Definition: For a set S with finite cardinality, we define the real valued function P on 2^S as $P(E) = \frac{|E|}{|S|}$ $\forall$ $E \subseteq S$ and refer to it as the discrete symmetric probability on S

THEOREM 1:

The discrete symmetric probability on S satisfies

(i) $P(S) = 1$

(ii) $P(E^c) = 1 - P(E)$

(iii) $\forall E \subseteq S$, $0 \leq P(E) \leq 1$

(iv) For $E_1, E_2 \subseteq S$, $E_1 \cap E_2 = \emptyset$ ~~[illegible]~~

$\Rightarrow P(E_1 \cup E_2) = P(E_1) + P(E_2)$

Definition: An experiment is called a random experiment if

(i) all possible outcomes are known in advance

(ii) any performance results in an outcome which cannot be known beforehand

(iii) the experiment can be repeated under identical and ideal conditions

Definition: A non empty collection of subsets S is called a σ-field and denoted $\mathcal{A}$ if it satisfies :

(i) $\phi \in \mathcal{A}$

(ii) $E \in \mathcal{A} \Rightarrow E^c \in \mathcal{A}$

(iii) $E_n \in \mathcal{A} \ \forall n \in \mathbb{N} \Rightarrow \bigcup_{n=1}^{\infty} E_n \in \mathcal{A}$

THEOREM 2:

A sigma field $\mathcal{A}$ associated with S satisfies

(i) $E_1, E_2 \in \mathcal{A} \Rightarrow E_1 \cup E_2 \in \mathcal{A}$

(ii) $S \in \mathcal{A}$

(iii) $E_n \in \mathcal{A} \ \forall n \in \mathbb{N} \Rightarrow \bigcap_{n=1}^{\infty} E_n \in \mathcal{A}$

Definition: A sample space of a random experiment is a pair $(S, \mathcal{A})$ where S is the set of all possible outcomes and $\mathcal{A}$ is the σ-field of [redacted] subsets of S

Definition: Let $(S, \mathcal{A})$ be a sample space. A real valued set map $\mathbb{P} : \mathcal{A} \to \mathbb{R}$ is a probability measure if

(i) $\mathbb{P}(E) \geq 0 \quad \forall E \in \mathcal{A}$

(ii) $\mathbb{P}(S) = 1$

(iii) For $E_1, E_2, E_3, \ldots$ mutually disjoint sets in $\mathcal{A}$, we have $\mathbb{P}\left(\bigcup_{n=1}^{\infty} E_n\right) = \sum_{n=1}^{\infty} \mathbb{P}(E_n)$

We further call $(S, \mathcal{A}, \mathbb{P})$ as probability space

Definition: For S being finite, $(S, \mathcal{A})$ is a sample space for $\mathcal{A} = 2^S =: \mathcal{A}_*$ and we define the symmetric probability measure $\mathbb{P}(E) = \frac{|E|}{|S|}$ for $E \subseteq S$

Definition: For $S = \{s_1, s_2, \cdots\}$, a countable set, we still consider $\mathcal{A}_* = 2^S$ and define the probability measure $\mathbb{P}(E) = \sum_{s_i \in E} p_i \quad \forall E \in \mathcal{A}_*$ where p_i is a number associated to $s_i \quad \forall i \in \mathbb{N}$ such that $p_i \geq 0 \ \forall i \in \mathbb{N}$, $\sum_{i=0}^{\infty} p_i = 1$.

We call this the discrete probability space

THEOREM 3:

Intersection of σ-fields is a σ-field.

Definition: Let C be some fixed collection of subsets of S. We define

$$\mathcal{S}_C = \{ \mathcal{A}_t \mid t \in \Lambda,\ \mathcal{A}_t \text{ is a } \sigma\text{-field containing } C \text{ i.e. } C \subseteq \mathcal{A}_t \}$$

where $\Lambda \subset \mathbb{R}$ is an indexing set.

The generated σ-field $\sigma(C)$ is then defined as $\sigma(C) = \bigcap_{t \in \Lambda} \mathcal{A}_t$

THEOREM 4:

$\sigma(\mathcal{C})$ is the smallest σ-field containing $\mathcal{C}$

Definition: Let $I = \{ (a,b] \mid 0 \leq a < b \leq 1 \}$. Then $\sigma(I)$ is called the Borel-σfield and is denoted $\mathcal{B}([0,1])$. Any $B \in \mathcal{B}([0,1])$ is called a borel set.

THEOREM 5:

A set is measurable iff it is a borel set from $\mathcal{B}(\mathbb{R})$

THEOREM 6:

$([0,1], \mathcal{B}_0, \hat{\mathbb{P}})$ is NOT a probability space where,

$$\mathcal{B}_0 = \{ \bigcup_{k \in \Lambda} I_k \mid \Lambda \subset \mathbb{N} \text{ is finite and } I_k \in I = \{(a,b] \mid 0 \leq a < b \leq 1\} \text{ are disjoint} \}$$

$$\hat{\mathbb{P}}\left(\bigcup_{k \in \Lambda} I_k \right) = \sum_{k \in \Lambda} |I_k|$$

However,

$$\sigma(\mathcal{B}_0) = \mathcal{B}([0,1]) \quad \text{(non-trivial)}$$

and $(S, \mathcal{B}([0,1]), \tilde{\mathbb{P}})$ forms a probability space where $\tilde{\mathbb{P}}$ is an extension of $\hat{\mathbb{P}}$

ie. $\tilde{\mathbb{P}}(x) = \hat{\mathbb{P}}(x)$ where $x \in \mathcal{B}_0$

(There is a unique such extension $\tilde{\mathbb{P}}$)
(non trivial)

THEOREM 7:

A probability measure $\mathbb{P}$ defined on $\mathcal{A}$ satisfies

(i) $\mathbb{P}(E_2 \setminus E_1) = \mathbb{P}(E_2) - \mathbb{P}(E_1)$ for $E_1 \subseteq E_2$ and $E_1, E_2 \in \mathcal{A}$

(ii) $\mathbb{P}(E^c) = 1 - \mathbb{P}(E)$

(iii) $\mathbb{P}(E_1 \cup E_2) = \mathbb{P}(E_1) + \mathbb{P}(E_2) - \mathbb{P}(E_1 \cap E_2)$ for $E_1, E_2 \in \mathcal{A}$

(iv) Principle of inclusion and exclusion:

$$\mathbb{P}\left(\bigcup_{k=1}^{n} E_k\right) = \sum_{k=1}^{n} \mathbb{P}(E_k) - \sum_{1 \le i < j \le n} \mathbb{P}(E_i \cap E_j) + \sum_{1 \le i < j < k \le n} \mathbb{P}(E_i \cap E_j \cap E_k) - \cdots + (-1)^{n+1}\, \mathbb{P}\left(\bigcap_{k=1}^{n} E_k\right)$$

for $n > 2$, $E_1, E_2, \ldots, E_n \in \mathcal{A}$

(v) $\sum_{k=1}^{n} \mathbb{P}(E_k) - \sum_{1 \le i < j \le n} \mathbb{P}(E_i \cap E_j) \le \mathbb{P}\left(\bigcup_{k=1}^{n} E_k\right)$

(vi) $\mathbb{P}\left(\bigcup_{k=1}^{n} E_k\right) \le \sum_{k=1}^{n} \mathbb{P}(E_k)$ (Boole's inequality)

THEOREM 8:

Let $E_n \in \mathcal{A}$ for $n = 1, 2, \ldots$

(i) if $E_1 \subset E_2 \subset E_3 \subset \cdots$

and $E = \bigcup_{n=1}^{\infty} E_n$

then $\lim_{n\to\infty} \mathbb{P}(E_n) = \mathbb{P}(E)$ — continuity from below

(ii) if $E_1 \supset E_2 \supset E_3 \supset \cdots$

and $E = \bigcap_{n=1}^{\infty} E_n$

then $\lim_{n\to\infty} \mathbb{P}(E_n) = \mathbb{P}(E)$ — continuity from above

Definition: Let $\{E_n\}$ be a collection of sets.

We define $\inf_{k\ge n} E_k = \bigcap_{k=n}^{\infty} E_k$

$\sup_{k\ge n} E_k = \bigcup_{k=n}^{\infty} E_k$

$\liminf_{n\to\infty} E_n = \bigcup_{n=1}^{\infty} \bigcap_{k=n}^{\infty} E_k$

$\limsup_{n\to\infty} E_n = \bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} E_k$

If $\liminf_{n\to\infty} E_n = \limsup_{n\to\infty} E_n$; we say $\lim_{n\to\infty} E_n$ exists & equals then

THEOREM 9 (Continuity of probability)

Let $\{E_n\}$ be a sequence of events so that $\lim_{n\to\infty} E_n = E$. Then,

$$\lim_{n\to\infty} \mathbb{P}(E_n) = \mathbb{P}(E)$$

THEOREM 10 (Fatou's lemma)

For any sequence $\{E_n\} \subset \mathcal{A}$,

$$\mathbb{P}\left(\liminf_{n\to\infty} E_n\right) \leq \liminf_{n\to\infty} \mathbb{P}(E_n)$$

$$\limsup_{n\to\infty} \mathbb{P}(E_n) \leq \mathbb{P}\left(\limsup_{n\to\infty} E_n\right)$$

Definition: For a probability space $(S, \mathcal{A}, \mathbb{P})$ with $G \in \mathcal{A}$ such that $\mathbb{P}(G) > 0$, we define the conditional probability of $E \in \mathcal{A}$ wrt G (ie. given G occured) as $\mathbb{P}\left(\frac{E}{G}\right) := \frac{\mathbb{P}(E \cap G)}{\mathbb{P}(G)}$

Definition: A collection $\{G_1, \ldots, G_N\} \subset \mathcal{A}$ is called a partition of S if

(i) $G_i \cap G_j = \emptyset \quad \forall \; i, j \in [N], \; i \neq j$

(ii) $\bigcup_{i=1}^{N} G_i = S$

ie. all G_i's are ~~mutually~~ pairwise disjoint and they 'span' the entire set S

THEOREM 11 (Total probability rule)

For $(S, \mathcal{S}, \mathbb{P})$, if $\{G_1, \ldots, G_N\}$ is a partition of S with $\mathbb{P}(G_i) > 0 \ \forall i \in [N]$, then $\forall E \in \mathcal{S}$,

$$\mathbb{P}(E) = \sum_{i=1}^{N} \mathbb{P}\left(\frac{E}{G_i}\right) \cdot \mathbb{P}(G_i)$$

THEOREM 12 (Baye's theorem)

For $(S, \mathcal{S}, \mathbb{P})$ if $\{G_1, \ldots, G_N\}$ is a partition of S with $\mathbb{P}(G_i) > 0 \ \forall i \in [N]$, then $\forall E \in \mathcal{S}$ with $\mathbb{P}(E) > 0$, we have,

$$\mathbb{P}\left(\frac{G_j}{E}\right) = \frac{\mathbb{P}\left(\frac{E}{G_j}\right) \mathbb{P}(G_j)}{\sum_{i=1}^{N} \mathbb{P}(G_i)\, \mathbb{P}\left(\frac{E}{G_i}\right)}$$

Definition: $E, G \in \mathcal{S}$ are said to be independent events if $\mathbb{P}(E) \cdot \mathbb{P}(G) = \mathbb{P}(E \cap G)$. Further, $H \subset \mathcal{S}$ is said to be mutually indepenpt collection if for every subcollection $T \subseteq H$,

$$\mathbb{P}\left(\bigcap_{E_i \in T} E_i\right) = \prod_{E_i \in T} \mathbb{P}(E_i).$$

$H \subset \mathcal{S}$ is said to be a pairwise independent collection if $\mathbb{P}(E_i \cap E_j) = \mathbb{P}(E_i)\, \mathbb{P}(E_j) \ \forall E_i, E_j \in H \ (i \neq j)$

Note: mutually independent $\underset{\not\Leftarrow}{\Longrightarrow}$ pairwise independent

THEOREM 13 (Borel Cantelli)

For a probability space $(S, \mathcal{L}, \mathbb{P})$, let $\{E_k\} \subset \mathcal{L}$ be a sequence of events. Denote $E = \limsup_{n\to\infty} E_n$; we have,

(i) $\sum_{n=1}^{\infty} \mathbb{P}(E_n) < \infty \Rightarrow \mathbb{P}(E) = 0$

(ii) If $\{E_k\}$ is mutually independent such that

$$\sum_{n=1}^{\infty} \mathbb{P}(E_n) = \infty, \text{ then } \mathbb{P}(E) = 1$$

Definition: Let $(S_k, \mathcal{L}_k, \mathbb{P}_k)$ be probability spaces for $k = 1, 2, \dots, n$.

We define the product space $(S, \mathcal{L}, \mathbb{P})$ where

$S = S_1 \times \cdots \times S_n$

and $\mathcal{L} = \begin{cases} 2^S & \text{all of } S_1, \dots, S_n \text{ are countable} \\ \sigma(C) & \text{at least one } S_i \text{ is uncountable} \end{cases}$

(where $C = \{E_1 \times \cdots \times E_n \mid E_k \in \mathcal{L}_k \ \forall\ k = 1, \dots, n\}$

and if all $S_1, \dots S_n$ are countable, we define

$$\mathbb{P}(E) = \sum_{s \in E} \prod_{k=1}^{n} \mathbb{P}_k(\{s_k\}), \quad s = (s_1, \dots, s_n)$$

If at least one S_k is uncountable, we define

$\hat{\mathbb{P}}(E_1 \times \cdots \times E_n) = \prod_{k=1}^{n} \mathbb{P}_k(E_k)$, (extension from borel)

Definition: A sequence of independent but identical stages is called a sequence of independent trials.

For the case of 2 outcomes at each stage, we refer to the sequence as Bernoulli trials

Definition: For a sample space $(S, \mathcal{A})$, $X : S \to \mathbb{R}$ is called a random variable if $\{s \in S \mid X(s) \in B\} \in \mathcal{A}$ $\forall\ B \in \mathcal{B}(\mathbb{R})$.

THEOREM 14

X is a random variable iff $X^{-1}(B) \in \mathcal{A}$ $\forall B \in \mathcal{B}(\mathbb{R})$

and $X^{-1}(T) = \{x \in S \mid X(x) \in T\}$

THEOREM 15

For $(S, \mathcal{A})$, $X : S \to \mathbb{R}$ is a random variable iff

$\forall\ x \in \mathbb{R}$ $\{X \leq x\} \in \mathcal{A}$.

where $\{X \leq x\} := \{s \in S \mid X(s) \leq x\} = X^{-1}((-\infty, x])$

THEOREM 16

If X, Y are random variables on $(S, \mathcal{A})$, so are

$aX + bY$, X^{m}, XY $\forall\ a, b \in \mathbb{R}$, $m \in \mathbb{N}$

Definition: A real valued function $F : \mathbb{R} \to [0,1]$ which is non decreasing, right continuous and has $F(-\infty) = 0$, $F(+\infty) = 1$ is called a distribution function

THEOREM 17

For a probability space $(S, \mathcal{A}, \mathbb{P})$ with a random variable X, the set map $\mathbb{P}_X(B) = \mathbb{P}(X^{-1}(B))$ $\forall\ B \in \mathcal{B}(\mathbb{R})$ is a probability measure on $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$

THEOREM 18

For a given r.v X on $(S, \mathcal{A}, P)$, the function $F_X : \mathbb{R} \to [0,1]$ as $F_X(x) = \mathbb{P}_X((-\infty, x])$

$$= \mathbb{P}(\{X \le x\})$$

is a distribution function

Definition: Motivated by the above theorem 18, we define the dist function of a r.v X on $(S, \mathcal{A})$ as

$F_X(x) = \mathbb{P}(\{X \le x\})$ $\forall\ x \in \mathbb{R}$

THEOREM 19

For a given distribution function F, $\exists!$ probability space and a unique random r.v X s.t. F is a distribution function of X.

Definition: If a r.v X on $(S, \mathcal{A})$ assumes only countably many values, we say X is a discrete random variable (ie. $\exists\ J$ countable s.t. $\mathbb{P}(X^{-1}(J)) = 1$)

Definition: For $(S, \mathcal{A}, \mathbb{P})$ and X, a discrete r.v, we define the real valued function $p_X : \mathbb{R} \to [0,1]$ as

$p_X(x) = \mathbb{P}(\{X = x\})$. We call this as the probability mass function of X

Examples of distributions

1) Bernoulli distribution : n independent reps of Bernoulli trials

$X_n \to$ denotes no. of success in n trials

$$p_{X_n}(x) = \begin{cases} \binom{n}{x} p^x (1-p)^{n-x} & x = 0, 1, \dots, n \\ 0 & \text{otherwise} \end{cases}$$

2) Hypergeometric distribution : r red balls and g green balls in a sack and $n (\leq r+g)$ are drawn without replacement

$X \to$ denotes no. of red balls in sample

$$p_X(x) = \begin{cases} \dfrac{\binom{r}{x}\binom{g}{n-x}}{\binom{r+g}{n}} & x = 0, 1, \dots, n \\ 0 & \text{otherwise} \end{cases}$$

3) Geometric distribution : Repeating bernoulli trials till first success occurs

X is a geometric r.v if $X(s) \in \{0, 1, 2, \ldots\}$ and

$$P_X(x) = \begin{cases} (1-p)^x \; p & x = 0, 1, \ldots \\ 0 & \text{otherwise} \end{cases}$$

for p being some real in $[0,1]$ ~~probability~~ ~~~~

4) Negative binomial distribution :

$$P_X(x) = \begin{cases} \binom{\alpha + x - 1}{x} p^\alpha (1-p)^x & x = 0, 1, \ldots \\ 0 & \text{otherwise} \end{cases}$$

5) Poisson distribution :

$$P_X(x) = \begin{cases} e^{-\lambda} \dfrac{\lambda^x}{x!} & x = 0, 1, \ldots \\ 0 & \text{otherwise} \end{cases}$$

Definition : A r.v X is said to be a continuous r.v if $\mathbb{P}(\{X = x\}) = 0 \quad \forall x \in \mathbb{R}$

THEOREM 20.

X is a continuous ~~~~ r.v on $(S, \mathcal{S})$ iff F_X (dist func of X) is continuous on $\mathbb{R}$.

Definition: A density function is a non negative integrable function $p : \mathbb{R} \to \mathbb{R}$ s.t. $\int_{-\infty}^{\infty} p(x)\, dx = 1$

X is said to be absolutely continuous if $\exists$ a density function p_X s.t. $F_X(x) = \int_{-\infty}^{x} p_X(t)\, dt$, $x \in \mathbb{R}$.

This p_X is called the probability dist func for X

Definition: $\phi : \mathbb{R} \to \mathbb{R}$ is a Borel measurable function on $\mathbb{R}$ if $\phi^{-1}((-\infty, y])$ is a Borel set in $\mathbb{R}$ $\forall y \in \mathbb{R}$

THEOREM 21

For a r.v on $(S, \mathcal{A})$ and ϕ, a borel measurable function, $Y = \phi \circ X$ is a r.v on $(S, \mathcal{A})$

Definition: Let $a < b$, $a, b \in \mathbb{R}$. We say X has uniform density if

$$p_X(x) = \begin{cases} \dfrac{1}{b-a} & a < x < b \\ 0 & \text{otherwise} \end{cases}$$

and consequently,

$$F_X(x) = \begin{cases} 0 & x < a \\ \dfrac{x-a}{b-a} & a \leq x < b \\ 1 & x \geq b \end{cases}$$

THEOREM 22 (Change of Variable)

Let $\phi : \mathbb{R} \to \mathbb{R}$ be strictly monotonic and differentiable on some $I \subset \mathbb{R}$. Let $\phi(I)$ be the range of ϕ. Let X be a continuous r.v with pdf P_X such that $P_X(x) = 0$ if $x \notin I$. Then $Y = \phi \circ X$ is continuous with pdf

$$P_Y(y) = \begin{cases} P_X(\phi^{-1}(y)) \cdot \left|\frac{d}{dy} \phi^{-1}(y)\right| & y \in \phi(I) \\ 0 & \text{otherwise} \end{cases}$$

Definition: Given a borel m'ble function ϕ, r.v X, the dist function of $Y = \phi \circ X$ is called as the induced distribution and is given by

$$F_Y(y) := \mathbb{P}(\{x \in \phi^{-1}((-\infty, y])\})$$

Examples of distributions

1) Normal distribution

$$P_X(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) \quad \forall x \in \mathbb{R}$$

and we write $X \sim N(\mu, \sigma^2)$

where μ, σ^2 are mean & variance and are effectively constant.

2) Standard normal distribution

$$P_X(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2} \quad \forall \; x \in \mathbb{R}$$

3) Exponential distribution

$$P_X(x) = \begin{cases} \lambda e^{-\lambda x} & x \geq 0 \\ 0 & x < 0 \end{cases}$$

4) Gamma distribution

$$P_X(x) = \begin{cases} \frac{1}{\Gamma(\alpha)\beta^{\alpha}} x^{\alpha-1} e^{-x/\beta} & x > 0 \\ 0 & x \leq 0 \end{cases}$$

$X \sim G(\alpha, \beta)$,

and $\Gamma(\alpha) := \int_0^\infty x^{\alpha-1} e^{-x} \, dx$

5) Cauchy distribution

$$P_X(x) = \frac{1}{\pi(1+x^2)} \quad \forall \; x \in \mathbb{R}$$

Definition: A density function f is said to be symmetric about $\alpha \in \mathbb{R}$ if $f(\alpha - x) = f(\alpha + x) \;\forall\; x \in \mathbb{R}$

X is symmetric about α if $\mathbb{P}(\{X \geq \alpha - x\}) = \mathbb{P}(\{X \leq \alpha + x\})$
$\forall\; x \in \mathbb{R}$

Definition: Set $I_n = \{ (x_1, \dots, x_n) : x_i \in (-\infty, a_i]$ for some real a_i's $\}$. $\underline{X}(s) = (X_1(s), \dots, X_n(s))$ is a random vector for real valued functions X_i defined on a set S if $X^{-1}(I_n) \in \mathcal{A}$ $\forall\ I_n \subseteq \mathbb{R}^n$

note: $X^{-1}(I_n) = \{ s \in S \mid X_i(s) \leq a_i \ \forall\ i = 1, 2, \dots, n \}$

THEOREM 23

If $X_1, \dots, X_n$ are R.V's on $(S, \mathcal{A})$, then $\underline{X}$ is a r.vec on $(S, \mathcal{A})$ where $\underline{X} = (X_1, \dots, X_n)$

Definition: For a random vector $\underline{X} = (X, Y)$, the function

$F_{\underline{X}}(x, y) = \mathbb{P}(\{X \leq x, Y \leq y\})$ $\forall\ x, y \in \mathbb{R}^2$ is the joint distribution function of $\underline{X}$. Further,

$F_X(x) = \mathbb{P}(\{X \leq x\}) = F_{\underline{X}}(x, \infty)$ and

$F_Y(y) = \mathbb{P}(\{Y \leq y\}) = F_{\underline{X}}(\infty, y)$ are called as the marginal distribution functions

Definition: $\underline{X}$ is a discrete random vector if it assumes countably infinite (or finite) values over $\mathbb{R}^2$

Definition: $\underline{X}$ is a discrete r-vec. $f_{\underline{X}} : \mathbb{R}^2 \to [0, 1]$ as $f_{\underline{X}}(x, y) = \mathbb{P}(\{X = x, Y = y\})$ is the joint probability mass function

THEOREM 24 :

The joint pmf of an r.vec $\underline{X} = (X, Y)$ over $(S, \mathcal{A}, \mathbb{P})$ satisfies

(i) $f_{\underline{X}}(x,y) \geq 0 \quad \forall\, x,y \in \mathbb{R}^2$

(ii) $$\sum_i f_{\underline{X}}(x_i, y_i) = 1$$

where $X(S) = \{ (x_i, y_i) \mid i = 1, 2, \dots \}$

(iii) $$A \subseteq \mathbb{R}^2 \Rightarrow \mathbb{P}(\{X \in A\}) = \sum_{(x,y) \in A} f_{\underline{X}}(x,y)$$

Definition: $f_X(x) = \mathbb{P}(\{X = x\})$
and $f_Y(y) = \mathbb{P}(\{Y = y\})$
are called the marginal probability mass functions

Definition: $\underline{X} = (X, Y)$, a r.vec over $(S, \mathcal{A}, \mathbb{P})$ is an absolutely continuous r.vec if $\exists\, f_{\underline{X}}$ for which

$$F_X(x,y) = \int_{-\infty}^{x} \int_{-\infty}^{y} f_{\underline{X}}(s,t) \quad \text{and } f_{\underline{X}} : \mathbb{R}^2 \to \mathbb{R}$$

is a non negative mapping. We then refer to $f_{\underline{X}}$ as the joint probability density function.

$f_X(x) = \int_{-\infty}^{\infty} f_{\underline{X}}(x,t)\,dt$, $f_Y(y) = \int_{-\infty}^{\infty} f_{\underline{X}}(s,y)\,ds$ are the marginal p.d.f's of X, Y respectively.

THEOREM 25

In accordance with our previously made definition:

(i) $F_{\underline{X}}(-\infty, y) = F_{\underline{X}}(x, -\infty) = 0$

(ii) $\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{\underline{X}}(s,t)\, dt\, ds = 1$

(iii) $\left.\begin{array}{l} \lim\limits_{t \downarrow y} F_{\underline{X}}(x,t) = F_{\underline{X}}(x,y) \\ \lim\limits_{s \downarrow x} F_{\underline{X}}(s,y) = F_{\underline{X}}(x,y) \end{array}\right\}$ right continuity in both variables

(iv) $F_{\underline{X}}(x_1, y) \le F_{\underline{X}}(x_2, y) \quad \forall\ x_1 \le x_2$

$F_{\underline{X}}(x, y_1) \le F_{\underline{X}}(x, y_2) \quad \forall\ y_1 \le y_2$

(v) $f_{\underline{X}}(x,y) = \dfrac{\partial^2}{\partial x \partial y} F_{\underline{X}}(x,y)$

(vi) $\mathbb{P}(\{(X,Y) \in A\}) = \iint_A f_{\underline{X}}(s,t)\, dt\, ds$

for $A \subseteq \mathcal{B}(\mathbb{R}^n)$, $\underline{X} = (X, Y)$

Definition: Two random vectors X, Y are said to be independent if for any two Borel sets A, B of $\mathbb{R}$, $\mathbb{P}(\{X \in A, Y \in B\}) = \mathbb{P}(\{X \in A\}) \cdot \mathbb{P}(\{Y \in B\})$

THEOREM 26

Let $\underline{X} = (X, Y)$ be a r.vec. X & Y are independent iff $F_{\underline{X}}(x,y) = F_X(x)\, F_Y(y) \quad \forall\ x, y \in \mathbb{R}^2$

THEOREM 27

Let $\underline{X} = (X, Y)$ be a r.vec. X and Y are independent iff $f_{\underline{X}}(x,y) = f_X(x) f_Y(y)$.

Definition : For X, Y both r.v's., we define the conditional density of Y given X as

$$f_{\frac{Y}{X}}(y|x) := \begin{cases} \dfrac{f_{\underline{X}}(x,y)}{f_X(x)}, & 0 < f_X(x) \\ 0 & \text{otherwise} \end{cases}$$

$$f_{\frac{X}{Y}}(x|y) := \begin{cases} \dfrac{f_{\underline{X}}(x,y)}{f_Y(y)} & f_Y(y) > 0 \\ 0 & \text{otherwise} \end{cases}$$

THEOREM 28

X, Y are independent r.v's iff $f_{\frac{Y}{X}}(y|x) = f_Y(y)$ (or equivalently $f_{\frac{X}{Y}}(x|y) = f_X(x)$) for $f_X(x) > 0$ and $y \in \mathbb{R}$

THEOREM 30 (Baye's rule)

The conditional density satisfies :

$$f_{\frac{X}{Y}}(x|y) = \frac{f_{\frac{Y}{X}}(y|x) f_X(x)}{f_Y(y)} = \frac{f_{\frac{Y}{X}}(y|x) f_X(x)}{\int_{-\infty}^{\infty} f_{\frac{Y}{X}}(y|s) f_X(s)\, ds}$$

THEOREM 31 (Change of variable)

Let $\underline{X} = (X, Y)$ be an absolutely continuous r.vec with joint pdf $P_{\underline{X}}(x, y)$. Let h_1, h_2 be functions mapping $\mathbb{R}^2 \to \mathbb{R}$ and $u = h_1(x, y)$, $v = h_2(x, y)$. Suppose (h_1, h_2) defines a 1-1 transform from a set R in the XY plaine to a set Q in the UV plane then there are functions $w_1, w_2 : \mathbb{R}^2 \to \mathbb{R}$ which invert the operation of h_1, h_2 ie. $w_1(u, v) = x$, $w_2(u, v) = y$. (Suppose h_1, h_2, w_1, w_2 exist ...) and further w_1, w_2 have continuous partial derivatives. Further let the Jacobian of $x = w_1(u, v)$, $y = w_2(u, v)$ is non zero at all $u, v \in Q$ then, $\underline{U} = (U, V)$ is continuous with the pdf

$$P_{\underline{U}}(u, v) = \begin{cases} P_{\underline{X}}(w_1(u,v), w_2(u,v))\,|J| & u, v \in Q \\ 0 & \text{otherwise} \end{cases}$$

Definition: For $X_1, \ldots, X_n$ which are random variables on $(S, \mathcal{S})$ we define $O_1, O_2, \ldots, O_n$ as

$O_k(s) := k^{th}$ smallest value in $\{X_1(s), \ldots, X_n(s)\}$

In particular, $O_1(s) = \min\{X_1(s), \ldots, X_n(s)\}$

$O_n(s) = \max\{X_1(s), \ldots, X_n(s)\}$

These O_k' random variables are known as order statistics and O_k is the k^{th} order statistic of $X_1, \ldots, X_n$

THEOREM 32

The distribution function G_k of O_k is given by

$$G_k(x) = \sum_{j=k}^{n} \binom{n}{j} (F(x))^j (1-F(x))^{n-j}$$

The pdf g_k of O_k is given by

$$g_k(x) = n \binom{n-1}{k-1} f(x) (F(x))^{k-1} (1-F(x))^{n-k}$$

where $F(x)$ is the distribution function of each X_i (hypothesis: all have the same dist. function) and $f(x)$ is the pdf of each X_i (again all same)

THEOREM 33

The r.v $R = O_n - O_1$ is called the sample range and we have:

$G_{1,n}(x,y)$ = joint distribution function of (O_1, O_n)

$$= (F(y))^n - (F(y)-F(x))^n$$

$g_{1,n}(x,y)$ = joint pdf of (O_1, O_n)

$$= \begin{cases} n(n-1)(F(y)-F(x))^{n-2} f(y) f(x) & x \leq y \\ 0 & \text{otherwise} \end{cases}$$

$$f_R(r) = \begin{cases} n(n-1) \int_{-\infty}^{\infty} (F(r+x)-F(x))^{n-2} f(r+x) f(x)\, dx & r \geq 0 \\ 0 & r < 0 \end{cases}$$

THEOREM 34 (sum of r.v's)

Let X, Y be independent r.v's and $Z = X + Y$. Let X, Y have pdf f_X, f_Y

If X, Y both are discrete,

$$f_Z(z) = \sum_x f_X(x)\, f_Y(z-x)$$

If both are continuous, then $\underline{X}$ is absolutely continuous and

$$f_Z(z) = \int_{-\infty}^{\infty} f_{\underline{X}}(x, z-x)\, dx$$

where $\underline{X} = (X, Y)$

Definition: For a given discrete r.v X, we define $E(X) = \sum_i x_i\, p_X(x_i)$ where $X(S) = \{x_1, x_2, \dots\}$

note: we define the infinite sum only if $\sum_{i=1}^{\infty} |x_i| p_X(x_i)$ converges. For a continuous r.v. X, we define $E(X) = \int_{-\infty}^{\infty} x\, p_X(x)\, dx$ provided $\int_{-\infty}^{\infty} |x|\, p_X(x)\, dx$ exists and is finite.

(absolute convergence will guarantee convergence but for borderline cases like $1 - \frac{1}{2} + \frac{1}{3} \dots$ it can have any value)

Examples of expectations

1) Binomial $\longrightarrow$ np
2) hypergeometric $\longrightarrow$ np
3) geometric $\longrightarrow$ $\frac{1-p}{p}$
4) negative binomial $\longrightarrow$ $\alpha\frac{(1-p)}{p}$
5) poisson $\longrightarrow$ λ
6) uniform dist $\longrightarrow$ $(a+b)/2$
7) exponential dist $\longrightarrow$ $1/\lambda$
8) gamma dist $\longrightarrow$ $\alpha\beta$
9) cauchy dist $\longrightarrow$ $\times$ (doesnt converge absolutly)

THEOREM 35

Let $\underline{X}$ be a r.vec having joint pmf $P_{\underline{X}}$, and let it take values $\{\vec{x}_1, \vec{x}_2, \ldots\}$. Let ϕ be a borel measurable map from $\mathbb{R}^n \to \mathbb{R}$. The r.v $Z = \phi(\underline{X})$ has finite expectation iff

$$\sum_{j=1}^{\infty} |\phi(\vec{x}_j)|\, P_{\underline{X}}(\vec{x}_j) < +\infty$$

and we have

$$E(Z) = \sum_{j=1}^{\infty} \phi(\vec{x}_j)\, P_{\underline{X}}(\vec{x}_j).$$

note: one can extend to integration from $-\infty$ to ∞ appropriately in the case of $\underline{X}$ being an absolutely continuous r. vector.

THEOREM 36

Let $c \in \mathbb{R}$. Let X, Y be discrete r.v's having finite expectations. Then,

(i) $E(cX) = c\,E(X)$,

(ii) $E(X+Y) = E(X) + E(Y)$

(iii) $\mathbb{P}(\{X \geq Y\}) = 1 \Rightarrow E(X) \geq E(Y)$

(iv) equality holds in (iii) iff $\mathbb{P}(\{X = Y\}) = 1$

(v) $|E(X)| \leq E(|X|)$

THEOREM 37

Let $\mathbb{P}(\{|X| \leq M\}) = 1$ for some M. Then X has a finite expectation and $|E(X)| \leq M$

THEOREM 38

If the rv's X, Y are independent, then,

$E(XY) = E(X)\,E(Y)$

THEOREM 39

Let X be a non negative, integer valued random variable. Then X has finite expectation iff $\sum_{k=1}^{\infty} \mathbb{P}(\{X \geq k\})$ converges and in this case, $E(X) = \sum_{k=1}^{\infty} \mathbb{P}(\{X \geq k\})$

Definition: We define the α^{th} moment of an r.v X as $E(X^{\alpha})$ (may or may not exist)

THEOREM 40

If X has a moment of order m, then all moments of orders $1, 2, \ldots, m-1$ exist for X.

Definition : If $E((X-c)^k)$ exists, we say that X has a moment of order k about point c. In particular if $c = E(X) = \mu$, we say X has a central moment of order k. In particular, the central moment of order 2 ie. $E((X-\mu)^2)$ is called the variance of X and the positive square root of variance is called the standard deviation of X

THEOREM 41

For any $t \neq \mu$, $Var(X) \leq E((X-t)^2)$ and we have $var(X) = E(X^2) - (E(X))^2$

Examples

1) $X \sim \Gamma(\alpha, \beta)$

$$E(X^m) = \beta^m \alpha(\alpha+1)\cdots(\alpha+m-1)$$

$$\sigma^2 = \beta^2 \alpha$$

2) $X \sim N(\mu, \sigma^2)$

$$E((X-\mu)^{2n}) = \frac{\sigma^{2n}}{\sqrt{2\pi}} 2^{n+1/2} \Gamma(n+\tfrac{1}{2})$$

$$E((X-\mu)^{2n+1}) = 0$$

THEOREM 42

For $a, b \in \mathbb{R}$, X an r.v,

$$Var(aX + b) = a^2 \, Var(X)$$

THEOREM 43

Let X be a discrete r.v that assumes non negative integer values with pdf p_X. Further let,

$$\sum_{x=0}^{\infty} p_X(x) \, t_0^x < +\infty \quad \text{for some } t_0 > 0.$$

If $\phi_X(t) = \sum_{x=0}^{\infty} t^x p_X(x) \quad \forall \, t \in [-t_0, t_0]$,

then $E(X) = \phi_X'(1)$, $Var(X) = \phi_X''(1) + \phi_X'(1) - (\phi_X'(1))^2$

Definition: Motivated by the above, we define the probability generating function of X as ϕ_X given by $\phi_X(t) = \sum_{x=0}^{\infty} t^x p_X(x)$; $t \in [-1, 1]$

Definition: For two r.v's X, Y having finite second moment, if $E((X - E(X))(Y - E(Y)))$ exists, we call it the covariance between X, Y and denote $Cov(X, Y) = E((X - \mu_X)(Y - \mu_Y))$

THEOREM 44

In accordance with the above,

$$Cov(X, Y) = E(XY) - E(X)E(Y)$$

THEOREM 45

we have $Var(X+Y) = Var(X) + Var(Y) + 2\,Cov(X,Y)$

and more generally,

$$Var\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} Var(X_i) + 2\sum_{i=1}^{n-1}\sum_{j=i+1}^{n} Cov(X_i, X_j)$$

Definition: For two r.v's X, Y if they have non zero variances and their covariance exists, we define the correlation coefficient $\rho(X,Y) = \dfrac{Cov(X,Y)}{\sqrt{Var(X)\,Var(Y)}}$

Definition: Let $\underline{X} = (X, Y)$ is a r.vec with joint pdf $f_{\underline{X}}$ and let $\phi : \mathbb{R} \to \mathbb{R}$ be a borel measurable function such that $E(\phi(Y))$ exists.

$$E(\phi(Y) \mid \{X = x\}) := \begin{cases} \dfrac{1}{f_X(x)} \displaystyle\sum_y \phi(y)\, f_{\underline{X}}(x,y) & X \text{ is discrete} \\ & \quad\text{if } f_X(x) > 0 \\ \dfrac{1}{f_X(x)} \displaystyle\int_{-\infty}^{\infty} \phi(y)\, f_{\underline{X}}(x,y)\,dy & X \text{ is cont} \\ 0 & \text{otherwise} \end{cases}$$

Definition ~~we define the moment generating function of X as $M_X(t) = E(e^{tX})$ ...~~

Definition: For a r.v X on $(S, \mathcal{A}, P)$, we define the moment generating function of X as $M_X(t) = E(e^{tX}) \; \forall \; t \in (-\delta, \delta)$; where $\delta > 0$ is such that $E(e^{tX})$ exists $\forall \; t \in (-\delta, \delta)$

THEOREM 46

Let X have mgf $M_X(t)$. Then,

$$E(X^n) = M_X^{(n)}(0)$$

THEOREM 47

Let $X_1, X_2, \ldots, X_n$ be independent r.v's with mgf's $M_{X_1}(t), M_{X_2}(t), \ldots, M_{X_n}(t)$ respectively. Let $Y = X_1 + \ldots + X_n$. Then,

$$M_Y(t) = M_{X_1}(t) \cdot M_{X_2}(t) \cdot \ldots \cdot M_{X_n}(t)$$

THEOREM 48

Two variables having the same mgf (ie. mgf agree on $(-\delta, \delta)$) have the same distribution necessarily

Definition: For X, an r.v on $(S, \mathcal{A}, P)$, the characteristic function of X is defined as $\phi_X(t) = E(e^{itX})$, $\forall \; t \in \mathbb{R}$ with $i^2 = -1$

note: $\phi_X(t)$ exists for any X and also for all $t \in \mathbb{R}$

THEOREM 49

If X has finite moments upto order n, then

$$i^k E(X^k) = \phi_X^{(k)}(0)$$

$\forall$ $k = 1, 2, \dots, n$

THEOREM 50

If two variables have the same characteristic function, then their distribution functions must agree and converse is also true.

THEOREM 51 (Inversion theorem)

Let X be a r.v with characteristic function ϕ_X and pdf f. Then,

$$f(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-itx} \phi_X(t)\, dt$$

holds for every x at which f is differentiable

THEOREM 52

For ϕ, a non negative Borel measurable function and X, a random variable, if $E(\phi(X))$ exists,

$$P(\{\phi(X) \geq \delta\}) \leq \frac{E(\phi(X))}{\delta} \quad \forall\ \delta > 0.$$

THEOREM 53 (Markov & chebyshev ineq)

- If X has finite moment of order r,

$$\mathbb{P}(\{ |X| \geq K \}) \leq \frac{E(|X|^r)}{K^r}$$

for some $K > 0$ (Markov inequality)

- If X has finite expectation μ, variance σ^2,

$$\mathbb{P}(\{ |X-\mu| \geq K\sigma \}) \leq \frac{1}{K^2}$$

for some $K > 0$ (Chebyshev inequality)

Moreover, this inequality is tight and $1/K^2$ is attainable.

Definition: Let X be an r.v on $(S, \mathcal{S}, \mathbb{P})$ and let

- $\{X_n\}$ also be a seq of r.v on $(S, \mathcal{S}, \mathbb{P})$.

Let $\Lambda = \{ s \in S \mid \lim_{n \to \infty} X_n(s) = X(s) \}$

We say $\{X_n\}$ converges almost surely to X if $\mathbb{P}(\Lambda) = 1$ and denote $X_n \xrightarrow{a.s.} X$

Definition: Let $X, \{X_n\}$ be r.v's on $(S, \mathcal{S}, \mathbb{P})$.

If $\forall \varepsilon > 0$ $\lim_{n \to \infty} \mathbb{P}(\{ |X_n - X| \geq \varepsilon \}) = 0$, we say $\{X_n\}$ converges to X in probability and denote by $X_n \xrightarrow{\mathbb{P}} X$

Definition: For a seq $\{F_n\}$ of distribution functions, if $\exists$ a distribution function F s.t. $\lim_{n\to\infty} F_n(x) = F(x)$ for every x at which F is continuous, we say $\{F_n\}$ converges weakly to F and denote $F_n \xrightarrow{w} F$

Definition: We say $\{X_n\}$, a sequence of rv's, converges ~~[illegible]~~ in distribution to X (another r.v) if the respective seq $\{F_n\}$ converges to F_X weakly and denote $X_n \xrightarrow{d} X$

Definition: We say a seq of r.vs $\{X_n\}$ converges in moment of order m to an r.v X if all X_n and also X have moments of order m and we have $\lim_{n\to\infty} E(|X_n - X|^m) = 0$ and we denote $X_n \xrightarrow[m]{moment} X$

THEOREM 54

For $\{X_n\}$, X r.v's in $(S, \mathcal{A}, P)$,

$$X_n \xrightarrow{a.s} X \Rightarrow X_n \xrightarrow{P} X \Rightarrow X_n \xrightarrow{d} X$$

but both converses are not true in general.

Definition: We say $\{X_n\}$ is independent and identically distributed (iid) if the seq of r.v's is independent sequence and all X_n's have the same distribution function

THEOREM 55 (Weak Law of large numbers)

Let $\{X_n\}$ be an iid seq with finite mean $\mu = E(X_n)$ and finite variance $\sigma^2 = var(X_n)$ for $n = 1, 2, \ldots$

we have that for any $\varepsilon > 0$,

$$\lim_{n \to \infty} P(|Y_n - \mu| \geq \varepsilon) = 0$$

where $Y_n = \dfrac{X_1 + X_2 + \cdots + X_n}{n}$.

i.e. $Y_n \xrightarrow{P} \mu$

THEOREM 56 (Strong law of large numbers)

Let $\{X_n\}$ be iid seq with finite mean μ, finite fourth moment about μ. We have,

$P(\{\lim_{n \to \infty} Y_n = \mu\}) = 1$ for same Y_n as above

i.e $Y_n \xrightarrow{a.s} \mu$

THEOREM 57 (Levy Cramer continuity)

Let $\{X_n\}$ be a seq of r.v's.

Assume that mgf M_n of X_n exists for $|t| \le \delta$

for every $n = 1, 2, \ldots$

If there exists X (rv) having mgf M such

that $M_n(t) \to M(t)$ as $n \to \infty$ $\forall\, t \in [-\delta_0, \delta_0]$,

then we have that $X_n \xrightarrow{d} X$

THEOREM 58 (Central limit theorem)

For $\{X_n\}$, an iid seq with mean μ, variance σ^2,

$$\lim_{n \to \infty} \mathbb{P}\left(\frac{S_n - n\mu}{\sigma \sqrt{n}} \le x \right) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-t^2/2}\, dt$$

where $S_n = X_1 + \cdots + X_n$

(and one should notice that RHS is distribution function of $Z \sim N(0,1)$, the standard normal variable)

Exercises 1.2

1) A company packs 1000 bulbs in 1 box. The probability of finding at least one defective bulb is 0.1 and that of finding at least two is 0.05. Find the probability of no defective bulb, exactly one faulty bulb

Ans Let X be the r.v denoting number of defective bulbs.

$\mathbb{P}(X \geq 1) = 0.1$, $\mathbb{P}(X \geq 2) = 0.05$

$\mathbb{P}(X = 0) = 1 - \mathbb{P}(X \geq 1)$

$= 0.9$

($\because X$ is a discrete r.v with range $\{0, 1, \ldots, 100\}$)

$\mathbb{P}(X = 1) = \mathbb{P}(X \geq 1) - \mathbb{P}(X > 1)$

$= \mathbb{P}(X \geq 1) - \mathbb{P}(X \geq 2)$

$= 0.1 - 0.05$

$= 0.05$

2) A fair coin is tossed three times. A player chooses an outcome with probability not 1. If the chosen outcome comes true, player wins, else loses. Suggest an outcome to maximize chance of winning

Ans $S = \{HHH, HTH, HHT, HTT, THH, TTH, THT, TTT\}$

$\mathcal{A} = 2^S = \mathcal{A}_x$

$\mathbb{P}(E) = \frac{|E|}{|S|} = \frac{|E|}{8}$

To maximize $\mathbb{P}(E)$ such that $\mathbb{P}(E) \neq 1$,

$|E| = 7$.

Now, choose E as the event in which the player gets at least one heads so that $|E| = 7$ and $\mathbb{P}(E) = \frac{7}{8}$ which is the maximum possible probability of any event (unless $\mathbb{P} = 1$ is allowed)

3) Let $S = \{0, 1, 2, \dots\}$. For $\lambda > 0$, show that $(S, \mathcal{A}, \mathbb{P})$ is a probability space (where $\mathcal{A} = \mathcal{A}_*$) and

$$\mathbb{P}(E) = \sum_{x \in E} \frac{e^{-\lambda}\lambda^x}{x!}, \quad \forall E \in \mathcal{A}$$

<u>Ans</u> $\mathcal{A}_* = 2^S$ is a σ-field since S is countably infinite

(i) $\mathbb{P}(E) = \sum_{x \in E} \frac{e^{-\lambda}\lambda^x}{x!} \geq 0 \quad \forall E \in \mathcal{A}$

(ii) $\mathbb{P}(S) = \sum_{x=0}^{\infty} \frac{e^{-\lambda}\lambda^x}{x!} = e^{-\lambda} e^{\lambda} = 1$

(iii) For disjoint events $E_1, E_2, \dots \in \mathcal{A}$,

$$\mathbb{P}(E_1 \cup E_2 \cup \dots) = \sum_{x \in E_1 \cup \dots} \frac{e^{-\lambda}\lambda^x}{x!}$$

($\because E_1, E_2, \dots$ are all disjoint)

$$= \sum_{x \in E_1} \frac{e^{-\lambda}\lambda^x}{x!} + \sum_{x \in E_2} \frac{e^{-\lambda}\lambda^x}{x!} + \dots$$

$$= \sum_{i} \mathbb{P}(E_i)$$

4) Let $S = \{0, 1, \dots\}$. $\exists$ $(S, \mathcal{A}, \mathbb{P})$ with

$$\mathbb{P}(E) = \begin{cases} 1 & E \text{ has finite no. of elements} \\ 0 & \text{otherwise} \end{cases}$$

a probability space ?

Ans using theorem 2 (i), and choosing

$E_1 = \{1, 2, 3\}$, $E_2 = \{4, 5\}$

since E_1, E_2 are disjoint,

$\mathbb{P}(E_1 \cup E_2) = \mathbb{P}(E_1) + \mathbb{P}(E_2)$ if $\mathbb{P}$ is a probability measure

But LHS $= \mathbb{P}(\{1, 2, 3, 4, 5\}) = 1$

RHS $= \mathbb{P}(\{1, 2, 3\}) + \mathbb{P}(\{4, 5\}) = 1+1 = 2$

$\therefore \mathbb{P}$ is not a probability measure

5) Suppose $1, 2, \dots, n$ are randomly permuted, what is the probability that there is an integer left unchanged ?

Ans Let E = event in which there is some integer left fixed by the permutation

Define E_i = event in which integer i is left fixed

Then $E = E_1 \cup E_2 \cup \dots \cup E_n$

Using theorem 7 (iv),

$$\mathbb{P}(E) = \sum_{i=1}^{n} \mathbb{P}(E_i) - \sum_{1 \le i < j \le n} \mathbb{P}(E_i \cap E_j) + \dots + (-1)^{n+1} \mathbb{P}(E_1 \cap \dots \cap E_n)$$

Now, $\mathbb{P}(E_{x_1} \cap E_{x_2} \dots \cap E_{x_k}) = \dfrac{(n-k)!}{n!}$

where $x_1, x_2, \dots, x_k$ are distinct integers since $|E_{x_1} \cap E_{x_2} \cdots \cap E_{x_k}| = (n-k)!$

and $|S| = n!$

$\therefore$ we get,

$$P(E) = \binom{n}{1}\frac{(n-1)!}{n!} - \binom{n}{2}\frac{(n-2)!}{n!} + \binom{n}{3}\frac{(n-3)!}{3!} - , \cdots$$

$$+ (-1)^{n+1}\frac{(n-n)!}{n!}$$

$$= \frac{1}{1!} - \frac{1}{2!} + \frac{1}{3!} - \frac{1}{4!} \cdots + \frac{(-1)^{n+1}}{n!}$$

6) For $\mathcal{E} = \{E_1, E_2, \ldots, E_n\}$, a collection of pairwise disjoint subsets of S such that $\bigcup_{i=1}^{n} E_i = S$, show that $\sigma(\mathcal{E})$ = set of all finite unions of E_i's.

Ans By theorem 4, $\sigma(\mathcal{E})$ is the smallest σ-field containing $\mathcal{E}$

$\therefore$ $E_1, E_2, \ldots, E_n \in \sigma(\mathcal{E})$

since $\sigma(\mathcal{E})$ is a σ-field, set of all finite unions will also be in $\sigma(\mathcal{E})$

Let the set of all finite unions of E_i's be B

Then we have just shown that $B \subseteq \sigma(\mathcal{E})$

Now we claim that B is a σ-field. Suppose our claim is true, then again by theorem 4, $\sigma(\mathcal{E}) \subseteq B$ since $\sigma(\mathcal{E})$ is smallest σ field, containing $\mathcal{E}$ and B clearly contains $\mathcal{E}$

$\phi \in \mathcal{B}$ is true by not taking any set in the union

Let $E \in \mathcal{B}$

then $E = E_{x_1} \cup E_{x_2} \cup \dots \cup E_{x_k}$

$E^c = E_{x_1}^c \cap E_{x_2}^c \dots \cap E_{x_k}^c$

But since $E \cup E^c = S$ ~~and~~ and all E_j's are mutually disjoint,

$E^c = E_{y_1} \cup E_{y_2} \cup \dots \cup E_{y_m}$ where

$\{x_1, x_2, \dots, x_k, y_1, y_2, \dots, y_m\} = \{1, 2, \dots, n\}$

$\therefore E^c \in \mathcal{B}$

Now for $E_j \in \mathcal{B}$ $\forall j \in \mathbb{N}$,

$\bigcup_{j=1}^{\infty} E_j$ is also a finite union of E_j's

since $\mathcal{B}$ itself is finite with 2^n elements

$\therefore \bigcup_{j=1}^{\infty} E_j \in \mathcal{B}$

Hence $\mathcal{B}$ is a σ-field.

7) Let $\{E_n\}$ be a seq of events s.t. $\mathbb{P}(E_n) = 1 \; \forall n$

Prove that $\mathbb{P}\left(\bigcap_{n=1}^{\infty} E_n\right) = 1$

Ans $\mathbb{P}\left(\bigcap_{n=1}^{\infty} E_n\right)$

$= \mathbb{P}\left(\left(\bigcup_{n=1}^{\infty} E_n^c\right)^c\right) = 1 - \mathbb{P}\left(\bigcup_{n=1}^{\infty} E_n^c\right)$

$\geq 1 - \sum_{n=1}^{\infty} \mathbb{P}(E_n^c)$ (Boole's inequality)

$= 1$

Since $\mathbb{P}(E) \leq 1$, we have $\mathbb{P}\left(\bigcap_{n=1}^{\infty} E_n\right) = 1$

8) For $(S, \mathcal{A}, \mathbb{P})$ a probability space, $E_1, E_2, \dots \in \mathcal{A}$, if $\sum_{n=1}^{\infty} \mathbb{P}(E_n) < \infty$, prove that $\mathbb{P}(\limsup_{n\to\infty} E_n) = 0$

Ans This is first part of theorem 13.

$$\limsup_{n\to\infty} E_n = \bigcap_{i=1}^{\infty} \bigcup_{j=i}^{\infty} E_j \subseteq \bigcup_{j=i}^{\infty} E_j \quad \forall i \in \mathbb{N}$$

$$\therefore \mathbb{P}(\limsup_{n\to\infty} E_n) \le \mathbb{P}\left(\bigcup_{j=i}^{\infty} E_j\right)$$

$$\le \sum_{j=i}^{\infty} \mathbb{P}(E_j) \quad \forall i \in \mathbb{N}$$

(Boole's inequality)

Since $\sum_{i=1}^{\infty} \mathbb{P}(E_i)$ converges,

we know $\sum_{j=m}^{\infty} \mathbb{P}(E_j) \to 0$ as $m \to \infty$

($\because$ Tail of converging series, converges to 0)

$$\therefore \lim_{i\to\infty} \mathbb{P}(\limsup_{n\to\infty} E_n) \le \lim_{i\to\infty} \sum_{j=i}^{\infty} \mathbb{P}(E_j) = 0$$

$$\therefore \mathbb{P}(\limsup_{n\to\infty} E_n) = 0$$

9) Two distinct teams T_1, T_2 are engaged in designing a new product. we know that the probability that

(i) T_1 meets deadline and succeeds in product is $2/3$

(ii) T_2 succeeds in product is $1/2$

(iii) at least one of them is successful is $3/4$

Find the probability that T_2 succeeded given that exactly one team was successful.

Ans $\mathbb{P}$ (at least one team successful) = $\mathbb{P}(T_1 \cup T_2)$

where T_1 denotes event of team T_1 succeeding

T_2 denotes event of team T_2 succeeding

$\therefore \mathbb{P}(T_1 \cup T_2) = \mathbb{P}(T_1) + \mathbb{P}(T_2) - \mathbb{P}(T_1 \cap T_2)$

$\therefore \frac{3}{4} = \frac{1}{2} + \frac{2}{3} - \mathbb{P}(T_1 \cap T_2)$

$\therefore \mathbb{P}(T_1 \cap T_2) = \frac{5}{12}$

$\mathbb{P}$ (exactly one team successful)

$= \mathbb{P}\{(T_1 \cap T_2^c) \cup (T_2 \cap T_1^c)\}$

$= \mathbb{P}(T_1 \cap T_2^c) + \mathbb{P}(T_2 \cap T_1^c)$

($\because$ they are disjoint events)

$= \mathbb{P}(T_1) - \mathbb{P}(T_1 \cap T_2) + \mathbb{P}(T_2) - \mathbb{P}(T_1 \cap T_2)$

$= \frac{2}{3} + \frac{1}{2} - 2 \times \frac{5}{12} = \frac{1}{3}$

$\therefore \mathbb{P}(T_2 \cap \text{exactly one team successful})$

$= \mathbb{P}(T_2 \cap T_1^c)$

$= \mathbb{P}(T_2) - \mathbb{P}(T_1 \cap T_2)$

$= \frac{1}{12}$

$\therefore \mathbb{P}\left(\frac{T_2}{\text{exactly one success}}\right) = \frac{1/12}{1/3} = \frac{1}{4}$

($\because$ by definition, $\mathbb{P}\left(\frac{A}{B}\right) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)}$)

10) A medical test is performed in a village. If a person is actually affected, the test gives positive with probability 0.99. It is also known that if a person is not affected, the test may give positive result with probability 0.05. Assume that 5% of the village is affected.

(i) Find P(affected and test gave negative)

(ii) Find P(unaffected and test gave positive)

(iii) Find P(affected given test gave positive)

Ans

$$P\left(\frac{\text{positive result}}{\text{affected}}\right) = 0.99$$

$$\therefore P\left(\frac{\text{negative result}}{\text{affected}}\right) = 0.01$$

$$P\left(\frac{\text{positive result}}{\text{not affected}}\right) = 0.05$$

$$\therefore P\left(\frac{\text{negative result}}{\text{not affected}}\right) = 0.95$$

(i) P(affected and negative result)

$$= P\left(\frac{\text{negative result}}{\text{affected}}\right) \times P(\text{affected})$$

$$= 0.01 \times 0.05$$

$$= 0.0005$$

(ii) P(unaffected and positive result)

$$= P\left(\frac{\text{positive}}{\text{unaffected}}\right) \times P(\text{unaffected})$$

$$= 0.05 \times 0.95 = 0.0475$$

(iii) using Bayes theorem,

$$P\left(\frac{\text{affected}}{\text{positive result}}\right) = \frac{P\left(\frac{\text{positive result}}{\text{affected}}\right) P(\text{affected})}{P\left(\frac{+ve}{aff}\right) P(aff) + P\left(\frac{+ve}{\text{not aff}}\right) P(\text{not aff})}$$

($\because$ since {affected, not affected} forms a partition)

$$= \frac{0.99 \times 0.05}{0.99 \times 0.05 + 0.05 \times 0.95}$$

$$= \frac{99}{194} \approx 0.5103$$

11) Consider m+1 boxes with k^{th} box having k red balls and m-k white balls (boxes labelled 0, 1, 2, ..., m). We choose a box at random and choose n balls from that box (with replacement, independently). Suppose a red ball was drawn each of the n times, what is the probability that if we drew one more ball it will be red ?

Ans ~~[illegible]~~

Let E_1 = event of drawing $(n+1)^{th}$ extra ball red

E_2 = event of drawing all n red balls

Let $E_1 \cap E_2 = E_3$

Then E_3 = event of drawing (n+1) red balls.

Let A_i = event of drawing (n+1) red balls from box i $(i = 0, 1, ..., m)$

and B_i : event of drawing n red balls from box i

We need $\mathbb{P}\left(\frac{E_1}{E_2}\right)$.

$$= \frac{\mathbb{P}(E_1 \cap E_2)}{\mathbb{P}(E_2)}$$

$$= \frac{\mathbb{P}(E_3)}{\mathbb{P}(E_2)}$$

$$= \frac{\mathbb{P}(A_0 \cup A_1 \cup \dots \cup A_m)}{\mathbb{P}(B_0 \cup B_1 \cup \dots \cup B_m)}$$

$$= \frac{\sum_{i=0}^{m} \mathbb{P}(A_i)}{\sum_{i=0}^{m} \mathbb{P}(B_i)} \quad \left(\because A_i\text{'s are disjoint},\ B_i\text{'s are disjoint} \right)$$

$$= \frac{\sum_{i=0}^{m} \mathbb{P}(\text{choosing box } i) \times \mathbb{P}\left(\frac{\text{drawing } n+1 \text{ red balls from box } i}{\text{box } i}\right)}{\sum_{i=0}^{m} \mathbb{P}(\text{choosing box } i) \times \mathbb{P}\left(\frac{\text{drawing } n \text{ red balls from box } i}{\text{box } i}\right)}$$

$$= \frac{\sum_{i=0}^{m} \frac{1}{m+1} \times \left(\frac{i}{m}\right)^{n+1}}{\sum_{i=0}^{m} \frac{1}{m+1} \times \left(\frac{i}{m}\right)^{n}}$$

$$= \frac{\sum_{i=0}^{m} \left(\frac{i}{m}\right)^{n+1}}{\sum_{i=0}^{m} \left(\frac{i}{m}\right)^{n}} = \frac{1}{m^{n+1}} \frac{\sum_{i=0}^{m} (i)^{n+1}}{\sum_{i=0}^{m} i^{n}}$$

12) A trader buys stocks S_1, S_2 & wants to sell them the next day. He knows,

- profit or loss of one stock doesnt affect the other
- $\mathbb{P}(S_1 \text{ gives profit}) = 3/4$
- $\mathbb{P}(S_2 \text{ gives loss}) = 1/5$
- $\mathbb{P}(\text{loss in } S_1 > \text{profit in } S_2) = 1/20$
- $\mathbb{P}(\text{loss in } S_2 > \text{profit in } S_1) = 3/20$

Find $\mathbb{P}(S_2 \text{ profited given trader lost some money})$

Ans
$$\mathbb{P}\left(\frac{S_2 \text{ profited}}{\text{overall loss}}\right) = \frac{\mathbb{P}(S_2 \text{ profited} \cap \text{overall loss})}{\mathbb{P}(\text{overall loss})}$$

$$= \frac{\mathbb{P}(\text{loss in } S_1 > \text{profit in } S_2)}{\mathbb{P}(\text{overall loss})}$$

Now $\mathbb{P}(\text{overall loss}) = \mathbb{P}((S_1 \text{loss} \cap S_2 \text{loss}) \cup (\text{loss in } S_1 > \text{profit in } S_2) \cup (\text{loss in } S_2 > \text{profit in } S_1))$

since these events are mutually exclusive, the probability adds up and using independence we split the $\cap$ as product.

$$\therefore \mathbb{P}(\text{overall loss}) = \frac{1}{4} \times \frac{1}{5} + \frac{1}{20} + \frac{3}{20} = \frac{1}{4}$$

$$\therefore \mathbb{P}\left(\frac{S_2 \text{ profited}}{\text{overall loss}}\right) = \frac{1/20}{1/4}$$

$$= 1/5 \;//$$

13) If A, B are independent events, show that A, B^c and A^c, B^c are independent pairs as well

Ans $\mathbb{P}(A \cap B) = \mathbb{P}(A)\,\mathbb{P}(B)$ by definition

$$\mathbb{P}(A \cap B^c) = \mathbb{P}(A \setminus B)$$
$$= \mathbb{P}(A) - \mathbb{P}(A \cap B)$$
$$= \mathbb{P}(A) - \mathbb{P}(A)\,\mathbb{P}(B)$$
$$= \mathbb{P}(A)\,(1 - \mathbb{P}(B))$$
$$= \mathbb{P}(A)\,\mathbb{P}(B^c)$$

$$\mathbb{P}(A^c \cap B^c) = 1 - \mathbb{P}(A \cup B)$$
$$= 1 - (\mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A)\,\mathbb{P}(B))$$
$$= (1 - \mathbb{P}(A))(1 - \mathbb{P}(B))$$
$$= \mathbb{P}(A^c)\,\mathbb{P}(B^c)$$

14) Let E_1, E_2, G all be independent with $\mathbb{P}(G) > 0$.

Prove that $\mathbb{P}\left(\frac{E_1 \cap E_2}{G}\right) = \mathbb{P}\left(\frac{E_1}{G}\right)\mathbb{P}\left(\frac{E_2}{G}\right)$

Ans $\mathbb{P}\left(\frac{E_1 \cap E_2}{G}\right) = \frac{\mathbb{P}(E_1 \cap E_2 \cap G)}{\mathbb{P}(G)}$

$$= \frac{\mathbb{P}(E_1)\,\mathbb{P}(E_2)\,\mathbb{P}(G)}{\mathbb{P}(G)}$$

$$= \frac{\mathbb{P}(E_1)\,\mathbb{P}(G)}{\mathbb{P}(G)} \cdot \frac{\mathbb{P}(E_2)\,\mathbb{P}(G)}{\mathbb{P}(G)}$$

$$= \frac{\mathbb{P}(E_1 \cap G)}{\mathbb{P}(G)}\,\frac{\mathbb{P}(E_2 \cap G)}{\mathbb{P}(G)}$$

$$= \mathbb{P}\left(\frac{E_1}{G}\right)\mathbb{P}\left(\frac{E_2}{G}\right)$$

15) For the given discrete probability spaces $(S_k, \mathcal{A}_k, \mathbb{P}_k)$ for $k = 1, 2, \ldots, n$ consider the product space $(S, \mathcal{A}, \mathbb{P})$

Let $E_k \in \mathcal{A}_k$ for $k = 1, 2, \ldots, n$ be given.

Show that the set of A_j's $(j = 1, 2, \ldots, n)$ given by $A_j = \{ s \in S \mid s_j \in E_j \} \in \mathcal{A}$, is a mutually independent set of events.

Ans For $s \in S$,

$$s = (s_1, s_2, \ldots, s_n) \text{ for some } s_i \in S_i$$
$$i = 1, 2, \ldots, n$$

$$\mathbb{P}(A_k) = \sum_{s \in A_k} \mathbb{P}(\{s\})$$

$$= \sum_{s \in A_k} \prod_{i=1}^{n} \mathbb{P}_i(\{s_i\})$$

$$= \sum_{\substack{(s_1, s_2, \ldots, s_n) \\ \text{with } s_k \in E_k}} \mathbb{P}_1(\{s_1\}) \times \cdots \times \mathbb{P}_n(\{s_n\})$$

$$= \sum_{\substack{(s_1, \ldots, s_n) \\ \text{with } s_k \in E_k}} \frac{1}{|S_1|} \times \frac{1}{|S_2|} \cdots \times \frac{1}{|S_n|} \times |S_k| \, \mathbb{P}_k(\{s_k\})$$

$$= \sum_{\substack{(s_1, \ldots, s_n) \\ s_1 \in S_1 \\ s_2 \in S_2 \\ \vdots \\ s_k \in E_k \\ \vdots \\ s_n \in F_n}} \mathbb{P}_1(\{s_1\}) \times \cdots \times \mathbb{P}_n(\{s_n\})$$

$$= \sum_{s_k \in E_k} \mathbb{P}_k(\{s_k\})$$

(others are 1)

$\therefore \mathbb{P}(A_1 \cap A_2 \cap \dots \cap A_n)$

$$= \sum_{\substack{s_i \in E_i \\ i=1,2,\dots n}} \mathbb{P}_1(\{s_1\}) \dots \mathbb{P}_n(\{s_n\})$$

$$= \left(\sum_{s_1 \in E_1} \mathbb{P}_1(\{s_1\})\right)\left(\sum_{s_2 \in E_2} \mathbb{P}_2(\{s_2\})\right) \dots \left(\sum_{s_n \in E_n} \mathbb{P}(\{s_n\})\right)$$

$$= \prod_{i=1}^{n} \mathbb{P}(A_i)$$

Similarly, we see that for any subset of $\{A_j \mid j = 1, 2, \dots, n\}$, the probability of the intersection splits as the product by using properties of summation.

$\therefore \{A_j \mid j = 1, 2, \dots n\}$ is a set of mutually independent events.

16) A man fires 12 shots independently at a target. Find the probability that he hits the target at least once if he has probability $9/10$ of hitting the target on any given shot

Ans considering Bernoulli trials and denoting hitting the target as 'success', we have $p = \mathbb{P}(\text{success}) = 9/10$

Let X denote no. of success in 12 shots.

X is a discrete r.v with range $\{0, 1, \ldots, 12\}$

$\mathbb{P}(X \geq 1)$

$= 1 - \mathbb{P}(X < 1)$

$= 1 - \mathbb{P}(X = 0)$

$= 1 - \binom{12}{0} p^0 (1-p)^{12}$

$= 1 - \left(\frac{1}{10}\right)^{12}$

Exercises 2.5

1) The distribution of a r.v is given by

$$F_X(x) = \begin{cases} 0 & x < 0 \\ x/4 & 0 \leq x < 1 \\ 1/2 & 1 \leq x < 2 \\ x/12 + 1/2 & 2 \leq x < 3 \\ 1 & 3 \leq x \end{cases}$$

Find the following

(i) $\mathbb{P}(\{X < 2\})$

(ii) $\mathbb{P}(\{X = 2\})$

(iii) $\mathbb{P}(\{1 \leq X < 3\})$

(iv) $\mathbb{P}(\{X > 3/2\})$

(v) $\mathbb{P}(\{X = 5/2\})$

(vi) $\mathbb{P}(\{2 < X \leq 7\})$

Ans (i) $\mathbb{P}(\{X<2\}) + \mathbb{P}(\{X=2\})$

$= \mathbb{P}(\{X \le 2\})$

$= F_X(2)$

$= \frac{1}{6} + \frac{1}{2} = \frac{2}{3}$

But $\mathbb{P}(\{X=2\})$.

$= \lim_{h\to 0} \mathbb{P}(\{2-h < x \le 2\})$

$= \lim_{h\to 0} F_X(2) - F_X(2-h)$

$= \lim_{h\to 0} \left(\frac{2}{3} - \frac{1}{2}\right)$

$= \frac{1}{6}$

$\therefore \mathbb{P}(\{X<2\}) = 1/2$

(ii) As above, $\mathbb{P}(\{X=2\}) = 1/6$

(iii) $\mathbb{P}(\{1 \le X < 3\})$

$= F_X(3^-) - F_X(1^-)$

$= \lim_{h\to 0^-} F_X(3+h) - \lim_{t\to 0^-} F_X(1+t)$

$= \lim_{h\to 0^+} \frac{1}{12}(3-h) + \frac{1}{2} - \lim_{t\to 0^+} \frac{1-t}{4}$

$= 1/2$

(iv) $\mathbb{P}(\{X > 3/2\}) = 1 - \mathbb{P}(\{X \le 3/2\})$

$$= 1 - F_X(3/2)$$

$$= 1 - 1/2$$

$$= 1/2$$

(v) $\mathbb{P}(\{X = 5/2\})$

$$= \lim_{h \to 0^+} F_X(5/2) - F_X(5/2 - h)$$

$$= \lim_{h \to 0^+} \frac{5}{24} + \frac{1}{2} - \lim_{h \to 0^+} \frac{1}{12}\left(\frac{5}{2} - h\right) + \frac{1}{2}$$

$$= 0$$

(vi) $\mathbb{P}(\{2 < X \le 7\})$

$$= F_X(7) - F_X(2)$$

$$= 1 - \left(\frac{1}{6} + \frac{1}{2}\right)$$

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

2) Let a pair of dice with 6 faces be rolled. Let X be r.v indicating sum of number of points on both the die. obtain dist. func, pmf.

Ans: X can only take values between 2 to 12 (integer values) making it a discrete r.v.

$P_X(2) = \mathbb{P}(\{X = 2\}) = 1/36$

$P_X(3) = \mathbb{P}(\{X = 3\}) = 2/36$

In general,

$P_X(x) = \dfrac{\text{no of the solutions of } a+b = x \text{ ; } a,b \in \{1,\dots,6\}}{36}$

Finding it explicitly one by one,

$$P_X(x) = \begin{cases} \dfrac{6-|x-7|}{36} & x = 2,3,\dots,12 \\ 0 & \text{otherwise} \end{cases}$$

and hence.

$$F_X(x) = \begin{cases} 0 & x < 2 \\ \sum_{i=2}^{\lfloor x \rfloor} \dfrac{6-|i-7|}{36} & 2 \le x \le 12 \\ 1 & x > 12 \end{cases}$$

3) An online exam stops when 50% or more servers fail. Let p be probability that any of the servers run successfully. Given that success is mutually independent, find values of p for which 4 servers are preferred over 2

Ans $\mathbb{P}$(exam stops in 4 server case)

$= \mathbb{P}(\{X \geq 2\})$

$= \mathbb{P}(\{X=2\}) + \mathbb{P}(\{X=3\}) + \mathbb{P}(\{X=4\})$

$= \binom{4}{2} p^2(1-p^2) + \binom{4}{3} p^1 (1-p)^3 + \binom{4}{4} p^0 (1-p)^4$

$\mathbb{P}$(exam stops with 2 servers) $= \binom{2}{1} p^1 (1-p)^1 + \binom{2}{2} (1-p)^2$

we require

$\mathbb{P}$ (exam stops with 4 servers) $<$ $\mathbb{P}$ (exam stops with 2)

solving the inequality

$$6p^2(1-p)^2 + 4p(1-p)^3 + (1-p)^4 < 2p(1-p) + (1-p)^2,$$

we get $\quad 1 > p > \frac{1}{3}$

4) Components of a 6 component system are to be picked from a bin of 20 components. The system will be function if 4 of 6 components are in working condition. Given that 15 out of 20 are functioning, find probability that system we make will function.

<u>Ans</u> $\mathbb{P}$ (system is functional)

$$= \mathbb{P}(\{x \geq 4\})$$

$$= \frac{\binom{15}{4}\binom{5}{2} + \binom{15}{5}\binom{5}{1} + \binom{15}{6}\binom{5}{0}}{\binom{20}{6}}$$

$$\approx 0.8687$$

5) Let x be no of births till first girl is born (including girl). If symmetric probability is taken, find pmf of x

Ans $\mathbb{P}(\{X = i\}) = \left(\frac{1}{2}\right)^{i-1} \times \frac{1}{2} = \left(\frac{1}{2}\right)^{i}$

$\forall\ i = 1, 2, \ldots$

and $\mathbb{P}(\{X = 0\}) = 0$

$$\therefore\ P_X(x) = \begin{cases} \left(\frac{1}{2}\right)^{x} & x = 1, 2, \ldots \\ 0 & \text{otherwise} \end{cases}$$

$$\therefore\ F_X(x) = \begin{cases} 0 & x < 1 \\ 1 - \left(\frac{1}{2}\right)^{n-1} & n-1 \leq x < n \quad n = 1, 2, \ldots \end{cases}$$

$(\because$ using GP $)$

6) A mathematician has matches in left & right pocket. He chooses a pocket randomly, uniformly and picks a match. Both pockets initially had N matchsticks. If he finds that one of his pocket is empty, find the probability of having k matches in the other pocket for $k = 0, 1, \ldots, N$

Ans Let H be the event of drawing from left pocket

T be the event of drawing from right pocket

Since one pocket was emptied, there were n H events occuring in the $2n - k$ no of events along with one extra H (assuming H got emptied)

$\therefore$ $\mathbb{P}$(left pocket empty)

$= \mathbb{P}(X = n) \times \frac{1}{2}$ ← extra toss to check left is empty

$= \binom{2n-k}{n} \times \left(\frac{1}{2}\right)^{2n-k} \times \frac{1}{2}$

But since it could have also been the right pocket i.e. T got emptied,

required probability $= \binom{2n-k}{n}\left(\frac{1}{2}\right)^{2n-k} \cancel{\frac{1}{2}} \cdot \cancel{2}$

7) An alarm goes off if it detects 5 or more radioactive particles within a second. If radiation emits particles according to poisson distribution with $\lambda = 0.5$, how likely it is for alarm to go off ?

Ans X = no. of particles registered

$$\mathbb{P}(\text{alarm goes off}) = \mathbb{P}(\{X \geq 5\})$$
$$= 1 - \left(\mathbb{P}(\{X = 0\}) + \cdots + \mathbb{P}(X = 4)\right)$$
$$= 1 - e^{-\lambda}\left(\frac{\lambda^0}{1} + \frac{\lambda^1}{1} + \frac{\lambda^2}{2} + \frac{\lambda^3}{6} + \frac{\lambda^4}{24}\right)$$
$$\approx 1.7 \times 10^{-4}$$

8) X has zeta distribution if

$$p_X(x) = \begin{cases} \dfrac{\zeta}{x^{\alpha+1}} & x = 1, 2, \ldots \\ 0 & \text{otherwise} \end{cases}$$; find ζ given $\alpha > 0$

Ans $\sum_{x} P_X(x) = 1$

$\Rightarrow \sum_{x=1}^{\infty} P_X(x) = 1$ (others it is 0)

$\Rightarrow \sum_{i=1}^{\infty} \frac{S}{i^{\alpha+1}} = 1$

$\Rightarrow S \cdot S_0(\alpha+1) = 1$

where S_0 is the Riemann Zeta function

$\therefore S = \frac{1}{S_0(\alpha+1)}$

9) X represents loss in investment in 1000's of rupees.

If $P_X(x) = \begin{cases} \alpha(2x - 3x^2) & x \in (-1, 0) \\ 0 & \text{otherwise} \end{cases}$

(i) find α

(ii) find P(loss at most 500)

Ans (i) $\int_{-\infty}^{\infty} P_X(x)\, dx = 1$

$\Rightarrow \int_{-1}^{0} \alpha(2x - 3x^2)\, dx = 1$

$\Rightarrow -2\alpha = 1$

$\Rightarrow \alpha = -1/2$

(ii) X represents loss and x is negative.

Considering Y to be a representation of magnitude of loss,

$$P_Y(y) = \begin{cases} \alpha\left(2(-x) - 3(-x)^2\right) & x \in (0,1) \\ 0 & \text{else} \end{cases}$$

where $\alpha = \frac{-1}{2}$

$\therefore \mathbb{P}(Y \leq 500)$

$= F_Y(500)$

$= \int_0^{0.5} \frac{-1}{2}(-2x - 3x^2)\,dx$

$= \frac{3}{16}$

Alternate :

$\mathbb{P}(\{0 \geq x \geq -0.5\})$

$= F_X(0) - F_X(0.5)$

$= \int_{-0.5}^{0} \frac{-1}{2}(2x - 3x^2)\,dx = \frac{3}{16}$

10) X = time b/w two customers visiting a shop

$$P_X(x) = \begin{cases} xe^{-x} & x \geq 0 \\ 0 & \text{otherwise} \end{cases}$$

(i) Find $\mathbb{P}$(once customer leaves, next customer comes after 2 min but before 5 min)

(ii) Find $\mathbb{P}$ (once customer leaves, it takes 7 min for next one to come).

Ans (i) $\mathbb{P}(2 \le X \le 5)$

$= F_X(5) - F_X(2^-)$

$= F_X(5) - F_X(2)$

($\because$ P_X is continuous $\Rightarrow$ F_X is continuous)

$F_X(x) = \int_0^x t e^{-t}\, dt$

$= 1 - e^{-x}(x+1)$

$\therefore F_X(5) - F_X(2) = 3e^{-2} - 6e^{-5}$

(ii) $\mathbb{P}(X \geq 7)$

$= 1 - \mathbb{P}(X < 7)$

$= 1 - F_X(7^-)$

$= 1 - F_X(7)$

$= 1 - (1 - 8e^{-7})$

$= 8e^{-7}$

11) Let X be a random point from $(0,1)$. Find the distribution of $Y = \frac{X}{X+1}$.

Ans X is randomly chosen in $(0,1)$

$\therefore p_X(x) = \begin{cases} \frac{1}{1-0} = 1 & a = 0 < x < 1 = b \\ 0 & \text{otherwise} \end{cases}$

$= \begin{cases} 1 & 0 < x \leq 1 \\ 0 & \text{otherwise} \end{cases}$

$\therefore F_X(x) = \begin{cases} 0 & x < a \\ x & a \leq x < b \\ 1 & x \geq b \end{cases}$

Let $\phi(t) = \frac{t}{t+1}$

$\phi : \mathbb{R} \setminus \{-1\} \to \mathbb{R}$, $\phi'(t) = \frac{1}{(t+1)^2} > 0$

$\therefore \phi$ is strictly monotone and differentiable on $\mathbb{R}$ with $\phi(\mathbb{R} \setminus \{-1\}) = \mathbb{R} \setminus \{1\}$

$\wedge$ $Y = \phi \circ X = \frac{X}{X+1}$

Y is continuous rv except when X takes value -1 but this is not possible since $\mathbb{P}(\{X = -1\}) = 0$

∴ By change of variable formula,

$$p_Y(y) = \begin{cases} p_X(\phi^{-1}(y)) \cdot \left| \frac{d}{dy} \phi^{-1}(y) \right| & y \in \phi(0,1) \\ 0 & \text{else} \end{cases}$$

$$\phi(y) = \frac{y}{y+1} = t$$

$$\Rightarrow \phi^{-1}(y) = \frac{y}{1-y}$$

Also, $\phi(0,1) = (0, 0.5)$

$$\therefore p_Y(y) = \begin{cases} p_X\left(\frac{y}{1-y}\right) \cdot \frac{1}{(1-y)^2} & y \in (0, 0.5) \\ 0 & \text{else} \end{cases}$$

$$= \begin{cases} \frac{1}{(1-y)^2} & y \in (0, 0.5) \\ 0 & \text{else} \end{cases}$$

$$\Rightarrow F_Y(y) = \begin{cases} 0 & y < 0 \\ \frac{y}{1-y} & 0 \le y \le 1 \\ 1 & y > 1 \end{cases}$$

12) Let X be uniformly distributed on $(0,1)$. Show that pdf of $Y = -\frac{\ln(1-X)}{\lambda}$ for $\lambda > 0$, is

$$p_Y(y) = \begin{cases} \lambda e^{-\lambda y} & y > 0 \\ 0 & y \le 0 \end{cases}$$

Ans $p_X(x) = \begin{cases} 1 & 0 \le x \le 1 \\ 0 & \text{else} \end{cases}$

$\phi : (0,1) \to \mathbb{R}$ as $\phi(x) = \dfrac{-\ln(1-x)}{\lambda}$

$\phi'(x) = \dfrac{1}{1-x}$

$\Rightarrow \phi$ is differentiable in $(0,1)$ and is also monotonic

$Y = \dfrac{-\ln(1-X)}{\lambda} = \phi \circ X$

Y is clearly continuous on $(0,1)$ since X takes values only in $(0,1)$

$\therefore$ ~~[illegible]~~ $\phi((0,1)) = \mathbb{R}^+$

$\phi(t) = \dfrac{-\ln(1-t)}{\lambda} \Rightarrow \phi^{-1}(m) = 1 - e^{-\lambda m}$

$\therefore \dfrac{d}{dm} \phi^{-1}(m) = \lambda e^{-\lambda m}$

$\therefore f_Y(y) = \begin{cases} p_X(1-e^{-\lambda y})(\lambda e^{-\lambda y}) & y \in \mathbb{R}^+ \\ 0 & \text{otherwise} \end{cases}$

$(\because 0 < 1 - e^{-\lambda y} < 1 \;\; \forall y \in \mathbb{R}^+)$ $= \begin{cases} \lambda e^{-\lambda y} & y \in \mathbb{R}^+ \\ 0 & \text{else} \end{cases}$

13) $X \sim N(0, \sigma^2)$. Find α, β so that $X^2 \sim G(\alpha, \beta)$

Ans $P_X(x) = \frac{1}{\sqrt{2\pi}\,\sigma} \exp\left(-\frac{x^2}{2\sigma^2}\right) \quad x \in \mathbb{R}$

$\phi(x) = x^2 \Rightarrow \phi$ is monotonic, diff on $x \geq 0$

$\phi^{-1}(x) = \sqrt{x}$ on $x \geq 0$

$\frac{d}{dx} \phi^{-1}(x) = \frac{1}{2\sqrt{x}}$ on $x \geq 0$

$\phi(\mathbb{R}^+ \cup \{0\}) = \mathbb{R}^+ \cup \{0\}$

$$\therefore P_{X^2}(y) = \begin{cases} \frac{1}{\sqrt{2\pi}\,\sigma} \exp\left(\frac{-y}{2\sigma^2}\right) \cdot \frac{1}{2\sqrt{y}} & y \geq 0 \\ 0 & \text{else} \end{cases}$$

$$\therefore P_{X^2}(y) = \begin{cases} \frac{1}{2\sqrt{2}\,\sigma\sqrt{\pi}}\, y^{-1/2}\, e^{-y/2\sigma^2} & y \geq 0 \\ 0 & \text{else} \end{cases}$$

$\therefore$ For $\alpha = 1/2$, $\beta = 2\sigma^2$, we have

$X^2 \sim G(1/2, 2\sigma^2)$

14) For a symmetric rv X about 0, show that $\forall\ x > 0$, the dist func F_X satisfies:

(i) $\mathbb{P}(\{|X| \le x\}) = 2F_X(x) - 1$

(ii) $\mathbb{P}(\{|X| > x\}) = 2(1 - F_X(x))$

(iii) $\mathbb{P}(\{X = x\}) = F_X(x) + F_X(-x) - 1$

Assume F_X is continuous

Ans (i) $\mathbb{P}(\{-x \le X \le x\})$

$= F_X(x) - F_X(-x)$

$= F_X(x) - (1 - F_X(x))$

$\Big($ $\because \mathbb{P}(\{X \geq -x\}) = \mathbb{P}(\{X \le x\})$

since X has symmetric prob about 0 $\Big)$

$= 2F_X(x) - 1$

(ii) $\mathbb{P}(\{|X| > x\}) = 1 - (2F_X(x) - 1)$

$= 2(1 - F_X(x))$

(iii) $\mathbb{P}(\{X = x\}) + \mathbb{P}(\{X < x\}) + \mathbb{P}(\{X > x\})$

$= 1$

$\therefore \mathbb{P}(\{X = x\}) = 1 - F_X(x) - F_X(-x)$

$= 1 - (1 - F_X(-x)) - (1 - F_X(x))$

$= F_X(x) + F_X(-x) - 1$

15) Consider $A = \{(1,1), (1,2), (2,2)\}$. Let X, Y be r.v's s.t. the joint pmf of $\underline{X} = (X, Y)$

i) $$f_{\underline{X}}(x,y) = \begin{cases} \frac{xy^2}{c} & (x,y) \in A \\ 0 & \text{otherwise} \end{cases}$$

(i) Find c

(ii) Find $\mathbb{P}(X+Y \leq 3)$

(iii) Find marginal pmf's of X, Y

(iv) Are X, Y independent

Ans (i) $\sum_{(x,y)\in A} f_{\underline{x}}(x,y) = 1$

$\Rightarrow f_{\underline{x}}(1,1) + f_{\underline{x}}(1,2) + f_{\underline{x}}(2,2) = 1$

$\Rightarrow \frac{1}{c} + \frac{4}{c} + \frac{8}{c} = 1$.

$\Rightarrow c = 13$.

(ii) $\mathbb{P}(X+Y \leq 3)$

$= \mathbb{P}(X+Y=3) + \mathbb{P}(X+Y=2)$

(since range$(X+Y) = \{2,3,4\}$)

$= \frac{4}{13} + \frac{1}{13} = \frac{5}{13}$.

$\Big[\; f_{X+Y}(3) = \mathbb{P}((X,Y) = (1,2))$

$f_{X+Y}(2) = \mathbb{P}((X,Y) = (1,1)) \;\Big]$

(iii) $f_X(x) = \sum_{y \in Y} f_{\underline{X}}(x,y)$

$= x \times \frac{1^2}{13} + x \times \frac{2^2}{13} + x \times \frac{2^2}{13}$

$= \frac{9x}{13}$

similarly, $f_Y(y) = \frac{1 \times y^2}{13} + \frac{1 \times y^2}{13} + \frac{2 \times y^2}{13} = \frac{4y^2}{13}$

(iv) $f_{\underline{X}}(x,y) = \begin{cases} \frac{xy^2}{13} & (x,y) \in A \\ 0 & \text{else} \end{cases}$

$f_X(x) = \begin{cases} \frac{9x}{13} & x = 1,2 \\ 0 & \text{else} \end{cases}$

$f_Y(y) = \begin{cases} \frac{4y^2}{13} & y = 1,2 \\ 0 & \text{else} \end{cases}$

$f_{\underline{X}}(x,y) \neq f_X(x)\, f_Y(y)$

$\therefore$ Not independent

16) Let the joint density function of $\underline{X} = (X, Y)$ for r.v's X, Y be

$$f_{\underline{X}}(x,y) = c \exp\left(\frac{-(y-\frac{x}{2})^2 - \frac{3}{4}x^2}{2}\right)$$

$\forall\ x, y \in \mathbb{R}^2$, $c > 0$.

Find c and comment if X, Y are independent.

Ans $f_{\underline{x}}(x,y) = C \exp\left(-\left(\frac{x^2+y^2}{2} - xy\right)\right)$

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{\underline{x}}(s,t)\,ds\,dt = 1$$

$$\Rightarrow C = \frac{\sqrt{3}}{4\pi}$$

$$f_x(x) = F_{\underline{x}}(x, +\infty)$$
$$= \int_{-\infty}^{\infty} \frac{\sqrt{3}}{4\pi} e^{\left(-\left(\frac{x^2+y^2-xy}{2}\right)\right)} dy$$
$$= \sqrt{\frac{3}{8\pi}}\, e^{-\frac{3x^2}{8}} \qquad x \in \mathbb{R}$$

Illy $f_y(y) = \sqrt{\frac{3}{8\pi}}\, e^{-\frac{3y^2}{8}} \qquad y \in \mathbb{R}$

Clearly $f_{\underline{x}}(x,y) \neq f_x(x)\, f_y(y)$

$\therefore$ Not independent

17) Let X, Y be independent r.v's and $g: \mathbb{R} \to \mathbb{R}$, $h: \mathbb{R} \to \mathbb{R}$ be borel measurable. Show that $g(X)$, $h(Y)$ are also independent rv's

Ans First we show $g(X)$ is r.v if X is r.v.

$\forall A \in \mathcal{B}(\mathbb{R})$, $(g(X))^{-1}(A) \in \mathcal{A}$ ($\because g$ is Borel func)

$$\Rightarrow X^{-1}(g^{-1}(A)) \in \mathcal{A}$$

$\therefore g^{-1}(A) \in \mathcal{B}(\mathbb{R})$ since X is an r.v

$(\Rightarrow)$ $\mathbb{P}(X = x, Y = y) = \mathbb{P}(X = x)\ \mathbb{P}(Y = y)$

$\therefore \mathbb{P}(X = g^{-1}(x), Y = h^{-1}(y))$

$= \mathbb{P}(X = g^{-1}(x))\ \mathbb{P}(Y = h^{-1}(y))$

$= \mathbb{P}(g(X) = x)\ \mathbb{P}(h(Y) = y)$

$(\Leftarrow)$ $\mathbb{P}(g(X) = x, h(Y) = y) = \mathbb{P}(g(X) = x)\ \mathbb{P}(h(Y) = y)$

$\therefore \mathbb{P}(X = g^{-1}(x), Y = h^{-1}(y)) = \mathbb{P}(X = g^{-1}(x))\ \mathbb{P}(Y = h^{-1}(y))$

$\therefore \mathbb{P}(X = x, Y = y) = \mathbb{P}(X = x)\ \mathbb{P}(Y = y)$

by choosing $x = g(x)$, $y = h(y)$

18) X is chosen at random uniformly in $(0,1)$. Then, Y is chosen from $(0, X)$. Find pdf of Y

Ans $X \sim U(0,1)$

$Y | X \sim U(0, x)$

$\therefore$ The pdf is clearly non zero $\forall$ $0 \le y \le x \le 1$ and 0 otherwise

$$f_{Y|X}\left(\frac{y}{x}\right) = \frac{f_{\underline{X}}(x, y)}{f_X(x)}$$

$$f_X(x) = \frac{1}{b-a} = \frac{1}{1-0} = 1$$

Let $Y | X$ be $W \sim U(0, x)$

$$f_W(\omega) = \frac{1}{x-0} = \frac{1}{x}$$

$$\therefore \frac{f_{\underline{X}}(x,y)}{f_X(x)} = f_W(\omega) = 1/x$$

$$\therefore f_{\underline{X}}(x,y) = \begin{cases} \frac{1}{x} & 0 \le y \le x \le 1 \\ 0 & \text{otherwise} \end{cases}$$

19) Let X be rv on $(S, \mathcal{A}, \mathbb{P})$ & X takes values in $\{1, 2, \dots\}$. If $0 < \mathbb{P}(\{X > k\}) < 1$. and if $\mathbb{P}(\{X > k+1 \mid X > k\}) = \mathbb{P}(\{X > 1\})$, show that X has geometric distribution with parameter $p = \mathbb{P}(\{X = 1\})$

Ans
$$\frac{\mathbb{P}(\{X > k+1\} \cap \{X > k\})}{\mathbb{P}(\{X > k\})} = \mathbb{P}(\{X > 1\})$$

$$\therefore \mathbb{P}(X > k+1) = \mathbb{P}(X > 1) \cdot \mathbb{P}(X > k)$$

By induction,

$$\mathbb{P}(X > k) = (\mathbb{P}(X > 1))^{k}$$

$$\therefore 1 - \mathbb{P}(X \le k) = (1-p)^{k}$$

$$\therefore 1 - F_X(k) = (1-p)^{k}$$

$\therefore F_X(k) = 1 - (1-p)^k$

This is indeed the dist func of geometric dist

since $\mathbb{P}(X \le k)$

$$= \sum_{i=1}^{k} (1-p)^{i-1} p$$

$$= 1 - (1-p)^k \quad (\because GP)$$

20) For X, Y r.v's, if cond. prob density func of X given $Y = y$ is

$$f_{X|Y}\left(\tfrac{x}{y}\right) = \begin{cases} \frac{x+y}{1+y} e^{-x} & x, y > 0 \\ 0 & \text{else} \end{cases}$$

find $\mathbb{P}(X<1 \mid Y=2)$

<u>Ans</u> $\cdot \dfrac{\mathbb{P}(\{X<1, Y=2\})}{\mathbb{P}(\{Y=2\})}$

$$f_{\frac{X}{Y}}(X=x \mid y=2) = \begin{cases} \left(\frac{x+2}{3}\right) e^{-x} & x > 0 \\ 0 & \text{otherwise} \end{cases}$$

$$\mathbb{P}(\{X \le 1 \mid Y=2\}) = \int_{-\infty}^{1} f_{\frac{X}{Y}}(X=s \mid y=2)\, ds$$

$$= \int_0^1 \left(\frac{x+2}{3}\right) e^{-x}\, dx$$

$$\approx 0.509$$

21) Let X, Y be continuous independent r.v's with pdfs F_X, F_Y.

Show that pdf and dist func of $U = X+Y$

are $$p_U(u) = \int_{-\infty}^{\infty} p_X(x)\, p_Y(u-x)\, dx$$

$$F_U(u) = \int_{-\infty}^{\infty} p_X(x)\, F_Y(u-x)\, dx$$

Ans Let $A_u = \{ (x,y) \mid x+y \leq u \}$

$$F_U(u) = \iint_{A_u} f_{\underline{X}}(s,t)\, ds\, dt$$

$$= \int_{-\infty}^{\infty} \int_{-\infty}^{y-x} f_{\underline{X}}(s,t)\, ds\, dt$$

after relevant substitutions

$$= \int_{-\infty}^{u} \int_{-\infty}^{\infty} f_{\underline{X}}(x, t-x)\, dx\, dt$$

$$\therefore p_U(u) = F_U'(u) = \int_{-\infty}^{\infty} f_{\underline{X}}(x, u-x)\, dx$$

$$= \int_{-\infty}^{\infty} f_X(x)\, f_Y(u-x)\, dx$$

$$F_U(u) = \int_{-\infty}^{u} p_U(s)\, ds$$

$$= \int_{-\infty}^{\infty} p_X(x) \left(\int_{-\infty}^{u} p_Y(s-x)\, ds \right) dx$$

$$= \int_{-\infty}^{\infty} p_X(x)\, F_Y(u-x)\, dx$$

Note: f and p have been used interchangeably

Exercises 3.3

1) In a quiz two questions are asked on Math and physics. If the first is answered then player proceeds to next one. Correctly answering math gives 50 Rs & correctly answering physics gives 100 Rs. P (answering math) = 0.8, P (answering physics) = 0.5, find expected value if

(i) math is asked first

(ii) physics is asked first

Ans (i) Math first.

Let X be r.v that denotes amount of money received overall. X only takes values 0, 50, 150

$$\therefore p_X(x) = \begin{cases} 0.2 & x=0 \\ 0.4 & x=50 \\ 0.4 & x=150 \\ 0 & \text{else} \end{cases}$$

$$E(X) = \sum_{i=1}^{3} x_i \, p_X(x_i) \quad \text{(by defn)}$$

$$= 0 \times 0.2 + 50 \times 0.4 + 150 \times 0.4$$

$$= \underline{\underline{80}}$$

(ii) Similarly $p_Y(y) = \begin{cases} 0.5 & x=0 \\ 0.5 \times 0.2 & x=100 \\ 0.5 \times 0.8 & x=150 \\ 0 & \text{else} \end{cases}$

$$E(X) = 0 \times 0.5 + 100 \times 0.1 + 150 \times 0.4$$

$$= 10 + 60$$

$$= \underline{\underline{70}}$$

2) A box has 10 discs of radii 1, 2, ..., 10. Find expected value of area of disc selected at random from the box (uniformly)

Ans X denotes area of disc selected

$$X \in \{\pi, 4\pi, 9\pi, \ldots, 100\pi\}$$

$$\therefore E(X) = \sum_{i=1}^{10} (i^2\pi) \times p_X(x)$$

$$= \sum_{i=1}^{10} (i^2\pi) \times \frac{1}{10} = \underline{\underline{38.5\pi}}$$

3) An urn has r red, b black balls. A ball is drawn at random and colour is noted. The ball is replaced with c more balls of same colour. This is repeated successively. X_n = no of red drawn in first n draws. Prove $E(X_n) = \frac{nr}{r+b}$

Ans Base case: $n = 1$

$$E(X_1) = \frac{r}{r+b}$$

$\therefore$ Claim is true for base case since

$$E(X) = E(\text{no. of red balls drawn in try 1})$$

$$= 0 \times \frac{b}{r+b} + 1 \times \frac{r}{r+b}$$

Assume true for a general k

$$E(X_k) = \frac{kr}{r+b}$$

$$E(X_{k+1}) = E(X_k) + E(R_{k+1})$$

where R_{k+1} = no of red balls in $k+1^{th}$ try

$$= \frac{kr}{r+b} + \frac{r}{r+b}$$

$\left(\because E(R_{k+1}) = \frac{r}{r+b}\right.$ from original Polya model$\left.\right)$

$$= \frac{(k+1)r}{r+b}$$

Hence proved.

4) X has the following distribution

$$F_X(x) = \begin{cases} 0 & x < -3 \\ 3/8 & -3 \le x < 0 \\ 1/2 & 0 \le x < 3 \\ 3/4 & 3 \le x < 4 \\ 1 & 4 \le x \end{cases}$$

Find $E(X)$, $E(X^2 - 2|X|)$, $E(X|X|)$

<u>Ans</u> using the concept of Riemann Stiljes integral,

$$E(X) = \int_{-\infty}^{3} x\,dF_1 + \int_{-3}^{0} x\,dF_2 + \int_{0}^{3} x\,dF_3 + \int_{3}^{4} x\,dF_4 + \int_{4}^{\infty} x\,dF_5$$

where $F_1 + F_2 + F_3 + F_4 + F_5 = F$ so that

$$E(X) = \int_{\mathbb{R}} x\,dF$$

we have jump discontinuities at $x = -3, 0, 3, 4$

$$\Rightarrow E(X) = -3\left(\frac{3}{8} - 0\right) + 0\left(\frac{1}{2} - \frac{3}{8}\right) + 3\left(\frac{3}{4} - \frac{1}{2}\right) + 4\left(1 - \frac{3}{4}\right) = \underline{\underline{\frac{5}{8}}}$$

$(\because dF_1 = dF_2 = \cdots = dF_5 = 0$ so we only need to account for the jump discontinuities$)$

Similarly,

$$E(X^2) = (-3)^2\left(\frac{3}{8}-0\right) + 0^2\left(\frac{1}{2}-\frac{3}{8}\right) + 3^2\left(\frac{3}{4}-\frac{1}{2}\right)$$

$$+ (4)^2\left(1-\frac{3}{4}\right) = \frac{77}{8}$$

$$E(|X|) = |-3|\left(\frac{3}{8}-0\right) + |0|\left(\frac{1}{2}-\frac{3}{8}\right) + |3|\left(\frac{3}{4}-\frac{1}{2}\right)$$

$$+ |4|\left(1-\frac{3}{4}\right)$$

$$= \frac{23}{8}$$

$$\Rightarrow E(X^2) - 2E(|X|) = \underline{\underline{\frac{31}{8}}}$$

$$E(X|X|) = -3|-3|\frac{3}{8} + 0 + 3|3|\frac{1}{4} + 4|4|\frac{1}{4}$$

$$= \underline{\underline{\frac{23}{8}}}$$

5) Let X have binomial distribution with parameters n, p. Find $E(X)$

<u>Ans</u>

$$P_X(x) = \begin{cases} \binom{n}{x} p^x (1-p)^{n-x} & x = 0, 1, \dots, n \\ 0 & \text{else} \end{cases}$$

$$\therefore E(x) = \sum_{x=0}^{n} x \binom{n}{x} p^x (1-p)^{n-x}$$

$$= \sum_{x=1}^{n} n\cdot\binom{n-1}{x-1} p^{x} (1-p)^{n-x}$$

$$= np \sum_{t=0}^{n-1} \binom{n-1}{t} p^{t} (1-p)^{(n-1)-t}$$

$$= np\,\left(p+(1-p)\right)^{n-1}$$

$$= np$$

6) Let X be uniformly distributed on $(0, \pi/4)$

Find $E(\cos^2 X)$

Ans
$$p_X(x) = \begin{cases} \frac{4}{\pi} & 0 \le x \le \frac{\pi}{4} \\ 0 & \text{else} \end{cases}$$

$$\therefore E(\cos^2 X) = \int_0^{\pi/4} \cos^2 x \cdot \frac{4}{\pi}\, dx$$

$$= \frac{1}{\pi} + \frac{1}{2}$$

7) For r.v's X, Y, prove the following

(i) $(E(XY))^2 \le E(X^2)\cdot E(Y^2)$

(ii) $\sqrt{E((X+Y)^2)} \le \sqrt{E(X^2)}\,\sqrt{E(Y^2)}$

(iii) $h(E(X)) \le E(h(X))$ for discrete r.v X and $h : (a,b) \to \mathbb{R}$ being convex

Ans (i) $E((tX-Y)^2) = t^2 E(X^2) - 2t\, E(XY) + E(Y^2)$

At extrema, $t_0 = \frac{E(XY)}{E(X^2)}$ $\left(\frac{d}{dt} = 0\right)$

from 2nd der test, t_0 is minima

$$E((t_0 X - Y)^2) = \frac{(E(XY))^2}{E(X^2)} - 2\frac{(E(XY))^2}{E(X^2)} + E(Y^2)$$

$$= E(Y^2) - \frac{(E(XY))^2}{E(X^2)}$$

But $E((tX-Y)^2) \geq 0$

$$\therefore E(Y^2) \geq \frac{(E(XY))^2}{E(X^2)}$$

(ii) $E((X+Y)^2) = E(X^2) + E(Y^2) + 2E(XY)$

$$\leq E(X^2) + E(Y^2) + 2\sqrt{E(X^2)E(Y^2)}$$

$$= \left(\sqrt{E(X^2)} + \sqrt{E(Y^2)}\right)^2$$

(iii) $h((1-\lambda)x + \lambda y) \leq (1-\lambda)\, h(x) + \lambda h(y)$

$$E(X) = \sum_{i=1}^{n} x_i p_i$$

$$h(E(X)) = h\left(\sum_{i=1}^{n} x_i p_i\right)$$

$$\leq \sum_{i=1}^{n} h(x_i)\, p_X(x_i)$$

$$= E(h(X))$$

8) Let X have poisson distribution with parameter λ. Find pgf of X. Find mean, variance of X using the pgf

Ans $\phi_X(t) = \sum_{x=0}^{\infty} t^x p_X(x)$ $\quad -1 \le t \le 1$

$$= \sum_{x=0}^{\infty} t^x \frac{e^{-\lambda} \lambda^x}{x!}$$

$$\therefore \phi_X(t) = e^{\lambda(t-1)} \quad t \in [-1, 1]$$

$$\mu = \phi_X'(1) = \lambda$$

$$\sigma^2 = \phi_X''(1) + \phi_X'(1) - (\phi_X'(1))^2$$

$$= \lambda^2 + \lambda - \lambda^2$$

$$= \lambda$$

9) For an r.v X with mean μ, $\sigma^2 = 0$, show that $\mathbb{P}(\{X = \mu\}) = 1$

Ans $\sigma^2 = E((X-\mu)^2) = 0$

But $E((X-\mu)^2) \geq 0$ always

$\therefore (X-\mu)^2 = 0$ (equality iff $(X-\mu)^2 = 0$)

$\therefore X = \mu$

$\therefore \mathbb{P}(\{X = \mu\}) = 1$ ($\because X$ is const. r.v)

10) An investor invested total money - Rs. N in two different assets and got back Rs M. Let $R = \frac{M-N}{N}$ be rate of return per total inv and R_1, R_2 be r.o.r's for 1st & 2nd investment. Prove $\text{var}(R) \leq \max(\text{var}(R_1), \text{var}(R_2))$

Ans WLOG, assume $\text{var}(R_1) \geq \text{var}(R_2)$

We want to prove that $\text{var}(R) \leq \text{var}(R_1)$

Let money N be split as cN, $(1-c)N$ into the investments for some $0 \leq c \leq 1$

$$\therefore R = c R_1 + (1-c) R_2$$

$$\text{var}(R) = \text{var}(cR_1 + (1-c) R_2)$$

$$\leq c \ \text{var}(R_1) + (1-c) \ \text{var}(R_2)$$

$$\leq c \ \text{var}(R_1) + (1-c) \ \text{var}(R_1)$$

$$= \text{var}(R_1)$$

And we are done

11) Let X, Y be r.v's with joint pdf

$$f_{\underline{X}}(x,y) = \begin{cases} 8xy & 0 < y < x < 1 \\ 0 & \text{else} \end{cases}$$

Find $E(X|Y)$, $E(Y|X)$

<u>Ans</u> $$f_X(x) = \int_{-\infty}^{\infty} f_{\underline{X}}(x,s)\,ds$$

$$= \int_0^x 8xy\,dy = 4x^3$$

$$\therefore f_X(x) = \begin{cases} 4x^3 & 0 < x < 1 \\ 0 & \text{else} \end{cases}$$

$$\therefore E\left(\frac{Y}{X=x}\right) = \frac{1}{f_X(x)} \int_{-\infty}^{\infty} y\, f_{\underline{X}}(x,y)\,dy$$

for $f_X(x) > 0$
ie. $x \in (0,1)$

$$\therefore E\left(\frac{Y}{X=x}\right) = \frac{1}{4x^3} \int_0^x y \times 8xy\,dy$$

$$= \frac{2x}{3}$$

$$\therefore E\left(\frac{Y}{X}\right) = \begin{cases} \frac{2X}{3} & X \in (0,1) \\ 0 & \text{else} \end{cases}$$

Illy $$f_Y(y) = \begin{cases} 4y(1-y^2) & 0 < y < 1 \\ 0 & \text{else} \end{cases}$$

and $E\left(\frac{X}{Y}\right) = \begin{cases} \frac{2}{3}\frac{(1-Y^3)}{(1-Y^2)} & Y \in (0,1) \\ 0 & \text{else} \end{cases}$

12) For any $h: \mathbb{R} \to \mathbb{R}$, prove that

$$E\left((Y - E(Y|X))^2\right) \leq E\left((Y - h(X))^2\right)$$

Ans

$$E\left((Y-h(X))^2\right)$$

$$= E\left(\left(Y - E\left(\frac{Y}{X}\right) + E\left(\frac{Y}{X}\right) - h(X)\right)^2\right)$$

$$= E\left(\left(Y - E\left(\frac{Y}{X}\right)\right)^2 + \left(E\left(\frac{Y}{X}\right) - h(X)\right)^2\right)$$

I II

$$+ 2E\left(\left(Y - E\left(\frac{Y}{X}\right)\right)\left(E\left(\frac{Y}{X}\right) - h(X)\right)\right)$$

III

$$\geq E\left(\left(Y - E\left(\frac{Y}{X}\right)\right)^2\right) + 0 + 0$$

13) Let X be an exponential r.v with parameter $\lambda > 0$. Using mgf, find μ, σ^2 of X

Ans $M_X(t) = E(e^{tX})$; $t \in (-\delta, \delta)$

X is an exponential r.v

$$\therefore \quad p_X(x) = \begin{cases} \lambda e^{-\lambda x} & x \geq 0 \\ 0 & x < 0 \end{cases}$$

$$\therefore E(e^{tx}) = \int_0^\infty e^{tx} \lambda e^{-\lambda x} \, dx$$

$$= \lambda \int_0^\infty e^{(t-\lambda)x} \, dx$$

$$= \frac{\lambda}{\lambda - t} \quad \text{for } t < \lambda$$

$$\mu = E(X')$$

$$= \mu'_{x}(0)$$

$$= \left. \frac{\lambda}{(\lambda - t)^2} \right|_{t=0} = \frac{1}{\lambda}$$

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

$$= \mu_x''(0) - \left(\frac{1}{\lambda}\right)^2$$

$$= \left. \frac{2\lambda}{(\lambda - t)^3} \right|_{t=0} - \frac{1}{\lambda^2}$$

$$= \frac{2}{\lambda^2} - \frac{1}{\lambda^2} = \frac{1}{\lambda^2}$$

$\therefore$ mean $= \frac{1}{\lambda}$, variance $= \frac{1}{\lambda^2}$

) For X a cont. r.v with pdf p_X, mgf M_X defined on $(-\delta, \delta)$, show that

$$\mathbb{P}(\{X \geq \alpha\}) \leq e^{-\alpha t} M_X(t) \quad \forall\, t \in (0, \delta)$$

Ans

$$M_X(t) = E(e^{Xt})$$

$$= \int_{-\infty}^{\infty} e^{xt} p_X(x)\, dx$$

$$= \int_{-\infty}^{\alpha} e^{xt} p_X(x)\, dx + \int_{\alpha}^{\infty} e^{xt} p_X(x)\, dx$$

$$\geq 0 + \int_{\alpha}^{\infty} e^{\alpha t} p_X(x)\, dx$$

$(\because e^{xt} > 0,\ p_X(x) > 0)$

$(\because \forall\, t \in (0, \delta),\ e^{xt} \geq e^{\alpha t}$ for $x \in [\alpha, \infty)$

$$= e^{\alpha t} \int_{\alpha}^{\infty} p_X(x)\, dx$$

$$= e^{\alpha t}\, \mathbb{P}(\{X \geq \alpha\})$$

$$\therefore \mathbb{P}(\{X \geq \alpha\}) \leq e^{-\alpha t} M_X(t)$$

15) Let X be an r.v with $M_X(t) = \dfrac{\delta^2}{\delta^2 - t^2}$, $t \in (-\delta, \delta)$

for some $\delta > 0$. Show that F_X is symmetric about origin

Ans $\mu = M_X'(0)$

$$= \left.\frac{2t\delta^2}{(\delta^2-t^2)^2}\right|_{t=0}$$

$$= 0$$

$E((X-\mu)^3)$

$= E(X^3)$

$= M_x'''(0)$

$= 0$

Since 3rd moment about mean is zero, f_X will be symmetric about the mean i.e. 0

claim: $E((X-\mu)^3) = 0 \Rightarrow F_X$ is symmetric about μ

$E((X-\mu)^3) = 0$

$$\Rightarrow \int_{-\infty}^{\infty} (x-\mu)^3 f(x)\,dx = 0$$

$$\Rightarrow \int_{-\infty}^{\infty} x^3 f(x+\mu)\,dx = 0$$

$$\Rightarrow \int_{-\infty}^{0} x^3 f(x+\mu)\,dx + \int_{0}^{\infty} x^3 f(x+\mu)\,dx = 0$$

$\Rightarrow \int_{\infty}^{0} (-u)^3 f(-u+\mu)\,(-du) + \int_{0}^{\infty} x^3 f(x+\mu)\,dx = 0$

$\Rightarrow \int_{0}^{\infty} x^3 f(x+\mu)\,dx = \int_{0}^{\infty} x^3 f(-x+\mu)\,dx$

$\Rightarrow f(x+\mu) = f(-x+\mu)$.

$\Rightarrow$ F_X is also symmetric about μ

b) Let X be normally distributed with $\mu = 0$, σ^2 variance. Find the characteristic function $\phi_X(t)$

Ans By definition,

$\phi_X(t) = E(e^{itX})$

$= \int_{-\infty}^{\infty} e^{itx}\, f_X(x)\,dx$

$= \int_{-\infty}^{\infty} e^{itx} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{x^2}{2\sigma^2}}\,dx$

$= \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{itx - \frac{x^2}{2\sigma^2}}\,dx$

$= \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\left[\frac{(x-\sigma^2 it)^2 + \sigma^4 t^2}{2\sigma^2}\right]}\,dx$

$= \frac{1}{\cancel{\sigma}\cancel{\sqrt{2\pi}}} \cancel{\sqrt{2\pi}}\, \cancel{\sigma}\, e^{-\frac{\sigma^2 t^2}{2}} = e^{-\frac{\sigma^2 t^2}{2}}$

17) Let $X_1, X_2, \ldots X_n$ be iid and $S_n = \sum_{k=1}^{n} X_n$. Show that $\phi_{S_n}(t) = (\phi_{X_1}(t))^n$. Also find pdf of S_n

Ans $\phi_{S_n}(t) = E(e^{itS_n})$

$$= E(e^{itX_1} e^{itX_2} \cdots e^{itX_n})$$

Since all X_i are iid & hence independent,

$$E(e^{itX_1} e^{itX_2} \cdots e^{itX_n})$$

$$= E(e^{itX_1})\, E(e^{itX_2}) \cdots E(e^{itX_n})$$

$$= \phi_{X_1}(t) \cdot \phi_{X_2}(t) \cdots \phi_{X_n}(t)$$

Since $X_1\ X_2 \ldots X_n$ are iid, they have the same distribution function and hence the same prob density function.

Thus, since $\phi_{X_i}(t) = \int_{-\infty}^{\infty} e^{itx}\, p_{X_i}(x)\, dx$,

~~which~~ will be the same for all i

$$\therefore \phi_{S_n}(t) = (\phi_{X_1}(t))^n$$

using the inversion function, the prob density function f of S_n can be written as

$$f(x) = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-itx}\, \phi_{S_n}(t)\, dt$$

$$= \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-itx}\, (\phi(t))^n\, dt$$

~~where g is the probability density~~

18) In a shopping mall, there are on an average 10k visitors per day. Find maximum probability that there will be at least 15k visitors on a fixed day, say tomorrow

Ans using Markov inequality,

$$\mathbb{P}(|X| > 15000) \le \frac{E(|X|)}{15000}$$

Note:

X denotes no of visitors

$\Rightarrow X = |X|$

$$= \frac{10000}{15000}$$

$$= \frac{2}{3}$$

19) The expected value of math test scores is 500, S.D = 100. If $\bar{x}$ is the mean of scores of random sample of 10 students, find lower bound for $\mathbb{P}(460 < \bar{x} < 540)$

Ans Using chebyshev,

$$P(|X-500| \geq K\cdot\sigma) \geq \frac{1}{K^2} \quad ; \sigma = 100$$

$$\Rightarrow P(|X-500| \geq 40) \geq \frac{1}{\left(\frac{40}{100}\right)^2} = \frac{100}{16}$$

But we have the data for 10 students

Ans we first find $E(\bar{x})$

$$E(\bar{x}) = E\left(\frac{x_1 + x_2 + \dots + x_n}{n}\right) = E(x_1) = \mu = 500$$

$(\because x_1, \dots, x_n$ all iid$)$

$$\text{var}(\bar{x}) = \text{var}\left(\frac{x_1 + \dots + x_n}{n}\right) = \frac{\sigma^2}{n} = \frac{(100)^2}{10} = 1000$$

$$\therefore P(|\bar{x} - 500| \geq 40) \leq \frac{1}{\left(\frac{40}{\sqrt{1000}}\right)^2}$$

$$= \frac{1000}{40 \times 40} = \frac{5}{8}$$

$$\therefore P(|\bar{x}-500| < 40) \geq 1 - \frac{5}{8} = \frac{3}{8}$$

20) Let $\bar{X}$ be a discrete R.V that takes values in $\{x_1, x_2, \ldots, x_n\}$.

$\bar{x} = \frac{1}{n} \sum_{k=1}^{n} x_k$, $s^2 = \frac{1}{n} \sum_{k=1}^{n} (x_k - \bar{x})^2$

Show that

$$P(\{X \in (\bar{x} - Ks, \bar{x} + Ks)\}) \geq 1 - \frac{1}{K^2}$$

Ans By definition, $\bar{x} = \mu$, $s^2 = \sigma^2$

$$\therefore P(|X - \mu| \geq K\sigma) \leq \frac{1}{K^2}$$

$$\therefore P(|X - \mu| < K\sigma)$$

$$= P(\{x \in (\mu - K\sigma, \mu + K\sigma)\})$$

$$\leq 1 - \frac{1}{K^2}$$

Exercises 4.4

1) Let $\{X_n\}$ be a seq of uniformly distributed r.v's on $(0, \frac{1}{n})$. Show that $X_n \to X_0$ in distribution as $n \to \infty$, where X_0 has distribution

$$F_0(x) = \begin{cases} 0 & x < 0 \\ 1 & 0 \leq x \end{cases}$$

Ans X_n is uniform on $(0, \frac{1}{n})$

$$\therefore F_n(x) = \begin{cases} 0 & x \leq 0 \\ nx & 0 \leq x \leq \frac{1}{n} \\ 1 & x \geq \frac{1}{n} \end{cases}$$

we need to verify that for every x,

$$F_n(x) \to F_0(x) \quad \text{as } n \to \infty$$

For $x \leq 0$,

$F_n(x) = 0 \quad \forall n = 1, 2, \ldots$

$F_0(x) = 0$

$\therefore \lim_{n \to \infty} F_n(x) = F_0(x)$

For $x \geq 1$,

$F_n(x) = 1 \quad \forall n = 1, 2, \ldots$

$F_0(x) = 1$

$\therefore \lim_{n \to \infty} F_n(x) = F_0(x)$

For some $x \in (0, 1)$, $\exists N_0$ s.t.

$\forall n > N_0, \ \frac{1}{n} \leq 1$

$\therefore \lim_{n \to \infty} F_n(x) = 1 = F_0(x)$

$\therefore \{F_n\} \xrightarrow{\omega} F_0$ (weakly)

Thus by definition, $X_n \to X_0$

2) Let $\{X_n\}$, X be r.v's on $(S, \mathcal{S}, P)$ with $E(X_n^2) < +\infty$ $\forall n = 1, 2, \ldots$, $E(X^2) < +\infty$.
If $E((X_n - X)^2) \to 0$ as $n \to \infty$, then show that $\{X_n\} \xrightarrow{P} X$ as $n \to \infty$.

<u>Ans</u> $E((X_n - X)^2) \to 0$ as $n \to \infty$

$\therefore X_n$ converges to X in moment of order 2

$$P(|X_n - X| \geq \varepsilon) = P(|X_n - X|^2 \geq \varepsilon^2)$$

$$\leq \frac{E(|X_n - X|^2)}{\varepsilon^2}$$

(By Chebyshev's inequality)

$$\therefore \lim_{n\to\infty} P(|X_n - X| \geq \varepsilon) \leq 0$$

$$\Rightarrow \lim_{n\to\infty} P(|X_n - X| \geq \varepsilon) = 0$$

$$\therefore X_n \xrightarrow{P} X$$

3) $X_n \xrightarrow{P} X$. Prove that $\forall \varepsilon > 0, \delta > 0$, $\exists N = N(\varepsilon, \delta)$ such that $\forall n, m \geq N$, $P(|X_n - X_m| \geq \varepsilon) < \delta$

$X_n \xrightarrow{P} X$ is given

Now $\mathbb{P}(|X_m - X_n| \geq \varepsilon)$

$$\leq \mathbb{P}\left(|X_m - X| \geq \frac{\varepsilon}{2}\right) + \mathbb{P}\left(|X_n - X| \geq \frac{\varepsilon}{2}\right)$$

can be made smaller than $\frac{\delta}{2}$ for $m \geq M$

can be made smaller than $\frac{\delta}{2}$ for $n \geq N$

$\therefore$ for $N_0 = \text{Max}\{M, N\} + 1$,

$\forall\, m, n > N_0$,

$$\mathbb{P}(|X_m - X_n| \geq \varepsilon) \leq \frac{\delta}{2} + \frac{\delta}{2} = \delta$$

4) $\{X_n\}_{n=1}^{n}$ is iid seq with mean 0, variance σ^2.

If third moment is finite, show that

$$\lim_{n \to \infty} E\left(\left(\frac{S_n}{\sigma\sqrt{n}}\right)^3\right) = 0 \quad \text{where}$$

$S_n = X_1 + \dots + X_n$

Ans from the CLT,

$$\lim_{n \to \infty} \mathbb{P}\left(\frac{S_n}{\sigma\sqrt{n}} \leq x\right) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{t^2}{2}}\, dt$$

But that doesnt apply here. LOL!

$$E(S_n^3) = E\left(\left(\sum_{i=1}^{n} X_i\right)^3\right)$$

$$= E\left(\sum_{i=1}^{n} X_i^3 + 3\sum_{1\le i<j\le n} X_i^2 X_j + 6\sum_{1\le i<j<k\le n} X_i X_j X_k\right)$$

$$= E\left(\sum_{i=1}^{n} X_i^3\right) + 0 + 0$$

$\left(\because E(X_i^2 X_j) = E(X_i^2)E(X_j) = 0 \right.$ since mean is $\left. 0 \right)$

$$= n\, E(X_1^3)$$

(all have same expectation)

$$\lim_{n\to\infty} \frac{n \cdot E(X_1^3)}{\sigma^3 \sqrt{n}\sqrt{n}\sqrt{n}} = \lim_{n\to\infty} \frac{E(X_1^3)}{\sigma^3\sqrt{n}} = 0$$

5) In 10000 tosses, find the approx. probability that no. of heads differs by less than 1% from 5000

<u>Ans</u> we want $\mathbb{P}(4950 < X < 5050)$

By the central limit theorem, we can approximate it using a normal distribution.

The normal distribution will have

$$\text{mean} = \mu = np = 10000 \times \frac{1}{2} = 5000$$

$$\text{and} \quad \sigma^2 = np(1-p) = 10000 \times \frac{1}{2} \times \frac{1}{2} = 2500$$

$$\therefore \sigma = 50$$

$$\therefore P(4950 < X < 5050)$$

$$\approx P(4950 < Y < 5050)$$

where X is the binomial distribution and

Y is the normal distribution

$$= P\left(\frac{4950-\mu}{\sigma} < \frac{Y-\mu}{\sigma} < \frac{5050-\mu}{\sigma}\right)$$

$$= P(-1 < Z < 1)$$

Now Z is the normal distribution which

is standard i.e. $\mu = 0$, $\sigma = \sigma^2 = 1$

$$\cancel{= 2P(Z < 1)}$$

$$\cancel{(\because Z \text{ is symmetric about } 0)}$$

$$\cancel{= 2 \times F_Z(1)}$$

$$\cancel{= 2 \times \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{1} e^{-\frac{t^2}{2}} dt}$$

$$\approx \cancel{1.6826}$$

$$= \frac{1}{\sqrt{2\pi}} \int_{-1}^{1} e^{-t^2/2} dt$$

$$\approx 0.68264$$

In-notes exercises

1.1) Show that the function P defined in the first definition satisfies the properties of theorem 1

Ans (i) $P(S) = \frac{|S|}{|S|} = 1$

(ii) $P(E^c) = \frac{|E^c|}{|S|} = \frac{|S| - |E|}{|S|} = 1 - P(E)$

(iii) $P(E_1 \cup E_2) = \frac{|E_1 \cup E_2|}{|S|} = \frac{|E_1| + |E_2|}{|S|}$ (E_1, E_2 are disjoint)

$= P(E_1) + P(E_2)$

(iv) $P(E) = \frac{|E|}{|S|} \leq \frac{|S|}{|S|} \leq 1$ and $\frac{|E|}{|S|} \geq 0$

1.2) Prove theorem 2

Ans (i) Choose $E_3 = E_4 = E_5 = \dots = \phi$ $(\phi \in \mathcal{A})$

$$\therefore \bigcup_{i=1}^{\infty} E_i = E_1 \cup E_2 \cup \bigcup_{i=3}^{\infty} E_i$$

$$= E_1 \cup E_2 \cup \phi$$

$$= E_1 \cup E_2 \in \mathcal{A} \text{ (by property iii)}$$

(ii) $\phi \in \mathcal{A} \Rightarrow \phi^c \in \mathcal{A} \Rightarrow S \in \mathcal{A}$

(iii) $E_n \in \mathcal{A} \quad \forall n \in \mathbb{N}$

$\Rightarrow E_n^c \in \mathcal{A} \quad \forall n \in \mathbb{N}$

$\Rightarrow \bigcup_{i=1}^{\infty} E_i^c \in \mathcal{A}$

$\Rightarrow \left(\bigcup_{i=1}^{\infty} E_i^c\right)^c \in \mathcal{A}$

$\Rightarrow \bigcap_{i=1}^{\infty} E_i \in \mathcal{A}$

1.3) Define a collection of subsets of S, $\mathcal{A}$ as a field if

(i) $\phi \in \mathcal{A}$

(ii) $E \in \mathcal{A} \Rightarrow E^c \in \mathcal{A}$

(iii) $E_1, E_2 \in \mathcal{A} \Rightarrow E_1 \cup E_2 \in \mathcal{A}$

Prove that every σ-field is a field. Analyse the converse also

Ans Let $\mathcal{A}$ be a σ-field

$\phi \in \mathcal{A}$

$E \in \mathcal{A} \Rightarrow E^c \in \mathcal{A}$ } defn of σ-field

$E_1, E_2 \in \mathcal{A} \Rightarrow E_1 \cup E_2 \in \mathcal{A}$ (prev problem)

$\therefore$ $\mathcal{A}$ is a field

Converse is not true. Let $S = [0,1]$

Consider $I = \{ (a,b] \mid 0 \le a < b \le 1 \}$

Consider $M = \{ \bigcup_{\lambda} I_\lambda \mid \lambda$ is a finite indexing set and $I_\lambda \in I \}$

Then M forms a field which is not a σ-field

$\phi \in M$ (don't choose any λ i.e. $\lambda = \phi$)

$E \in M \Rightarrow E^c \in M$

since if $E = (a_1, b_1] \cup \dots\dots \cup (a_m, b_m]$,

then $E^c = (0, a_1] \cup (b_1, a_2] \cup \dots \cup (b_m, 1]$

$\in M$

$E_1, E_2 \in M \Rightarrow E_1 \cup E_2 \in M$

since if $E_1 = (a_1, b_1] \cup \dots \cup (a_m, b_m]$

$E_2 = (c_1, d_1] \cup \dots \cup (c_n, d_n]$

we can algorithmically evaluate $E_1 \cup E_2$

and express it as $(p_1, q_1] \cup \dots \cup (p_k, q_k]$

But M is not a σ-field since

$(0, 1-\frac{1}{n}] \in M \quad \forall n = 2, 3, \dots$

but $\bigcup_{n=2}^{\infty} (0, 1-\frac{1}{n}] = (0,1) \notin M$

1.4) Prove $\mathbb{P}(E^c) = 1 - \mathbb{P}(E)$

<u>Ans</u> We know $\mathbb{P}(E_2 \setminus E_1) = \mathbb{P}(E_2) - \mathbb{P}(E_1)$ if $E_1 \subseteq E_2$

$\Rightarrow \mathbb{P}(S \setminus E) = \mathbb{P}(S) - \mathbb{P}(E)$

for $E \subseteq S$, $E, S \in \mathcal{A}$

$\Rightarrow \mathbb{P}(E^c) = 1 - \mathbb{P}(E)$

1.5) Prove Bonferroni's inequality (Theorem 7 (v), (vi))

<u>Ans</u> we first prove Boole's inequality

ie. $\mathbb{P}\left(\bigcup_{k=1}^{n} E_k\right) \le \sum_{k=1}^{n} \mathbb{P}(E_k)$

For $n=1$, it is trivially true with equality

Assume it to be true for a general $n = r$

we prove it for $n = r+1$

$\mathbb{P}\left(\bigcup_{k=1}^{r} E_k \cup E_{r+1} \right)$

$= \mathbb{P}\left(\bigcup_{k=1}^{r} E_k \right) + \mathbb{P}(E_{r+1}) - \mathbb{P}\left(\left(\bigcup_{k=1}^{r} E_k\right) \cap E_{r+1} \right)$

(by theorem 7 (iii))

$\leq \mathbb{P}\left(\bigcup_{k=1}^{r} E_k \right) + \mathbb{P}(E_{r+1})$

$(\because \mathbb{P}(E) \geq 0 \quad \forall E \in \mathcal{A})$

$= \sum_{k=1}^{r+1} \mathbb{P}(E_k)$ and hence we are

done by mathematical induction.

Now we prove the other side ie.

$$\sum_{k=1}^{n} \mathbb{P}(E_k) - \sum_{1\leq i<j\leq n} \mathbb{P}(E_i \cap E_j) \leq \mathbb{P}\left(\bigcup_{k=1}^{n} E_k \right)$$

It is trivially true with equality for $n=1$

Assume true for $n = r$

Now, $\mathbb{P}\left(\bigcup_{k=1}^{r} E_k \cup E_{r+1} \right)$

$= \mathbb{P}\left(\bigcup_{k=1}^{r} E_k \right) + \mathbb{P}(E_{r+1}) - \mathbb{P}\left(\bigcup_{k=1}^{r} E_k \cap E_{r+1} \right)$

$\geq \sum_{k=1}^{r+1} \mathbb{P}(E_k) - \sum_{1\leq i<j\leq r} \mathbb{P}(E_i \cap E_j) - \mathbb{P}(\;\; " \;\;)$

$= \sum_{k=1}^{n+1} \mathbb{P}(E_k) - \sum_{1 \le i < j \le n+1} \mathbb{P}(E_i \cap E_j)$

1.6) $(S, \mathcal{A}, \mathbb{P})$ is a prob space and $G \in \mathcal{A}$ with $\mathbb{P}(G) > 0$.

Show that $(S, \mathcal{A}, \mathbb{P}_G)$ is a prob space where

$\mathbb{P}_G(E) = \mathbb{P}(E \mid G) \quad \forall E \in \mathcal{A}$

<u>Ans</u> we just verify that $\mathbb{P}_G$ is a probability measure.

$$\mathbb{P}_G(E) = \mathbb{P}(E \mid G) = \frac{\mathbb{P}(E \cap G)}{\mathbb{P}(G)} \geq 0$$

$(\because \mathbb{P}(G) > 0)$

$$\mathbb{P}_G(S) = \mathbb{P}(S \mid G) = \frac{\mathbb{P}(S \cap G)}{\mathbb{P}(G)} = \frac{\mathbb{P}(G)}{\mathbb{P}(G)} = 1$$

$$\mathbb{P}\Big(\underbrace{\bigcup_{n=1}^{\infty} E_n}_{\text{mutually disjoint}}\Big) = \frac{\mathbb{P}\left(\bigcup_{n=1}^{\infty} (E_n \cap G)\right)}{\mathbb{P}(G)}$$

if $E_1, E_2, \ldots$ are mutually disjoint then so are $E_1 \cap G$, $E_2 \cap G$, ...

hence we get $\displaystyle \sum_{n=1}^{\infty} \frac{\mathbb{P}(E_n \cap G)}{\mathbb{P}(G)}$

$$= \sum_{n=1}^{\infty} \mathbb{P}(E_n \mid G)$$

7) Prove that for a sequence of events $\{E_n\}$ show that

$$\limsup_{n\to\infty} E_n = \lim_{n\to\infty} \bigcup_{k\geq n} E_k$$

$$\liminf_{n\to\infty} E_n = \lim_{n\to\infty} \bigcap_{k\geq n} E_k$$

Ans we prove only the first one. The second one follows directly/similarly

$$\limsup_{n\to\infty} E_n = \bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} E_k$$

$$\bigcup_{k=1}^{\infty} E_k \supseteq \bigcup_{k=2}^{\infty} E_k \supseteq \bigcup_{k=3}^{\infty} E_k \supseteq \cdots\cdots$$

$$\lim_{n\to\infty} \left(\bigcup_{k\geq n} E_k \right) = \limsup_{n\to\infty} \left(\bigcup_{k\geq n} E_k \right)$$

(by defn of limit of sequence of sets)

$$= \bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} \bigcup_{k\geq n} E_k$$

$$= \bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} E_k$$

$$= \limsup_{n\to\infty} E_n$$

PD

1.8) Show that $\sum_{k=0}^{n} \mathbb{P}(H_k) = 1$ where $\mathbb{P}$ is the binomial probability and H_k is event that all n tuples consist of exactly k- heads

Ans $\mathbb{P}(H_k) = \binom{n}{k} p^k (1-p)^{n-k}$

$\therefore \sum_{k=0}^{n} \binom{n}{k} p^k (1-p)^{n-k} = 1$ by the binomial theorem

2.1) For a real valued $X : S \to \mathbb{R}$ (S being non empty), show that $\sigma(X) = \{ X^{-1}(B) \mid B \in \mathcal{B}(\mathbb{R}) \}$ is a σ-field

Ans (i) $\emptyset \in \mathcal{B}(\mathbb{R}) \Rightarrow X^{-1}(\emptyset) = \emptyset \in \sigma(X)$

(ii) Let $E \in \sigma(X)$

$\therefore X(E) \in \mathcal{B}(\mathbb{R})$

and $E = X^{-1}(B)$ for some borel set B

$\Rightarrow E^c = (X^{-1}(B))^c$

$= X^{-1}(B^c)$

B is borel set $\Rightarrow B^c$ is borel set

$\Rightarrow E^c \in \sigma(X)$

(iii) $E_n \in \sigma(X)\ \forall n$ (say)

Then $E_i = X^{-1}(B_i)$ for borel sets B_i.

$$\bigcup_{i=1}^{\infty} E_i = \bigcup_{i=1}^{\infty} X^{-1}(B_i)$$

$$= X^{-1}\left(\bigcup_{i=1}^{\infty} B_i\right)$$

$$= X^{-1}(C) \quad \text{for borel measurable } C$$

$$\therefore \bigcup_{i=1}^{\infty} E_i \in \sigma(X)$$

$\therefore \sigma(X)$ is a σ-field

2.2) Prove that for a random variable X wrt $(S, \mathcal{A})$,

(i) $\{X = a\}$

(ii) $\{X < a\}$

(iii) $\{a < X \le b\}$

(iv) $\{a \le X < b\}$

(v) $\{a < X < b\}$

(vi) $\{a \le X \le b\}$

all belong to $\mathcal{A}$

Ans (i) $\{X = a\} = \{s \in S \mid X(s) = a\}$

X is an r.v, $B = \{a\}$ is a borel set

$\therefore \{s \in S \mid X(s) \in B\} \in \mathcal{A}$ by definition of a random variable

$\therefore \{X = a\} \in \mathcal{A}$

(ii) Choose $B = [a, b)$

(iii) Choose $B = (-\infty, a)$

(iv) $B = [a, b)$ (vi) $[a, b]$

(v) (a, b)

2.3) Prove the following.

(i) $\mathbb{P}(\{a < X \leq b\}) = F_X(b) - F_X(a)$

(ii) $\mathbb{P}(\{a < X < b\}) = F_X(b^-) - F_X(a)$

(iii) $\mathbb{P}(\{a \leq X < b\}) = F_X(b^-) - F_X(a^-)$

(iv) $\mathbb{P}(\{a \leq X \leq b\}) = F_X(b) - F_X(a^-)$

(v) $\mathbb{P}(\{X > b\}) = 1 - F_X(b)$

Ans (i) $\mathbb{P}(\{a < X \leq b\})$

$= \mathbb{P}(\{X \leq b\} \setminus \{X \leq a\})$ $\left(\because a < b \Rightarrow \{X \leq a\} \subseteq \{X \leq b\}\right)$

$= \mathbb{P}(X \leq b) - \mathbb{P}(X \leq a)$

$= F_X(b) - F_X(a)$

(ii) $\mathbb{P}(X < b) - \mathbb{P}(X \leq a)$

$= F_X(b^-) - F_X(a)$

(iii) $\mathbb{P}(X < b) - \mathbb{P}(X < a)$

$= F_X(b^-) - F_X(a^-)$

(iv) $\mathbb{P}(X \leq b) - \mathbb{P}(X < a)$

$= F_X(b) - F_X(a^-)$

(v) $\mathbb{P}(X > b) = 1 - \mathbb{P}(X \leq b)$

$= 1 - F_X(b)$

2.4) Let $\{p_k\}$ be a sequence of real numbers such that $\sum_{n=1}^{\infty} p_i = 1$ and $\{x_n\}$ be a seq of reals.

Show that $\exists$ $(S, \mathcal{S}, \mathbb{P})$ and X on $(S, \mathcal{S})$ such that

$$f(x) = \begin{cases} p_k & x = x_k \quad k = 1, 2, \dots \\ 0 & \text{otherwise} \end{cases}$$

is the pmf of X

<u>Ans</u> Construct $S = \{x_1, x_2, \cdots\}$

$\mathcal{S} = \mathcal{S}_* = 2^S$

$\mathbb{P}$ is defined as follows:

$$\mathbb{P}(E) = \sum_{x_i \in E} p_i$$

This is the same as discrete prob space discussed before theorem 3

Now define $X : S \to \mathbb{R}$ as

$$X(x_i) = x_i$$

Then clearly

$$f(x_k) = \mathbb{P}(\{X = x_k\})$$

$$= \mathbb{P}(\{x_k\})$$

$$= x_k \quad \forall \; k = 1, 2, \dots$$

and since X takes values only $x_1, x_2, \cdots$, outside this it will be 0 $\blacksquare$

2.5) Prove that $(-1)^x \binom{-\alpha}{x} = \binom{\alpha+x-1}{x}$

Ans By definition,

$$\binom{-\alpha}{x} = \frac{(-\alpha)(-\alpha-1)(-\alpha-2)\cdots(-\alpha-x+1)}{x!}$$

$$= (-1)^x \frac{(\alpha)(\alpha+1)(\alpha+2)\cdots(\alpha+x+1)}{x!}$$

$$= (-1)^x \frac{(\alpha+x-1)!}{(\alpha-1)!\, x!}$$

$$\therefore \binom{-\alpha}{x} = (-1)^x \binom{\alpha+x-1}{x}$$

$$\therefore (-1)^x \binom{-\alpha}{x} = \binom{\alpha+x-1}{x} \cancel{(-1)^{2x}} \rightarrow 1$$

2.6) show that $\lim_{n\to\infty} P_{X_n}(x) = \frac{e^{-\lambda}\lambda^x}{x!}$ for

all $x = 0, 1, 2, \ldots$ where $p_n \in (0,1)\ \forall n$,

$\lim_{n\to\infty} p_n = 0$, $\lim_{n\to\infty} np_n = \lambda$, P_{X_n} is binomial

distribution with parameters (n, p_n)

Ans
$$P_{X_n}(x) = \begin{cases} \binom{n}{x} p_n^x (1-p_n)^{n-x} & x = 0, 1, 2, \ldots \\ 0 & \text{otherwise} \end{cases}$$

we only wish to analyse it for $x = 0, 1, 2, \ldots$

$$\therefore \lim_{n\to\infty} \binom{n}{x} p_n^{\,x} (1-p_n)^{n-x}$$

$$= \lim_{n\to\infty} \frac{n!}{x!\,(n-x)!} p_n^{\,x} (1-p_n)^{n-x}$$

$$= \frac{1}{x!} \lim_{n\to\infty} \frac{n!\, p_n^{\,x} (1-p_n)^{n}}{(n-x)!\,(1-p_n)^{x}}$$

$$= \frac{1}{x!} \lim_{n\to\infty} \frac{n!}{(n-x)!} p_n^{\,x} (1-p_n)^{n}$$

$$= \frac{1}{x!} \lim_{n\to\infty} n(n-1)\cdots(n-x+1)\, p_n^{\,x} (1-p_n)^{n}$$

$$= \frac{1}{x!} \lim_{n\to\infty} (n p_n)^{x} \left(1-\frac{1}{n}\right)\left(1-\frac{2}{n}\right)\cdots\left(1-\frac{(x-1)}{n}\right)(1-p_n)^{n}$$

$$= \frac{1}{x!} \lim_{n\to\infty} (n p_n)^{x} \lim_{n\to\infty} (1-p_n)^{n} \times 1$$

$$= \frac{1}{x!} \lambda^{x} e^{-\lambda} = \frac{e^{-\lambda}\lambda^{x}}{x!}$$

2-7) Prove that X is a continuous r.v iff F_X is continuous

<u>Ans</u> X is continuous

$\Rightarrow \exists\, p_X(x)$ s.t. $F_X(x) = \int_{-\infty}^{x} p_X(t)\,dt \;\; \forall x$

But for integrals, we know that $p_X(x) = F_X'(x)$

and hence F_X is differentiable and hence continuous

F_X is continuous

By definition F_X is increasing

$\therefore F_X$ is differentiable almost everywhere and hence

$$F_X(x) = \int_{-\infty}^{x} g(t)\, dt \quad \text{for some function } g$$

and thereby, we have that X is continuous.

Simpler approach:

$$F(x_-) = \lim_{y<x,\, y\to x} F(y)$$

F is continuous at x iff $F(x) = F(x_-)$ since F is anyway right continuous

$$F(x_-) = P(X < x)$$

$$\therefore F(X) - F(X_-) = P(\{X = x\})$$

$\therefore F$ is continuous iff $P(\{X = x\}) = 0 \ \forall x$

iff X is continuous by definition.

2.8) Show that $F(x) = \int_{-\infty}^{x} p(t)\, dt$ is a dist func (if p is a density function)

Ans $F(x_-) = \lim_{y<x,\, y\to x} F(y)$

$$= \lim_{h\to 0^+} \int_{-\infty}^{x-h} p(t)\, dt = \int_{-\infty}^{x} p(t)\, dt$$

$\therefore F$ is right continuous

for $a < b$,

$\int_a^b p(t)\, dt > 0$ since p is non negative

Also, $\lim_{x \to -\infty} F(x) = 0$

$\lim_{x \to +\infty} F(x) = \int_{-\infty}^{\infty} p(t) = 1$

($\because p$ is density funct)

$\therefore F$ is a dist func

2.9) Show that an absolutely continuous r.v is cont

Ans Let X be absolutely continuous

$\therefore \exists\ p_X$ such that

$F_X(x) = \int_{-\infty}^{x} p_X(t)\, dt$

$\int_{-\infty}^{x} p_X(t) dt$ is continuous.

$\Rightarrow F_X(x)$ is continuous

$\Rightarrow X$ is continuous (2.7)

2.10) Show that every continuous function is Borel measurable

Ans (X, τ), (Y, β) are topological spaces. σ_1, σ_2 are the respective Borel-σ-algebras

Let $f : X \to Y$ be a continuous map

$\forall\, S \in \beta,\ f^{-1}(\beta) \in \tau$ since f is continuous

By definition, the σ-algebra $\sigma_1 = \sigma(\tau)$, the generated σ-algebra and hence $\tau \subset \sigma_1$

$$\Rightarrow f^{-1}(\beta) \in \tau \subseteq \sigma_1$$

mapping is measurable iff it is measurable on the generator and hence

f is measurable mapping

2.11) Let X be a continuous rv with pdf p_X. For $a, b \in \mathbb{R}$ and $b \neq 0$, show that the pdf of $Y = a + bX$ is given by $p_Y(y) = \frac{1}{|b|} p_X\left(\frac{y-a}{b}\right),\ y \in \mathbb{R}$

Ans $\phi : \mathbb{R} \to \mathbb{R}$ is defined by, say,

$$\phi(t) = a + bt$$

Then $\phi(X) = Y$

i.e. $Y = \phi \circ X$

ϕ is differentiable everywhere and hence $I = \mathbb{R}$, $\phi(I) = \phi(\mathbb{R}) = \mathbb{R}$

$$\phi^{-1}(y) = \frac{y-a}{b},\ (\phi^{-1}(y))' = \frac{1}{b}$$

$\therefore p_Y(y) = p_X\left(\frac{y-a}{b}\right) \times \left|\frac{1}{b}\right| \quad \forall \ y \in \phi(I) = \mathbb{R}$

$\therefore p_Y(y) = \frac{1}{|b|} \ p_X\left(\frac{y-a}{b}\right) \ , \ y \in \mathbb{R}$

2.12) missing ??

2.13) X and Y are independent iff $f_{\underline{X}}(x,y) = f_X(x) f_Y(y)$
$\forall (x,y) \in \mathbb{R}^2$ where $\underline{X} = (X, Y)$ is an r. vec.

Ans we have shown (or at least know that) X is indep iff $F_{\underline{X}}(x,y) = F_X(x) F_Y(y)$

Now $f_{\underline{X}}(x,y) = f_X(x) f_Y(y)$

$\Leftrightarrow \frac{\partial^2}{\partial x \partial y} F_{\underline{X}}(x,y) = \frac{d F_X(x)}{dx} \frac{d}{dy} F_Y(y)$

$\Leftrightarrow F_{\underline{X}}(x,y) = F_X(x) \cdot F_Y(y)$

$\Leftrightarrow$ X, Y are independent

2.14) Show that X, Y are independent iff

$f_{Y|X}(y|x) = f_Y(y) \ ; \ 0 < f_X(x) < +\infty, \ y \in \mathbb{R}$

Ans $f_{Y|X}(y|x) = f_Y(y)$

$\Leftrightarrow \frac{f_{\underline{X}}(x,y)}{f_X(x)} = f_Y(y) \quad (\because 0 < f_X(x))$

$\Leftrightarrow \; f_{\underline{X}}(x,y) = f_X(x)\, f_Y(y)$

$\Leftrightarrow$ X, Y are independent.

2.15) missing ??

2.16) Let X, Y be independent r.v's with the same pdf $f(x) = \begin{cases} e^{-x} & x > 0 \\ 0 & x \leq 0 \end{cases}$

Find joint pdf of $\left(X+Y, \frac{X}{Y} \right)$

<u>Ans</u> $h_1 : \mathbb{R}^2 \to \mathbb{R}$

as $h_1(x,y) = x+y$

$h_2 : \mathbb{R}^2 \setminus \{(t,0) : t \in \mathbb{R}\} \to \mathbb{R}$ as

$h_2(x,y) := \frac{x}{y}$

Then, $U = h_1(X,Y) = X+Y$

$V = h_2(x,y) = \frac{X}{Y}$

$$\left.\begin{aligned} u &= x+y \\ v &= \frac{x}{y} \end{aligned}\right\} \Rightarrow \quad \begin{aligned} x &= \frac{uv}{v+1} \\ y &= \frac{u}{v+1} \end{aligned}$$

$$J = \det \begin{pmatrix} \frac{v}{v+1} & \frac{u}{(v+1)^2} \\ \frac{1}{v+1} & \frac{-u}{(v+1)^2} \end{pmatrix}$$

$$\therefore J = \frac{-uv}{(v+1)^3} - \frac{u}{(v+1)^3}$$

$$= -\frac{u}{(v+1)^2}$$

Q is set of points where $J \neq 0$

which is $\mathbb{R}^2 \setminus$ x axis

$$\therefore p_{\underline{U}}(u,v) = \begin{cases} p_X\left(\frac{uv}{v+1}, \frac{u}{v+1}\right) \left|\frac{u}{(v+1)^2}\right| & u, v \in Q \\ 0 & \text{otherwise} \end{cases}$$

$$= \begin{cases} e^{-\frac{uv}{v+1}} e^{-\frac{u}{v+1}} \cdot \frac{|u|}{(v+1)^2} & \frac{uv}{v+1} > 0, \frac{u}{v+1} > 0 \\ 0 & \text{otherwise} \end{cases}$$

(But $\frac{uv}{v+1} > 0$, $\frac{u}{v+1} > 0 \Rightarrow u > 0, v > 0$)

2.17) Not in syllabus but anyways...

Find sample range R's pdf for $X_1, \ldots, X_n$ being independent & uniform distributed r.v's on $(0,1)$

Ans $$f_R(r) = \begin{cases} n(n-1) \int_{-\infty}^{\infty} (F(r+x) - F(x))^{n-2} f(r+x) f(x)\, dx & (r > 0) \\ 0 & (\text{otherwise}) \end{cases}$$

For $r \geq 0$,

$$f_R(r) = n(n-1) \int_{-\infty}^{\infty} \left(F(r+x) - F(x)\right)^{n-2} f(r+x) f(x)\, dx$$

Now $$f(x) = \begin{cases} 1 & 0 \leq x \leq 1 \\ 0 & \text{otherwise} \end{cases}$$

and $$F(x) = \begin{cases} 0 & x < 0 \\ x & 0 \leq x \leq 1 \\ 1 & x > 1 \end{cases}$$

$$\therefore f_R(r) = n(n-1) \int_0^1 \left(F(r+x) - F(x)\right)^{n-2} f(r+x) f(x)\, dx$$

$$= n(n-1) \int_0^{1-r} (r+x-x)^{n-2} \cdot 1 \cdot 1 \; dx \quad \text{for } r \leq 1$$

$$= n(n-1) \times r \times (1-r) \quad \text{for } r \leq 1$$

$$\therefore f_R(r) = \begin{cases} n\, r\, (n-1)(1-r) & r \in [0,1] \\ 0 & \text{otherwise} \end{cases}$$

3.1) show that $E(X) = \int_{-\infty}^{\infty} x \, dF_X(x)$ for an absolutely continuous r.v

Ans

$$\int_{-\infty}^{\infty} x \, dF_X(x)$$

$$= \int_{-\infty}^{\infty} x \, F_X'(x) \, dx$$

$$= \int_{-\infty}^{\infty} x \, p_X(x) \, dx$$

$$= E(X)$$

3.2) If X has a moment of order α, then it has all moments of order $1, 2, \ldots, \alpha - 1$

Ans Let $g: \mathbb{R} \to \mathbb{R}$ as $g(x) = |x|^r + 1$ and

$h: \mathbb{R} \to \mathbb{R}$ as $h(x) = |x|^s$

Then for all $r \geq s$, $|x|^r \geq |x|^s$ and

hence $E(g(X)) \geq E(h(X))$

since

$$E(g(X)) = \int_{\mathbb{R}} |x|^r + 1 \, dF$$

$$= \int_{\mathbb{R}} |x|^r dF + \int_{\mathbb{R}} 1 \cdot dF > \int_{\mathbb{R}} |x|^s dF$$

$\therefore \infty > E(|X|^r) + 1 > \int_{\mathbb{R}} |x|^s dF = E(|X|^s)$

$\therefore E(|X|^s)$ also exists

Hence $E(X^s)$ also exists since it is lower than $E(|X|^s)$

3.3) Show that variance is always positive for a continuous r.v

Ans $\sigma^2 = \int_{-\infty}^{\infty} (x-\mu)^2 \, p_X(x) \, dx$

> 0

since $(x-\mu)^2 > 0$

$p_X(x)$ is non negative

$\therefore \sigma^2 \geq 0$

But if $\sigma^2 = 0$, then $\int_{-\infty}^{\infty} (x-\mu)^2 p_X(x) dx$

$= 0 \Rightarrow (x-\mu)^2 p_X(x)$ is overall 0

in $-\infty, \infty$ which is not possible since

$\int_{-\infty}^{a} p_X(x) = 1$

4.1) $X_n \xrightarrow{a.s} X \Rightarrow X_n \xrightarrow{P} X$ (prove it)

$X_n \xrightarrow{a.s.} X$

Ans

Let $A = \{ \lim_{n\to\infty} X_n = X \}$

$= \{ s \in S \mid \lim_{n\to\infty} X_n(s) = X(s) \}$

$= \{ s \in S \mid \forall \varepsilon > 0 \ \exists m \in \mathbb{N}$ s.t. $|X_n(s) - X(s)| < \varepsilon \ \forall n \geq m \}$

$\subseteq \bigcap_{\varepsilon > 0} \bigcup_{n=1}^{\infty} \bigcap_{k=n}^{\infty} \{ |X_k - X| < \varepsilon \}$

$\subseteq \bigcup_{n=1}^{\infty} \bigcap_{k=n}^{\infty} \{ |X_k - X| < \varepsilon \} \ \forall \varepsilon > 0$

$\mathbb{P}(A) = 1$ (given)

$\therefore 1 \leq \mathbb{P}\left(\bigcup_{n=1}^{\infty} \bigcap_{k=n}^{\infty} \{ |X_k - X| < \varepsilon \} \right) \ \forall \varepsilon > 0$

$\therefore \mathbb{P}\left(\bigcup_{n=1}^{\infty} \bigcap_{k=n}^{\infty} \{ |X_k - X| < \varepsilon \} \right) = 1$

$= \mathbb{P}\left(\bigcup_{n=1}^{\infty} \inf_{k \geq n} \{ |X_k - X| < \varepsilon \} \right)$

$\therefore \mathbb{P}\left(\liminf_{n\to\infty} \{ |X_n - X| < \varepsilon \} \right) = 1$

By Fatou's lemma,

$\mathbb{P}\left(\limsup_{n\to\infty} \{ |X_n - X| < \varepsilon \} \right) = 1$

$\therefore \mathbb{P}\left(\lim_{n\to\infty} \{ |X_n - X| < \varepsilon \} \right) = \lim_{n\to\infty} \mathbb{P}(\{ |X_n - X| < \varepsilon \})$

$= 1$

Finally,

$$\lim_{n\to\infty} \mathbb{P}(\{|X_n - X| < \varepsilon\}) = 1 \quad \forall \varepsilon > 0$$

$$\therefore \lim_{n\to\infty} \mathbb{P}(\{|X_n - X| \geq \varepsilon\}) = 0 \quad \forall \varepsilon > 0$$

$$\therefore X_n \xrightarrow{\mathbb{P}} X$$

The nontrivial step was showing that

$$A \subseteq \bigcap_{\varepsilon>0} \bigcup_{n=1}^{\infty} \bigcap_{k=n}^{\infty} \{|X_k - X| < \varepsilon\}$$

<u>Alternate method:</u>

$$X_n \xrightarrow{a.s.} X$$

$\therefore$ $A = \{s \mid \lim_{n\to\infty} X_n(s) \neq X(s)\}$ has

measure 0 (or probability 0)

For a fixed $\varepsilon > 0$,

$$B_n = \bigcup_{m=n}^{\infty} \{|X_m - X| > \varepsilon\}$$

$$B_n \supseteq B_{n+1} \supseteq \cdots \supseteq B_\infty$$

where $B_\infty = \bigcap_{n=1}^{\infty} \bigcup_{m=n}^{\infty} \{|X_m - X| > \varepsilon\}$

$\lim_{n\to\infty} \mathbb{P}(B_n) = \mathbb{P}(B_\infty)$ (convergence from above)

for all $s \notin A$, $\lim_{n\to\infty} X_n(s) = X(s)$ and

hence $|X_n(s) - X(s)| < \varepsilon \ \forall\ n \geq N$ for some N

$\therefore s \notin B_n \Rightarrow s \notin B_\infty$

$\therefore B_\infty \subseteq A$

$\Rightarrow \mathbb{P}(B_\infty) \leq \mathbb{P}(A) = 0$

$\Rightarrow \lim_{n\to\infty} \mathbb{P}(B_n) = \mathbb{P}(B_\infty) = 0$

$\therefore \mathbb{P}(|X_n - X| > \varepsilon) \leq \mathbb{P}(B_n) \xrightarrow[n\to\infty]{} 0$

4.2) Prove that $X_n \xrightarrow{P} X \Rightarrow X_n \xrightarrow{d} X$

Ans lemma:

$\mathbb{P}(Y \leq \alpha) \leq \mathbb{P}(X \leq \alpha + \varepsilon) + \mathbb{P}(|Y - X| > \varepsilon)$

proof of lemma:

$\mathbb{P}(Y \leq \alpha) = \mathbb{P}(Y \leq \alpha, X \leq \alpha + \varepsilon)$
$+ \mathbb{P}(Y \leq \alpha, X > \alpha + \varepsilon).$

$\le \mathbb{P}(X \le \alpha+\varepsilon) + \mathbb{P}(Y-X \le \alpha - X, \alpha - X < -\varepsilon)$

$\le \mathbb{P}(X \le \alpha+\varepsilon) + \mathbb{P}(Y-X < -\varepsilon)$

$\le \mathbb{P}(X \le \alpha+\varepsilon) + \mathbb{P}(Y-X < -\varepsilon) + \mathbb{P}(Y-X > \varepsilon)$

$= \mathbb{P}(X \le \alpha+\varepsilon) + \mathbb{P}(|Y-X| > \varepsilon)$

Let a be a random point in $\mathbb{R}$ for which F_X is continuous. We want to show $F_n(a) \to F(a)$ as $n \to \infty$

we have $\forall \varepsilon > 0$,

$$\mathbb{P}(X_n \le a) \le \mathbb{P}(X \le a+\varepsilon) + \mathbb{P}(|X_n - X| > \varepsilon)$$

and $\mathbb{P}(X \le a-\varepsilon) \le \mathbb{P}(X_n \le a) + \mathbb{P}(|X_n - X| > \varepsilon)$

$\therefore$ $\mathbb{P}(X \le a-\varepsilon) - \mathbb{P}(|X_n - X| > \varepsilon)$

$\le \mathbb{P}(X_n \le a)$

$\le \mathbb{P}(X \le a+\varepsilon) + \mathbb{P}(|X_n - X| > \varepsilon)$

$$\Rightarrow F_X(a-\varepsilon) \le \lim_{n\to\infty} \mathbb{P}(X_n \le a) \le F_X(a+\varepsilon)$$

(taking $\lim_{n\to\infty}$)

F_X is continuous at a by assumption

$\therefore$ $F_X(a-\varepsilon)$, $F_X(a+\varepsilon) \to F_X(a)$ as $\varepsilon \to 0^+$

$$\therefore \lim_{n \to \infty} \mathbb{P}(X_n \le a) = \mathbb{P}(X \le a) = F_X(a)$$

$$\therefore \lim_{n \to \infty} F_{X_n}(a) = F_X(a) .$$

$$\therefore X_n \xrightarrow{d} X$$