UNIT II
Some Special Probability Distributions

1 Binomial Distribution

Definition

Let A be some event associated with a random experiment E, such that $P(A) = p$ and $P(\bar{A}) = 1 - p = q$. Assuming that p remains the same for all repetations, if we consider n independent repetations (or trials) of E and if the random variable (RV) X denotes the number of times the event A has occured, then X is called a *binomial random varaiable* with parameters n and p or we say that X follows a *binomial distribution* with parameters n and p, or symbolically $B(n, p)$. Obviously the possible values that X can take, are $0, 1, 2, ..., n$.
By the theorem under Bernoulli's trials in Chapter 1, the probability mass function of a binomial RV is given by

$$P(X = x) = nC_r p^x q^{n-x}; x = 0, 1, 2, ..., n \quad \text{where} \quad p + q = 1$$

Examples:

1. Tossing coin once results is head or tail.
2. Toss a dice once, the outcomes can be classified as 'even number' and 'odd number'.
3. Firing at a target results in hit or miss.

1.1 Moment Generating Function (MGF) of Binomial distribution

Binomial distribution is

$$p(x) = P(X = x) = nC_x p^x q^{n-x}; x = 0, 1, 2, ..., n$$

$$\text{MGF is} \, \mathrm{M}_X(t) = E[e^{tX}]$$

$$= \sum_{x=0}^{n} e^{tx} p(x)$$

$$= \sum_{x=0}^{n} e^{tx} nC_x p^x q^{n-x}$$

$$= \sum_{x=0}^{n} nCx (pe^t)^x q^{n-x}$$

$$= (q + pe^t)^n$$

Mean and Variance of Binomial distribution

We know that MGF is

$$\begin{aligned} M_X(t) &= (q+pe^t)^n \\ \therefore M'_X(t) &= n(q+pe^t)^{n-1}.pe^t \\ M''_X(t) &= np\left\{(n-1)(q+pe^t)^{n-2}(pe^t).e^t+(q+pe^t)^{n-1}.e^t\right\} \\ &= np(q+pe^t)^{n-2}e^t(n-1)pe^t+q+pe^t \end{aligned}$$

$$\begin{aligned} \text{putting } t=0, \text{we get } M'_X(0) &= n(q+p)^{n-1}.p \\ \implies mean &= \mu'_1 = np \\ M''_X(0) &= np(n-1)p+q+p \\ \implies \mu'_2 &= np(np+q) = n^2p^2+npq \\ \text{Var(X)} &= E(X^2)-[E(X)]^2 = \mu'_2-(\mu'_1)^2 \\ &= n^2p^2+npq-n^2p^2 = npq \end{aligned}$$

Problems

1. The mean of a binomial distribution 20 and standard deviation is 4. Find the parameters of the distribution.
Solution:
Let (n,p) be the parameters of the binomial distribution, then

$$\text{mean} = np \text{ and S.D}= \sqrt{npq}$$
$$\text{Given, } np=20 \text{ and } \sqrt{npq}=4$$
$$npq=16$$
$$\therefore 20.q = 16$$
$$\implies q = \frac{4}{5}$$
$$\therefore p = 1-q = 1-\frac{4}{5} = \frac{1}{5}$$
$$\therefore n.\frac{1}{5} = 20 \implies n = 100$$
$$\therefore \text{the parameters are } 100, \frac{1}{5}.$$

2. Six dice are thrown 729 times. How many times do you expect atleast 3 dice to show a five or a six?
Solution:
Success is getting 5 or 6 in a die. Let X denote the numer of success when 6 dice are thrown.
$\therefore X$ is a binomial R.V with parameter (n,p)

$$\therefore P(X=x) = nC_xp^xq^{n-x}, x=0,1,2,...,n$$

Given $n=6$ and $p = P(5\,\text{or}\,6) = \frac{2}{6} = \frac{1}{3}$

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

$$\therefore P(X=x) = 6C_x\left(\frac{1}{3}\right)^x\left(\frac{2}{3}\right)^{6-x}, x=0,1,2,...6$$

$$
\begin{aligned}
P(X \geq 3) &= 1 - P(X < 3) \\
&= 1 - \left\{ \left(\frac{2}{3}\right)^6 + 6C_1 \frac{1}{3} \left(\frac{2}{3}\right)^5 + 6C_2 \left(\frac{1}{3}\right)^2 . \left(\frac{2}{3}\right)^4 \right\} \\
&= 1 - \frac{1}{3^6} \{64 + 192 + 240\} \\
&= 1 - \frac{496}{729} = \frac{233}{729}
\end{aligned}
$$

When 6 dice are thrown 729 times, the number of times atleast 3 dice show 5 or 6 is $729 \times \frac{233}{729}$=233.

3.Out of 800 families with 4 children each, how many families would be expected to have **(a)** 2 boys and 2 girls,**(b)** atleast 1 boy, **(c)** at most 2 girls, and **(d)** children of both sexs. Assume equal probabilities for boys and girls.

Solution:

Considering each child as a trial, $n = 4$. Assuming that birth of a boy is success, $p = \frac{1}{2}$ and $q = \frac{1}{2}$. Let X denote the number os successes (boys)

$$
\begin{aligned}
\text{(a)}P(2\text{ boys and}\,2\,\text{girls}) &= P(X = 2) \\
&= 4C_2 . \left(\frac{1}{2}\right)^2 . \left(\frac{1}{2}\right)^{4-2} \\
&= 6 \times \left(\frac{1}{2}\right)^4 = \frac{3}{8}
\end{aligned}
$$

$\therefore$ No.of families having 2 boys and 2 girls

$$
\begin{aligned}
&= N.P(X = 2)(\text{where}\,N\,\text{is the total no.of families considered}) \\
&= 800 \times \frac{3}{8} \\
&= 300.
\end{aligned}
$$

$$
\begin{aligned}
\text{(b)}P(\text{atleast 1 boy}) = P(X \geq 1)& \\
&= P(X = 1) + P(X = 2) + P(X = 3) + P(X = 4) \\
&= 1 - P(X = 0) \\
&= 1 - 4C_0 . \left(\frac{1}{2}\right)^0 . \left(\frac{1}{2}\right)^4 \\
&= 1 - \frac{1}{16} = \frac{15}{16}
\end{aligned}
$$

$\therefore$ No.of families having atleast 1 boy

$$
= 800 \times \frac{15}{16} = 750.
$$

$$\begin{aligned}
(c)P(\text{at most 2 girls}) &= P(\text{exactly 0 girl, 1 girl or 2 girl}) \\
&= P(X=4, X=3 \text{or} X=2) \\
&= 1-\{P(X=0)+P(X=1)\} \\
&= 1-\left\{4C_0.\left(\frac{1}{2}\right)^4+4C_1.\left(\frac{1}{2}\right)^4\right\} \\
&= \frac{11}{16}
\end{aligned}$$

$\therefore$ No.of families having atmost 2 girls

$$= 800\times\frac{11}{16}=550.$$

$$\begin{aligned}
(d)P(\text{children in both sexs}) & \\
&= 1-P(\text{children in same sex}) \\
&= 1-\{P(\text{all are boys})+P(\text{all are girls})\} \\
&= 1-\{P(X=4)+P(X=0)\} \\
&= 1-\left\{4C_4.\left(\frac{1}{2}\right)^4+4C_0.\left(\frac{1}{2}\right)^4\right\} \\
&= 1-\frac{1}{8}=\frac{7}{8}
\end{aligned}$$

$\therefore$ No.of families having children of both sexs

$$= 800\times\frac{7}{8}=700.$$

4. An irregular 6-faced die is such that the probability that is given 3 even numbers in 5 throws is twice the probability that it gives 2 even numbers in 5 throws. How many sets of exactly 5 trials can be expected to give no even number out of 2500 sets?

Solution:

Let the probability of getting an event number with the unfair die be p.

Let X denote the number of even numbers obtained in 5 trials (throws).

$$\begin{aligned}
P(X=3) &= 2\times P(X=2) \\
\text{i.e.,}\quad 5C_3p^3q^2 &= 2\times 5C_2p^2q^3 \\
\text{i.e.,}\quad p &= 2q=2(1-p)
\end{aligned}$$

$$\therefore\qquad 3p=2 \text{ or} p=\frac{2}{3} \text{ and } q=\frac{1}{3}$$

$$\begin{aligned}
\text{Now,}P(\text{getting no even number}) & \\
&= P(X=0) \\
&= 5C_0.p^0.q^5=\left(\frac{1}{3}\right)^5=\frac{1}{243}
\end{aligned}$$

$\therefore$ Number of sets having no success (even number) out of N sets $= N\times P(X=0)$

$\therefore$ Required number of sets $= 2500\times\dfrac{1}{243}=10,$ nearly

5. Two dice are thrown 120 times. Find the average number of times in which the number of the first dice exceeds the number on the second dice.

Solution:

The number on the first dice exceeds that on the second dice, in the following combinations: (2,1);(3,1),(3,2);(4,1),(4,2),(4,3);(5,1),(5,2),(5,3),(5,4);(6,1);(6,2),(6,3),(6,4),(6,5), where the numbers in the parentheses represent the numbers in the first and second dice respectively.

$\therefore$ P(success) = P(number in the first dice exceeds the number in the second dice)

$$= \frac{15}{36} = \frac{5}{12}$$

This probability remains the same in all the throws that are independent.

If X is the number of success, then X follows a binomial distributions with parameters $n(=120)$ and $p(=\frac{5}{12})$.

$\therefore E(X) = np$=120 $\times \frac{5}{12} = 50$

6. Fit a binomial distribution for the following data:

x:	0	1	2	3	4	5	6	Total
f:	5	18	28	12	7	6	4	80

Fitting a binomial distribution means assuming that the given distribution is approximately binomial and hence finding the probability mass function and then finding the theoretical frequencies.

To find the binomial frequency distribution $N(q+p)^n$, which fits the given data, we required N, n and p. We assume N=total frequency = 80 and n= no. of trials =6 from the given data.

To find p, we compute the mean of the given frequency distribution and equate it to np (means of the binomial distribution).

x:	0	1	2	3	4	5	6	Total
f:	5	18	28	12	7	6	4	80
fx:	0	18	56	36	28	30	24	192

$$\bar{x} = \frac{\sum fx}{\sum f} = \frac{192}{80} = 2.4$$
$$\text{i.e.,} \quad np = 2.4 \text{or } 6p = 2.4$$
$$\therefore \quad p = 0.4 \text{and } q = 0.6$$

If the given distribution is nearly binomial, the theoretical frequencies are given by the successive terms in the expansion of $80(0.6+0.4)^6$. Thus we get,

x:	0	1	2	3	4	5	6
Theoretical f:	3.73	14.93	24.88	22.12	11.06	2.95	0.33

Converting these values into whole numbers consistent with the condition that the total frequency is 80, the corresponding binomial frequency distribution is as follows:

x:	0	1	2	3	4	5	6	Total
f:	4	15	25	22	11	3	0	80

2 Poisson Distribution

Definition

If X is a discrete RV that can assume the values $0, 1, 2, ...,$ such that its probability mass function is given by

$$P(X = x) = \frac{e^{-\lambda}\lambda^x}{x!}; \quad x = 0, 1, 2, ...; \qquad \lambda > 0$$

then X is said to follow a *Poisson distribution* with parameter λ or symbolicaly X is said to follow $P(\lambda)$.

Poisson Distribution on Limiting form of Binomial Distribution

Poisson distribution is a limiting case of binomial distribution under the following conditions:

1. n, the number of trials is indefinitely large, i.e., $n \to \infty$,
2. p, the constant probability of success in each trial is very small, i.e., $p \to 0$.
3. $np (= \lambda)$ is finite or $\frac{\lambda}{n}$ and $q = 1 - \frac{\lambda}{n}$, where λ is a positive real number.

Examples

1. The number of defective items produced in a factory in one day.
2. The number of deaths due to rare disease.
3. The no. of mistakes in a page of a book.
4. The number of earthquakes in a given time interval.

2.1 Moment Generating Function of Poisson distribution

Let X be a Poisson randon variable with parameter λ
$\therefore$ Poisson distribution is given by $p(x) = P(X = x) = \dfrac{e^{-\lambda}\lambda^x}{x!}, x = 0, 1, 2, ...$
Moment generating function of X is

$$\begin{aligned} M_X(t) = &\quad E[e^{tX}] = \textstyle\sum_{x=0}^{\infty} e^{tx} p(x) \\ = &\quad \textstyle\sum_{x=0}^{\infty} e^{tx} \dfrac{e^{-\lambda}\lambda^x}{x!} \\ = &\quad e^{-\lambda} \textstyle\sum_{x=0}^{\infty} \dfrac{(\lambda e^t)^x}{x!} \\ = &\quad e^{-\lambda}.e^{\lambda e^t} \\ M_X(t) = &\quad e^{\lambda(e^t - 1)} \end{aligned}$$

Mean and Variance of Poisson distribution

We have

$$\begin{aligned}
M_X(t) &= e^{\lambda(e^t-1)} \\
\therefore\ M_X'(t) &= e^{\lambda(e^t-1)}.\lambda e^t \\
M_X''(t) &= \lambda\left\{e^{\lambda(e^t-1)}.e^t + e^t.e^{\lambda(e^t-1)}.\lambda e^t\right\} \\
&= \lambda e^{\lambda(e^t-1)}.e^t\{1+\lambda e^t\} \\
\text{putting } t=0,\ M_X'(0) &= \lambda \implies \mu_1' = \lambda \\
M_X''(0) &= \lambda(1+\lambda) \implies \mu_2' = \lambda(1+\lambda) \\
\therefore\ \text{Mean} &= \mu_1' = \lambda \\
\text{Var(X)} &= \mu_2' - (\mu_1')^2 \\
&= \lambda + \lambda^2 - \lambda^2
\end{aligned}$$

Note: The mean and variance of Poisoon distribution are equal.

Problems

1. If X is a poisson R.V such that $P(X=2) = 9P(X=4) + 90P(X=6)$, then find (i) the variance, (ii) the mean , $E(X^2), P(X \geq 2)$.
Solution:

$$\text{Poisson distribution is } P(X=x) = \frac{e^{-\lambda}\lambda^x}{x!}, x=0,1,2,...$$

$$\text{Given } P(X=2) = 9P(X=4) + 90P(X=6)$$

$$\implies e^{-\lambda}\frac{\lambda^2}{2!} = 9.\frac{e^{-\lambda}\lambda^4}{4!} + 90.\frac{e^{-\lambda}\lambda^6}{6!}$$

$$\implies \frac{1}{2} = \frac{3}{8}\lambda^2 + \frac{1}{8}\lambda^4 \qquad [\text{Dividing by} e^{-\lambda}\lambda^2]$$

$$\implies \lambda^4 + 3\lambda^2 - 4 = 0$$

$$\implies (\lambda^2+4)(\lambda^2-1) = 0$$

$$\implies \lambda = 1 \qquad [\therefore \lambda > 0]$$

$$\text{We know that Mean} = \lambda = 1;\ \text{Var(X)}\lambda = 1$$

$$E(X^2) = \lambda^2 + \lambda = 1+1 = 2$$

$$P(X \geq 2) = 1 - P(X<2)$$

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

2. Find the probability that at most 4 defective fuses will be found in a box of 200 fuses if experience shows that 2% of such fuses are defective.(given $e^{-4} = 0.0183$).
Solution:

Let X denote the number of defectives in a box of 200 fuses then
p = probability of defective fuses.

$$= \frac{2}{100} = 0.012$$

$$\text{Given } n = 200 \text{ and } p \text{ is very small.}$$
$$\therefore X \text{ is a Poisson R.V with parameter } \lambda$$
$$\therefore \lambda = np = 200 \times \frac{2}{100} = 4.$$
$$\text{So, the poisson distribution is } P(X = x) = \frac{e^{-\lambda}\lambda^x}{x!}, x = 0, 1, 2, ...$$
$$= \frac{e^{-4}4^x}{x!}, x = 0, 1, 2, 3, ...$$
$$\text{Required } P(X \leq 4) = P(X = 0) + P(X = 1) + P(X = 2) + P(X = 3) + P(X = 4)$$
$$= e^{-4} + e^{-4}.\frac{4}{1!} + e^{-4}.\frac{4^2}{2!} + e^{-4}.\frac{4^3}{3!} + e^{-4}.\frac{4^4}{4!}$$
$$= e^{-4}\left(1 + 4 + 8 + \frac{32}{3} + \frac{32}{3}\right)$$
$$= e^{-4}\left(13 + \frac{64}{3}\right)$$
$$= 0.0183 \times \frac{103}{3} = 0.628$$

3.The number of accidents in a year attributed to taxi drivers in a city follows a poisson distribution with mean equal to3. Out of 1000 taxi drivers, find approximately the number of drivers with (i) no accident in a year, (ii) more than 3 accidents in a year.

Solution:

Let X denote the number of accidents in a yaer due to taxi drivers.
Given X is a Poisson distribution. Let λ be the parameter.
Given mean =3 $\implies \lambda = 3$

$$\therefore \text{Poisson distribution is given by } P(X = x) = \frac{e^{-\lambda}\lambda^x}{x!}, \quad x = 0, 1, 2, ...$$
$$= \frac{e^{-3}.3^x}{x!}, \quad x = 0, 1, 2, ...$$

(i)$P(X = 0) = e^{-3} = 0.0498$
$\therefore$ number of drivers with no accidents =$100 \times 0.0498 = 49.8$
i.e., number of drivers with no accidents =50.

$$\text{(ii) } P(X > 3) = 1 - P(X \leq 3)$$
$$= 1 - \{P(X = 0) + P(X = 1) + P(X = 2) + P(X = 3)\}$$
$$= 1 - e^{-3}\left(1 + 3 + \frac{3^2}{2!} + \frac{3^3}{3!}\right)$$
$$=1-0.0498 \times 13$$
$$=1-0.6474=0.3526$$
$$\therefore \text{number of drivers with more than 3 accidents}$$
$$=1000 \times 0.3526=353.$$

4.The number of monthly breakdown of a computer is a random variable having a poisson distribution with mean equal to 1.8. Find the probability that this computer will function for a month
(i)without breakdown

(ii) with only one breakdown and
(iii)with atleast one breakdown
Solution:
Let X denote the number of breakdowns of a computer in a month.
Given X follows Poisson distribution with mean $=\lambda$=1.8.
$\therefore$ Poisson distribution is given by

$$\begin{aligned} P(X=x) &= \frac{e^{-\lambda}\lambda^x}{x!}, \quad x=0,1,2,... \\ &= \frac{e^{-1.8}(1.8)^x}{x!}, \quad x=0,1,2,... \end{aligned}$$

(1)$P(X=0)=e^{-1.8}=0.1653$
(2)$P(X=1)=e^{-1.8}(1.8)=0.2975$
(3)$P(X\geq 1)=1-P(X=0)=1-0.1653=0.8347$

5. Fit a poisson distribution to the data:

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

Solution:
Poisson distribution

$$P(X=x)=\frac{e^{-\lambda}\lambda^x}{x!}, \quad x=0,1,2,...$$

Mean=$\lambda=\frac{\sum fx}{\sum f}$

x:	0	1	2	3	4	5	Total
f:	142	156	69	27	5	1	400
fx:	0	156	138	81	20	5	400

$\therefore$ Mean = $\frac{\sum fx}{\sum f}$=$\frac{400}{400}=1$

$$\lambda=1$$

Fitting of Poisson distribution:

x:	0	1	2	3	4	5
observed frequency:	142	156	69	27	5	1
Expected frequency:	147.15	147.15	73.58	24.53	6.13	1.23
Expected(Approx):	147	147	74	25	6	1

3 Geometric Distribution

Definition

Let the RV X denote the number of trials of a random experiment required to obtain the first success (occurence of an event A).Obviously, X can assume the values $1,2,3,...$

Now, $X = x$, if and only if the first $(x-1)$ trials result in failure(occurance of $\overline{A}$) and the xth trial results in success (occurrence A). Hence,

$$P(X = x) = q^{x-1}p; x = 1,2,3,...,\infty$$

where $P(A) = p$ and $P(\overline{A}) = q$.
If X is a discrete RV that can assume the values $1,2,3,...,\infty$ such that its probability mass function is given by

$$P(X = x) = q^{x-1}p; x = 1,2,3,...,\infty \quad \text{where } p+q = 1$$

then X is said to follow a *geometric distribution*.

3.1 Moment generating Function of geometric distribution

The geometric distribution with parameter p is

$$P(X = x) = q^{x-1}p, \qquad x = 1,2,3,...$$

$$M_X(t) = E(e^{tX})$$

$$= \sum_{x=1}^{\infty} e^{tx}p(x)$$

$$= \sum_{x=1}^{\infty} e^{tx}q^{x-1}p$$

$$= \sum_{x=1}^{\infty} e^{t}.e^{t(x-1)}q^{x-1}p$$

$$= pe^t \sum_{x=1}^{\infty}(qe^t)^{x-1}$$

$$= pe^t[1+(qe^t)+(qe^t)^2+...\infty]$$

$$pe^t \frac{1}{1-qe^t} \text{ if } \quad qe^t < 1$$

$$= \frac{pe^t}{1-qe^t}$$

Mean and variance

$$
\begin{aligned}
M_X(t) = &\ \frac{pe^t}{1-qe^t} \\
\therefore M_X'(t) =& p\left\{\frac{(1-qe^t)e^t - e^t(-qe^t)}{(1-qe^t)^2}\right\} \\
=& pe^t\left\{\frac{1-qe^t+qe^t}{(1-qe^t)^2}\right\} = \frac{pe^t}{(1-qe^t)^2} \\
\text{Now } M_X''(t) =& p\left\{\frac{(1-qe^t)^2.e^t - e^t.2(1-qe^t)(-qe^t)}{(1-qe^t)^4}\right\} \\
=&\ pe^t(1-qe^t)\left\{\frac{1-qe^t+2qe^t}{(1-qe^t)^4}\right\} \\
=& \frac{pe^t(1+qe^t)}{(1-qe^t)^3} \\
\therefore M_X'(0) = &\ \frac{p}{(1-q)^2} = \frac{p}{p^2} = \frac{1}{p} \implies \text{mean} = \frac{1}{p} \\
M_X''(0) =& \frac{p(1+q)}{(1-q)^3} = \frac{p(1+q)}{p^3} \\
\implies \mu_2' = &\ \frac{1}{p^2} + \frac{q}{p^2} \\
\text{Var(X)} = &\ \mu_2' - (\mu_1')^2 \\
=& \frac{1}{p^2} + \frac{q}{p^2} - \frac{1}{p^2} = \frac{q}{p^2}
\end{aligned}
$$

Property of Geometric Distribution:

[Geometric Random Variable has memoryless property]
Definition:
A non-negative random variable X is memoryless if

$$P(X > s+t/X > s) = P(X > t) \text{for all } s,t \geq 0.$$

Proof: The geometric distribution is

$$
\begin{aligned}
P(X = x) = &\ \ q^{x-1}p,\ x = 1,2,3,... \\
P(X > m) = &\ \ P(X = m+1) + P(X = m+2) + ...\infty \\
= &\ \ q^m p + q^{m+1}p + q^{m+2}p + ...\infty \\
= &\ \ q^m p[1+q+q^2+...+\infty] \\
= &\ \ \frac{q^m p}{1-p} = \frac{q^m p}{p} = q^m \\
\therefore P(X > s+t/X > s) = &\ \ \frac{P(X > s+t \cap X > s)}{P(X > s)}
\end{aligned}
$$

Since the event $\{X > s+t\} \subset \{X > s\}$,

$$\begin{aligned} \{X > s+t\} \cap \{X > s\} &= \{X > s+t\} \\ \therefore P(X > s+t/X > s) &= \frac{P(X > s+t)}{P(X > s)} \\ &= \frac{q^{s+t}}{q^s} = q^t = P(X > t) \end{aligned}$$

Note: This property is also known as Markov property and ageless property.

Problems

1.If the propability that a target is destroyed by anyone shot is 0.6 what is the probability that it should be destroyed on the 5th attempt?

Solution:

Let X denote the number of attempts required for the first success (success is distribution)

Then X follows geometric distribution

$$P(X = x) = q^{x-1}p, \quad x = 1, 2, ...$$

$$\text{Given } p = 0.6 \text{ and } q = 1 - p = 0.4$$

$$\text{Here first success is the 5th attempt}$$

$$\therefore x = 5$$

Required $P(X = 5) = (0.4)^4(0.6) = 0.01536$

2. In a particular manufacturing process it is found that, on the average, 1% of the items is defective. What is the probability that the fifth item inspected is the first defective item?

Solution:

Here defective is success. First defective is the 5th one.

So, geometric distribution with $p = \frac{1}{100} = 0.01$, $x = 5$.

$$P(X = x) = q^{x-1}p, \qquad x = 1, 2, 3, ...$$

$$\therefore P(X = 5) = \left(\frac{99}{100}\right)^4 . \frac{1}{100} = 0.0096.$$

3. A die is thrown until 6 appears. What is the probability that it must be thrown more than four times?

Solution:

Getting 6 is success. For the first success it should be thrown more than 4 times. So, geometric distribution is used.

$$P(X = x) = q^{x-1}p, \quad x = 1, 2, ...$$

$$\text{Here } p = \text{probability of } 6 = \frac{1}{6} \qquad \therefore q = 1 - p = \frac{5}{6}$$

Required

$$\begin{aligned}
P(X>4) &= 1-P(X\le 4)\\
&= 1-\{P(X=1)+P(X=2)+P(X=3)+P(X=4)\}\\
&= 1-\left\{p+qp+q^2p+q^3p\right\}\\
&= 1-\frac{p(1-q^4)}{1-q}\ [\text{sum of G.P } = \frac{a(1-r^n)}{1-r}]\\
&= 1-(1-q^4)=r^4\\
&= \left(\frac{5}{6}\right)^4=\frac{625}{1296}=0.48.
\end{aligned}$$

4 Uniform or Rectangular Distribution

Definiton

A continuous RV X is said to follow a *uniform* or *rectangular distribution* in any finite interval, if its probability density function is a constant in that interval.
If X follows a uniform distribution in $a<x<b$, then its pdf is given by $f(x)=\dfrac{1}{b-a}$, $a<x<b$.

Moment generating function, Mean and Variance

$$\begin{aligned}
M_X(t)=E(e^{tX}) &= \int_a^b e^{tx}f(x)dx\\
&= \int_a^b e^{tx}\frac{1}{b-a}dx\\
&= \frac{1}{b-a}\left[\frac{e^{tx}}{t}\right]_a^b=\frac{1}{t(b-a)}(e^{tb}-e^{ta}),t\neq 0\\
\text{Now}, M_X(t) &= \frac{1}{t(b-a)}(e^{tb}-e^{ta})\\
&= \frac{1}{t(b-a)}\left[\left(1+bt+\frac{b^2t^2}{2!}+\frac{b^3t^3}{3!}+\dots\right)-\left(1+at+\frac{a^2t^2}{2!}+\frac{a^3t^3}{3!}+\dots\right)\right]\\
&= \frac{1}{t(b-a)}\left[t(b-a)+\frac{(b^2-a^2)t^2}{2!}+\frac{(b^3-a^3)t^3}{3!}+\dots\right]\\
&= 1+\frac{b+a}{2}.\frac{t}{1!}+\frac{b^2+ab+a^2}{3}.\frac{t^2}{2!}+\dots
\end{aligned}$$

$$
\begin{aligned}
\text{Mean} = \mu_1' &= \text{coefficient of } \frac{t}{\Gamma 1} \text{ in } M_X(t) = \frac{a+b}{2} \\
\mu_2' &= \text{coefficient of } \frac{t^2}{\Gamma 2} \text{ in } M_X(t) = \frac{a^2+ab+b^2}{3} \\
\text{Now,Var(X)} &= \mu_2' - (\mu_1')^2 \\
&= \frac{a^2+ab+b^2}{3} - \frac{(a+b)^2}{4} \\
&= \frac{1}{12}(4a^2+4ab+4b^2-3a^2-3b^2-6ab) \\
&= \frac{1}{12}(a-b)^2
\end{aligned}
$$

Problems

1.If X is uniformly distributed over (0,10) find $(i)P(X<4),(ii)P(X>6),(iii)P(2<X<5)$.
Solution:
X is uniformly distributed over (0,10)

$$\therefore \text{pdf is } f(x) = \begin{cases} \frac{1}{10}, \ 0<x<10 \\ 0, \text{ otherwise} \end{cases}$$

$$P(X<4) = \int_0^4 \frac{1}{10}dx = \frac{4}{10} = \frac{2}{5}$$

$$P(X>6) = \int_6^{10} \frac{1}{10}dx = \frac{4}{10} = \frac{2}{5}$$

$$P(Z<X<5) = \int_2^5 \frac{1}{10}dx = \frac{3}{10}$$

2.Show that for the rectangular distribution $dF = dx,\ o \le x < 1,\ \mu_1' = \frac{1}{2}$ and variance $=\frac{1}{12}$

Solution:

Given $\frac{dF}{dx} = 1,\ 0 \le x < 1,$

where F is cumulative function.

$$\therefore F'(x) = 1 \implies f(x) = \begin{cases} 1, \ 0 \le x < 1 \\ 0, \text{ otherwise} \end{cases}$$

$$\mu_1' = E(X) = \int_0^1 x.f(x)dx = \int_0^1 x\,dx = \frac{1}{2}$$

$$\mu_2' = E(X^2) = \int_0^1 x^2\,f(x)dx = \left[\frac{x^3}{3}\right]_0^1 = \frac{1}{3}$$

$$\therefore \text{Var}(X) = \mu_2' - \left(\mu_1'\right)^2 = \frac{1}{3} - \frac{1}{4} = \frac{1}{12}$$

3. Buses arrive at a specified stop at 15 minutes interval starting at 6 AM i.e. they arrive at 6 AM, 6.15 AM., 6.30 AM. and so on. If a passenger arrives at the stop at a time that is uniformly distributed between 6 and 6.30 AM, find the probability that he waits (i) less than 5 minutes for a bus. (ii) more than 10 minutes for a bus.

Solution:

Let X denote the number of minutes after 6 AM that a passenger arrives at the stop.

X is a uniform R.V over the interval (0,30)

$\therefore$ its pdf is $f(x) = \frac{1}{30},\ 0 < x < 30$

(i) Since the buses arrive at 15 minutes interval strating with 6 AM, a passenger has to wait less than 5 minutes if he comes to the stop between 6.10 and 6.15 or between 6.25 and 6.30 AM.

$\therefore$ Required probability = $P(10 < X < 15) + P(25 < X < 30)$

$$=\int_{10}^{15} \frac{1}{30}dx \int_{25}^{30} \frac{1}{30}dx$$

$$=\frac{5}{30} + \frac{5}{30} = \frac{1}{3}$$

(ii) The passenger has to wait more than 10 minutes if he comes to the stop between 6 and 6.5 AM or between 6.15 and 6.20 AM.

$\therefore$ Required probability = $P(0 < X < 5) + P(15 < X230)$

$$=\int_{0}^{5} \frac{1}{30}dx \int_{15}^{20} \frac{1}{30}dx = \frac{1}{3}$$

4. A random variable X has a uniform distribution over the interval (-3,3). Compute (1) $P(X = 2)$ (2) $P(|X-2| < 2)$ (3) Find k such that $P(X > k) = \frac{1}{3}$.

Solution:

Given X has a uniform distribution over the interval (-3,3).

$\therefore$ the pdf of X is $f(x) = \begin{cases} \frac{1}{6}, & -3 < x < 3 \\ 0, & \text{otherwise} \end{cases}$

Since X is continuous random variable, $P(X = 2) = 0$

$$\begin{aligned} P(|X-2| < 2) &= P(-2 < X-2 < 2) \\ &= P(-2+2 < X < 2+2) = P(0 < X < 4) \\ &= \int_0^4 f(x)\,dx = \int_0^3 \frac{1}{6}dx \\ &= \frac{1}{6}[x]_0^3 = \frac{3}{6} = \frac{1}{2} \end{aligned}$$

(3) $P(X > k) = \int_k^\infty f(x)dx = \int_k^3 \frac{1}{6}dx = \frac{1}{6}[x]_k^3 = \frac{3-k}{6}$

Given$P(X > k) = \frac{1}{3}$

$\implies \dfrac{3-k}{6} = \dfrac{1}{3}$
$\implies 3-k=2$
$\therefore k=1$

5 Exponential Distribution

Definition:

A continuous RV X is said to follow an *exponential distribution* or *negative exponential distribution* with prameter $\lambda > 0$, if its probability density function is given by

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

We note that $\int_0^\infty f(x)dx = \int_0^\infty \lambda e^{\lambda x}dx = 1$ and hence $f(x)$ is a legitimate density function.

5.1 Moment generating function, Mean and variance

$$\begin{aligned}
M_X(t) = & \quad E\left[e^{tX}\right] = \int_{-\infty}^{\infty} e^{tx} f(x)dx \\
= & \quad \int_0^\infty e^{tx}\lambda e^{-\lambda x}dx \\
= & \quad \lambda \int_0^\infty e^{-(\lambda-t)x}dx \\
= & \quad \lambda \left[\frac{e^{-(\lambda-t)x}}{-(\lambda-t)}\right]_0^{-\infty}, \quad \text{if} \lambda - t > 0 \implies \lambda > t \\
= & \quad \frac{\lambda}{-(\lambda-t)}[0-1] \\
M_X(t) = & \quad \frac{\lambda}{\lambda-t}, \lambda > t \\
M_X(t) = & \quad \frac{1}{1-\frac{t}{\lambda}} = \left(1-\frac{t}{\lambda}\right)^{-1} = 1 + \frac{t}{\lambda} + \frac{t^2}{\lambda^2} + \frac{t^3}{\lambda^3} + \ldots
\end{aligned}$$

$$\begin{aligned}
\text{Mean} = \mu_1' = \text{coefficient of } \frac{t}{\Gamma 1} \text{ in } M_X(t) = \frac{1}{\lambda} \\
\mu_2' = \text{coefficient of } \frac{t^2}{\Gamma 2} \text{ in } M_X(t) = \frac{2}{\lambda^2}
\end{aligned}$$

$$\begin{aligned}
\text{Now,Var(X)} = \mu_2' - (\mu_1')^2 \\
= \frac{2}{\lambda^2} - \frac{1}{\lambda^2} = \frac{1}{\lambda^2}
\end{aligned}$$

5.2 Memoryless Property of the Exponential Distribution

If X is exponentially distributed, then

$$\begin{aligned}
P(X > s+t/X > s) &= P(X > t), \text{for any} s,t > 0 \\
P(X > k) &= \int_k^\infty \lambda e^{-\lambda x} dx \\
&= (-e^{\lambda x})_k^\infty \\
&= e^{-\lambda k} \qquad (1)
\end{aligned}$$

$$\begin{aligned}
\text{Now} \, P(X > s+t/X > s) = & \frac{P\{X > s+t \text{ and } X > s\}}{P\{X > s\}} \\
= & \frac{P\{X > s+t\}}{P\{X > s\}} \\
= & \frac{e^{-\lambda(s+t)}}{e^{-\lambda s}}, [\text{by } (1)] \\
= & e^{-\lambda t} = P(X > t).
\end{aligned}$$

Problems

1.Find the moment generating function of the exponential distribution $f(x) = \frac{1}{c} e^{\frac{-x}{c}}, x \geq 0$ and $c > 0$. Hence find its mean and S.D.

Solution:

We know for exponential distribution

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

$$M_X(t) = \frac{1}{1 - \frac{t}{\lambda}}$$

$$\text{Mean} = \frac{1}{\lambda} \quad \text{Var(X)} = \frac{1}{\lambda^2}$$

$$\text{In the given problem } \lambda = \frac{1}{c}$$

$$\therefore M_X(t) = \frac{1}{1 - ct}$$

$$\therefore \text{Mean} = c$$

$$\text{Var(X)} = c^2$$

$$\therefore \text{S.D.} = c$$

2.Suppose the duration X in minutes of long distance calls from your home follows exponential law with pdf

$$f(x) = \begin{cases} \frac{1}{5}e^{-x/5} & \text{for } x > 0 \\ 0 & \text{otherwise} \end{cases}$$

find $P(X > 5), P(3 \leq X \leq 6)$, mean and variance of X.

Solution:

We know that for an exponential distribution given by

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

mean $=\frac{1}{\lambda}$, Var(X)=25

$$\begin{aligned} P(X > 5) &= \int_5^\infty f(x)dx = \frac{1}{5}\int_5^\infty e^{-x/5}dx \\ &= \frac{1}{5}\left[\frac{e^{-x/5}}{\frac{-1}{5}}\right]_5^\infty \\ &= -(0 - e^{-1}) \\ &= \frac{1}{e} = 0.36788 \end{aligned}$$

$$\begin{aligned} P(3 \leq X \leq 6) &= \int_3^6 f(x)dx \\ &= \frac{1}{5}\int_3^6 e^{-x/5}dx \\ &= \frac{1}{5}\left[\frac{e^{-x/5}}{\frac{-1}{5}}\right]_3^6 \\ &= -\left[e^{\frac{-6}{5}} - e^{\frac{-3}{5}}\right] \\ &= e^{\frac{-3}{5}} - e^{\frac{-6}{5}} \\ &= 0.5488 - 0.3012 = 0.2476 \end{aligned}$$

3.Suppose the length of life of an appliance has an exponential distribution with mean 10 years. A used appliance is bought by a person. What is the probability that it will not fail in the next 5 years?

Solution:

Let X represent the length of time of the appliance

Given X follows exponential distribution with mean =10
$\therefore$its pdf $f(x) = \lambda e^{-\lambda x}, x \geq 0$
Mean =$\frac{1}{\lambda} = 10$
$\implies \lambda = \frac{1}{10}$
$\therefore f(x) = \frac{1}{10}e^{\frac{-x}{10}}, x \geq 0$
Since exponential distribution has memoryless property, the length of life t years before the purchase is irrelevant.

$$\begin{aligned}\therefore P(X > t+5|X > t) &= P(X > 5)\\ &= \int_5^\infty f(x)dx\\ &= \frac{1}{10}\int_5^\infty e^{\frac{-x}{10}}dx\\ &= \frac{1}{10}\left[\frac{e^{\frac{-x}{10}}}{\frac{-1}{10}}\right]_5^\infty\\ &= e^{\frac{-1}{2}} = 0.6065\end{aligned}$$

4.The time required to repair a machine is exponentially distributed with parameter $\frac{1}{2}$. What is the conditional probability that the repair exceeds 2 hours? What is the conditional probability that the repair takes atleast 10 hours given that the duration exceeds 9 hours?
Solution:
Let X denote the repairing time.
Given X follows exponential distribution with parameter $\lambda = \frac{1}{2}$
$\therefore$ The exponential distribution is given by

$$\begin{aligned}f(x) &= \lambda e^{-\lambda x}, x \geq 0\\ &= \frac{1}{2}e^{\frac{-1}{2}x}, x \geq 0\\ P(X > 2) &= \int_2^\infty f(x)dx\\ &= \frac{1}{2}\int_2^\infty e^{-x/2}dx\\ &= \frac{1}{2}\left[\frac{e^{-x/2}}{-1/2}\right]\\ &= -\left[e^{-\infty} - e^{-1}\right]\\ &= e^{-1} = \frac{1}{e}\\ \text{Required}P(X \geq 10|X > 9) &= \frac{P(X \geq 10 \cap X > 9)}{P(X > 9)} = P(X > 1)\end{aligned}$$

$$= \int_1^\infty f(x)dx$$

$$= \int_1^\infty \frac{1}{2}e^{-\frac{1}{2}x}dx$$

$$= \frac{1}{2}\left[\frac{e^{-\frac{1}{2}x}}{\frac{-1}{2}}\right]_1^\infty$$

$$= -\left(e^{-\infty} - e^{-1/2}\right)$$

$$= e^{-0.5} = 0.6$$

6 Normal distribution

Definition

A continuous RV X is said to follow *normal distribution* or *Gaussian distribution* with parameter μ and σ, if its pdf is given by

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

Note:

1. Moment generating function of Normal distribution is $M_X(t) = e^{\mu t + \frac{t^2\sigma^2}{2}}$.
2. Mean = μ.
3. Variance, Var(X) = σ^2.

Properties of Normal distribution

1. The mean, median and mode coincide at $x = \mu$.
2. It is symmetric about $x = \mu$.
3. The maximum value of $f(x)$ is at $x = \mu$ and the maximum value $= \frac{1}{\sigma\sqrt{2\pi}}$.
4. The points of inflection are at $x = \mu - \sigma$ and $x = \mu + \sigma$.
5. The curve approaches the horizontal axis asymptotically on either side of $x = \mu$.
6. The total area under the curve is 1.
7. Coefficient of skewness is 0 and coefficient of kurtosis is 3.
8. All odd moments about mean vanish.

Standard Normal Distribution

$$
\begin{aligned}
Z &= \frac{X-\mu}{\sigma} \\
E(Z) &= \frac{1}{\sigma}(E(X)-\mu) = \frac{1}{\sigma}(\mu-\mu) = 0 \\
\text{Var(Z)} &= \frac{1}{\sigma^2}\text{Var(X)} = \frac{1}{\sigma^2}\sigma^2 = 1 \\
\therefore \text{Its pdf is given by } \phi(z) &= \frac{1}{\sqrt{2\pi}}e^{\frac{-z^2}{2}},\ -\infty < z < \infty.
\end{aligned}
$$

Its MGF is given by $M_Z(t) = e^{\frac{1}{2}t^2}$.

Problems

1. X is a normal distribution with mean 16 and S.D. 3. Find $(i) P(X \geq 19), (ii) P(12.5 < X < 19), (iii) P(10 < X < 25), (iv) K$ if $P(X > K) = 0.24$

Solution:

Given $\mu = 16$, $\sigma = 3$

Put $Z = \dfrac{X-\mu}{\sigma} = \dfrac{X-16}{3}$

(i) when $X = 19, Z = \dfrac{19-16}{3} = 1$

$\therefore\ P(X \geq 19) = P(Z \geq 1) = 0.5 - P(0 < Z < 1) = 0.5 - 0.3413 = 0.1587$

(ii) $P(12.5 < X < 19)$

when $X = 12.5, Z = \dfrac{12.5-16}{3} = -1.16$

when X=19, Z= $\dfrac{19-16}{3} = 1$

$$
\begin{aligned}
\therefore P(12.5 < X < 19) &= P(-1.16 < Z < 1) \\
&= P(-1.16 < Z < 0) + P(0 < Z < 1) \\
&= P(0 < Z < 1.16) + P(0 < Z < 1), \text{(by symmetry)} \\
&= 0.3770 + 0.3413 = 0.7183
\end{aligned}
$$

(iii) $P(10 < X < 25)$

when X=10, Z= $\dfrac{10-16}{3} = -2$

when X=25, Z= $\dfrac{25-16}{3} = 3$

$$
\begin{aligned}
\therefore P(10 < X < 25) &= P(-2 < Z < 3) \\
&= P(-2 < Z < 0) + P(0 < Z < 3) \\
&= P(0 < Z < 2) + P(0 < Z < 3), \text{(by symmetry)} \\
&= 0.4772 + 0.4987 = 0.9759
\end{aligned}
$$

$(iv) P(X > K) = 0.24$

when X=K, Z= $\dfrac{K-\mu}{\sigma} = \dfrac{K-16}{3} = z_1\text{(say)}$

$$
\begin{aligned}
\therefore P(X > K) &= 0.24 \\
\implies P(Z > z_1) &= 0.24 \\
\implies P(0 < Z < z_1) &= 0.5 - 0.24 = 0.26
\end{aligned}
$$

$\therefore z_1$ is the value of Z corresponding to the area 0.26.

and $z_1 = 0.7$(from table)

$\therefore \dfrac{K-16}{3} = 0.7$

$\implies K = 16 + 3 \times 0.7 = 16 + 2.1 = 18.1$

2. The average percentage of marks of candidates in an university examination is 42 with S.D of 10. The maximum for a pass is 50%. If 1000 candidates appear for the examination, how many can be expected to pass it assuming normality of the distribution of marks? If it is required that double that number should pass, what should be the minimum percentage of marks?

Solution:

Given $\mu = 42, \sigma = 10$

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

$$Z = \frac{X-\mu}{\sigma} = \frac{X-42}{10}$$

A student is passed when scored greater than 50 i.e. $X \geq 50$

$\therefore$ when $X = 50$, then $Z = \dfrac{50-42}{10} = 0.8$

to $\therefore P(X \geq 50) = P(Z \geq 0.8)$

=0.5-P(0¡Z¡0.8)

=0.5-0.2881

=0.2119

Number of students expected to pass =1000 × 0.2119=211.9 ≈212

If the no. of passed student is 424, then proportion out of 1000=$\dfrac{424}{1000} = 0.424$

Let $X = x$, is minimum pass marks and corresponding $Z = z_1$, then

$$\begin{aligned}
P(Z > z_1) &= 0.424\\
0.5 - P(0 < Z < z_1) &= 0.424\\
\therefore P(0 < Z < z_1) &= 0.5 - 0.424 = 0.076\\
\text{Now from the table}, z_1 &= 0.19\\
\frac{X - 42}{10} &= 0.19\\
X = 0.19 \times 10 + 42 = 43.9 \approx 44
\end{aligned}$$

$\therefore$ The minimum pass marks to double pass student is 44%.

3. In a distribution exactly normal 7% of the items are under 35 and 89% are under 63. What are the mean and S.D. of the distribution?(Draw diagram if required)

Solution:

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

$$P(X < 35) = 7\% = 0.07$$

$$P(X < 63) = 89\% = 0.89$$

Let $Z = \dfrac{X - \mu}{\sigma}$

When $X = 35$, $X = \dfrac{35 - \mu}{\sigma} = -Z_1$

$$35 - \mu = -\sigma Z_1 \qquad (1)$$

Now,

$$\begin{aligned}
P(X < 35) &= P(Z < -Z_1) = 0.07\\
P(Z > Z_1) &= 0.07\\
0.5 - P(0 < Z < Z_1) &= 0.07\\
\therefore P(0 < Z < Z_1) &= 0.43\\
\therefore Z_1 &= 1.48 \quad \text{From table}
\end{aligned}$$

When $X = 63$, $Z = \dfrac{63 - \mu}{\sigma} = Z_2$, since area below $X = 63$ is 0.69.

The value of Z is positive

$$63 - \mu = \sigma Z_2 \quad (2)$$

$$\begin{aligned}
P(X < 63) &= 0.89\\
P(Z < Z_2) &= 0.89 = 0.5 + 0.39\\
\therefore P(0 < Z < Z_2) &= 0.39
\end{aligned}$$

Z_2 is the value of Z corresponding to the area =0.39

$$\text{From table} \quad Z_2 = 1.23$$

$$\text{From (1)}, 35 - \mu = 1.48\sigma \quad (3)$$

$$\text{From (2)}, 63 - \mu = 1.23\sigma \quad (4)$$

Solving (3) and (4), we have

$$\mu = 50.3, \quad \sigma = 10.33$$

4. In a normal distribution, 31% of the items are under 45 and 8% are over 64. Find the mean and variance of the distribution.(Draw diagram if required)

Solution:

We know that, $Z = \dfrac{X-\mu}{\sigma}$

Let $Z = Z_1$, where $X = 45$ and $Z = Z_2$, when $X = 64$

$$\begin{aligned}
P(0 < Z < Z_1) &= 0.31 \\
P(Z_1 < Z < 0) &= 0.19 \\
\therefore \text{ From tables}, Z_1 &= -0.49 \\
\text{i.e.,} \frac{45-\mu}{\sigma} &= -0.49 \quad \text{or} \\
45-\mu &= -0.49\sigma \qquad (1) \\
P(Z > Z_1) &= 0.8 \text{ or} \\
P(0 < Z < Z_2) &= 0.42 \\
\therefore \text{ From tables, } Z_2 &= 1.40 \\
\therefore \frac{64-\mu}{\sigma} &= 1.40 \\
64-\mu &= 1.40\sigma \qquad (2) \\
(1)-(2) \implies -19 &= -1.89\sigma \\
\sigma &= \frac{-19}{-1.89} = 10
\end{aligned}$$

From (1) $\mu = 45 + (0.49 \times 10) = 45 + 4.9 = 49.9 = 50$

$$\text{Mean} = 50 \text{ and S.D} = 10$$

$$\text{i.e.,} \mu = 50,\ \sigma = 10,\ \text{Variance}(\sigma^2) = 100$$