel elasticity for the specific commodity can be obtained from the areas under these two curves.

(4 + 4 + 8 + 4) (20)

3. EITHER

Describe briefly the method used by Murti and Sastry for estimating the value of capital in the major manufacturing industries of India. What are the main assumptions in their method and are they realistic enough? (10 + 10) (20)

OR

Write short notes on any THREE of the following:

i) The quality elasticity
ii) The CES production function
iii) The present state of income distribution data in India
iv) Multicollinearity (20)

GROUP B

Maximum marks: 40

Answer any two questions out of the following:

1. Explain concepts of efficiency locus and social welfare function and their relevance to planning. Consider a two-sector economy with Cobb-Douglas production functions and illustrate the problem of optimal allocation of resources for a one-period plan. (6 + 14) (20)

2. Discuss the nature of planning decisions which can be answered in the framework of the Mahalanobis two-sector model. (20)

3. Derive from the static Leontief system the various collections of producible net outputs, given the labour supply in the economy. How will your analysis be affected if there are capacity limitations in different sectors of the economy ? (12 + 8) (20)

INDIAN STATISTICAL INSTITUTE
Research and Training School
Annual Examination, 1965
B. Stat. IV Year
ECONOMICS II

Duration 3 hours. Maximum marks: 100 Date: 18.5.1965

1. The following table is condensed from a 36 X 36 - sector inter-industry transactions matrix of the Indian economy, 1955–56. Four broad sectors are distinguished: A. Primary production, B. Large scale manufacturing, C. Small scale manufacturing, and D. Other activities.

Table 1. Interindustry transactions of the Indian economy, 1955–56, (at Market prices).

Rs. crores

Sectors producing	Sectors consuming				Final demand
	Primary production	Large scale mfg.	Small scale mfg.	Other activities	
(0)	(1)	(2)	(3)	(4)	(5)
Primary production	1,842.22	461.03	377.88	367.09	3,825.15
Large scale manufacturing	66.12	649.77	282.07	445.31	1,187.14
Small scale manufacturing	56.93	14.26	78.50	158.24	1,224.77
Other activities	106.27	334.56	137.45	601.71	4,920.56
Labour input	4,768.83	1,627.21	656.80	4,528.20	
Gross output	6,840.37	3,086.83	1,532.70	6,000.55	

Work out the effect on the outputs of the given sectors if there is an increase of 5 per cent, 20 per cent, 15 per cent and 10 per cent simultaneously in the final demand for the products of the sectors A, B, C and D respectively.

Evaluate also the direct and indirect employment generated in each sector by a unit expansion in the final demand. (50)

2. The distribution of per capita monthly expenditure (X) in rural India as estimated by the National Sample Survey, 13 round (September 1957–May 1958) is given in Table 2.

Table 2. Distribution of monthly per capita expenditure in rural India, Sept. September 1957 – May 1958.

Per capita monthly expenditure (X)	Estimated per cent of population (100p)
Below Rs. 8	12.44
" Rs.11	28.79
" Rs.13	38.93
" Rs.15	48.89
" Rs.18	62.44
" Rs.21	71.61
" Rs.24	79.05
" Rs.28	85.68
" Rs.34	91.33
" Rs.43	95.84
" Rs.55	97.60
Above Rs.55	2.40
all	100.00

Please Turn Over

By plotting t_p defined by

$$p = \int_{-\infty}^{t_p} \frac{1}{\sqrt{2\pi}} e^{-\frac{u^2}{2}} du, \quad 0 < p < 1$$

against log X, examine graphically whether the distribution of log X is normal.

Estimate graphically the mean and the standard deviation of the distribution of log X for rural India (assuming normality). Obtain also estimates of the corresponding parameters of the distribution of X.

Derive an estimate of the Lorenz ratio and also calculate the proportion of total consumption accruing to the bottom and top deciles. (50)

INDIAN STATISTICAL INSTITUTE
Research and Training School
Annual Examination, 1965
B. Stat. IV Year

EDUCATIONAL STATISTICS

Duration: 3 hours | Maximum marks: 100 | Date: 20.5.1965.

Answer question 1 and any other five of the following.

1. Describe briefly the method of computation, properties and interpretation of any two of the following methods of standardizing test scores. (10 + 10) = (20)

a) standard or z score
b) percentile ranks
c) normalised scores

2. Write short notes on any four of the following

a) Item difficulty
b) Item discrimination
c) Correction for attenuation
d) Correction for guessing
e) Relevant validity. (4 + 4 + 4 + 4) = (16)

3.a) Give the formula for the rank correlation coefficient and explain why it is used in preference to the product moment correlation for some of the psychological data.

b) The reliability of a test is .85 and the s.d of the obtained scores is 8.0. Find

i) standard error measurement
ii) standard deviation of true scores. (8 + 8) = (16)

4.a) What is the effect of increasing the test length K times on the variance of the observed scores and the standard error of measurement.

b) A test in mathematics has reliability of .80 and a test in English has reliability of .64. The intercorrelation of the two tests is .50.

Estimate the degree of intercorrelation if

i) the Mathematics test alone is made perfectly reliable
ii) the English test alone is made perfectly reliable. (8 + 8) = (16)

5.a) What are the assumptions underlying Kuder Richardson Formula 20

$$\frac{n}{n-1}\left(1-\frac{\sum_{i=1}^{n} p_i q_i}{\sigma_t^2}\right)$$

used for estimating the reliability of a test where

n = no. of items
p_i = the difficulty value of the ith item
$q_i = 1 - p_i$
σ_t = the standard deviation of the whole score.

b) A test of 50 items has a reliability of .7 and validity of .5. If another 150 comparable items are added what will be the validity ?

(8 + 8) = (16)

Please Turn Over

6.c) Show how the method of analysis of variance can be used to estimate the reliability of a test. (8 + 8) = (16)

b) For item criterion correlation, when items are scored 0 and 1 and the criterion is a continuous variable which of the following coefficient would you find and why.

i) Biserial, ii) Point biserial, iii) Phi,
iii) Tetra-choric, v) any other

7. Describe the errors of measurement, substitution and prediction in psychological tests, giving the equations defining them and their standard deviations. Explain the meaning of symbols used in the equation. (16)

8.a) Describe a function of reliability which is invariant with regard to test length.

b) Why should the split-half method of reliability be used for 'power test' only ? (5 + 6 + 5) = (16)

c) What is meant by 'item analysis' of a test ?

9.a) It is known that a limited group has a standard deviation of 8.0 and a reliability coefficient of .85 for a test, what will be the reliability coefficient in a more variable group whose standard deviation is 10.0 ?

b) When the validity coefficients of several tests are known, how should the test scores be combined so as to maximise the prediction of the criterion ? Briefly describe the procedure. (8 + 8) = (16)

INDIAN STATISTICAL INSTITUTE
Research and Training School
Annual Examination, 1965
B. Stat. IV Year

GENETICS

Duration: 2 hours | Maximum marks : 50 | Date: 26.5.1965

Attempt all questions. The figures in the margin indicate marks allotted to different questions

1. EITHER

Describe the mechanism of inheritance of sex-linked characters. Why is it said that in respect of inheritance of a sex linked recessive malady, human males are at a disadvantage ? (10)

OR

What is polymorphism ? Describe the genetics of the human blood-group system O – A – B – AB. (10)

2. Show that frequencies of the three phenotypes controlled by a single pair of genes stabilise in one generation of pan-mixia (random mating). Given that in a random saple of size n from a population the frequencies of the phenotypes AA, Aa and aa are f_1, f_2 and f_3 respectively, $f_1 + f_2 + f_3 = n$, how will you examine whether the frequencies are in agreement with the hypothesis of pan-mixia? If so, how will you estimate the frequency of the gene a ? (10)+(10)=(20)

3. Calculate the expected frequencies of the phenotypes AB, Ab, aB and ab amongst the offsprings of a cross AaBb × AaBb, when there is linkage, and both the parents are in the coupling phase.

Estimate the coefficient of linkage from the following F_2 data:

AB	Ab	aB	ab
75	14	14	11

Also obtain the standard error of the estimate. (8) + (6) + (6) = (20)

Central Statistical Orgainsation

Training Course in Official Statistics for B. Stat. & M. Stat. students of the I.S.I. Calcutta, 1964.

Final Examination

Attempt any five questions choosing at least three from Section A.

All questions carry equal marks.

7 August 1964 — Time 3 hours

SECTION A

1. Describe the need for and the functions of a National Statistical System.

Or

Give a detailed account of the functions and organisation of the Central Statistical Organisation.

2. Discuss the salient features of the 1961 Population Census of India.

3. Write a comprehensive note on the recent improvements in Agricultural Statistics with special reference to Crop Estimation.

4. Describe the coverage, content and the methods of collection of statistics of Manufacturing Industries in India.

5. Describe the method of construction of either

   a) Index of Industrial Production

   Or

   b) Index of Wholesale Prices

6. Write a detailed note on any two of the following :

   1. Foreign Trade Statistics
   2. Working Class Consumer Price Index
   3. Financial Statistics
   4. Educational Statistics

7. Describe the method used for computing National Income in India from any one of the following sectors :

   1. Agriculture
   2. Unorganised Service sector
   3. Mining and Manufacturing

SECTION B

1. Distinguish between the following :

   1. The directly standardised death rate and the indirectly standardised death rate.
   2. Total fertility rate and gross reproduction rate.

2. What are the various methods of Population Projection ? Describe fully the component method.

(P.T.O.)

3. Describe how you would use an econometric model, an inter-industry model and a linear programming model for national economic planning.

4. Distinguish between the fixed base and chain base methods of index number construction. Give their relative merits and demerits.

5. Discuss the comparative merits of sample surveys and complete enumeration in any scheme of data collection.