COUNCIL OF SCIENTIFIC & INDUSTRIAL RESEARCH

Mathematical Science Code:4

12 Probability & Statistics

12.1 Random Experiment, Event Space

12.2 Statistical Regularity and definition of Probability

12.3 Random Variable and Distribution Function

12.4 Discrete random variable and distribution

12.5 Continuous random variable and distribution

12.6 Some important definitions

12.7 Some discrete and continuous distributions

12 Probability & Statistics

12.1 Random Experiment, Event Space

12.1.1 <u>Definition (Experiment/ Trial)</u>: A process which results in some well-defined out comes is known as experiment/ trial.

12.1.2 <u>Definition (Random Experiment)</u>: An experiment E is said to be a random experiment if

(i) All possible outcomes of E are known in advance
(ii) It is impossible to predict which outcome will occur at a particular performance of E
(iii) E can be repeated under identical conditions of infinite numbers of times.

<u>Examples (12.1):</u>

(i) The experiment of tossing a coin.
(ii) The experiment of throwing a die.
(iii) The experiment of drawing a card from a full pack of 52 cards.

12.1.3 <u>Definition (Event space)</u>: The set of all outcomes of a given random experiment E is called the event space of the experiment and it will be denoted by S.

<u>Examples (12.2):</u>

i. The event space S of the random experiments of tossing a coin is {H, T}, where H denote the outcome 'head' and T denote the outcome 'tail'.
ii. The he event space of throwing a die is S= {1, 2, 3, 4, 5, 6}.

12.1.4 <u>Definition (Events)</u>: A subset of the event space of a random experiment is said to be an event.

<u>Examples (12.3):</u>

i. Let E be the random experiment of throwing a die. Here S= {1, 2, 3, 4, 5, 6}. Let A= {2, 4, 6} be event which can be described as 'even number appears in throwing a die'.

12.1.5 <u>Impossible event:</u> An event of a given random experiment is called an impossible event if it can never happen in any performance of the random experiment under identical conditions. Impossible event is described by the empty set Φ.

<u>Example (12.4):</u> For the random experiment E of throwing a die, the event 'face marked 8' is an impossible event.

12.1.6 Certain event: An event of a given random experiment is called certain event if it happens in every performance of the corresponding random experiment under identical conditions. Certain event is described by the empty set S.

Example (12.5): For the random experiment E of tossing a coin, the event space S={H, T} is certain event.

12.1.7 Simple and Composite Event: An event A is called a simple event or an elementary event if A contains exactly one element.

An event A is called a composite event or compound event if A contains more than one element.

Example (12.6): In connection with the random experiment of throwing a die, the event space S is given by S= {1, 2, 3, 4, 5, 6}. Let A, B be the events defined by A= {2, 4, 6}, B= {3}. The event A is a composite event and B is a simple event.

12.1.8 Mutually Exclusive event: Two events A, B connected to a given random experiment E are said to be mutually exclusive event if A, B can never happen simultaneously in any performance of E, i.e. if $A \cap B = \Phi$.

Example (12.7): In connection with throwing a die, the events 'even number' and 'odd number' are mutually exclusive.

12.2 Statistical Regularity and definition of Probability

12.2.1 Statistical Regularity: Let a random experiment E be repeated N times under identical conditions, in which we note that an event A of E occurs N(A) times. Then the ratio $\frac{N(A)}{N} = f(A)$ is called the frequency ratio of A and is denoted by f(A). Now if E is repeated a very large number of times, it is seen that f(A) is gradually stabilizes to a constant. This tendency of stability of frequency ratio is called statistical regularity.

12.2.2 Classical definition of Probability: Let S be the event space of a given random experiment E be finite. If all the simple events connected to E be equally likely then the probability of an event A is defined as P(A)=$\frac{m}{n}$, where n is the total number of simple events connected to E, i.e. *n* is the number of distinct elements of S and *m* is the distinct elements of A.

12.2.3 Some Important Results:

i. $0 \leq P(A) \leq 1$
ii. $P(S) = 1$
iii. $P(\Phi) = 0,$
iv. $P(\bar{A}) = 1 - P(A)$
v. $P(A \cup B) = P(A) + P(B) - P(A \cap B)$
vi. $P(\cup A_j) = \cup\ P(A_j),\ A_i \cap A_j = \phi\ \forall\, i, j\ and\ i \neq j.$

12.2.4 Frequency Definition of Probability: Let a random experiment E be repeated N times under identical conditions, in which we note that an event A of E occurs N(A) times. Then on the basis of statistical regularity we can assume that $\lim_{n\to\infty} \frac{N(A)}{N}$ exists finitely and the value of this limit is called the probability of the event A, denoted by $P(A)$, i.e. $P(A) = lim_{n\to\infty} \frac{N(A)}{N}$.

Let A be an event.

$$f_{rel}(A) = \frac{f(A)}{n} = \frac{Number\ \ of\ times\ A\ occurs}{Number\ \ of\ trials}$$

$f(A)$ is called the frequency of A clearly $0 \leq f_{rel}(A) \leq 1$.

(i) $f_{rel}(A) = 0$ if A does not occur in a sequence of trials.
(ii) $f_{rel}(A) = 1$ if A occurs in each trial of sequence.
(iii) $f_{rel}(A \cup B) = \frac{f(A)+f(B)}{n} = f_{rel}(A) + f_{rel}(B) = 0$, if $A \cap B = \phi$

12.2.5 Probability as the counter part of relative frequency

Definition (Probability): - Given a sample space S, with each event A of S there is associated a number P(A), called the probability of A, such that the following anions of probability are satisfied.

(i) For all $A \subset S,\ 0 \leq P(A) \leq 1$
(ii) $P(S) = 1$
(iii) $P(\cup A_i) = \cup_i P(A_i)\ ,\ A_i \cap A_j = \phi,\ j \neq i$

Some results:

(i) $P(A) = 1 - P(A)$
(ii) $P(A \cup B) = P(A) + P(B) - P(A \cap B)$

12.2.6 **Definition (Conditional Probability):** To find the probability of an event B under the condition that an event A occurs. This probability is called the conditional probability of B *given* A and is denoted by $P(A|B)$. In this case A serves as a new (reduced) sample space, and that probability is the fraction of $P(A)$ which corresponds to $A \cap B$. Thus, $P(B|A) = \frac{P(A\cap B)}{P(A)}$, *provided* $P(A) \neq 0$.

(i) Multiplication rule: $P(A \cap B) = P(A)\, P(B|A) = P(B)\, P(A|B),\ P(A) \neq 0,\ P(B) \neq 0$.

(ii) Independent Events: $P(A \cap B) = P(A)\, P(B)$.

12.2.7 **Permutation (Arrangement in any order):**

(i) n different things – then permutation is given by $n! = 1 \cdot 2 \cdot 3 \cdots n$

(ii) $n = n_1 + n_2 + \ldots\ldots + n_r,\ n_i$ are the number of same things in $i-th$ class. Then the number of permutation of these things taken all at a time is $\frac{n!}{n_1!\, n_2! \ldots\ldots n_r!}$

(iii) The number of permutations of n different things taken k at a time $nP_r = \frac{n!}{(n-k)!} = n(n-1)(n-2)\ldots\ldots(n-k+1)$ repetitions and with repetitions is n^k

12.2.8 **Combinations (Selection of one or more things without an order):**

The number of different combinations of n different things, k at a time, without repetitions, is $\binom{n}{k} = \frac{n!}{k!(n-k)!}$

Stirling formula:- $n! \sim \sqrt{2\pi n}\, \left(\frac{n}{e}\right)^n$ (for large n)

12.3 Random Variable and Distribution Function

12.3.1 **Definition of Random Variable**: Let E be the random experiment and S be the event space of E. Let Þ(S) be the collection of all subsets of S forming the class of all events connected to E. A mapping $X: S \to R$ is called a *random variable* or a *stochastic variable* or a *variate*, if for any $x \in R$, the set

$$\{\omega \in S: -\infty < X(\omega) \leq x\} \in Þ(S)$$

i.e., if $\{\omega \in S: -\infty < X(\omega) \leq x\}$ is an *event* connected to E.

In short form we write $\{\omega \in S: -\infty < X(\omega) \leq x\} = \{-\infty < X \leq x\}$

The range of the mapping is called the *spectrum* of the random variable X.

Example (12.8): A coin is tossed twice. Here,

$$S = \{\omega_1 = (H, H), \omega_2 = (H, T), \omega_3 = (T, H), \omega_4 = (T, T)\}$$

A mapping X:S→R is defined by $X(\omega_i) = k$, where k is number of heads, $i = 1, 2, 3, 4$. Then $X(\omega_1) = 2, X(\omega_2) = 1, X(\omega_3) = 1, X(\omega_4) = 0$. Here X is the random variable defined in the domain S and the spectrum of $X = \{0, 1, 2\}$. Here, according to the notation $X = 0$ represents the event $\{(T, T)\}$, $\{\omega \, \varepsilon \, S : 0 \leq X(\omega) \leq 2\}$ is a certain event and $\{\omega \, \varepsilon \, S : 1 < X(\omega) < 2\}$ is an impossible event Φ.

12.3.2 Definition of Distribution Function: Let $P: Þ(S) \to R$ be a probability function, where Þ(S) is the collection of subsets of S forming the class of events. The ordered 3 tuple $(S, Þ(S), P)$ is called a probability space.

Let X be a random variable defined on the event space S connected to a random experiment E. The *distribution function* of the random variable X with respect to the probability space $(S, Þ(S), P)$ is a real valued function F(x) of real variable x, defined in (-∞, ∞), where

$$F(x) = P(-\infty < X \leq x), \text{for all } x \, \varepsilon \, (-\infty, \infty).$$

It is evident that the range of the distribution function is a subset of [0, 1].

12.3.3 Properties of Distribution Function:

(i) $0 \leq F(x) \leq 1$

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

(iii) $F(x)$ is monotonically increasing function.

(iv) $F(x)$ is continuous to the right at every point a, i.e., $\lim_{X \to a+0} F(x) = F(a)$ or, $F(a + 0) = F(a)$

(v) $P(X = a) = F(a) - F(a - 0)$

(vi) $F(\infty) = 1$

(vii) $F(-\infty) = 0$

The set of points of discontinuity of a distribution function is at most enumerable.

Random Variable
- (a) Discrete random variable.
- (b) Continuous random variable.

12.4 Discrete random variable and distribution:

12.4.1 Probability Mass Function(pmf): A random variable X defined on an event space S is said to be discrete if the spectrum of X is at most countable (finite or enumerable). In this case the probability distribution of X is called a discrete distribution.

Let the spectrum of X be $\{x_i : i = 0, \pm 1,, \pm 2, \dots \}$, where

$$\cdots\cdots < x_{-2} < x_{-1} < x_0 < x_1 < x_2 < \cdots\cdots$$

Let $P(X = x_i) = f_i$, x_i being the spectrum point. A function $f : R \to [0,1]$ is defined as follows:

$f(x) = f_i$, if $x = x_i$, which is a point of the spectrum,

= 0, elsewhere.

The function f *defined* above is called the probability mass function (pmf) of the random variable X.

The distribution function F(x) of a discrete random variable is defined as:

$$F(x) = \sum_{j=-\infty}^{i} P(X = x_j) = \sum_{j=-\infty}^{i} f_j \text{ , } if \ x_i \le x < x_i + 1 \ (i = 0, \pm 1,, \pm 2, \dots)$$

12.4.2 Some Important Results on Discrete Distributions:

1. $\sum_{j=-\infty}^{\infty} f_j = 1$
2. At each non-spectrum point a, $P(X = a) = 0$.
3. $P(a < X \le b) = \sum_{a < x_j \le b} f_j$

Note: $F(x)$ is a step function.

Example (12.9):

$X =$ so that is fair die turns up $f(x = x_i) = \frac{1}{6}, x_i = 1, 2, \dots\dots \ ..,6$

$F(x) = \sum_{x_i \le x} f(x_i)$. $F(x = 3) = \frac{1}{6} + \frac{1}{6} + \frac{1}{6} = \frac{1}{2}$ (step function) i.e.,

X	1	2	3	4	5	6
$f(x)$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$
$F(x)$	$\frac{1}{6}$	$\frac{2}{6}$	$\frac{3}{6}$	$\frac{4}{6}$	$\frac{5}{6}$	$\frac{6}{6}$

Result: $P(a < X \leq b) = F(b) - F(a) = \sum_{a<x_i\leq b} P_i$ (For discrete case).

12.5 Continuous random variable and distribution:

A random variable X and its distribution are called continuous, if the corresponding distribution function $F(x)$ can be given by an integral in the form $F(x) = \int_{-\infty}^{x} f(u)\, du$, with the density f of the distribution being continuous (possible except for finitely many values of u) and non-negative

(i) $F'(x) = f(x)$ for every x for which $f(x)$ is continuous.

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

(iii) $P(a < X \leq b) = P(a \leq X \leq b) = P(a \leq X < b) = P(a < X < b) = \int_{a}^{b} f(x)\, dx$

12.6 Some important definitions

12.6.1 <u>Mean(μ):</u> Let X be a random variable defined on a given event space S. Then mean of the distribution of X, denoted by μ is defined by μ = $\sum_i x_i\ f_i$ (discrete)

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

Note: $\mu = E(X)$, expectation of X

12.6.2 <u>Variance(σ^2):</u> Let X be a random variable defined on a given event space S. Then variance of the distribution of X, denoted by *Var(X)*= σ^2, is defined by

$\sigma^2 = Var(X) = \mathrm{E}(x - \mu)^2 = \sum_i (x_i - \mu)^2 f_i$ (discrete)

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

12.6.3 Standard deviation(σ): Let X be a random variable defined on a given event space S. Then standard deviation of the distribution of X, denoted by $\sigma(X)$ and is defined by

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

12.6.4 Transformation of mean and variable:

$X^* = c_1X + c_2,\ (c_1 \neq 0,\ c_1,\ c_2 \in \mathbb{R})$

(a) $m^* = c_1m + c_2$

(b) $\sigma^{*2} = {c_1}^2\sigma^2$

(c) $Z = \frac{X-m}{\sigma},\ m_2 = 0, {\sigma_Z}^2 = 1$

i.e., $Z \sim N(0,1)$

$$E\big(H(X)\big) = \sum H(x)P_X(x) = \int_{-\infty}^{\infty} H(x)f_X(x)\,dx$$

Note: Let X is continuous of $a < x < b$ and if $y = H(x)$ is continuous and strictly monotone on (a, b) then $Y = H(X)$ has the density function $f_y(y) = f(x)\left|\frac{dx}{dy}\right|$ $H(a) < y < H(b)$ $or, H(b) < y < H(a)$ accending as H is monotonic increasing or monotonic decreasing.

12.6.5 Median: Let F(X) be the distribution function of a random variable X, defined on a given event space S. If there exists a real number m such that $F(m\text{-}0) \leq \frac{1}{2}$ and $F(m) \geq \frac{1}{2}$, then m is called median of the distribution of X.

12.6.6 Mode: Let X be a random variable defined on a given event space S. Mode of the distribution of X is a point where the p.d.f $f(x)$ or p.m.f f_i is relatively maximum.

12.6.7 Moments: Let X be a random variable defined on a given event space S and α be a real number. The value $E\{(X-\alpha)^2\}$ if it exists, is called the k-th order moments of X about α where k is a positive integer.

12.6.8 Central moments: Let X be a random variable defined on a given event space S. The value $E\{(X-\mu)^2\}$ if it exists, is called the k-th order central moments of X where μ is the mean of X.

$\mu_1 = E(X-\mu) = E(X) - \mu = 0$

$\mu_2 = E\{(X-\mu)^2\} = Var(X) = \sigma^2$

12.6.9 <u>Skewness:</u> Let X be a random variable defined on a given event space S. The third order central moments μ_3 is known as measure of skewness.

$\beta_1 = \frac{\mu_3}{\sigma^3}$ is called coefficient of skewness.

$\gamma_1 = \beta_1{}^2$ is the measure of skewness.

12.6.10 <u>Kurtosis:</u> Let X be a random variable defined on a given event space S. The fourth order central moments μ_4 is known as measure of kurtosis.

$\gamma_2 = \frac{\mu_4}{\sigma^4}$ is called coefficient of kurtosis.

$\beta_2 = \gamma_2 - 3$ is the coefficient of excess.

<u>Note:</u> A distribution is said to be mesokurtic, leptokurtic or platykurtic according as $\beta_2 = 0,\ \beta_2 > 0\ or\ \ \beta_2 < 0.$

<u>Example (12.10)</u>: If X be a Poisson-4 distribution, then find the mean, standard deviation, skewness, kurtosis.

<u>Solution:</u> Since X be a Poisson-4 distribution, we have

mean(μ) = 4 = standard deviation

skewness(γ_1) = $\frac{1}{4}$, kurtosis(γ_2) = $3 + \frac{1}{4}$

12.7 Some discrete and continuous distributions:

12.7.1 <u>Binomial Distribution:</u>

Let A be an event, and $P(A)$ for single trial and $P(A) = 1 - P = q$. Let us consider the experiment in n times and $X = number\ of\ times\ A\ occur$. Then $X = 0, 1, 2, \dots \dots n$ and

$$Probability\ Mass\ Function(p.m.f.) = f(x) = \begin{cases} \binom{n}{x}\ p^x\ q^{n-x}, & x = 0, 1, 2, \dots . n \\ 0, & otherwise \end{cases}$$

$m = np,\ \sigma^2 = npq,\ M_X(t) = (p\ e^t + q)^n = E(e^{tX})$

$M_X(f) = E(e^{tX}) = \sigma = \left(1 + tX + \frac{tX^2}{2!} + \dots .\right) = 1 + E(X)t + E(X^2)\frac{t^2}{2!} + \dots . \dots$ (Moment – generating function)

Some properties:

(i) Binomial distribution has two parameters *p or q*, the probability of 'success' or 'failure' and *n*, the number of trials. The parameter *n* is always integer.
(ii) The mean (μ), standard deviation (σ), variance (σ^2) and the coefficient of dispersion (C.D.) of a binomial distribution depend on *n* and proportions *p* and *q* of the populations.

$$\text{Mean } (\mu) = np$$
$$\text{Standard deviation } (\sigma) = \sqrt{(npq)}$$
$$\text{Variance } (\sigma^2) = npq$$
$$\text{Coefficient of dispersion (C.D.)} = \frac{npq}{np} = q$$

(iii) Skewness and kurtosis of binomial distribution depends on the proportion of *p* and *q* in the population.

$$\text{Skewness } (\gamma_1) = \frac{(p-q)^2}{npq}$$
$$\text{Kurtosis } (\gamma_2) = 3 + \frac{1-6pq}{npq}$$

(iv) Binomial distribution is symmetrical of $p = q = 0.5$
(v) It is positive skewed if $p < 0.5$ and negatively skewed if $p > 0.5$.
(vi) The binomial coefficients are given by the Pascal's triangle.

Example (12.11): Calculate the probability of getting head three times when a coin is tossed 5 times.

Solution: There are two outcomes either head (H) or tail (T). So it is binomial distribution.

If the probability of getting head is *p* and of tail is *q*. The probability of getting head 3 times can be calculated in the following way.

$$n=5,\ r=3,\ p=\frac{1}{2}=q$$

$$P(X=3)=\binom{5}{3}\left(\frac{1}{2}\right)^3\left(\frac{1}{2}\right)^{5-3}=10\times\frac{1}{2^5}=0.3125$$

Example (12.12): Calculate the probability of having two males and two female children in a family of four children.

Solution: In a particular time there are two outcomes either children may be male (M) or female (F). So it is binomial distribution.

If the probability of having a male child is *p* and of female children is *q*. The probability of having two male children can be calculated in the following way.

$$n=4,\ r=2,\ p=\frac{1}{2}=q$$

$$P(X=2)=\binom{4}{2}\left(\frac{1}{2}\right)^2\left(\frac{1}{2}\right)^{4-2}=6\times\frac{1}{2^4}=0.375$$

Example (12.13): There are 64 beds in a garden and 3 seeds of a particular type of flower are shown in each bed. The probability of a flower being white is $\frac{1}{4}$. Find the number of beds with 3, 2, 1, 0 white flower.

Solution: There are two outcomes either a flower may be white (W) or other colour (O). So it is binomial distribution.

The probability p of a flower being white is $\frac{1}{4}$. So, $q = 1 - p = 1 - \frac{1}{4} = \frac{3}{4}$

Here, $n = 3, N = 64, r = 0, 1, 2, 3.$

(i) Number of beds with no white flower are

$$64 \times P(X = 0) = 64 \times \binom{3}{0}(\frac{1}{4})^0(\frac{3}{4})^{3-0} = 64 \times \frac{27}{64} = 27$$

(ii) Number of beds with one white flowers are

$$64 \times P(X = 1) = 64 \times \binom{3}{1}(\frac{1}{4})^1(\frac{3}{4})^{3-1} = 64 \times 3 \times \frac{1}{4} \times \frac{9}{16} = 27$$

(iii) Number of beds with two white flowers are

$$64 \times P(X = 2) = 64 \times \binom{3}{2}(\frac{1}{4})^2(\frac{3}{4})^{3-2} = 64 \times 3 \times \frac{3}{64} = 9$$

(iv) Number of beds with three white flowers is

$$64 \times P(X = 3) = 64 \times \binom{3}{3}(\frac{1}{4})^3(\frac{3}{4})^{3-3} = 64 \times \frac{1}{64} = 1$$

Example (12.14): The probability that a screw manufactured by a machine to be defective is $\frac{1}{50}$.

A lot of 6 screws are taken at random. Find the probability that (i) there are exactly 2 defective screws in the lot. (ii) no defective screw and (iii) at most 2 defective screws.

Solution: We consider the event 'getting a screw defective' as 'success'. Then p= probability of success=$\frac{1}{50}$, n=6. So if X be the random variable corresponding to the number of defective screws, then X is a binomial (6, p) variate.

(i) The required probability=$P(X = 2) = \binom{6}{2}(\frac{1}{50})^2(1 - \frac{1}{50})^{6-2}$ =15×$\frac{(49)^4}{(50)^6}$

(ii) The required probability=$P(X = 0)$=$\binom{6}{0}(\frac{1}{50})^0(1 - \frac{1}{50})^{6-0}$ =$\frac{(49)^6}{(50)^6}$

(iii) The required probability= $P(X = 0) + P(X = 1) + P(X = 2)$=$\frac{(49)^6}{(50)^6}$ +

$$\binom{6}{1}(\frac{1}{50})^1(1 - \frac{1}{50})^{6-1} + 15\times\frac{(49)^4}{(50)^6} = \frac{(49)^6}{(50)^6} + 6 \times \frac{(49)^5}{(50)^6} + 15\times\frac{(49)^4}{(50)^6}$$

12.7.2 <u>Poisson Distribution</u>: It is the limiting case of binomial distribution with $P \to 0$ and $n \to \infty$ and $\mu = np \to finite\ value\ (\mu)$

$$p.m.f. = f(x) = \begin{cases} \frac{\mu^x}{x!}\ e^{-\mu}\ , & x = 0, 1, 2, \dots. \\ 0, & otherwise \end{cases}$$

$$m = \mu,\ \sigma^2 = \mu,\ M_X(t) = e^{\mu(e^t - 1)}$$

Some properties:

(i) Poisson distribution has only one parameter □ which is a non-negative real number.

(ii) The mean (μ), standard deviation (σ) and variance (σ^2) of Poisson distribution depend on λ.

Mean (μ) = λ

Standard deviation (σ) = $\sqrt{\lambda}$

Variance (σ^2) = λ

(iii) Skewness and kurtosis of Poisson distribution depends on □ and are given by

Skewness (γ_1) = $\frac{1}{\lambda}$

Kurtosis (γ_2) = $3 + \frac{1}{\lambda}$

<u>Example (12.15)</u>: A box contains 1 white and 99 black balls. If 1000 drawing are made with replacements, then what is the probability that 2 drawings will yield white balls?

<u>Solution:</u> Here $p = \frac{1}{100} = 0.01, n = 1000,\ \lambda = n \times p = 1000 \times 0.01 = 10.$

So, p is small and n is large. Therefore, we apply Poisson distribution.

Required probability is $P(2) = P(X=2) = \frac{e^{-10}10^2}{2!} = 0.0023$

<u>Example (12.16)</u>: In a hospital, the mortalities rate for malignant malaria is 7 in 1000. What is the probability for just 3 on account of the disease in a group of 400 people?

<u>Solution:</u> Here $p = \frac{7}{1000} = 0.007, n = 400,\ \lambda = n \times p = 400 \times 0.007 = 2.8.$

So, p is small and n is large. Therefore, we apply Poisson distribution.

Required probability is $P(3) = P(X = 3) = \frac{e^{-2.8}\ 2.8^2}{3!} = 0.22$

Example (12.17): The probability of a product produced by a machine to be defective is 0.01.

If 30 products are taken at random, find the probability that exactly 2 will be defective.

Solution: We consider the event 'getting a defective product' as 'success'. Then *p*= probability of success=0.01, *n*=30, large. So if X be the random variable corresponding to the number of defective product ,then X is a Poisson $\lambda = n \times p = 30 \times 0.01 = 0.3$ distribution. Hence the probability of getting exactly 2 defective products is

$$P(2) = P(X = 2) = \frac{e^{-0.3} 0.3^2}{2!} = 0.03337$$

12.7.3 Hyper-geometric Distribution:- If a box contains N things and M are defective, then the probability of drawing a defective object in a trial is $P = \frac{M}{N}$ · Hence in drawing a sample of n objects without replacement, the probability of obtaining precisely x defective objective(s) is

$$p.m.f. = f(x) = \frac{\binom{M}{x}\binom{N-M}{n-x}}{\binom{N}{n}}, x = 0, 1, 2, \ldots \ldots . . n$$

[which is known as hyper-geometric distribution]

with replacement, the probability of obtaining precisely x defective object(s) is

$$f(x) = \binom{n}{x}\left(\frac{M}{N}\right)^x \left(1 - \frac{M}{N}\right)^{n-x}, x = 0, 1, \ldots \ldots, n$$

$$m = n\frac{M}{N}, \quad \sigma^2 = \frac{n\, M\, (N-M)(N-n)}{N^2(N-1)}$$

12.7.4 Geometric Distribution:

$$p.m.f. = f(x) = \begin{cases} q^{x-1}, & x = 1,2,3, \ldots \ldots \\ 0, & otherwise \end{cases}$$

$$Mean\ (m) = \frac{1}{p},\ Variance\ (\sigma^2) = \frac{q}{p^2}$$

Moment – generating function is $M_X(t) = \frac{Pe^t}{1-qe^t}$

12.7.5 Pascal Distribution: This is an extension of the geometric distribution. In this case random variable X denotes the trial don which the r – th success occurs, $r \in \mathbb{z}$. The probability mass function of X is given by

$$f(x) = \begin{cases} \binom{x-1}{r-1} p^r q^{x-r}, & x = r, r+1, r+2, \ldots.. \\ 0, & otherwise \end{cases}$$

$$m = \frac{n}{p}, \quad \sigma^2 = \frac{rq}{p^2}, M_X(t) = \left(\frac{Pe^t}{1-qe^t}\right)^r$$

Median: Where the probability is $\frac{1}{2}$ $(m \in \mathbb{R})$ is called median of a distribution if $P(X \le m) \ge \frac{1}{2}$ and $P(X \ge m) \ge \frac{1}{2}$

i.e., (For continuous case) $\int_{-\infty}^{m} f(x)\,dx \ \ge \frac{1}{2}$ and $\int_{m}^{\infty} f(x)\,dx \ge \frac{1}{2}$

For absolutely continuous distribution $P(X \le m) = \ P(X \ge m) = \int_{-\infty}^{m} f(x)\,dx = \frac{1}{2}$

Example (12.18):

(i) $N(m, \sigma^2) \rightarrow median = m$

(ii) Uniform distribution $(in\ a\ to\ b), median\ = \frac{a+b}{2}$

Mode: The mode of a discrete probability distribution (or continuous probability distribution) $x \in \mathbb{R}$ for which $p.m.f.(resp. pdf) has\ its\ maximum\ value.$

Example: $N(m, \sigma^2) \rightarrow mode = 0$

Skewness: (Measure of the symmetry of the probability distribution)

Define as $\gamma_1 = E\left[\left(\frac{X-m}{\sigma}\right)^3\right] = \frac{E(X^3) - 3m\sigma^2 - m^3}{\sigma^3} = \frac{m_3}{\sigma^3}$

Kurtosis: (kurtosis is any measure of the 'peakedness' of the probability distribution)

Define as $\gamma_2 = E\left(\frac{X-m}{\sigma}\right)^4 - 3 = \frac{m_4}{\sigma^4} - 3$