Module-4

PROBABILITY DISTRIBUTIONS

- Binomial distribution
- Poisson distribution
- Normal distribution
- Gamma distribution
- Exponential distribution
- Weibull distribution

Bernoulli's trials:

Suppose, associated with random trial there is an event called 'success' and the complementary event is called 'failure'. Let the probability for success be p and probability for failure be q. Suppose the random trials are prepared n times under identical conditions. These are called Bernoullian trials.

Bernoulli's Distribution:

A random variable X which takes two values 0 and 1 with probability $q\ and\ p$ respectively. That is $P(X = 0) = q\ and\ P(X = 1) = p,\ \ q = 1 - p$ is called a Bernoulli's discrete random variable. The probability function of Bernoulli's distribution can be written as

$$P(X) = p^X q^{1-X} = p^X (1-p)^{1-X}\ \ ; X = 0,1$$

Note:

1. Mean of Bernoulli's distribution discrete random variable X

$$\mu = E(X) = \sum X_i . P(X_i) = (0 \times q) + (1 \times p) = p$$

2. Variance of X is

$$V(X) = E(X^2) - E(X)^2 = \sum {X_i}^2\, P(X_i) - \mu^2$$

$$= (0^2 \times q) + (1^2 \times p) - p^2 = p - p^2 = p(1-p) = pq$$

The standard deviation is $\sigma = \sqrt{pq}$

Probability Binomial Distribution:

Let a random experiment be performed repeatedly and let the occurrence of an event A is any trial be called success and non-occurrence $P(\bar{A})$, a failure (Bernoulli trial). Consider a series of n independent Bernoulli trials (n being finite) in which the probability of success $P(A) = p$ or $P(\bar{A}) = 1 - p = q$ in any trial is constant for each trial.

$$P(X = x) = n_{C_x}\, p^x q^{n-x}, \qquad x = 0,1,2,3, \dots, n$$

Since the probabilities of 0,1,2, 3,…,n successes, namely $q^n, n_{C_1}\ \ q^{n-1}p, n_{C_2}\ \ q^{n-2}p^2, \dots, p^n$ are the successive terms of the Binomial expansion of $(q+p)^n$, the probability distribution so obtained is called Binomial probability distribution.

Definition:

A random variable X is said to be follow Binomial distribution denoted by B (n, p), if it assumes only non-negative values and its probability mass function is given by

$$P(X = x) = \begin{cases} n_{C_x}\, p^x q^{n-x}, & x = 0,1,2,3, \dots, n \\ 0, & otherwise \end{cases}$$

Where $n\ and\ p$ are known as parameters.

Note:

- If n is also sometimes known as the degree of the distribution
- $\sum_{x=0}^{n} n_{C_x}\, p^x q^{n-x} = (q+p)^n = 1$
- The Binomial distribution is important not only because of its wide range applicability, but also because it gives rise to many other probability distributions.

- Any variable which follows Binomial distribution is known as Binomial variate.

Conditions for Binomial Experiment:

The Bernoulli process involving a series of independent trials, is based on certain conditions as under:

- There are only two mutually exclusive and collective exhaustive outcomes of the random variable and one of them is referred to as a success and the other as a failure.
- The random experiment is performed under the same conditions for a fixed and finite (also discrete) number of times, say n. Each observation of the random variable in a random experiment is called a trial. Each trial generates either a success denoted by p or a failure denoted by q.
- The outcome (i.e., success or failure) of any trial is not affected by the outcome of any other trial.
- All the observations are assumed to be independent of each of each other. This means that the probability of outcomes remains constant throughout the process.

Example:

To understand the Bernoulli process, consider the coin tossing problem where 3 coins are tossed. Suppose we are interested to know the probability of two heads. The possible sequence of outcomes involving two heads can be obtained in the following three ways:
HHT, HTH, THH.

Binomial Probability Function:

In general, for a binomial random variable, X the probability of success (occurrence of desired outcome) r number of times in n independent trials, regardless of their order of occurrence is given by the formula:

$$P(X = r\ successes) = n_{C_r}\, p^r q^{n-r} = \frac{n!}{(n-r)!\, r!} p^r q^{n-r}, r = 0,1,2,3, \dots, n$$

where

n = number of trials (specified in advance) or sample size

p = probability of success

$q = (1 - p)$, probability of failure

x = discrete binomial random variable

r = number of successes in n trials

Relationship between mean and variance:

Mean of a Binomial distribution:

The Binomial probability distribution is given by

$$p(r) = n_{C_r}\, p^r q^{n-r}\ ;\ r = 0,1,2, \dots, n, \qquad and\ q - 1 - p$$

Mean of X is

$$\mu = E(X) = \sum_{r=0}^{n} r\, p(r) = \sum_{r=0}^{n} r\, n_{C_r}\, p^r q^{n-r}$$

$$= 0 \times q^n + 1 \times n_{C_1} p q^{n-1} + 2 \times n_{C_2} p^2 q^{n-2} + 3 \times n_{C_3} p^3 q^{n-3} + \cdots + n\, p^n$$

$$= npq^{n-1} + 2.\frac{n(n-1)}{2!} p^2 q^{n-2} + 3.\frac{n(n-1)(n-2)}{3!} p^3 q^{n-3} + \cdots + n\, p^n$$

$$= np\left[q^{n-1} + (n-1)pq^{n-2} + \frac{(n-1)(n-2)}{2!} p^2 q^{n-3} + \cdots + p^{n-1}\right]$$

$$= np(q + p)^{n-1}$$

$$= np$$

$$\mu = E(X) = np$$

Variance of a Binomial distribution:

Variance $V(X) = E(X^2) - E(X)^2$

$$= \sum_{r=0}^{n} r^2 p(r) - \mu^2$$

$$= \sum_{r=0}^{n} [r(r-1) + r].p(r) - \mu^2$$

$$= \sum_{r=0}^{n} r(r-1)\, n_{C_r}\, p^r q^{n-r} + \sum_{r=0}^{n} r\, p(r) - \mu^2$$

$$= \left[2.n_{C_2}\, p^2 q^{n-2} + 3.2.n_{C_3} p^3 q^{n-3} + \cdots + n(n-1)\, p^n\right] + \mu - \mu^2$$

$$= \left[2.\frac{n(n-1)}{2!} p^2 q^{n-2} + 6.\frac{n(n-1)(n-2)}{3!} p^3 q^{n-3} + \cdots + n(n-1)\, p^n\right] + \mu - \mu^2$$

$$= n(n-1)p^2 \left[q^{n-2} + (n-2)pq^{n-3} + \frac{(n-2)(n-3)}{2!} p^2 q^{n-4} + \cdots + \ p^{n-2}\right] + \mu - \mu^2$$

$$= n(n-1)p^2 (q+p)^{n-2} + \mu - \mu^2$$

$$= n(n-1)p^2 + np - (np)^2$$

$$= n^2p^2 - np^2 + np - n^2p^2 = np - np^2 = np(1-p) = npq$$

$$V(X) = npq$$

Problem 1:

A fair coin is tossed six times, then find the probability of getting four heads.

Solution:

p= probability of getting a head=1/2

q= probability of not getting a head=1/2

$n = 6, r = 4$

$p(r) = 6_{C_4} p^r q^{n-r}$

$p(4) = 6_{C_4} \left(\frac{1}{2}\right)^4 \left(\frac{1}{2}\right)^2$

$= \frac{6!}{4!\,2!} \left(\frac{1}{2}\right)^6 = \frac{15}{64}$

Problem 2:

The incidence of an occupational disease in an industry is such that the workers have a 20% chance of suffering from it. What is the probability that out of 6 workers chosen at random, four or more will suffer from disease?

Solution:

The probability of a worker suffering from disease= p=20%=0.2

The probability that of no worker suffering from disease= q=80%=0.8

The probability that four or more workers suffer from disease $= P(X \geq 4)$

$$P(X \geq 4) = P(X = 4) + P(X = 5) + P(X = 6)$$

$$= 6_{C_4}(0.2)^4(0.8)^2 + 6_{C_5}(0.2)^5(0.8) + 6_{C_6}(0.2)^6 = 0.0175$$

Problem 3:

Six dice are thrown 729 times. How many times do you except at least three dice to show a 5 or 6.

Solution:

$p =$ probability of occurrence of 5 or 6 in one throw$= \frac{2}{6} = \frac{1}{3}$

$q = 1 - p = 1 - \frac{1}{3} = \frac{2}{3}$

$n = 6$

The probability of getting at least three dice to show a 5 or 6

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

$= 6_{C_3}\left(\frac{1}{3}\right)^3\left(\frac{2}{3}\right)^3 + 6_{C_4}\left(\frac{1}{3}\right)^4\left(\frac{2}{3}\right)^2 + 6_{C_5}\left(\frac{1}{3}\right)^5\left(\frac{2}{3}\right)^1 + 6_{C_6}\left(\frac{1}{3}\right)^6$

$= \frac{1}{(3)^6}(160 + 60 + 12 + 1) = \frac{233}{729}$

The expected number of such cases in 729 times

$= 729\left(\frac{233}{729}\right) = 233$

Problem 4:

If the probability of a detective bolt is 0.2, find

(i) Mean and
(ii) Standard deviation for the bolts in a total of 400.

Solution:

Given $n = 400, p = 0.2, q = 0.8$

(i) Mean is $np = 400 \times 0.2 = 80$
(ii) Standard deviation is $\sqrt{npq} = \sqrt{80 \times 0.8} = \sqrt{64} = 8$

Problem 5:

Find the maximum n such that the probability of getting no head in tossing a fair coin n times is greater than 0.1.

Solution:

$p =$ probability of getting a head = ½

$q =$ probability of not getting a head = 1- ½ = ½

Probability of getting no head in tossing a fair coin n times is greater than 0.1 is

$$P(X = 0) > 0.1$$

$$n_{C_0} p^0 q^n > 0.1$$

$$q^n > 0.1$$

$$\left(\frac{1}{2}\right)^n > 0.1$$

$$2^n < 10, then\ n > 3.$$

Hence the required maximum $n = 3$.

Problem 6:

Fit a binomial distribution to the following frequency distribution

x	0	1	2	3	4	5	6
f	13	25	52	58	32	16	4

Solution:

The number of trials is $n = 6$

$N = \sum f_i =$ total frequency

Mean $= \frac{\sum f_i x_i}{\sum f_i} = \frac{25+104+174+128+8+24}{200} = 2.675$

Mean $= np$

$$np = 6p, \quad then\ p = \frac{2.675}{6} = 0.446$$

$$q = 1 - 0.446 = 0.554$$

Binomial distribution to be fitted is $N(q+p)^n = 200(0.554 + 0.446)^6$

$$= 200\big[6_{C_0}(0.554)^6 + 6_{C_1}(0.554)^5(0.446) + 6_{C_2}(0.554)^4(0.446)^2 + 6_{C_3}(0.554)^3(0.446)^3$$
$$+ 6_{C_4}(0.554)^2(0.446)^4 + 6_{C_5}(0.554)^1(0.446)^5 + 6_{C_6}(0.446)^6\big]$$
$$= 200[0.02891 + 0.1396 + 0.2809 + 0.3016 + 0.1821 + 0.05864 + 0.007866]$$
$$= 5.782 + 27.92 + 56.18 + 60.32 + 36.42 + 11.728 + 1.5732$$

The successive terms in the expansion give the expected or theoretical frequencies which are

x	0	1	2	3	4	5	6
f (expected or theoretical	6	28	56	60	36	12	2

frequencies)							

Home Work:

1. A die is tossed thrice. A success is getting 1 or 6 on a toss. Find the mean and variance of the number of successes.
2. The mean and variance of a binomial distribution are 4 and 4/3 respectively. Then find $P(X \geq 1)$.
3. Fit a binomial distribution to the following frequency distribution

x	0	1	2	3	4	5
f	2	14	20	34	22	8

Moment Generating Function of Binomial distribution:

Let $X \sim B(n, p)$, then

$$M(t) = M_X(t) = E(e^{tx})$$

$$= \sum_{x=0}^{n} e^{tx} \, n_{C_x} p^x q^{n-x}$$

$$= \sum_{x=0}^{n} n_{C_x} (pe^t)^x \, q^{n-x} = (q + pe^t)^n$$

Characteristic Function of Binomial distribution:

$$\emptyset_X(t) = E\left(e^{itx}\right)$$

$$= \sum_{x=0}^{n} e^{itx} \, n_{C_x} p^x q^{n-x}$$

$$= \sum_{x=0}^{n} n_{C_x} (pe^{it})^x q^{n-x} = (q + pe^{it})^n$$

Problem:

If the moment generating function of a random variable X is of the form $(0.4\ e^t + 0.6)^8$, find the moment generating function of $3X + 2$.

Solution:

Moment generating function of a random variable X is

$$M_X(t) = E(e^{tx}) = (q + pe^t)^n = (0.6 + 0.4\ e^t)^8$$

X follows the Binomial distribution with $q = 0.4, p = 0.6, n = 8$

MGF of $3X + 2$ is

$$M_{3X+2}(t) = E\left(e^{t(3x+2)}\right) = E(e^{t3x}e^{2t}) = e^{2t}E(e^{t3x})$$

$$= e^{2t}E\left(e^{(3t)x}\right)$$

$$= e^{2t}(0.4\ e^{3t} + 0.6)^8$$

Cumulative Binomial distribution:

$$B(x; n, p) = P(X \leq x) = \sum_{k=0}^{x} b(k; n, p) = \sum_{k=0}^{x} n_{C_k}\, p^k q^{n-k}\ , \qquad x = 0,1,2,3, \dots, n$$

The Binomial probabilities can be obtained from cumulative distribution as follows

$$b(x; n, p) = B(x; n, p) - B(x - 1; n, p)$$

Note: B(-1)=0

By using the Binomial table these can also be obtained.

Example:

The manufacture of large high-definition LCD panels is difficult, and a moderately high proportion have too many defective pixels to pass inspection. If the probability is 0.3 that an LCD panel will not pass inspection, what is the probability that 6 of 18 panels, randomly selected from production will not pass inspection?

Solution:

X: LCD panel not pass in inspection.

n=18, p=0.30 and x=6

$$b(x; n, p) = B(x; n, p) - B(x - 1; n, p)$$

$$b(6; 18,0.30) = B(6; 18, 0.30) - B(5; 18, 0.30) = 0.7217 - 0.5344 = 0.1873$$

Table A.1 (continued) Binomial Probability Sums $\sum_{x=0}^{r} b(x; n, p)$

		p									
n	r	0.10	0.20	0.25	0.30	0.40	0.50	0.60	0.70	0.80	0.90
17	0	0.1668	0.0225	0.0075	0.0023	0.0002	0.0000				
	1	0.4818	0.1182	0.0501	0.0193	0.0021	0.0001	0.0000			
	2	0.7618	0.3096	0.1637	0.0774	0.0123	0.0012	0.0001			
	3	0.9174	0.5489	0.3530	0.2019	0.0464	0.0064	0.0005	0.0000		
	4	0.9779	0.7582	0.5739	0.3887	0.1260	0.0245	0.0025	0.0001		
	5	0.9953	0.8943	0.7653	0.5968	0.2639	0.0717	0.0106	0.0007	0.0000	
	6	0.9992	0.9623	0.8929	0.7752	0.4478	0.1662	0.0348	0.0032	0.0001	
	7	0.9999	0.9891	0.9598	0.8954	0.6405	0.3145	0.0919	0.0127	0.0005	
	8	1.0000	0.9974	0.9876	0.9597	0.8011	0.5000	0.1989	0.0403	0.0026	0.0000
	9		0.9995	0.9969	0.9873	0.9081	0.6855	0.3595	0.1046	0.0109	0.0001
	10		0.9999	0.9994	0.9968	0.9652	0.8338	0.5522	0.2248	0.0377	0.0008
	11		1.0000	0.9999	0.9993	0.9894	0.9283	0.7361	0.4032	0.1057	0.0047
	12			1.0000	0.9999	0.9975	0.9755	0.8740	0.6113	0.2418	0.0221
	13				1.0000	0.9995	0.9936	0.9536	0.7981	0.4511	0.0826
	14					0.9999	0.9988	0.9877	0.9226	0.6904	0.2382
	15					1.0000	0.9999	0.9979	0.9807	0.8818	0.5182
	16						1.0000	0.9998	0.9977	0.9775	0.8332
	17							1.0000	1.0000	1.0000	1.0000
18	0	0.1501	0.0180	0.0056	0.0016	0.0001	0.0000				
	1	0.4503	0.0991	0.0395	0.0142	0.0013	0.0001				
	2	0.7338	0.2713	0.1353	0.0600	0.0082	0.0007	0.0000			
	3	0.9018	0.5010	0.3057	0.1646	0.0328	0.0038	0.0002			
	4	0.9718	0.7164	0.5187	0.3327	0.0942	0.0154	0.0013	0.0000		
	5	0.9936	0.8671	0.7175	0.5344	0.2088	0.0481	0.0058	0.0003		
	6	0.9988	0.9487	0.8610	0.7217	0.3743	0.1189	0.0203	0.0014	0.0000	
	7	0.9998	0.9837	0.9431	0.8593	0.5634	0.2403	0.0576	0.0061	0.0002	
	8	1.0000	0.9957	0.9807	0.9404	0.7368	0.4073	0.1347	0.0210	0.0009	
	9		0.9991	0.9946	0.9790	0.8653	0.5927	0.2632	0.0596	0.0043	0.0000
	10		0.9998	0.9988	0.9939	0.9424	0.7597	0.4366	0.1407	0.0163	0.0002
	11		1.0000	0.9998	0.9986	0.9797	0.8811	0.6257	0.2783	0.0513	0.0012
	12			1.0000	0.9997	0.9942	0.9519	0.7912	0.4656	0.1329	0.0064
	13				1.0000	0.9987	0.9846	0.9058	0.6673	0.2836	0.0282
	14					0.9998	0.9962	0.9672	0.8354	0.4990	0.0982
	15					1.0000	0.9993	0.9918	0.9400	0.7287	0.2662
	16						0.9999	0.9987	0.9858	0.9009	0.5497
	17						1.0000	0.9999	0.9984	0.9820	0.8499
	18							1.0000	1.0000	1.0000	1.0000

Poisson Distribution:

S.D. Poisson (1837) introduced Poisson distribution as a rare distribution of rare events.

i.e. The events whose probability of occurrence is very small but the no. of trials which could lead to the occurrence of the event, are very large.

Ex:

1. The no. of printing mistakes per page in a large text
2. Number of suicides reported in a particular city
3. Number of air accidents in some unit time

4. Number of cars passing a crossing per minute during the busy hours of a day, etc.

Definition:

A random variable X taking on one of the non-negative values 0,1,2,3,4, … (i.e. which do not have a natural upper bound) with parameter λ, λ >0, is said to follow Poisson distribution if its probability mass function is given by

$$P(x;\lambda) = P(X = x) = \begin{cases} \dfrac{\lambda^x e^{-\lambda}}{x!}, & x = 0,1,2,3,\ldots \\ 0, & otherwise \end{cases}$$

Then X is called the Poisson random variable and the distribution is known as Poisson distribution.

And the Poisson parameter, $\lambda = \mathrm{np}$>0

Conditions to follow in Poisson Distribution:

- The no. of trials 'n' is very large
- The probability of success 'p' is very small
- $\lambda = \mathrm{np}$ is finite.

Mean of Poisson distribution:

$$\mu = E(X) = \sum_{x=0}^{\infty} x\, P(x;\lambda) = \sum_{x=0}^{\infty} x\, \frac{\lambda^x e^{-\lambda}}{x!}$$

$$\mu = E(X) = \lambda = \mathrm{np}$$

Variance of Poisson distribution:

Variance $V(X) = E(X^2) - E(X)^2$

$$= \sum_{x=0}^{\infty} x^2 p(x; \boldsymbol{\lambda}) - \mu^2$$

$$V(X) = \sigma^2 = \lambda$$

Cumulative Poisson distribution:

$$F(x; \lambda) = P(X \leq x) = \sum_{k=0}^{x} P(k; \lambda) = \sum_{k=0}^{x} \frac{\lambda^k e^{-\lambda}}{k!}$$

Moment generating function:

$$M(t) = E(e^{tX}) = \sum_{x=0}^{\infty} e^{tx}.P(x; \lambda) = \sum_{x=0}^{\infty} e^{tx} \frac{\lambda^x e^{-\lambda}}{x!} = e^{\lambda(e^t - 1)}$$

$$M(t) = e^{\lambda(e^t - 1)}$$

Characteristic function:

$$\emptyset(t) = E\left(e^{itX}\right) = \sum_{x=0}^{\infty} e^{itx}.P(x; \lambda) = \sum_{x=0}^{\infty} e^{itx} \frac{\lambda^x e^{-\lambda}}{x!} = e^{\lambda(e^{it} - 1)}$$

$$\emptyset(t) = e^{\lambda(e^{it} - 1)}$$

Problem 1:

A hospital switch board receives an average of 4 emergency calls in a 10-minute interval. What is the probability that

i. there at most 2 emergency calls in a 10-minute interval

ii. there are exactly 3 emergency calls in a 10-minute interval.

Solution:

Mean $= \lambda = 4$

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

i. P(at most 2 calls) = $P(X \leq 2)$

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

$$= \frac{1}{e^4} + 4.\frac{1}{e^4} + 8.\frac{1}{e^4}$$

$$= \frac{1}{e^4}(1 + 4 + 8) = 0.2381$$

ii. P(Exactly 3 calls) = $P(X = 3) = \frac{1}{e^4}.\frac{16}{3!} = 0.1954$

Problem 2:

If a random variable has a Poisson distribution such that P (1) =P (2). Find

i. Mean of the distribution
ii. P(4)
iii. $P(X \geq 1)$
iv. $P(1 < X < 4)$

Solution:

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

$$\lambda^2 = 2\lambda$$

$$\lambda = 0 \text{ or } 2$$

$$\text{But } \lambda \neq 0 \text{ or } 2$$

$$\text{Therefore } \lambda = 2$$

i. Mean of the distribution is $\lambda = 2$

ii. $\mathrm{p(x)} = \frac{e^{-\lambda}\lambda^x}{x!}$

$$\mathrm{p}(4) = \frac{e^{-2}2^4}{4!} = 0.09022$$

iii. $P(X \geq 1) = 1 - P(X < 1) = 1 - P(X = 0)$

$$= 1 - \frac{e^{-2}2^0}{0!} = 0.8647$$

iv. $P(1 < X < 4) = P(X = 2) + P(X = 3)$

$$= \frac{e^{-2}2^2}{2!} + \frac{e^{-2}2^3}{3!} = 0.4511$$

Problem 3:

Fit a Poisson distribution to the following data

x	0	1	2	3	4	5
f	142	156	69	27	5	1

Solution:

Mean $= \frac{\sum f_i x_i}{\sum f_i} = \frac{0+156+138+81+20+5}{400} = 1$

Mean of the distribution is $\lambda = 1$

So, theoretical frequency for x successes are given by $N\ P(x)$.

$$N\ P(x) = 400 \times \frac{e^{-1}1^x}{x!}, \quad x = 0,2,3,4,5$$

i.e., $400 \times e^{-1}, 400 \times e^{-1}, 200 \times e^{-1}, 66.67 \times e^{-1}, 16.67 \times e^{-1}, 3.33 \times e^{-1}$

i.e., $147.15, 147.15, 73.58, 24.53, 6.13, 1.23$

The expected frequencies are

x	0	1	2	3	4	5
Theoretical frequency	142	156	69	27	5	1
Expected frequency	147	147	74	25	6	1

Problem 4:

If the moment generating function of the random variable is $e^{4(e^t-1)}, find\ P(X=\mu+\sigma)$, where μ and σ^2 are the mean and variance of the Poisson random variable X.

Solution:

$$M(t) = e^{\lambda(e^t-1)} = e^{4(e^t-1)}$$

Mean=Variance= $\lambda = 4$

Standard deviation = $\sqrt{4}$ =2

$$P(X=\mu+\sigma) = P(X=4+2) = P(6)$$

$$P(X=x) = P(X=6) = \frac{e^{-4}4^6}{6!} = 0.1042$$

Try yourself:

1. The distribution of typing mistakes committed by a typist is given below. Assuming the distribution to be Poisson, find the expected frequencies

x	0	1	2	3	4	5
f	42	33	14	6	4	1

Discrete Bivariate Distributions:

Covariance:

We are often interested in the inter-relationship, or association, between two random variables. The covariance of two random variables X and Y is

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

Or

$$Cov(X,Y) = E[(X - E(X))(Y - E(Y))]$$

Note:

Covariance is an obvious extension of variance

$$Cov(X,X) = Var(X)$$

Moments:

1. **The moments for the Binomial Distribution**:

By the definition of a moment $\mu_r = E[X - E(X)]^r$

The first four central moments of the Binomial distribution are:

we know that $\mu_0 = 1, \mu_1 = 0$

$$Mean = np$$
$$\mu_2 = npq$$
$$\mu_3 = npq(q - p)$$
$$\mu_4 = npq[1 + 3pq(n - 2)]$$

2. **The moments for the Poisson Distribution**:

The first four central moments of the Poisson distribution are:

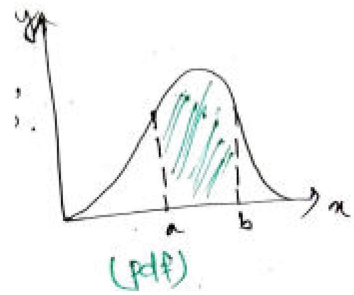
$$\mu_1 = 0$$
$$\mu_2 = \lambda$$
$$\mu_3 = \lambda$$
$$\mu_4 = 3\lambda^2 + \lambda$$

Normal probability distribution:

Probability density or distribution function (PDF):

Let X be a continuous random variable, then the probability density function of X is a function of $f(X)$ such that for any two numbers a and b with $a \leq b$.

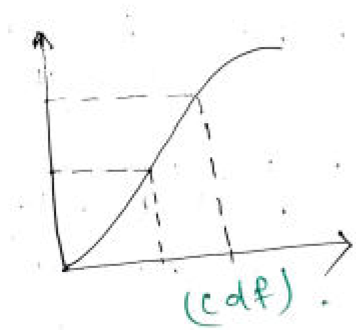
$P(a \leq X \leq b) = \int_a^b f(X)dX$



Cumulative distribution function (CDF):

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

$$P(a \leq X \leq b) = F(b) - F(a)$$



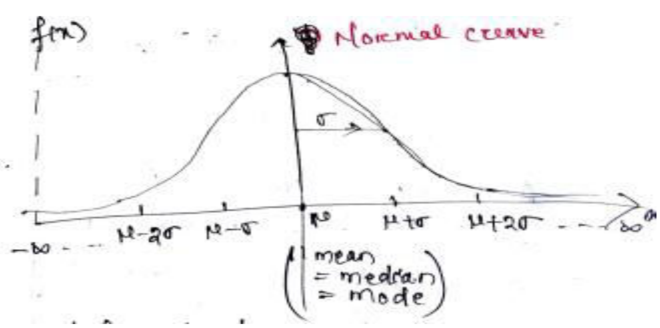
Normal Distribution:

A random variable X is said to have a normal distribution, if its density function or probability distribution is given by

$$f(x;\mu,\sigma) = \frac{1}{\sigma\sqrt{2\pi}}\, e^{\frac{-(x-\mu)^2}{2\sigma^2}}, -\infty < x < \infty, -\infty < \mu < \infty, \qquad \sigma > 0.$$

Where, μ is the mean and σ is the standard deviation of x.

- As can be seen, the function, called probability density function of the normal distribution, depends on two values μ and σ. These are referred as the two parameters of the normal distribution.
- The curve has maximum value at μ and tapers off on either side but never touches the horizontal line.
- The curve on the left side goes up to $-\infty$, and on the right side it goes up to $+\infty$. However, as much as 99.73% of the area under the curve lies between $(\mu - 3\sigma)$ and $(\mu + 3\sigma)$ and only 0.277% of the area lies beyond these points.
- The random variable x is then said to a normal random variable or normal variate. The curve representing the normal distribution is called the normal curve and the total area bounded by the curve and the x -axis is one. i.e., $\int f(x)dx = 1$



Normal distribution is applicable in the following situations:

1. Life of items subjected to wear and tear like tyres, batteries, bulbs, currency notes, etc.
2. Length and diameter of certain products like pipes, screws and discs.
3. Height and weight of baby at birth.
4. Aggregate marks obtained by students in an examination.
5. Weekly sales of an item in store.

Standard Normal Distribution:

The Normal Distribution with mean (μ) =0 and S.D. (σ)=1, is known as Standard Normal Distribution.

The random variable that follows this distribution is denoted by z. If a variable x follows normal distribution with mean μ and s.d. σ, the variable z defined as

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

has standard normal distribution with mean 0 and s.d. as 1. This is also referred as z-score.

Uses of Normal distribution:

1. The Normal distribution can be used to approximate Binomial and Poisson distributions.
2. It has extensive use in sampling theory. It helps us to estimate parameter from statistic and to find confidence limits of the parameter.
3. It has a wide use in testing Statistical Hypothesis and Tests of significance in which it is always assumed that the population from which the samples have been drawn should have normal distribution.
4. It serves as a guiding instrument in the analysis and interpretation of statistical data.

Mean of Normal Distribution:

Consider the Normal Distribution with b, σ as the parameters. Then

$$f(x; b, \sigma) = \frac{1}{\sigma\sqrt{2\pi}}\, e^{\frac{-(x-b)^2}{2\sigma^2}}$$

The mean μ= E(X) is given by

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

$$= \int_{-\infty}^{\infty} \frac{1}{\sigma\sqrt{2\pi}}\, e^{\frac{-(x-b)^2}{2\sigma^2}}\, dx.$$

put z = $\frac{x-b}{\sigma}$ => σz+ b= x

dz = $\frac{dx}{\sigma}$

$$\mu = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} (\sigma z + b) e^{\frac{-z^2}{2}} dz.$$

$$\mu = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} (\sigma z) e^{\frac{-z^2}{2}} dz + \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} (b) e^{\frac{-z^2}{2}} dz$$

here, $ze^{\frac{-z^2}{2}}$ *is odd function so* $\int_{-\infty}^{\infty} (z) e^{\frac{-z^2}{2}} dz = 0$

$$\mu = 0 + \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} (b) e^{\frac{-z^2}{2}} dz.$$

$$\mu = \frac{b}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{\frac{-z^2}{2}} dz.$$

$e^{\frac{-z^2}{2}}$ *is even function so* $\int_{-\infty}^{\infty} e^{\frac{-z^2}{2}} dz = 2\int_{0}^{\infty} e^{\frac{-z^2}{2}} dz$

$$\mu = \frac{2b}{\sqrt{2\pi}} \int_{0}^{\infty} e^{\frac{-z^2}{2}} dz.$$

We know that, $\int_{0}^{\infty} e^{\frac{-z^2}{2}} dz = \sqrt{\frac{\pi}{2}}$

$$\mu = \frac{2b}{\sqrt{2\pi}} \times \sqrt{\frac{\pi}{2}}$$

$$\mu = b$$

Variance of N.D:

$$\sigma_X{}^2 = Var(X) = \int_{-\infty}^{\infty} (X-\mu)^2 dX = E[(X-\mu)^2]$$

Let X has a normal distribution i.e., $X \sim N(\mu, \sigma^2)$ with mean μ and standard deviation σ, we can standardize to a standard normal random variable

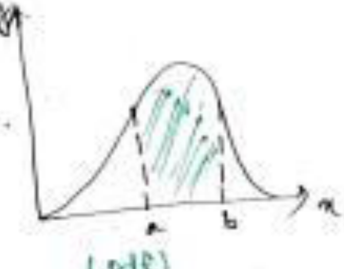
$$Z = \frac{X-\mu}{\sigma}$$

Normal Distribution

Probability density / distribution function :→

Let X be a cont. r.v. Then the probability density function (pdf) of X is a function 'f(x)' such that for any two numbers 'a' & 'b' with a≤b.

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

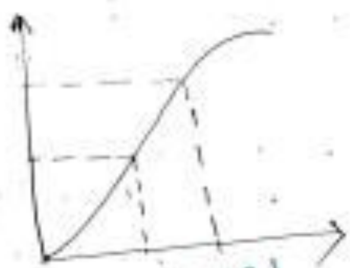


(pdf)

CDF

$$F(x) = P(X \le x) = \int_{-\infty}^{x} f(y)\,dy$$

$$P(a \le X \le b) = F(b) - F(a)$$



(cdf).

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

Variance $\sigma_x^2 = V[X] = \int_{-\infty}^{\infty} (x-\mu)^2 \cdot f(x)\,dx = E[(X-\mu)^2]$.

Normal distribution

A cont. r.v. 'X' is said to have the normal distribution with parameter 'μ' and 'σ' if its pdf is given by

$$\boxed{f(x) = \frac{1}{\sqrt{2\pi}\,\sigma} e^{-(x-\mu)^2/2\sigma^2}}, \quad -\infty < x < \infty$$

where $\mu = E(x)$, σ = std. dev.

- Thus the normal distribution is characterised by mean 'μ' & std dev 'σ'

*. It is used to study the height of person, the velocity in any direction of a gas molecule & error made in measuring physical quantity. etc..

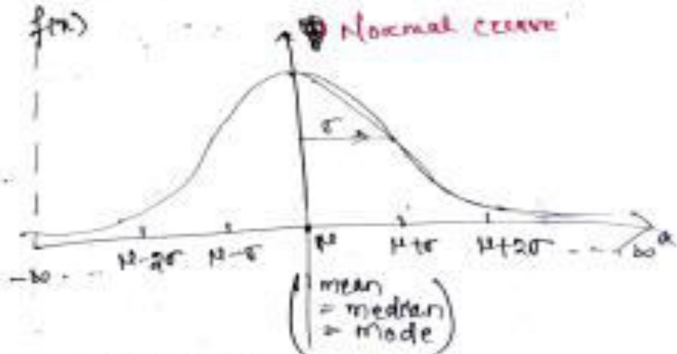
Properties

→ Applied to single variable cont. data. i.e. height, weight, length etc.

→ ~~The normal curve is bell~~ used to calculate the probability 'less than' 'greater than', & in betⁿ.

→ Since f(x) being the probability, can never be negative, no portion of the curve lies below x-axis.

graphical property



→ The normal density curve is bell shaped.

→ It is symmetric about mean. (mean = median = mode)

→ Spread of the curve is determined by the std deviation 'σ'.

→ Location is determined by the mean 'μ'.

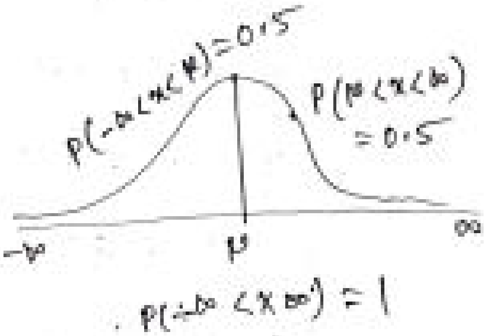
13(4)-
4

The ~~curve~~ normal curve is symmetric about mean

The total area under the curve is 1.

i.e. $P(-\infty < X < \infty) = 1$, also $P(-\infty < Z < \infty) = 1$

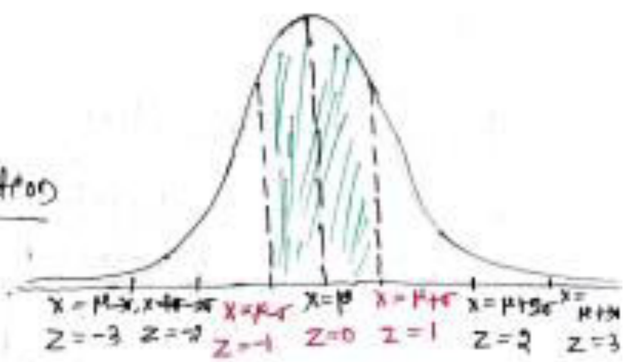
and since the curve is symmetric about mean μ so half is above (right) the mean and half is below (left) the mean



Standardizing Normal RV.

If 'X' has a normal distribution (i.e. $X \sim N(\mu, \sigma^2)$) with mean μ and standard deviation σ, we can standardize to a standard normal r.v.

$$\boxed{Z = \frac{X-\mu}{\sigma}}$$



Standard normal distribution

$(\mu = 0, \sigma^2 = 1)$

then,

$$\boxed{f(z) = \frac{1}{\sqrt{2\pi}\,\sigma} e^{-\frac{z^2}{2}}}$$

or

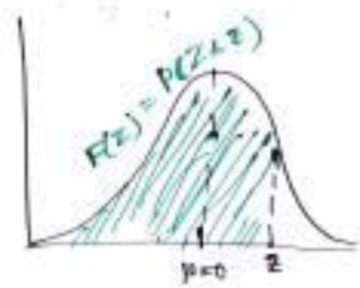
$f(z;0,1)$ or $\phi(z)$

CDF of Z

$$F_Z(z) = P(Z \le z) = \int_{-\infty}^{z} f(y;0,1)\,dy$$

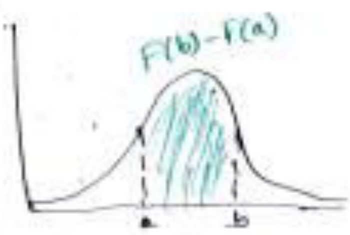
or $\Phi(z)$

$$= \int_{-\infty}^{z} \frac{1}{\sigma\sqrt{2\pi}} \cdot e^{-\frac{y^2}{2}}\,dy$$



The standard normal probability in the form of CDF.

$$P(a \le Z \le b) = F(b) - F(a)$$



* $F(-z) = 1 - F(z)$. Note: CDF of n.d. does not have any analytic soln & its values must be looked up in $N(0,1)$ table.

MGF

$$\boxed{M(t) = \exp\left(\mu t + \frac{t^2\sigma^2}{2}\right)}$$

Normal probabilities

To find probabilities concerning X, we need to convert its values to z scores using

$$\boxed{Z = \frac{X-\mu}{\sigma}}$$

* → when X has the normal distribution with mean μ & std. deviation σ.

$$P(a < X \le b) = P\left(\frac{a-\mu}{\sigma} < \frac{X-\mu}{\sigma} \le \frac{b-\mu}{\sigma}\right)$$

$$= P\left(\frac{a-\mu}{\sigma} < Z \le \frac{b-\mu}{\sigma}\right)$$

$$= \Phi\left(\frac{b-\mu}{\sigma}\right) - \Phi\left(\frac{a-\mu}{\sigma}\right)$$

Ex. Find the probabilities that a r.v. having the std. normal distribution will take on a value

(i) between 0.87 and 1.28

(ii) between −0.34 & 0.62

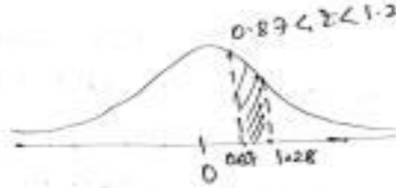
(iii) greater than 0.85

(iv) greater than −0.65.

5

Sol

(i) $P(0.87 < Z < 1.28)$

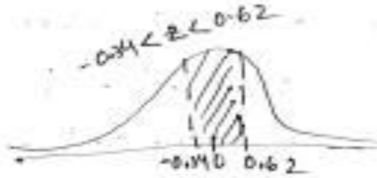


$= \Phi(1.28) - \Phi(0.87)$

$= 0.8997 - 0.8078$

$= 0.0919.$

(ii) $P(-0.34 < Z < 0.62)$



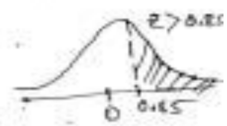
$= \Phi(0.62) - \Phi(-0.34)$

$= \Phi(0.62) - [1 - \Phi(0.34)]$

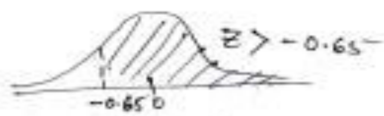
$= 0.7324 - 0.3669 = 0.3655$

(iii) $P(Z > 0.85) = 1 - P(Z < 0.85)$

$= 1 - \Phi(0.85)$



(iv) $P(Z > -0.65)$



$= 1 - P(Z < -0.65)$

$= 1 - \Phi(-0.65)$

$= 1 - [1 - \Phi(0.65)]$

$= \Phi(0.65) = 0.7422.$

Ex - 2 X is normally distributed and the mean of X is 12 and S.D. is 4.

(a) Find out the probability of the following

(i) $X \geq 20$ (ii) $X \leq 20$ (iii) $0 \leq X \leq 12$

(b) Find x', when $P(X > x') = 0.24$

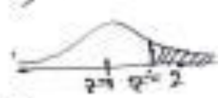
(c) Find x_0' & x_1', when $P(x_0' < X < x_1') = 0.50$ and $P(X > x_1') = 0.25$

Sol we have $\mu = 12$, $\sigma = 4$, i.e. $X \sim N(\underset{\mu,\ \sigma^2}{12, 16})$

(a)

(i) $P(X \geq 20)$

$\Rightarrow P\left(\frac{X-\mu}{\sigma} \geq \frac{20-\mu}{\sigma}\right) = P\left(Z \geq \frac{20-12}{4}\right)$



$\Rightarrow P(Z \geq 2) = 1 - P(Z \leq 2)$
$= 1 - 0.9772 = 0.0228$

(ii) $P(X \leq 20)$
$= P(Z \leq 2) = 0.9772$

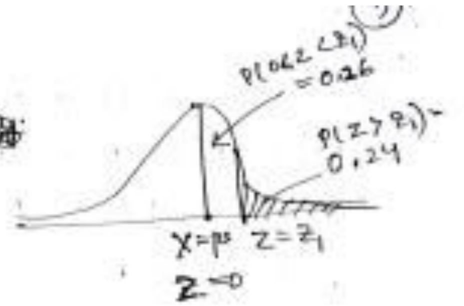
(iii) $P(0 \leq X \leq 12) = P(-3 \leq Z \leq 0)$
$= \Phi(0) - \Phi(-3)$
$= \Phi(0) - (1 - \Phi(3))$
$= 0.5 - (1 - 0.9987)$
$= 0.4987$

(b) Given $P(X > x') = 0.24$

$\Rightarrow P\left(\frac{X-\mu}{\sigma} > \frac{x'-\mu}{\sigma}\right) = P(Z > z_1) = 0.24$

Since, $P(Z > z_1) = 0.24$

then $P(0 < Z < z_1) = 0.26$.



$\Rightarrow \Phi(z_1) - \Phi(0) = 0.26$

$\Rightarrow \Phi(z_1) = 0.26 + 0.5$

$\Rightarrow \Phi(z_1) = 0.76$.

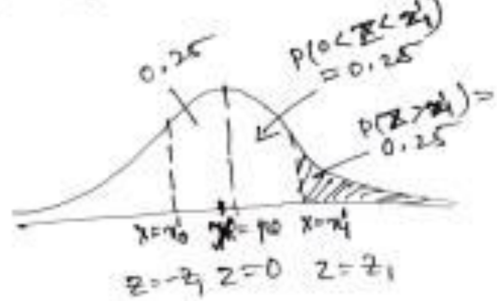
$\Rightarrow z_1 \sim .71$. (from normal table).

hence, $\frac{x_1' - 12}{4} = 0.71 \Rightarrow x' = 12 + 4 \times 0.71 = 14.84$

(c) We are given $P(x_0' < X < x_1') = 0.5$ & $P(X > x_1') = 0.25$

when $X = x_1'$

$$Z = \frac{x_1' - 12}{4} = z_1' \text{ (say)}$$



$X = x_0'$.

$$Z = \frac{x_0' - 12}{4} = -z_1'$$

we have $P(Z > z_1) = 0.25 \Rightarrow P(0 < Z < z_1) = 0.25$

$z_1 = 0.67$ (From normal table).

hence $\frac{x_1' - 12}{4} = 0.67 \Rightarrow x_1' = 12 + 4 \times 0.67 = 14.68$

$\frac{x_0' - 12}{4} = -0.67 \Rightarrow x_0' = 12 - 4 \times 0.67 = 9.32$.

HW $X \sim N(\underset{\mu}{30}, \underset{\sigma^2}{25})$, Find the probabilities that

(i) $26 \le X < 40$ (ii) $X \geq 45$ & (iii) $|X - 30| > 5$.

Exponential Probability Distribution:

A continuous random variable X is said to follow an exponential distribution with parameter $\lambda > 0$, if its probability density function is given by

$$f(x) = \begin{cases} \lambda e^{-\lambda x}, & x \geq 0 \\ 0, & otherwise \end{cases}$$

The general form of the exponential distribution is

$f(x) = \frac{1}{a} e^{-\frac{x}{a}}, \ \ a > 0, \ x \geq 0$ with parameter a.

Momentum generating Function (MGF) of Exponential Distribution:

The MGF is $M_X(t) = \frac{\lambda}{\lambda - t}$

Mean $= \frac{1}{\lambda}$

Variance$= \frac{1}{\lambda^2}$

The cumulative distribution function is

$$F(x) = P(X \leq x) = \int_0^x f(x)dx = \int_0^x \lambda e^{-\lambda x} dx = 1 - e^{-\lambda x}$$

$$F(x) = \begin{cases} 1 - e^{-\lambda x}, & x \geq 0 \\ 0\,, & x < 0 \end{cases}$$

Exponential Distribution possesses memoryless property:

$P(X > s + t\ /X > t) = P(X > s)$, for any $s, t > 0$

The probability density function of X is

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

$P(X > k) = \int_k^\infty \lambda e^{-\lambda x} dx = e^{-\lambda k}$

$$P(X > s + t/\ X > t) = \frac{P(X > s + t\ \cap\ X > t)}{P(X > t)} = \frac{P(X > s + t)}{P(X > t)}$$

$$= \frac{e^{-\lambda(s+t)}}{e^{-\lambda t}} = e^{-st} = P(X > s)$$

Therefore, exponential distribution possesses memoryless property.

Problem 1:

If X has an exponential distribution with mean is 2, find $P(X < 1/\ X < 2)$.

Solution:

Mean of the exponential distribution is

Mean $= \frac{1}{\lambda} = 2$

$$\lambda = \frac{1}{2} = 0.5$$

The probability density function is

$$f(x) = \lambda e^{-\lambda x} = 0.5e^{-0.5\,x}, x \geq 0$$

$$P(X < 1/\ X < 2) = \frac{P(X < 1 \cap X < 2)}{P(X < 2)}$$

$$= \frac{P(X < 1)}{P(X < 2)}$$

$$P(X < 1) = \int_{-\infty}^{1} f(x)dx = \int_{0}^{1} 0.5e^{-0.5x}dx = 0.5\left[\frac{e^{-0.5} - 1}{-0.5}\right] = 0.393\,4$$

$$P(X < 2) = \int_{0}^{x} f(x)dx = \int_{0}^{2} 0.5e^{-0.5x}dx = 0.5\left[\frac{e^{-1} - 1}{-0.5}\right] = 0.6321$$

$P(X < 1/X < 2) = \frac{0.3934}{0.6321} = 0.6223$

Problem 2:

The time (in hours) required to repair a watch is exponentially distributed with parameter $\lambda = \frac{1}{2}$

i. What is the probability that the repair the time exceeds 2 hours?

ii. What is the probability that a repair takes 11 hours given that duration exceeds 8 hours?

with mean 120 days, find the probability that such a watch

iii. Will have to set in less than 24 days, and

iv. Not have to reset in a least 180 days.

Solution:

Let X be the random variable which denotes the time to repair the watch.

The probability density function of the exponential distribution is

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

Given that $\lambda = \frac{1}{2}$

$$f(x) = \lambda e^{-\lambda x} = \frac{1}{2} e^{-\frac{1}{2}x}, x \geq 0$$

i. $P(X > 2) = \int_2^\infty \frac{1}{2} e^{-\frac{1}{2}x} dx = e^{-1}$

ii. Using the memoryless property, we have

$$P(X \geq 11/ X > 8) = P(X > 3)$$

$P(X > 3) = \int_3^\infty \frac{1}{2} e^{-\frac{1}{2}x} dx = e^{-1.5}$

In the second case, given

Mean=120 i.e., Mean $= \frac{1}{\lambda} = 120$

$$\lambda = \frac{1}{120}$$

The probability density function is given by

$$f(x) = \lambda e^{-\lambda x} = \frac{1}{120} e^{-\frac{1}{120}x}, x \geq 0$$

iii. $P(X < 24) = \int_0^{24} \frac{1}{120} e^{-\frac{1}{120}x} dx = 1 - e^{-0.2} = 0.1813$

iv. $P(X > 180) = \int_{180}^\infty \frac{1}{120} e^{-\frac{1}{120}x} dx = e^{-1.5} = 0.2231$

Problem 3:

The time line in hours required to repair a machine is exponentially distributed with parameter $\lambda = \frac{1}{2}$. What is the probability that the required time

i. Exceeds 2 hours

ii. Exceed 5 hours

Solution:

Let X be the random variable which denotes the time to repair the machine. Then the density function of X is given by

$$f(x) = \lambda e^{-\lambda x} = \frac{1}{2} e^{-\frac{1}{2}x}, \qquad x > 0$$

i. $P(X > 2) = \int_2^\infty \frac{1}{2} e^{-\frac{1}{2}x} dx = e^{-1}$

ii. $P(X > 5) = \int_5^\infty \frac{1}{2} e^{-\frac{1}{2}x} dx = e^{-\frac{5}{2}} = 0.082$

Try yourself:

Problem 4:

A component has an exponentially time of failure distribution with mean 10,000 hours

i. The component has already been in operation for its mean life. What is the probability that it will fail by 15,000 hours?

ii. At 15,000 hours the component is still in operation life. What is the probability that it operates for another 5,000 hours?

Problem 5:

The mileage which car owners get with certain kind of radial tyre is a random variable having an exponential distribution with mean 40,000 km. Find the probabilities that one of these tyres will last

i. At least 20,000 km

ii. At most 30,000 km.

Gamma Distribution:

A continuous random variable X is said to follow general Gamma distribution with two parameters $\boldsymbol{\lambda} > \mathbf{0}$ **and** $\boldsymbol{k} > \mathbf{0},$ if its probability density function is given by

$$\boldsymbol{f(x)} = \begin{cases} \dfrac{\boldsymbol{\lambda^k x^{k-1} e^{-\lambda x}}}{\boldsymbol{\Gamma(k)}}, \boldsymbol{x \geq 0} \\ \mathbf{0}, \qquad\qquad \boldsymbol{otherwise} \end{cases}$$

Note:

1. When $\boldsymbol{k = 1}$, the distribution is called exponential distribution
2. $\int_{-\infty}^{\infty} f(x)dx = 1$ $\left(\text{Since, } \int_0^{\infty} x^{k-1}e^{-ax}\, dx = \frac{\Gamma(k)}{a^k}\right)$

Momentum generating Function of Gamma Distribution:

The probability density function of the general Gamma random variable X is

$$f(x) = \begin{cases} \dfrac{\lambda^k x^{k-1} e^{-\lambda x}}{\Gamma(k)}, x \geq 0 \\ 0, \qquad\qquad otherwise \end{cases}$$

Where $\lambda\ and\ k$ are the parameters.

The MGF is

$$\boldsymbol{M_X(t)} = \left(\frac{\boldsymbol{\lambda}}{\boldsymbol{\lambda - t}}\right)^{\boldsymbol{k}}$$

Mean $= \frac{\boldsymbol{k}}{\boldsymbol{\lambda}}$

Variance $=\frac{\boldsymbol{k}}{\boldsymbol{\lambda^2}}$

Problem 1:

The lifetime (in hours) of a certain piece of equipment is a continuous random variable having range $0 < x < \infty$ and the PDF is $f(x) = \begin{cases} xe^{-kx}, & 0 < x < \infty \\ 0, & otherwise \end{cases}$. Determine the constant k and evaluate the probability that the lifetime exceeds 2 hours.

Solution:

Let X denote the lifetime of a certain piece of equipment with PDF

$$f(x) = \begin{cases} xe^{-kx}, & 0 < x < \infty \\ 0, & otherwise \end{cases}$$

Now we have to find k

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

$$\int_0^{\infty} xe^{-kx}dx = \int_0^{\infty} x^{2-1}\, e^{-kx}dx = 1$$

Using

$$\int_0^{\infty} x^{n-1}e^{-ax}dx = \frac{\Gamma(n)}{a^n}$$

$$\frac{\Gamma(2)}{k^2} = 1$$

$$k^2 = 1$$

$$k = 1$$

Then

$$f(x) = \begin{cases} xe^{-x}, & 0 < x < \infty \\ 0, & otherwise \end{cases}$$

P (lifetime exceeds 2 hours) $= P(X > 2) = \int_2^{\infty} f(x)dx = \int_2^{\infty} xe^{-x}dx = 0.4060$

Problem 2:

The daily consumption of milk in a city, in excess of 20,000 liters, is approximately distributed as a Gamma variate with parameters $k = 2 \; and \; \lambda = \frac{1}{10{,}000}$. The city has a daily stock of 30,000 liters. What is the probability that the stock is insufficient on a particular day?

Solution:

If the random variable X denotes the daily consumption of milk (in liters) in a city, then the random variable $Y = X - 20{,}000$ has a Gamma distribution with probability density function

$$f(x) = \frac{\lambda^k x^{k-1} e^{-\lambda x}}{\Gamma(k)}, x \geq 0$$

$$f(y) = \frac{\lambda^2 y^{2-1} e^{-\lambda y}}{\Gamma(2)}, y \geq 0$$

$$= \frac{\left(\frac{1}{10{,}000}\right)^2 y^{2-1} e^{-\left(\frac{1}{10{,}000}\right)y}}{\Gamma(2)}, y \geq 0$$

Since the daily stock of the city is 30,000 liters, the required probability that the stock is insufficient on a particular day is given by

$$P(X > 30{,}000) = P(Y > 10{,}000) = \int_{10{,}000}^{\infty} f(y)dy$$

$$= \int_{10{,}000}^{\infty} \left(\frac{1}{10{,}000}\right)^2 \frac{y^{2-1} e^{-\frac{y}{10{,}000}}}{\Gamma(2)} dy = \int_{1}^{\infty} ze^{-z} dz$$

Taking $z = \frac{y}{10{,}000}$

$$\int_{1}^{\infty} ze^{-z} dz = e^{-1} + e^{-1} = 2e^{-1} = 0.7357$$

Problem 3:

In a certain city, the daily consumption of electric power (in millions of kilowatt-hours) can be treated as a random variable having Gamma distribution with parameters $\lambda = \frac{1}{2}$ and $k = 3$. If the power plant of this city has a daily capacity of 12 million kilowatt hours, what is the probability that this power supply will be adequate on any day?

Solution:

Let X be the random variable denoting the daily consumption of electric power (in millions of kilowatt hours).

Also, given $\lambda = \frac{1}{2}$ and $k = 3$.

Gamma distribution with PDF is

$$f(x) = \frac{\lambda^k x^{k-1} e^{-\lambda x}}{\Gamma(k)}$$

$$= \frac{\left(\frac{1}{2}\right)^3 x^{3-1} e^{-\frac{x}{2}}}{\Gamma(3)}, x \geq 0$$

The daily capacity of the power plant is 12 million kilowatt hours. The power supply is more than 12 million on any day.

$$P(X > 12) = \int_{12}^{\infty} f(x)dx = \int_{12}^{\infty} \frac{\left(\frac{1}{8}\right) x^2 e^{-\frac{x}{2}}}{\Gamma(3)} dx$$

$$= \int_{12}^{\infty} \frac{\left(\frac{1}{8}\right) x^2 e^{-\frac{x}{2}}}{\Gamma(3)} dx = 0.0625$$

Try yourself:

Problem 4:

Consumer demand for milk in a certain locality per month is known to be a general Gamma distribution random variable. If the average demand is a liters and the most likely demand is b liters $(b < a)$, what is the variance of the demand?

Weibull distribution:

The random variable X is said to follow Weibull distribution, if its probability distribution is given by

$$f(x) = \begin{cases} \alpha\beta x^{\beta-1} e^{-\alpha x^{\beta}}, & x > 0 \\ 0, & otherwise \end{cases}$$

Where $\alpha > 0$ and $\beta > 0$ are two parameters of the Weibull distribution.

Note:

When $\beta = 1$, the Weibull distribution reduces to the exponential distribution with parameter α.

Mean and Variance of Weibull distribution:

The probability density function of Weibull distribution is given by

$$f(x) = \begin{cases} \alpha\beta x^{\beta-1} e^{-\alpha x^{\beta}}, & x > 0 \\ 0, & otherwise \end{cases}$$

Where $\alpha > 0$ and $\beta > 0$ are two parameters.

Mean = $E(X) = \mu = \alpha^{-\frac{1}{\beta}}\Gamma\left(1 + \frac{1}{\beta}\right)$

Variance = $\sigma^2 = \alpha^{-2/\beta}\left\{\Gamma\left(1 + \frac{2}{\beta}\right) - \left[\Gamma\left(1 + \frac{1}{\beta}\right)\right]^2\right\}$

Cumulative distribution function:

$$F(x;\alpha,\beta)=\begin{cases}1-e^{-\alpha x^{\beta}}, & x\geq 0\\ 0, & x<0\end{cases}$$

Problem 1:

Suppose that the lifetime of a certain kind of an emergency backup battery (in hours) is a random variable X having Weibull distribution with $\alpha = 0.1\ and\ \beta = 0.5$. Find

i. The mean lifetime of these batteries

ii. The probability that such battery will last more than 300 hours.

Solution:

i. Mean is $\mu = \alpha^{-\frac{1}{\beta}}.\Gamma\left(1+\frac{1}{\beta}\right)$

$$= (0.1)^{-\frac{1}{0.5}}\,\Gamma\left(1+\frac{1}{0.5}\right) = \frac{2}{\left(\frac{1}{10}\right)^{2}} = 200\ hours$$

ii. $P(X > 300) = \int_{300}^{\infty} \alpha\beta x^{\beta-1}e^{-\alpha x^{\beta}}dx$

$$= \int_{300}^{\infty}(0.1)(0.5)x^{-0.5}e^{-0.1(x)^{0.5}}dx$$

$$= 0.1769$$

Problem 2:

Suppose that the time to failure (in minutes) of certain electronic components subjected to continuous vibrations may be looked upon as a random variable having the Weibull distribution with $\alpha = \frac{1}{5}\ and\ \beta = \frac{1}{3}$

i. How long can such a component be expected to last?

ii. What is the probability that such a component will fail in less than 5 hours.

Solution:

i. Mean is $\mu = \alpha^{-\frac{1}{\beta}}\,\Gamma\left(1+\frac{1}{\beta}\right)$

$$= \left(\frac{1}{5}\right)^{-\frac{1}{\left(\frac{1}{3}\right)}} \Gamma\left(1 + \frac{1}{\left(\frac{1}{3}\right)}\right) = 5^3 3! = 750\ minutes$$

ii. $P(X < 5\ hours) = P(X < 300\ minutes)$

$$= \int_0^{300} \alpha\beta x^{\beta-1} e^{-\alpha x^\beta} dx = \int_0^{300} \left(\frac{1}{5}\right)\left(\frac{1}{3}\right) x^{\frac{1}{3}-1} e^{-\left(\frac{1}{5}\right)(x)^{\frac{1}{3}}} dx = 0.7379$$

Or

$$F\left(300; \frac{1}{5}, \frac{1}{3}\right) = \begin{cases} 1 - e^{-\alpha x^\beta}, & x \geq 0 \\ 0, & x < 0 \end{cases}$$

$$= 1 - e^{-\frac{1}{5}(300)^{\frac{1}{3}}} = 0.7379$$

The failure rate of the Weibull distribution:

- The Weibull distribution helps to determine the failure rate (or hazard rate) in order to get a sense of deterioration of the component.
- Consider the reliability of a component or product as the probability that it will function property for at least a specified time to under specified experimental conditions.
- Then the **failure rate** at time ‘t’ for the Weibull distribution is given by

$$\boldsymbol{Z(t) = \alpha\beta t^{\beta-1}, t > 0}$$

Interpretation of the failure rate;

1. If $\beta = 1$, the failure rate=α, a constant. This is the special case of the exponential distribution in which lack of memoryless property.
2. If $\beta > 1$, $Z(t)$ is an increasing function of time ‘t’, which indicates that the component **wears over time**.

3. If $\beta < 1$, $Z(t)$ is decreasing function of time 't' and hence the component strengthens or **hardness over time**.

Problem:

The length of life X, in hours of an item in a machine shop has a Weibull distribution with $\alpha = 0.01 \; and \; \beta = 2$

i. What is the probability that it fails before eight hours of usage?

ii. Determine the failure rates.

$$\boldsymbol{F(x;\alpha,\beta)} = \begin{cases} \mathbf{1} - \boldsymbol{e^{-\alpha x^{\beta}}}, & \boldsymbol{x \geq 0} \\ \mathbf{0}, & \boldsymbol{x < 0} \end{cases}$$

Solution:

i. $P(X < 8) = F(8) = 1 - e^{-(0.01)8^2} = 1 - 0.257 = 0.473$

ii. Here $\beta = 2$, and hence it wears over time and the failure rate is given by

$$Z(t) = 0.02 \; t$$

If $\beta = \frac{3}{4} \; and \; \alpha = 2$, then

$$Z(t) = 1.5 \; t^{1/4}$$

Hence the component gets stronger over time.