Random Variables

A random variable, usually written X, is a variable whose possible values are numerical outcomes of a random phenomenon. There are two types of random variables, discrete and continuous.

A discrete random variable is one which may take on only a countable number of distinct values such as 0, 1, 2, 3, 4 ..... Discrete random variables are usually (but not necessarily) counts. If a random variable can take only a finite number of distinct values, then it must be discrete. Examples of discrete random variables include the number of children in a family, ~~the~~ the Friday night attendance in a cinema, the number of patients in a doctor's surgery, the number of defective light bulbs in a box of ten.

The probability distribution of a discrete random variable is a list of probabilities associated with each of its possible values. It is also sometimes called the probability function or the probability mass function.

Suppose a random variable X may take k different values, with probability that $X=x_i$ defined to be $P(X=x_i)=p_i$. The probabilities p_i must satisfy the following:

(i) $0 \leq p_i \leq 1$ for each i

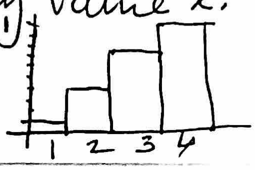
(ii) $p_1 + p_2 + \cdots + p_k = 1$.

ex: Suppose a random variable X takes the values 1, 2, 3 or 4. The probabilities associated with each outcome are as follows:

outcome:	1	2	3	4
probability:	0.1	0.3	0.4	0.2

Cummulative distribution function is a function giving the probability that the random variable X is less than or equal to x, for every value x.

outcome:	1	2	3	4
Cum. dist. fun:	0.1	0.4	0.8	1



Random Variable

A random variable is a rule that assigns a numerical value to each possible outcomes of a probabilistic experiment.

We denote a random variable by a capital letter, say X.

example

X: the age of a randomly selected student in a class.

A discrete random variable can take only distinct, separate values.

example

X: number of heads when tossing 3 coins.

A continuous random variable can take any value in some interval.

example.

X: time a customer spends waiting in line at the store.

Probability function / probability mass function

~~A function f whose value for each real number x is given by $f(x) = P(X=x)$, is called the probability function of the random variable X.~~

example :

Consider an experiment of tossing 3 coins.

Sample space = {HHH, HHT, HTH, HTT, TTT, THT, TTH, THH}

X be defined as the number of heads.

$$X = \{0, 1, 2, 3\}$$

$$f(0)\ P(X=0) = \frac{1}{8}$$

$$f(1)\ P(X=1) = \frac{3}{8}$$

$$f(2)\ P(X=2) = \frac{3}{8}$$

$$f(3)\ P(X=3) = \frac{1}{8}$$

2, 4, 4, 4, 5, 5, 7, 9

$\mu = 5$

$\sigma^2 = 4$

s.d = 2

0, 0, 2, 2, 3, 3, 15, 15

$\mu = 5$

$\sigma^2 = 34.5$

$\sigma = 5.87$

A six-sided die is tossed. win \$2 → 1,
win \$1 → 6.
lose \$1 → 2,3,4,5.

x	2	1	-1
$p(x)$	1/6	1/6	4/6

Mean $= \Sigma x\, p(x) = -0.17$

variance $= \Sigma x^2 p(x) - \mu^2 = 1.47$

s.d $= \sqrt{\text{variance}} = 1.21$

example

Consider an example of rolling two dies.

Sample space = $\{(i,j) \mid i,j = 1,2,3,4,5,6\}$

X be defined as the sum of the number on the two dies.

$$X = \{2,3,4,5,6,7,8,9,10,11,12\}$$

$$f(2) = P(X=2) = P\{(1,1)\} = \frac{1}{36}$$

$$f(3) = P(X=3) = P\{(1,2),(2,1)\} = \frac{2}{36}$$

$\therefore$

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

Note: The probabilities of X in the table add upto 1.

Probability Distribution Function.

The function f is called a probability distribution function of the random variable X, if it satisfies the conditions:

(i) $f(x) \geq 0$ $\quad (P(X=x_i) \geq 0)$

(ii) $\sum f(x) = 1$ $\quad (\sum P(X=x_i) = 1)$

The Cummulative distribution function (CDF) of the random variable X is defined as below:

$$F_X(t) = P(X \leq x_i)$$

Mean or Expected Value, Variance, Standard Deviation

The mean or expected value of a random variable X is defined as

$$E(X) = \mu = \sum P(X=x_i)\, x_i = \sum p_i x_i$$

The variance of a random variable X is defined as

$$Var(X) = \sigma^2 = \sum P(X=x_i)(x_i-\mu)^2$$

$$= \sum p_i (x_i-\mu)^2$$

$$\left[\sum p_i(x_i-\mu)^2 = \sum p_i(x_i^2-2\mu x_i+\mu^2) = \sum p_i x_i^2 - 2\mu \sum p_i x_i + \mu^2 \sum p_i = \sum p_i x_i^2 - 2\mu\times\mu + \mu^2\times 1 = \sum p_i x_i^2 - \mu^2\right]$$

Standard Deviation of a random variable X is defined as

$$SD(X) = \sigma$$

$$= \sqrt{\sigma^2}$$

$$= \sqrt{Var(X)}$$

Suppose X is a discrete random variable.
Let the pmf of X be defined as $f(x) = \frac{5-x}{10}$, $x = 1, 2, 3, 4$.
Find the cdf $F_X(t)$ for $t = 1, 2, 3, 4$.

* Let X be the number that comes up on a roll of one die. Compute the mean, variance and standard deviation of X.

Sol^n $X = \{1, 2, 3, 4, 5, 6\}$.

$f(x_1) = f(1) = P(X=1) = \frac{1}{6} = p_1$ $\quad f(x_4) = f(4) = P(X=4) = \frac{1}{6} = p_4$

$f(x_2) = f(2) = P(X=2) = \frac{1}{6} = p_2$ $\quad f(x_5) = f(5) = P(X=5) = \frac{1}{6} = p_5$

$f(x_3) = f(3) = P(X=3) = \frac{1}{6} = p_3$ $\quad f(x_6) = f(6) = P(X=6) = \frac{1}{6} = p_6$

$$\text{mean} = \sum_i p_i x_i = p_1 x_1 + p_2 x_2 + p_3 x_3 + p_4 x_4 + p_5 x_5 + p_6 x_6$$

$$= \frac{1}{6} \times 1 + \frac{1}{6} \times 2 + \frac{1}{6} \times 3 + \frac{1}{6} \times 4 + \frac{1}{6} \times 5 + \frac{1}{6} \times 6$$

$$\mu = 3.5.$$

$$\text{variance} = \sum p_i (x_i - \mu)^2$$

$$= p_1(x_1 - 3.5)^2 + p_2(x_2 - 3.5)^2 + p_3(x_3 - 3.5)^2 + p_4(x_4 - 3.5)^2 + p_5(x_5 - 3.5)^2 + p_6(x_6 - 3.5)^2$$

$$= \frac{1}{6}(1-3.5)^2 + \frac{1}{6}(2-3.5)^2 + \frac{1}{6}(3-3.5)^2 + \frac{1}{6}(4-3.5)^2 + \frac{1}{6}(5-3.5)^2 + \frac{1}{6}(6-3.5)^2$$

$$\sigma^2 = 2.917.$$

$$\text{Standar deviation} = \sqrt{\text{variance}}$$

$$= \sqrt{2.917}$$

$$\sigma = 1.708$$

* Find the mean and variance of the no of tails when 2 coins are tossed.

Soln.

Sample space = {HH, HT, TH TT}

$X = \{0, 1, 2\}$

$f(0) = P(X=0) = P\{TT\} = \frac{1}{4} = P_1$, $|||^{ly}$ $f(1) = \frac{2}{4} = P_2$, $f(2) = \frac{1}{4} = P_3$

x_i	0	1	2
$P_i = f(x_i)$	$\frac{1}{4}$	$\frac{2}{4}$	$\frac{1}{4}$

$$\mu = \Sigma p_i x_i$$

$$= \frac{1}{4} \times 0 + \frac{2}{4} \times 1 + \frac{1}{4} \times 2$$

mean $\mu = 1$

$$\sigma^2 = \Sigma p_i (x_i - \mu)^2$$

$$= \Sigma p_i x_i^2 - \mu^2$$

$$= \frac{1}{4} \times 0^2 + \frac{2}{4} \times 1^2 + \frac{1}{4} \times 2^2 - 1^2$$

variance $\sigma^2 = \frac{1}{2}$

S.D $\sigma = \frac{1}{\sqrt{2}}$

* If X is a random variable with $P(X=x) = \frac{1}{2^x}$ (4)

where $x = 1, 2, 3, \ldots \infty$, find.

(i) $P(X)$ (ii) $P(X = \text{even})$ (iii) $P(X = \text{divisible by } 3)$

(i) $P(X) = \frac{1}{2^1} + \frac{1}{2^2} + \frac{1}{2^3} + \cdots + \cancel{\frac{1}{2}} \cdots$

which is a geometric progression

with $a = \frac{1}{2}$ & $r = \frac{1}{2}$

$\frac{1/2^2}{1/2} = \frac{1/4}{1/2}$

$S_\infty = \frac{a}{1-r} = \frac{1/2}{1-1/2} = 1$

$= \frac{1}{4} \times \frac{2}{1}$

$= \frac{1}{2}$

(ii) $P(X = \text{even}) = \frac{1}{2^2} + \frac{1}{2^4} + \frac{1}{2^6} + \cdots$

here $a = \frac{1}{2^2}$, $r = \frac{1}{2^2}$

$\therefore S_\infty = \frac{a}{1-r} = \frac{1/2^2}{1 - 1/2^2} = \frac{1/4}{3/4} = \frac{1}{3}$

(iii) $P(X = \text{divisible by } 3) = \frac{1}{2^3} + \frac{1}{2^6} + \frac{1}{2^9} + \cdots$

here $a = \frac{1}{2^3}$, $r = \frac{1}{2^3}$

$\therefore S_\infty = \frac{a}{1-r} = \frac{1/2^3}{1 - \frac{1}{2^3}} = \frac{1/8}{7/8} = \frac{1}{7}$

* Verify if the following function can be a probability distribution function. Find the mean if they are so.

(a) $f(x) = \dfrac{x}{24}$ where $x = \{0, 1, 2, 3, 4\}$

(b) $f(x) = \dfrac{15-x^2}{24}$ where $x = \{0, 1, 2, 3, 4\}$

Sol^n

(a) $f(x) = \dfrac{x}{24}$

(i) $f(x) = \left\{0, \dfrac{1}{24}, \dfrac{2}{24}, \dfrac{3}{24}, \dfrac{4}{24}\right\}$ for $x = \{0,1,2,3,4\}$

$f(x) \geq 0$, for every x

(ii) $\Sigma f(x) = 0 + \dfrac{1}{24} + \dfrac{2}{24} + \dfrac{3}{24} + \dfrac{4}{24}$

$= \dfrac{10}{24} \neq 1$

$\therefore$ It is not a PDF

(b) $f(x) = \dfrac{15-x^2}{24}$

(i) $f(x) = \left\{\dfrac{15}{24}, \dfrac{14}{24}, \dfrac{11}{24}, \dfrac{6}{24}, -\dfrac{1}{24}\right\}$ for $x = \{0,1,2,3,4\}$

$f(x)$ is negative when $x = 4$.

ie, $f(x) < 0$ $(\ngtr 0)$

Hence it is not a PDF.

* Show that the distribution represents a probability distribution. Find the mean & S.D. (5)

x	0	1	2	3	4	5
$f(x)$	$\frac{1}{32}$	$\frac{5}{32}$	$\frac{5}{16}$	$\frac{5}{16}$	$\frac{5}{32}$	$\frac{1}{32}$

Sol^n:

(i) $f(x) \geq 0$ for every x.

(ii) $\sum f(x) = \frac{1}{32} + \frac{5}{32} + \frac{5}{16} + \frac{5}{16} + \frac{5}{32} + \frac{1}{32}$

$= \frac{1+5+10+10+5+1}{32}$

$= 1$

Hence the distribution represents a PDF.

$\mu = \sum p_i x_i$

$= \frac{1}{32} \times 0 + \frac{5}{32} \times 1 + \frac{5}{16} \times 2 + \frac{5}{16} \times 3 + \frac{5}{32} \times 4 + \frac{1}{32} \times 5$

$= \frac{80}{32}$

mean $\mu = 2.5$

$\sigma^2 = \sum p_i x_i^2 - \mu^2$

$= \frac{1}{32} \times 0 + \frac{5}{32} \times 1 + \frac{5}{16} \times 4 + \frac{5}{16} \times 9 + \frac{5}{32} \times 16 + \frac{1}{32} \times 25 - (2.5)^2$

var $\sigma^2 = 1.25$

$\sigma = \sqrt{1.25}$

S.D $\sigma = 1.12$

* The PDF of random variable X is given by the table

x	0	1	2	3	4	5	6
$P(X=x) = f(x)$	k	$3k$	$5k$	$7k$	$9k$	$11k$	$13k$

For what value of k is it a valid probability distribution. Also find (a) $P(X \geq 5)$ (b) $P(3 \leq X \leq 6)$

For a valid PDF $\sum P(X) = 1$

ie, $k + 2k + 5k + 7k + 9k + 11k + 13k = 1$

$\Rightarrow 49k = 1 \Rightarrow k = 1/49$

$\therefore$ The table can be written as

x	0	1	2	3	4	5	6
$f(x)$	$\frac{1}{49}$	$\frac{3}{49}$	$\frac{5}{49}$	$\frac{7}{49}$	$\frac{9}{49}$	$\frac{11}{49}$	$\frac{13}{49}$

(a) $P(X \geq 5) = P(X=5) + P(X=6)$

$= \frac{11}{49} + \frac{13}{49}$

$= \frac{24}{49}$

(b) $P(3 \leq X \leq 6) = P(X=3) + P(X=4) + P(X=5) + P(X=6)$

$= \frac{7}{49} + \frac{9}{49} + \frac{11}{49} + \frac{13}{49}$

$= \frac{40}{49}$

* A discrete random variable has the probability function as follows. (6)

X	0	1	2	3	4	5	6	7
$P(X)$	0	k	$2k$	$2k$	$3k$	$3k^2$	$2k^2$	$7k^2+k$

Find (a) k, (b) $P(X<3)$ (c) $P(2<x\leq 5)$

Sol^n

(a) For a PDF $\sum P(X) = 1$

i.e, $P(X=0)+P(X=1)+P(X=2)+P(X=3)+P(X=4)+P(X=5)+P(X=6)$
$+P(X=7)=1$

or, $0+k+2k+2k+3k+3k^2+2k^2+7k^2+k=1$

$\Rightarrow 12k^2+9k-1=0$

$\Rightarrow k = \dfrac{-9\pm\sqrt{129}}{24}$

$k = 0.098$

Hence the table takes the form,

X	0	1	2	3	4	5	6	7
$P(X)$	0	0.098	0.196	0.196	0.294	0.0288	0.0192	0.1652

(b) $P(X<3) = P(X=0)+P(X=1)+P(X=2)$

$= 0 + 0.098 + 0.196$

$= 0.294$

(c) $P(2<x\leq 5) = P(X=3)+P(X=4)+P(X=5)$

$= 0.196 + 0.294 + 0.028812$

$= 0.5188$

* Two cards are drawn randomly, simultaneously from a well shuffled deck of 52 cards. Find the variance for the no of aces.

Sol^n Sample sapace = 52 cards.

X : drawing ~~2~~ aces ~~together~~.

$X = \{0, 1, 2\}$

$$P(X=0) = \frac{{}^{48}C_2 \times {}^4C_0}{{}^{52}C_2} = \frac{1128}{1326}$$

no aces

$$P(X=1) = \frac{{}^{48}C_1 \times {}^4C_1}{{}^{52}C_2} = \frac{192}{1326}$$

1 ace

$$P(X=2) = \frac{{}^{48}C_0 \times {}^4C_2}{{}^{52}C_2} = \frac{6}{1326}$$

2 aces

X	0	1	2
$P(X)$	$\frac{1128}{1326}$	$\frac{192}{1326}$	$\frac{6}{1326}$

variance $\sigma^2 = \sum p_i x_i^2 - \mu^2$

$$= \sum p_i x_i^2 - \left(\sum p_i x_i\right)^2$$

$$= \left[\frac{1128}{1326}\times 0^2 + \frac{192}{1326}\times 1^2 + \frac{6}{1326}\times 2^2\right] - \left[\frac{1128}{1326}\times 0 + \frac{192}{1326}\times 1 + \frac{6}{1326}\times 2\right]^2$$

$$= \frac{216}{1326} - \left(\frac{204}{1326}\right)^2 = 0.1392$$

* 5 defective bulbs are accidently mixed with twenty good ones. It is not possible to just look at a bulb & tell whether or not it is defective. Find the mean of the number of defective bulbs if 4 bulbs are drawn at random from this lot. (7)

Sol^n:-

X : no of defective bulbs.

$X = \{0, 1, 2, 3, 4\}$.

Sample space = 5 defective bulbs + 20 good bulbs
= 25 bulbs.

$P(X=0) = \frac{{}^{20}C_4 \times {}^5C_0}{{}^{25}C_4} = \frac{969}{2530}$

4 nondefective + 0 defective

$P(X=1) = \frac{{}^{20}C_3 \times {}^5C_1}{{}^{25}C_4} = \frac{1140}{2530}$

3 non defective + 1 defective

$P(X=2) = \frac{{}^{20}C_2 \times {}^5C_2}{{}^{25}C_4} = \frac{380}{2530}$

2 non defective + 2 defective

$P(X=3) = \frac{{}^{20}C_1 \times {}^5C_3}{{}^{25}C_4} = \frac{40}{2530}$

1 non defective + 3 defective

$P(X=4) = \frac{{}^{20}C_0 \times {}^5C_4}{{}^{25}C_4} = \frac{1}{2530}$

0 non defective + 4 defective

mean $\mu = \Sigma p_i x_i = \frac{969}{2530}\times 0 + \frac{1140}{2540}\times 1 + \frac{380}{2530}\times 2 + \frac{40}{2530}\times 3 + \frac{1}{2530}\times 4$

$= \frac{2024}{2530} = 0.8$

* If X is a discrete random variable taking value $1, 2, 3, \ldots$ with $p(x) = \frac{1}{2}\left(\frac{2}{3}\right)^x$.
Find $P(X$ being an odd no$)$ by first establishing that $P(x)$ is a probability function.

$$\Sigma P(X = x_i) = \sum_{x=1}^{\infty} \frac{1}{2}\left(\frac{2}{3}\right)^x$$

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

is a G.P with $a = \frac{2}{3}$, $r = \frac{2}{3}$

$$= \frac{1}{2}\left[\frac{\frac{2}{3}}{1 - \frac{2}{3}}\right]$$

$$= 1$$

$\therefore$ $p(x)$ is a probability function.

$$P(X = x \text{ being an odd no}) = \sum_{x=1,3,5} \frac{1}{2}\left(\frac{2}{3}\right)^x$$

$$= \frac{1}{2}\left[\frac{2}{3} + \left(\frac{2}{3}\right)^3 + \left(\frac{2}{3}\right)^5 + \ldots\right]$$

is a G.P with $a = \frac{2}{3}$, $r = \left(\frac{2}{3}\right)^2$

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

$$= \frac{3}{5}$$

A Continuous Random Variables is one which takes an infinite number of possible values. Continuous random variables are usually measurements. Examples include, height, weight, the amount of sugar in an orange, the time required to run a mile.

A continuous random variable is not defined at specific values. Instead, it is defined over an interval of values, and is represented by the area under a curve. The probability of observing any single value is equal to 0, since the number of values which may be assumed by the random variable is infinite.

Suppose a random variable X may take all values over an interval of real numbers. Then the probability that X is in the set of outcomes A, P(A) is defined to be the area above A and under a curve. The curve, which represents a function $p(x)$, must satisfy the following:

(i) the curve has no negative values ($p(x) \geq 0, \forall x$)

(ii) the total area under the curve is equal to 1.

A curve meeting these requirements is known as a density curve.

Continuous random variable is a random variable which takes an uncountably infine number of possible value or which can take any value in some interval.

example!
Rainfall in a particular area

Probability density function of a continuous random variable X with sample space $S=(-\infty,\infty)$ is an integrable function $f(x)$ satisfying the following:

(i) $f(x)$ is positive everywhere in $S=(-\infty,\infty)$

i.e., $f(x) > 0$ for all x in $S=(-\infty,\infty)$

(ii) The area under the curve $f(x)$ in $S_{(-\infty,\infty)}$ is 1.

i.e., $\int_{-\infty}^{\infty} f(x)\,dx = 1$ $\left[\text{or } \int_S f(x)\,dx = 1\right]$

If $f(x)$ is the p.d.f (probability density function) of x, then the probability that x belongs to A, where A is some interval (a,b), is given by the integral of $f(x)$ over that interval,

i.e., $P(X \in A) = \int_A f(x)\,dx = \int_a^b f(x)\,dx$

Cumulative density function (Cumulative distribution ~~function~~ function) of a continuous random variable X is defined as

$$p(X<t) = F(x) = \int_{-\infty}^{x} f(t)\,dt \quad \text{for } -\infty < x < \infty .$$

Mean, Variance, Standard deviation

The expected value or mean of a continuous random variable X is: $\mu = E(X) = \int_{-\infty}^{\infty} x f(x)\,dx$

The variance of a continuous random variable X is

$$Var(X) = \sigma^2 = E[(X-\mu)^2] = \int_{-\infty}^{\infty} (x-\mu)^2 f(x)\,dx$$

Alternatively, you can still use the shortcut formula for the variance, $\sigma^2 = E(X^2) - \mu^2$ with $E(X^2) = \int_{-\infty}^{\infty} x^2 f(x)\,dx$.

The Standard deviation of a continuous random variable X is:

$$\sigma = \sqrt{var(X)} .$$

$$P(x>t) = P(x \geq t) = \int_{t}^{\infty} p(x)\,dx$$

If $F(t) = P(x<t) = P(x \leq t)$, then $F(t) = \int_{-\infty}^{t} p(x)\,dx = \int_{-\infty}^{\infty} p(x)\,dx - \int_{t}^{\infty} p(x)\,dx$

$$= 1 - \int_{t}^{\infty} p(x)\,dx$$

* Verify whether $f(x) = \begin{cases} 6x(1-x) & 0<x<1 \\ 0 & \text{elsewhere} \end{cases}$

is a probability density function and hence find $P\left(\frac{2}{3} < x < 1\right)$

<u>Soln</u>

(i) $f(x) \geq 0$ in the given interval.

(ii) $$\int_{-\infty}^{\infty} f(x)\,dx = \int_{-\infty}^{0} f(x)\,dx + \int_{0}^{1} f(x)\,dx + \int_{1}^{\infty} f(x)\,dx$$

$$= 0 + \int_{0}^{1} (6x - 6x^2)\,dx + \int 0$$

$$= \left. \frac{6x^2}{2} - \frac{6x^3}{3} \right]_0^1$$

$$= 1$$

Hence verified.

$$P\left(\frac{2}{3} < x < 1\right) = \int_{\frac{2}{3}}^{1} f(x)\,dx$$

$$= \int_{2/3}^{1} (6x - 6x^2)\,dx$$

$$= \left. \frac{6x^2}{2} - \frac{6x^3}{3} \right]_{2/3}^{1}$$

$$= 3 - 2 - \frac{4}{3} + \frac{16}{27} = \frac{7}{27}.$$

* A continuous random variable has the density function

$$f(x) = \begin{cases} k x^2 & -3 < x < 3 \\ 0 & \text{elsewhere.} \end{cases}$$

Find k and hence find $P(x<3)$, $P(x>1)$

<u>Sol?</u> For a density function, we have $\int_{-\infty}^{\infty} f(x)\,dx = 1$.

i.e, $\int_{-\infty}^{-3} f(x)\,dx + \int_{-3}^{3} f(x)\,dx + \int_{3}^{\infty} f(x)\,dx = 1$.

i.e, $0 + \int_{-3}^{3} k x^2\,dx + 0 = 1$

ie, $\left.\frac{k x^3}{3}\right]_{-3}^{3} = 1 \Rightarrow k\left[\frac{27}{3} - \frac{(-27)}{3}\right] = 1 \Rightarrow k = \frac{1}{18}$

$P(x<3) = \int_{-\infty}^{3} f(x)\,dx = \cancel{\int_{-\infty}^{3} k x^2\,dx = \frac{k x^3}{3}}$

$= \int_{-\infty}^{-3} f(x)\,dx + \int_{-3}^{3} f(x)\,dx$

$= 0 + \int_{-3}^{3} k x^2\,dx$

$= \left.\frac{k x^3}{3}\right]_{-3}^{3}$

$= \frac{1}{18}\left[\frac{27}{3} - \frac{(-27)}{3}\right]$

$= 1$

$P(x>1) = \int_{1}^{\infty} f(x)\,dx$

$= \int_{1}^{3} f(x)\,dx + \int_{3}^{\infty} f(x)\,dx$

$= \int_{1}^{3} k x^2\,dx + 0$

$= \left.\frac{k x^3}{3}\right]_{1}^{3}$

$= \frac{1}{18}\left[\frac{27}{3} - \frac{(+1)}{3}\right]$

$= \frac{13}{27}$

* Verify that $f(x) = 3x^2$ is a probability density function. for $0 < x < 1$ (10)

Sol^n: (i) $f(x) > 0$ for $0 < x < 1$.

(ii) $\int_0^1 3x^2 dx = x^3 \Big]_0^1 = 1 - 0 = 1$.

Hence verified.

* Let X be a continuous random variable whose probability desity function is: $f(x) = \frac{x^3}{4}$ for an ~~y~~ interval $0 < x < c$. What is the value of the value of the constant that makes $f(x)$ a valid probability density function?

Sol^n: ~~(i) $\int_0^c \frac{x^3}{4} dx = \frac{1}{4} \cdot \frac{x^4}{4}\Big]_0^c$~~

For $f(x)$ to be a p.d.f, we should have $\int_0^c f(x)\, dx = 1$

ie, $\int_0^c \frac{x^3}{4} dx = 1$

ie, $\frac{1}{4} \times \frac{x^4}{4}\Big]_0^c = 1$

$\Rightarrow \frac{c^4}{16} = 1$

$c^4 = 16$

$c = \pm 2$

but $c = -2$ is not possible

$\therefore c = 2$

* Suppose X is a continuous random variable with the following probability density function; $f(x) = 3x^2$ for $0<x<1$. Find the mean and variance of X.

Sol:

$$\mu = \int_{-\infty}^{\infty} x . f(x) dx$$

$$= \int_{-\infty}^{0} x . f(x) dx + \int_{0}^{1} x . f(x) dx + \int_{1}^{\infty} x . f(x) dx$$

$$= 0 + \int_{0}^{1} x \times 3x^2 dx + 0$$

$$= \int_{0}^{1} 3x^3 dx$$

$$= \left. \frac{3x^4}{4} \right]_0^1$$

$$\mu = \frac{3}{4}$$

$$\sigma^2 = \int_{-\infty}^{\infty} x^2 f(x) dx - \mu^2$$

$$= \int_{0}^{1} x^2 f(x) dx - \mu^2$$

$$= \int_{0}^{1} x^2 \times 3x^2 dx - \left(\frac{3}{4}\right)^2$$

$$= \int_{0}^{1} 3x^4 dx - \left(\frac{3}{4}\right)^2$$

$$= \left. \frac{3x^5}{5} \right]_0^1 - \frac{9}{16}$$

$$\sigma^2 = \frac{3}{5} - \frac{9}{16}$$

$$\sigma^2 = \frac{3}{80}$$

* Find the mean & variance for

$$f(x) = \begin{cases} 2e^{-2x} & x > 0 \\ 0 & x \le 0 \end{cases}$$

Sol^n

$$\mu = \int_{-\infty}^{\infty} x \cdot f(x)\,dx$$

$$= \int_{-\infty}^{0} x\, f(x)\,dx + \int_{0}^{\infty} x\, f(x)\,dx$$

$$= 0 + \int_{0}^{\infty} 2x e^{-2x}\,dx$$

$$= 2\left[x \times \frac{e^{-2x}}{-2} - 1 \times \frac{e^{-2x}}{4}\right]_0^{\infty}$$

$$= 2\left[0 - 0 + 0 + \frac{1}{4}\right]$$

$$\mu = \frac{1}{2}$$

$$\sigma^2 = \int_{-\infty}^{\infty} x^2 \cdot f(x)\,dx - \mu^2$$

$$= \int_{-\infty}^{0} x^2 \cdot f(x)\,dx + \int_{0}^{\infty} x^2 \cdot f(x)\,dx - \mu^2$$

$$= 0 + \int_0^{\infty} x^2 \cdot 2e^{-2x}\,dx - \left(\frac{1}{2}\right)^2$$

$$= 2\left[x^2 \times \frac{e^{-2x}}{-2} - (2x) \times \frac{e^{-2x}}{4} + (2) \times \frac{e^{-2x}}{-8}\right]_0^{\infty} - \frac{1}{4}$$

$$\sigma^2 = 2 \times \frac{2}{+8} - \frac{1}{4}$$

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

$$\sigma^2 = \frac{1}{4}$$

* Find the mean and variance of the probability density function: $f(x) = \frac{1}{2}e^{-|x|}$

Sol^n

$$f(x) = \begin{cases} \frac{1}{2}e^{-(-x)} & x < 0 \\ \frac{1}{2}e^{-x} & x \geq 0 \end{cases} = \begin{cases} \frac{1}{2}e^{x} & x < 0 \\ \frac{1}{2}e^{-x} & x \geq 0. \end{cases}$$

$$\mu = \int_{-\infty}^{\infty} x\, f(x)\,dx$$

$$= \int_{-\infty}^{0} x\, f(x)\,dx + \int_{0}^{\infty} x\, f(x)\,dx$$

$$= \int_{-\infty}^{0} x \times \frac{1}{2}e^{x}\,dx + \int_{0}^{\infty} x \times \frac{1}{2}e^{-x}\,dx$$

$$= \frac{1}{2}\left[x \times e^{x} - 1 \times e^{x}\right]_{-\infty}^{0} + \frac{1}{2}\left[x \times \frac{e^{-x}}{-1} - 1 \times \frac{e^{-x}}{1}\right]_{0}^{\infty}$$

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

$$\underline{\mu = 0}$$

$$\sigma^2 = \frac{1}{2}\int_{-\infty}^{\infty} x^2 f(x)\,dx - \mu^2$$

$$= \int_{-\infty}^{0} x^2 f(x)\,dx + \int_{0}^{\infty} x^2 f(x)\,dx - 0$$

$$= \int_{-\infty}^{0} x^2 \times e^{x}\,dx + \int_{0}^{\infty} x^2 \times e^{-x}\,dx$$

$$= \left[x^2 \times e^{x} - 2x \times e^{x} + 2 \times e^{x}\right]_{-\infty}^{0}$$

$$+ \left[x^2 \times \frac{e^{-x}}{-1} - 2x \times \frac{e^{-x}}{1} + 2 \times \frac{e^{-x}}{-1}\right]_{0}^{\infty}$$

$$= \frac{1}{2}(2 + 2)$$

$$\underline{\sigma^2 = 2}$$

* In a certain city, the daily consumption of electric power (in million kwatt/hr) is random variable X having the probability density function

$$f(x) = \begin{cases} \frac{1}{9} x e^{-x/3} & x > 0 \\ 0 & x < 0 \end{cases}.$$

If the cities power plant has a daily capacity of 12 million kw/hr, what is the probability that this power supply will be insufficient on any given day.

Sol^n:

The power supply will be insufficient on any given day means the power consumption is more than the supply. i.e, $P(X > 12)$

$$P(X > 12) = \int_{12}^{\infty} f(x)\,dx$$

$$= \int_{12}^{\infty} \frac{1}{9} x e^{-x/3}\,dx$$

$$= \frac{1}{9}\left[x \times \frac{e^{-x/3}}{\left(-\frac{1}{3}\right)} - 1 \times \frac{e^{-x/3}}{\left(\frac{1}{9}\right)} \right]_{12}^{\infty}$$

$$= \frac{1}{9}\left[36 e^{-4} + 9 e^{-4}\right]$$

$$= 5 e^{-4}$$

* The length of time (in minutes) that a certain lady speaks on telephone is found to be a random variable with probability function

$$f(x) = \begin{cases} A e^{-x/5} & \text{for } x \geq 0 \\ 0 & \text{otherwise} \end{cases}$$

(a) Find A

(b) Find the probability that she will speak on the phone (b) more than 10 min.

(c) less than 5 min

(d) between 5 & 10 min.

Sol^n: (a) Given $f(x)$ is p.d.f $\int_{-\infty}^{\infty} f(x)\,dx = 1$

$$\Rightarrow \int_{-\infty}^{0} f(x)\,dx + \int_{0}^{\infty} f(x)\,dx = 1 \Rightarrow 0 + \int_{0}^{\infty} A e^{-x/5}\,dx = 1$$

$$\Rightarrow A \left.\frac{e^{-x/5}}{\left(-\frac{1}{5}\right)}\right]_0^{\infty} = 1 \Rightarrow 5A = 1 \Rightarrow A = \frac{1}{5}$$

(b) $P(x > 10) = \int_{10}^{\infty} f(x)\,dx = \int_{10}^{\infty} \frac{1}{5} e^{-x/5}\,dx = \frac{1}{5} \times \left.\frac{e^{-x/5}}{-\frac{1}{5}}\right]_{10}^{\infty} = e^{-2} = 0.1353$

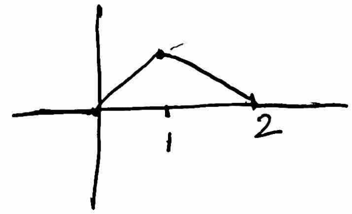
(c) $P(x < 5) = \int_{-\infty}^{5} f(x)\,dx = \int_{0}^{5} \frac{1}{5} e^{-x/5}\,dx = \frac{1}{5} \left.\frac{e^{-x/5}}{-\frac{1}{5}}\right]_{0}^{5} = -e^{-1} + 1 = 0.6322$

(d) $P(5 < x < 10) = \int_{5}^{10} f(x)\,dx = \int_{5}^{10} \frac{1}{5} e^{-x/5}\,dx = \frac{1}{5} \left.\frac{e^{-x/5}}{-1/5}\right]_{5}^{10} = -e^{-2} + e^{-1}$

$= 0.2325$

* The probability density function of a random variable x is given by

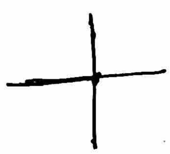
$$p(x) = \begin{cases} x, & 0 \le x \le 1 \\ 2-x, & 1 < x \le 2 \\ 0, & \text{elsewhere.} \end{cases}$$



Find (i) cumulative distribution function, and
(ii) $P(x \geq 1.5)$

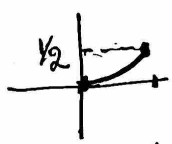
Sol^n. $F(t) = P(x \le t) = \int_{-\infty}^{t} p(x)\,dx$

$-\infty < t < 0$



$$= \int_{-\infty}^{0} p(x)\,dx + \cancel{\int_{0}^{t} p(x)\,dx} = \int_{-\infty}^{0} 0.\,dx = 0$$

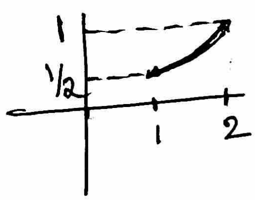
$0 \le t \le 1$



$$= \int_{-\infty}^{0} p(x)\,dx + \int_{0}^{t} p(x)\,dx = \int_{-\infty}^{0} 0.\,dx + \int_{0}^{t} x\,dx = \frac{1}{2}t^2$$

$1 < t \le 2$

$$= \int_{-\infty}^{0} 0.\,dx + \int_{0}^{1} x\,dx + \int_{1}^{t} (2-x)\,dx$$

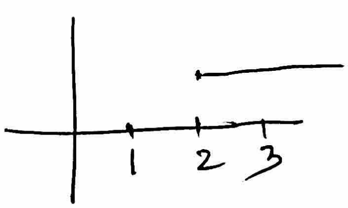


$$= 0 + \left[\frac{x^2}{2}\right]_0^1 + \left[\frac{(2-x)^2}{-2}\right]_1^t = 0 + \frac{1}{2} - \frac{(2-t)^2}{2} + \frac{1}{2}$$

$$= 1 - \frac{(2-t)^2}{2}$$

$t > 2$,

$$= \int_{-\infty}^{0} 0.\,dx + \int_{0}^{1} x\,dx + \int_{1}^{2} (2-x)\,dx + \int_{2}^{t} 0.\,dx$$



$$= 0 + \left.\frac{x^2}{2}\right]_0^1 + \left.\frac{(2-x)^2}{-2}\right]_1^2 + 0$$

$$= \frac{1}{2} - 0 + \frac{1}{2} + 0 = 1$$

required cumulative distribution function is

$$F(t) = \begin{cases} 0, & -\infty < t < 0 \\ \frac{1}{2}t^2, & 0 \le t \le 1 \\ 1 - \frac{1}{2}(t-2)^2, & 1 \le t \le 2 \\ 1, & t > 2 \end{cases}$$

$$P(x \ge 1.5) = 1 - P(x < 1.5)$$
$$= 1 - F(1.5)$$
$$= 1 - \left\{1 - \frac{1}{2}(1.5-2)^2\right\}$$
$$= \frac{1}{8}$$

* Suppose the p.d.f of x is given by $f(x) = 3x^2$ on $0 < x \le 1$. What is the cummulative distribution function $F(t)$.

* pdf is $f(x) = \frac{x^3}{4}$ on $0 < x < 2$. what is cdf.

* pdf is $f(x) = \begin{cases} x+1 & -1 < x < 0 \\ 1-x & 0 \le x < 1 \end{cases}$ what is cdf

Markov inequality

For any nonnegative random variable X with finite $E[X]$, and any $k>0$, the following inequality holds:

$$P[X \geq k] \leq \frac{E[X]}{k}$$

example

A biased coin, with probability of tossing a head being $\frac{1}{5}$, is tossed 10 times. Estimate the probability of getting at least 8 heads in 10 tosses.

$\underline{Sol^n}$ $E[X] = np = 10 \times \frac{1}{5} = 2$.

By Markov inequality,

$$P[X \geq 8] \leq \frac{2}{8} = \frac{1}{4} = 0.25$$

The probability of atleast 8 heads is,

$n = 10 \quad p = \frac{1}{5}, \quad P[X \geq 8] = P[X=8] + P[X=9] + P[X=10]$

${}^{n}C_{x} p^{x} q^{n-x}$

$$= {}^{10}C_{8}\left(\tfrac{1}{5}\right)^{8}\left(\tfrac{4}{5}\right)^{2} + {}^{10}C_{9}\left(\tfrac{1}{5}\right)^{9}\left(\tfrac{4}{5}\right)^{1} + {}^{10}C_{10}\left(\tfrac{1}{5}\right)^{10}\left(\tfrac{4}{5}\right)^{0}$$

$$= 0.0000779 < 0.25$$

example

A random variable X has the following PMF:

$$P_X(x) = \begin{cases} 1/25, & x=5 \\ 24/25, & x=0 \end{cases}$$

Estimate using Markov inequality a bound on the probability that X is at least 5.

$\underline{Sol^n}$ $E[X] = \sum x P_X(x) = 5 \times \frac{1}{25} + 0 \times \frac{24}{25} = \frac{1}{5}$.

$$P[X \geq 5] = \frac{1/5}{5} = \frac{1}{25}$$

and this is exactly the probability of $X=5$.

Markov inequality: Any random variable with expectation $\mu = E[X]$, the probability that X is atleast k times its expected value is at most $\frac{1}{k}$

Chebyshev's inequality

For any real-valued random variable X, with finite mean μ and finite variance $Var[X]$, and $k>0$, the following inequality holds: $P[|X-\mu| \geq k] \leq \frac{Var[X]}{k^2}$.

If we set $k = n\sigma$, where σ is the standard deviation of X, then we get $P[|X-\mu| \geq n\sigma] \leq \frac{Var[X]}{n^2\sigma^2} = \frac{1}{n^2}$.

example.

Let X be a random variable with mean 4 and variance 2. Use Chebyshev's inequality to obtain an upper bound on $P[|X-4| \geq 2]$.

soln $P[|X-4| \geq 2] \leq \frac{2}{2^2} = \frac{1}{2}$.

example

A random variable X is exponentially distributed with parameter λ. Find an upper bound on $P[|X-E[X]| \geq 1]$ using Chebyshev's inequality.

soln For $\lambda>0$, exponential distribution function is given by $f_X(x) = \lambda e^{-\lambda x}$, $x \geq 0$, $E[X] = \frac{1}{\lambda}$ and $Var[X] = \frac{1}{\lambda^2}$.

$\therefore P[|X-E[X]| \geq 1] \leq \frac{1/\lambda^2}{1^2} = \frac{1}{\lambda^2}$

Chebyshev's inequality: Any random variable X with expectation $\mu = E[X]$ and variance $\sigma^2 = var[X]$ belongs to the interval $\mu \pm k = [\mu-k, \mu+k]$ with probability of atleast $1-\left(\frac{\sigma}{k}\right)^2$

ie, $P\{|X-\mu| \geq k\} \leq \left(\frac{\sigma}{k}\right)^2$