A introduction to
M.Sc. Statistics

Mrinal Nandi

Department of Statistics, West Bengal State University, Barasat.
e-mail : mrinal.nandi1@gmail.com

June 4, 2023

Contents

Chapter 1

Real Analysis

Syllabus:

1. Real number system, cluster points of sets, closed and open sets, compact sets.

2. Bolzano-Weierstrass property, Heine-Borel property and its applications.

3. Sequences and Series of functions, pointwise convergence, uniform convergence, absolute convergence. Some tests of convergence.

4. Continuity, uniform continuity.

5. Differentiability of univariate and multivariate functions. Mean value Theorem.

6. Reimann integral and its properties. Reimann-Stieltjes integral.

7. Review of sequence and series of functions. Uniform convergence: term by term differentiation and integration Power series. Taylor series expansion.

1.1 Exercise

1. Prove that a convergent sequence is bounded. Is the converse true? Justify your answer.

2. Prove that an uniformly continuous function is continuous but the converse is not true.

3. Give an example of real valued differentiable function defined on $\mathbb{R}$ such that the derivative is not continuous at 0. Justify your answer.
4. Define Open Set in R. Prove that intersection of two open set is also an open set.
5. Prove that a finite set has no limit point.
6. Give an example of set which is closed as well as open. Justify.
7. Give an example of infinite set which have no limit point.
8. Prove that $\mathbb{Q}$ is dense in $\mathbb{R}$.
9. Prove that the sequence $\{1/n, n \in \mathbb{N}\}$ is a Cauchy sequence.
10. Prove that product of two odd function is an even function.
11. Define directional derivatives of a real valued function of n variables.
12. Find the directional derivatives of the following function along $(1, 1, \ldots, 1)$ $f : \mathbb{R}^n \to \mathbb{R}$, such that, $f(x_1, x_2, \ldots, x_n) = x_1 + 2x_2 + \cdots + nx_n$.
13. Optimize, $x^2 + y^2 + z^2$ subject to $x + y + z = 3$.

1. Give an example of real valued differentiable function defined on $\mathbb{R}$ such that the derivative is not continuous at 0. Justify your answer. 5
2. a) Define a power series and the radius of convergence of the power series? 3
   b) Find the radius of convergence of the following series $\sum_{n=1}^{\infty} \frac{x^n}{n!}$ and $\sum_{n=1}^{\infty} \frac{(-x)^n}{n!}$ 4
   c) Find the dimensions of a box of largest volume that can be inscribed in a unit sphere. 3
3. Prove that a convergent sequence is bounded. Is the converse true? Justify your answer. 5
4. Give an example of open set which has no limit points, Justify.
5. a) A set E subset of X is called dense in X if every point of X is a limit point of E. Prove that $\mathbb{Q}$ is dense in $\mathbb{R}$. 4
   b) Prove that an uniformly continuous function is continuous but the converse is not true. 3
   c) Prove that product of two odd function is an even function. 3

6. a) Prove by definition, without using any standard result, $sinx \to 1$ as $x \to \pi/2$ 3
   b) Prove that the sequence $\{1/n, n \in \mathbb{N}\}$ is a Cauchy sequence. 3
   c) Find the extreme points of the function $f = x_1^3 + x_2^3 + 2x_1^2 + 4x_2^2 + 6$. 4

7. a) Define directional derivatives of a real valued function of n variables. 3
   b) Find the directional derivatives of the following function along $(1, 1, \ldots, 1)$ $f : \mathbb{R}^n \to \mathbb{R}$, such that, $f(x_1, x_2, \ldots, x_n) = x_1 + 2x_2 + \cdots + nx_n$. 4
   c) Optimize, $x^2 + y^2 + z^2$ subject to $x + y + z = 3$. 3

8. a) Prove that a finite set has no limit point. 4
   b) Give an example of set which is closed as well as open. Justify. 3
   c) Give an example of infinite set which have no limit point. 3

9. a) A set E subset of X is called dense in X if every point of X is a limit point of E. Prove that $\mathbb{R} \setminus \mathbb{Q}$ is dense in $\mathbb{R}$. 3
   b) Prove that a monotonically decreasing sequence which is bounded below converges. 3
   c) Prove that, uniform continuity implies continuity. Give an example of continuous function which is not uniformly continuous. 4

10. a) Prove by definition, $f(x) = x^2$, $\infty < x < \infty$, is a continuous function. 3
    b) Prove that the sequence $\{2/n^2, n \in \mathbb{N}\}$ is a Cauchy sequence. 3
    c) Prove that the union of finite closed sets is a closed set. 4

11. a) Prove that, the open interval $(0, 1)$ is an open set. 3
    b) Give an example of set which has only two limit points. Justify your answer. 4
    c) Prove that the harmonic series diverge to ∞. 3

12. a) Give an example of set which is not closed as well as not open. 2
    b) State Heine Borel property of compact subsets of real numbers. 2
    c) Give an example of two non-convex subset in $\mathbb{R}^2$ such that whose union is a convex set. Justify your answer. 3
    d) Prove that the sequence $\{1/n^2, n \in \mathbb{N}\}$ is a convergent sequence. 3

13. a) Define open set. Give two examples of open set? 3
    b) Find points of inflection, max. and min of the following function

$$f(x) = \begin{cases} x^4 & \text{if } -4 < x \leq 2, \\ x & \text{if } 2 < x \leq 4, \\ 4 & \text{if } 4 < x \leq 6, \end{cases}$$

3
c) Prove that an open interval is an open set. 4

14. a) Prove that any closed interval can be expressed as intersection of infinitely many open intervals. 4
b) Define a series. When it is called convergent? 3
c) Prove that the sequence $\{2/n, n \in \mathbb{N}\}$ is a convergent sequence. 3

15. a) Give an example of power series whose radius of convergence is 1. 2
b) Give an example, with Justification, of closed set which has no limit points. 2
c) Prove that, the open interval (a, b) is an open set. 3
d) Give an example of infinite set has only one limit point. Justify. 3

Group B: Answer any 4 questions. $5 \times 4 = 20$

16. a) State sandwitch theorem on limit.
Using this theorem prove that $\lim_{n\to\infty}(3^n + 2^n)^{\frac{1}{n}} = 3$. 3
b) Give an example of two convex subset in $\mathbb{R}^2$ such that whose union is a not a convex set. Justify your answer. 2

17. a) Define compact set. Give examples. 2
b) Write $[a, b]$ as union and complement of the following type of intervals $(a, b]$, Where $a < b$ two real numbers. 3

18. a) Give an example of set which is not closed also not open. 2
b) Prove that a monotonically decreasing sequence which is bounded below converges. 3

19. Find the infinite sum: $\frac{1}{1.2.3} + \frac{1}{2.3.4} + \frac{1}{3.4.5} + \cdots + \cdots \infty$.

20. Let $f : I \to \mathbb{R}$ is a differentiable function and $c \in I$. If $f'(c) > 0$ then prove that f is increasing at c. Show that the converse is not true.

21. When a series is called convergent. Find the lim $\frac{1.2.3...(2n-1)}{2.4.6....2n}$ as $n \to \infty$.

Chapter 2

Complex Analysis

1. a) What do you mean by an analytic function? State and prove a necessary condition for a complex valued function to be analytic. Give an example to show that the condition is not sufficient. 7
   b) Define open disc in a complex plane and give an example. 3

2. a) If $f(z)$

$$= \begin{cases} \frac{|z|^4}{z^3} & \text{if } z \neq 0 \\ 0 & \text{if } z = 0, \end{cases}$$

   then show that the Cauchy-Riemann equations are satisfied at $z = 0$ although the derivative of f does not exist at $z = 0$. 7

   b) Find modulus and amplitude of the Complex number $-2i - 2$. Find the value of i^i. 3

3. Deducing necessary results prove that an analytic function $f(z)$ must be a constant if the real part of $f(z)$ is a constant. 5

4. a) Illustrate Cauchy's Integral formula with an example. Hence deduce Cauchy's inequality. 7
   b) Show that $\{z^n\}$ is a null sequence for $|z| < 1$. 3

5. a) Expand $f(z) = sinz$ in a Taylor series about $z = 0$. Hence or otherwise expand $f(z) = \frac{sin(z)}{z^3}$ in a Laurent series, valid for $|z| > 0$ and evaluate

$$\int_{|z|=1} f(z)dz \qquad 7$$

   b) Discuss with examples: Singularity and Poles. 3

6. If $f(z)$ is continuous in a domain D admitting an anti-derivative evaluate $\int_\gamma f(z)dz$ for any closed contour γ lying inside D. Evaluate $\int_\gamma (z+2)e^z dz$ when γ is the parabola $\pi^2 y = x^2$ from $(0,0)$ to $(\gamma, 1)$. 7
   b) Discuss with examples: Laurent Series Expansion of a complex valued function. 3

7. State a necessary and sufficient condition for a continuous function $f(z)$ to be analytic in a region D. Prove that the complex conjugate of an analytic function is also analytic. Define a harmonic function. Prove that if $f(z)$ is analytic then the real and imaginary parts of $f(z)$ are harmonic. 10

8. (a) Evaluate $\int_\gamma \frac{|z|^2}{z} dz$ when (i) γ is a straight line from $z = 0$ to $z = 1 + i$ and (ii) γ is the arc of the parabola $y = x^2$, $z = 0$ to $z = 1 + i$.
   (b) Define limit of a complex valued function. What will be the value of $lim_{z \to 0} \frac{Re(z)}{Im(z)}$ when (i) z approaches 0 along the line $y = x$ and (ii) z approaches 0 along the imaginary axis. Hence comment on the existence of the limit. 10

9. State the residue theorem and discuss with examples the different methods of obtaining the residues. Identifying the poles of the function $f(z) = \frac{1}{z(z^2-3z+2)}$ evaluate $\int_\gamma f(z)dz$ when $(i)\gamma = \{z : |z| = \frac{1}{2}\}$ $(ii)\gamma = \{z : |z| = \frac{4}{3}\}$ and $(iii)\gamma = \{z : |z| = \frac{5}{2}\}$. [1+3+6]

10. Consider $f(z) = \frac{\bar{z}}{z}, z \neq 0$. Examine whether $\lim_{z \to 0} f(z)$ exist or not.

11. Define continuity of a complex function. Consider the following function

    $f(z)$

$$= \begin{cases} \frac{z^2-3z-10i}{z+(1+2i)} & \text{if } z \neq -(1+2i) \\ 0 & \text{otherwise ,} \end{cases}$$

    Check the continuity of f at $-(1+2i)$.

Chapter 3

Linear Algebra

Syllabus:

1. Vector spaces with real field. Basis dimension of vector space.
2. Orthogonal vectors, Gram-Schmidt orthogonalization.
3. Linear transformation of matrices. Matrix operations.
4. Elementary matrices and their uses.
5. Rank of a matrix and related results. Determinants. Inverse of a matrix.
6. System of linear equations: homogenous and non-homogenous system.
7. Generalized inverse: Moore-Penrose.
8. Idempotent matrices and its properties.
9. Characteristic roots and vectors.
10. Quadratic forms and canonical reduction.
11. Singular value decomposition.

3.1 Exercise

1.

Chapter 4

Probability I

Syllabus:

1. Random Variables: Definition of discrete and continuous random variables. Cumulative distribution function and its properties, probability mass function and probability density function. Expectation and moments, Dispersion, Skewness, Kurtosis and Quantiles. [5] Bivariate probability distributions. Marginal and conditional distributions. Independence. Conditional moments. [7] Correlation and Regression. Generating Functions : probability generating function and moment generating function in univariate and bivariate cases. [5] Probability Inequalities : Chebyshev's lemma, Markov's Chebyshev's inequalities. Some common univariate distributions. Bivariate Normal distributions and its properties. [6] Limit Theorems: Convergence in distribution: De-Moivre-Laplace limit theorem and Normal approximation to the Poisson distribution. [2] Sigma fields in probabily. Borel sigma field. Measures and its properties. Probability as a measure. Measurable functions and its properties. Random variable as a measurable function . Integration of a measurable functions. [10] Sequence of measurable functions: Monotone convergence theorem, Fatou's lemma and Dominated convergence theorem and their probabilistic aspects. Radon-Nikodym theorem and its applications. Distribution functions: application of Lebesgue-Stieltje's measure. Expectation and inequalities.

4.1 Exercise GEC Probability

1. For two random variables X and Y, $E(X) = 8$, $E(Y) = 6$, $var(Y) = 36$ and $r_{XY} = 0.5$. Find i) $E(XY)$, ii) $cov(X, X + Y)$, iii) $var(2X - 2Y)$

2. Define probability density function of a random variable X. Is the following a probability density function?
$$\begin{aligned} f(x) &= x/2 \quad , \quad 0 < x \leq 1 \\ &= \tfrac{1}{2} \quad , \quad 1 < x \leq 2 \\ &= \tfrac{3-x}{2} \quad , \quad 2 < x \leq 3 \\ &= 0 \quad , \quad \text{otherwise} \end{aligned}$$

3. For mutually exclusive events $A_1, A_2, \dots, A_n$, prove that $P(\cup_{i=1}^n A_i) = \sum_{i=1}^n P(A_i)$. What will be the value of $P(\cup_{i=1}^n A_i)$ if $A_1, A_2, \dots, A_n$ are mutually exclusive and exhaustive events.

4. Let $B_1, B_2, \dots, B_n$ be exhaustive and mutually exclusive events with $P(B_i) > 0, i = 1, 2, \dots, n$. Show that for any event A, $P(A) = \sum_{i=1}^n P(B_i)P(A|B_i)$. 5+5

5. A random variable X has pmf $f(x) = \begin{cases} \frac{x}{21} & , \text{ for } x = 1, 2, \dots, 6 \\ 0 & , \text{ otherwise} \end{cases}$. Find $P(\frac{1}{2} < X < \frac{5}{2}|X > 1)$.

6. Write down the p.d.f of Normal(μ, σ^2) distribution. Show that this distribution is symmetric. Calculate its median and mode.

7. Find the mean deviation about mean of X where $X \sim N(0, \sigma^2)$. 5+5

8. Define CDF of a random variable. Write down the properties to be satisfied by a CDF.

9. For a Binomial distribution prove that $cov(X, n - X) = -npq$, where the notations have their usual meaning. 5+5

10. Suppose $P(A) = p_1$, $P(B) = p_2$ and $P(A \cap B) = p_3$. Show that $P(A^c \cap B^c) = 1 - p_1 - p_2 + p_3$.

11. For a Binomial distribution with parameters n and p, establish the following relationship
$$\mu_{r+1} = pq(nr\mu_{r-1} + \frac{d\mu_r}{dp})$$

12. For a normal distribution with mean 3 and variance 16, find the value of y of the variate such that the probability of the variate lying in the interval $(3, y)$ is 0.4772. [You are given $P(Z \leq 2) = 0.9772$].

13. If A and B are independent events and $P(A) = 0.4, P(B) = 0.8$, then find $P(A - B)$.

14. State Boole's inequality for 3 events A, B and C.

15. If A and B are independent events with $P(A) > 1/2, P(B) > 1/2, P(A \cap B^c) = 3/25$ and $P(B \cap A^c) = 8/25$, find out the values of $P(A)$ and $P(B)$.

16. State and prove Bayes' theorem.

17. For any four events A_1, A_2, A_3, prove the following inequalities
(i) $P(A_1 \cup A_2 \cup A_3) \leq P(A_1) + P(A_2) + P(A_3)$
(ii) $P(A_1 \cup A_2 \cup A_3) \geq P(A_1) + P(A_2) + P(A_3) - P(A_1 \cap A_2) - P(A_2 \cap A_3) - P(A_3 \cap A_1)$
Give two examples where above two equality hold.

18. Two balls are drawn without replacement from a bag containing 6 red, 8 green and 10 black balls. What is the Probability of drawing one red and one green ball?

19. If A and B are independent events with $P(A) > 1/2, P(B) > 1/2, P(A \cap B^c) = 3/25$ and $P(B \cap A^c) = 8/25$, find out the values of $P(A)$ and $P(B)$.

20. The p.m.f. $f(x)$ of a discrete rv X assuming values $0, 1, 2, \ldots$ satisfies the relation
$f(x + 1) = \frac{a}{x+1} f(x), a > 0, x = 0, 1, 2, \ldots$ determine a and $f(x)$. Find the expectation of X.

21. (a) The C.V. of a Poisson distribution is 25%. Find its mean and s.d.
(b) For a nonnegative random variable X, show that $\sqrt{(E(X))} \geq E(\sqrt{X})$
(c) If $P(A) = 1/2, P(B^c) = 1/4$ and $P(A^c \cap B) = 5/11$, find $P(A|B^c)$.

22. Two fair dice are thrown. If the sum of two numbers obtained is 8, then the Probability that the first number is 6 will be (a) 1/6 (b) 5/36 (c) 1/4 (d) none of these.

23. If the occurrence of an event A implies that of B, then
(a) $P(B) \leq P(A)$ (b) $P(B) < P(A)$ (c) $P(A) \leq P(B)$ (d) $P(A) < P(B)$.

24. The probability of having at most one tail in 3 tosses of a fair coin is
(a) 3/8 (b) 1/8 (c) 1/2 (d) 7/8.

25. What is the Probability of drawing one white ball from a bag containing 6 red, 8 green and 10 black balls?

26. If $P(A|B) = 1/4$, then what is the value of $P(A^c|B)$.

27. A problem in mathematics is given to 3 students A_1, A_2, A_3, whose chances of solving it are $1/4, 1/2, 3/4$ respectively. What is the Probability that the problem will be solved by at least one of the students?

28. For three events A_1, A_2, A_3, state and prove Bonferroni's inequality.
Or, If an integer X is randomly selected from the first 50 positive integers, then find the value of $P\left(X + \frac{96}{X} > 50\right)$.

29. If $P(A) = 1/2, P(B) = 1/3$ and $P(A^c \cap B^c) = 5/12$, find $P(A|B)$.

30. If X is a continuous random variable, then the value of $P(X = 3)$ is..........

31. A random variable X takes only two values 0 and 1 where $P(X = 1) = 2/3$. Find $E(X)$.

32. A man gets n^2 rupees for the n dots appearing on rolling a single unbiased die. Determine expected gain from a single throw.

33. Three coins, whose faces are marked as 1 and 2, are tossed. What is the expectation of the total value (sum) of numbers on their faces?

34. Define expectation of a random variable. X and Y are two independent random variables with $E(X) = E(Y)$, show that $E(X(X - Y)) = Var(X)$.

35. If the standard deviation of a Poisson variable is 2.5, its mode is

    (a) 2 (b) 1 (c) 6 (d) 3.

36. If the mean of a binomial distribution $B(n, p)$ is $n/2$, then skewness of distribution is

    (a) positive (b) negative (c) symmetric (d) none of these.

37. For a normal distribution with parameters μ and σ^2, β_2 is

    (a) 0 (b) 3 (c) 2 (d) 1.

38. The third order central moment of Poisson distribution having parameter 4 is

    (a) 4 (b) 2 (c) 16 (d) none of these.

39. Let X follow normal distribution with mean 45 and standard deviation 5 and $\Phi(1) = 0.84$, then $P(X \leq 40)$ is

    (a) 0.16 (b) 0.84 (c) 0.68 (d) none of these.

40. State true or False: The mean and variance of a binomial distribution are 5 and 16 respectively.

41. Fill in the gap: The S.D. of standard normal variable is ....

42. Write true or false: If X is a negative random variable such that $|E(X)| = 4$, then give the values of $E(X)$.

43. The C.V. of a Poisson distribution is 25%. Find its mean.

44. If the mean of a Poisson distribution is 1.5, its mode is.........

45. A random variable X is such that $P(X = 1) = P(X = 0) = 1/2$. Find $E(X^2)$.

46. If X is a Poisson variable such that $P(X = 2) = 9P(X = 4) + 90P(X = 6)$, find $P(X \geq 2)$.

47. The p.m.f. $f(x)$ of a discrete rv X assuming values $0, 1, 2, \ldots$ satisfies the relation

    $f(x + 1) = \frac{a}{x+1} f(x), a > 0, x = 0, 1, 2, \ldots$ determine $f(x)$.

48. If $E(X) = 3, E(X(X - 1)) = 22$, then $\text{Var}(7 - 2X)$ is equal to

    (a) 16 (b) 64 (c) 32 (d) none of these.

49. Write true or false: Values of a random variable are always positive.

50. Write true or false: Expectation of a random variable cannot be negative.

51. Write true or false: For a negative random variable X, $E(X)$ may be positive.

52. Write true or false: Binomial distribution can never be symmetric.

53. Write true or false: If $V(X) = 0$, then all values of X are equal.

54. What do you mean by probability distribution of a discrete random variable.

55. When is a random variable said to have a binomial distribution?

56. For a random variable X, show that $(E(X^2))^{\frac{1}{2}} \geq E(X)$.

57. Define binomial distribution.

58. Let f be a function such that $f(x) = ax, (x = 1, 2, \ldots, l)$. For what value of a will this $f(x)$ be the probability mass function of a discrete random variable x? Find the expectation of x.

59. Arithmetic mean and standard deviation of a binomial distribution are respectively 4 and $\sqrt{8/3}$. Find the values of n and p.

60. Write down the probability mass function of a Poisson distribution. State its arithmetic mean and standard deviation.

61. If the 3rd quartile of standard normal variable is 0.675. Find the quartile deviation of that variable.

62. Write down the p.m.f. of Poison(5) distribution.

4.2 Exercise

1. If A_n is an increasing sequence of subset of probability space $(\Omega, \mathcal{F}, P)$, prove that $\lim P(A_n) = P(\lim A_n)$. [5]

2. Let X be a random variable defined on $(\Omega, \mathcal{F})$, and a, b be reals. Then show that $aX + b$ is also an random variable on $(\Omega, \mathcal{F}, P)$. [5]

3. Prove that the distribution function F is right continuous and $\lim_{x\to\infty} F(x) = 1$. [5]

4. Show that the minimum sigma-field containing the class $C = \{(a, b] : \infty < a < b < \infty\}$ generates all Borel sets of $\mathbb{R}$. [5]

5. What is meant by sigma-finite measure? Show that the Lebesgue measure defined on the class of Borel sets of the real line is sigma-finite. [5]

6. Define measurable function on measure space $(X, \mathcal{B})$. Show that if $\{f_n\}$ be a sequence of measurable functions, then $sup f_n$ is measurable. [5]

7. Suppose $2r$ balls are distributed at random into r boxes. Let X_i denote the number of balls in box i.
(a) Find the joint distribution of $X_1, X_2, \ldots, X_r$.
(b) Find the probability that each box contains exactly 2 balls.
(c) Find the the distribution of $X_1|X_2$. 3+3+4

8. Let X and Y be independent random variables having geometric distribution with 0.3 and 0.4 respectively. Find
(a) $P(X \geq Y)$
(b) $P(X = Y)$
(c) distribution of $\min(X, Y)$. 3+3+4

9. (a) For a probability space $(\Omega, \mathcal{F}, P)$ and the sequence of sets $\{A_n\}$ in $\mathcal{F}$, if $\Sigma_n P(A_n) < \infty$, show that $P(lim\ sup\ A_n) = 0$.

   (b) For a probability space $(\Omega, \mathcal{F}, P)$, let $\mathcal{F}_0(\subset \mathcal{F})$ be a field of subsets of Ω and consider any set $A \in \sigma(\mathcal{F}_0)$ (= minimum σ-field generated by the set A). Then show that for any $\epsilon > 0$, there exists a set $T_\epsilon \in \mathcal{F}_0$ such that $P(A \Delta T_\epsilon) < \epsilon$, Δ being the symmetric difference. What is the implication of this result? [4+6]

10. (a) Let $(\Omega, \mathcal{B}, \mu)$ be a measure space with finite measure μ. Show that for any sequence of set $\{A_n\}$ in $\mathcal{B}$:
    (i) $\liminf \mu(A_n) \geq \mu(\liminf A_n)$ and
    (ii) $\limsup \mu(A_n) \leq \mu(\limsup A_n)$.
    Clearly mention where finiteness of the measure are required.
    Hence or otherwise show that $A_n \to A$ implies $\mu(A_n) \to \mu(A)$

    (b) Show that extended real-valued measureable function$f$ is the limit of a sequence $\{f_n\}$ of simple finite valued functions. [6+4]

Chapter 5

Generalized Inverse

1. Define G-inverse. Prove that every matrix has a g-inverse. Is this unique? Justify.

2. Let G be a g-inverse of A. Then the system $Ax = b$ is consistent if and only if $AGb = b$.

3. What is the Moore-Penrose inverse of a null matrix of order 2×3.

4. Let G be a g-inverse of A. Then column space of B is a subspace of column space of A if and only if $AGB = B$.

5. If A, G be matrices of order $m \times n$ and $n \times m$ respectively. Then prove that G is a g-inverse of A if and only if for any $y \in \mathcal{C}(A), x = Gy$ is a solution of $Ax = y$.

6. Prove that if G is a g-inverse of A then Rank(A)= Rank(GA)= Rank(AG).

7. Define reflexive g-inverse of a matrix. Prove that if G is a g-inverse of A then Rank(A)= Rank(G).

8. Define minimum norm g-inverse of A. Prove that G is a minimum norm g-inverse of A if and only if for any $y \in C(A), x = Gy$ is a solution of $Ax = y$ with minimum norm.

9. Define least square g-inverse of A. Prove that G is a least square g-inverse of A if and only if for any $x, y, ||AGy - y|| \leq ||Ax - y||$.

10. Let $A = BC$ be a rank factorization. Let B_l^- be a left inverse of B and C_r^- be a right inverse of C. Show that $C_r^- B_l^-$ is a g-inverse of A.

11. Show that a matrix which is not a square matrix or not a nonsingular matrix admits infinitely many g-inverses.

12. Let A be a matrix of order $m \times n$ and G be a g-inverse of A, and let $y \in \mathcal{C}(A)$. Prove that the class of solutions of $Ax = y$ is given by $Gy + (I - GA)z$, where z is arbitrary.

13. Define Moore-Penrose inverse. Show that it exist and unique.

14. Let A be a matrix of order $m \times n$ and G be a g-inverse of A. Prove that the class of all g-inverse of A is given by $G + (I - GA)U + V(I - AG)$, where U, V are arbitrary.

15. Find the set of all g-inverses of the following matrix:

$$\begin{bmatrix} 1 & 2 & 1 \\ 0 & 1 & 1 \\ 1 & 3 & 2 \end{bmatrix}$$

Find a g-inverse of the above matrix which does not contain any zero entry.

16. Show that the class of g-inverses of

$$\begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}$$

is given by

$$\begin{bmatrix} 1 + a + c & a + d \\ b + c & b + d \end{bmatrix}$$

where a, b, c, d are arbitrary.

17. Let A be a matrix of order $m \times n$, let rank $A = r$, and let $r \leq k \leq \min\{m, n\}$. Show that A has a g-inverse of rank k. In particular, show that any square matrix has a nonsingular g-inverse.

18. Find the minimum norm solution of the system of equations:

$$2x + y - z = 1, x - 2y + z = -2, x + 3y - 2z = 3.$$

19. Find the Moore-Penrose inverse of

$$\begin{bmatrix} 2 & 4 \\ 3 & 6 \end{bmatrix}.$$

20. Let x be an $n \times 1$ vector. find the g-inverse of x which is closest to the origin.

21. Let X be a matrix of order $m \times n$ and let $y \in \mathbb{R}^n$. Show that the orthogonal projection of y onto $\mathcal{C}(X)$ is given by $X(X'X)^{-}X'y$ for any choice of the g-inverse.

22. For any matrix X, show that $X^{+} = (X'X)^{+}X'$ and $X(X'X)^{-}X' = XX^{+}$.

23. Let A be a matrix of order $m \times n$, let P, Q be matrices of order $r \times m$. Show that $PA = QA$ if and only if $PAA' = QAA'$.

24. Let A, G be matrices of order $m \times n, n \times m$ respectively. Show that G is a mimimum norm g-inverse of A if and only if $GAA' = A'$.

Chapter 6

Linear Model

Content

1. Gauss-Markov model. Estimable functions. Best linear unbiased estimator (BLUE).

2. Gauss Markov Theorem. Estimation space and error space.

3. Sum of squares due to a set of linear functions.

4. Estimation with correlated observations. Least Square estimation with linear restriction on the parameters.

5. General linear hypothesis.

6.1 Preliminaries

Let $\boldsymbol{y}$ be a column vector with components $y_1, y_2, \ldots, y_n$. We call $\boldsymbol{y}$ a *random vector* if each y_i is a random variable. The expectation of $\boldsymbol{y}$, denoted by $E(\boldsymbol{y})$, is the column vector with components $E(y_1), E(y_2), \ldots, E(y_n)$. Clearly,

$$E(B\boldsymbol{x} + C\boldsymbol{y}) = BE(\boldsymbol{x}) + CE(\boldsymbol{y}) \ldots (1)$$

where $\boldsymbol{x}$, $\boldsymbol{y}$ are random vectors and B, C are constant nonrandom matrices of proper order.

Quick Exercise

1. What is the order of B and C?
2. Prove (1).

If $\boldsymbol{x}$ and $\boldsymbol{y}$ are random vectors of order m and n respectively, then the *covariance matrix* between $\boldsymbol{x}$ and $\boldsymbol{y}$, denoted by $\operatorname{cov}(\boldsymbol{x},\boldsymbol{y})$, is an $m \times n$ matrix whose (i, j)-entry is $\operatorname{cov}(x_i, y_j)$.

The *dispersion matrix*, or the *variance-covariance matrix* of $\boldsymbol{y}$, denoted by $D(\boldsymbol{y})$, is defined to be $\operatorname{cov}(\boldsymbol{y}, \boldsymbol{y})$.

If b and c are constant nonrandom column vectors of proper order then

$$\operatorname{cov}(b'\boldsymbol{x}, c'\boldsymbol{y}) = b'\operatorname{cov}(\boldsymbol{x}, \boldsymbol{y})c \ldots (2)$$

Setting $\boldsymbol{x} = \boldsymbol{y}$ and $b = c$ gives

$$\operatorname{var}(b'\boldsymbol{x}) = b'D(\boldsymbol{x})b \ldots (3).$$

Since varience is nonnegative we conclude that $D(\boldsymbol{x})$ is positive semidefinite.

Recall that $\operatorname{var}(b'\boldsymbol{x}) = 0$ if and only if there exists a linear combination $b'\boldsymbol{x}$ which is constant with probability one.

From (2) we get

$$\operatorname{cov}(B\boldsymbol{x}, C\boldsymbol{y}) = B\operatorname{cov}(\boldsymbol{x}, \boldsymbol{y})C' \ldots (4).$$

Quick Exercise

1. What is the order of b and c?
2. Prove (2) and (4).
3. Prove that the dispersion matrix is symmetric.
4. When $D(\boldsymbol{x})$ is positive definite?

6.2 Linear Model

Consider the random variables $y_1, y_2, \ldots, y_n$ such that the distribution of these random variables is controlled by some unknown parameters. In a linear model, the basic assumption is that $E(y_i)$ is a linear function of the parameters $\beta_1, \beta_2, \ldots, \beta_n$ with known coefficients. In matrix notation this can be expressed as

$$E(\boldsymbol{y}) = X\beta \ldots (5)$$

where $\boldsymbol{y}$ is the $n \times 1$ column vector with components $y_1, y_2, \ldots, y_n$; X is known nonrandom matrix of order $n \times p$ and β is a $p \times 1$ column vector with components $\beta_1, \beta_2, \ldots, \beta_p$.

The matrix X called as *Disign matrix*.

Homoscedasticity assumptions:
(i) $y_1, y_2, \ldots, y_n$ are uncorrelated and (ii) $\mathrm{var}(y_i) = \sigma^2$ for all i.

Hence

$$D(\boldsymbol{y}) = \sigma^2 I \ldots (6)$$

Another way to write the model (5) is

$$\boldsymbol{y} = X\beta + \boldsymbol{e} \ldots (7)$$

where the random vector $\boldsymbol{e}$ satisfies the homoscedasticity assumptions $E(\boldsymbol{e}) = 0$ and $D(\boldsymbol{e}) = \sigma^2 I$.

Our objective is to find estimates of $\beta_1, \beta_2, \ldots, \beta_p$ and thier linear combinations; and to find estimates of σ^2.

Quick Exercise

1. Prove that (7) $\implies$ (5).

2. Prove (6).

3. Prove that (5) $\implies$ (7).

4. Consider the model:
   $E(y_1) = \alpha + \gamma$
   $E(y_2) = \alpha - 4\beta + 2\gamma$
   $E(y_3) = -6\alpha + 2\beta + \gamma$
   where y_1, y_2, y_3 are independent with common variance σ^2 and α, β, γ are parameters. Write the linear model in matrix notation and also describe the order of all matrices and vectors.

   Is the design matrix invertible?

5. Consider the linear model

$$E(y_1) = \beta_1 + \beta_2, E(y_2) = 2\beta_1 - \beta_2, E(y_3) = \beta_1 - \beta_2,$$

where y_1, y_2, y_3 are uncorrelated with a common variance σ^2.

(i) Find two different linear functions of y_1, y_2, y_3 which are unbiased for β_1. Determine their variances and covariance between the two.

(ii) Find two linear functions of y_1, y_2, y_3 which are unbiased for β_2 and uncorrelated.

(iii) Write the model in terms of the new parameters $\theta = \beta_1 + 2\beta_2, \theta_2 = \beta_1 - 2\beta_2$.

6.3 Estimability

Consider the linear model

$$E(\boldsymbol{y}) = X\beta, \quad D(\boldsymbol{y}) = \sigma^2 I \ldots (8)$$

where $\boldsymbol{y}$ is the $n \times 1$, X is $n \times p$ and β is $p \times 1$.

The linear parametric function $l'\beta$ is said to be *estimable* if there exists a linear function $c'\boldsymbol{y}$ of the observations $y_1, y_2, \ldots, y_n$, such that $E(c'\boldsymbol{y}) = l'\beta$ for all $\beta \in \mathbb{R}^p$.

Now the condition $E(c'\boldsymbol{y}) = l'\beta$ is equivalent to the condition $c'X\beta = l'\beta$.

Now we have the following facts:
Fact (1): If $Ax = Bx$ for all $x \in \mathbb{R}^p$ then $A = B$.
Fact (2): $c'X = l'$ if and only if $l' \in \mathcal{R}(X)$.

Hence $l'\beta$ is estimable if and only if $l' \in \mathcal{R}(X)$.
Notation: $\mathcal{R}(X)$ is the row space of X and $\mathcal{C}(X)$ is the column space of X

Quick Exercise

1. Prove the fact (1) and (2).

2. Consider the model:
   $E(y_1) = \alpha + \beta + 2\gamma$
   $E(y_2) = 2\alpha + \beta - \gamma$
   $E(y_3) = 2\alpha + 2\beta + \gamma$
   where y_1, y_2, y_3 are independent with common variance σ^2.

   Find a, b, c such that $a\alpha + b\beta + c\gamma$ is estimable.

3. Consider the linear model $E(y) = X\beta, D(y) = \sigma^2 I$ and let x_i be the i-th column of X, for $i = 1, 2, \ldots, p$. Prove that the function $l_1\beta_1 + l_2\beta_2$, where l_1, l_2 are not both zero, is estimable if and only if x_1, x_2 do not belong to the linear span of $l_2x_1 - l_1x_2, x_3, \ldots, x_p$.

4. Consider the model $E(y) = X\beta, D(y) = \sigma^2 I$. If X has full column rank, prove the following:
(a) Every function $l'\beta$ is estimable.
(b) $X'X$ is non singular.

5. Consider three independent random variables, Y_1, Y_2, Y_3 having common variance σ^2 and expectations $EY_1 = \mu_2 + \mu_3$, $EY_2 = \mu_1 + \mu_2$ and $EY_3 = \mu_1 + \mu_3$. Determine the condition of estimability of parametric function $l_1\mu_1 + l_2\mu_2 + l_3\mu_3$.

6. Consider the linear model

$$E(y_1) = 2\beta_1 - \beta_2 - \beta_3, E(y_2) = \beta_2 - \beta_4, E(y_3) = \beta_2 + \beta_3 - 2\beta_4$$

with usual assumptions. Determine the estimable functions.

Some results from generalized inverse

The following facts concerning generalized inverse are important in linear model:

1. For any matrix X, $\mathcal{R}(X) = \mathcal{R}(X'X)$ and $\mathcal{C}(X') = \mathcal{C}(X'X)$.

2. The matrix AC^-B is invariant under the choice of the g-inverse C^- of C if $\mathcal{C}(B) \subset \mathcal{C}(C)$ and $\mathcal{R}(A) \subset \mathcal{R}(C)$.

3. The matrix $X(X'X)^-X'$ is invariant under the choice of the ginverse of X.

4. The matrix $X(X'X)^-X'X = X$ and $X'X(X'X)^-X' = X'$.

5. Any choice of g-inverse, $(X'X)^-X'$ is a least square g-inverse of X.

6.4 Best Linear Unbiased Estimate

Let $l'\beta$ be an estimable function. We say a linear function $c'y$ is *BLUE(best linear unbiased estimate)* if $E(c'\boldsymbol{y}) = l'\beta$ i.e., $c'\boldsymbol{y}$ is an unbiased estimate of $l'\beta$ and has minimum variance among all unbiased estimates of $l'\beta$.

Theorem 6.4.1 *Let $l'\beta$ be an estimable function and let G be a least square g-inverse of X. Then $l'G\boldsymbol{y}$ is a BLUE of $l'\beta$ and var$(l'G\boldsymbol{y}) = \sigma^2 l'(X'X)^- l$.*

Proof: Since $l'\beta$ is estimable, $l' = u'X$ for some u. Then

$$E(l'G\boldsymbol{y}) = u'XGX\beta = u'X\beta = l'\beta \ldots (9)$$

and hence $l'G\boldsymbol{y}$ is an unbiased estimate for$l'\beta$.

Let $c'\boldsymbol{y}$ be any other unbiased estimate for $l'\beta$, i. e., $E(c'\boldsymbol{y}) = l'\beta$.

Put $w' = c' - l'G$, hence $E(w'\boldsymbol{y}) = 0$. Which implies $w'X\beta = 0$.

Since the above relation true for all β, we have $w'X = 0$.

So any unbiased estimate of $l'\beta$ is of the form $(l'G + w')\boldsymbol{y}$, where $w'X = 0$.

Now using (3) we get

$$\text{var}\,((l'G + w')\boldsymbol{y}) = (l'G + w')D(\boldsymbol{y})(l'G + w')' = \sigma^2(l'G + w')(G'l + w).$$

Since $l' = u'X$ we have

$$\text{var}\,((l'G + w')\boldsymbol{y}) = \sigma^2(u'XG + w')(G'X'u + w) \ldots (10).$$

Since G is a least square g-inverse of X, we have $G'X' = (XG)' = XG$, and hence

$$u'XGw = uG'X'w = 0 \ldots (11)$$

and therefore from (10) and (11) we get

$$\text{var}\,((l'G + w')\boldsymbol{y}) = \sigma^2\,(l'GG'l + w'w) = \text{ var}(l'G\boldsymbol{y}) + \sigma^2 w'w \ldots (12).$$

Hence

$$\text{var}\,((l'G + w')\boldsymbol{y}) \geq \text{var}(l'G\boldsymbol{y}) \ldots (13).$$

Now equality holds in (13) if and only if $w'w = 0$ i.e., $w = 0$, hence BLUE is unique.

Therefore $l'G\boldsymbol{y}$ is the BLUE of $l'\beta$.

We know if $(X'X)^+$ be the Moore-Penrose g-inverse of $X'X$ then $(X'X)^+X'$ is a least square g-inverse of X and hence

$$\text{var}(l'G\boldsymbol{y}) = \text{var}(l'(X'X)^+X'\boldsymbol{y}) = \sigma^2 l'(X'X)^+X'(l'(X'X)^+X')' = \sigma^2 l'(X'X)^+ l \ldots (14).$$

Now $l'(X'X)^- l = u'X(X'X)^- X'u$, which is invariant with respect to the choice of g-inverse, hence $l'(X'X)^- l = u'X(X'X)^- X'u = u'X(X'X)^+X'u = l'(X'X)^+ l$.

Hence the theorem.

Example Consider the model $E(y_{i,j}) = \alpha_i + \beta_j, \quad i = 1,2; j = 1.2$. The model can be expressed in standard form as $E(\boldsymbol{y}) = X\beta$, where

$$\boldsymbol{y} = \begin{bmatrix} y_{11} \\ y_{12} \\ y_{13} \\ y_{14} \end{bmatrix}, \quad X = \begin{bmatrix} 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{bmatrix} \quad \text{and} \quad \beta = \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \beta_1 \\ \beta_2 \end{bmatrix}.$$

Let S be the set of all vectors (l_1, l_2, m_1, m_2) such that $l_1 + l_2 = m_1 + m_2$. Now we have the following observations:

Observation 1: $\mathcal{R}(X) \subset S$

Observation 2: $\dim(S) = 3$ and rank of X is 3.

From Observation 1 and 2 we get

Observation 3: $\mathcal{R}(X) = S$.

Hence $l_1\alpha_1 + l_2\alpha_2 + m_1\beta_1 + m_2\beta_2$ is estimable if and only if $l_1 + l_2 = m_1 + m_2$.

Next we have the following results:

Result 1:

$$X'X = \begin{bmatrix} 2 & 0 & 1 & 1 \\ 0 & 2 & 1 & 1 \\ 1 & 1 & 2 & 0 \\ 1 & 1 & 0 & 2 \end{bmatrix}$$

Result 2:

$$(X'X)^{-} = \frac{1}{4}\begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 4 & -2 & -2 \\ 0 & -2 & 3 & 1 \\ 0 & -2 & 1 & 3 \end{bmatrix}$$

Result 3:

$$X(X'X)^{-}X' = \frac{1}{4}\begin{bmatrix} 3 & 1 & 1 & -1 \\ 1 & 3 & -1 & 1 \\ 1 & -1 & 3 & 1 \\ -1 & 1 & 1 & 3 \end{bmatrix}$$

Now the BLUE of any estimable function $u'X\beta$ is $u'X(X'X)^{-}X'y$.

For example, if $u = (1,0,0,0)'$ then $u'X\beta = \alpha_1 + \beta_1$ and $u'X\beta$ is $u'X(X'X)^{-}X'\boldsymbol{y} = \frac{1}{4}(3y_{11} + y_{12} + y_{21} + y_{22}) \ldots (*)$.

Hence the BLUE of $\alpha_1 + \beta_1$ is $\frac{1}{4}(3y_{11} + y_{12} + y_{21} + y_{22})$.

Now the variance of the BLUE is $\sigma^2 u'X(X'X)^{-}X'u = \frac{3\sigma^2}{4} \ldots (**)$.

Quick Exercise

1. Prove (12) and (13).

2. Prove every steps of (14).

3. Prove Observation 1, 2 and 3.

4. Consider the model:
$E(y_1) = \alpha - \beta + \gamma$
$E(y_2) = \alpha + \beta + 2\gamma$
$E(y_3) = -\alpha + 2\beta + \gamma$
where y_1, y_2, y_3 are independent with common variance σ^2. Find a, b, c such that $a\alpha + b\beta + c\gamma$ is estimable. Find the BLUE for the estimable functions. Find variance for the BLUE.

5. Analyze the model $E(y_{i,j}) = \alpha_i + \beta_j, i = 1, 2; j = 1.2$. Find the BLUE of $\alpha_1 + \beta_1$.

6. Show that the BLUE of an estimable function is unique i.e., if $l'\beta$ is an estimable function and if $c'y, d'y$ are both BLUE of the function, then $c = d$.

7. Prove Results 1, 2 and 3.

8. Prove (*) and (**) in the above example.

9. Find the BLUE and the variance of the BLUE of $\alpha_1 + \alpha_2 + \beta_1 + \beta_2$.

10. Consider the linear model $E(y) = X\beta, D(y) = \sigma^2 I$ and suppose y, X are partition as
$$y = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}, X = \begin{bmatrix} X_1 \\ X_2 \end{bmatrix}.$$
The model thus broken into two models: $E(y_1) = X_1\beta, D(y_1) = \sigma^2 I$ and $E(y_2) = X_2\beta, D(y_2) = \sigma^2 I$. How does one combine the BLUE of an estimable function under the two models to arrive at the BLUE under the original model.

11. Consider the model $E(y_1) = 2\beta_1 + \beta_2, E(y_2) = \beta_1 - \beta_2, E(y_3) = \beta_1 + \alpha\beta_2$ with usual assumptions. Determine α so that the BLUEs of β, β_2 are uncorrelated.

6.5 Regression Model

The model (8) is said to be a full rank model or regression model if X has full column rank, i.e., rank$X = p$.

For a regression model we have the following results:

(A) $\mathcal{R}(X) = \mathbb{R}^p$.

(B) Every function $l'\beta$ is estimable.

(C) $X'X$ is nonsingular.

Theorem 6.5.1 (Gauss-Markov Theorem) *Consider the regression model $E(\boldsymbol{y}) = X\beta, D(\boldsymbol{y}) = \sigma^2 I$. Let $\hat{\beta}_i$ be the BLUE of β_i and $\hat{\beta}$ be the column vector with components $\hat{\beta}_1, \hat{\beta}_2, \ldots, \hat{\beta}_p$. Then $\hat{\beta} = (X'X)^{-1}X'\boldsymbol{y}$ and dispersion matrix of $\hat{\beta}$ is $\sigma^2(X'X)^{-1}$ and the BLUE of $l'\beta$ has variance $\sigma^2 l'(X'X)^{-1}l$.*

Proof: Since $X'X$ is nonsingular, $(X'X)^- = (X'X)^{-1}$. We have the BLUE of $l'\beta$ is $l'(X'X)^{-1}X'\boldsymbol{y}$.

Put l_i be the $p \times 1$ column vector whose i-th component is 1 and rest are 0. Then $\beta_i = l_i'\beta$ and hence $\hat{\beta}_i = l_i'(X'X)^{-1}X'\boldsymbol{y}$ for all $i = 1, 2, \ldots, p$.

Therefore, $\hat{\beta} = I_p(X'X)^{-1}X'y = (X'X)^{-1}X'\boldsymbol{y}$.

Next the dispersion matrix of $\hat{\beta}$ is

$$(X'X)^{-1}X'D(\boldsymbol{y})\left((X'X)^{-1}X'\right)' = (X'X)^{-1}X'(\sigma^2 I_n)X(X'X)^{-1} = \sigma^2(X'X)^{-1} \ldots (15).$$

Finally the BLUE of $l'\beta$ has variance $\sigma^2 l'(X'X)^{-1}l$.

A result from linear algebra

Ler X be an $n \times n$ matrix and suppose $|x_{ij}| \leq 1$ for all i, j. Then $\det(X'X) \leq n^n$.

6.5.1 Weighing Design

Suppose four objects are to be weighed using an ordinary balance (without bias) with two pans. We are allowed four weighings. In each weighing we may put some of the objects in the right pan and rest in left pan. Let $\beta_1, \beta_2, \beta_3, \beta_4$ be the true weights of the objects. Define $x_{ij} = 1$ or -1 depending upon whether we put the j-th object in the right pan or in the left pan in the i-th weighing. Let y_i denote the weight needed to achieve balance in the i-th weighing. If the sign of y_i is positive, then the weight is required in the left pan, otherwise in the right pan.

Then we have the model $E(\boldsymbol{y}) = X\beta$, where $X = (x_{ij})$ is the 4×4 design matrix, $\boldsymbol{y} = (y_1, y_2, y_3, y_4)'$ and $\beta = (\beta_1, \beta_2, \beta_3, \beta_4)$.

Assumption: (a) y_i's are uncorrelated with common variance σ^2.

(b) $X'X$ is nonsingular.

Now observe the followings:

(a) $X'X$ is positive semidefinite.

(b) $|X'X| \leq 4^4$.

(c)

$$\text{For} \quad X = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 \\ 1 & -1 & 1 & -1 \end{bmatrix} \quad \text{we have} \quad |X'X| = 4^4$$

(d) X is a Hadamard matrix, i.e., each entry is either 1 or -1 and the rows are orthogonal.

The dispersion matrix of $\hat{\beta}$ is $\sigma^2(X'X)^{-1}$.

Hence the minimum possible determinant of the dispersion matrix is $\frac{\sigma^2}{4^4}$.

Quick Exercise

1. Prove the results (A), (B) and (C).
2. Prove every steps of (15).
3. Prove the observations (a), (b), (c) and (d).
4. Given a linear equation of x and y, discuss the regression model?

6.6 Residual Sum of Squares

Consider the linear model given in (7), $\boldsymbol{y} = X\beta + \boldsymbol{e}$. The vector $\boldsymbol{e}$ is known as *error vector*.

Let y and e be an observation from $\boldsymbol{y}$ and $\boldsymbol{e}$ respectively.

Hence we have $y = X\beta + e$. The *least square method* is to minimized $e'e$ with respect to β.

Observed that $X\beta$ is a vector in the column space of X and y is not a vector in that space unless $e = 0$. So idea is that $e'e$ will be minimum when e is orthogonal to the column space of X, that is e is orthogonal to every columns of X.

Which implies $X'e = 0 \ldots (\#)$.

Hence we have $X'X\beta = X'y \ldots (16)$.

The above equation known as the *Normal equation.*

Since $\mathcal{C}(X') = \mathcal{C}(X'X)$, the normal equation is consistent.

Now if $\hat{\beta}$ is a solution of the normal equation then $\hat{\beta} = (X'X)^{-}X'y$ for some choice of the g-inverse.

The *residual sum of squares* (RSS) is defined to be

$$y - X\hat{\beta})'(y - X\hat{\beta}).$$

Observe that, the RSS is invariant under the choice of the g-inverse $(X'X)^{-}$ although $\hat{\beta}$ may depend on the choice ... (observation @).

Theorem 6.6.1 *The minimum of $(y - X\beta)'(y - X\beta)$ is attained at $\hat{\beta}$.*

Proof: We have

$$(y - X\beta)'(y - X\beta) = (y - X\hat{\beta} + X\hat{\beta} - X\beta)'(y - X\hat{\beta} + X\hat{\beta} - X\beta).$$

Now

$$(X\hat{\beta} - X\beta)'(y - X\hat{\beta}) = (\hat{\beta} - \beta)'X'(y - X(X'X)^{-}X'y) = 0 \ldots (+)$$

Therefore

$$(y - X\beta)'(y - X\beta) = (y - X\hat{\beta})'(y - X\hat{\beta}) + (\hat{\beta} - \beta)'X'X(\hat{\beta} - \beta) \ldots (++)$$
$$\implies (y - X\beta)'(y - X\beta) \geq (y - X\hat{\beta})'(y - X\hat{\beta})$$

and equality holds when $\beta = \hat{\beta}$.

Theorem 6.6.2 *If Rank$(X) = r$, then $E(\boldsymbol{y} - X\hat{\beta})'(\boldsymbol{y} - X\hat{\beta}) = (n - r)\sigma^2$.*

Proof: We have

$$E\left((\boldsymbol{y} - X\beta)'(\boldsymbol{y} - X\beta)\right) = D(\boldsymbol{y}) = \sigma^2 I.$$

Thus

$$E(\boldsymbol{y}'\boldsymbol{y}) = E(\boldsymbol{y})\beta'X' + X\beta E(\boldsymbol{y}') - X\beta\beta'X' + \sigma^2 I = X\beta\beta'X' + \sigma^2 I.$$

Put $P = I - X(X'X)^{-}X'$.

Observe that P is symmetric, idempotent and $PX = 0 \ldots (17)$.
Now

$$E(\boldsymbol{y} - X\hat{\beta})'(\boldsymbol{y} - X\hat{\beta}) = E(\boldsymbol{y} - X(X'X)^{-}X'\boldsymbol{y})'(\boldsymbol{y} - X(X'X)^{-}X'\boldsymbol{y}) = E(\boldsymbol{y}'P'P\boldsymbol{y}).$$

We have $\boldsymbol{y}'P'P\boldsymbol{y} = \boldsymbol{y}'P^2\boldsymbol{y} = \boldsymbol{y}'P\boldsymbol{y} = trace(\boldsymbol{y}'P\boldsymbol{y}) = trace(P\boldsymbol{y}\boldsymbol{y}')$. Hence

$$E(\boldsymbol{y} - X\hat{\beta})'(\boldsymbol{y} - X\hat{\beta}) = E(trace(P\boldsymbol{y}\boldsymbol{y}')) = trace(E(P\boldsymbol{y}\boldsymbol{y}')) = trace(PE(\boldsymbol{y}\boldsymbol{y}')).$$

Now since $PX = 0$ we have

$$PE(\boldsymbol{y}\boldsymbol{y}') = P(X\beta\beta'X' + \sigma^2 I) = PX\beta\beta'X' + \sigma^2 P = \sigma^2 P.$$

Therefore $E(\boldsymbol{y} - X\hat{\beta})'(\boldsymbol{y} - X\hat{\beta}) = \sigma^2 traceP$.

Finally,

$$traceP = n - trace(X(X'X)^{-}X' = n - trace((X'X)^{-}X'X) = n - rank((X'X)^{-}X'X)$$

since $(X'X)^{-}X'X$ is idempotent.

However $rank((X'X)^{-}X'X) = rank(X'X) = rankX = r$.

Hence $traceP = n - r$ and the proof is complete.

Conclusion (1): RSS$/(n - r)$ is an unbiased estimator of σ^2.

Conclusion (2): RSS$=\boldsymbol{y}'\boldsymbol{y} - \hat{\beta}'X'X\hat{\beta} = \boldsymbol{y}'\boldsymbol{y} - \boldsymbol{y}'X\hat{\beta}$.

Quick Exercise

1. Prove every steps of (#).

2. Prove observation @.

3. Prove conclusion (1) and (2).

4. Prove (+) and (++).

5. Prove (17).

6. Consider the model:
   $E(y_1) = \alpha + \beta + \gamma$
   $E(y_2) = \alpha + \beta + 2\gamma$
   $E(y_3) = 2\alpha + 2\beta + \gamma$
   where y_1, y_2, y_3 are independent with common variance σ^2.

   Find a, b, c such that $a\alpha + b\beta + c\gamma$ is estimable. Find the BLUE for the estimable functions. Find variance for the BLUE. Find RSS and an unbiased estimator of σ^2.

7. Analyze the one way classified model $E(y_{ij}) = \alpha_i + \epsilon_{ij}, \quad i = 1, \dots, k; j = 1, \dots, n_i$, where ϵ_{ij} are independent with mean 0 and variance σ^2. Find RSS and an unbiased estimator of σ^2.

8. Suppose $x_i, y_i, z_i, i = 1, 2, \dots, n$ are $3n$ independent observations with common variance σ^2 and expectations given by $E(x_i) = \theta_1, E(x_2) = \theta_2, E(z_i) = \theta_1 - \theta_2, i = 1, 2, \dots, n$. Find BLUEs of θ_1, θ_2 and compute the RSS.

9. Consider the model $E(y_i) = \theta_i, i = 1, 2, 3, 4$; where y_i are uncorrelated with variance σ^2. Suppose we have the restriction $\theta_1 + \theta_2 + \theta_3 + \theta_4 = 0$ on the parameters. Find RSS. Moreover assume that $\theta_1 = \theta_2$. Then show that the RSS is $4\bar{y}^2 + \frac{1}{2}(y_1 - y_2)^2$.

10. Consider the model $E(y_1) = \beta_1 + \beta_2, E(y_2) = 2\beta_2, E(y_3) = \beta_1 - \beta_2$ with usual assumptions. Find the RSS.

6.7 One-way Classification

Consider the model

$$y_{ij} = \alpha_i + e_{ij}, \quad i = 1, 2, \dots, k; \quad j = 1, 2, \dots, n_i,$$

where e_{ij} are independent with mean 0 and variance σ^2.

Now the model can be written as

$$y = X\beta + e,$$

where $y = (y_{11}, \dots, y_{1n_1}, \dots, y_{k1}, \dots, y_{kn_k})'$;

$e = (e_{11}, \dots, e_{1n_1}, \dots, e_{k1}, \dots, e_{kn_k})'$;

$\beta = (\alpha_1, \dots, \alpha_k)'$ and

X is a $(n = \sum_{i=1}^k n_i) \times k$ matrix whose first n_1 rows are $(1, 0, 0, \dots, 0)$; next n_2 rows are $(0, 1, 0, \dots, 0)$; and so on. Thus

$$X'X = diag(n_1, \dots, n_k) \dots (18)$$

and

$$X'y = (y_{1\cdot}, \dots, y_{k\cdot})' \dots (19)$$

where

$$y_{i\cdot} = \sum_{j=1}^{n_i} y_{ij}, \quad i = 1, 2, \dots, k.$$

Now observe that (i) the model is full column rank and (ii) the BLUE of α_i are given by the components of

$$\hat{\alpha} = (X'X)^{-1}X'y = (\bar{y_1}, \ldots, \bar{y_k})'$$

where $\bar{y_i} = \frac{y_i}{n_i}, \quad i = 1, \ldots, k.$

Now

$$RSS = y'y - \hat{\alpha}'X'y = \sum_{i=1}^{k}\sum_{j=1}^{n_i} y_{ij}^2 - \sum_{i=1}^{k} \frac{y_{i\cdot}^2}{n_i} \ldots (20).$$

Note that (i) the rank of X is k and hence (ii) $E(RSS) = (n-k)\sigma^2$, where $n = \sum_{i=1}^{k} n_i$.

Quick Exercise

1. Prove every steps of (18) and (19).
2. Prove the observations (i), and (ii).
3. Prove every steps of (20).
4. Prove notes (i) and (ii).
5. Consider the one-way model:

$$y_{ij} = \mu + \alpha_i + e_{ij}, \quad i = 1, \ldots, k; j = 1, \ldots, n_i;$$

   where e_{ij} are independent with mean 0 and variance σ^2. What are the estimable functions? Is it correct to say that the grand mean $\bar{y}_{..}$ is an unbiased estimator of μ?

6.8 RSS under restriction

Consider the linear model $E(y) = X\beta, D(y) = \sigma^2 I$, where y is $n \times 1$, X is $n \times p$.

Suppose we have a linear restriction $L\beta = z$ on the parameters. We assume that $\mathcal{R}(L) \subset \mathcal{R}(X)$ and $z \in \mathcal{C}(L)$, that is the equation $L\beta = z$ is consistent.

Let $\hat{\beta} = (X'X)^{-}X'y$ for a fixed g-inverse $(X'X)^{-}$ and let

$$\bar{\beta} = \hat{\beta} - (X'X)^{-}L'(L(X'X)^{-}L')^{-}(L\hat{\beta} - z).$$

Theorem 6.8.1 *The minimum of $(y - X\beta)'(y - X\beta)$ subject to $L\beta = z$ is attained at $\beta = \bar{\beta}$.*

Proof: Since $\mathcal{R}(L) \subset \mathcal{R}(X)$ and $\mathcal{R}(X'X) = \mathcal{R}(X)$, we have $L = WX'X$ for some W. Let $T = WX'$. Now

$$L(X'X)^{-}L' = TT' \ldots (21).$$

Since $z \in \mathcal{C}(L)$, $Lv = z$ for some v. Thus

$$L(X'X)^{-}L'(L(X'X)^{-}L')^{-}z = TT'(TT')^{-}TXv = z \ldots (22),$$

$$L\hat{\beta} = Ty \ldots (23),$$

and

$$L(X'X)^{-}L'(L(X'X)^{-}L')^{-}L\hat{\beta} = TT'(TT')^{-}Ty = Ty \ldots (24).$$

Using (22), (23) and (24) we get $L\bar{\beta} = z$.

Now for any β satisfying $L\beta = z$,

$$(y - X\beta)'(y - X\beta) = \left(y - X\bar{\beta} + X(\bar{\beta} - \beta)\right)' \left(y - X\bar{\beta} + X(\bar{\beta} - \beta)\right)$$

Claim: $(\bar{\beta} - \beta)'X'(y - X\bar{\beta}) = 0$.
Hence

$$(y - X\beta)'(y - X\beta) = (y - X\bar{\beta})'(y - X\bar{\beta}) + (\bar{\beta} - \beta)'X'X(\bar{\beta} - \beta)$$

Therefore

$$(y - X\beta)'(y - X\beta) \geq (y - X\bar{\beta})'(y - X\bar{\beta})$$

and equality holds if and only if $\beta = \bar{\beta}$.

To proof the claim we first observe that

$$X'X\bar{\beta} = X'X\hat{\beta} - X'X(X'X)^{-}L'(L(X'X)^{-}L')^{-}(L\hat{\beta} - z).$$

$$= X'y - L'(L(X'X)^{-}L')^{-}(L\hat{\beta} - z) \text{ since } L' = X'XW' \ldots (25).$$

Hence

$$X'(y - X\bar{\beta}) = L'(L(X'X)^{-}L')^{-}(L\hat{\beta} - z).$$

Since $L\bar{\beta} = L\hat{\beta} = z$, it follows that

$$(\bar{\beta} - \beta)'X'(y - X\bar{\beta}) = (\bar{\beta} - \beta)'L'(L(X'X)^{-}L')^{-}(L\bar{\beta} - z) = 0 \ldots (26).$$

Which complete the proof. *An example* Consider the model $E(y_i) = \theta_i,\ i = 1, 2, 3, 4;$

where y_i are uncorrelated with variance σ^2.

Suppose we have the restriction $\theta_1 + \theta_2 + \theta_3 + \theta_4 = 0$ on the parameters.

The model in standard form has $X = I_4$. The restriction on the parameters can be written as $L\theta = 0$, where $L = (1, 1, 1, 1)$ and $\theta = (\theta_1, \theta_2, \theta_3, \theta_4)'$. Thus

$$\hat{\theta} = (X'X)^{-1}X'y = y$$

and

$$\bar{\theta} = \hat{\theta} - (X'X)^{-1}L'\left(L'(X'X)^{-1}L'\right)^{-1}L\hat{\theta} = y - (\bar{y}, \bar{y}, \bar{y}, \bar{y})' \ldots (27).$$

Thus RSS$=(y - X\bar{\theta})'(y - X\bar{\theta}) = 4\bar{y}^2 \ldots (28)$.

Quick Exercise

1. Prove (21) to (28).
2. Consider the model $E(y_1) = \beta_1 + 2\beta_2, E(y_2) = 2\beta_1, E(y_3) = \beta_1 + \beta_2$ with the usual assumptions. Find the RSS subject to the restriction $\beta_1 = \beta_2$.
3. Consider the one-way model with $k \geq 2$

$$y_{ij} = \alpha_i + e_{ij}, \quad i = 1, \ldots, k; j = 1, \ldots, n_i;$$

where e_{ij} are independent with mean 0 and variance σ^2. Find the RSS subject to the restriction $\alpha_1 = \alpha_2$.

6.9 General Linear Model

The general linear model is

$$E(y) = X\beta, D(y) = \sigma^2 V,$$

where Y is $n \times 1, X$ is $n \times p$ and V is a known positive semidefinite $p \times p$ matrix. In previous section $V = I$. We do not assume any condition on the rank of X.

Consider the transformation $z = V^{-\frac{1}{2}}y$, then we have the model $E(z) = V^{-\frac{1}{2}}X, D(z) = \sigma^2 I$, which is the linear model as discussed in previous section.

Quick Exercise

1. What is Estimation space and error space.

6.10 Practical Problems

1. Consider the relation $y = 3 + 2x + \epsilon$ where ϵ follows $N(0, 1)$.
   Step 1: Choose 100 values of x arbitrary, as your choice. You may choose same value for x maximum four times. For each value of x draw a value of ϵ from $N(0, 1)$ using R-code or any other package. Then find the value of y from the given relation.
   Step 2: Now we have 100 pairs of data (x_i, y_i) for $i = 1, 2, \ldots, 100$. Now do the following:
   (a) Draw a scattered diagram y vs. x.
   (b) Fit a linear model $y = a + bx + \epsilon$ with $E(\epsilon) = 0$ and $V(\epsilon) = \sigma^2$.
   (c) Find least square estimation for a, b and σ^2.
   (d) Find goodness of fit, RSS, R^2 and adjusted R^2.
   (e) Plot all residuals in a scattered diagram. Comments on this diagram.
   Step 3: Next fit a linear model $y = a + bx + cx^2 + \epsilon$ with $E(\epsilon) = 0$ and $V(\epsilon) = \sigma^2$ with the same set of 100 data and do the following:
   (a) Find least square estimation for a, b, c and σ^2.
   (b) Find goodness of fit, RSS, R^2 and adjusted R^2.
   (c) Plot all residuals in a scattered diagram. Comments on this diagram.
   Comment on the fact the which fit is better and why?

2. (a) Do the same as previous problem with ϵ follows $U(0, 6)$.
   (b) Do the same as previous problem with ϵ follows $Bin(6, 0.5)$.

3. Solve the following approximate system using R:
   $3a + 4b + c = 3.4 + \epsilon_1$
   $3a + 4b + c = 3.5 + \epsilon_2$
   $4a + 3b + 2c = 10.1 + \epsilon_3$
   $4a + 3b + 2c = 9.8 + \epsilon_4$
   $6a + 5b + 2c = 5.6 + \epsilon_5$.
   State what is the design matrix and rank of the design matrix.

4. Go to the page number 17 of the attached file. This is a data from NSSO site. Fit a appropriate linear model stating all the assumption and concept clearly.

5. Go through the exel file of world bank data. Suggest what type of data analysis can be done for **India** through linear model.

6. Assume that the data below satisfy the simple linear model. Find the m.l.e. of the coefficients and the variance.

Chapter 7

Design of Experiments I

Syllabus:

1. Basic principles of experimental design: randomization, replication, and local control.

2. Uniformity trials. Shapes and sizes of plots and blocks.

3. Standard designs and their analyses: completely randomised design (CRD), randomised block design (RBD), latin square design (LSD), split plot design, and strip arrangements.

4. Comparison of efficiencies. Applications of the techniques of analysis of variance to the analysis of the above designs.

5. General block designs and its information matrix. Concepts of connectedness, orthogonality and balance. Resolvable designs. Properties of BIB designs. Designs derived from BIB designs.

6. Intrablock analysis of orthogonal (CRD, RBD, LSD) and non-orthogonal designs (Balanced incomplete block design and Youden square design).

7.1 Exercise

1. Obtain the layouts of a CRD with three treatments, A, B and C, the replication numbers being 6, 5 and 10 respectively.

2. Obtain the layouts of a RBD with five treatments in four blocks.

3. Construct the model and ANOVA of CRD, RBD and LSD for $n \times n$ design explicitly written all the steps.

4. Discuss the following terms with examples in connection with Design of Experiments: Replication, Randomization, Local Control, Experimental Error, Block, Plot, Block size, Guard area and Treatment.

5. Define Contrast and Orthogonal contrast. What do you mean by Sum of Squares of a contrast.

6. Prove that given a set of n values there are maximum $n-1$ numbers of mutually orthogonal contrast. Write down explicitly two such set of $n-1$ mutually orthogonal contrast.

7. The sum total of the s.s. due to all the $n-1$ mutually orthogonal contrasts among n observations is equal to the s.s. obtained by summing the squares of their deviations from the mean.

8. Whether the following BIBD is constructable or not with the parameter b, v, r, k, λ. If constructable then construct the design and if not, then give reason. All the symbols have usual meaning.
   $v = 5, r = 6,\ k = 3, b = 10,\ \lambda = 3.$
   $v = b = 22,\ r = k = 7,\ \lambda = 2.$

9. Write down the complete set of mutually orthogonal contrast among y_1, y_2, y_3, y_4 using (a) 2 and -2 as coefficients and (b) 1 and -2 as coefficients.

10. What are orthogonal data and non orthogonal data. Analyze these two types of data.

11. What is standard Latin square design. Give two examples of 4×4 non standard LSD.

12. Prove that for any natural number n there is a LSD of order n.

13. An $r \times n$ Latin rectangle is an $r \times n$ array made out of the integers $\{1, 2, \ldots, n\}$ such that no integer is repeated in any row or in any column. Give an example of 3×5 Latin rectangle and extend it to a Latin Square of the order 5. In general can any latin rectangle extended to a Latin square.

14. Let D be a BIBD and D^c, it's complementary design. Show that in the design formed by taking the union of the blocks of D and D^c, every triplet occurs together in a constant number of blocks. Using this or otherwise prove that for a BIBD $b \geq 3(r - \lambda)$. Hence show that if $r = 3\lambda$ then $b \geq 2r$.

15. Show that if in a BIBD $b = 3r - 2\lambda$ then $r > 2\lambda$.

16. Prove that $Cov(Q, P) = 0$ if and only if $N = \frac{rk}{n}$, for a block design, where the symbols are in usual meaning.

17. If N is the incidence matrix of a symmetric BIBD, prove that $(NN')(N'N) = (r - \lambda)(N'N) + k^2\lambda E_{vv}$.

18. Let N be the incidence matrix of a BIBD with $b = v$. Show that $(NN')^{-1} = (r - \lambda)^{-1}I - \frac{\lambda}{r^2(r-\lambda)}JJ'$. Hence show that for a symmetric BIBD, any two blocks have λ treatments in common.

19. Let B_0 be the Block of a BIBD and x_i be the number of treatments common between B_0 and the ith of the remaining blocks. Show that $\sigma^2 = \sum x_i^2 - \frac{(\sum x_i)^2}{b-1} = \frac{(r-k)(r-\lambda)(v-k)}{(b-1)}$. Hence prove the Fisher's inequality.

20. Give a example of BIBD with parameter $v = 9, b = 12, k = 3, r = 4, \lambda = 1$.

21. Prove that, the necessary conditions for the existence of a BIBD is $\lambda(v - 1) \equiv 0$ [mod $(k - 1)$] and $\lambda v(v - 1) \equiv 0$ [mod $k(k - 1)$].

22. Define a BIBD with parameter v, b, k, r, λ. State and prove Fisher's inequality.

23. Prove that for a BIBD, the inequality $b \geq v + r - k$ holds. Is this inequality equivalent to Fisher's inequality?

24. Let N be the incidence matrix of a BIBD. Show that (i) $\det(NN') = 0$ if $b > v$ and (ii) the eigenvalues of NN' are rk and $r - \lambda$ with multiplicities unity and $v - 1$ respectively.

25. Prove that, for a symmetrical BIB design with v even, $r - \lambda$ must be a perfect square. Also Prove that, for symmetrical BIB design $(N^t)^{-1} = \frac{N - \frac{\lambda}{r}J_v}{r-\lambda}$ and hence show that, $NN^t = N^tN$.

26. Construct a of BIBD with parameter $v = s^2, b = s^2 + s, k = s, r = s + 1, \lambda = 1$ with $s = 4$.

27. State which BIBD exist (with brief construction) and which BIBD does not (with proof).
$a)v = 103 = b, r = 18 = k, \lambda = 3$
$b)v = 13, b = 26, k = 3, r = 6, \lambda = 1$
$c)v = 106 = b, r = 21 = k, \lambda = 5$
$d)v = 4, b = 12, k = 2, r = 6, \lambda = 2$
$e)v = 137 = b, k = 17 = r, \lambda = 2$
$f)v = 53 = b, k = 13 = r, \lambda = 3$

$g)v = 46 = b, k = 10 = r, \lambda = 2$
$h)v = 106 = b, r = 20 = k, \lambda = 5$

28. Show that starting from the BIBD $v = 4 = b, k = 3 = r, \lambda = 2$, we can get the solution of the following BIBD:
$a)v = 16 = b, r = 6 = k, \lambda = 2$
$b)v = 16 = b, r = 10 = k, \lambda = 6$
$c)v = 10, b = 15, k = 4, r = 6, \lambda = 2$
$d)v = 6, b = 15, k = 2, r = 5, \lambda = 1$

29. Show that the BIBD with parameters
$(i)\ v = b = 22,\ r = k = 7,\ \lambda = 2$ $(ii)\ v = b = 34,\ r = k = 12,\ \lambda = 4$ do not exist.

30. Show that in a symmetric BIBD, any two blocks have the same number λ of treatments in common.

31. Show that if from a BIBD, we delete one block and all treatments belonging to that block, then the remainder is again a BIBD, find its parameters.

32. Show that, if x is the number of treatments common between any two blocks of a BIBD with parameters v, b, r, k and λ then $\frac{2k\lambda + r(r-k-\lambda)}{r} \geq x \geq -(r-k-\lambda)$.

33. If N is the incidence matrix of a BIBD, prove that $(NN')(N'N) = (r-\lambda)(NN') + rk\lambda E_{vv}$.

34. A square matrix with single entry of unity in each row and in each column with all other elements zero is called permutation matrix. If N is the incidence matrix of a symmetric BIBD, show that N can be express as a sum of k permutation matrix.

35. Show that the following BIBD can be constructed for all integral values of $m(\geq 2)$: $(i)\ \ v = 2^m - 1 = b, r = 2^{m-1} = k, \lambda = 2^{m-2}$
$(ii)\ \ v = 2^{m-1}, b = 2(2^{m-1} - 1), r = 2^{m-1} - 1, k = 2^{m-2}, \lambda = 2^{m-2} - 1$
$(iii)\ \ v = 2^{m-1} - 1, b = 2(2^{m-1} - 1), r = 2(2^{m-2} - 1), k = 2^{m-2} - 1, \lambda = 2(2^{m-3} - 1)$
$(iv)\ \ v = 2^{m-1} - 1, b = 2(2^{m-1} - 1), r = 2^{m-1}, k = 2^{m-2}, \lambda = 2^{m-2}$

36. Show that, for a BIBD if $b = 4(r - \lambda)$ then $k = (v + \sqrt{2})/2$ or $(v - \sqrt{2})/2$.

37. Show that if there exist a BIBD with $r = 2k + 1$ and $\lambda = 1$, then we can construct a BIBD with parameter $v^* = 4k^2 – 1 = b^*, k^* = 2k^2 = r^*, \lambda^* = k^2$.

38. Write down the fixed effect model for the Block Design and find expression for the **C** matrix.

39. When is a block design called connected. Prove that a block design is connected iff Rank $(\mathbf{C}) = v - 1$.

40. Prove that for a resolvable BIBD with incidence matrix N, Rank$(N) \leq b - (r-1)$.

41. If Q denotes the vector of adjusted treatment totals, show that
(i) $Q'J = 0$, (ii) $E(Q) = C\tau$ (iii) $D(Q) = \sigma^2 C$.

42. Show that for a connected design, the diagonal elements of the **C**-matrix are all positive.

43. Prove that a BIBD is always connected unless $k = 1$.

44. Show that a necessary and sufficient condition for a connected block design to be variance balanced is that it's **C**-matrix is of the form **C**$= (a-b)I + bJJ'$.

45. Show that $C + aJJ'$ is non-singular for a connected design.

46. Show that for a connected design, $(C + rr'/n)$ is non-singular. Hence show that the variance of the estimated contrast $l'\tau$ is $\sigma^2 l'(C + rr'/n)^{-1} l$.

47. For a connected, equireplicate design, let $P^{-1} = K - NN'/r + KK'vr$. Show that a solution of the normal equation $C\tau = Q$ is $\tau = (I/r + NPN'/r^2)Q$.

48. An equireplicate, proper block design is called balanced ternary design if its incidence matrix $N = (n_{ij})$ has the following properties:
(i) $n_{ij} = 0, 1$ or 2 and (ii) $\sum_j n_{ij} n_{mj} = \lambda$ for all $i \neq m$, $i, m = 1, 2, \ldots, v$.
Show that for such a design $\sum_{j=1}^{b} n_{ij}^2 = rk - \lambda(v-1)$.

49. Suppose D is a connected block design. Show that the variance of the BLUE of an elementary treatment contrast lies between $2\sigma^2\lambda_{max}^{-1}$ and $2\sigma^2\lambda_{min}^{-1}$, where λ_{max} and λ_{min} denotes the largest and the smallest positive eigenvalue of C respectively.

50. In a connected design, show that the average variance of all elementary treatment contrasts is $2\sigma^2/H$, where$H$ is the H.M. of the non-zero eigenvalues of C.

Chapter 8

Design of Experiments II

Syllabus:

1. Recovery of inter block information in BIB designs. Missing plot techniques.
2. Elementary ideas of Lattice and PBIB designs.
3. Construction of Mutually Orthogonal Latin Squares (MOLS).
4. Construction of BIBD using MOLS and Boses's fundamental method of difference.
5. Factorial experiment: Confounding and balancing in symmetric factorial experiments, Analysis.
6. Response survey designs.

8.1 Recovery of Intra-block Informations

When we discuss the intra-block analysis of block design, the block and the treatment effects are fixed. Now we analyzed the block designs with block effects are random.

We consider only binary and proper designs. Consider the model,

$$y_{ij} = \mu + \tau_i + \beta_j + e_{ij},$$

β_j' s are random variables such that,

$$(i)\ E(\beta_j) = 0,\ Var(\beta_j) = \sigma_b^2,\ Cov(\beta_j, \beta_{j'}) = 0,\ \text{for } j \neq j' = 1, 2, ..., b.$$

(ii) β_j's are uncorrelated with the errors e_{ij}'s.

Now the block totals B_j's are given by

$$B_j = k\mu + \sum n_{ij}\tau_i + k\beta_j + \sum e_{ij}$$

The new error terms denoted by $d_j = k\beta_j + \sum e_{ij}$. Then

$$E(d_j) = 0, \; Var(d_j) = k^2\sigma_b^2 + k\sigma^2$$

The intra-block estimates are then obtained by minimizing the sum of squares due to new errors, i.e.,

$$\sum d_j^2 = \sum (B_j - k\mu - \sum n_{ij}\tau_i)^2 = (B - k\mu J - N^{'}\tau)^{'}(B - k\mu J - N^{'}\tau)$$

The normal equations are

$$-2kB^{'}J + 2k^2\mu b + 2kJ^{'}N^{'}\tau = 0, \quad \text{which implies}$$

$$k\mu b + J^{'}R\tau = G \quad \text{and} \quad kRJ\mu + NN^{'}\tau = NB$$

Combining two equation we can write $A(\mu \quad \tau)^{'} = (G \quad NB)^{'}$

$$\text{Where,} \;\; A = \begin{pmatrix} bk & J'R \\ kRJ & NN' \end{pmatrix}$$

Premultipling both side by the non-singular matrix

$$\begin{pmatrix} I & O \\ -RJ/b & I \end{pmatrix}$$

We obtain

$$A_1(\mu \quad \tau)^{'} = (G \quad NB - GRJ/b)^{'}$$

$$\text{Where,} \;\; A_1 = \begin{pmatrix} bk & J'R \\ O & NN' - RJJ'R/b \end{pmatrix}$$

Assuming $J^{'}R\tau = 0$ and $NN^{'}$ is non-singular, we get $NN^{'}\tau = NB - GRJ/b$, i.e.,

$$\tau = (NN^{'})^{-1}(NB - GRJ/b) = (NN^{'})^{-1}N(B - GJ/b)$$

$$= (NN^{'})^{-1}(NB - (G/bk)NkJ) = (NN^{'})^{-1}NB - GJ/bk$$

and $bk\mu + J^{'}R\tau = G$, i.e., $\mu = G/bk$

We now have two estimators of τ, the intra-block estimator, given by $\tau_1 = C^{-}Q$ and the interblock estimator.

Let $\chi = p'\tau$ be a contrast of treatment effects.
The intra-block estimator of χ, say, χ_1 is

$$\chi_1 = p'\tau_1 = p'C^{-}Q \text{ with variance, } \ Var(\chi_1) = \sigma^2 p'C^{-}p.$$

The inter-block estimator of χ, say, χ_2 is

$$\chi_2 = p'(NN')^{-1}N(B - GJ/b) = p'(NN')^{-1}NB - p'J(G/bk) = p'(NN')^{-1}NB$$

with variance, $Var(\chi_2) = \sigma_d^2 p'(NN')^{-1}p$ where $\sigma_d^2 = k(k\sigma_b^2 + \sigma^2)$.

Two estimators χ_1 and χ_2 are uncorrelated, because $Cov(Q_i, B_j) = 0$ for all i, j.

Now, we want to combine these two estimators, to obtain an estimator with smaller variance, the combined estimator is obtained by taking a weighted average of χ_1 and χ_2, weights being the inverse of the variances of the two estimators. Thus the combined estimator χ^* is given by

$$\chi^* = \frac{\theta_1\chi_1 + \theta_2\chi_2}{\theta_1 + \theta_2} \quad \text{where} \quad \theta_1^{-1} = (p'C^{-}p\sigma^2), \theta_2^{-1} = p'(NN')^{-1}p\sigma_d^2$$

But σ^2, σ_d^2 are usually unknown.

8.2 Alternative methods due to Bose

As before, the model is

$$y = \mu J + D_1'\tau + D_2'\beta + e = (J \quad D_1')(\mu \quad \tau)' + D_2'\beta + e$$

with the assumptions

$$E(e) = 0, \ D(e) = \sigma^2 I, \ E(\beta) = 0, \ D(\beta) = \sigma_b^2 I, \ Cov(e, \beta) = 0.$$

If we order the n observations, such that the first k come from the first block, the next k come from the second block, etc. then we have

$$D_2'D_2 = \text{diag}(J_{KK}, J_{KK}, ..., J_{KK}).$$

Thus, $D(y) = \sum = \sigma^2 I + \sigma_b^2 D_2'D_2 = \text{diag}(L, L, ..., L)$ where, $L = \sigma^2 I_k + \sigma_b^2 J_{KK}$

Now, $\sum^{-1} = \text{diag}(L^{-1}, L^{-1}, ..., L^{-1})$ where, $L^{-1} = \alpha I_k + \beta J_{KK}$, $\alpha = \sigma^{-2}$, $\beta = \frac{-\sigma_b^2}{\sigma^2(\sigma^2 + k\sigma_b^2)}$.

Define, $w_1 = \alpha = \sigma^{-2},\ w_2 = \frac{1}{\sigma^2 + k\sigma_b^2}$ then, $\beta = -(w_1 - w_2)/k$,

so that, $\sum^{-1} = w_1 I_n - \frac{w_1 - w_2}{k}\text{diag}(J_{KK}, J_{KK}, ..., J_{KK}) = w_1 I_n - \frac{w_1-w_2}{k} D_2' D_2$.
Now the normal equation for estimating linear functions of $\mu, \tau_1, \tau_2, ..., \tau_v$, are

$(J' \quad D_1)' \sum^{-1} (J' \quad D_1)(\mu \quad \tau)' = (J' \quad D_1)' \sum^{-1} y$.

Now, $J' \sum^{-1} J = nw_1 - (w_1 - w_2)nk/k = nw_2$.

Similarly, $J' \sum^{-1} D_1' = w_2 r',\ D_1 \sum^{-1} D_1' = w_1 R - (w_1 - w_2)NN'/k,\ J' \sum^{-1} y = w_2 G$

and $D_1 \sum^{-1} y = w_1 Q + w_2(T - Q)$.

Therefore, the normal equation reduces to

$$A_2(\mu \quad \tau)' = (w_2 G \quad w_1 Q + w_2(T - Q))'$$

$$\text{Where,} \quad A_2 = \begin{pmatrix} nw_2 & w_2 r' \\ w_2 r & w_1 R - (w_1 - w_2)NN'/k \end{pmatrix}$$

Eliminating μ we get

$$(w_1(R - NN'/k) + w_2(NN'/k - rr'/n))\tau = w_1 Q + w_2(T - Q - Gr/n).$$

Let, $C^* = NN'/k - rr'/n$ and $Q^* = T - Q - Gr/n$, then, $(w_1 C + w_2 C^*)\tau = w_1 Q + w_2 Q^*$.

This is known as adjusted intra-block normal equation.

Also, $E(Q) = C\tau$ and $E(Q^*) = C^*\tau$.
Now, $C^* J = 0$, since the best estimate of an estimable function of treatment effects is of the form $q'(w_1 Q + w_2 Q^*)$, it follows that if a linear function of the treatment effects is estimable, then it must be a contrast. In practice, w_1 and w_2 will be known.

8.3 Exercise

1. Show that, for $n \geq 2$, there can be at most $n - 1$ mutually orthogonal Latin square of order n. Why and when Latin square design is used. 5

2. Let $L^{(k)} = \left(a_{ij}^{(k)}\right)$, where $a_{ij}^{(k)} = i + jk \ (mod\ 9)$. Which of $L^{(k)}, 1 \leq k \leq 8$, are Latin squares? Are $L^{(2)}$ and $L^{(5)}$ orthogonal?

3. Take two pair of MOLS of order 3 and use Kronecker product to construct a pair of MOLS of order 9.

4. Deduce the normal equations for Recovery of Inter- Block information correctly stating the model.

5. The following are three key-blocks before randomization for a 2^4 experiments with factors A, B, C, D

   Replication I : $(1), abc, abd, cd$
   Replication II : $(1), abc, acd, bd$
   Replication III : $(1), abd, acd, bd$

   Find out the effects confounded. 5

6. Construct a 2^6 design confounding the interaction effects ABC, CDE, ADF.

7. Construct a 2^5 design in 2^3 blocks confounding the interaction effects ABC, CDE and keeping the treatment acd in the key block.

8. Construct a 3^3 design in blocks per replicate partially confounding ABC and ABC^2. 5

9. Construct a 3^3 – factorial design in 9 blocks of 3 plots each completely confounding ABC and ABC^2.

10. Show that, for the 3^3 factorial design contrasts belonging to A and ABC^2 are orthogonal to each other. Define and give examples of generalized interactions.

11. Show that in the s^n factorial the treatment groups corresponding to any interaction contain an equal number of treatments from each of the groups of any other interaction and hence show that the contrasts for the two interaction are orthogonal.

12. What is the maximum number of factors to save interactions up to a given order for a given block size. Give argument to justify your answer.

13. Show that in 2^n experiment the main effects and interaction are mutually orthogonal contrast.

14. Write down the independent treatments in the key block of size 3^2 of 3^{13} factorial such that no main effect and two factor interactions confounded. Obtain the independent interactions confounded.

15. Construct an initial block of $(3^4, 3^2)$ factorial design such that ABC^2 is confounded and ab^2c is an treatment combination applied in the initial block.

16. Given the following key blocks of 2^8 plot factorial in 2^3 blocks, find the interaction confounded.

$$\begin{bmatrix} A & B & C & D & E & F & G & H \\ 0 & 1 & 1 & 0 & 1 & 1 & 0 & 1 \\ 1 & 1 & 0 & 0 & 0 & 0 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 1 & 0 & 1 & 1 & 0 \\ 1 & 0 & 1 & 0 & 1 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 & 1 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

17. Find the confounded interactions separately for five replications of the factorial 2^5 in 2^3 plot such that confounding is balanced for the second and third order interactions.

18. Given that the following three independent interactions .$ABCD, BEFH, ABGH$ are confounded in 2^8 factorial having blocks of 2^5 plots, Find the key block.

19. Show that the total number of solutions of the following k equations is s^{n-k} for every choice of $(\alpha_1, \alpha_2, \ldots, \alpha_k)$:

$$\begin{aligned} p_{11}x_1 + p_{12}x_2 + \cdots + p_{1n}x_n &= \alpha_1 \\ p_{21}x_1 + p_{22}x_2 + \cdots + p_{2n}x_n &= \alpha_2 \\ \cdots \quad \cdots \quad \cdots \quad \cdots \quad & \cdots \\ p_{k1}x_1 + p_{k2}x_2 + \cdots + p_{kn}x_n &= \alpha_k \end{aligned}$$

where the coefficient p_{ij}'s and α_i, $(i = 1, 2, \ldots, k)$ and the solutions are in the finite field $GF(s)$, where $s = p^m$ and p is a prime.

20. Let M be the module consisting of residue classes $\mod 5$ and to every element a of M, let their correspond three treatments a_1, a_2, a_3. Consider the set of following seven blocks:
$\{(0_1, 1_1, 0_2), (0_2, 1_2, 2_3), (0_3, 1_3, 2_1), (0_1, 2_1, 3_2), (0_2, 0_3, 3_2), (0_3, 2_3, 0_1), (0_1, 2_2, 1_3)\}$. Verify the conditions of Bose's first fundamental theorem on method of differences. Hence construct a suitable BIBD using these blocks as initial blocks.

21. Construct a BIBD with parameter $v = 15 = b, k = 7 = r, \lambda = 3$, using Hadamard matrix.

22. Describe the method of construct a BIB design with parameter $v = 4t - 1 = b, k = 2t - 1 = r, \lambda = t - 1$, using Hadamard matrix.

23. Test whether the following design is variance-balanced or not:
$(1,2); (1,3); (1,4); (2,3); (2,4); (3,4); (1,2,5,5); (1,3,5,5); (1,4,5,5); (2,3,5,5); (2,4,5,5); (3,4,5,5)$.

24. Using finite geometries construct the following BIBD:
$v = 15, b = 35, k = 3, r = 7, \lambda = 1$
$v = 40, b = 40, k = 13, r = 13, \lambda = 4$
$v = 27, b = 39, k = 9, r = 13, \lambda = 4$.

25. Let $v = 4t+3$ be prime or prime power and x be a primitive element of GF(v). Suppose $x^0 + x = x^{2a}$ an even power of x, then show that the initial blocks $(0, x^j, x^{2a+i}), i = 0, 2, \ldots, 4t$ provide a solution for BIBD with parameters $v = 4t+3, b = (2t+1)(4t+3), r = 3(2t+1), k = 3 = \lambda$. Can we construct the same series of BIBD design through the method of difference if $x^0 + x = x^{2a+1}$, an odd power of x.

26. Let the incidence matrix of a design with $v = 5, b = 8$ as below:

$$\begin{bmatrix} 11102000 \\ 11010200 \\ 10110020 \\ 01110002 \\ 00001111 \end{bmatrix}.$$

Show that the design is variance-balanced.

27. A $v \times b$ matrix D with entries $-1, 0, 1$ is said to be a balanced orthogonal design (BOD) if (i) the inner product of any two rows of D is zero, and (ii) when the -1's in D are replaced by $+1$'s the resultant matrix becomes the incidence matrix of a BIBD with parameters v, b, r, k, λ. Show that a necessary condition for the existence of a BOD is that λ is even.

28. Describe the method to construct a BIB design with parameters $v = 4t - 1 = b, k = 2t - 1 = r, \lambda = t - 1$ using Hadamard matrix.

29. State Bose's first and second fundamental theorem on method of differences.

30. Define A- optimality in connection with intra block model and the inference problem $P : \eta = L\tau$ with $LJ = 0$. Prove that, for a given $b, v, k(< v)$, a BIBD is A- optimal for estimating all elementary contrasts in the class of all connected incomplete block design.

31. Prove that for a resolvable BIBD with incidence matrix N, Rank$(N) \leq b - (r-1)$. Hence obtain Bose inequality for resolvable BIBD.

32. Let $v = 4t + 1$ be a prime or a prime power and let x be a primitive element of $GF(v)$. show that the t initial blocks $(x^i, x^{i+1}, x^{i+2}, \ldots, x^{i+(k-1)t})$, $i = 0, 1, 2, \ldots, t-1$ provide a solution to the BIBD with parameters $v = tk+1, b = t(tk+1), r = tk, k, \lambda = k-1$.

33. Let $v = 10t + 1$ be a prime or a prime power and let x be a primitive element of $GF(v)$. show that the t initial blocks $(x^i, x^{2t+i}, x^{4t+i}, x^{6t+i}, x^{8t+i})$, $i = 0, 1, 2, \ldots, t-1$ provide a solution to the BIBD with parameters $v = 10t+1, b = t(10t+1), r = 5t, k = 5, \lambda = 2$.

Chapter 9

Stochastic Process

Syllabus:

1. Markov chain with finite state space and countable state space, Classification of states, Chapman-Kolmogorov equation, Calculation of n-step transition probability matrix and its limit, Stationary distribution of Markov chain. Random walk and Gambler's ruin problem and reversibility. [14] Discrete state space continuous time Markov chains, Poisson process. [6] Renewal theory: Elementary Renewal theorem, Stopping time, Statement and uses of Key Renewal theorem.

9.1 Exercise

1. Define autoregressive process of order 1 on stationary time series and derive their autocovariances for different lag values.

2. Show that $g(0) > |g(h)|$ and $g(h) = g(-h)$ for all h and $g(h)$ is the auto covariance function of lag h. 5

3. For a Markov chain $X_n, n \geq 0$, with a finite state space, prove that $P(X_0 = x_0 | X_n = x_n, \ldots, X_1 = x_1) = P(X_0 = x_0 | X_1 = x_1)$. 5

4. Suppose we have two boxes and $c+d$ balls, of which c are black and d are red. Initially, c balls are placed in box 1, and remainder of the balls are placed in box 2. At each trial a ball is chosen at random from each of the boxes, and the two balls are put back in the opposite boxes. Let X_0 denotes the number of black balls initially in box 1 and for $n \geq 1$, let X_n denotes the number of

black balls in box 1 after the n-th trial. Find the 2 step transition function of the Markov chain $X_n, n \geq 0$. 5

5. Consider a Markov chain having state space $\{0, 1, \ldots, 6\}$ and transition matrix

$$\begin{pmatrix} 0.2 & 0 & 0.3 & 0.4 & 0.1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0.4 & 0 & 0.6 \\ 0 & 0 & 0 & 0 & 0.3 & 0.7 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0.5 & 0.5 \end{pmatrix}$$

Determine which states are transient and which states are recurrent. 5

6. Define waiting time distribution. Show that waiting time S_n has gamma distribution. 5

7. Let x and y are distinct states of a Markov chain having $d(< \infty)$ states and suppose that x leads to y. Let n_0 be the smallest positive integer such that $P^{n_0}(x, y) > 0$ and let $x_1, \ldots, x_{n_0-1}$ be states such that

$$P(x, x_1)P(x_1, x_2) \cdots P(x_{n_0-2}, x_{n_0-1})P(x_{n_0-1}, y) > 0.$$

Show the followings:
(a) $x, x_1, \ldots, x_{n_0-1}, y$ are distinct states.
(b) $n_0 \leq d - 1$
(c) $P_x(T_y \leq d - 1) > 0$. 4+2+4

8. Suppose that η_n particles are added to a box at time $n = 1, 2, \ldots,$ where η_n's are independent and have a Poisson distribution with common parameter λ. Suppose that each particle in the box at time n, independently of all other particles in the box and independently of how particles are added to the box, has probability $p < 1$ of remaining in the box at time $n + 1$ and probability $q = 1 - p$ of being removed from the box at a time $n + 1$. Let $X_n, n \geq 0$ denotes the number of particles in the box at time n.
(a) Show that $X_n, n \geq 0$ is a Markov chain.
(b) Let X_0 have a Poisson distribution with parameter t. Show that X_n has a Poisson distribution with parameter $tp^n + \frac{\lambda}{q}(1 - p^n)$.
(c) Further show that $E(X_n|X_0 = x) = xp^n + \frac{\lambda}{q}(1 - p^n)$. 3+3+4

9. (a) Find the two-step transition matrix $P^{(2)}$ for the following problem: a person starts a game with one rupee. If he has loses, he has to pay one rupee and if

he wins he will get one rupee. If he have either two rupees or no rupee, then he quits the game. 3

(b) Let π be a stationary distribution of a Markov chain. Suppose that y and z are two states such that for some constant c, $P(x,y) = cP(x,z)$ for all $x \in \mathcal{S}$. Show that $\pi(y) = c\pi(z)$. 4

(c) Prove that

$$P_x(T_y \leq n+1) = P(x,y) + \sum_{x \neq y} P(x,z) P_z(T_y \leq n), \quad n \geq 0$$

$$\text{and} \quad \rho_{xy} = P(x,y) + \sum_{z \neq y} P(x,z)\rho_{zy}. \qquad 3$$

10. (a) Define stationary time series, describe clearly strict and weak stationarity in this context.

(b) Consider the Ehrenfest chain with $d = 3$. Find the stationary distribution and $\lim_{n\to\infty} P^n$, where P is the transition matrix. 5+5