MAR GREGORIOS COLLEGE
OF ARTS & SCIENCE

Block No.8, College Road, Mogappair West, Chennai – 37

Affiliated to the University of Madras
Approved by the Government of Tamil Nadu
An ISO 9001:2015 Certified Institution

DEPARTMENT OF MATHEMATICS

SUBJECT NAME: MATHEMATICAL STATISTICS I

SUBJECT CODE: BMA-CSA04

SEMESTER: III

PREPARED BY: PROF.S.C.PREMILA

UNIVERSITY OF MADRAS

B.Sc. DEGREE COURSE IN MATHEMATICS

SYLLABUS WITH EFFECT FROM 2020-2021

BMA-CSA04 ALLIED: MATHEMATICAL STATISTICS–I

(Common to B.Sc. Maths with Computer Applications)

Learning outcomes:

Students will acquire knowledge of

- The laws of Probability and Baye's theorem.
- Measures of Location, Dispersion, Correlation and Regression The Discrete and Continuous Probability Distributions.

UNIT I

Concept of sample space- Events- Definition of Probability (Classical,Statstical& Axiomatic)- Addition and Multiplication laws of Probability- Independence- Conditional Probability- Baye's theorem – Simple Problems.

UNIT II

Random Variables (Discrete and Continuous) Distribution function- Expected values and Moments- Moment generating function – Probability generating function- Examples.

UNIT III

Characteristic function- Uniqueness and Inversion theorems (Statements and applications only)Cumulants - Chebychev's Inequality – Simple Problems.

UNIT IV

Concepts of bivariate distributions- Correlation and Regression- Linear Prediction- Rank Correlation coefficient- Concepts of partial and multiple correlation coefficients- Simple problems.

UNIT V

Standard Distributions – Binomial- Poisson- Normal- Uniform distributions- Geometric- Exponential-Gamma -Beta distributions- Inter relationship between distributions.

Reference:

- S.C.Gupta&V.K.Kapoor : Elements of Mathematical Statistics, Sultan Chand & Sons, NewDelhi.
- Hogg R.V. & Craig A.T. (1988) : Introduction to Mathematical Statistics, McMillan. Mood A.M. &Graybill F.A. &Boes D.G. (1974): Introduction to theory of Statistics, McGraw Hill.
- Snedecor G.W. & Cochran W.G(1967) : Statistical Methods, Oxford and IBH.

UNIT I

PROBABILITY

Introduction:

In this chapter we develop the mathematical theory of probability and introduce the concept of random variables which form the basis for various types of theoretical distributions.

Definition:

An experiment is defined as an action which we conceive and do or intend to do.

Each experiment ends with an outcome. For example, a research student in "statistics" when undertaking a pre election sample survey.

An experiment is called a random experiment if, when repeated under the same conditions, it is such that the outcome cannot be predicted with certainty but all possible outcomes can be determined prior to the performance of the experiment.

Each performance of the random experiment is called trail. The collection of all possible outcomes of a random experiment is called the sample space **S**. The elements of sample space are called sample points.

Example:

When two cons are tossed at a time the outcome is an ordered pair (H,H) or (H, T) or (T,H) or (T,T). Hence for the random experiment of tossing two coins, sample space S={(H,H), (H,T), (T,H), (T,T)}.

Definition:

Any subset A of a sample space S is called an event.

The event S is called a sure event and the event φ is called an impossible event.

Definition:

Let S be a sample space associated with an experiment. Let A be event suppose an experiment is repeated N times and suppose the event A happens f times. Then f/N is called the relative frequency of the event A. clearly $0 \leq f/N \leq 1$.

The following desirable properties to be satisfied by a probability function P.

a) P(A) ≥ 0 for all events
b) P(A) $\leq 1$for all events
c) P(S) = 1
d) If A and B are disjoint events

P(AUB) = P(A) +P(B)

Example:

When two cons are tossed at a time the outcome is an ordered pair (H,H) or

(H, T) or (T,H) or (T,T). Hence for the random experiment of tossing two coins, sample space S={(H,H), (H,T), (T,H), (T,T)}.

Definition:

Any subset A of a sample space S is called an event.

The event S is called a sure event and the event φ is called an impossible event.

Definition:

Let S be a sample space associated with an experiment. Let A be event suppose an experiment is repeated N times and suppose the event A happens f times. Then f/N is called the relative frequency of the event A. clearly $0 \leq f/N \leq 1$.

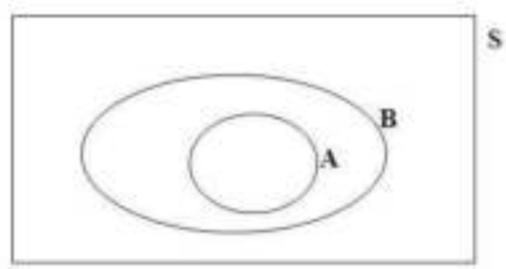
The following desirable properties to be satisfied by a probability function P.

a) P(A)$\leq$ 1for all events
b) P(S) = 1
c) If A and B are disjoint events then P(B)>=P(A)

Proof:

Let $A \subseteq B$.

Then B=A∪($\bar{A}$ ∩B).A and$\bar{A} \cap B$ disjoint events of S.



Hence $P(B) = P(A) + P(\bar{A} \cap B)$

But$P(\bar{A} \cap B) \geq P(A)$

Hence $P(B) \geq (PA)$

Corollary:

For each A⊆S, 0≤ P(A)≤1.

Proof:

$\emptyset \subseteq A \subseteq S$

Hence P($\emptyset$)≤P(A) ≤P(S)

Hence $0 \leq P(A) \leq 1$

.

ADDITION THEOREM:

If A and B are any two events of a sample space S then

$P(A \cup B) = P(A) + P(B) - P(A \cap B)$

Proof:

$A \cap B = A \cup (\bar{A} \cap B)$ and A and $\bar{A} \cap B$ are disjoint sets.

$\therefore$ $P(A \cup B) = P(A) + P(\bar{A} \cap B)$ ……………..(1)

Now, B = $(A \cap B) \cup (\bar{A} \cap B)$ and $A \cap B$ and $\bar{A} \cap B$ are disjoint sets.

$\therefore P(B) = P(A \cap B) + P(\bar{A} \cap B)$

$\therefore P(\bar{A} \cap B) = P(B) - P(A \cap B)$………………(2)

(1)$\Rightarrow P(A \cup B) = P(A) + P(B) - P(A \cap B)$ [(2) in (1)]

Example:

Let S = {(H,H), (H,T) (T,H),(T,T)}.

Let us assign the probability of ¼ to each element of the sample space S.

Let A={(H,H),(H,T)}; B={(H,H), (T,H)}

$A \cup B$ = { (H,H) (H,T), (T,H)}

$A \cap B$ = {(H,H)}

$P(A) = \frac{1}{2}$, $P(B) = \frac{1}{2}$, $P(A \cup B) = \frac{3}{4}$

And $P(A \cap B) = \frac{1}{4}$

We have, $P(A \cup B) = P(A) + P(B) - P(A \cap B)$

$= \frac{1}{2} + \frac{1}{2} - \frac{1}{4}$

$= \frac{2+2-1}{4} = \frac{3}{4}$

Hence it is verified.

Example:

Let S= {(i,j) /i,j $\in$N, 1$\leq$i$\leq 6, 1 \leq$j$\leq$6} be the sample space of the random experiment of throwing two dice. We assign the uniform probably of 1/36 to each of the 36 sample points in the sample space S.

Let A denote the event of getting 1 in the second die.

Then A = {(1,1) (2,1), (3,1), (4,1), (5,1) (6,1) }

P(A) = $\frac{6}{36} = \frac{1}{6}$

Let B = {(1,2), (2,2) (3,2)}

Then P(B) = $\frac{3}{36} = \frac{1}{12}$

A$\cap$B = $\emptyset$, P(A$\cap B) = 0$

P(A$\cup B) = P(A) + P(B) - P(A \cap B)$

$$= \frac{6}{36} + \frac{3}{36} = \frac{9}{36} = \frac{1}{4}$$

Definition:

Let S denote a sample space associated with an experiment. Let $A_1, A_2, A_3 \dots, A_n, \dots$ be a sequence of Subsets of S.

If $A_i \cap A_j = \emptyset$ for all i, j with i $\neq$ j then the sequence of subsets is said to be mutually disjoint.

If $\cup_{n=1}^{\infty} A_n$ then the sequence of events is said to be exhaustive.

Example:

Let S= {(i,j) / i, j $\in$ N, 1$\leq i \leq 6, 1 \leq j \leq 6$}

Let A be an event of getting the sum i + j and odd number as B be an event of getting the sum as an even number.

Clearly A$\cap$B = $\emptyset$ and A$\cup$B = S Hence A and B are mutually exclusive and exchaustive events.

Theorem:

For any two events A and B, P $(\bar{A} \cap B) = P(B) - P(A \cap B)$

Proof:

Let $\bar{A} \cap B$ and $A \cap B$ are disjoint events and

$$(\bar{A} \cap B) \cup (\bar{A} \cap B) = B$$

P(B) = P [(A$\cap B) \cup (\bar{A} \cap B)$]

P(B) = P(A$\cap B) + P(\bar{A} \cap B)$

$\therefore$ P ($\bar{A} \cap$B) = P (B) – P (A$\cap B)$

Remark:

Similarly we shall get P (A$\cap \bar{B}$) = P(A) – P (A$\cap B)$

Theorem:

If B$\subset$A, then (i) P(A$\cap \bar{B}) = P(A) - P(B)$

(ii) P(B) $\leq$P(A

Proof:

i).When B$\subset$A, B and (A$\cap \bar{B})$ are mutually exclusive events and their union is A.

$\therefore$ P(A) = P[B$\cup (A \cap \bar{B}$)]

P(A) = P(B) + P(A$\cap \bar{B}$)

$\therefore$ P(A$\cap \bar{B}$) = P(A) – P(B).

ii) Using axiom (i)

P(A$\cap \bar{B}$)$\geq 0 \Rightarrow P(A) - P(B) \geq 0$.

$P(B) \leq P(A)$.

Corollary:

If (A$\cap$B) $\subset$ A and (A$\cap$B) $\subset$(B) then $P(A \cap B) \leq P(A)$ and $P(A \cap B) \leq P(B)$

Law of addition of probabilities:

Statement:

If A and B are any two events (subsets of a sample space S) and are not disjoint then, P(A$\cup$B) = P(A) + P(B) – P(A$\cap$B)

Proof:

We have, A$\cup B = A \cup (\bar{A} \cap B)$

Since A and ($\bar{A} \cap B)$ *are disjoint*.

$P(A \cup B) = P(A) + P(\bar{A} \cap B)$

$= P(A) + [P(\bar{A} \cap B) + P(A \cap B) - P(A \cap B)]$

$= P(A) + P[(\bar{A} \cap B) \cup (A \cap B)] - P(A \cap B)$

$[\because (\bar{A} \cap B)$ *and* $(A \cap B)$ *are disjoint*$]$

$\therefore P(A \cup B) = P(A) + P(B) - P(A \cap B)$

Theorem: (Extension of General law of addition of probabilities]

For n events $A_1, A_2 \dots \dots \dots, A_n$, wehave

$P(\bigcup_{i=1}^{m} A_i) = \sum_{i=1}^{n} p(A_i) - \sum\sum P(A_i \cap A_j) + \cdots.$

$+ (-1)^{n-1} P(A_1 \cap A_2 \cap \dots . \cap A_n)$ for all $1 \le i \le j \le n$.

Proof:

For two events A_1 *and* A_2

We have,
$P(A_1 \cup A_2) = P(A_1) + (P(A_2) - P(A_1 \cap A_2)$

It is true for n=2

Suppose that, it is true for n=r

Then, $P(\bigcup_{i=1}^{r} A_i) = \sum_{i=1}^{r} P(A_i) - \sum\sum P(A_i \cap A_j) + \cdots.$

$+ (-1)^{r-1} P(A_1 \cap A_2 \cap \dots . \cap A_r)$

for all $1 \le i \le j \le r$.

Now ,

$P(\bigcup_{i=1}^{r+1} A_i) = P[(\bigcup_{i=1}^{r} A_i) \cup A_{r+1}]$

$$=P(\cup_{i=1}^{r} A_i) + P(A_{r+1}) - P[(\cup_{i=1}^{r} A_i) \cap A_{r+1}]$$

$$= P(\cup_{i=1}^{r} A_i) + P(A_{r+1}) - P[\cup_{i=1}^{r}(A_i \cap A_{r+1})]$$

(Distributive Law)

$$=\sum_{i=1}^{r} P(A_i) - \sum\sum P(A_i \cap A_j) + \ldots + (-1)^{r-1} P(A_1 \cap A_2 \cap \ldots \cap A_r) + P(A_{r+1}) - P[\cup_{i=1}^{r} A_i \cap A_{r+1}]$$

For all $1 \leq i \leq j \leq r$

$$=\sum_{i=1}^{r+1} P(A_i) - \sum\sum P(A_i \cap A_j) + \cdots + (-1)^{r-1} P(A_1 \cap A_2 \cap \ldots \cap A_r) - \left[\sum_{i=1}^{r} P(A_i \cap A_{r+1}) - \sum\sum P(A_i \cap A_j \cap A_{r+1}) + (-1)^{r-1} P(A_1 \cap A_2 \cap \ldots \cap A_r \cap A_{r+1})\right]$$

For all $1 \leq i \leq j \leq r$

$$=\sum_{i=1}^{r+1} P(A_i) - \left[\sum\sum P(A_i \cap A_j) + \sum_{i=1}^{r} P(A_i \cap A_{r+1})\right] + (-1)^{r} P(A_1 \cap A_2 \cap \ldots \cap A_r \cap A_{r+1})$$

For all $1 \leq i \leq j \leq r$

$$=\sum_{i=1}^{r+1} P(A_i) - \sum\sum P(A_i \cap A_j) + \cdots + (-1)^{r} P(A_1 \cap A_2 \cap \ldots \cap A_r \cap A_{r+1})$$

For all $1 \leq i \leq j \leq r+1$

Hence it is true for n=r and it is also true for n=r+1.

Hence by the principle of mathematical induction, it is true for all positive integral values of n.

Theorem: [Boole's inequality]

For n events $A_1, A_2, \ldots A_n$ we have

a) $P(\cap_{i=1}^{n} A_i) \geq \sum_{i=1}^{n} P(A_i) - (n-1)$

b) $P(\cup_{i=1}^{n} A_i) \leq \sum_{i=1}^{n} P(A_i)$

Proof:

The result is now prove by mathematical introduction

$P(A_1 \cup A_2) = P(A_1) + P(A_2) - P(A_1 \cap A_2) \leq 1$

$P(A_1 \cup A_2) \geq P(A_1) + P(A_2) - 1$

Hence it is true for n=2

Suppose that it is true for n=r such that

$P(\cup_{i=1}^{r} A_i) \geq \sum_{i=1}^{r} P(A_i) - (r-1)$

Then $P(\cap_{i=1}^{r+1} A_i) = P(\cap_{i=1}^{r} A_i \cap A_{r+1})$

$\geq$ P $(\cap_{i=1}^{r} A_i)$+P(A_{r+1})-1

$$\geq \sum_{i=1}^{r} P(A_i) - (r-1) + P(A_{r+1}) - 1$$

$\Rightarrow$P$(\cap_{i=1}^{r+1} Ai) \geq \sum_{i=1}^{r+1} P(A_i) - r$

$\therefore$ It is true for n=r+1 also

b) Let $\bar{A}_1\bar{A}_2$……..$\bar{A}_n$ be the events

We get

P($\bar{A}_1 \cap \bar{A}_2 \ldots \cap \bar{A}$n) $\geq$[P($\bar{A}_1$) + P($\bar{A}_2$) + ……..+P ($\bar{A}_n$)] – (n-1)

= [1-P(A_1) + [1-P(A_2)] + …….+ [1-P(A_n)] – (n-1)

= 1-P(A_1) – P(A_2) - …….-P (A_n)

$\Rightarrow$ P(A_1) + P(A_2) +…..P(A_n)$\geq$1-P ($\bar{A}_1 \cap \bar{A}_2 \cap \ldots \cap \bar{A}_n$)

= 1-P$(\overline{A_1 \cup A_2 \cup \ldots \cup A_n)}$

= P($A_1 \cup A_2 \cup \ldots \cup A_n$)

$\Rightarrow$ P($A_1 \cup A_2 \cup \ldots \cup A_n$)$\leq$P($A_1$) + P($A_2$) +…..P($A_n$)

Theorem:

For n events $A_1, A_2, \ldots A_n$.

we have P[$\cup_{i=1}^{n} A_i] \geq \sum_{i=1}^{n} P(A_i) - \sum\sum P\left(A_i \cap A_j\right), for\ all\ 1 \leq i \leq j \leq n$

Proof:

We shall prove this theorem by the method of induction.

We have,

P(A_1+A_2+A_3) = P(A_1) +P(A_2) + P(A_3)- [P($A_1 \cap A_2$) + P($A_2 \cap A_3$) + P($A_2 \cap A_1$)] + P($A_1 \cap A_2 \cap A_3$)

P$(\cup_{i=1}^{3}) \geq \sum_{i=1}^{3} P(A_i) - \sum\sum_{1 \leq i \leq j \leq 3} P\left(A_i \cap A_j\right)$

Thus the result is true for n=3

Suppose that the result is true for n=r

Then P$(\cup_{i=1}^{r} A_i) \geq \sum_{i=r} P(A_i)$ - $\sum\sum_{1 \leq i \leq j \leq r} P\left(A_i \cap A_j\right)$

Now,

P$(\cup_{i=1}^{r+1} A_i)$ =P$[\cup_{i=1}^{r} A_i \cup A_{r+1}]$

$= \text{P}(\cup_{i=1}^{r}(A_i) + P(A_{r+1}) - P\,[(\cup_{i=1}^{r} A_i) \cap A_{r+1})]$

$= \text{P}(\cup_{i=1}^{r}(A_i) + P(A_{r+1}) - P\,[(\cup_{i=1}^{r}(A_i \cap A_{r+1})]$

$$\geq \left[\sum_{i=1}^{r} P(A_i) - \sum\sum_{1\leq i\leq j<r} P(A_i \cap A_j)\right] + P(A_{r+1}) - P[\cup_{i=1}^{r}(A_i \cap A_{r+1})]$$

From Boole"s inequality we get

P$[\cup_{i=1}^{r}(A_i \cap A_{r+1})] \leq \sum_{i=1}^{r} P(A_i \cap A_{r+1})$

$\Rightarrow$ -P $[\cup_{i=1}^{r}(A_i \cap A_{r+1})] \geq -\sum_{i=1}^{r} P(A_i \cap A_{r+1})$

$(\cup_{i=1}^{r+1} A_i \geq \sum_{i=1}^{r+1} P(A_i) - \sum\sum_{1\leq i\leq j\leq r} P(A$ $\therefore P$)

$_i \cap A_j) - \sum_{i=1}^{r} P(A_i \cap A_{r+1})$

$\geq \sum_{i=1}^{r+1} P(A_i) - \sum\sum_{1\leq i\leq j\leq r+1} P(A_i \cap A_j)$ $P(\cup_{i=1}^{r+1} A_i$

)

Hence, if the theorem is true for n=r, it is also true for n=r+1 Hence by

mathematical induction,

the result is true for all positive integral values of n.

Multiplication law of probability and conditional probability:

Theorem:

For two events A and B

P(A∩B) = P(A). P(B | A), P(A)>0

= P(B).P(A | B), P(B)>0

Where P(B | A) represents the conditional probability of occurrence of B when the event A has already happened and P(A | B) is the conditional probability of happening of A, given that B has already happened.

Proof:

Suppose the sample space contains N occurrences of which n_A occurrences belong to the event A and n_B occurrences belong to the event B.

Let n_{AB} be the number of occurrences favorable to the compound event A∩B then, the unconditional probabilities are given by

P(A) = $\frac{n_A}{N}$, P(B) = $\frac{n_B}{N}$ and P($A \cap B) = \frac{n_{AB}}{N}$

Now, the conditional probability P(A | B) refers to the sample space of n_B occurances, out of which n_{AB} occurrences pertain to the occurrence of A, when B has already happened.

P(A | B) = $\frac{n_{AB}}{n_B}$

Similarly P(B | A) = $\frac{n_{AB}}{n_A}$

Now, P(A$\cap B) = \frac{n_{AB}}{N}$

$= \frac{n_{AB}}{n_A} \cdot \frac{n_A}{N}$

$= P(B \mid A)\, P(A)$

And $P(A \cap B) = \frac{n_{AB}}{N} = \frac{n_{AB}}{n_B} \cdot \frac{n_B}{N}$

$= P(A \mid B)\, P(B)$

$\therefore P(B \mid A) = \frac{P(A \cap B)}{P(A)}$ and $P(A \mid B) = \frac{P(A \cap B)}{P(B)}$

Thus the conditional probabilities $P(B \mid A)$ and $P(A \mid B)$ are defined iff $P(A) \neq 0$ and $P(B) \neq 0$ respectively.

Extension of multiplication law of probability :

Theorem:

For n events $A_1, A_2, \ldots A_n$, we have $P(A_1 \cap A_2 \cap \ldots \cap A_n) = P(A_1)\, P(A_2 \mid A_1)$ $P(A_3 \mid A_1 \cap A_2) \ldots P(A_n \mid A_1 \cap A_2 \cap \ldots\ldots \cap A_{n-1})$ …..

Where,

$P(A_i \mid A_j \cap A_k \cap \ldots\ldots \cap A_i)$ represents the conditional probability of the event A_i, given that the events A_j, $A_k, \ldots A_i$ have already happened.

Proof:

We have for three events A_1, A_2, and A_3

$P(A_1 \cap A_2 \cap A_3) \quad = P(A_1 \cap \overbrace{A_2 \cap A_3})$

$= P(A_1)\, P(A_2 \cap A_3 \mid A_1)$

$= P(A_1)\, P(A_2 \mid A_1) P(A_3 \mid A_1 \cap A_2)$

It is true for n=2 and n=3

Suppose that it is true for n=k.

So that,

$P(A_1 \cap A_2 \cap \ldots \cap A_k) = P(A_1)\, P(A_2 \mid A_1) P(A_3 \mid A_1 \cap A_2) \ldots$
$P(A_k \mid A_1 \cap A_2 \cap \ldots \cap A_{k-1})$

Now,

$P(\overbrace{A_1 \cap A_2 \cap \ldots \cap A_k} \cap A_{k+1}) = P(A_1 \cap A_2 \cap \ldots \cap A_k)\, P(A_{k+1} \mid A_1 \cap A_2 \cap \ldots \cap A_k)$

$= P(A_1)\, P(A_2 \mid A_1) \ldots P(A_k \mid A_1 \cap A_2 \cap \ldots \cap A_{k-1} \times P(A_{k+1} \mid A_1 \cap A_2 \cap \ldots \cap A_k)$

It is true for n=k+1 also

Since it is true for n=2 and n=3, by the principle of mathematical induction, it is true for all integral values of n.

Theorem:

For any three events A,B,C; $P(A \cup B | C) = P(A|C) + P(B|C) - P(A \cap B|C)$

Proof:

We have

$P(A \cup B) = P(A) + P(B) - P(A \cap B)$

$\Rightarrow P[(A \cap C) \cup (B \cap C)] = P(A \cap C) + P(B \cap C) - P(A \cap B \cap C)$

Dividing both sides by P(C) we get

$$\frac{P(A\cap C)\cup(B\cap C)}{P(C)} = \frac{P(A\cap C) + P(B\cap C) - P(A\cap B\cap C)}{P(C)}, P(C) > 0$$

$$= \frac{P(A\cap C)}{P(C)} + \frac{P(B\cap C)}{P(C)} - \frac{P(A\cap B\cap C)}{P(C)}$$

$$\Rightarrow \frac{P[(A\cup B)\cap C]}{P(C)} = P(A|C) + P(B|C) - P(A \cap B)|C)$$

$$\Rightarrow P(A \cup B|C) = P(A|C) + P(B|C) - P(A \cap B|C)$$

Theorem:

For any three events A, B and C $P(A \cap \bar{B}|C) + P(A \cap B|C) = P(A|C)$

Proof:

$$P(A \cap \bar{B}|C) + P(A \cap B|C) = \frac{P(A\cap \bar{B}\cap C)}{P(C)} + \frac{P(A\cap B\cap C)}{P(C)}$$

$$= \frac{P(A\cap \bar{B}\cap C) + P(A\cap B\cap C)}{P(C)}$$

$$= \frac{P(A\cap C)}{P(C)} = P(A|C)$$

Theorem:

For a fixed B with $P(B)>0$, $P(A|B)$ is a probability function

Proof:

i. $P(A|B) = \frac{P(A\cap B)}{P(B)} \geq 0$

ii. $A(S|B) = \frac{P(s\cap B)}{P(B)} = \frac{P(B)}{P(B)} = 1$

iii. If {An} is any finite or infinite sequence of disjoint events, then

$$P[\cup_n A_n | B] = \frac{P[(\cup_n A_n)\cap B]}{P(B)}$$

$= \frac{P[\cup_n A_n B]}{P(B)}$

$= \frac{\sum_n P(A_n B)}{P(B)} = \sum_n \frac{P(A_n B)}{P(B)}$

$= \sum_n P(A_n \mid B)$

Hence the theorem

Theorem:

For any three events A,B and C defined on the sample space S such that $B \subset C$ and

$P(A>0), P(B|A) \leq P(C|A)$

Proof:

$P(C|A) = \frac{P(C \cap A)}{P(A)}$

$= \frac{P[(B \cap C \cap A) \cup (\bar{B} \cap C \cap A)]}{P(A)}$

$= \frac{P(B \cap C \cap A)}{P(A)} + \frac{P(\bar{B} \cap C \cap A)}{P(A)}$

$= P(B \cap C \mid A) + P(\bar{B} \cap C \mid A)$

Now, $B \subset C \Rightarrow B \cap C = B$

$\therefore P(C|A) = P(B|A) + P(\bar{B} \cap C|A)$

$\Rightarrow P(C|A) \geq P(B|A)$

Theorem:

If A and B are independent events then A and $\bar{B}$ are also independent events

Proof:

WE have $P(A \cap \bar{B}) = P(A) - (A \cap B)$

$= P(A) - P(A)\, P(B)$ [A and B are independent]

$= P(A)\, [1-P(B)]$

$= P(A)\, P(\bar{B})$

$\Rightarrow$ A and $\bar{B}$ are independent events

Theorem:

If A and B are independent events then $\bar{A}$ and $\bar{B}$ are also independent events

Proof:

We are given $P(A \cap B) = P(A)P(B)$

Now $P(\bar{A} \cap \bar{B}) = P(\overline{A \cup B})$

$= 1 - P(A \cup B)$

$= 1 - [P(A) + P(B) - P(A \cap B)]$

$= 1 - [P(A) + P(B) - P(A)\,P(B)]$

$= 1 - P(A) - P(B) + P(A)\,P(B)$

$= [1 - P(B)] - P(A)\,[1 - P(B)]$

$= [1 - P(B)]\,[1 - P(A)]$

$= [1 - P(A)]\,[1 - P(B)]$

$= P(\bar{A})\,P(\bar{B})$

$\therefore$ $\bar{A}$ and $\bar{B}$ are independent events

Theorem:

If A,B,C are mutually independent events when $A \cup B$ *and C are also independent*.

Proof:

We are required to prove

$P[(A \cup B) \cap C] = P(A \cup B)P(C)$

L.H.S. $= P[(A \cap C) \cup (B \cap C)]$ $\quad [\text{Distributive Law}]$

$= P(A \cap C) + P(B \cap C) - P(A \cap B \cap C)$

$= P(A)\,P(C) + P(B)\,P(C) - P(A)\,P(B)\,P(C)$

[A, B, and C mutually independent)

$= P(C)\,[P(A) + P(B) - P(\cap B)]$

$= P(C)\,P(A \cup B) = R.H.S$

Hence $(A \cup B)$ *and C are independent*.

Theorem:

Prove that if A,B and C are random events in a sample space and if A,B,C are pair wise independent and A is independent of $(B \cup C)$, *then A, B and C are mutually independent*

Proof:

We are given,

$$\left.\begin{array}{l} P(A \cap B) = P(A)\,P(B) \\ P(B \cap C) = P(B)\,P(C) \\ P(A \cap C) = P(A)\,P(C) \\ P[A \cap (B \cup C)] = P(A)\,P(B \cup C) \end{array}\right\} \quad (1)$$

Now,

$P[A \cap (B \cup C)] = p[(A \cap B) \cup (A \cap C)$

$= P(A \cap B) + P(A \cap C) - P[(A \cap B) \cap (A \cap C)]$

$= P(A)\,P(B) + P(A)\,P(C) - P(A \cap BC)$………………..(2)

And

$P(A)P(B \cup C) = P(A)[P(B) + P(C) - P(B \cap C)]$

$= P(A)\,P(B) + P(A)\,P(C) - P(A)\,P(B \cap C)$…………(3)

From (2) and (3) on using (1) we get

$P(A \cap B \cap C) = P(A)P(B \cap C)$

$= P(A)\,P(B)\,P(C)$

Hence A, B, C are mutually independent

Theorem:

For any two events A and B, $P(A \cap B) \le P(A) \le P(A \cup B)P(A) + P(B)$

Proof:

We have

$A = (A \cap \bar{B}) \cup (A \cap B)$

We have $P(A) = P\{ (A \cap \bar{B}) \cup (A \cap B)$

$= P(A \cap \bar{B}) + P(A \cap B)$

But $P(A \cap \bar{B}) \ge 0$

$\therefore P(A) \ge P(A \cap B)$

Similarly $P(B) \ge P(A \cap B)$

$\Rightarrow P(B) - P(A \cap B) \ge = o$

Now $P(A \cup B) = P(A) + [P(B) - P(A \cap B)]$

$P(A \cup B) \ge P(A)$

$\Rightarrow P(A) \leq P(A \cup B)$

Also P (A$\cup B) \leq P(A) + P(B)$ From (2)

Hence from (1),(2) and (3) we get

P(A$\cap B) \leq P(A) \leq P(A \cup B) \leq P(A) + P(B)$

Example:

Two dice, one green and the other red, are thrown. Let A be the event that the sum of the points on the faces shown is odd and B the event of at least one ace (number „1")

a. Describe the

i) complete sample space.

ii) events A,B, $\bar{B}$, A$\cap$B, A$\cup$B, and A$\cap$ $\bar{B}$ and find their probabilities assuming all the 36 saple points have equal probabilities.

a. Find the probabilities of the events

i. $(\bar{A} \cup \bar{B})$ ii. $(\bar{A} \cap \bar{B})$ iii. $(A \cap \bar{B})$ iv. $(A \cap B)$

v. $\overline{A \cap B}$ vi. $\bar{A} \cup B$ vii. $(\overline{A \cup B})$ viii. $\bar{A} \cap (A \cup B)$ ix. $A \cup (\bar{A} \cap B)$

x. (A | B) and (B | A) and (A | $\bar{B}$) and ($\bar{B}$ | $\bar{A}$)

Solution:

a.

The sample space S, consists of the 36 elementary events

{(1,1); (1,2); (1,3); (1,4); (1,5); (1,6)
(2,1); (2,2); (2,3); (2,4); (2,5); (2,6)
(3,1); (3,2); (3,3); (3,4); (3,5); (3,6)
(4,1); (4,2); (4,3); (4,4); (4,5); (4,6)
(5,1); (5,2); (5,3); (3,4); (3,5); (3,6)
(4,1); (4,2); (4,3); (4,4); (4,5); (4,6)
(5,1); (5,2); (5,3); (5,4); (5,5); (5,6)
(6,1); (6,2); (6,3); (6,4); (6,5); (6,6)}

for example, the ordered pair (4,5) refers to the elementary event that the green die shows 4 and the red die shows 5.

A= the event that the sum of the numbers shown by the two dice is odd.

= {(1,2); (2,1); (1,4); (2,3); (3,2); (4,1); (1,6); (2,5); (3,4); (4,3); (5,2); (6,1); (3,6); (4,5); (5,4); (6,3); (5,6); (6,5)}

Therefore,

P(A) = $\frac{n(A)}{n(S)} = \frac{18}{36}$

B= the event that at least one face is 1

= {(1,1); (1,2); (1,3); (1,4); (1,5); (1,6); (2,1); (3,1) (4,1); (5,1); (6,1)} therefore

$$P(B) = \frac{n(B)}{n(S)}$$

$$= \frac{11}{36}$$

$\bar{B}$ = the event that each of the face obtained is not an ace

= {(2,2); (2,3); (2,4); (2.5); (2,6); (3,2); (3,3); (3,4); (3,5); (3,6); (4,2); (4,3); (4,4); (4,5); (4,6); (5,2); (5,3); (5,4); (5,5); (5,6); (6,2); (6,3); (6,4); (6,5); (6,6)} therefore

$$P(\bar{B}) = \frac{n(\bar{B})}{n(S)}$$

$$= \frac{25}{36}$$

A∩B = the event that sum is odd and atleast one face is an ace. = {(1,2); (2,1); (1,4); (4,1); (1,6); (6,1)}

$$\therefore P(A \cap B) = \frac{n(A \cap B)}{n(s)} = \frac{6}{36} = \frac{1}{6}$$

A∪B = {(1,2); (2,1); (1,4); (2,3); (3,2); (4,1); (1,6), (2,5); (3,4); (4,3); (5,2); (6,1); (3,6); (4,5); (5,4); (6,3); (5,6); (6,5); (1,1); (1,3); (1,5); (1,5) (3,1), (5,1)}

$$\therefore P(A \cup B) = \frac{n(A \cup B)}{n(s)} = \frac{23}{36}$$

A∩ $\bar{B}$ = {(2,3); (2,5); (3,2); (3,4); (3,6); (4,1); (4,5); (5,2); (5,4) (5,6), (6,3) (6,5)}

$$P(A \cap \bar{B}) = \frac{n(A \cap \bar{B})}{n(S)}$$

$$= \frac{12}{36}$$

$$= \frac{1}{3}$$

b. i. $P(\bar{A} \cup \bar{B}) = P(\overline{A \cap B})$

$= 1 - P(A \cap B)$

$= 1 - \frac{1}{6} = \frac{5}{6}$ ii) $P(\bar{A} \cap \bar{B}) = P(\overline{A \cup B})$

$= 1 - P(A \cup B)$

$= 1 - P(A) - P(B) + P(A \cap B)$

$= 1 - \frac{18}{36} - \frac{11}{36} + \frac{6}{36}$

$= \frac{13}{36}$ iii.

$P(A \cap \bar{B}) = P(A) - P(A \cap B)$

$= \frac{18}{36} - \frac{6}{36}$

$= \frac{12}{36}$

$= \frac{1}{3}$

iv) $P(\bar{A} \cap B) = P(B) - P(A \cap B)$

$= \frac{11}{36} - \frac{6}{36}$

$= \frac{5}{36}$

V) $P(A \cap \bar{B}) = 1 - P(A \cap B)$

$= 1 - \frac{1}{6}$

$= \frac{5}{6}$

vi) $P(\bar{A} \cup B) = P(\bar{A}) + P(B) - P(\bar{A} \cap B)$

$= (1 - \frac{18}{36}) + \frac{11}{36} - \frac{5}{36}$

$= \frac{2}{3}$

vii) $P(\overline{A \cup B}) = 1 - P(A \cup B)$

$= 1 - \frac{23}{36} = \frac{13}{36}$

viii) $P(\bar{A} \cap (A \cup B)] = P[(A \cap \bar{A}) \cup (\bar{A} \cap B)$

$= P(\bar{A} \cap B)$ $[A \cap \bar{A} = \emptyset]$

$= \frac{5}{36}$

ix) $P(A \cup (\bar{A} \cap B)] = P(A) + P(\bar{A} \cap B) - P(A \cap \bar{A} \cap B)$

$= P(A) + P(\bar{A} \cap B)$

$= \frac{18}{36} + \frac{5}{36}$

$= \frac{23}{36}$

x. $P(A | B) = \frac{P(A \cap B)}{P(B)}$

$= \frac{6/36}{11/36}$

$= \frac{6}{11}$

$P(B \mid A) = \frac{P(A \cap B)}{P(A)}$

$= \frac{6/36}{18/36}$

$= \frac{6}{18}$

$= \frac{1}{3}$

xi. $P(\bar{A} \mid B) = \frac{P(\bar{A} \cap B)}{P(B)}$

$= \frac{13/36}{25/36}$

$= \frac{13}{25}$

$P(\bar{B}/\bar{A}) = \frac{P(\bar{A} \cap \bar{B})}{P(\bar{A})}$

$= \frac{13/36}{25/36} = \frac{13}{18}$

Example:

If two dice are thrown, what is the probability that the sum is a) greater than 8 and b) neither 7 nor 11?

Solution:

a.) If S denotes the sum on the two dice then we want P(S>8)
i. S=9, (ii) S=10, iii) S=11 iv) S=12

Hence by addition theorem of probability P(S>8) = P(S=9) + P(S=10) +

P(S=11) + P(S=12)

n(S) = 36

The number of favorable cases can be enumerated as follows S=9: (3,6), (6,3), (4,5),

(5,4)

i.e. 4 sample points

$P(S = 9) = \frac{4}{36}$

S= 10; (4,6), (6,4), (5,5) i.e. 3 sample points

$P(S=10) = \frac{3}{36}$

S=11: (5,6), (6,5) i.e. 2 sample points

$P(S=11) = \frac{2}{36}$

S=12: (6,6) i.e. 1 sample point

$$P(S = 12) = \frac{1}{36}$$

$$P(S>8) = \frac{4}{36} + \frac{3}{36} + \frac{2}{36} + \frac{1}{36} = \frac{10}{36} = \frac{5}{18}$$

b. Let A denotes the event of getting the sum of 7 and B, the event of getting the sum of 1 with a pair of dice.

S=7; (1,6), (6,1), (2,5), (5,2), (3,4), (4,3)

ii. i.e. 6 distinct sample points

$P(A) = P(S=7) = \frac{6}{36} = \frac{1}{6}$

S=11; (5,6), (6,5)

$P(B) = P(S=11) = \frac{2}{36} = \frac{1}{18}$

Required probability = $P(\bar{A} \cap \bar{B})$

$= 1 - P(A \cup B)$

$= 1 - [P(A) + P(B)]$ $(\because A \text{ and } B \text{ are disjoint events})$ $= 1 - \frac{1}{6} - \frac{1}{18}$

$= \frac{7}{9}$

Example:

An urn contains 4 tickets numbered 1,2,3,4 and another contains 6 tickets numbered 2,4,,6,7,8,9. If one of the two urns is chosen at random and a ticket is drawn at random from the chosen urn, find the probability that the ticket drawn bears of the number

i. 2 or 4 (ii) 3 (iii) 1 or 9

Solution:

1. Required event can happen in the following mutually exclusive ways,

I. First urn is chosen and then a ticket is drawn

II. Second urn is chosen and then a ticket is drawn

Since the probability of choosing any urn is ½ the required probability P is given by

$$P = P(I) + P(II)$$

$$= \frac{1}{2} \times \frac{2}{4} + \frac{1}{2} \times \frac{2}{6} = \frac{5}{12}$$

ii) Required probability = $\frac{1}{2} \times \frac{1}{4} + \frac{1}{2} \times 0$

$= \frac{1}{8}$

iii) Required probability = $\frac{1}{2} \times \frac{1}{4} + \frac{1}{2} \times \frac{1}{6} = \frac{5}{24}$

Example:

A card is drawn from a well – shuffled pack of playing cards. What is the probability that it is either a spade or on ace.

Solution:

Let A and B denote the events of events drawing a spade card and an ace respectively. Then A consists of 13 sample points and B consists of A sample points.

i.e. P (A) = $\frac{13}{52}$ and P (B) = $\frac{4}{52}$

The compound event A$\cap B$consists of only one sample point

i.e. P (A$\cap B) = \frac{1}{52}$

The probability that the card drawn is either a spade or an ace is given by

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

$$= \frac{13}{52} + \frac{4}{52} - \frac{1}{52}$$

$$= \frac{4}{13}$$

Example:

A box contains 6 red, 4 white and 5 black balls. A person draws 4 balls from that among the balls drawn there is at least one ball of each color

Solution:

The required event E that in a draw of 4 balls from the box at random there is at least one ball of each color can materialize the following mutually disjoint ways.

i) 1 red, 1 white, 2black balls ii) 2 red, 1 white, 1 black balls iii) 1 red, 2 white, 1 black balls

Hence by the addition theorem of probability the required probability is given by,

P (E) = P(i) + P(ii) + P(iii)

$$= \frac{6c_1 \times 4c_1 \times 5c_2}{15c_4} + \frac{6c_2 \times 4c_1 \times 5c_1}{15c_4} + \frac{6c_1 \times 4c_2 \times 5c_1}{15c_4}$$

$$= \frac{1}{15c_4}[\, 6\times4\times10+15\times4\times5+6\times6\times5]$$

$$= \frac{4!}{15\times14\times13\times12}[\, 240+300+180]$$

$$= \frac{24 \times 720}{15\times14\times13\times12} = 0.5275$$

Example:

A problem is statistic is given to the three students A,B and C whose chances of solving it are ½, -3/4, and ¼ respectively. What is the probability that the problem will be solved?

Solution:

Let A ,B ,C denote the events that the problem is solved by the students A,B,C respectively.

Then, P (A) = ½, P(B) = ¾ and P(C) = 1/4

$$P(A \cup B \cup C) = P(A) + P(B) + P(C) - P(A \cap B) - P(A \cap C) - P(B \cap C) + P(A \cap B \cap C)$$

$$= P(A) + P(B) + P(C) - P(A).P(B) - P(A).P(C) - P(B).P(C) + P(A).P(B).P(C)$$

$$= \frac{1}{2} + \frac{3}{4} + \frac{1}{4} - \frac{1}{2}.\frac{3}{4} - \frac{3}{4}.\frac{1}{4} - \frac{1}{2}.\frac{1}{4} + \frac{1}{2}.\frac{3}{4}.\frac{1}{4}$$

$$= \frac{29}{32}$$

Example:

A bag contains 6 white 9 black balls. Four balls are drawn at a time. Find the probability for the first draw to give 4 white and the second to give 4 black balls in each of the following cases.

i. The balls are replaced before the second draw

ii. The balls are not replaced before the second draw

Solution:

1. The experiment of drawing 4 balls from a bag containing 6 white and 9 black balls result in $15C_4$ ways and hence the sample space consist $15C_4$ sample points.

Let A be the event that the first drawing gives 4 white balls and B be the event that the second drawing gives 4 black balls.

The event A consists of $6C_4$ sample points as there are 6 white balls and 4 are to be chosen from them

P(A) = $\frac{6C_4}{15C_4}$

Now, if the drawn balls are not replaced our sample space is reduced to $11C_4$ points only. The event B that the second draw results in 4 black balls.

$P(\frac{B}{A}) = \frac{9C_4}{11C_4}$

Hence, $P(A \cap B) = P(A) \times P(B|A)$

$$= \frac{6C_4}{15C_4} \times \frac{9C_4}{11C_4}$$

$$= \frac{3}{715}$$

ii) consider the same experiment with replacement

P(A) = $\frac{6C_4}{15C_4}$

Whether A has occurred or not, the probability of drawing 4 black ball in the second draw is $9C_4 \mid 15C_4$

P(A$\cap B) = P(A) \times P(B \mid A)$

= P(A).P(B), as B is independent of A

$$= \frac{6C_4}{15C_4} \times \frac{9C_4}{15C_4} = \frac{6}{5926}$$

Exercise:

1. A bag contains 6 balls of different colors and a ball is drawn from its. A speaks truth thrice out of 4 times and B speaks truth 7 times out of times. If both A and B say that a red ball was drawn, find the probability of their joint statement being true (Ans : 7/15)

2. A and B are two very weak students of statics and their chances of solving a problem correctly are 1/8 and 1/12 respectively if the probability of their making a common mistake is 1/1001 and they obtain the same answer, find the chance that their answer is correct (Ans : 13/14)

3. A bag contains 10 balls, two of which are red three blue and 5 black. Three balls are drawn at random from the bag, that is every ball has an equal chances of being included is the three what is the probability that

   i) The three balls are of different colors

   ii) Two balls are of the same colors and

   iii) The balls are all of the same color?

Ans: $\frac{30}{120}$; $ii)\ \frac{79}{120}$ $iii)\ \frac{11}{120}$

Bayes Theorem Statement

Let $E_1, E_2, \ldots, E_n$ be a set of events associated with a sample space S, where all the events $E_1, E_2, \ldots, E_n$ have nonzero probability of occurrence and they form a partition of S. Let A be any event associated with S, then according to Bayes theorem,

$P(E_i \mid A) = P(E_i)P(A \mid E_i)\sum_{k=1}^{n}P(E_k)P(A \mid E_k)$
for any k = 1, 2, 3, …., n

Bayes Theorem Proof

According to the conditional probability formula,

$$P(E_i|A) = \frac{P(E_i \cap A)}{P(A)} \cdots\cdots\cdots\cdots\cdots(1)$$

Using the multiplication rule of probability,

$$P(E_i \cap A) = P(E_i)P(A|E_i) \cdots\cdots\cdots\cdots\cdots(2)$$

Using total probability theorem,

$$P(A) = \sum_{k=1}^{n} P(E_k)P(A|E_k) \cdots\cdots\cdots\cdots\cdots(3)$$

Putting the values from equations (2) and (3) in equation 1, we get

$$P(E_i|A) = \frac{P(E_i)P(A|E_i)}{\sum_{k=1}^{n} P(E_k)P(A|E_k)}$$

Note:

The following terminologies are also used when the Bayes theorem is applied:

Hypotheses: The events E_1, E_2,… E_n is called the hypotheses

Priori Probability: The probability $P(E_i)$ is considered as the priori probability of hypothesis E_i

Posteriori Probability: The probability $P(E_i|A)$ is considered as the posteriori probability of hypothesis E_i

Bayes Theorem Formula

If A and B are two events, then the formula for Bayes theorem is given by:

P(A|B) = P(A∩B)/P(B)

Where P(A|B) is the probability of condition when event A is occurring while event B has already occurred.

P(A ∩ B) is the probability of event A and event B

P(B) is the probability of event B

Bayes Theorem Derivation

Bayes Theorem can be derived for events and random variables separately using the definition of conditional probability and density.

From the definition of conditional probability, Bayes theorem can be derived for events as given below:

P(A|B) = P(A ∩ B)/ P(B), where P(B) ≠ 0

P(B|A) = P(B ∩ A)/ P(A), where P(A) ≠ 0

Here, the joint probability P(A ∩ B) of both events A and B being true such that,

P(B ∩ A) = P(A ∩ B)

P(A ∩ B) = P(A | B) P(B) = P(B | A) P(A)

P(A|B) = [P(B|A) P(A)]/ P(B), where P(B) ≠ 0

Similarly, from the definition of conditional density, Bayes theorem can be derived for two continuous random variables namely X and Y as given below:

$$f_{X|Y=y}(x) = \frac{f_{X,Y}(x,y)}{f_Y(y)} \quad f_{Y|X=x}(y) = \frac{f_{X,Y}(x,y)}{f_X(x)}$$

Therefore,

$$f_{X|Y=y}(x) = \frac{f_{Y|X=x}(y) f_X(x)}{f_Y(y)}$$

Examples and Solutions

Some illustrations will improve the understanding of the concept.

Example 1:

A bag I contain 4 white and 6 black balls while another Bag II contains 4 white and 3 black balls. One ball is drawn at random from one of the bags, and it is found to be black. Find the probability that it was drawn from Bag I.

Solution:

Let E_1 be the event of choosing the bag I, E_2 the event of choosing the bag II, and A be the event of drawing a black ball.
Then, $P(E_1) = P(E_2) = 12$
Also, $P(A|E_1) = P$(drawing a black ball from Bag I) = $610 = 35$
$P(A|E_2) = P$(drawing a black ball from Bag II) = 37
By using Bayes' theorem, the probability of drawing a black ball from bag I out of two bags,

$P(E_1|A) = P(E_1)P(A|E_1)P(E_1)P(A|E_1)+P(E_2)P(A|E_2)$
$= 12 \times 3512 \times 35 + 12 \times 37 = 712$

Random Variable Definition

A random variable is a rule that assigns a numerical value to each outcome in a sample space. Random variables may be either discrete or continuous. A random variable is said to be discrete if it assumes only specified values in an interval. Otherwise, it is continuous. We generally denote the random variables with capital letters such as X and Y. When X takes values 1, 2, 3, …, it is said to have a discrete random variable.

As a function, a random variable is needed to be measured, which allows probabilities to be assigned to a set of potential values. It is obvious that the results depend on some physical variables which are not predictable. Say, when we toss a fair coin, the final result of happening to be heads or tails will depend on the possible physical conditions. We cannot predict which outcome will be noted. Though there are other probabilities like the coin could break or be lost, such consideration is avoided.

Variate

A variate can be defined as a generalization of the random variable. It has the same properties as that of the random variables without stressing to any particular type of probabilistic experiment. It always obeys a particular probabilistic law.

- A variate is called discrete variate when that variate is not capable of assuming all the values in the provided range.
- If the variate is able to assume all the numerical values provided in the whole range, then it is called continuous variate.

Types of Random Variable

As discussed in the introduction, there are two random variables, such as:

- Discrete Random Variable
- Continuous Random Variable

Let's understand these types of variables in detail along with suitable examples below.

Discrete Random Variable

A discrete random variable can take only a finite number of distinct values such as 0, 1, 2, 3, 4, … and so on. The probability distribution of a random variable has a list of probabilities compared with each of its possible values known as probability mass function.

In an analysis, let a person be chosen at random, and the person's height is demonstrated by a random variable. Logically the random variable is described as a function which relates the person to the person's height. Now in relation with the random variable, it is a probability distribution that enables the calculation of the probability that the height is in any subset of likely values, such as the likelihood that the height is between 175 and 185 cm, or the possibility that the height is either less than 145 or more than 180 cm. Now another random variable could be the person's age which could be either between 45 years to 50 years or less than 40 or more than 50.

Continuous Random Variable

A numerically valued variable is said to be continuous if, in any unit of measurement, whenever it can take on the values a and b. If the random variable X can assume an infinite and uncountable set of values, it is said to be a continuous random variable. When X takes any value in a given interval (a, b), it is said to be a continuous random variable in that interval.

Formally, a continuous random variable is such whose cumulative distribution function is constant throughout. There are no "gaps" in between which would compare to numbers which have a limited probability of occurring. Alternately, these variables almost never take an accurately prescribed value c but there is a positive probability that its value will rest in particular intervals which can be very small.

Random Variable Formula

For a given set of data the mean and variance random variable is calculated by the formula. So, here we will define two major formulas:

- Mean of random variable
- Variance of random variable

Mean of random variable: If X is the random variable and P is the respective probabilities, the mean of a random variable is defined by:

Mean (μ) = $\sum XP$

where variable X consists of all possible values and P consist of respective probabilities.

Variance of Random Variable: The variance tells how much is the spread of random variable X around the mean value. The formula for the variance of a random variable is given by;

$Var(X) = \sigma^2 = E(X^2) - [E(X)]^2$

where $E(X^2) = \sum X^2P$ and $E(X) = \sum XP$

Functions of Random Variables

Let the random variable X assume the values x_1, x_2, …with corresponding probability $P(x_1), P(x_2)$,… then the expected value of the random variable is given by:

Expectation of X, $E(x) = \sum x\, P(x)$.

A new random variable Y can be stated by using a real Borel measurable function $g:\mathbb{R}\rightarrow\mathbb{R}$, to the results of a real-valued random variable X. That is, Y = f(X). The cumulative distribution function of Y is then given by:

$F_Y(y) = P(g(X) \le y)$

If function g is invertible (say $h = g^{-1}$)and is either increasing or decreasing, then the previous relationship can be extended to obtain:

$$F_Y(y) = P(g(X) \leq y) = \begin{cases} P(X \leq h(y)) = F_X(h(y)), & \text{if } h = g^{-1} \text{ increasing,} \\ P(X \geq h(y)) = 1 - F_X(h(y)), & \text{if } h = g^{-1} \text{ decreasing.} \end{cases}$$

Now if we differentiate both the sides of the above expressions with respect to y, then the relation between the probability density functions can be found:

$f_Y(y) = f_x(h(y))|dh(y)/dy|$

Random Variable and Probability Distribution

The probability distribution of a random variable can be

- Theoretical listing of outcomes and probabilities of the outcomes.
- An experimental listing of outcomes associated with their observed relative frequencies.
- A subjective listing of outcomes associated with their subjective probabilities.

The probability of a random variable X which takes the values x is defined as a probability function of X is denoted by f (x) = f (X = x)

A probability distribution always satisfies two conditions:

- $f(x) \geq 0$
- $\sum f(x)=1$

Probability Generating Functions — Introduction A polynomial whose coefficients are the probabilities associated with the different outcomes of throwing a fair die .

G(η) = 0η 0 + 1 6 η 1 + 1 6 η 2 + 1 6 η 3 + 1 6 η 4 + 1 6 η 5 + 1 6 η 6

The coefficients of the square of this polynomial give rise to a higher-order polynomial whose coefficients are the probabilities associated with the different sums which can occur when two dice are thrown. There is nothing special about a fair die.

The set of probabilities associated with almost any discrete distribution can be used as the coefficients of G(η) whose general form is: G(η) = P(X = 0)η 0 + P(X = 1)η 1 + P(X = 2)η 2 + P(X = 3)η 3 + P(X = 4)η 4 + · · ·

) This is a power series which, for any particular distribution, is known as the associated probability generating function.

Commonly one uses the term generating function, without the attribute probability, when the context is obviously probability. Generating functions have interesting properties and can often greatly reduce the amount of hard work which is involved in analysing a distribution. The crucial point to notice, in the power series expansion of G(η), is that the coefficient of η r is the probability P(X = r).

Properties of Generating Functions

It is to be noted that G(0) = P(X = 0) and, rather more importantly, that: G(1) = P(X = 0) + P(X = 1) + P(X = 2) + · · · = X r P(X = r) = 1

In the particular case of the generating function for the fair die: G(1) = 0 + 1 6 + 1 6 + 1 6 + 1 6 + 1 6 + 1 6 = 1

Next, consider G(η) together with its first and second derivatives G′ (η) and G′′(η) (the differentiation is with respect to η of course):

G(η) = P(X = 0)η 0 + P(X = 1)η 1 + P(X = 2)η 2 + P(X = 3)η 3 + P(X = 4)η 4 + · · · G′ (η) = 1 P(X = 1)η 0 + 2 P(X = 2)η 1 + 3 P(X = 3)η 2 + 4 P(X = 4)η 3 + · · · G′′(η) = 2.1 P(X = 2)η 0 + 3.2 P(X = 3)η 1 + 4.3 P(X = 4)η 2 + · · ·

Now, consider G(1), G′ (1) and G′′(1): G(1) = P(X = 0) + P(X = 1) + P(X = 2) + P(X = 3) + P(X = 4) + · · ·

G′ (1) = 1 P(X = 1) + 2 P(X = 2) + 3 P(X = 3) + 4 P(X = 4) + · · · G′′(1) = 2.1 P(X = 2) + 3.2 P(X = 3) + 4.3 P(X = 4) + · · · At this stage, recall the general formula for the expectation of an arbitrary function of a random variable: E

Mathematical expectation.

Definition:

Let x be a discrete random variable which can assume any of the values $x_1, x_2, .. x_n$ with corresponding probabilities $P_i = P(x=x_i)$ i=1,2,… Then the mathematical expectation of x, denoted by E(x) is defined by

$E(x) = \sum_i P_i X_i$ provided the series is absolutely convergent.

Example:

1.Let X be the discrete random variable taking the values 1,2,…, 6 with corresponding probabilities P_i = 1/6

Then $E(x) = \sum_{i=1}^{6} P_i X_i$

$$= \frac{1}{6}(1) + \frac{1}{6}(2) + \dots\dots + \frac{1}{6}(6)$$
$$= \frac{1}{6}(1+2+3+4+5+6) = \frac{7}{2}$$

2. Let x be a random variable having the p.d.f

$$P(x) = \begin{cases} \frac{x}{6} & if\ x = 1,2,3 \\ 0 & otherwise \end{cases}$$

Then $E(x^3+2x^2) = E(x^3) + 2E(x^2)$

$$= \sum_{i=1}^{3} x^3 P(x_i) + 2\sum_{i=1}^{3} x^2 P(x_i)$$
$$= \left(\frac{1}{6} + \frac{16}{6} + \frac{81}{6}\right) + 2\left(\frac{1}{6} + \frac{8}{6} + \frac{27}{6}\right)$$
$$= \frac{98}{6} + 2 \times \frac{36}{6}$$
$$= \frac{170}{6}$$
$$= \frac{85}{3}$$

Mathematical expectation of continuous random variable

Definition:

If x is a continuous random variable with p.d.f f (x) then mathematical expectation of x is defined to be $E(X) = \int_{-\infty}^{\infty} x f(x) dx$

Provided the integral is absolutely convergent.

E $(\psi(x)) = \int_{-\infty}^{\infty} \psi(x) f(x)\, dx$ where $\psi(x)$ is a function of r.v. X.

Example: Let x have a p.d.f f(x) = $\begin{cases} \frac{x+2}{18} & if -2 < x < 4 \\ 0 & \text{otherwise} \end{cases}$

Solution:

i. $E(x) = \int_{-2}^{4} x\left(\frac{x+2}{18}\right) dx$

$= \frac{1}{18}\left[\frac{x^3}{3} + x^2\right]_{-2}^{4}$

$= \frac{1}{18}\left[\left(\frac{64}{3} + 16\right) - \left(\frac{-8}{3} + 4\right)\right]$

$= \frac{1}{18}\left[\frac{108}{3}\right] \quad = 2$

ii. $E[(x+2)^2] = E(x^2+4x+4)$

$= E(x^2) + 4E(x) + 4$

$= \int_{-2}^{4} x^2\left(\frac{x+2}{18}\right) dx + 4E(x) + 4$

$= \frac{1}{18}\left[\frac{x^4}{4} + \frac{2x^3}{3}\right]_{-2}^{4} + (4x2 + 4)$

$= \frac{1}{18}\left[\left(64 + \frac{128}{3}\right) - \left(4 - \frac{16}{3}\right)\right] + 12$

$= \frac{1}{18}\left[\frac{320}{3} + \frac{4}{3}\right] + 12$

$= 18 + 12 = 30$

Definition:

Let X be a r.v E(x) is called the mean value of x and is denoted by μ.

Hence $\bar{X} = \mu = E(X)$

$E(X^r)$, $r \geq 1$ is called the r^{th} moment of X about the origin and is denoted by μ_r^1. Hence $\mu_r' = E(X^r)$ and

$\bar{x} = \mu_1'$

$E(X-\mu)^2$ is called the variance of X and is denoted by σ^2. The positive square root σ of the variance is called the standard deviation of X.

$E(X-\mu)^r$ is called the r^{th} central moment of X and is denoted by μ_r.

Hence $\mu_r = E(X-\mu)^r$

Lemma:

σ^2= E(X²) – $[E(x)]^2 = \mu_2 - \mu_1'^2$

Proof:

$$\sigma^2 = E[(x-\mu)^2]$$

$$= E(x^2 - 2\mu x + \mu^2)$$

$$= E(x^2) - 2\mu E(x) + \mu^2$$

$$= E(x^2) - 2[E(x)]^2 + [E(x)]^2$$

$$= E(x^2) - [E(x)]^2$$

$$= \mu_2' - \mu_1'^2$$

Lemma:

$$\mu_r = \mu_r' - r_{c_1}\mu\mu_{r+1} + r_{c_2}\mu^2\mu_{r-2}..........$$

Proof:

$\mu_r = E[(X-\mu)^r]$

$E(X^r - r_{c_1}\mu X^{r-1}$ = $+ \cdots..)$

$=\mu_r' - r_{c_1}\mu\mu_{r-1} + r_{c_2}\mu^2\mu_{r-2}..........$

In particular, $\mu_1 = E(X-\mu) = E(X) - \mu = \mu - \mu = 0$.

$$\mu_2 = \mu_2' - 2\mu_1'\mu + \mu^2\mu_0$$

$\therefore \mu_2 = \mu_2' - \mu_1'^2$ (Since $\mu = \mu_1'$ and $\mu_0' = 1$)

$$\mu_3 = \mu_3' - 3\mu_2'\mu_1' + 2(\mu_1')^2$$

Problem:

A random variable X is defined as the sum of the numbers on the faces when two dice are thrown. Find the expected value of X.

Solution:

The probability distribution of X is given by the following table.

x_i	2	3	4	5	6	7	8	9	10	11	12
P(x_i)	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{3}{36}$	$\frac{4}{36}$	$\frac{5}{36}$	$\frac{6}{36}$	$\frac{5}{36}$	$\frac{4}{36}$	$\frac{3}{36}$	$\frac{2}{36}$	$\frac{1}{36}$

$E\ X =$

$x_i p(x_i)$

=2(1/36)+3(2/36)+4(3/36)+5(4/36)+6(5/36)+7(6/36)+8(5/36)+9(4/36)+10(3/36)+11(2/36)+12(1/36)=252/36=7.

Problem:

A random variable X has the following probability functions.

x_i	-2	-1	0	1	2	3
P(x_i)	0.1	K	0.2	2k	0.3	k

Find (i) the value of k (ii) mean (iii) variance (iv) $p(x \geq 2)$ (v) $p(x < 2)$

(vi) $p(-1 < x < 3)$.

Solution:

(i) $\sum p_i = 1$

$\Rightarrow$ 0.1+k+0.2+2k+0.3+k=1

$\Rightarrow$4k=0.4
$\Rightarrow$k=0.1.

x_i	-2	-1	0	1	2	3
P(x_i)	0.1	0.1	0.2	0.2	0.3	0.1

Hence the probability function is

(ii) Mean=$E(X) = \sum x_i p(x_i)$.

=(-2)(0.1)+(-1)(0.1)+(0)(0.2)+(1)(0.2)+2(0.3)+3(0.1)

=0.8.

(iii) variance=$E(X^2) - [E(X)]^2$

=$\sum x_i^2 p(x_i)$-0.8²

=[$4(0.1) + 1(0.1) + 1(0.2) + 4(0.3) + 9(0.1)$

=2.8-0.64=2.16.

(iv) $p(x \geq 2) = p(X = 2) + p(= 3)$=0.3+0.1=0.4.

(v) $p(x < 2) = 1 - p(x \geq 2) = 0.6$.

(vi) $p(-1 < x < 3) = p(X = 0) + p(X = 1) + p(X = 2)$

=0.2+0.2+0.3=0.7.

Problem:

Let X have the p.d.f $f(x)= \begin{cases} \frac{x+1}{2}, & \text{if } -1<x<1 \\ 0, & \text{otherwise.} \end{cases}$

Find the mean and standard deviation of x.

Solution:

$$\mu = E(x) = \int_{-\infty}^{\infty} x f(x)dx = \int_{-1}^{1} x\left(\frac{x+1}{2}\right)dx$$

$$=1/2\left[\frac{x^3}{3}+\frac{x^2}{2}\right]_{-1}^{1}$$

$$\mu = \frac{1}{3}$$

$$\sigma^2 = E(X^2)-[E(X)]^2$$

$$= \int_{-1}^{1} \frac{x^2}{2}(x+1)dx - \left(\frac{1}{3}\right)^2$$

$$=1/2\left[\frac{x^4}{4}+\frac{x^3}{3}\right]_{-1}^{1} - \frac{1}{9}$$

$$\sigma^2 = \frac{2}{9}$$

Moment Generating Function:

Definition:

The moment generating function (m.g.f) for any random variable X about the origin is defined by $M_X(t) = E(e^{tx}) =$

$$f(x) = \begin{cases} \int (e^{tx}) f(x)dx & \text{if } X \text{ is a continuos r.v with p.d.f } f(x) \\ \sum_x e^{tx} P(x), & \text{if } X \text{ is a discrete r.v with p.d.f } P(x) \end{cases}$$

Where the integration or summation is taken over the entire range of X and t is a real parameter.

Definition:

More generally the m.g.f of a random X about a point a denoted by $M_{X=a}(t)$ *is defined by* $M_{X=a}(t) = E\left(e^{t(x-a)}\right)$.

Example :

Define $P(x) = \begin{cases} \frac{6}{\pi^2 x^2}, & \text{if } x = 1,2,3, \ldots \\ 0, & \text{otherwise} \end{cases}$

Now, $\sum P(x) = \sum \frac{6}{\pi^2 x^2} = \frac{6}{\pi^2}\sum_{n=1}^{\infty}\frac{1}{n^2}$

$=\frac{6}{\pi^2} \times \frac{\pi^2}{6}$ (Since $1 + \frac{1}{2^2} + \frac{1}{3^3} \dots += \frac{\pi^2}{6}$)$=1$

$\therefore P(x) \text{ is a probability distribution of the r.v } X$.

Now, the m.g.f of X, if it exists, is given by $M_X(t) = E(e^{tx}) = \sum e^{tx} P(x)$

$$=\sum e^{tx}\left(\frac{6}{\pi^2 x^2}\right)$$

$$=\frac{6}{\pi^2}\sum_{n=1}^{\infty}\left(\frac{e^{tn}}{n^2}\right)$$

The above series, by ratio test ,diverges for t>0.

$\therefore$ *Moment generating function of X does not exist*.

Properties of moment generating function:

1. The r^{th} derivative of m.g.f of r.v X at t=0 is μ_r'.

Proof:

Let $M_X(t)$ be the m.g.f of the r.v X

Then $M_X(t) = 1 + \mu_1' + \frac{\mu_2'}{2!}t^2 + \cdots + \frac{\mu_r'}{r!}t^r + \cdots$

$$\therefore \frac{d^r}{dt^r}(M_X(t)) = \frac{\mu_r' \, r!}{r!} + \frac{t\mu_{r+1}'}{(r+1)!} + \dots$$

At t=0, $\frac{d^r}{dt^r}(M_X(t)) = \mu_r'$

Problem:

A random variable X has the probability function

P(x)=$\frac{1}{2^x}$; $x = 1,2,3 \dots$ *Find its (i) m.g.f*

(ii) Mean and (iii) Variance.

Solution:

X is a discrete random variable

(i) ($M_X(t)$=E(e^{tx}) $= \sum_x P(x)$

$= \sum_{x=1}^{\infty} e^{tx}\left(\frac{1}{2^x}\right)$

$=\sum_{x=1}^{\infty}\left(\frac{e^t}{2}\right)^x$

$=\frac{e^t}{2}+\left(\frac{e^t}{2}\right)^2+\ldots$

$=\frac{e^t}{2}\left[1+\frac{e^t}{2}+\left(\frac{e^t}{2}\right)^2\right]$

$=\frac{e^t}{2}\left[\frac{1}{1-\frac{e^t}{2}}\right]$

$M_X(t) = \frac{e^t}{2-e^t}$

(ii) $\frac{d}{dt}M_X(t) = \frac{(2-e^t)e^t+e^te^t}{(2-e^t)^2} = \frac{2e^t}{(2-e^t)^2}$

$\therefore \mu_1' = \left[\frac{d}{dt}M_X(t)\right]_{t=0} = 2$

$\frac{d^2}{dt^2}(M_X(t) = \frac{(2-e^t)^2 2e^t+2e^t 2(2-e^t)}{(2-e^t)^4}$

$=\frac{8e^t-2e^{2t}}{(2-e^t)^3}$

$\therefore \mu_2' = \left[\frac{d^2}{dt^2}M_X(t)\right]_{t=0} = 6$

(iii) Variance $\mu_2 = \mu_2' - (\mu_1')^2$

$=6-4 = 2$

Problem :

Find the m.g.f of the r.v X having the p.d.f

$$f(x)=\begin{cases}\frac{-1}{3}, & -1<x<2\\ 0, & otherwise\end{cases}$$

Solution:

X is a continuous random variable.

$$M_x(t) = E(e^{tx}) = \int_{-\infty}^{\infty} e^{tx} f(x)dx$$

$=\frac{1}{3}\int_{-1}^{2} e^{tx}\, dx = \frac{1}{3}\left[\frac{e^{tx}}{t}\right]_{-1}^{2}$ when $t\neq 0$

$=\frac{e^{2t}-e^{-t}}{3t}$ when $t \neq 0$

when $t = 0$, $M_x(t) = \frac{1}{3}\int_{-1}^{2} dx = \frac{1}{3}[x]_{-1}^{2} = 1$

HAR $M_x(t) = \begin{cases} \frac{e^{2t}-e^{-t}}{3t}, & \text{when } t \neq 0 \\ 1 & \text{when } t = 0 \end{cases}$

UNIT III

(Characteristic Functions) The characteristic function of a random variable X is defined as

$$\Phi_X(t) = Ee^{itX}, t \in \mathbb{R}.$$

(where $Ee^{itX} = E\cos tX + iE\sin tX$)

Example 7.0.43 Let $X \sim \text{Bernoulli}(p)$. Then

$$\phi_X(t) = (1-p) + pe^{it}.$$

$X \sim \text{exponential}(\lambda)$. Then

$$\begin{aligned} \Phi_X(t) = Ee^{itX} &= \lambda \int_0^\infty e^{itx} e^{-\lambda x} dx \\ &= \lambda(\int_0^\infty \cos tx e^- \lambda x dx \\ &= i\int_0^\infty \sin tx e^{-\lambda x} dx) \\ &= \lambda(\frac{\lambda}{\lambda^2+t^2} + i\frac{t}{\lambda^2+t^2}) \\ &= \frac{\lambda(\lambda+it)}{\lambda^2+t^2}, t \in \mathbb{R}. \end{aligned}$$

Theorem For any random variable X, its characteristic function $\phi_X(\cdot)$ is uniformly continuous on $\mathbb{R}$ and satisfies

(i) $\Phi_X(0) = 1$

(ii) $|\Phi_X(t)| \leq 1$

(iii) $\Phi_X(-t) = \overline{\Phi_X(t)}$, where for z a complex number, $\bar{z}$ denote the conjugate.

Proof: We prove (iii), (i) and (ii) are exercises.

$$\begin{aligned} \Phi_X(-t) = Ee^{-itX} &= E\cos tX - iE\sin tX \\ &= \overline{E\cos tX + iE\sin tX} \\ &= \overline{\Phi_X(t)}. \end{aligned}$$

$$\begin{aligned} |\Phi_X(t+h) - \Phi_X(t)| &= |E(e^{i(t+h)X} - e^{itX})|, \\ &\leq E|e^{ihX} - 1| \\ &= E\sqrt{2(1-\cos(hX))} \end{aligned}$$

Using Dominated Convergence theorem, uniformly in as $h \to 0$. This imply that Φ_X is uniformly continuous.

Theorem If the random variable X has finite moments upto order n. Then Φ has continuous derivatives upto order n.More over

$$i^k EX^k = \Phi_X^{(k)}(0), \ k = 1, 2, \ldots, n.$$

Proof. Consider

$$\frac{\Phi_X(t+h) - \Phi_X(t)}{h} = E[e^{itX}\frac{(e^{ihX} - 1)}{h}]$$

since $|e^{ihx} - 1| \leq |hx|)$, we get

$$E[|e^{itX}\frac{(e^{ihX} - 1)}{h}|] \leq E|X| < \infty.$$

Hence by Dominated Convergence theorem

$$\lim_{h \to 0} E[e^{itX}\frac{(e^{ihX} - 1)}{h}] = E[iXe^{itX}].$$

Therefore

$$\Phi'_X(t) = E[iXe^{itX}].$$

Put $t = 0$, we get

$$\Phi_X^{(1)}(0) = i\,EX.$$

For higher order derivatives, repeat the above arguments.

(Inversion theorem) Let X be a random variable with distribution function F and characteristic function $\phi_X(\cdot)$. Then

$$F(b) - F(a) = \frac{1}{2\pi}\int_{-\infty}^{\infty} \frac{e^{-ita} - e^{-itb}}{it}\Phi_X(t)dt,$$

whenever a, b are points of continuity of F.

Proof. Consider

$$\begin{aligned}\frac{1}{2\pi}\int_{-\infty}^{\infty}\frac{e^{-ita}-e^{-itb}}{it}\Phi_X(t)dt &= \frac{1}{2\pi}\int_{-\infty}^{\infty}\frac{e^{-ita}-e^{-itb}}{it}Ee^{itX}dt \\ &= \frac{1}{2\pi}E\int_{-\infty}^{\infty}\frac{e^{-ita}-e^{-itb}}{it}e^{itX}dt \\ &= E\int_{-\infty}^{\infty}\frac{e^{it(X-a)}-e^{it(X-b)}}{2\pi it}dt\,.\end{aligned} \qquad (7.0.1$$

$$\int_{-\infty}^{0}\frac{e^{it(X-a)}-e^{it(X-b)}}{2\pi it}dt \;=\; \int_{0}^{\infty}\frac{e^{-it(X-a)}-e^{-it(X-b)}}{2\pi it}dt$$

Hence, using $2iSin\theta \;=\; e^{i\theta}-e^{-i\theta}$, we have

$$\int_{-\infty}^{\infty}\frac{e^{it(X-a)}-e^{it(X-b)}}{2\pi it}dt \;=\; \frac{1}{\pi}\int_{0}^{\infty}\frac{Sin\,t(X-a)}{t}dt \;-\; \frac{1}{\pi}\int_{0}^{\infty}\frac{Sin\,t(X-b)}{t}dt\,.$$

$$\int_{0}^{\infty}\frac{Sin\,\alpha x}{x}dx \;=\; \frac{\pi}{2}sgn(\alpha)$$

$$\frac{1}{\pi}\int_{0}^{\infty}\frac{Sin\,t(X-a)}{t}dt \;=\; \begin{cases}\frac{1}{2} & \text{if} \quad X>a \\ 0 & \text{if} \quad X=a \\ -\frac{1}{2} & \text{if} \quad X<a\,,\end{cases}$$

$$sgn(\alpha) \;=\; \begin{cases}-1 & \text{if} \quad \alpha<0 \\ 0 & \text{if} \quad \alpha=0 \\ 1 & \text{if} \quad \alpha>0\,.\end{cases}$$

where
Similarly, the other integral. Combining (7.0.1), (7.0.3) and (7.0.4), we complete the proof.

Theorem (Uniqueness Theorem)

Let X_1, X_2 be two random variables such that $\Phi_{X1} \equiv \Phi_{X2}$. Then X_1, X_2 have same distribution.

Proof: Using Inversion theorem, we have

$$F_1(b) - F_1(a) = F_2(b) - F_2(a)$$

for all $a, b \in \mathbb{R}$ F_1, F_2 a b

Now let $a \to -\infty$, we have

$$F_1(b) = F_2(b)$$

for all b at which F_1 and F_2 are continuous. Therefore

$$F_1 \equiv F_2 \text{ (Exercise)}$$

Properties

- The characteristic function of a real-valued random variable always exists, since it is an integral of a bounded continuous function over a space whose measure is finite.
- A characteristic function is uniformly continuous on the entire space
- It is non-vanishing in a region around zero: $\varphi(0) = 1$.
- It is bounded: $|\varphi(t)| \leq 1$.
- It is Hermitian: $\varphi(-t) = \varphi(t)$. In particular, the characteristic function of a symmetric (around the origin) random variable is real-valued and even.
- There is a bijection between probability distributions and characteristic functions. That is, for any two random variables X_1, X_2, both have the same probability distribution if and only if .
- If a random variable X has moments up to k-th order, then the characteristic function φ_X is k times continuously differentiable on the entire real line. In this case

- If a characteristic function φ_X has a k-th derivative at zero, then the random variable X has all moments up to k if k is even, but only up to $k-1$ if k is odd.[11]

- If $X_1, ..., X_n$ are independent random variables, and $a_1, ..., a_n$ are some constants, then the characteristic function of the linear combination of the X_i 's is

- Let $\phi(t)$ be the characteristic function, defined as the Fourier transform of the probability density function $P(x)$ *using Fourier transform parameters* $a=b=1$,

$$\phi(t) = \mathcal{F}_x[P(x)](t) \tag{1}$$

$$= \int_{-\infty}^{\infty} e^{itx} P(x)\, dx. \tag{2}$$

- The cumulants κ_n are then defined by

$$\ln \phi(t) \equiv \sum_{n=1}^{\infty} \kappa_n \frac{(it)^n}{n!} \tag{3}$$

- (Abramowitz and Stegun 1972, p. 928). Taking the Maclaurin series gives

$$\begin{aligned}\ln \phi(t) = {} & (it)\mu_1' + \tfrac{1}{2}(it)^2(\mu_2' - \mu_1'^2) + \tfrac{1}{3!}(it)^3(2\mu_1'^3 - 3\mu_1'\mu_2' + \mu_3') + \\ & \tfrac{1}{4!}(it)^4(-6\mu_1'^4 + 12\mu_1'^2\mu_2' - 3\mu_2'^2 - 4\mu_1'\mu_3' + \mu_4') + \\ & \tfrac{1}{5!}(it)^5[24\mu_1'^5 - 60\mu_1'^3\mu_2' + 20\mu_1'^2\mu_3' - 10\mu_2'\mu_3' + 5\mu_1'(6\mu_2'^2 - \mu_4') + \mu_5'] + \ldots,\end{aligned} \tag{4}$$

- where μ_n' are raw moments, so

$$\kappa_1 = \mu_1' \tag{5}$$

$$\kappa_2 = \mu_2' - \mu_1'^2 \tag{6}$$

$$\kappa_3 = 2\mu_1'^3 - 3\mu_1'\mu_2' + \mu_3' \tag{7}$$

$$\kappa_4 = -6\mu_1'^4 + 12\mu_1'^2\mu_2' - 3\mu_2'^2 - 4\mu_1'\mu_3' + \mu_4' \tag{8}$$

$$\kappa_5 = 24\mu_1'^5 - 60\mu_1'^3\mu_2' + 20\mu_1'^2\mu_3' - 10\mu_2'\mu_3' + 5\mu_1'(6\mu_2'^2 - \mu_4') + \mu_5'. \tag{9}$$

- These transformations can be given by `CumulantToRaw`[*n*] in the Mathematica application package `mathStatica`.
- In terms of the central moments μ_n,

$$\kappa_1 = \mu \tag{10}$$

$$\kappa_2 = \mu_2 \tag{11}$$

$$\kappa_3 = \mu_3 \tag{12}$$

$$\kappa_4 = \mu_4 - 3\,\mu_2^2 \tag{13}$$

$$\kappa_5 = \mu_5 - 10\,\mu_2\,\mu_3, \tag{14}$$

- where μ is the mean and $\sigma^2 \equiv \mu_2$ is the variance. These transformations can be given by `CumulantToCentral`[*n*].
- Multivariate cumulants can be expressed in terms of raw moments, e.g.,

$$\kappa_{1,1} = -\mu'_{0,1}\,\mu'_{1,0} + \mu'_{1,1} \tag{15}$$

$$\kappa_{2,1} = 2\,\mu'_{0,1}\,{\mu'_{1,0}}^2 - 2\,\mu'_{1,0}\,\mu'_{1,1} - \mu'_{0,1}\,\mu'_{2,0} + \mu'_{2,1}, \tag{16}$$

- and central moments, e.g.,

$$\kappa_{1,1} = \mu_{1,1} \tag{17}$$

$$\kappa_{2,1} = \mu_{2,1} \tag{18}$$

$$\kappa_{3,1} = -3\,\mu_{1,1}\,\mu_{2,0} + \mu_{3,1} \tag{19}$$

$$\kappa_{4,1} = -6\,\mu_{2,0}\,\mu_{2,1} - 4\,\mu_{1,1}\,\mu_{3,0} + \mu_{4,1} \tag{20}$$

$$\kappa_{5,1} = 30\,\mu_{1,1}\,\mu_{2,0}^2 - 10\,\mu_{2,1}\,\mu_{3,0} - 10\,\mu_{2,0}\,\mu_{3,1} - 5\,\mu_{1,1}\,\mu_{4,0} + \mu_{5,1} \tag{21}$$

- using `CumulantToRaw`[{*m*, *n*, ...}] and `CumulantToCentral`[{*m*, *n*, ...}], respectively.
- The *k*-statistics are unbiased estimators of the cumulants.

Chebyshev Inequality

The Chebyshev inequality enables us to obtain bounds on probability when both the mean and variance of a random variable are known. The inequality can be stated as follows:

Chebyshev's Theorem states:

$$P\left(|X-\mu| \leq k\sigma\right) \geq 1-\frac{1}{k^2} \quad \text{or} \quad P\left(|X-\mu| > k\sigma\right) \leq \frac{1}{k^2}$$

$\mu = \textit{mean}$ $\quad \sigma = \textit{standard deviation}$ $\quad k = \#\ \textit{of standard deviations we are willing to go from the mean}$

Chebyshev's Inequality Formula

Remember, Chebyshev's inequality implies that it is unlikely that a random variable will be far from the mean. Hence, our k-value is our limit that we set, stating the number of standard deviations we are willing to go away from the mean.

Example

So now let's look at an example. Suppose 1,000 applicants show up for a job interview, but there are only 70 positions available. To select the best 70 people amongst the 1,000 applicants, the employer gives an aptitude test to judge their abilities. The mean score on the test is 60, with a standard deviation of 6. If an applicant scores an 84, can they assume they are getting a job?

$$\text{If} \quad \mu = 60 \quad \sigma = 6 \quad k = ? \quad P(x > 84) = ?$$

$$|X - \mu| \rightarrow |84 - 60| = 24 \quad \textit{and} \quad 24 = \underbrace{4(6)}_{k\sigma} \quad \textit{so} \quad k = 4$$

$$\text{If } P(|X - \mu| > k\sigma) \leq \frac{1}{k^2}, \text{ then } P(X > 84) \leq \frac{1}{(4)^2} = 0.0625$$

$$1{,}000(0.0625) = 62.5$$

Chebyshev's Inequality Example

Our results show that about 63 people scored above a 60, so with 70 positions available, an applicant who scores an 84 can be assured they got the job.

UNIT IV

CONCEPTS OF BIVRIATE DISTRIBUTIONS

Bivariate Distributions

- extend the definition of a probability distribution of one random variable to the **joint probability distribution** of two random variables
- learn how to use the **correlation coefficient** as a way of quantifying the extent two which two random variables are linearly related
- extend the definition of the conditional probability of events in order to find the **conditional probability distribution** of a random variable *X* given that *Y* has occurred
- investigate a particular joint probability distribution, namely the **bivariate normal distribution**

CORRELATION AND REGRESSION

CORRELATION:

Definition:

Consider a set of bivariate data (x_i, y_i) i=1,2….n. If there is a change in one variable corresponding to a change in the other variable we say that the variables are correlated.

If the two variable two variables deviate in the same direction the correlation is said to be direct or positive.

Definition:

The covariance between x and y is defined by

Cov (x,y) = $\frac{\sum(x_i-\bar{x})(y_i-\bar{y})}{n}$

Hence $\gamma_{xy} = \frac{cov\ (x,y)}{\sigma_x \sigma_y}$

Example:

The heights and weights of five students are given below.

Height in c.m x	160	161	162	163	164
Weight in kgs y	50	53	54	56	57

$Here\ x = 162; y = 54$; $\varsigma_x = 2\ \ and\ \varsigma_y = 6$

Now $\sum(x_i - \bar{x})(y_i - \bar{y}) = (-2)(-4) + (-1)(-1) + 0 + (1 \times 2) + (2 \times 3) = 17$

$$\therefore \gamma_{xy} = \frac{\sum(x_i-\bar{x})(y_i-\bar{y})}{n\sigma_x\sigma_y} = \frac{17}{5\sqrt{2}\sqrt{6}} = \frac{17 \times \sqrt{12}}{60} = \frac{17 \times 3.46}{60} = 0.98$$

Theorem:

$$\gamma_{xy} = \frac{n\sum x_i y_i - \sum x_i \sum y_i}{[n\sum x_i^2 - (\sum x_i)^2]^{1/2}[n\sum y_i^2 - (\sum y_i)^2]^{1/2}}$$

Proof:

$$\gamma_{xy} = \frac{\sum(x_i - \bar{x})(y_i - \bar{y})}{n\sigma_x\sigma_y} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (1)$$

$$\sum(x_i - \bar{x})(y_i - \bar{y}) = \sum x_i y_i - \bar{x}\sum y_i - \bar{y}\sum x_i + n\bar{x}\bar{y}$$

$$= \sum x_i y_i - \bar{x}(n\bar{y}) - \bar{y}(n\bar{x}) + n\bar{x}\bar{y}$$

$$= \sum x_i y_i - n\bar{x}\bar{y}$$

$$= \sum x_i y_i - (\tfrac{1}{n}) \sum x_i y_i$$

$$= \tfrac{1}{n}[\mathrm{n}\sum x_i y_i - \sum x_i \sum y_i] \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots . . (2)$$

Also,

$$\sigma_x^2 = \frac{1}{n}\sum(x_i - \bar{x})^2$$

$$= \tfrac{1}{n}\ [\sum x_i^2 - 2\bar{x}\sum x_i + n(\bar{x})^2]$$

$$= \tfrac{1}{n}\ [\sum x_i^2 - 2n(\bar{x})^2 + n\ (\bar{x})^2]$$

$$= \tfrac{1}{n}\left[\sum x_i^2 - \tfrac{1}{n}(\sum x_i)^2\right]$$

$$= \tfrac{1}{n^2}[n\sum x_i^2 - (\sum x_i)]^2$$

$$\therefore \sigma_x = \tfrac{1}{n}\ [n\sum x_i^2 - (\sum x_i)^2]^{1/2} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots . (3)$$

Similarly,

$$\sigma_y = \tfrac{1}{n}\ [n\sum y_i^2 - (\sum y_i)^2]^{1/2} \quad \ldots\ldots\ldots\ldots . (4)$$

Substituting (2),(3) and (4) in (1) we get the required result.

Theorem:

-1 $\leq \gamma \leq 1$ **Proof**:

$$\gamma_{xy} = \frac{\sum(x_i - \bar{x})(y_i - \bar{y})}{n\sigma_x\sigma_y}$$

$$= \frac{\left(\frac{1}{n}\right)\sum(x_i-\bar{x})(y_i-\bar{y})}{\left[\frac{1}{n}\sum(x_i-\bar{x})^2\right]^{1/2}\left[\frac{1}{n}\sum(y_i-\bar{y})^2\right]^{1/2}}$$

Let $a_i = x_i - \bar{x}$ and $b_i = y_i - \bar{y}$

$$\therefore \gamma_{xy}^2 = \frac{(\sum a_i b_i)^2}{(\sum a_i^2)(\sum b_i^2)}$$

By Schwartz inequality we have,

$$\left(\sum a_i b_i\right)^2 \le \left(\sum a_i\right)^2 \left(\sum b_i\right)^2$$

Hence $\gamma_{xy}^2 \le 1$

$$\therefore |\gamma_{xy}| \le 1$$

$$\therefore -1 \le \gamma \le 1$$

Note: 1

If $\gamma = 1$ *the correlation is perfet and positive*.

Note: 2

If $\gamma = -1$ the correlation is perfect and negative.

Note: 3

If $\gamma = 0$ the variables are uncorrelated.

Note: 4

If the variables x and y are uncorrelated then cov (x,y) = 0

Problem: **1**

Ten students obtained the following percentage of marks in the college internal test (x) and in the final university examination (y). Find the correlation coefficient between the marks of the two tests.

x	51	63	63	49	50	60	65	63	46	50

y	49	72	75	50	48	60	70	48	60	56

Solution:

Choosing the origin A = 63 for the variable x and B= 60 for y and taking $u_i = x_i - A$ and $v_i = y_i - B$.

We have the following table:

x_i	u_i	y_i	v_i	u_i^2	v_i^2	
51	-12	49	-11	144	121	

63	0	72	12	0	144
63	0	75	15	0	225
49	-14	50	-10	196	100
50	-13	48	-12	169	144
60	-3	60	0	9	0
65	2	70	10	4	100
63	0	48	-12	0	144
46	-17	60	0	289	0
50	-13	56	-4	169	16
Total	-70	-	-12	980	994

u_iv_i
132
0
0
140
156
0
20
0
0
52
500

$$\gamma_{xy} = \gamma_{uv}$$

$$= \frac{n\sum u_i v_i - \sum u_i \sum v_i}{\left[n\sum u_i^2 - (\sum u_i)^2\right]^{1/2}\left[n\sum v_i^2 - \sum(v_i)^2\right]^{1/2}}$$

$$= \frac{10\times500-(-70)\times(-12)}{[10\times980-(-70)^2]^{1/2}[10\times994-(-12)^2]^{1/2}}$$

$$= \frac{4160}{70\times98.97}$$

=0.6

Problem: 2

If x and y are two variable. Prove that the correlation coefficient between $ax + b$and $cy + d$is

$$\gamma_{ax+b,cy+d} = \frac{ac}{|ac|}\gamma_{xy} \quad if\ a, c \neq 0.$$

Proof:

Let $u = ax + b$and $v = cy + d$

$$\therefore \bar{u} = a\bar{x} + b \ and\ \bar{v} = c\bar{y} + d$$

$$\sigma_u^2 = \frac{1}{n}\sum(u - \bar{u})2$$

$$= \frac{a^2}{n}\sum(x_i - \bar{x})2$$

$$= a^2\sigma_x^2$$

Similarly,

$$\sigma_v^2 = c^2\sigma_y^2$$

Now,

$$\gamma_{uv} = \frac{\sum(u-\bar{u})(v-\bar{v})}{n\sigma_u\sigma_v}$$

$$= \frac{\sum a(x-\bar{x})c(y-\bar{y})}{n|ac|\sigma_x\sigma_y}$$

$$= \frac{ac}{|ac|}\gamma_{xy}.$$

Problem: 3

A programmer while writing a program for correlation coefficient between two variable x and y from 30 pairs of observations obtained the following results $\sum x = 300;\ \sum x^2 = 3718, \sum y = 210;\ \sum y^2 = 2000;\ \sum xy = 2100$. At the time of checking it was found that he had copied down two pairs (x_i, y_i) as (18, 20) and (12, 10) instead of the correct values (10,15) and (20,15). Obtain the correct value of the correlation coefficient.

Solution:

Corrected $\sum x = 300 - 18 - 12 + 10 + 20$

= 300

Corrected $\sum y = 210 - 20 - 10 + 15 + 15$

= 210

Corrected $\sum x^2 = 3718 - 18^2 - 12^2 + 10^2 + 20^2$

= 3750

Corrected $\sum y^2 = 2000 - 20^2 - 10^2 + 15^2 + 15^2$

= 1950

Corrected$\sum xy = 2100 - (18 \times 20) - (12 \times 100) + (10 \times 15) + (20 \times 15) =$ 2070

After correction the correlation coefficient is,

$$\gamma_{xy} = \frac{n\sum xy - \sum x \sum y}{[n\sum x^2 - (\sum x)^2]^{1/2}[n\sum y^2 - (\sum y)^2]^{1/2}}$$

$$\gamma_{xy} = \frac{30 \times 2070 - 300 \times 210}{[30 \times 3750 - 300^2]^{1/2}\left[[30 \times 1950 - 210^2]^{\frac{1}{2}}\right]}$$

$$= \frac{62100 - 63000}{(112500 - 9000)^{1/2}(58500 - 44100)^{1/2}}$$

$$= \frac{-900}{(22500)^{\frac{1}{2}}(14400)^{1/2}}$$

$$= - \frac{900}{150 \times 120}$$

$$= - \frac{1}{20}$$

$$= - 0.05$$

Problem: 4

$If\ x, y\ and\ z$ are uncorrelated variable each having same standard deviation obtain the coefficient of correlation between $x + y\ and\ y + z.$

Solution:

$Given\ \sigma_x = \sigma_y = \sigma_z = \sigma$.

$X\ and\ y$ are uncorrelated $\Rightarrow \sum(x - \bar{x})(y - \bar{y}) = 0$

$Y\ and\ z$ are uncorrelated $\Rightarrow \sum(y - \bar{y})(z - \bar{z}) = 0$

$Z\ and\ x$ are uncorrelated $\Rightarrow \sum(z - \bar{z})(x - \bar{x}) = 0$

Let$u = x + y\ and\ v = y + z$

$\bar{u} = \bar{x} + \bar{y}$ and $\bar{v} = \bar{y} + \bar{z}$

Now ,

$$\sigma_u^2 = \frac{1}{n}\sum(u - \bar{u})^2$$

$$= \frac{1}{n}\sum[(x - \bar{x}) + (y - \bar{y})]^2$$

$= \frac{1}{n} [\sum(x-\bar{x})^2 + \sum(y-\overline{y})^2 + 2\sum(x-\bar{x})(y-\bar{y})]$

$= \sigma_x^2 + \sigma_y^2$ [since $\sum(x-\bar{x})(y-\bar{y}) = 0$]

$= 2\sigma^2$

Similarly,

$$\sigma_v^2 = 2\sigma^2$$

Now, $\sum(u-\bar{u})(v-\bar{v}) = \sum[\{x-\bar{x}) + (y-\bar{y})\}\{(y-\bar{y}) + (z-\bar{z})\}]$

$= \sum(x-\bar{x})(y-\bar{y}) + \sum(y-\bar{y})^2 + \sum(x-\bar{x})(z-\bar{z}) + y{-}y(z{-}z)$

$= 0 + n\sigma_y^2 + 0 + 0$

$= n\sigma^2$

$$\gamma_{uv} = \frac{\sum(u-\bar{u})(v-\bar{v})}{n\sigma_u\sigma_v}$$

$$= \frac{n\sigma^2}{n(2\sigma^2)}$$

= ½.

Problem: 5

Show that the variable $u = x\cos\alpha + y\sin\alpha$ and $v = y\cos\alpha - x\sin\alpha$ are uncorrelated. If $\alpha = \frac{1}{2}\tan^{-1}\left(\frac{2\gamma_{xy}\sigma_x\sigma_y}{\sigma_x^2-\sigma_y^2}\right)$

Solution:

$$u_i = x_i\cos\alpha + y_i\sin\alpha \text{ and } v_i = y_i\cos\alpha - x_i\sin\alpha$$

$$\bar{u} = \bar{x}\cos\alpha + \bar{y}\sin\alpha \text{ and } \bar{v} = \bar{y}\cos\alpha - \bar{x}\sin\alpha$$

$$u_i - \bar{u} = (x_i - \bar{x})\cos\alpha + (y_i - \bar{y})\sin\alpha$$

The variable u_i and v_i are uncorrelated if $\sum(u_i-\bar{u})(v_i-\bar{v}) = 0$

$$\sum[(x_i-\bar{x})cos\alpha+(y_i-\bar{y})sin\alpha][(y_i-\bar{y})cos\alpha-(x_i-\bar{x})sin\alpha]=0$$

$$\therefore \sum(x_i-\overline{x})(y_i-\bar{y})\cos^2\alpha-\sum(x_i-\bar{x})(y_i-\bar{y})\sin^2\alpha-cos\alpha sin\alpha[\sum(x_i-\bar{x})^2-yi-y2]=0$$

$$\therefore n\gamma_{xy\sigma_x\sigma_y}(\cos^2\alpha-\sin^2\alpha)=\ ncos\alpha sin\alpha(\sigma_x^2-\sigma_y^2)$$

$$\therefore \gamma_{xy\sigma_x\sigma_y}\cos 2\alpha=\frac{1}{2}sin2\alpha(\sigma_x^2-\sigma_y^2).$$

$$\therefore tan2\alpha=\frac{2\gamma_{xy}\sigma_x\sigma_y}{\sigma_x^2-\sigma_y^2}$$

$$\therefore\ \alpha=\frac{1}{2}\tan^{-1}\left(\frac{2\gamma_{xy}\sigma_x\sigma_y}{\sigma_x^2-\sigma_y^2}\right)$$

Rank correlation:

Suppose that a group of n individuals are arranged in the order of merit or efficiency with respect to some characteristics. Then the rank is a variable which takes only the values 1,2,3…n. assuming that there is no tie.

Hence $\bar{x}=\frac{1+2\ldots\ldots\ldots\ldots+n}{n}=\frac{n+1}{2}$ and the variance is given by $\sigma_x^2=\frac{1}{12}(n^2-1)$.

Theorem:

Rank correlation ρ is given by

$\rho=1-\frac{6\sum(x-y)^2}{n(n^2-1)}$.

Proof:

Consider a collection of n individuals

let x_i and y_i be the ranks of the i^{th} individual in the two different rankings.

$$\therefore \bar{x}=\frac{1}{2}(n+1)=\bar{y}\ \text{and}\ \sigma_x^2=\frac{1}{12}(n^2-1)=\sigma_y^2$$

.

Now, $\sum(x-y)^2=\sum[(x-\bar{x})-(y-\bar{y})]^2(since\ \bar{x}=\bar{y})$

$=\sum(x-\bar{x})^2+\sum(y-\bar{y})^2-2\sum(x-\bar{x})(y-\bar{y})$

$=n\sigma_x^2+n\sigma_y^2-2n\rho\sigma_x\sigma_y$

$=2n\sigma_x^2(1-\rho)\quad(since\ \sigma_x^2=\sigma_y^2)$

$$=\frac{1}{6}\, n\,(n^2-1)(1-\rho)$$

$$1-\rho=\frac{6\sum(x-y)^2}{n(n^2-1)}$$

$$\rho=1-\frac{6\sum(x-y)^2}{n(n^2-1)}.$$

Problem: 6

Find the rank correlation coefficient between the height in cm and weight in kg of 6 soldiers in Indian Army.

Height	165	167	166	170	169	172
Weight	61	60	63.5	63	61.5	64

Solution:

Height	**Rank in height x**	**Weight**	**Rank in weight y**	**x-y**	$(x-y)^2$
165	6	61	5	1	1
167	4	60	6	-2	4
166	5	63.5	2	3	9
170	2	63	3	-1	1
169	3	61.5	4	-1	1
172	1	64	1	0	0
Total	-	-	-	-	-16

$$\rho=1-\frac{6\sum(x-y)^2}{n(n^2-1)}=1-\frac{6\times16}{6\times35}$$

= 1-0.457

= 0.543.

Problem: 7

From the following data of marks obtained by 10 students in physics and chemistry. Calculate the rank correlation coefficient.

Physics (P)	35	56	50	65	44	38	44	50	15	26

Chemistr y (Q)	50	35	70	25	35	58	75	60	55	35

Solution:

We rank the marks of physics and chemistry and we have the following table.

P	Rank in p x	Q	Rank in Q y	x-y	$(x-y)^2$
35	8	50	6	2	4
56	2	35	8	-6	36
50	3.5	70	2	1.5	2.25
65	1	25	10	-9	81
44	5.5	35	8	-2.5	6.25
38	7	58	4	3	9
44	5.5	75	1	4.5	20.25
50	3.5	60	3	0.5	0.25
15	10	55	5	5	25
26	9	35	8	1	1
Total	-	-	-	-	185

We observe that in the values of x the marks 50 and 44 occurs twice. In the values of y the mark 35 occurs thrice.

Hence in the calculation of the rank correlation coefficient $\sum(x-y)^2$ is to be corrected by adding the following correction factors

$$\left[\frac{2(2^2-1)}{12}+\frac{2(2^2-1)}{12}\right]+\frac{3(3^2-1)}{12}=3$$

After correction $\sum(x-y)^2 = 188$.

Now,

$$\rho = 1 - \frac{6\sum(x-y)^2}{n(n^2-1)}$$

$$= 1 - \frac{6\times188}{10\times99}$$

$$= 1 - \frac{1128}{990}$$

= 1-1.139

= -0.139.

Problem: 8

Three judges assign the ranks to 8 entries in a beauty contest.

Judge Mr.x	1	2	3	4	5	6	7	8
Judge Mr.y	3	2	1	5	4	7	6	8

Judge Mr.y	1	2	3	4	5	7	8	6

Which pair of Judges has the nearest approach to common in beauty?

Solution:

Table for the rank correlation coefficients $\rho_{xy}, \rho_{yz}, \rho_{zx}$

x	**y**	**Z**	**x-y**	$(x-y)^2$	**y-z**	$(y-z)^2$	**z-x**	$(z-x)^2$
1	3	1	-2	4	2	4	0	0
2	2	2	0	0	0	0	0	0
3	1	3	3	9	-2	4	-1	1
4	5	4	-2	4	1	1	1	1
5	4	5	3	9	-1	1	-2	4
6	7	7	-1	1	0	0	1	1
7	6	8	-1	1	-2	4	3	9
8	8	6	0	0	2	4	-2	4
Total			-	28	-	18	-	20

ρxy = 1- 6 ______ $n\ (\ nX2 - - y1\ )2$

$= 1 - \frac{6\times 28}{8\times(8^2-1)}$

$= 1 - \frac{168}{504}$

=1-0.333= 0.667.

$\rho_{yz} = 1 - \frac{6\times 18}{8\times 63}$

$= 1 - \frac{108}{504}$

= 1-

---

o.214 =

0.786.

ρ_{zx} = 1- $\frac{6\times20}{8\times63}$

= 1- $\frac{120}{504}$

= 1-0.238

= 0.762.

Since ρ_{yz} is greater than ρ_{xy} ρ_{xz} and the judges Mr. Y and Mr. Z have nearest approach to common taste in beauty.

Problem: 9

The coefficient of rank correlations of marks obtained by 10 students in mathematics and physics was found to be 0.8. It was later discovered that the differences in rank in two subjects obtained by one of the students was wrongly taken as 5 instead of 8. Find the correct coefficient of rank correlation.

Solution:

ρ_{xy} = 1- $\frac{6\sum(x-y)^2}{n(n^2-1)}$.

Given ρ_{xy} = 0.8 and n=10

0.8 =

- $\frac{6\sum(x-y)^2}{10(10^2-1)}$1

= 1- $\frac{6\sum(x-y)^2}{990}$

$\frac{6\sum(x-y)^2}{990}$ = 1-0.8 = 0.2

$6\sum(x-y)^2 = 990 \times 0.2$

= 198

$\therefore \sum(x-y)^2 = 33$

Corrected $\sum(x-y)^2 = 33 - 5^2 + 8^2 = 72$

Now, after correction ρ_{xy} = 1- $\frac{6\times72}{10(10^2-1)}$

= 1- $\frac{432}{990}$

= 1-0.486

= 0.564

The correct coefficient of rank correlation is 0.546.

Exercises:

1. Ten students got the following percentage of marks in two subjects.

Economics	78	65	36	98	25	75	82	90	62	39
Statistics	84	53	51	91	60	68	62	86	58	47

2. The following table shows how 10 students were ranked according to their achievements in the laboratory and lecture portions of a biology course. Find the coefficient of rank correlation.

Laboratory	8	3	9	2	7	10	4	6	1	5
Lecture	9	5	10	1	8	7	3	4	2	6

REGRESSION:

If there is a functional relationship between the two variable x_i and $y_i$the points in the scatter diagram will cluster around some curve called a line of regression. If the curve is a straight line it is called a line of regression between the two variables.

Definition:

It we fit a straight line by the principle of least squares to the points of the scatter diagram in such a way that the sum of the squares of the distance parallel to the y- axis from the points to the line is minimized we obtain a line of best fit for the data and its is called the regression line of y and x.

Similarly we can define the regression line of$x \ on \ y$.

Theorem:

The equation of the regression line of $y \ on \ x$ is given by $y - \bar{y} = \gamma \frac{\sigma_x}{\sigma_y} (x - \bar{x}$)

Proof:

Let $y = ax + b$ be the line of regression of on x.

According to the principle of least square the constants a and b are to be determined in such a way that S= $\sum[y_i - (ax_i + b)]^2 is \ minimum$

$\frac{ds}{da} = 0 \Rightarrow -2 \sum[(y_i - (ax_i + b)]x_i = 0$

$\Rightarrow \sum x_i y_i = a \sum x_i^2 + b \sum x_i$ ....(1)

$\frac{\partial s}{\partial b} = 0 \Rightarrow -2\left[\sum(y_i - (ax_i + b)\right] = 0$

$\Rightarrow \sum y_i = a\sum x_i + nb$ ……………(2)

Equations(1) and(2) are called normal equations.

From (2) we obtain $\bar{y} = a\bar{x} + b$………(3)

∴The line of regression passes through the point $(\bar{x}, \bar{y})$.

Now, shifting the origin to this point $(\bar{x}, \bar{y})$ by means of the transformation

$X_i = x_i - \bar{x}$ and $Y_i = y_i - \bar{y}$.

We obtain $\sum x_i = 0 = \sum y_i$ and the equation of the line of regression becomes

$y = ax$ …………(4)

Corresponding to this $line\ y = ax$ the constant a can be determined from the normal equation .

$a\sum X_i^2 = \sum x_i y_i$

a= $\frac{\sum X_i Y_i}{\sum X_i^2}$

$= \frac{\sum(x_i - \bar{x})(y_i - \bar{y})}{\sum(x_i - \bar{x})2}$

$= \frac{\gamma \sigma_x \sigma_y}{\sigma_x^2}$

$= \gamma \frac{\sigma_x}{\sigma_y}$

The required regression line (4) becomes $Y = \left(\gamma \frac{\sigma_y}{\sigma_x}\right) X$

$\therefore y - \bar{y} = \gamma \frac{\sigma_y}{\sigma_x} (x - \bar{x}).$

Definition:

The scope of the regression line of y on x is called the regression coefficient of y on x and it is denoted by b_{yx}.

Hence $b_{yx} = \gamma \frac{\sigma_y}{\sigma_x}$

The regression coefficient of x on y is given by $b_{xy} = \gamma \frac{\sigma_x}{\sigma_y}$.

Theorem:

Arithmetic mean of the regression coefficients is greater than or equal to the correlation coefficient.

Proof:

Let b_{xy} and b_{yx} be the regression coefficients.

we have to prove ½ $(b_{xy} + b_{yx}) \geq \gamma$

Now, ½ $(b_{xy} + b_{yx})) \geq \gamma$

$$\Leftrightarrow b_{yx} + b_{xy} \geq 2\gamma$$

$$\Leftrightarrow \gamma\frac{\sigma_y}{\sigma_x} + \gamma\frac{\sigma_x}{\sigma_y} \geq 2\gamma$$

$$\Leftrightarrow \sigma_x^2 + \sigma_y^2 - 2\sigma_x\sigma_y \geq 0$$

$$(\sigma_x - \sigma_y)^2 \geq 0.$$

This is always true.

Hence the theorem.

Theorem:

Regression coefficient are independent of the change of origin but dependent on change of scale.

Proof:

Let $u_i = \frac{x_i - A}{h}$ and $v_i = \frac{y_i - B}{k}$

Let $x_i = A + hu_i$ and $y_i = B + kv_i$

We know that, $\sigma_x = h\sigma_u, \sigma_y = k\sigma_v \; and \; \gamma_{xy} = \gamma_{uv}$

Now $b_{yx} = \gamma_{xy}\frac{\sigma_y}{\sigma_x}$

$= \gamma_{uv}\left(\frac{k\sigma_v}{h\sigma_u}\right)$

$= \frac{k}{h}\, b_{uv} \ldots\ldots\ldots\ldots\ldots\ldots .(1)$

similarly $b_{xy} = (\frac{h}{k})\, b_{uv}$..............(2)

From (1) and (2) $\Rightarrow b_{yx}$ and b_{xy} depend upon the scales h and k, but not on the origins A and B .

Hence the theorem.

Theorem:

The angle between two regression line is given by $\theta = \tan^{-1}\left[\left(\frac{1-\gamma^2}{\gamma}\right)\left(\frac{\sigma_x\sigma_y}{\sigma_x^2+\sigma_y^2}\right)\right]$

Proof:

The equations of lines of regression of y on x and x on y respectively are

$$y - \bar{y} = \gamma \frac{\sigma_y}{\sigma_x}(x - \bar{x}) \text{...............................}(1)$$

$$x - \bar{x} = \gamma \frac{\sigma_x}{\sigma_y}(y - \bar{y}) \dots\dots\dots\dots(2)$$

(2) can also be written as

$$y - \bar{y} = \frac{1}{\gamma}\frac{\sigma_y}{\sigma_x}(x - \bar{x}) \dots\dots\dots\dots(3)$$

Slopes of the two lines (1) and (2) are $\gamma \frac{\sigma_y}{x}$ and $\frac{\sigma_y}{\gamma \sigma_x}$.

Let θ be the acute angle between the two lines of regression.

$$\therefore tan\theta = \frac{\gamma \frac{\sigma_y}{\sigma_x} - \frac{\sigma_y}{\gamma \sigma_x}}{1 + \left(\gamma \frac{\sigma_y}{\sigma_x}\right)\left(\frac{y}{\gamma \sigma_x}\right)}$$

$$= \frac{\gamma^2 - 1}{\gamma}\left(\frac{\sigma_x \sigma_y}{\sigma_x^2 + \sigma_y^2}\right)$$

$$= \frac{1 - \gamma^2}{\gamma}\left(\frac{\sigma_x \sigma_y}{\sigma_x^2 + \sigma_y^2}\right)$$

(since $\gamma^2 \leq 1$ *and* θ *is acute*).

$$\therefore \theta = \tan^{-1}\left[\left(\frac{1 - \gamma^2}{\gamma}\right)\left(\frac{\sigma_x \sigma_y}{\sigma_x^2 + \sigma_y^2}\right)\right]$$

Problem: 10

The following data relate to the marks of 10 students in the internal test and the university examination for the maximum of 50 in each.

Internal marks	25	28	30	32	35	36	38	39	42	45

University marks	20	26	29	30	25	18	26	35	35	46

i) Obtain the two regression equations and determine.

ii) The most likely internal mark for the university mark of 25. iii) the most likely university mark for the internal mark of 30.

Solution:

(i) Let the marks of internal test and university examination be denoted by x and y respectively.

We have $\bar{x} = \frac{1}{10}\sum x_i = 35 \text{ and } \bar{y} = \frac{1}{10}\sum y_i = 29$.

For the calculation of regression we have the following table.

x_i	x_i-35	$(x_i - 35)^2$	y_i	y_i-29	$(y_i - 29)^2$	(x_i-35)($y_i - 29$)
25	-10	100	20	-9	81	90
28	-7	49	26	-3	9	21
30	-5	25	29	0	0	0
32	-3	9	30	1	1	-3
35	0	0	25	-4	16	0
36	1	1	18	-11	121	-11
38	3	9	26	-3	9	-9
39	4	16	35	6	36	24
42	7	49	35	6	36	42
45	10	100	46	17	289	170
Total	0	358	-	0	598	324

$\varsigma x2 =$ ______ ___ $xin - x$ 2 = 101 $xi - 35$ 2 = 35.8

$\sigma_y^2 = \frac{\sum(y_i-\bar{y})^2}{n} = \frac{1}{10}\sum(y_i - 29)^2 = 59.8$

$$\sigma_x = 5.98 \ and \ \sigma_y = 7.73$$

$$\therefore \gamma = \frac{\sum(x_i - \bar{x})(y_i - \bar{y})}{n\sigma_x\sigma_y}$$

$$= \frac{324}{10\times5.98 \times 7.73}$$

$$= \frac{324}{462.254}$$

= 0.7 (approximately)

Now the regression of y on x is $y - \bar{y} = \gamma\frac{\sigma_y}{\sigma_x}(x - \bar{x})$

$$\therefore \gamma\frac{\sigma_y}{\sigma_x} = \frac{\sum(x_i-\bar{x})(y_i-\bar{y})}{n\sigma_x^2}$$

$$= \frac{\sum(x_i-\bar{x})(y_i-\bar{y})}{\sum(x_i-\bar{x})^2}$$

$$= \frac{324}{358} = 0.905$$

Similarly, $\gamma\frac{\sigma_x}{\sigma_y} = \frac{\sum(x_i-\bar{x})(y_i-\bar{y})}{\sum(y_i-\bar{y})^2} = \frac{324}{598} = 0.542$

The regression line $of \ y \ on \ x$ is $y = 29 \ = \ 0.905\ (x - 35)$

(ie), $y = 0.905\,x - 2.675$ ....(1)

The regression line $of \ x \ on \ y \ is \ x - 35 \ = \ 0.542\ (y - 29)$

(ie),$x = 0.542y \ + \ 19.282$.............(2)

(1) and (2) are the required regression equations.

ii) the most likely internal mark for the university mark of 25 is got from the regression equation $of \ x \ on \ y$ by putting $y \ = \ 25$

(2)$\Rightarrow x = 0.542 \ \times 25 + 19.282 = 32.83$

iii) The most likely university mark for the internal mark of 30 is got from the regression equation $of \ y \ on \ x$ by putting$x \ = \ 30$

(1)$\Rightarrow y = 0.905 \ \times 30 - 2.675 = 24.475$

Problem: 11

The two variable x and y have the regression lines 3x+2y-26 = 0 and 6x+y-31=0. Find

i)The mean values of x and y
ii)The correlation coefficient between x and y iii)The variance of y if the variance of x is 25

Solution:

(i) Since the two lines of regression pass through ($\bar{x}\ y$,_)

we have $3\bar{x}+\bar{y}=26$......(1)

$6\bar{x}+\bar{y}=31$......(2)

Solving(1) and(2) we get $\bar{x} = 4 \; and \; \bar{y} = 67$

(ii) As in the previous problem we can prove that $y = \frac{-3}{2}x + 13 \; and \; x = \frac{-1}{6}y + \frac{31}{6}$ represent the regression lines of y on x and x on y by respectively.

Hence we get the regression coefficients $as \; b_{yx} = \frac{-3}{2}$

$$and \; b_{xy} = -1/6$$

Now,

$$\gamma^2 = \left(\frac{-3}{2}\right) \times \left(\frac{-1}{6}\right) = 1/4$$

$$\gamma = \pm\frac{1}{2}$$

Since both the regression coefficients are negative we take $\gamma = -\frac{1}{2}$

iii) Given $\sigma_x = 5$

We have $b_{yx} = \gamma \frac{\sigma_y}{\sigma_x}$

$$\therefore \frac{-3}{2} = \left(\frac{-1}{2}\right)\left(\frac{\sigma_y}{5}\right)$$

$$\sigma_y = 15 \,.$$

Problem: 12

If θ is the acute angle between the two regression lines.
Show that $\theta \leq 1 - \gamma^2$

Solution:

We know that if θ is the acute angle between the two regression on lines we have,

$\tan\theta = \left(\frac{1-\gamma^2}{\gamma}\right)\left(\frac{\sigma_x\sigma_y}{\sigma_x^2+\sigma_y^2}\right)$ ....(1)

We claim that $\sigma_x^2 + \sigma_y^2 \geq 2\sigma_x\sigma_y$.

Suppose not, then $\sigma_x^2 + \sigma_y^2 < 2\sigma_x\sigma_y$

(i.e.) $\sigma_x^2 + \sigma_y^2 - 2\sigma_x\sigma_y < 0$

$(\sigma_x - \sigma_y)^2 < 0$. This is impossible.

Hence

$\sigma_x^2 + \sigma_y^2 \geq 2\sigma_x\sigma_y$

$\therefore \frac{\sigma_x\sigma_y}{\sigma_x^2+\sigma_y^2} \leq \frac{1}{2}$.

$(1) \Rightarrow \tan\theta \leq \left(\frac{1-\gamma^2}{\gamma}\right)\left(\frac{1}{2}\right)$

$$\therefore \tan\theta \leq \left(\frac{1-\gamma^2}{2\gamma}\right)$$

Hence $\sin\theta \leq \left(\frac{1-\gamma^2}{1+\gamma^2}\right)$

$\sin\theta \leq 1-\gamma^2$.

Exercise:

1.calculate the coefficient of correlation of correlation and obtain the lines of regression for the following data.

X	1	2	3	4	5	6	7	8	9
Y	9	8	10	12	11	13	14	16	15

Correlation coefficient for a bivariate frequency distribution:

The correlation coefficient between x and y is given by $\gamma_{xy} = \frac{\cos(x,y)}{\sigma_x\sigma_y}$

$$\therefore \gamma_{xy} = \frac{\sum_{i=1}^{n}\sum_{j=1}^{m} f_{ij}x_iy_j - \frac{1}{N}\left(\sum_{i=1}^{n} g_ix_i\right)\left(\sum_{j=1}^{m} f_jy_j\right)}{\sqrt{\sum_{i=1}^{n} g_ix_i^2 - \frac{1}{N}\left(\sum_{i=1}^{n} g_ix_i\right)^2} \times \sqrt{\sum_{j=1}^{m} f_jy_j^2 - \frac{1}{N}\left(\sum_{j=1}^{m} f_jy_j\right)^2}}$$

Note: Since correlation coefficient is independent of origin and scale if x and y are transformed to u and v by the formula u = $\frac{x-A}{h}$ and v = $\frac{y-B}{k}$ then we *have* $\gamma_{xy} = \gamma_{uv}$.

Problem: 13

Find the correlation coefficient between x and y from the following table:

x	5	10	15	20
y				
4	2	4	5	4
6	5	3	6	2
8	3	8	2	3

Solution:

Y \ X		x1	x2	x3	x4	Total
		5	10	15	20	
y1	4	2	4	5	4	F1=15
y2	6	5	3	6	2	F2=16
y3	8	3	8	2	3	F3=16
Total		g1=10	g2=15	g3=13	g4=9	N=47

Correlation coefficient between x and y is given by

$$\therefore\ \gamma_{xy} = \frac{\sum_{i=1}^{n}\sum_{j=1}^{m} f_{ij}x_iy_j - \frac{1}{N}\left(\sum_{i=1}^{n} g_ix_i\right)\left(\sum_{j=1}^{m} f_jy_j\right)}{\sqrt{\sum_{i=1}^{n} g_ix_i^2 - \frac{1}{N}\left(\sum_{i=1}^{n} g_ix_i\right)^2} \times \sqrt{\sum_{j=1}^{m} f_jy_j^2 - \frac{1}{N}\left(\sum_{j=1}^{m} f_jy_j\right)^2}}$$

Where i=1,2,3,4 and j=1,2,3.

$$\sum g_ix_i = 50 + 150 + 195 + 180 = 575$$

$$\sum f_jy_j = 60 + 96 + 128 = 284$$

$$\sum g_ix_i^2 = 250 + 1500 + 2925 + 3600 = 8275$$

$$\sum f_jy_j^2 = 240 + 576 + 1024 = 1840$$

$$\sum\sum f_{ij}x_iy_j = (40 + 160 + 300 + 320) + (150 + 180 + 540 + 240) + (120 +$$

$640+240+480=3410$

$$\gamma_{xy} = \frac{3410 - \frac{1}{47}(575 \times 284)}{\sqrt{8275 - \frac{1}{47}(575)^2} \times \sqrt{1840 - \frac{1}{47}(284)^2}}$$

$$= \frac{3410 \times 47-(575\times284)}{\sqrt{8275\times47-575^2}\times\sqrt{1840\times47-284^2}}$$

$$= \frac{160270-163300}{\sqrt{388925-330625}\times\sqrt{86480-80656}}$$

$$= \frac{-3030}{\sqrt{58300}\times\sqrt{5824}} = \frac{-3030}{241.5\times76.3}$$

$$= \frac{-3030}{18426.5}$$

= -0.16

Problem: 14

Find the correlation coefficient between the heights and weight of 100 students which are distributed as follows.

Height in c.m	Weight in Kgs					Total
	30-40	40-50	50-60	60-70	70-80	
150-155	1	3	7	5	2	18
155-160	2	4	10	7	4	27
160-165	1	5	12	10	7	35
165-170	-	3	8	6	3	20
Total	4	15	37	28	16	100

Solution:

Let x_i denote the mid value of the classes of weights and y_j denote the mid value of the classes of heights.

$$Let\ u_i = \frac{x_i - 55}{10}\ and\ v_j = \frac{y_j - 157.5}{5}$$

Then the 2 -way frequency table is given below.

	35	45	55	65	75 (-4)	f_j	v_j	$f_j v_j$	$f_j v_j^2$	$f_{ij} u_i v_j$

152.5	(2) 1	(3) 3	(0) 7	(-5) 5	2	18	-1	-18	18	(-4)
157.5	(0) 2	(0) -4	(0) 10	(0) 7	(0) 4	27	0	0	0	(0)
162.5	(2) 1	(-5) 5	(0) 12	(10) 10	(14) 7	35	1	35	35	(17)
167.5	-	(-6) 3	(0) 8	(12) 6	(12) 3	20	2	40	80	(18)
g_i	4	15	37	28	16	100	-	57	133	(31)
u_i	-2	-1	0	1	2	-				
g_iu_i	-8	-15	0	28	32	37				
$g_iu_i^2$	16	15	0	28		123				
$f_{ij}u_iv_j$	(0)	(-8)	(0)	(17)	64 (22)	(31)				

$$\gamma_{xy} = \gamma_{uv} = \frac{\sum_{i=1}^{n}\sum_{j=1}^{m} f_{ij} u_i v_j - \frac{1}{N}(\sum_{i=1}^{n} g_i u_i)(\sum_{j=1}^{m} f_j v_j)}{\sqrt{\sum_{i=1}^{n} g_i u_i^2 - \frac{1}{N}(\sum_{i=1}^{n} g_i u_i)^2} \times \sqrt{\sum_{j=1}^{m} f_j v_j^2 - \frac{1}{N}(\sum_{j=1}^{m} f_j v_j)^2}}$$

$$= \frac{31 - \frac{1}{100}(37 \times 57)}{\sqrt{123 - \frac{1}{100}37^2} \times \sqrt{133 - \frac{1}{100}57^2}}$$

$$= \frac{3100 - 37 \times 57}{\sqrt{12300 - 37^2} \times \sqrt{13300 - 57^2}}$$

$$= \frac{991}{104.5 \times 100.25}$$

$$= 0.09.$$

Multiple and Partial Correlation

I. With only two predictors

A. The beta weights can be computed as follows:

B.

$$\boxed{\beta_{Y1.2} = \frac{r_{Y1} - r_{Y2}\, r_{12}}{1 - r_{12}^2}} \quad (1)$$

$$\boxed{\beta_{Y2.1} = \frac{r_{Y2} - r_{Y1}\, r_{12}}{1 - r_{12}^2}} \quad (2)$$

A. Multiple R can be computed several ways. From the simple correlations, as

B.

$$\boxed{R_{Y.12} = \sqrt{\frac{r_{Y1}^2 + r_{Y2}^2 - 2 r_{Y1} r_{Y2} r_{12}}{1 - r_{12}^2}}} \quad (3)$$

or from the beta weights and validities as

$$\boxed{R_{Y.12} = \sqrt{\beta_{Y1.2}\, r_{Y1} \beta_{Y2.1}\, r_{Y2}}} \quad (4)$$

A. SemipR^2ial correlations in general equal the square root of R^2 complete minus reduced. These are called semipartial correlations because th e variance of the other controlled variable(s) is removed from the predictor, but not from the criterion. Therefore, in the two predictor case, they are equal

B.
$$r^2_{Y(1.2)} = R^2_{Y.12} - r^2_{Y2} \tag{5}$$

Using Equation 3 above and some algebra

$$r^2_{Y(1.2)} = \frac{(r_{Y1} - r_{Y2}\, r_{12})^2}{1 - r^2_{12}} \tag{6}$$

$$r^2_{Y(1.2)} = \frac{r_{Y1} - r_{Y2}\, r_{12}}{\sqrt{1 - r_{12}}} \tag{7}$$

$$r^2_{Y(2.1)} = \frac{r_{Y2} - r_{Y1}\, r_{12}}{\sqrt{1 - r_{12}}} \tag{8}$$

So the relationship between the semipartial correlation and the beta weight from Equations 1 and 7 is

$$r_{Y(1.2)} = \beta_{Y1.2}\sqrt{1 - r^2_{12}} \tag{9}$$

A. Partial correlations differ from semipartial correlations in that the partialled (or covaried) variance is removed from both the criterion and the predictor. The squared partial correlation is equal to R^2 complete minus R^2 reduced divided by 1 minus R^2 reduced. In the two variable case the equation is

B.
$$r^2_{Y1.2} = \frac{R^2_{Y.12} - r^2_{Y2}}{1 - r^2_{Y2}} \tag{10}$$

Again using Equation 3 and some more algebra

$$r^2_{Y1.2} = \frac{(r_{Y1} - r_{Y2}\, r_{12})^2}{(1 - r^2_{12})(1 - r^2_{Y2})} \tag{11}$$

$$r_{Y1.2} = \frac{r_{Y1} - r_{Y2}\, r_{12}}{\sqrt{(1 - r^2_{12})(1 - r^2_{Y2})}} \tag{12}$$

$$r_{Y2.1} = \frac{r_{Y2} - r1_{Y1}\, r_{12}}{\sqrt{(1 - r^2_{12})(1 - r^2_{Y1})}} \tag{13}$$

The relation between partial correlations and beta weights for the two predictor problem turns out to be

$$r_{Y1.2} = \sqrt{\beta_{Y1.2}\,\beta_{1Y.2}} \tag{14}$$

So semipartial correlations are directional but partial correlations are nondirectional.

The semipartial correlations are:

$$r^2_{Y(1.2)} = R^2_{Y.12} - r^2_{Y2} = a$$

$$r^2_{Y(2.1)} = R^2_{Y.12} - r^2_{Y1} = b$$

And the partial correlations are:

$$r^2_{Y1.2} = \frac{R^2_{Y.12} - r^2_{Y2}}{1 - r^2_{Y2}} = \frac{a}{a+e}$$

$$r^2_{Y2.1} = \frac{R^2_{Y.12} - r^2_{Y1}}{1 - r^2_{Y1}} = \frac{b}{b+e}$$

II. With more than two predictors

A. First the relation between a multiple R and various partial r's.

B.

$$R^2_{Y.123..P} = 1 - (1 - r^2_{Y2.1})(1 - r^2_{Y3.12}) \ldots (1 - r^2_{YP.123..}) \qquad (15)$$

This should remind the reader of stepwise multiple regression where each new variable is entered while controlling the variance explained by earlier entered variables. Therefore, if we could compute the higher order partial correlations, we could do multiple regression by hand. A recurrence relationship allows us to do just that, which is

$$r_{Y1.23..P} = \frac{r_{Y1.23..P-1} - (r_{YP.23..P-1})(r_{1P.23..P-1})}{\sqrt{(1 - r_{YP.23..P-1})(1 - r_{1P.23..P-1})}} \qquad (16)$$

Unfortunately, the work involved in solving all the necessary partial correlations is about the same as the work required to solve the normal equations in the first place, but at least each step is interpretable. Again in the general case the rel ation between partial correlations and beta weights is

$$r_{Y1.23..P} = \sqrt{\beta_{1.23..P}\,\beta_{1Y.23..P}} \qquad (17)$$

STANDARD DISTRIBUTIONS

Introduction:

In this chapter we discuss some important distribution of random variable which are frequently used is statistics. We make a detailed study of binomial distribution, Poisson distribution which are of discrete type and normal distribution which is of continuous type.

Binomial Distribution:

Definition:

Let n be any positive integer and let $0 < p < 1,\ Let\ q = 1 - p$

Define $p(x) = \begin{cases} n_{c_x} p^x q^{n-x} & if\ x = 0,1,2, \dots .n \\ 0 & otherwise \end{cases}$

A discrete random variable with the above p.d.f. is said to have binomial distribution and the p.d.f itself is called a binomial distribution.

Note: 1

The two independent constants n and p in the distributions are known as the parameters of the distribution. If x is a binomial variate with parameters n and p we write as

$$X \sim B(n, P).$$

Note:2

In this experiment is repeated N times (say) then the frequency function of the binomial distribution is given $by\ f(x) = NP(x) = N n_{c_x} p^x q^{n-x}, x = 0,1,2, \dots \dots ..n$.

Theorem: 1

m.g.f of a binomial distribution about the origin is

$(q + pe^t)^n$.

Proof:

$$M_X(t) = E(e^{Xt}) = \sum e^{Xt} p(x) = \sum_{x=0}^{n} \left(e^{tx} n_{c_x} p^x q^{n-x}\right)$$

$$= \sum_{x=0}^{n} n_{c_x} (pe^t)^x q^{n-x}$$

$$= (q + pe^t)^n.$$

Moments of binomial distribution:

We know that for any random variable x the m.g.f is $M_X(t) = 1 + \mu_1' + \frac{\mu_2'}{2!}t^2 + \dots + \frac{\mu_r'}{r!}t^r + \cdots \dots$

Theorem: 2 (Addition property of binomial distribution)

If $X_1 \sim B(n_1, p), X_2 \sim B(n_2, p)$ are independent random variable then $X_1 + X_2$ is $B(n_1 + n_2, p)$.

Proof:

Given X_1, and X_2 are independent random variable with parameters $n_1, p \; and \; n_2, p$ respectively.

Let us consider m.g.f of X_1 and X_2 about origin.

$\therefore M_{X_1} = (q + pe^t)^{n_1} M_{X_2}(t) = (q + pe^t)^{n_2}$

Now, $M_{X_1 + X_2}(t) = M_{X_1}(t) + M_{X_2}(t)$ (since X_1, and X_2 are independent)

$$= (q + pe^t)^{n_1} + (q + pe^t)^{n_2}$$

$$= (q + pe^t)^{n_1 + n_2}$$

= m.g.f of the binomial $X_1 + X_2$ with parameters $n_1 + n_2$ and P.

Hence the uniqueness theorem $X_1 + X_2$ is a binomial variable with parameters

$n_1 + n_2$ and P.

Theorem: 3

Characteristic function of binomial distribution is $(q + pe^{it})^n$.

Proof:

Let $X \sim B(n, P)$. Hence $\varphi_x(t) = E(e^{itx})$

$\sum_{x=0}^{n} e^{itx} p(x) = \sum_{x=0}^{n}$ = $e^{itx} n_{c_x} p^x q^{n-x}$

$\sum_{x=0}^{n} n_{c_x} (pe^{it})^x q^{n-x}$
=

$(q + Pe^{it})^n$
= .

Mode of Binomial Distribution:

Let $X \sim B(n,P).\ then\ P(x) = n_{c_x} p^x q^{n-x}$.Let x be the mode of the binomial distribution.

Example:

The unbiased coins are tossed and number of heads noted. The experiment is repeated 64 times and the following distribution is obtained.

No.of heads	0	1	2	3	4	5	Total
Frequencies	3	6	24	26	4	1	64

Solution:

Here n=5 and N=64

Since the coins are unbiased p=1/2 =q. So that p/q =1

Now $p(0)=q^n=(1/2)^5 =1/32$

Hence $f(0)=N\, q^n =64\times1/32 = 2$

Using the recurrence formula $p(x+1) = \left(\frac{n-x}{x+1}\right)\left(\frac{p}{q}\right)p(x)$

P(1)= 5(1/32).

Hence f(1)=10

x	Probabilities $p(x)$	Expected frequencies $f(x) = Np(x)$	Observed frequencies
0	p(0)=1/32	2	3
1	P(1)=5/32	10	6
2	P(2)=10/32	20	24
3	P(3)=10/32	20	26
4	P(4)=5/32	10	4
5	P(5)=1/32	2	1
Total		64	64

Problem:

In a binomial distribution the mean is 4 and the variance is 8/3.Find the mode of the distribution.

Solution:

Given mean =4 and Variance=8/3

$\therefore np = 4$ and npq=8/3

$\frac{npq}{np} = \frac{8}{3\times4} = \frac{2}{3}$

q = 2/3,p=1/3

$\therefore np = 4$

n=12

Consider (n+1)p=13/3=4.3

Hence the mode is 4.

Problem:

A discrete random variable X has the mean 6 and variance 2. If it is assumed that the distribution is binomial. Find the probability that $5 \leq x \leq 7$.

Solution:

Given np=6 and npq=2

Hence q=1/3 and p=2/3

Also n=9

Now, p($5 \leq x \leq 7$)=p(x=5)+p(x=6)+p(x=7)

$$=9C_5\left(\frac{2}{3}\right)^5\left(\frac{1}{3}\right)^4 + 9C_6\left(\frac{2}{3}\right)^6\left(\frac{1}{3}\right)^3 + 9C_7\left(\frac{2}{3}\right)^7\left(\frac{1}{3}\right)^2$$

$$= \frac{126}{3^9} x 2^5 + 84 \times \frac{2^6}{3^9} + 36 \times \frac{2^7}{3^9}$$

$$= \frac{2^5}{3^9}[126+168+144]$$

$$= \frac{2^5}{3^9} \text{ x } 438$$

$$= \frac{32\times438}{19683}$$

= 0.712

Problem :

An insurance agent accepts policies of 5 men all of identity age and in good health. The probability that man of this age will be alive 30 years hence is 2/3. Find the probability that is 30 years (i) all five men (ii) at least one man (ii) almost three will be alive.

Solution:

Here n=5, p=2/3, q=1-p=1-2/3=1/3

It is a binomial distribution and x~B (5,2/3) Hence

$P(X=x) = nC_x P^x q^{n-x}$

Probability of all 5 will be alive is

$P(x=5) = 5C_5\,(2/3)^5 = 32/243$

ii) Probability of at least one being alive = 1- probability of no one being alive. Probability of no one being alive

$P(x=0) = 5C_0\,(1/3)^5 = 1/243$

Probability of at least one being alive

= 1-

1/243 =

242/24

3

iii) Probability of almost 3 being alive = probability of one man being alive or probability of 2 men being alive (or) probability of 3 men being a live.

= 1- [Prob. Of 4 men being a live (or) probability of 5 men being alive]

$P(x \leq 3) = 1 - P(X > 3)$

= 1- [P(x=4) + P(x=5)]

$=1-[5C_4\,(2/3)^4\,(1/3) + 5C_5\,(2/3)^5]$

= 1-[5(16/243)+ 32/243]

= 1-112/243 = 131/243

Problem :

Six dice are thrown 729 times. How many times do you except at least 3 dice to show a five or six.

Solution:

Here n = 6, N= 729

P=prob. Of getting 5 or 6 with one dice = 2/6 = 1/3

Q = 1-1/3 = 2/3

The expected frequency of 0,1,2,..6. successor are the successive terms of

$729\ (1/3 + 2/3)^6$

Excepted number of times at least 3 dice showing five or six

$= 729\ (6C_3\ (1/3)^3 + 6C_4(1/3)^4\ (2/3)^2 + 6C_5\ (1/3)^5\ 2/3 + 6C_6\ (1/3)^6]$

$= 729/3^6 \times 233$

$= 169857/729$

$=233.$

Problem :

If the m.g.f. of a r.v.x is of the form $M_x(t) = (0.4e^t+0.6)^8$ find i) E(x)

ii) the m.g.f of the r.v Y= 3X+2

Solution:

We have the m.g.f of the binomial variable X ~ B(n,p) is $(q+pe^t)^n$

Here p=0.4, q=0.6 and n=8

i) We have E(X) = $\mu_1 = np$

E(X) = 8x0.4

= 3.2

ii) $M_y\ (t) = M_{3X+2}\ (t)$

$= e^{2t}\ M_x\ (3x)$

$= e^{\ 2t}\ (0.6+0.4e^{\ 3t})^{\ 8}$

Poisson distribution:

We have the binomial distribution is determined by 2 parameters p and n. If the number of trails is in definitely large and the probability p of success is in definitely small such that np=λ ,where λ is a constant then the limiting case of the binomial distribution when n $\rightarrow \infty$ *and* $p \rightarrow 0$ becomes a distribution known as Poisson distribution

Definition:

A discrete random variable X is said to be follow a Poisson distribution if it assumes only nonnegative integer values and its probability density function is given by

$$p(x) = p(X = x) = f(x) = \begin{cases} \dfrac{e^{-x}\lambda^x}{x!}, & if\ x = 0,1,2 \ldots \\ 0, & otherwise \end{cases}$$

Where $\lambda$is the parameter of the distribution if X is a Poisson variate with parameter λ.we write X~$p(\lambda)$

Example:

If fit a Poisson distribution to the following data

x	0	1	2	3	4	Total
f	123	9	14	3	1	200

Solution:

To fit a Poisson distribution we have to calculate all the excepted frequencies

Here the mean $\lambda = \frac{\Sigma f_i x_i}{\Sigma f_i}$

$= \frac{59+28+9+4}{200} = 0.5$

$\lambda = 0.5$

Hence the p.d.f of poisson distribution is

$$p(x) = \frac{e^{-0.5}(0.5)^x}{x!}$$

$\therefore p(0) = e^{-0.5}$ = 0.6065

Hence f(0)=Np(0)

=200$\times$ (0.6065)

x	Probabilities using p(x+1)=(λ/(x+1))p(x)	Expected frequencies f(x)=Np(x)	Observed frequencies
0	P(0)=0.6065	121.3	123

=1213

Using recurrence formula

VP(x+1)= $\left(\frac{\lambda}{(x+1)}\right) p(x)$

We have p(1) = $\frac{0.5}{p(0)}$

=0.3033

=60.66

1	P(1)=0.3033	60.66	59
2	P(2)=0.0758	15.16	14
3	P(3)=0.0126	2.52	3
4	P(4)=0.0027	0.54	1

Problem :

The probabilities of a Poisson variable taking the values 3 and 4 are equal calculate the probabilities of variates taking the values 0 and 2

Solution

Take x be a Poisson variate with parameter λ

∴P(X=x) = $\frac{e^{-\lambda}\lambda^x}{x!}$

Given P(X=3)=P(X=4)

To find P(X=0) and P(X=1)

P(X=3)=P(X=4) $\Rightarrow \frac{e^{-\lambda}\lambda^3}{3!}$

$=\frac{e^{-\lambda}\lambda^4}{4!} \Rightarrow \lambda=4$

P(X=0)= $\frac{e^{-4}4^0}{0!} = e^{-4} = 0.0183$

P(X=2)= $\frac{e^{-4}4^2}{2!} = 0.146.$

Problem :

Assuming that one in 80 births in a case twins. Calculate the probability of 2 or more birth of twins on a day when 30 births occur using (i) binomial distribution (ii) Poisson approximation.

Solution:

Assuming X to be a binomial variate with p=probability of twin births=1/80=0.125.

(q=0.9875) where n=30 we get

$$p(x) = 30_{c_x}(0.0125)^x(0.9875)^{30-x}$$

Probability of 2 or more births of twins on a day is

$$p(X \geq 2) = 1 - p(X < 2)$$

=1-[p(X=0)+p(X=1)]

=1-[$(0.9875)^{30}$+30(0.0125)$(0.9875)^{29}$

=1-0.6943(1.3625)

=0.054

Assuming X to be a Poisson variate with λ=np=0.375

We get $p(x) = \frac{e^{-\lambda}\lambda^x}{x!} = \frac{e^{-0.375}0.375^x}{x!}$

$$p(X \geq 2) = 1 - [p(X=0) + p(X=1)]$$

=1-$e^{-0.375}$+$e^{-0.375}$(0.375)

=1-0.6873(1.375)

=0.0550.

NORMAL DISTRIBUTION:

Normal distribution is one of the most widely used distribution in application of statistical methods.

We have $\int_{-\infty}^{\infty} e^{-y^2/2}\, dy = \sqrt{2\pi}$

Hence $\int_{-\infty}^{\infty} \frac{e^{-y^2/2}}{\sqrt{2\pi}}\, dy$ =1

Put $y = \frac{x-\mu}{\sigma}$ then we have $\int_{-\infty}^{\infty} \frac{e^{-(x-\mu)^2/2\sigma^2}}{\sigma\sqrt{2\pi}}\, dx = 1$

Definition:

A continuous random variable X is said to follow a normal distribution if its probability density function is given by f(x)= $\frac{e^{-(x-\mu)^2/2\sigma^2}}{\sigma\sqrt{2\pi}}$ where $-\infty < x < \infty$

μ *and* σ are constants and σ>0 and are called the parameters of the distribution and we write $X \sim N(\mu, \sigma^2)$.

Fitting of normal distribution:

To fit a normal distribution to given date we first calculate the mean μ *and* $s.d$ σ.Thus the normal curve fitted to the given data is given by

f(x)= $\frac{e^{-(x-\mu)^2/2\sigma^2}}{\sigma\sqrt{2\pi}}$ where $-\infty < x < \infty$

Properties of normal distribution:

1. The normal probability curve is symmetrical about the ordinate at $x = \mu$.The ordinate decreases rapidly as x increases. The curve extends to infinity on either side of the mean. The X-axis is an asymptote to the curve.

2. The mean, median and mode coincide the maximum ordinate at $x = \mu$ is given by $\frac{1}{\sigma\sqrt{2\pi}}$.

3. $\mu \pm \sigma$ are the points of inflexion of the normal curve and hence the points of inflexion are also equidistant from the median.

4.The area under normal curve is unity. The ordinate at $x = \mu$divides the area under the normal curve into two equal parts symmetry also ensures that the first and third quartiles of normal distribution are equidistant from the median of course on either side.

5. $p(\mu - \sigma < x < \mu + \sigma) = 0.6826$

   $p(\mu - 2\sigma < x < \mu + 2\sigma) = 0.9544$

   $p(\mu - 3\sigma < x < \mu + 3\sigma) = 0.9973$

6. Q.D:M.D:S.D=10:12:15.

Problem :

If X is normal distributed with zero mean and unit variance.

Find the expectation of X^2.

Solution:

Given $X \sim N(0,1)$

Hence the normal is $f(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$

Now $E(X^2) = \int_{-\infty}^{\infty} x^2 f(x) dx$

$= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} x^2 e^{-x^2/2} dx$

$= \frac{1}{\sqrt{2\pi}} [-xe^{-x^2/2}]_{-\infty}^{\infty} + \int_{-\infty}^{\infty} e^{-x^2/2} dx$

$= \frac{1}{\sqrt{2\pi}} [0 + \sqrt{2\pi}]$

$=1.$

Problem :

If X is normally distributed with mean 8 and s.d 4. Find

(i)P($5 \le X \le 10)$ $(ii) P(10 \le X \le 15)$

(iii) P($X \ge 5)$ (iv) $P(|X-5| \le 15$.

Solution:

Given

$X \sim N(8,4)$

Hence the standard normal variate $Z = \frac{X-8}{4}$

When X=5;z=-0.75

When X=10;Z=0.50

When X=15;Z=1.75

When X=20 ;Z=3

(i) P($5 \le X \le 10) = P(-0.75 \le Z \le 0.50)$

=P(-0.75$\le Z \le 0) + P(0 \le Z \le 0.50)$

=P(0$\le Z \le 0.75) + P(0 \le Z \le 0.50)$

=0.2734+0.1915

=0.4649.

$(ii) P(10 \leq X \leq 15)$=P(0.5$\leq Z \leq 1.75)$

=P(0$\leq Z \leq 1.75) - P(0 \leq Z \leq 0.5)$

= 0.4599 – 1.1915 =

0.2684 iii) P (x≥5) = p(Z

≥1.75)

= 0.5 – p ($0 \leq z \leq 1.75)$

= 0.5 – 0.4599

= 0.0401

iv) p(x $\leq 5)$ = p(z ≤-0.75)

= 0.5 – p(-0.75 $\leq z \leq 0)$

= 0.5 – p(0$\leq z \leq 0.75)$

= 0.5 – 0.2734

= 0.2266

Problem :

The marks of 1000 students in a university are found to be normally distributed with mean to and s.d. 5. Estimate the number of students whose marks will be i) between 60 and 75 (ii) more than 75 (iii) less than 68.

Solution :

Let x denote the marks of students .

Hence X~N (70, 25)

The standard normal variate is

$$Z = \frac{X-\mu}{\sigma} = \frac{X-70}{5}$$

i) To find p(60 < x < 75)

When x = 60; z= -2 and when x=75; z=1

P (60 < x < 75) = p (-2 < z < 1)

= p (-2 <z < 0) + p (0 < z < 1)

= p (0 < z < 2) + p (0 < z < 1)

= 0.4772 + 0.3413

= 0.8185

∴ The number of students whose marks is between 60 and 75 is 1000 x 0.8185 = 819.

(ii) To find p (x> 75)

When x = 75; z =1

∴p (x > 75) = p(z > 1)

= 0.5 – p (0 <z < 1)

= 0.5 – 0.3413

= 0.1587

∴ The number of students whose marks is more than 75 is 159.

(iii)To find p (x < 68)

When x = 68, $z = \frac{68-70}{5} = -0.4$

∴p (x < 68) = p (z < -0.4)

= p(z > 0.4)

= 0.5 – p(0 < z < 0.4)

= 0.5 – 0.1554 = 0.3546.

Problem :

Assume the mean height of soldiers to be 68.22 inches with variance of 10.8 inches. How many soldiers in a regiment of 2000 soldiers would you expect to be over six feet tall. Assume heights to be normally distributed.

Solution :

Let the variable x denote the height in inches of the solders.

Mean $\mu = 68.22$; $\sigma^2 = 10.8, \sigma = 3.286$

Hence X ~N (68.22, 3.286)

∴p(x > 6 feet) = p (x > 72)

When x = 72; $z = \frac{x-\mu}{\sigma}$

$= \frac{72-68.22}{3.286}$

$= \frac{3.78}{3.286}$

= 1.15

$\therefore$p (x > 72) = p(z > 1.15)

= 0.5 – p (0 $\leq z \leq 1.15$)

= 0.5 – 0.3749

= 0.1251

$\therefore$ The number of soldiers in the regimet of 2000 over 6 feet tall is 2000 x 0.1251= 250

Problem :

A set of examination marks is approximately distributed with mean 75 and S.D. of 5. If the top 5% of students get grade A and the bottom 25% get grade B what mark is the lowest A and what mark is the highest B?

Solution :

It x denote the marks in the examination.

Given x is normally distributed with mean $\mu = 75$ and $\varsigma = 5$

ie) X ~N (75, 25)

Let x_1 be the lowest marks for A and x_2 be the highest marks for B. Given p(x >x_1) = 0.05 and p(x <x_2) = 0.25

The standard normal variate

$$z = \frac{x_{1-\mu}}{\sigma}$$

$$= \frac{x_1 - 75}{5} = z_1$$

$$Z = \frac{x_2 - \mu}{\sigma} = \frac{x_2 - 75}{5} = -z_2 \qquad \ldots\ldots\ldots\ldots (1)$$

P(0 < z <z_1) = 0.45

z_1 = 0.45

P(-$z_2 < z < 0$) = 0.25

P(0 < z <z_2) = 0.25

z_2 = 0.675

(1)$\Rightarrow$ x_1 = 75 + 5z_1

x_1 = 83.225

$x_1 \approx 83$

x_2= 75 – 5z_2

$x_2 = 71.625$

$x_2 \approx 72$

Hence the lowest mark for grade A is 83 and the highest mark for B is 72.

Problem :

In a normal distribution 31% of the items are under 45 and 8% are over 64. Find the mean and standard deviation.

Solution :

Let x denote the normal variate with mean μ *and* $S.D.\varsigma$

Given p(x < 45) = 0.31 and p (x > 64) = 0.08

When x = 45, $z = \frac{45-\mu}{\sigma} = -z_1$

When x = 64, $z = \frac{45-\mu}{\sigma} = z_2$

P(0 < z < z_2) = 0.42 and p($-z_1 < z < 0$)=0.19 From the area table we get

$z_1 = 0.496$ *and* $z_2 = 1.405$

(1)⇒ $45 - \mu = 0.496\ \varsigma$

$64 - \mu = 1.405\ \varsigma$

$\varsigma = 9.99 \approx 10$

$\mu = 49.96 \approx 50$

Problem :

Find the probability of getting between 3 heads to 6 heads in 10 tosses of a fair coin using (i) binomial distribution (ii) the normal approximation to the binomial distribution.

Solution :

(i). Take x as the binomial variate,

P = $\frac{1}{2}$, $q = \frac{1}{2}, n = 10$

We have X $\sim B\ (10, \frac{1}{2})$

Probability of getting at least 3 heads

= p (x $\geq 3)$

= p(x=3) + p(x=4) + p(x=5) + p(x=6)

$$= 10c_3\left(\frac{1}{2}\right)^3\left(\frac{1}{2}\right)^7 + 10c_4\left(\frac{1}{2}\right)^4\left(\frac{1}{2}\right)^6 + 10c_5\left(\frac{1}{2}\right)^5\left(\frac{1}{2}\right)^5 + 10\,c_6\left(\frac{1}{2}\right)^6\left(\frac{1}{2}\right)^4$$

$$= \frac{1}{2^{10}}[120 + 210 + 252 + 210]$$

$$= \frac{792}{2^{10}}$$

= 0.7734

(ii). Taking the data as continuous it follows that 3 to 6 heads can be considered as 2.5 to 6.5 heads.

Mean $\mu = np$

$$= 10\left(\frac{1}{2}\right)$$

= 5

$$\sigma = \sqrt{npq} = 1.58$$

$$\therefore\ X \sim N\,(5, 1.58)$$

The standard normal variate is $z = \frac{X-\mu}{\sigma}$

For X = 2.5 ; $z = \frac{2.5-5}{1.58}$

= -1.58

X = 6.5 ; $z = \frac{6.5-5}{1.58}$

Z = 0.95

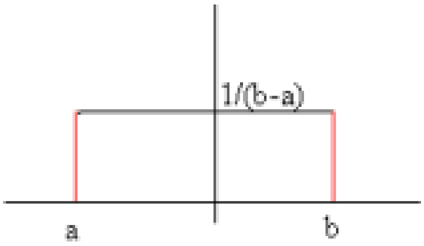
The Uniform Distribution

This page covers Uniform Distribution, Expectation and Variance, Proof of Expectation and Cumulative Distribution Function.

A continuous random variable X which has probability density function given by:

f(x) = $\frac{1}{b - a}$ for a £ x £ b

(and f(x) = 0 if x is not between a and b) follows a **uniform distribution** with parameters a and b. We write X ~ U(a,b)



Remember that the area under the graph of the random variable must be equal to 1 (see continuous random variables).

Expectation and Variance

If X ~ U(a,b), then:

- $E(X) = \frac{1}{2}(a + b)$
- $Var(X) = (1/12)(b - a)^2$

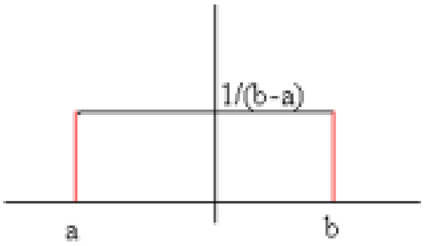
Proof of Expectation

$$\int_a^b \frac{x}{b-a}\,dx = \frac{1}{b-a}\left[\frac{x^2}{2}\right]_a^b = \frac{b^2 - a^2}{2(b-a)} = \frac{b+a}{2}$$

Cumulative Distribution Function

The cumulative distribution function can be found by integrating the p.d.f between 0 and t:

$$F(t) = \int_a^t \frac{1}{b-a}\,dx = \frac{t-a}{b-a}$$



Remember that the area under the graph of the random variable must be equal to 1 (see continuous random variables).

Expectation and Variance

If X ~ U(a,b), then:

- $E(X) = \frac{1}{2}(a + b)$
- $Var(X) = (1/12)(b - a)^2$

Proof of Expectation

$$\int_a^b \frac{x}{b-a}\,dx = \frac{1}{b-a}\left[\frac{x^2}{2}\right]_a^b = \frac{b^2 - a^2}{2(b-a)} = \frac{b+a}{2}$$

Cumulative Distribution Function

The cumulative distribution function can be found by integrating the p.d.f between 0 and t:

$$F(t) = \int_{a}^{t} \frac{1}{b-a}\,dx = \frac{t-a}{b-a}$$

The Geometric Distribution

Geometric distribution - A discrete random variable X is said to have a geometric distribution if it has a probability density function (p.d.f.) of the form:

- $P(X = x) = q^{(x-1)}p$, where $q = 1 - p$

If X has a geometric distribution with parameter p, we write X ~ Geo(p)

Expectation and Variance

If X ~ Geo(p), then:

- $E(X) = 1/p$
- $Var(X) = q/p^2$, where $q = 1 – p$

Gamma Distribution

The gamma distribution is another widely used distribution. Its importance is largely due to its relation to exponential and normal distributions. Here, we will provide an introduction to the gamma distribution. In Chapters 6 and 11, we will discuss more properties of the gamma random variables. Before introducing the gamma random variable, we need to introduce the gamma function.

Gamma function: The gamma function [10], shown by $\Gamma(x)\Gamma(x)$, is an extension of the factorial function to real (and complex) numbers. Specifically, if $n\in\{1,2,3,...\}n\in\{1,2,3,\ldots\}$, then

$$\Gamma(n)=(n-1)!$$

Note that if $\alpha=n$, where n is a positive integer, the above equation reduces to

$$n!=n\cdot(n-1)!$$

Properties of the gamma function

For any positive real number α:

1. $\Gamma(\alpha)=\int_0^\infty x^{\alpha-1}e^{-x}dx$;
2. $\int_0^\infty x^{\alpha-1}e^{-\lambda x}dx=\frac{\Gamma(\alpha)}{\lambda^\alpha}, \text{for } \lambda>0$;
3. $\Gamma(\alpha+1)=\alpha\Gamma(\alpha)$;
4. $\Gamma(n)=(n-1)!, \text{ for } n=1,2,3,\cdots$;
5. $\Gamma(\frac{1}{2})=\sqrt{\pi}$.

Example
Answer the following questions:

1. Find $\Gamma(\frac{7}{2})$.
2. Find the value of the following integral:

$$I=\int_0^\infty x^6e^{-5x}dx.$$

- **Solution**

---

for some values of α and λ.

Example
Using the properties of the gamma function, show that the gamma PDF integrates to 1, i.e., show that for $\alpha,\lambda>0$, we have

$$\int_0^\infty \frac{\lambda^\alpha x^{\alpha-1}e^{-\lambda x}}{\Gamma(\alpha)}dx=1.$$

- **Solution**

A gamma distribution is a general type of statistical distribution that is related to the beta distribution and arises naturally in processes for which the waiting times between Poisson distributed events are relevant. Gamma distributions have two free parameters, labeled α and θ, a few of which are illustrated above.

Consider the distribution function $D(x)$ of waiting times until the hth Poisson event given a Poisson distribution with a rate of change λ,

$$\begin{aligned} D(x) &= P(X \le x) && (1)\\ &= 1 - P(X > x) && (2)\\ &= 1 - \sum_{k=0}^{h-1} \frac{(\lambda x)^k e^{-\lambda x}}{k!} && (3)\\ &= 1 - e^{-\lambda x} \sum_{k=0}^{h-1} \frac{(\lambda x)^k}{k!} && (4)\\ &= 1 - \frac{\Gamma(h, x\lambda)}{\Gamma(h)} && (5) \end{aligned}$$

for $x \in [0, \infty)$, where $\Gamma(x)$ is a complete gamma function, and $\Gamma(a, x)$ an incomplete gamma function. With h an integer, this distribution is a special case known as the Erlang distribution.

The corresponding probability function $P(x)$ of waiting times until the hth Poisson event is then obtained by differentiating $D(x)$,

$$\begin{aligned} P(x) &= D'(x) && (6)\\ &= \lambda e^{-\lambda x} \sum_{k=0}^{h-1} \frac{(\lambda x)^k}{k!} - e^{-\lambda x} \sum_{k=0}^{h-1} \frac{k(\lambda x)^{k-1}\lambda}{k!} && (7)\\ &= \lambda e^{-\lambda x} + \lambda e^{-\lambda x} \sum_{k=1}^{h-1} \frac{(\lambda x)^k}{k!} - e^{-\lambda x} \sum_{k=1}^{h-1} \frac{k(\lambda x)^{k-1}\lambda}{k!} && (8)\\ &= \lambda e^{-\lambda x} - \lambda e^{-\lambda x} \sum_{k=1}^{h-1} \left[\frac{k(\lambda x)^{k-1}}{k!} - \frac{(\lambda x)^k}{k!}\right] && (9)\\ &= \lambda e^{-\lambda x} \left\{1 - \sum_{k=1}^{h-1} \left[\frac{(\lambda x)^{k-1}}{(k-1)!} - \frac{(\lambda x)^k}{k!}\right]\right\} && (10)\\ &= \lambda e^{-\lambda x} \left\{1 - \left[1 - \frac{(\lambda x)^{h-1}}{(h-1)!}\right]\right\} && (11)\\ &= \frac{\lambda(\lambda x)^{h-1}}{(h-1)!} e^{-\lambda x}. && (12) \end{aligned}$$

Now let $\alpha \equiv h$ (not necessarily an integer) and define $\theta \equiv 1/\lambda$ to be the time between changes. Then the above equation can be written

$$P(x) = \frac{x^{\alpha-1} e^{-x/\theta}}{\Gamma(\alpha)\theta^{\alpha}} \qquad (13)$$

---

for $x \in [0, \infty)$. This is the probability function for the gamma distribution, and the corresponding distribution function is

$$D(x) = P\left(\alpha, \frac{x}{\theta}\right), \tag{14}$$

where $P(a, z)$ is a regularized gamma function.

It is implemented in the Wolfram Language as the function `GammaDistribution`[*alpha*, *theta*].

The characteristic function describing this distribution is

$$\phi(t) = \mathcal{F}_x \left\{ \frac{x^{-x/\theta}\, x^{\alpha-1}}{\Gamma(\alpha)\,\theta^\alpha} \left[\tfrac{1}{2}(1 + \operatorname{sgn} x)\right] \right\}(t) \tag{15}$$

$$= (1 - i\, t\, \theta)^{-\alpha}, \tag{16}$$

where $\mathcal{F}_x[f](t)$ is the Fourier transform with parameters $a = b = 1$, and the moment-generating function is

$$M(t) = \int_0^\infty \frac{e^{t x}\, x^{\alpha-1}\, e^{-x/\theta}\, d x}{\Gamma(\alpha)\,\theta^\alpha} \tag{17}$$

$$= \int_0^\infty \frac{x^{\alpha-1}\, e^{-(1-\theta t)x/\theta}\, d x}{\Gamma(\alpha)\,\theta^\alpha}. \tag{18}$$

giving moments about 0 of

$$\mu_r' = \frac{\theta^r\, \Gamma(\alpha + r)}{\Gamma(\alpha)} \tag{19}$$

(Papoulis 1984, p. 147).

In order to explicitly find the moments of the distribution using the moment-generating function, let

$$y \equiv \frac{(1 - \theta t)x}{\theta} \tag{20}$$

$$d y = \frac{1 - \theta t}{\theta}\, d x, \tag{21}$$

so

$$M(t) = \int_0^\infty \left(\frac{\theta y}{1 - \theta t}\right)^{\alpha-1} \frac{e^{-y}}{\Gamma(\alpha)\,\theta^\alpha}\, \frac{\theta\, d y}{1 - \theta t} \tag{22}$$

$$= \frac{1}{(1 - \theta t)^\alpha\, \Gamma(\alpha)} \int_0^\infty y^{\alpha-1}\, e^{-y}\, d y \tag{23}$$

$$= \frac{1}{(1 - \theta t)^\alpha}, \tag{24}$$

giving the logarithmic moment-generating function as

$$R(t) = -\alpha \ln(1 - \theta t) \tag{25}$$

$$R'(t) = \frac{\alpha\theta}{1 - \theta t} \tag{26}$$

$$R''(t) = \frac{\alpha\theta^2}{(1-\theta t)^2}. \tag{27}$$

The mean, variance, skewness, and kurtosis excess are then

$$\mu = \alpha\theta \tag{28}$$
$$\sigma^2 = \alpha\theta^2 \tag{29}$$
$$\gamma_1 = \frac{2}{\sqrt{\alpha}} \tag{30}$$
$$\gamma_2 = \frac{6}{\alpha}. \tag{31}$$

The gamma distribution is closely related to other statistical distributions. If X_1, X_2, ..., X_n are independent random variates with a gamma distribution having parameters (α_1, θ), (α_2, θ), ..., (α_n, θ), then $\sum_{i=1}^{n} X_i$ is distributed as gamma with parameters

$$\alpha = \sum_{i=1}^{n} \alpha_i \tag{32}$$
$$\theta = \theta. \tag{33}$$

Also, if X_1 and X_2 are independent random variates with a gamma distribution having parameters (α_1, θ) and (α_2, θ), then $X_1/(X_1 + X_2)$ is a beta distribution variate with parameters (α_1, α_2). Both can be derived as follows.

$$P(x_1, x_2) = \frac{1}{\Gamma(\alpha_1)\Gamma(\alpha_2)} e^{x_1+x_2} x_1^{\alpha_1-1} x_2^{\alpha_2-1}. \tag{34}$$

Let

$$u = x_1 + x_2 \quad x_1 = u\,v \tag{35}$$
$$v = \frac{x_1}{x_1 + x_2} \quad x_2 = u(1-v), \tag{36}$$

then the Jacobian is

$$J\left(\frac{x_1, x_2}{u, v}\right) = \begin{vmatrix} v & u \\ 1-v & -u \end{vmatrix} = -u, \tag{37}$$

so

$$g(u, v)\,d\,u\,d\,v = f(x, y)\,d\,x\,d\,y = f(x, y)\,u\,d\,u\,d\,v. \tag{38}$$
$$g(u, v) = \frac{u}{\Gamma(\alpha_1)\Gamma(\alpha_2)} e^{-u} (u\,v)^{\alpha_1-1} u^{\alpha_2-1} (1-v)^{\alpha_2-1} \tag{39}$$
$$= \frac{1}{\Gamma(\alpha_1)\Gamma(\alpha_2)} e^{-u} u^{\alpha_1+\alpha_2-1} v^{\alpha_1-1} (1-v)^{\alpha_2-1}. \tag{40}$$

The sum $X_1 + X_2$ therefore has the distribution

$$f(u) = f(x_1 + x_2) = \int_0^1 g(u, v)\, d v = \frac{e^{-u}\, u^{\alpha_1 + \alpha_2 - 1}}{\Gamma(\alpha_1 + \alpha_2)}, \tag{41}$$

which is a gamma distribution, and the ratio $X_1 / (X_1 + X_2)$ has the distribution

$$h(v) = h\left(\frac{x_1}{x_1 + x_2}\right) \tag{42}$$

$$= \int_0^\infty g(u, v)\, d u \tag{43}$$

$$= \frac{v^{\alpha_1 - 1}\, (1 - v)^{\alpha_2 - 1}}{B(\alpha_1, \alpha_2)}, \tag{44}$$

where B is the beta function, which is a beta distribution.

If X and Y are gamma variates with parameters α_1 and α_2, the X / Y is a variate with a beta prime distribution with parameters α_1 and α_2. Let

$$u = x + y \quad v = \frac{x}{y}, \tag{45}$$

then the Jacobian is

$$J\left(\frac{u, v}{x, y}\right) = \begin{vmatrix} 1 & 1 \\ \frac{1}{y} & -\frac{x}{y^2} \end{vmatrix} = -\frac{x + y}{y^2} = -\frac{(1 + v)^2}{u}, \tag{46}$$

so

$$d x\, d y = \frac{u}{(1 + v)^2}\, d u\, d v \tag{47}$$

$$g(u, v) = \frac{1}{\Gamma(\alpha_1)\,\Gamma(\alpha_2)}\, e^{-u} \left(\frac{u\, v}{1 + v}\right)^{\alpha_1 - 1} \left(\frac{u}{1 + v}\right)^{\alpha_2 - 1} \frac{u}{(1 + v)^2} \tag{48}$$

$$= \frac{1}{\Gamma(\alpha_1)\,\Gamma(\alpha_2)}\, e^{-u}\, u^{\alpha_1 + \alpha_2 - 1}\, v^{\alpha_1 - 1}\, (1 + v)^{-\alpha_1 - \alpha_2}. \tag{49}$$

The ratio X / Y therefore has the distribution

$$h(v) = \int_0^\infty g(u, v)\, d u = \frac{v^{\alpha_1 - 1}\, (1 + v)^{-\alpha_1 - \alpha_2}}{B(\alpha_1, \alpha_2)}, \tag{50}$$

which is a beta prime distribution with parameters (α_1, α_2).

The "standard form" of the gamma distribution is given by letting $y \equiv x / \theta$, so $d y = d x / \theta$ and

---

$$P(y)\, d y = \frac{x^{\alpha - 1}\, e^{-x/\theta}}{\Gamma(\alpha)\, \theta^\alpha}\, d x \tag{51}$$

$$= \frac{(\theta y)^{\alpha-1} e^{-y}}{\Gamma(\alpha)\,\theta^{\alpha}} (\theta\, d\, y) \tag{52}$$
$$= \frac{y^{\alpha-1} e^{-y}}{\Gamma(\alpha)}\, d\, y, \tag{53}$$

so the moments about 0 are

$$\nu_r = \frac{1}{\Gamma(\alpha)} \int_0^{\infty} e^{-x}\, x^{\alpha-1+r}\, d\, x \tag{54}$$
$$= \frac{\Gamma(\alpha+r)}{\Gamma(\alpha)} \tag{55}$$
$$= (\alpha)_r, \tag{56}$$

where $(\alpha)_r$ is the Pochhammer symbol. The moments about $\mu = \mu_1$ are then

$$\mu_1 = \alpha \tag{57}$$
$$\mu_2 = \alpha \tag{58}$$
$$\mu_3 = 2\,\alpha \tag{59}$$
$$\mu_4 = 3\,\alpha^2 + 6\,\alpha. \tag{60}$$

The moment-generating function is

$$M(t) = \frac{1}{(1-t)^{\alpha}}, \tag{61}$$

and the cumulant-generating function is

$$K(t) = \alpha \ln(1-t) = \alpha \left(t + \tfrac{1}{2} t^2 + \tfrac{1}{3} t^3 + \ldots\right), \tag{62}$$

so the cumulants are

$$\kappa_r = \alpha \Gamma(r). \tag{63}$$

If X is a normal variate with mean μ and standard deviation σ, then

Exponential Distribution

In probability theory, the exponential distribution is defined as the probability distribution of time between events in the Poisson point process. The exponential distribution is considered as a special case of the gamma distribution. Also, the exponential distribution is the continuous analogue of the geometric distribution. In this article, we will discuss what is exponential distribution, its formula, mean, variance, memoryless property of exponential distribution, and solved examples.

Table of Contents:

What is Exponential Distribution?

In Probability theory and statistics, the exponential distribution is a continuous probability distribution that often concerns the amount of time until some specific event happens. It is a process in which events happen continuously and independently at a constant average rate. The exponential distribution has the key property of being memoryless. The exponential random variable can be either more small values or fewer larger variables. For example, the amount of money spent by the customer on one trip to the supermarket follows an exponential distribution.

Exponential Distribution Formula

The continuous random variable, say X is said to have an exponential distribution, if it has the following probability density function:

$$f_X(x|\lambda)=\begin{cases}\lambda e^{-\lambda x} & \text{for } x>0\\ 0 & \text{for } x\le 0\end{cases}$$

Where

λ is called the distribution rate.

Mean and Variance of Exponential Distribution

Mean:

The mean of the exponential distribution is calculated using the integration by parts.

Mean = E[X] = $\int_0^\infty x\lambda e^{-\lambda x}dx$

$=\lambda\left[\left|\frac{-xe^{-\lambda x}}{\lambda}\right|_0^\infty+\frac{1}{\lambda}\int_0^\infty e^{-\lambda x}dx\right]$

$=\lambda\left[0+\frac{1}{\lambda}\frac{-e^{-\lambda x}}{\lambda}\right]_0^\infty$

$=\lambda\frac{1}{\lambda^2}$

$=\frac{1}{\lambda}$

Hence, the mean of the exponential distribution is 1/λ.

Variance:

To find the variance of the exponential distribution, we need to find the second moment of the exponential distribution, and it is given by:

$$E[X^2]=\int_0^\infty x^2\lambda e^{-\lambda x}=\frac{2}{\lambda^2}$$

Hence, the variance of the continuous random variable, X is calculated as:

Var (X) = E(X^2)- E(X)^2

Now, substituting the value of mean and the second moment of the exponential distribution, we get,

Var (X) $=2\lambda_2-1\lambda_2=1\lambda_2$
Thus, the variance of the exponential distribution is $1/\lambda^2$.

Memoryless Property of Exponential Distribution

The most important property of the exponential distribution is the memoryless property. This property is also applicable to the geometric distribution.

An exponentially distributed random variable “X” obeys the relation:

$P_r(X > s+t \mid X>s) = P_r(X>t)$, for all s, t ≥ 0

Now, let us consider the the complementary cumulative distribution function:

$P_r(X > s+t \mid X>s)$ $=P_r(X>s+t\cap X>s)P_r(X>s)$

$=P_r(X>s+t)P_r(X>s)$

$=e^{-\lambda(s+t)}e^{-\lambda s}$
$= e^{-\lambda t}$

$= P_r\ (X>t)$

Hence, $P_r(X > s+t \mid X>s) = P_r(X>t)$

This property is called the memoryless property of the exponential distribution, as we don’t need to remember when the process has started.

Sum of Two Independent Exponential Random Variables

The probability distribution function of the two independent random variables is the sum of the individual probability distribution functions.

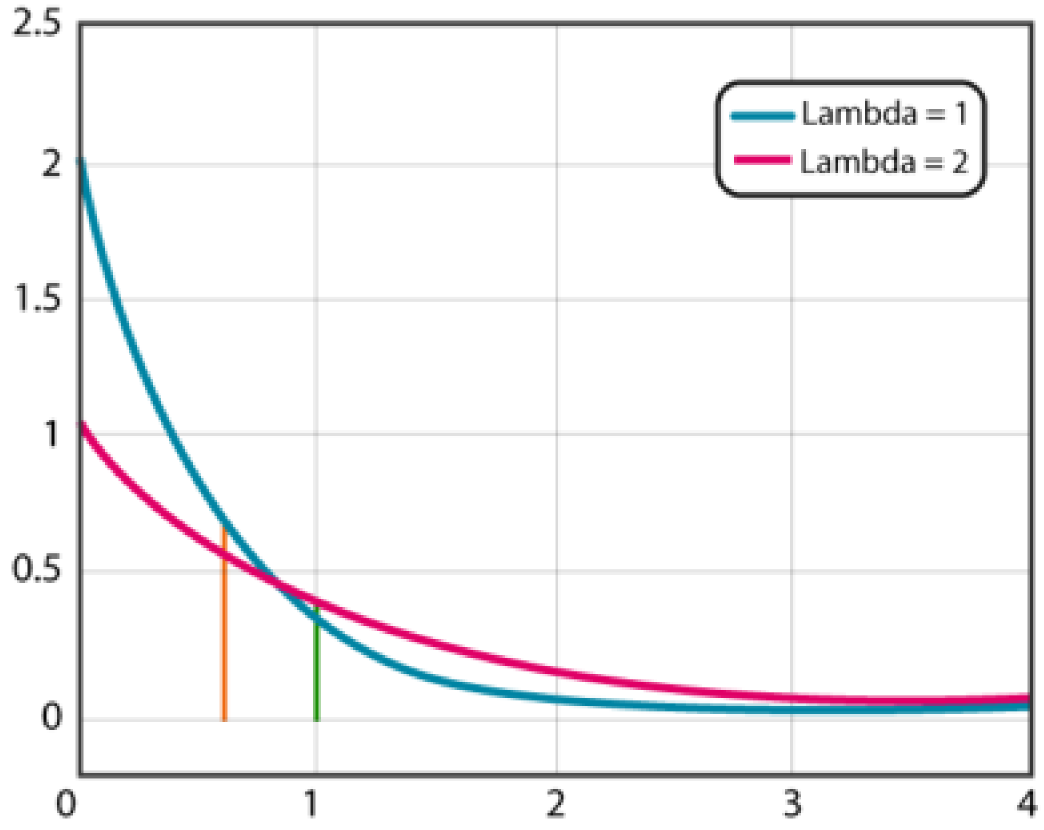
If X_1 and X_2 are the two independent exponential random variables with respect to the rate parameters λ1 and λ2 respectively, then the sum of two independent exponential random variables is given by Z = X_1 + X_2.

$fzz=\int_{\infty-\infty} f_{X_1}(x1)f_{X_2}(z-x1)dx1$
$=\int_{z0}\lambda1e^{-\lambda_1x_1}\lambda2e^{-\lambda_2(z-x_1)}dx1$
$=\lambda1\lambda2e^{-\lambda_2z}\int_{z0}e^{(\lambda_2-\lambda_1)x_1}dx1$
$=\{_{\lambda_1\lambda_2\lambda_2-\lambda_1}(e^{-\lambda_1z}-e^{-\lambda_2z})\lambda2ze^{-\lambda z}\text{if } \lambda1\neq\lambda2\text{if } \lambda1=\lambda2=\lambda$

- Exponential Distribution Formula
- Exponential Distribution Calculator
- Poisson Distribution Formula

Exponential Distribution Graph

The exponential distribution graph is a graph of the probability density function which shows the distribution of distance or time taken between events. The two terms used in the exponential distribution graph is lambda (**λ**)and x. Here, lambda represents the events per unit time and x represents the time. The following graph shows the values for λ=1 and λ=2.



Exponential Distribution Applications

One of the widely used continuous distribution is the exponential distribution. It helps to determine the time elapsed between the events. It is used in a range of applications such as reliability theory, queuing theory, physics and so on. Some of the fields that are modelled by the exponential distribution are as follows:

- Exponential distribution helps to find the distance between mutations on a DNA strand
- Calculating the time until the radioactive particle decays.
- Helps on finding the height of different molecules in a gas at the stable temperature and pressure in a uniform gravitational field
- Helps to compute the monthly and annual highest values of regular rainfall and river outflow volumes

Exponential Distribution Problem

Example:

Assume that, you usually get 2 phone calls per hour. calculate the probability, that a phone call will come within the next hour.

Solution:

It is given that, 2 phone calls per hour. So, it would expect that one phone call at every half-an-hour. So, we can take

λ = 0.5

So, the computation is as follows:

$p(0\leq X\leq 1)=\sum_{x=0}^{1} 0.5e^{-0.5x}$
= 0.393469

Therefore, the probability of arriving the phone calls within the next hour is 0.393469

Stay tuned with BYJU'S – The Learning App and download the app to learn with ease by exploring more Maths-related videos.

Frequently Asked Questions on Exponential Distribution

What is meant by exponential distribution?

The exponential distribution is a probability distribution function that is commonly used to measure the expected time for an event to happen.

What is the difference between the Poisson distribution and exponential distribution?

Poisson distribution deals with the number of occurrences of events in a fixed period of time, whereas the exponential distribution is a continuous probability distribution that often concerns the amount of time until some specific event happens.

What is the mean and the variance of the exponential distribution?

The mean of the exponential distribution is 1/λ and the variance of the exponential distribution is $1/\lambda^2$.

Why is the exponential distribution memoryless?

The key property of the exponential distribution is memoryless as the past has no impact on its future behaviour, and each instant is like the starting of the new random period.

What does lambda mean in the exponential distribution?

The lambda in exponential distribution represents the rate parameter, and it defines the mean number of events in an interval.

Beta functions are a special type of function, which is also known as Euler integral of the first kind. It is usually expressed as B(x, y) where x and y are real numbers greater than 0. It is also a symmetric function, such as B(x, y) = B(y, x). In Mathematics, there is a term known as special functions. Some functions exist as solutions of integrals or differential equations.

What are the Functions?

Functions play a vital role in Mathematics. It is defined as a special association between the set of input and output values in which each input value correlates one single output value. We know that there are two types of Euler integral functions. One is a beta function, and another one is a gamma function. The domain, range or codomain of functions depends on its type. In this page, we are going to discuss the definition, formulas, properties, and examples of beta functions.

Example:

Consider a function $f(x) = x^2$ where inputs (domain) and outputs (co-domain) are all real numbers. Also, all the pairs in the form (x, x^2) lie on its graph.

Let's say if 2 be input; then we would get an output as 4, and it is written as f(2) = 4. It is said to have an ordered pair (2, 4).

Beta Function Definition

The beta function is a unique function where it is classified as the **first kind of Euler's integral**. The beta function is defined in the domains of real numbers. The notation to represent the beta function is "**β**". The beta function is meant by B(p, q), where the parameters p and q should be real numbers.

The beta function in Mathematics explains the association between the set of inputs and the outputs. Each input value the beta function is strongly associated with one output value. The beta function plays a major role in many mathematical operations.

Beta Function Formula

The beta function formula is defined as follows:

$$B(p,q)=\int_0^1 t^{p-1}(1-t)^{q-1}\,dt$$

Where p, q > 0

The beta function plays a major role in calculus as it has a close connection with the gamma function, which itself works as the generalisation of the factorial function. In calculus, many complex integral functions are reduced into the normal integrals involving the beta function.

Relation with Gamma Function

The given beta function can be written in the form of gamma function as follows:

$$B(p,q)=\frac{\Gamma p.\Gamma q}{\Gamma(p+q)}$$

Where the gamma function is defined as:

$$\Gamma(x)=\int_0^\infty t^{x-1}e^{-t}\,dt$$

Also, the beta function can be calculated using the factorial formula:

$$B(p,q)=\frac{(p-1)!(q-1)!}{(p+q-1)!}$$

Where, p! = p. (p-1). (p-2)… 3. 2. 1

Beta Function Properties

The important properties of beta function are as follows:

- This function is symmetric which means that the value of beta function is irrespective to the order of its parameters, i.e B(p, q) = B(q, p)
- B(p, q) = B(p, q+1) + B(p+1, q)
- B(p, q+1) = B(p, q). [q/(p+q)]
- B(p+1, q) = B(p, q). [p/(p+q)]
- B (p, q). B (p+q, 1-q) = π/ p sin (πq)
- The important integrals of beta functions are:
    - $B(p,q)=\int_0^\infty \frac{t^{p-1}}{(1+t)^{p+q}}dt$
    - $B(p,q)=2\int_0^{\pi/2} \sin^{2p-1}\theta \cos^{2q-1}\theta \, d\theta$

Incomplete Beta Functions

The generalized form of beta function is called incomplete beta function. It is given by the relation:

$$B(z:a,b)=\int_0^z t^{a-1}(1-t)^{b-1}dt$$

It is also denoted by $B_z(a, b)$. We may notice that when z = 1, the incomplete beta function becomes the beta function. i.e. B(1 : a, b) = B(a, b). The incomplete beta function has many implementations in physics, functional analysis, integral calculus etc.

Beta Function Examples

Question: Evaluate: $\int_0^1 t^4(1-t)^3dt$
Solution:

$$\int_0^1 t^4(1-t)^3dt$$

The above form can also be written as:

$$\int_0^1 t^{5-1}(1-t)^{4-1}dt$$

Now, compare the above form with the standard beta function: $B(p,q)=\int_0^1 t^{p-1}(1-t)^{q-1}dt$
So, we get p= 5 and q = 4

Using the factorial form of beta function: $B(p,q)=\frac{(p-1)!(q-1)!}{(p+q-1)!}$, we get
B (p, q) = (4!. 3!) / 8!

= (4!. 6) /8! = 1/ 280

Therefore, the value of the given expression using beta function is 1/ 280

Beta Function Applications

In Physics and string approach, the beta function is used to compute and represent the scattering amplitude for Regge trajectories. Apart from these, you will find many applications in calculus using its related gamma function also.

Hypergeometric distribution is a random variable of a hypergeometric probability distribution. Using the formula of you can find out almost all statistical measures such as mean, standard deviation, variance etc.

$$P(x|N,m,n) = \frac{\left(\binom{m}{x}\binom{N-m}{n-x}\right)}{\binom{N}{n}}$$

Where,

N: The number of items in the population.
n: The number of items in the sample.
x: The number of items in the sample that are classified as successes.
P(x| N, n, k): hypergeometric probability – the probability that an n-trial hypergeometric experiment results in exactly x successes, when the population consists of N items, k of which are classified as successes.

Solved Examples

Question 1: Calculate the probability density function of the hypergeometric function if N, n and m are 50, 10 and 5 respectively ?

Solution:
Given parameters are,
N = 50
n = 10
m = 5

Formula for hypergeometric distribution is,

P(x|N,m,n) = $$\frac{\binom{m}{x}\binom{N-m}{n-x}}{\binom{N}{n}}$$

P(x|N,m,n) = $$\frac{\binom{5}{x}\binom{50-5}{10-x}}{\binom{50}{10}}$$

So, the probability distribution function is,

P(x|50, 5, 10) = $$\frac{\binom{5}{x}\binom{50-5}{10-x}}{\binom{50}{10}}, [0 \leq x \leq 10]$$


$f_X(x)$