21st March, 2023

UNIT-2 Random Variables and Probability Distribution

Random variable

* Let S be the sample space. A random variable that assigns a real no. to every sample point in the sample space of a random exp.

(or)

* A random variable X on a sample space S is a function $X: S \to R$ from S to the set of real nos. R, which assigns a real no. $X(s)$ to each sample point w of the sample space.

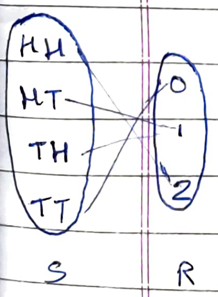
(or)

* Random variable Y is a func. from sample space S into the real nos.

$$S = \{H, T\}$$
$$\quad\downarrow\ \ \downarrow$$
$$\quad 1\ \ \ 0$$
$$X = \{0, 1\}$$

F

$$S = \{HH, HT, TH, TT\}$$
$$\quad 2 \quad\ 4 \quad\ 1 \quad\ 0$$



X = no. of heads in the outcome
$= \{0, 1, 2\}$

Y = no. of tails in the outcome
$= \{0, 1, 2\}$

Note: 1. Random variables are usually denoted by X, Y, Z.

2. Different Random variables may be associated the same sample space.

→ Types of random variables

- Discrete
- Continuous

1. Discrete random variable (counting):
(Integer)

If a random variable takes finite or countably infinite no. of values, then it is discrete

eg: tossing a coin & observing its otcome
no. of people born in Jan

(No gap b/w values)

2 Continuous random variables (measuring):
(Real no.)

If a random variable takes non-counting infinite no. of values.

eg: Weight of apples
life of bulbs
Temp, time
Height

Discrete probability distribution

Probability function :

If for each value x_i of a discrete random variable X, we assign $p(x_i)$ such that

i) $p(x_i) \geq 0$

ii) $\Sigma_i p(x) = 1$

$p(x)$ – probability func.

eg: $S = \{HT, TH, HH, TT\}$

x = no. of heads turning up

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

$p(x=0) = \frac{1}{4} > 0$

$p(x=1) = \frac{2}{4} = \frac{1}{2} > 0$

$p(x=2) = \frac{1}{4} > 0$

* The set of values of x_i together with their probabilities $p(x_i)$ i-e $[x_i, p(x_i)]$ is called a discrete probability distribution of X.

The func. $p(x) = P(X=x)$ is called probability mass func

$x = x_i$	0	1	2
$p(x) = p(x_i)$	1/2	1/2	1/4

→ Cumulative distribution func. of a discrete random variable X (c.d.f)

Cumulative distribution func. $F(x)$ specifies the probability that a random variable is less than or equal to a given value.

(OR)

Cumulative distribution func. $F(x)$ is defined by

$$F(x) = P(X \le x) = \sum_{i=1}^{n} p(x_i), \quad x \text{ being an integer.}$$

→ Mean and Variance of Discrete Probability distribution

Let X be a discrete random variable

Mean $\mu = \sum_i x_i \, p(x_i)$

Variance $\sigma^2 = \sum_i (x_i - \mu)^2 p(x_i)$

or $\sigma^2 = \sum_i x_i^2 - \mu^2$

Standard Deviation $\sigma = \sqrt{\text{Variance}}$

Q: A coin is tossed three times. Let X denote the no. of heads showing up. Find the discrete probability distribution for X. Also find its mean and variance

Sol. $S = 2^3 = 8$

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

3 2 2 1 2 1 0

THT

1

X = no. of heads it outcome

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

$P(X=0) = 1/8$

$P(X=1) = 3/8$

$P(X=2) = 3/8$

$P(X=3) = 1/8$

$x = x_i$	0	1	2	3
$p(x) = p(x_i)$	1/8	3/8	3/8	1/8

The above distribution is called DPD since each $p(x_i) \geq 0$ and

$$\sum p(x_i) = \frac{1}{8} + \frac{3}{8} + \frac{3}{8} + \frac{1}{8} = 1.$$

mean $\mu = \sum x_i \, p(x_i)$

$$= 0\left(\frac{1}{8}\right) + 1\left(\frac{3}{8}\right) + 2\left(\frac{3}{8}\right) + 3\left(\frac{1}{8}\right) = \frac{3}{2}$$

$$\boxed{\mu = 1.5}$$

variance $\sigma^2 = \sum (x_i - \mu)^2 \, p(x_i)$

$$= (0-1.5)^2(1/8) + (1-1.5)^2(3/8) + (2-1.5)^2(3/8) + (3-1.5)^2(1/8)$$

$$= 0.75$$

24th March, 2023

q. A die is tossed thrice. A success is "getting 1 or 6" on a toss. Find the mean and variance of the no. of successes.

Sol X = no. of successes

6^3 $S = \{(x, y, z)$ where $x = 1, 2 \dots 6$, $y = 1, 2 \dots 6$, $z = 1, 2 \dots$

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

$$P(\text{getting success}) = \underbrace{\frac{1}{6}}_{\text{P of getting 1}} \underset{\text{or}}{+} \underbrace{\frac{1}{6}}_{\text{Prob of getting 6}} = \frac{1}{3}$$

$$P(\text{failure}) = 1 - \frac{1}{3} = \frac{2}{3}$$

$$P(X = 0) = \frac{2}{3}\left(\frac{2}{3}\right)\frac{2}{3} = \frac{8}{27}$$

$$P(X = 1) = \frac{1}{3}\left(\frac{2}{3}\right)\left(\frac{2}{3}\right) + \frac{2}{3}\left(\frac{1}{3}\right)\left(\frac{2}{3}\right) + \frac{2}{3}\left(\frac{2}{3}\right)\left(\frac{1}{3}\right)$$

$$= \frac{4}{9}$$

$$P(X = 2) = \frac{1}{3}\left(\frac{1}{3}\right)\left(\frac{2}{3}\right) + \frac{1}{3}\left(\frac{2}{3}\right)\left(\frac{1}{3}\right) + \frac{2}{3}\left(\frac{1}{3}\right)\left(\frac{1}{3}\right)$$

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

$$P(X = 3) = \frac{1}{3}\left(\frac{1}{3}\right)\left(\frac{1}{3}\right) = \frac{1}{27}$$

$\Sigma P(x_i) = \frac{8}{27} + \frac{4}{9} + \frac{2}{9} + \frac{1}{27} = 1$

Mean $\mu = \Sigma x_i \, p(x_i)$

$= 0 \times \frac{8}{27} + 1 \times \frac{4}{9} + 2 \times \frac{2}{9} + 3 \times \frac{1}{27}$

$= 1$

Variance $\sigma^2 = \Sigma (x_i - \mu)^2 \, p(x_i)$

$= (0-1)^2 \frac{8}{27} (.0) + (0) + (2-1)^2 \left(\frac{2}{9}\right) + (3-1)^2 \left(\frac{1}{27}\right)$

$= \frac{2}{3}$

Q: The number of flaws in a 1 inch length of copper wire manufactured by a certain process varies from wire to wire. Overall 48% of the wires produced have no flaws, 39% have 1 flaw, 12% have two flaws and 1% have three flaws. Let X be the no. of flaws in a randomly selected piece of wire. Find the cumulative distribution function $F(x)$ of the random variable X. Also plot $F(x)$.

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

Let us compute $F(x)$ for each of the possible values of X which are 0, 1, 2, 3

$P(X \leq 0)$

$F(0) = \frac{48}{100} = 0.48$

$F(1) = P(X \leq 1) = P(X=0) + P(X=1)$

$= \frac{48}{100} + \frac{39}{100} = 0.87$

$F(2) = P(X \leq 2) = P(0) + P(X=1) + P(X=2)$

$= 0.87 + 0.12$

$= 0.99$

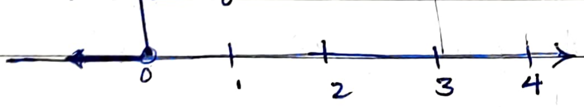
$F(3) = P(X \leq 3) = P(X=0) + P(X=1) + P(X=2) + P(X=3)$

$= 0.99 + 0.01$

$= 1$

For plotting:

$$F(x) = \begin{cases} 0 & x < 0 \\ 0.48 & 0 \leq x < 1 \\ 0.87 & 1 \leq x < 2 \\ 0.99 & 2 \leq x < 3 \\ 1 & x \geq 3 \end{cases}$$



q. Probability density

X:	0	1	2	3	4	5	6
$p(x)$:	k	$3k$	$5k$	$7k$	$9k$	$11k$	$13k$

Find the value of k and Evaluate

Date
Page

i) $P(X<4)$; $P(X \geq 5)$; $P(3 < x \leq 6)$

ii) What will be the max value of k so that $P(X \leq 2) > 0.3$

Sol. For a probability density func.

$$\Sigma P(x_i) = 1$$

$$k + 3k + 5k + 7k + 9k + 11k + 13k = 1$$

$$k = \frac{1}{49}$$

i) $P(X<4) = P(X=0) + P(X=1) + P(X=2) + P(X=3)$

$= k + 3k + 5k + 7k$

$= 16k$

$= \frac{16}{49}$

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

$= 11k + 13k$

$= \frac{24}{49}$

$P(3 < x \leq 6) = P(4) + P(5) + P(6)$

$= 9k + 11k + 13k = \frac{33}{49}$

ii) $P(X \leq 2) > 0.3$ | $k = ?$

$k + 3k + 5k > 0.3$

$9k > 0.3$

$k > \frac{0.3}{9}$

$k > 0.03$

27th March, 2023

q. The probability density function of a random variable X is given

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

Find k. Evaluate $P(X<6)$; $P(X\geq 6)$ and $P(0<X<5)$

Sol. Here we must have $\underset{①}{P(X)\geq 0}$ & $\underset{②}{\Sigma P(X)=1}$

① is satisfied when $k\geq 0$

$$0+k+2k+2k+3k+k^2+2k^2+7k^2+k=01$$
$$k=0.1$$

$$P(X<6) = P(0)+P(1)+P(2)+P(3)+P(4)+P(5)$$
$$= 0.81$$

$$P(X\geq 6) = P(6)+P(7)$$
$$= 0.19$$

$$P(0<x<5) = P(1)+P(2)+P(3)+P(4)$$
$$= 0.8$$

Continuous Probability Distribution

* Here Probability means area.

Definition: If for every x belonging to the range of a continuous random variable X, we assign a real no. $f(x)$ satisfying the condition.

Date
Page

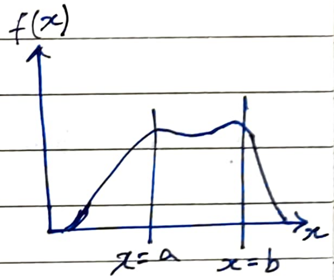
1) $f(x) \geq 0$

2) $\int_{-\infty}^{\infty} f(x)\, dx = 1$ (geometrically means the total area bounded by the curve $f(x)$ and the axis is unity).

then $f(x)$ is called a continuous probability func. or probability density func. (p.d.f)

Note: 1. If (a, b) is a subinterval of the range space of X then the probability that x lies in (a, b) is defined to be the integral of $f(x)$ b/w a and b. That is

$$P(a \leq x \leq b) = \int_a^b f(x)\, dx$$



$p(a \leq x \leq b)$ is = region bounded by curve $f(x)$, the x-axis and the ordinates $x = a$ & $x = b$

→ Cumulative distribution func.

If X is a continuous random variable with probabily density func. $f(x)$ then func. $F(x)$ is defined by

$$F(x) = P(X \leq x) = \int_{-\infty}^{x} f(x)\, dx$$

is called c.d.f of x.

* $F(x) = P(X \leq x) = P(-\infty < X < x)$

* If r is any real no. then

i) $P(X \geq r) = \int_{r}^{\infty} f(x)\, dx$

ii) $P(X < r) = 1 - P(X \geq r)$

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

→ Mean and Variance of Continuous Probability Distribution

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

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

Q. Is the func. defined as follows a density func.?

$$f(x) = e^{-x} \quad (x \geq 0)$$
$$= 0 \quad (x < 0)$$

If so, determine the random variable X having this density will fall in interval (1, 2).

Sol. clearly $f(x) \geq 0$

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

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

$$= \left[-e^{-x}\right]_0^{\infty}$$

$$= -0 + 1 = 1$$

$\therefore f(x)$ is a probability ~~distribution~~ density func.

$$P(1 < x < 2) = \int_1^2 f(x)\, dx$$

$$= \int_1^2 e^{-x}\, dx$$

$$= \left(-e^{-x}\right)_1^2$$

$$= -e^{-2} + e^{-1}$$

$$= \frac{1}{e} - \frac{1}{e^2}$$

$$= 0.2325$$

Q: Let X be continuous random variable with probability density func. given by :

$$f(x) = kx \quad (0 \le x \le 2)$$
$$= 2k \quad (2 \le x \le 4)$$
$$= -k + 6k \quad (4 \le x \le 6)$$

Find k, mean value of X.

Sol. For a probability density $f(x) \ge 0$

$$\int_{-\infty}^{\infty} f(x)\, dx = \cancel{\int_{-\infty}^{0} f(x)\, dx}^{\,0} + \int_0^2 f(x) dx + \int_2^4 f(x)\, dx + \int_4^6 f(x)\, dx + \cancel{\int_6^{\infty} f(x)\, dx}^{\,0}$$

func not defined

func. not defined

$$= 1$$

$$= \int_0^2 kx\, dx + \int_2^4 2k\, dx + \int_4^6 (-kx + 6k)\, dx = 1$$

$$= k\left[\left(\frac{x^2}{2}\right)_0^2 + (2x)_2^4 + \left(-\frac{x^2}{2} + 6x\right)_4^6\right] = 1$$

$$k = 0.125$$

$\left({}^{\cdot}/\!\!\!\phi'', \ da\right)$

$\partial a \rightarrow$ adds

$\partial P \rightarrow$ ada

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

$$= \int_0^2 x(kx)\, dx + \int_2^4 x(2k)\, dx + \int_4^6 (-kx+6k)\, x\, dx$$

$$= k\left[\left(\frac{x^3}{3}\right)_0^2 + 2\left(\frac{x^2}{2}\right)_2^4 + \left(\frac{-x^3}{3} + \frac{6x^2}{2}\right)_4^6\right]$$

$$= k(24) = \frac{1}{8}(24) = 3$$

29th March, 2023

Discrete Probability Distribution
- → Bernoulli
- → Binomial
- → Poisson

Continuous Probability Distribution
- → Exponential (Not in syllabus)
- → Normal

The Bernoulli Distribution

→ Bernoulli trial : A random experiment that can only have an outcome of either success or failure (1 or 0) is known as Bernoulli trial. Such an experiment is used in a Bernoulli distribution. The trials' success is ~~defined~~ denoted as p $(x=1)$, and failure is expressed as $1-p$ $(x=0)$

Date
Page

eg: i) Tossing a coin
ii) A team will win a game or not
iii) A student will pass/ fail

→ Bernoulli distribution

X	0	1
$P(X=x)$	$1-p$	p

$$P(X=x) = \begin{cases} p & x=1 \\ 1-p & x=0 \end{cases} \qquad \text{for } 0<p<1$$

random variable X is said to have, Bernoulli Distribution with parameter p and is denoted by $X \sim \text{Bernoulli}(p)$

→ Mean and Variance of Bernoulli distribution

Mean $\mu_x = 0(1-p) + 1(p)$
$= p$

DPB
mean $= \sum x_i P(x_i)$
var. $= \sum (x_i - \mu)^2 P(x_i)$

Variance $\sigma^2 = (0-p)^2(1-p) + (1-p)^2 p$
$= p(1-p)$

Q: A coin has a probability of 0.5 of landing head when tossed. Let $X=1$ if the coin comes up heads and $X=0$ if coin comes up tails. What is the distribution of X?

Sol. The success probability is $p = P(X=1) = 0.5$
$X \sim \text{Bernoulli}(0.5)$

X is said to have Bernoulli dist. with success probability p

q. A die has a probability 1/6 of coming up 6 when rolled. Let $X=1$ if the die comes up and $X=0$ otherwise. What is the distribution of X?

Sol.

X	1	0
$P(X=x)$	1/6	5/6

The success probability is $p = P(X=1) = 1/6$

$\therefore X \sim \text{Bernoulli}\left(\frac{1}{6}\right)$ means X is said to have Bernoulli Distribution with success probability 1/6

q. Ten percent of components manufactored by a certain process are defective. A component is chosen at random. Let $X=1$ if the component is defective, and $X=0$ otherwise.
What is the distribution of X? Fing μ_x and σ_x^2

Sol.

X	1	0
$P(X=x)$	0.1	0.9

The success probability is $p = P(X=1) = 0.1$

$\therefore X \sim \text{Bernoulli}(0.1)$

$\mu_x = p = 0.1$ $\qquad$ $\sigma_x^2 = p(1-p) = 0.1(0.9)$

$= 0.09$

Date
Page

Binomial Distribution

* Here a series of n Bernoulli trials is conducted, each with the same success probability p. Let the random variable X equals the no. of successes in these n trials. Then X is said to have the binomial distribution with parameters n and p.
The notation $X \sim Bin(n, p)$, X is a discrete random variable and its possible values are 0, 1, 2, ... n.

→ Probability mass function for a binomial random variable.

If $X \sim Bin(n, p)$, the probability mass func. of X is

$$p(x) = P(X = x) = \begin{cases} {}^nC_x \, p^x (1-p)^{n-x} & x = 0, 1, 2 \ldots n \\ 0 & \text{otherwise} \end{cases}$$

→ Mean and Variance

$$\mu_x = np \qquad \sigma_x^2 = np(1-p) = npq$$

* $X \sim B(n, p)$ means that the discrete random variable X has a binomial probability with n trials and probability of success p.

X = no. of ~~independent~~ success in a n independent trials

Q. The probability that a pen manufactured by the company will be defective is 1/10. If 12 pens are manufactured, find the probability that
i) exactly 2 pens are defective
ii) at least 2 pens are defective
iii) none of pens are defective

Sol. X = no. of success = no. of defective pens

$p = P(\text{success}) = P(\text{defective pens}) = 0.1$
$q = P(\text{failure}) = P(\text{non-defective pens}) = 1 - 0.1 = 0.9$

$\therefore p = 0.1 \; ; \; q = 0.9 \; ; \; n = 12$

i) $P(X=2) = P(\text{exactly 2 pens defective}) = 12C_2 (0.1)^2 (0.9)^{12-2}$

$= \frac{12(11)\cancel{(10!)}}{\cancel{(10!)}\,2} (0.1)^2 (0.9)^{8}$

$= 0.23$

ii) $P(\text{at least } 2) = P(X \geq 2) = 1 - P(X < 2)$

$= 1 - (P(X=0) + P(X=1))$

$= 1 - [{}^{12}C_0 (0.1)^0 (0.9)^{12} + {}^{12}C_1 (0.1)^1 (0.9)^{11}]$

$= 0.34$

iii) $P(X=0) = {}^{12}C_0 (0.1)^0 (0.9)^{12}$

$= 0.28$

Date
Page

q. Out of 800 families with 5 children each, how many have i) 3 boys ii) 5 girls iii) either 2/3 boys

Sol. $N = 800$
$n = 5$

X = no. of success = having a boy

$p = P(\text{having a boy}) = 1/2$
$q = P(\text{not having a boy}) = 1/2$

i) P(getting 3 boys) $= P(X=3)$
$= {}^5C_3 (0.5)^3 (0.5)^2 = 0.31$

$\therefore$ no. of families having 3 boys $= 800 \times 0.31 = 250$

ii) $P(X=0) = {}^5C_0 (0.5)^0 (0.5)^5$
$= 0.031$

$\therefore$ no. of families having 5 girls $= 0.031 \times 800$
$= 25$

iii) P(either 2 or 3 boys) $= P(X=2) + P(X=3)$
$= {}^5C_2 (0.5)^2 (0.5)^3 + 0.31$
$= 0.625$

$\therefore$ no. of families having 2/3 boys $= 800 \times 0.625$
$= 500$

q. In 256 sets of 12 tosses of a coin, how many sets can expect 8 heads and 4 tails ?

Sol. $N = 256$

$n = 12$

X = no. of success = getting a head

p = P(getting a head) $= 1/2$

q = P(not getting a head) $= 1/2$

P(getting 8 heads & 4 tails) $= P(X=8)$

$= {}^{12}C_8 \, (0.5)$

Population

To estimate the success probability p, we compute the same proportion $\hat{p}$

$$\hat{p} = \frac{\text{no. of successes}}{\text{no. of trials}} = \frac{X}{n}$$

q. A quality engineer in testing the calliberation of a machine that packs ice cream. In a sample of 20 containers, 3 are unfilled. Estimate the probability p that the machine underfills a container.

Sol. The sample proportion of underfilled condition is

$$\hat{p} = \frac{X}{n} = \frac{3}{20} = 0.15$$

Date
Page

* The uncertainty $\hat{p}$ is $\sigma_{\hat{p}} = \sqrt{\frac{p(1-p)}{n}}$

Note: When computing $\sigma_{\hat{p}}$, we substitute $\hat{p}$ for p, since p is unknown.

q. The safety commission in a large city wants to estimate the proportion of building in city that is in violation of fire codes. A random sample of 40 buildings are chosen for inspection and 4 of them are found to have fire code violation. Estimate the proportion of building in the city that have fire code violations and find the uncertainty in estimate

Sol. $\hat{p} = \frac{X}{n} = \frac{4}{40} = 0.1$

$$\text{uncertainty} = \sqrt{\frac{p(1-p)}{n}} = \sqrt{\frac{0.1(1-0.1)}{40}} = 0.047$$

3rd April, 2023

Poisson Distribution

The poisson probability mass function is defined by -

$$p(x) = P(X=x) = \begin{cases} \dfrac{e^{-\lambda}\lambda^{x}}{x!} & \text{if } x \text{ is a non-negative integer} \\ 0 & \text{otherwise} \end{cases} \quad -①$$

If X is a random variable whose probability mass func. is given by ① then X is said to have the Poisson distribution with parameter λ.
The notation is

$$X \sim \text{Poisson}(\lambda)$$

* X is a random variable following a Poisson Distribution.
* x is the no. of times an event occurs.
* $P(X=x)$ is the probability that an event will occur x times.
* e is euler's constant (≈ 2.718)
* λ is avg. no. of times an event occurs.
* ! is factorial function.

- Some important results

* Poisson distribution is the discrete probability distribution of a discrete random variable X, which has no upper bound ($\to \infty$).

* Each trial is a ~~Binomial~~ Bernoulli trial.

* no. of trials n is very large i.e $n \to \infty$

* The constant probability of success p, is very small ($p \to 0$).
* $\lambda = np$ is a finite +ve real no.

Note : Poisson Distribution is very close to Binomial when $n \to \infty$ & $p \to 0$ such that $np = \lambda$ = constant

- Mean and Variance of Poisson Distribution

Mean $\mu_x = \lambda = np$

Variance $\sigma^2 = \lambda = np$

here mean = variance

q: If $X \sim$ Poisson $(\underset{\lambda}{3})$, compute $P(X=2)$, $P(X=0)$, $P(X=-1)$ & $P(X=0.5)$.

Sol. $\lambda = 3$

$$P(X=x) = \frac{e^{-\lambda} \lambda^x}{x!}$$ if x is non negative I

$$P(X=2) = \frac{e^{-3} 3^2}{2!} = 0.224$$

$$P(X=0) = \frac{e^{-3} 3^0}{0!} = 0.049$$

$P(X=-1)$ & $P(X=0.5) = 0$ as x is negative & $\notin$ I

q: $X \sim$ Poisson(4). compute $P(X \leq 2)$ & $P(X > 1)$

Sol. $\lambda = 4$

$$P(X = x) = \frac{e^{-\lambda}\lambda^x}{x!}$$

$$P(X \le 2) = P(X=0) + P(X=1) + P(X=2)$$

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

$$= 0.238$$

$$P(X > 1) = 1 - P(X \le 1)$$

$$= 1 - [P(X=0) + P(X=1)]$$

$$= 0.909$$

Q: Particles are suspended in a liquid medium at a concentration of 6 particles per ml. A large volume of the suspension is thoroughly agitated, and then 3 ml are withdrawn. What is the probability that exactly 15 particles are withdrawn?

Sol. X = no. of particles withdrawn

λ = mean no. of particles in 3 ml

$= 6 \times 3 = 18$

$$P(X = x) = \frac{e^{-18}18^x}{x!}$$

$$P(X = 15) = \frac{e^{-18}(18)^{15}}{15!} = 0.079$$

Date
Page

q. Assume that the no. of cars that pass through a certain intersection during a fixed time interval follows a Poisson Distribution. Assume that mean rate is 5 cars per min. Find probability that exactly 17 cars will pass through the intersection in next 3 mins.

Sol. X = no. of cars passing through intersection in next 3 minutes.

$\lambda = 5 \times 3 = 15$ = mean no. of cars for 3 mins

$$P(X=x) = \frac{e^{-15}\, 15^{x}}{x!}$$

$$P(X=17) = \frac{e^{-15} \cdot 15^{17}}{17!} = 0.085$$

q. Suppose 2% of the items produced by a factory are defective. Find the probability that there are 5 defective items in a sample of 100 items.

Sol. $n = 100$

$p = p(\text{success}) = p$ (no. of defective items)

$p = 0.02$

$\lambda = np = 100\,(0.02) = 2$

$$P(X=5) = \frac{e^{-2}\, 2^{5}}{5!} = 0.036$$

Binomial $P(X=x) = {}^{n}C_{x}\, p^{x} q^{n-x}$

$= {}^{100}C_{5}\, (0.02)^{5}\, (0.98)^{95}$

$= 0.035$

q. Suppose 300 misprints are distributed randomly throughout a book of 500 pages. Find probability that a given page contains

i) Exactly 2 misprints
ii) Two or more misprints

Sol. X = no. of misprints in a given page.
$n = 300$

$$p = p(\text{no. of misprints}) = \frac{300}{500} = 0.6$$

$$\lambda = np = 300(0.6) = 180$$

$$P(X=2) = \frac{e^{-180} (180)^2}{2!} = 1.08 \times 10^{-74} = 0$$

$\Rightarrow$ almost 0

$$P(X \geq 2) = 1 - P(X<2)$$
$$= 1 - [P(X=0) + P(X=1)]$$
$$= 1 - \left[\frac{e^{-180}(180)^0}{0!} + \frac{e^{-180}(180)^1}{1!}\right]$$
$$= 1$$

8th April, 2023

q. An average of 0.61 soldiers died by horse kick per year in each Prussian army. Calculate the probability that exactly two soldiers died in army assuming that death per year follows Poisson's distribution.

Sol. $\lambda = 0.61$ $\quad P(X=2) = \frac{e^{-0.61}(0.61)^2}{2!} = 0.101$

Date
Page

q. At a small walk-in clinic, on avg. of 5 patients arrive at clinic per hour during opening hours. What is the probability that exactly 3 patients arrive in the next hour.

Sol. $\lambda = 5$ $\quad P(X=3) = \dfrac{e^{-5}\, 5^3}{3!} = 0.14$

q. A communication channel receives independent pulses at the rate of 12 pulses per micro second. The probability of transmission error is 0.001 for each micro second. Compute Probability of

i) no error during a microsecond
ii) one error per microsecond
iii) atleast one error per microsecond
iv) 2 errors
v) ~~atleast~~ atmost 2 errors

Sol. X = no. of transmission errors

$n = 12$

$p = 0.001$ (error)

$\lambda = np = 0.012$

$P(X=0) = e^{-0.012} = 0.988$

$P(X=1) = \dfrac{e^{-0.012}(0.012)}{1} = 0.01$

$P(X \geq 1) = 1 - P(X=0) = 0.012$

$P(X=2) = \dfrac{e^{-0.012}(0.012)^2}{2!} = 7.11 \times 10^{-15}$

$P(X \leq 2) = P(0) + P(1) + P(2) = 0.988 + 0.01 + 7.11 \times 10^{-15}$

$= 0.998$

8th April, 2023

Normal Distribution

* The probability density func. of a random variable with mean (μ) and variance (σ^2) is given by,

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \quad \text{where } \sigma > 0$$

Note: If X is a continuous random variable whose probability density function is equal with mean (μ) and variance (σ^2), $X \sim N(\mu, \sigma^2)$

If $X \sim N(\mu, \sigma^2)$ then the mean and variance of X are given by

$$\mu_x = \mu \; ; \; \sigma_x^2 = \sigma^2$$

$$P(a \le x \le b) = \int_a^b f(x)\,dx = \int_a^b \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx$$

Standard Normal Distribution

$$z = \frac{x-\mu}{\sigma}$$

(or) $x = \mu + \sigma z$

$dx = \sigma\, dz$

$$z_1 = \frac{a-\mu}{\sigma} \qquad z_2 = \frac{b-\mu}{\sigma}$$

$$P(a \le x \le b) = P(z_1 \le z \le z_2) = \int_{z_1}^{z_2} \frac{1}{\sigma\sqrt{2\pi}} e^{-z^2/2} \cdot \sigma\,dz$$

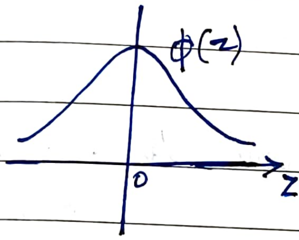
Standard normal curve

If $F(z) = \frac{1}{\sqrt{2\pi}} e^{-z^2/2}$, then $F(x)$ is called standard normal density function.
Normal density function with $\mu = 0$ & $\sigma = 1$ is called standard normal probability function.

Here $z = \frac{x - \mu}{\sigma}$ is called normal variate

and $F(x)$ is called normal curve which is symmetrical about line $z = 0$

curve:



$$\phi(z) = \frac{1}{\sqrt{2\pi}} \int_0^z e^{-z^2/2} \, dz$$

represents area under standard normal curve from $z=0$ to $z=z$

Note: 1. $P(z_1 \le x \le z_2) = \int_{z_1}^{z_2} F(z) \, dz$

represents area bounded by normal curve $f(z)$ which is bell shaped curve b/w $z = z_1$ & $z = z_2$

2. Area under standard normal curve is unity.
3. Since total area $= 1$, area on either side of $z = 0$ is 0.5
4. We use tables to find area b/w co-ordinates $z = z_1$ & z_2.
5. Here the probabilities are found by looking up tables of pre calculated values
6. The standard normal distribution is also called as z distribution and the values are called z-values.

Results: 1. $P(-\infty \le z \le \infty) = 1$

2. $P(-\infty \le z \le 0) = 0.5$

3. $P(0 \le z \le \infty) = 0.5$

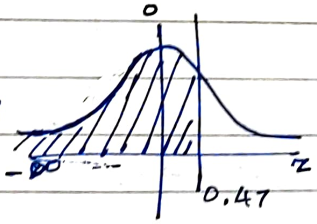
q. Ball bearings manufactured for a cetrain application have diameter (in mm) that are normally distributed with mean 5 and standard deviation 0.08. A particular ball has a diameter of 5.06mm. Find Z

$$Z = \frac{x - \mu}{\sigma} = \frac{5.06 - 5}{0.08} = 0.75$$

q. Find area under curve to left of $z = 0.47$

area from $-\infty$ to $0 = 0.5$

area from 0 to $0.47 = 0.78$



Area $= 0.5 + 0.78$

$= 1.28$.

q. Lifetime of batteries in a certain application are normally distributed with mean 50 hours and standard deviation 5 hours. Find the probability that a randomly chosen battery lasts b/w 42 & 52 hours. Find the 40th percentile of the battery lifetimes

Date
Page

Sol. $z = \frac{x-\mu}{\sigma}$ $\mu = 50, \sigma = 5$

$x = 42$ $z_1 = \frac{42-50}{5} = -1.6$

$x = 52$ $z_2 = \frac{52-50}{5} = 0.4$

$P(42 < x < 52) = P(-1.6 \leq z \leq 0.4) = 0.6$

From the z-table, the closest area to 0.4 is 0.4013, corresponding to a z-score of −0.25. The population of lifetimes has mean & S.D 5. The 40th percentile is the point 0.25 standard deviations below ~~the~~ mean.

$z = \frac{x-\mu}{\sigma}$ $\Rightarrow -0.25 = \frac{x-50}{5}$

$x = 48.75$

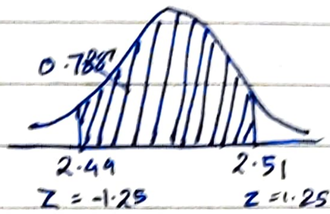
∴ 40th percentile of battery lifetime is 48.75 hours

Q. A process of manufacturers ball bearings whose diameters are normally distributed with mean 2.505 cm and s.d 0.008 cm. Specifications call for the diameter to be in the internal 2.5 ± 0.01 cm. i) What is the probability that the ball bearings will meet specifications?

ii) The process can be recalibrated so that the mean will equal 2.5cm, the center of specification interval. The s.d of the process remains 0.008 cm. What proportion of the diameters will meet the specifications?

iii) Assume that the process has been recalibrated so that the mean diameter is now 2.5 cm. To what value.

Sol i) Let x



ii) $\mu = 2.5, \quad \sigma = 0.008$

when $x = 2.49$, $\quad z = \dfrac{x-\mu}{\sigma} = \dfrac{2.49-2.5}{0.008} = -1.25$

$x = 2.51$, $\quad z = \dfrac{2.51-2.5}{0.008} = 1.25$

$P(2.49 < x < 2.51) = P(-1.25 < z < 1.25) = 0.89435 - 0.10565$

$= 0.7887$

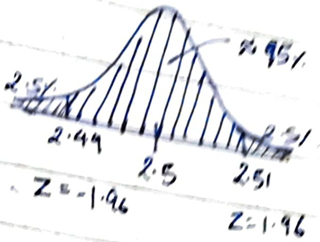
On recalibrating the mean, the proportion to meet the specification is increased to 78.87%

iii) The z-score corresponding to 0.025 to the left and to the right is −1.96 (from calc) to the right, z score is +1.96

$\therefore z = \dfrac{x-\mu}{\sigma}$

$1.96 = \dfrac{2.51 - 2.5}{\sigma}$

$\sigma = 0.005$



q. In a test on 2000 electric bulbs it was found that life of a particular bulb with an average life of 2040 hours and standard deviation of 60 hours. Find the no. of bulbs likely to burn for.

i) More than 2150 hours
ii) Less than 1950 hours
iii) More than 1920 hours and less than 2160 hours.

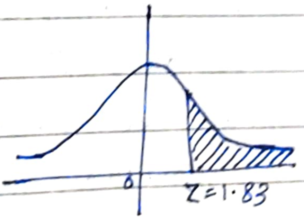
Sol. $\mu = 2040$; $\sigma = 60$; $x \rightarrow$ Life of an electric bulb.

i) $P(x > 2150)$

$x = 2150$; $z = \dfrac{x-\mu}{\sigma} = \dfrac{2150-2040}{60} = 1.83$

$= P(z > 1.83)$

$= 1 -$ area to the left of $z = 1.83$
(from calci)



$= 1 - 0.9667$

$= 0.03362$

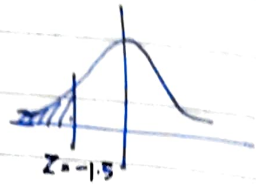
$\therefore$ no. of bulbs likely to burn for 2150 hours is

$0.03362 \times 2000 = 67$

ii) $P(X < 1950)$

$P(z < -1.5)$

$= 0.0668$



The no. of bulbs that are likely burn for less than 1950 hours is

$2000 \times 0.0688 = 133$

iii) $P(1920 < x < 2160)$

when $x = 1920$; $z = -2$

$x = 2160$; $z = 2$

$P(-2 < z < 2) = 0.9545$

no. of bulbs $= 2000 \times 0.9545$

$= 1909$

9. The lifetime of a battery in a certain application is normally distributed with mean 16 hours and s.d 2 hours.

i) What is the probability that a battery will last more than 19 hours?

ii) Find the 10th percentile of the lifetimes?

iii) A particular battery lasts 14.5 hours. What is percentile is its lifetime on?

iv) What is the probability that the lifetime of a battery is b/w 14.5 and 17 hours?

12th April, 2023

Central Limit Theorem

* Let $X_1, \dots X_n$ be a simple random sample from a population with mean μ and variance σ^2

Let $\bar{X} = \dfrac{X_1 + \dots + X_n}{n}$ be the sample mean

Let $S_n = X_1 + \dots + X_n$ be sum of sample observations

Then if n is sufficiently large $\bar{X} \sim N\left(\mu, \dfrac{\sigma^2}{n}\right)$ approx.

& $S_n \sim N(n\mu, n\sigma^2)$ approx.

$\bar{X} = \dfrac{S_n}{n}$

$S_n = n\bar{X}$

* CLT says that if we

⋮

Note: 1. For any sample mean

$$\mu_{\bar{X}} = \mu$$

$$\sigma_{\bar{X}}^2 = \sigma^2/n$$

2. $S_n = n\bar{X}$

$\therefore \mu_{S_n} = n\mu$ and $\sigma_{S_n}^2 = n\sigma^2$

3. For most population, if the sample size is greater than 30, the CLT approximation is good.

q. A simple random sample of 100 men is chosen from a population with mean height of 70 inches and a standard deviation of 2.5 inches. What is the probability that the average height of the men in the sample is greater than 69.5 inches?

Sol. $\mu = 70$; $\sigma = 2.5$; $n = 100$

$p(\bar{x})$ $p(\bar{x} > 69.5)$

$$\bar{x} = 69.5$$

$$\mu_{\bar{x}} = \mu = 70$$

$$\sigma_{\bar{x}}^2 = \frac{\sigma^2}{n} = \frac{(2.5)^2}{100} = 0.0625$$

$$\sigma_{\bar{x}} = 0.25$$

when $\bar{x} = 69.5$ $z = ?$

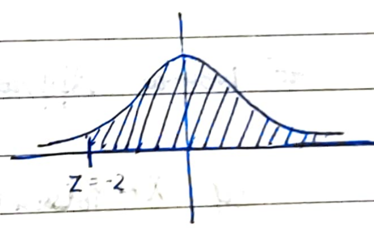
$$z = \frac{\bar{x} - \mu_{\bar{x}}}{\sigma_{\bar{x}}} = \frac{69.5 - 70}{0.25} = -2$$

$\therefore p(\bar{x} > 69.5) = p(z > -2)$

$= 1 - p(p < -2)$

value gives area from left

$= 0.9772$



q. At a large uni, the mean age of the students is 22.3 years, and s.d is 4 years. A random sample of 64 students is drawn. What is the probability that avg. age of these students is > 23 years.

Sol. $\mu = 22.3$ $\sigma = 4$ $n = 64$

$P(\bar{x} > 23)$

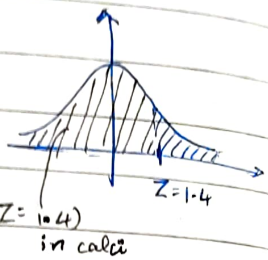
$$\mu_{\bar{x}} = \mu = 22.3$$

$$\sigma_{\bar{x}}^2 = \frac{\sigma^2}{n} = 0.25 \qquad \sigma_{\bar{x}} = 0.5$$

when $\bar{x} = 23$ $\quad z = \dfrac{\bar{x} - \mu_{\bar{x}}}{\sigma_{\bar{x}}}$

$z = \dfrac{23 - 22.3}{0.5} = 1.4$

$P(z > 1.4) = 1 - P(z < 1.4)$
$= 0.08$



Normal Approximation to Binomial
(provided n is very large)

If $X \sim Bin(n, p)$, and if $np > 10$ and $n(1-p) > 10$
then,

using normal to find binomial

mean $\quad$ variance

$X \sim N(np,\ np(1-p))$ approximately

$\hat{p} \sim N\left(p,\ \dfrac{p(1-p)}{n}\right)$ approximately

# Normal Approximation to Poisson
mean = variance

If $X \sim Poisson(\lambda)$, where $\lambda > 10$ then,

mean $\quad$ variance

$X \sim N(\lambda, \lambda)$

q. Let X denote the no. of flaws in a 1 inch length of copper wire. The probability mass func. of X is given below.

x	$P(X=x)$
0	0.48
1	0.39
2	0.12
3	0.01

One hundred wires are sampled from this population. What is the probability that avg. no. of flaws < 0.5 ?

Sol pop. mean $\mu = 0.25$ $\left(\frac{0.48 + 0.39 + 0.12 + 0.01}{4}\right)$

pop. variance $\sigma^2 =$