Module-1

Discrete Probability distributions

Probability distributions.

It is a way to shape the sample data to make prediction and draw conclusions about an entire population. is probability distributions is a mathematical function that gives the probabilities of occurence of different possible outcomes for an experiment.

Discrete Probability distributions.

- A discrete probability distribution relates to discrete data.
- It is often used to model uncertain events were the possible values for the variables are either attribute or countable.
- The commonly used discrete probability distributions are uniform, geometric, binomial, Poisson distributions.

Uniform distributions.

- The probability distributions in which all outcomes are equally likely. Those events in a sample space who have same chance of occurence are called equally likely outcomes.

Eg: Throwing a die.

X :	1	2	3	4	5	6
$P(x)$:	$1/6$	$1/6$	$1/6$	$1/6$	$1/6$	$1/6$

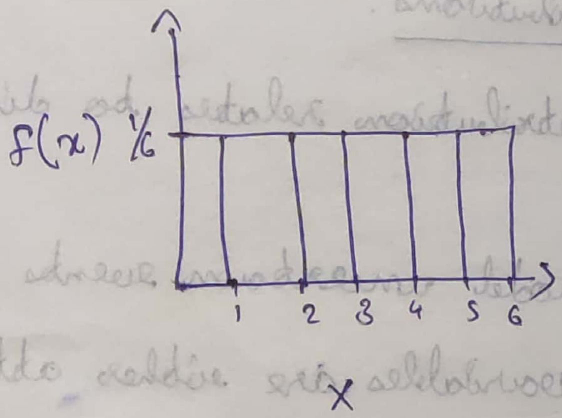
: Tossing a coin

X :	H	T
$P(x)$:	$1/2$	$1/2$

Probability density function

$$f(x) = \frac{1}{n} \quad ; x = 1, 2, \ldots, n$$

n = number of chances.



Mean of Uniform distributions.

$$E(x) = \sum_{x=1}^{n} x \cdot f(x)$$

$$= \sum_{x=1}^{n} x \times \frac{1}{n}$$

$$= \frac{1}{\cancel{n}} \times \frac{\cancel{n}(n+1)}{2} = \frac{n+1}{2}$$

Variance of Uniform distributions.

$$Var(x) = E(x^2) - [E(x)]^2$$

$$E(x^2) = \sum_{x=1}^{n} x^2 \cdot f(x)$$

$$= \frac{1}{n} \times \sum_{x=1}^{n} x^2$$

$$= \frac{1}{\cancel{n}} \times \frac{\cancel{n}(n+1)(2n+1)}{6}$$

$$= \frac{(n+1)(2n+1)}{6}$$

$$Var(x) = \frac{(n+1)(2n+1)}{6} - \left[\frac{n+1}{2}\right]^2$$

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

$$= \left[\frac{n+1}{2}\right]\left[\frac{n-1}{6}\right]$$

$$= \frac{n^2-1}{12}$$

Eg: If x denotes the number shown by an unbiased die then write down the pdf of x and derive the mean and variance of x.

Ans: $x = 1, 2, 3, 4, 5, 6.$

$$f(x) = \frac{1}{6} \quad \text{for } x = 1, 2, \ldots, 6$$

$$\text{Mean} = \frac{n+1}{2} = \frac{7}{6}$$

$$\text{Variance} = \frac{n^2-1}{12} = \frac{35}{12}$$

Examples of Discrete Random variable..

1) Number of children in a family.

2) The number of pages in statistics books.

3) The number of customer who visit a bank during any given hour.

4) The number of heads obtained in three tosses of a coin.

5) The friday night attendence at a cinema.

6) The number of houses in a certain block.

7) The number of patients a doctor sees in one day.

8) The number of defective light bulb in a box of ten.

Examples of continuous random variable.

1) The time taken to complete an examination.

2) Interest rates.

3) Time required to perform a job

4) Volume of soft drink in a 16-oun can

5) Product weights.

6) Financial ratios.

7) Temperature of a cleaning solution.

8) Income levels.

9) Distance between two points.

10) The height and weight of a person.

Geometric distribution.

*. Conditions for geometric distributions:

- Only two categories [ie, Success or failure] [Bernolli trials]
- A trial is repeated until a success occurs.
- The repeated trials are independent of each other.
- The probability of success P is constant for each trial
- x represents the number of trials in which the first success occurs.

Scanned with CamScanner

Distribution formula

$$f(x) = q^{x-1}p \quad ; \quad x = 1, 2, \ldots$$
$$= 0 \quad ; \text{ otherwise.}$$

p - probability of success

q - probability of failure

$p + q = 1$

Eg: The products produced by a machine has a 3% defective rate. What is the probability that the first defective occurs in the fifth item inspected?

Ans: $p = P(\text{Success}) = 0.03$, $q = 1 - p = 0.97$

$$P(X=5) = P(1^{st}\ 4 \text{ non-defectives})\, P(5^{th} \text{ defective})$$
$$= q^4 p = (0.97)^4 \times 0.03$$
$$= \underline{\underline{0.027}}$$

Mean of geometric distribution

$$E(X) = \sum_x x \cdot f(x)$$
$$= \sum_{x=1}^{\infty} x\, p q^{x-1}$$
$$= p \times \sum_{x=1}^{\infty} x\, q^{x-1}$$
$$= p \times (1 + 2q + 3q^2 + \cdots)$$
$$= p(1-q)^{-2} = p\,p^{-2} = \frac{p}{p^2} = \underline{\underline{\frac{1}{p}}}$$

$$V(x) = E(x^2) - [E(x)]^2$$

$$= \frac{2q+p}{p^2} - \frac{1}{p^2}$$

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

$$= \frac{2q-q}{p^2} = \underline{\underline{\frac{q}{p^2}}}.$$

1) Consider the event of throwing darts at a board untill you hit the center area. Let the probability of hitting the center area is ~~p = 0.17~~ 0.17. Find the probability that it takes eight throws until you hit the center.

Ans. ~~p hitting the center.~~

p = Probability of hitting the center.

q = Probability of not hitting the center.

$p = 0.17$; $q = 1 - p = 0.83$; $x = 8$

$P(X=8) = q^{8-1} p$

$= q^7 p = (0.83)^7 \times 0.17$

$= \underline{\underline{0.046}}$

2) Consider the event of finding a store that carries a special printer ink. Of the stores that carry printer ink, only 10% of them carry the special ink. The stores are randomly called. Find the probability that the 3rd store contains the special ink.

Ans. p = probability that the store contains special ink

q = probability that the stores do not contain special ink

~~p = 0.01 ; q = 0.99~~

$p = 0.1$; $q = 0.9$; $x = 3$.

$$P(x=3) = q^{3-1}p$$
$$= q^2 p$$
$$= (0.9)^2 \times 0.1$$
$$= \underline{\underline{0.081}}$$

3) An instructor feels that 15% of students get below a C on their final exams. Find the probability that the 5th answer sheet contains a C grade.

Ans Let p = Probability of getting ~~an answer~~ answer sheet with C grade

q = Probability of getting answer sheet without C grade.

$p = 0.15 \quad q = 0.85 \quad x = 5$

$$p(x=5) = q^4 p$$
$$= (0.85)^4 \times 0.15$$
$$= \underline{\underline{0.092}}$$

4) In an amusement fair, a competitor is entitled for a prize if he throws a ring on a peg from a certain distance. Only 80% of competitors are able to do it. If someone is given 5 chances, what the probability of him winning the prize at the 5th chance

Ans p = Probability of winning the prize.

q = Probability of losing the prize

$P = 0.3$; $q = 0.7$; $x = 5$

$P(x=5) = q^4 P$ $= (0.7)^4 \times 0.3$

$\cancel{= (0.3)^4 \times 0.7}$ $= \underline{0.072}$

$\cancel{= 0.81 \times 0.7}$

$\cancel{= 0.567}$

~~8) A baseball player has a batting average of 0.320. Wh~~
~~is the general probability ...~~

5) A baseball player has a batting average of 0.320. What is the probability that he gets ~~a~~ his first hit ~~...~~ in the third trip to bat.

Ans $P = 0.320$; $q = 0.680$; $x = 3$

$P(x=3) = q^2 P$

$= (0.68)^2 \times 0.32$

$= \underline{0.148}$

<u>Binomial distribution</u>

It allows to compute the probability of the number of successes for a given number of trials.

P - probability of success

q - probability of failure.

X - number of success in 'n' trials.

Scanned with CamScanner

Situations where Binomial distribution can be applied:

- There are only two possible outcomes to each trial. [Success and failure].
- The number of trials is fixed.
- The probability of success is identical for all trials.
- The trials are independent [i.e carrying out one trial has no effect on any other trials].

Mean = $E(x) = np$

Variance = $V(x) = npq = np(1-p)$

Standard deviation = $\sigma = \sqrt{npq}$

n = number of trials.

p = probability of success.

q = probability of failure.

$$\therefore f(x) = {}^{n}C_{x}\, p^{x}\, q^{n-x} \quad ; x = 0, 1, 2, \ldots, n \,;\; p + q = 1$$

$$= 0 \quad ; \text{ otherwise.}$$

Where

- p is the probability of success on any trial.
- $q = 1 - p$ = the probability of failure.
- n - number of trials.
- x - the number of successes ; it can take value $0, 1, 2, \ldots n$
- ${}^{n}C_{x} = \dfrac{n!}{x!(n-x)!}$ denotes number of combinations of n elements taken x at a time.

2) If 20% of the articles produced by a machine are defective, determine the probability that out of 4 articles choosen at random at most two are defective.

Ans. Let x denote the number of defective articles in the sample $x = 0, 1, 2, 3, 4$.

$n = 4$

$p = P(\text{success}) = P(\text{getting defective article}) = 0.2$.

$q = 1 - 0.2 = 0.8$.

pdf of x is $f(x) = {}^{n}C_{x}\, p^{x} q^{n-x} \quad : x = 0, 1, 2, 3, 4.$

$P(\text{at most two are defective}) = P(x \leq 2)$

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

$= {}^4C_0 (0.8)^4 \cdot (0.2)^0 + {}^4C_1 (0.8)^3 \cdot (0.2)^1$

$+ {}^4C_2 (0.8)^2 \cdot (0.2)^2$

$= 0.4096 + 0.4096 + 0.1536$

$= \underline{\underline{0.9728}}$

3) The probability of man hitting the target is 1/4. If he fire 7 times, what is the probability of hitting the target at least twice? How many times should he fire if the probability of his hitting the target at least once is equal to 2/3?

Ans. Let X denote the number of times the target is hit.

$X = 0, 1, 2, 3, 4, 5, 6, 7.$

$n = 7$; $p = 1/4$; $q = 3/4$

pdf of X is $f(x) = {}^nC_x \, p^x q^{n-x}$; $x = 0, 1, 2, \ldots, n$

$= {}^7C_x \, p^x q^{7-x}$; $x = 0, 1, 2, \ldots, 7.$

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

$= 1 - f(0) - f(1)$

$= 1 - {}^7C_0 \left(\frac{3}{4}\right)^7 - {}^7C_1 \left[\frac{1}{4}\right]^1 \left[\frac{3}{4}\right]^6$

$= 1 - 0.1334 - 0.3114$

$= \underline{\underline{0.5552}}$

5) Eight unbiased coins were tossed simultaneously. Find the probability of getting.

a) exactly 4 heads

b) no heads at all

c) 6 or more heads

d) at most two heads

e) number of heads ranging from 3 to 8

Ans. Let x denote the total number of heads that can occur.

$x = 0, 1, 2, 3, 4, 5, 6, 7, 8.$

$\therefore n = 8.$

$p = P(\text{success}) = P(\text{getting head}) = \frac{1}{2}.$

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

pdf of x is $f(x) = {}^nC_x\, p^x q^{n-x}$; $x = 0, 1, 2, \ldots, n$

Here $f(x) = {}^8C_x\, p^x q^{8-x}$; $x = 0, 1, 2, \ldots, 8.$

a) P(exactly 4 heads) = $P(x=4)$

$$= {}^8C_4 \left[\frac{1}{2}\right]^4 \times \left[\frac{1}{2}\right]^4$$

$$= 70 \times \left[\frac{1}{2}\right]^8$$

$$= 0.2734$$

b) P(no heads at all) = $P(x=0)$

$$= {}^8C_0 \times \left[\frac{1}{2}\right]^0 \times \left[\frac{1}{2}\right]^8 = \left[\frac{1}{2}\right]^8 = 0.0039$$

c) $P(6 \text{ or more heads}) = P(x=6) + P(x=7) + P(x=8)$

$$= {}^8C_6 \times \left[\frac{1}{2}\right]^8 + {}^8C_7 \times \left[\frac{1}{2}\right]^8 + {}^8C_8 \times \left[\frac{1}{2}\right]^8$$

$$= \left[\frac{1}{2}\right]^8 \times \left[{}^8C_6 + {}^8C_7 + {}^8C_8\right]$$

$$= \left[\frac{1}{2}\right]^8 \times \left[28 + 8 + 1\right]$$

$$= \left[\frac{1}{2}\right]^8 \times 37.$$

$$= 0.1445$$

d) $P(\text{at most 2 heads}) = P(X \leq 2)$

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

$$= {}^8C_2 \times \left[\frac{1}{2}\right]^8 + {}^8C_1 \times \left[\frac{1}{2}\right]^8 + {}^8C_0 \times \left[\frac{1}{2}\right]^8$$

$$= \left[\frac{1}{2}\right]^8 \times \left[{}^8C_2 + {}^8C_1 + {}^8C_0\right]$$

$$= \left[\frac{1}{2}\right]^8 \times \left[28 + 8 + 1\right]$$

$$= \left[\frac{1}{2}\right]^8 \times 37$$

$$= 0.1445$$

c) P(number of heads ranging from 3 to 5)

$= P(x=3) + P(x=4) + P(x=5)$

$= {}^8C_3 \times \left[\frac{1}{2}\right]^8 + {}^8C_4 \times \left[\frac{1}{2}\right]^8 + {}^8C_5 \times \left[\frac{1}{2}\right]^8$

$= \left[\frac{1}{2}\right]^8 \times \left[{}^8C_3 + {}^8C_4 + {}^8C_5\right]$

$= \left[\frac{1}{2}\right]^8 \times [56 + 70 + 56]$

$= \left[\frac{1}{2}\right]^8 \times 182.$

$= \underline{0.7109}$

Q) Consider families with 4 children each. What percentage of families would you expect to have.

a) two boys ; b) at least one boy

c) no girls ; d) at most two girls

Ans. Let X be probability of getting boy

ie $X = 0, 1, 2, 3, 4$

$n = 4$

$p = P(\text{getting a boy}) = \frac{1}{2}$

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

$f(x) = {}^nC_x \, q^{n-x} \times p^x \; ; \; x = 0, 1, 2, \ldots, n.$

$= {}^4C_x \, q^{4-x} \times p^x \; ; \; x = 0, 1, 2, 3, 4$

a) P (two boys) = $P(x=2)$

$$= {}^4C_2 \times \left[\frac{1}{2}\right]^2 \times \left[\frac{1}{2}\right]^2$$

$$= \frac{\cancel{4!}\, 3\times 4^2}{\cancel{2!} \times \cancel{2!}} \times \left[\frac{1}{2}\right]^4$$

$$= 6 \times \left[\frac{1}{2}\right]^4$$

$$= \underline{\underline{0.375}}$$

b) P (at least one boy) = $P(x \geq 1)$

$$= 1 - P(x < 1)$$

$$= 1 - f(0)$$

$$= 1 - {}^4C_0 \times \left[\frac{1}{2}\right]^4$$

$$= 1 - \frac{1}{16}$$

$$= \underline{\underline{0.9875}}$$

c) P (no girls) = $P(x=4)$

$$= {}^4C_4 \times \left[\frac{1}{2}\right]^4 \times \left[\frac{1}{2}\right]^0$$

$$= \left[\frac{1}{2}\right]^4$$

$$= \underline{\underline{0.0625}}$$

Scanned with CamScanner

d) P(atmost two girls) = P(atleast two boys)

$$= P(x \geq 2)$$

$$= f(2) + f(3) + f(4)$$

$$= {}^4C_2 \times \left[\frac{1}{2}\right]^4 + {}^4C_3 \times \left[\frac{1}{2}\right]^4 + {}^4C_4 \times \left[\frac{1}{2}\right.$$

$$= \left[\frac{1}{2}\right]^4 \times \left[{}^4C_2 + {}^4C_3 + {}^4C_4\right]$$

$$= \left[\frac{1}{2}\right]^4 \times \left[6 + 4 + 1\right]$$

$$= \left[\frac{1}{2}\right]^4 \times 11$$

$$= 0.6875$$

Poisson Distribution

- In Binomial distribution, n is finite. So it is possible to count the number of times an event occur and the number of times it does not occur. But when p is small and n is very large or n is not finite, the use of Binomial distribution is not logical. In such situations we use Poisson distributions.

Scanned with CamScanner

- The Poisson distribution is used in those situations where the probability of happening of an event is small. i.e. the event rarely occurs.

The pdf formula is.

$$f(x) = \frac{e^{-\lambda} \lambda^{x}}{x!} \; ; \quad x = 0, 1, 2, \ldots \quad \lambda > 0$$

λ is the rate of occurrence of the event in unit time.

Examples for Poisson Distribution

→ To count the number of accidents taking place in a day on a busy road.

→ To count the number of errors per page in a typed material.

→ To count the number of incoming telephone calls in an office.

→ To count the number of persons dying due to a rare disease such as heart attack in a year.

Mean of Poisson distribution

$$E(x) = \sum_{x} x \cdot f(x)$$

$$= \sum_{x=0}^{\infty} x \cdot \frac{e^{-\lambda} \times \lambda^{x}}{x!} = \lambda$$

Variance, $V(x) = \lambda$

1) It is known from past experience that in a certain plant there are on the average 4 industrial accidents per month. Find the probability that in a given year there will be less than 3 accidents.

Ans. $f(x) = \frac{e^{-\lambda} \lambda^{x}}{x!} = \frac{e^{-4} 4^{x}}{x!}$ [∵ Here $\lambda = 4$]

$e^{-4} = 0.0183$

∴ $P(x < 3) = p(x=0) + P(x=1) + P(x=2)$

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

$= e^{-4} \left[1 + 4 + 8 \right]$

$= 0.0183 \times 13$

$= \underline{\underline{0.2382}}$

2) In a town 10 accidents took place in a span of 100 days. Assuming that the number of accidents follow Poisson, find the probability that there will be 3 or more accidents in a day.

∴ λ = Average number of accidents $= 10/100 = 0.1$

$\cancel{f(x) = \frac{e^{-x} \cdot \lambda^{x}}{x!}} = \cancel{e^{-0.1} \times}$

$$\therefore f(x) = \frac{e^{-0.1} \cdot (0.1)^x}{x!}$$

$$e^{-0.1} = 0.9048$$

$$P(x \geq 3) = 1 - P(x < 3)$$

$$= 1 - P(x=0) - P(x=1) - P(x=2)$$

$$= 1 - \frac{e^{-0.1} \cdot (0.1)^0}{0!} - \frac{e^{-0.1} \cdot (0.1)^1}{1!} - \frac{e^{-0.1} \cdot (0.1)^2}{2!}$$

$$= 1 - e^{-0.1}\left[1 + 0.1 + 0.005\right]$$

$$= 1 - e^{-0.1} \left[\cancel{0.895}\right] \left[1.105\right]$$

$$= 1 - 0.9048 \cdot \cancel{0.895}\ 1.105$$

$$= 1 - 0.9998$$

$$= \underline{\underline{0.0002}}$$

Scanned with CamScanner

4) During two hours between 8 am and 10 am on an average 1.5 number of phone calls per minute are reported on the switch board of a company. Find the probability that during particular minute there will be

a) no phone calls at all.

b) exactly 8 calls.

c) at least 4 calls.

Ans: Here $\lambda = 1.5$

$$P(x) = \frac{e^{-\lambda} \cdot \lambda^x}{x!} = \frac{e^{-1.5} \cdot (1.5)^x}{x!}$$

a) $P(x=0) = \dfrac{e^{-1.5} * (1.5)^0}{0!}$

$= e^{-1.5} * 1$

$= 0.2231$

b) $P(x=3) = \dfrac{e^{-1.5} * (1.5)^3}{3!}$

$= \dfrac{0.2231 * 3.375}{6}$

$= 0.1254$

c) $P(x \geq 4) = 1 - P(x<4)$

$= 1 - [P(x=0) + P(x=1) + P(x=2) + P(x=3)]$

$= 1 - e^{-1.5}\left[\dfrac{(1.5)^0}{0!} + \dfrac{(1.5)^1}{1!} + \dfrac{(1.5)^2}{2!} + \dfrac{(1.5)^3}{3!}\right]$

$= 1 - e^{-1.5}[1 + 1.5 + 1.125 + 0.5628]$

$= 1 - e^{-1.5} * 4.1878$

$= 1 - 0.9343$

$= 0.0657$

5) If a random variable X follows a Poisson distribution such that $P(X=1) = P(X=2)$, find $P(X=0)$

Ans. $P(X=x) = \dfrac{e^{-\lambda} \times \lambda^{x}}{x!}$

$$P(X=1) = \frac{e^{-\lambda} \times \lambda}{1!}$$

$$P(X=2) = \frac{e^{-\lambda} \times \lambda^{2}}{2!}$$

Given $P(X=1) = P(X=2)$

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

$$\lambda = \frac{2!}{1!} = \underline{\underline{2}}$$

$$P(X=0) = \frac{e^{-\lambda} \times \lambda^{0}}{0!}$$

$$= \frac{e^{-2} \times 2^{0}}{0!} = e^{-2} = \underline{\underline{0.1353}}$$

6) Bank clerks in a certain bank are found to make errors in entering figure in their ~~bat~~ books at the rate of 3 errors in 4 pages. What is the probability that 4 randomly choosen pages will show 2 or more errors.

Ans. Here $\lambda = 3$

$$f(x) = \frac{e^{-\lambda} \times \lambda^{x}}{x!} = \frac{e^{-3} \times 3^{x}}{x!}$$

$$e^{-\lambda} = e^{-3} = \underline{\underline{0.0498}}$$

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

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

$$= 1 - \left[\frac{e^{-3} \times 3^{0}}{0!} + \frac{e^{-3} \times 3}{1!}\right]$$

$$= 1 - [1 + 3] \times e^{-3}$$

$$= 1 - 4 \times 0.0498$$

$$= 1 - 1992$$

$$= \underline{\underline{0.8008}}$$