Date: 25h March 2021

VIRTUAL COACHING CLASSES
ORGANISED BY BOS (ACADEMIC), ICAI

FOUNDATION LEVEL
PAPER 3: BUSINESS MATHEMATICS, LOGICAL REASONING & STATISTICS

Faculty: CA Arijit Chakraborty

Theoretical Probability Distributions-Discussion flow :

- **Probability weightage = 3 to 4 problems in exam : v. imp**
- **Theoretical distribution weightage = 4-5 problems in exam, v.imp**
- Random variable
- Continuous & Discrete random variable
- Probability distribution
- Theoretical distribution – 3 types

  Binomial , Poisson, Normal
- Probability Mass Function (PMF), Probability Density Function (PDF)
- Mean, SD, Mode
- Examples, Ex 17 set A, B,C & additional qs bank

Discussion flow :

- Chap 17 : **Learning Objective : Binomial & Poisson Distribution**
- Binomial distribution is applicable when the **trials are independent and each trial has just 2 outcomes :** success and failure. **2 alternative possibilities:**
- Heads or tails
- Girl or boy
- **applied in:** coin tossing experiments,
- sampling inspection plan,
- genetic experiments
- **profitably employed to make short term projections for the future.**

Random variables

- Often we take **measurements which have different values on different occasions.**
- The values are subject to random variation - **they are not completely predictable, and so are not deterministic.** They are *random variables.*
- Examples -
- **crop yield,**
- **maximum temperature,**
- **number of cyclones in a season,**
- **rain/no rain.**

17.1 Probability distributions

- If we **measure a random variable many times, we can build up a distribution of the values it can take.**
- Imagine an underlying distribution of values which we would get if it was possible **to take more and more measurements under the same conditions.**
- This **gives the probability distribution for the variable.**
- A probability distribution also possesses all the characteristics of an observed distribution. : **mean median, mode, standard deviation etc.**

Discrete probability distributions

- A discrete probability distribution associates a probability with each value of a discrete random variable.
  - ***Example 1. Random variable has two values Rain/No Rain.*** *P(Rain) = 0.2, P(No Rain) = 0.8 gives a probability distribution.*
  - ***Example 2. Let X = Number of sunny days in a 10 day period.*** *P(X=0) = 0.1074, P(X=1) = 0.2684, P(X=2) = 0.3020, … P(X=6) = 0.0055, ...*
- Note that P(rain) + P(No Rain) = 1; P(X=0) + P(X=1) + P(X=2) + … +P(X=6) + … P(X=10) = 1.
- **Two important discrete probability distributions are**
- **(a) Binomial Distribution and (b) Poisson distribution.**

Continuous probability distributions

- Because continuous random variables **can take all values in a range, it is not possible to assign probabilities to individual values.**
- Instead we have a **continuous curve, called a probability density function**, which **allows us to calculate the probability a value within any interval.**
- **This probability is calculated as the area under the curve between the values of interest. The total area under the curve must equal 1.**
- **Continuous probability distributions = Normal Distribution**

Binomial distribution

- f (x) = p (X = x) = n Cx p^x q^{n-x} for x = 0, 1, 2, …., n ……….(17.1)
- = 0, otherwise
- As n >0, p, q $\geq$ 0, it follows that f(x) $\geq$ 0 for every x
- Also å sigma f(x) = f(0) + f(1) + f(2) + …..+ f(n) = 1……….(17.2)
- Binomial distribution is known as biparametric distribution as it is characterised by two parameters n and p.
- This means that if the values of n and p are known, then the distribution is known completely.

- The mean of the binomial distribution is m (mew) = np …. (17.3)
- The variance of the binomial distribution is given by
- sigma ^{2} = npq (17.5)
- Since p and q are numerically less than or equal to 1,
- <u>npq < np = variance of a binomial variable is always less than its mean.</u>
- Also <u>variance of X attains its maximum value at p = q = 0.5 and this maximum value is n/4.</u>

- Depending on the values of the two parameters, binomial distribution may be unimodal or bi- modal. mew 0 , the mode of binomial distribution, is given by :
- mew 0 = the largest integer contained in (n+1)p if (n+1)p is a non-integer
- and (n+1)p – 1 if (n+1)p is an integer ….(17.4)
- Additive property of binomial distribution.
- If X and Y are two independent variables such that X~B (n1, P)
- and Y~B (n2, P)
- Then (X+Y) ~B (n1 + n2 , P) …… (17.6)

- **Example 17.1:** A coin is tossed 10 times. Assuming the coin to be unbiased, what is the probability of getting
- 4 heads?
- at least 4 heads?
- at most 3 heads?
- **Solution:** We apply binomial distribution as the tossing are independent of each other. With every toss, **just two outcomes** either a head, which we call a success or a tail, **probability of a success (or failure) remains constant** throughout.
- Let X denotes the no. of heads. Then X follows binomial distribution with parameter n = 10 and p = 1/2 (since the coin is unbiased). Hence q = 1 – p = ½
- The **probability mass function** of X is given by $f(x) = nCx\ p^{x}\ q^{n-x}$
- $= 10Cx\ .\ (1/2)^{x}\ .\ (1/2)10\text{-}x$
- $= 10Cx\ /\ 1024$ for x = 0, 1, 2, 10

- i) probability of getting 4 heads
- = f (4)
- = 10c4 / 1024
- = 210 / 1024
- = 105 / 512

- **ii) probability of getting at least 4 heads**
- = P (X >= 4)
- = P (X = 4) + P (X = 5) + P (X = 6) + P(X = 7) +P (X = 8)
- = 10c4 / 1024 + 10c5 / 1024 + 10c6 / 1024 + 10c7 / 1024 + 10c8 /1024
- = 837 / 1024

- **(iii) probability of getting at most 3 heads**
- = P (X < =3)
- = P (X = 0) + P (X = 1) + P (X = 2) + P (X = 3)
- = f (0) + f (1) + f (2) + f (3)
- = 10c0 / 1024 + 10c1 / 1024 + 10c2 / 1024 +10c3 / 1024
- 176 / 1024 = 11/64

- **Example 17.2:** If 15 dates are selected at random, what is the probability of getting two Sundays?
- **Solution:** If X denotes the number of Sundays, then it is obvious that X follows binomial distribution with parameter n = 15 and p = probability of a Sunday in a week = 1/7 and q = 1 – p = 6 / 7.
- Then f(x) = 15cx (1/7)x. (6/7)15−x., for x = 0, 1, 2, 15
- Hence the probability of getting two Sundays
- = f(2)
- = $15c2\ (1/7)^2 \cdot (6/7)^{15-2}$
- 0.29

- **Example 17.3:** The incidence of occupational disease in an industry is such that the workmen have a 10% chance of suffering from it. What is the probability that out of 5 workmen, 3 or more will contract the disease?
- **Solution:** Let X denote the number of workmen in the sample. X follows binomial with parameters n = 5 and p = probability that a workman suffers from the occupational disease = 0.1
- Hence q = 1 – 0.1 = 0.9.
- Thus f (x) = PMF = 5cx . (0.1)x. (0.9)5-x For x = 0, 1, 2,…….,5.
- The probability that 3 or more workmen will contract the disease
- = P (x >= 3) = f (3) + f (4) + f (5)
- = 5c3 (0.1)3 (0.9)5-3 + 5c4 (0.1)4. (0.9) 5-4 + 5c5 (0.1)5
- = 10 x 0.001 x 0.81 + 5 x 0.0001 x 0.9 + 1 x 0.00001
- = 0.0081 + 0.00045 + 0.00001 = 0.0086.

- **Example 17.5:** Find the binomial distribution for which mean and standard deviation are 6 and 2 respectively.
- **Solution:** Let x ~ B (n, p)
- Given that mean of x = np = 6 … (1) and SD of x = 2
- Þ variance of x = npq = 4 ….. (2)
- Dividing (2) by (1), we get q= 2/3
- So p = 1/3
- Replacing p by 1/3 in equation (1), we get Þ n = 18
- Thus the probability mass function of x is given by
- f(x) = ncx px q n–x
- = 18cx (1/3)x . (2/3)18–x
- for x = 0, 1, 2,…… ,18

- **Example 17.9:** What is the mode of the distribution for which mean and SD are 10 and **sq root 5** respectively.
- **Solution:** As given np = 10 (1)
- npq = 5 (2)
- on solving (1) and (2), we get n = 20 and p = 1/2
- Hence mode = Largest integer contained in (n+1)p
- = Largest integer contained in (20+1) × 1/2
- = Largest integer contained in 10.50
- = 10.

- **Example 17.10:** If x and y are 2 independent binomial variables with parameters 6 and 1/2 and 4 and 1/2 respectively, what is P (x + y >= 1)?
- **Solution:** Let z = x + y.
- It follows that z also follows binomial distribution with parameters (6 + 4) and ½
- i.e. 10 and ½
- $1 – P (z < 1)$
- $= 1 – P (z = 0)$
- $= 1 – 10c0 (1/2)^{0} . (1/2)^{10–0}$
- $= 1 – 1 / 2^{10}$
- $= 1023 / 1024$

Poisson - Descriptive measures

Given a random variable X in an experiment, we have denoted $f(x) = P(X = x)$, the probability that $X = x$. For discrete events $f(x) = 0$ for all values of x except $x = 0, 1, 2, \ldots\ldots$

Properties of discrete probability distribution

1. $0 \le f(x) \le 1$
2. $\sum f(x) = 1$
3. $\mu = \sum x.f(x)$ [is the *mean*]
4. $\sigma^2 = \sum (x-\mu)^2 . f(x)$ [is the *variance*]

In 2, 3 and 4, summation is extended for all possible discrete values of x.

Note: *For discrete uniform distribution,* $f(x) = \frac{1}{n}$ with $x = 1, 2, \ldots\ldots, n$

$$\mu = \frac{1}{n}\sum_{i=1}^{n} x_i$$

$$\text{and } \sigma^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \mu)^2$$

The Poisson Distribution- example

Poisson distributions are often used ***to describe the number of occurrences of a 'rare' event.***

There are some experiments, ***which involve the occurring of the number of outcomes during a given time interval (or in a region of space).***

Such a process is called Poisson process.

- The **main assumptions are**
- **Events occur**
  - ***1. at random (the occurrence of an event doesn't change the probability of it happening again)***
- **2. at a constant rate**
- Poisson distributions **also arise as approximations to binomials** when n is large and p is small.

Definition of Poisson Distribution

- The probability of getting **x successes in a relatively long time interval T containing k small time intervals t i.e. T = kt. is given by**
- for x = 0, 1, 2, ¥ ...... (17.7) $\frac{e^{-kt}.(kt)^x}{x!}$
- Taking kT = m, the above form is reduced to $\frac{e^{-m}.m^x}{x!}$
- for x = 0, 1, 2, ... ¥ (17.8)
- A random variable X is defined to follow Poisson distribution with parameter m, to be denoted by X ~ P (m) if the PMF of x is given by

$$f(x) = P(X = x) = \frac{e^{-m}.m^x}{x!} \text{ for } x = 0, 1, 2, \ldots \infty$$

$$= 0 \text{ otherwise}$$

Properties

- Since e−m = 1/em >0, whatever may be the value of m, m > 0, it follows that f (x) >= 0 for every x.
- Also it can be established that å sigma f(x) = 1
- i.e. f(0) + f(1) + f(2) +……. = 1 (17.10)
- Poisson distribution is known as a uniparametric distribution as it is characterised by only one parameter m.
- The mean of Poisson distribution is given by m i,e mew = m. (17.11)
- The variance of Poisson distribution is given by sigma2 = m (17.12)
- Like binomial distribution, Poisson distribution could be also unimodal or bimodal depending upon the value of the parameter m.

- We have mew 0 = The **largest integer contained in m** if m is a non-integer
- and **m–1 if m is an integer** (17.13)
- **Poisson approximation to Binomial distribution**
- If **n, the number of independent trials of a binomial distribution, tends to infinity** and **p, the probability of a success, tends to zero**, so that
- **m = np remains finite**, then a **binomial distribution with parameters n and p can be approximated by a Poisson distribution with parameter m (= np).**
- In other words when **n is rather large and p is rather small so that m = np is moderate then**
- **b (n, p) @ P (m) (17.14)**

- **Additive property of Poisson distribution**
- If X and y are two independent variables following Poisson distribution with parameters **m1 and m2** respectively, **then Z = X + Y also follows Poisson distribution with parameter (m1 + m2).**
- i.e. if X ~ P (m1) and Y ~ P (m2)
- and X and Y are independent, then
- Z = X + Y ~ P (m1 + m2) (17.15)

Poisson distributions –example 1

- Suppose that we can assume that the number of cyclones, X, in a particular area in a season has a Poisson distribution with a mean (average) of 3. Then P(X=0) = 0.05, P(X=1) = 0.15, P(X=2) = 0.22, P(X=3) = 0.22, P(X=4) = 0.17, P(X=5) = 0.10, … Note:
  - *There is no upper limit to X, unlike the binomial where the upper limit is n.*
  - *Assuming a constant rate of occurrence, the number of cyclones in 2 seasons would also have a Poisson distribution, but with mean 6.*

- **Example 17.11:** Find the mean and standard deviation of x where x is a Poisson variate satisfying the condition P (x = 2) = P (x = 3).
- **Solution:** Let x be a Poisson variate with parameter m. The probability max function of x is then given by f (x) = e-*m* .m x / x!
- for x = 0, 1, 2, ¥
- now, P (x = 2) = P (x = 3)
- Þ f(2) = f(3)
- e-m .m2/ 2!= e -m .m3/3 !
- 1 – m / 3 = 0 (as e–m > 0, m > 0)
- Þ m = 3
- Thus the mean of this distribution is m = 3 and standard deviation = sq root of 3= @ 1.73.

- **Example 17.12:** The probability that a random variable x following Poisson distribution would assume a positive value is (1 – e–2.7). What is the mode of the distribution?
- **Solution:** If x ~ P (m), then its probability mass function is given by
- The probability that x assumes a positive value
- = P (x > 0)
- = 1– P (x £ 0)
- = 1 – P (x = 0)
- = 1 – f(0)
- = 1 – e–m
- As given, 1 – e–m = 1 – e–2.7
- Þ e–m = e–2.7 , hence m = 2.7
- Thus mode mew 0 = largest integer contained in 2.7 , = 2

- **Example 17.14:** X is a Poisson variate satisfying the following relation: P (X = 2) = 9P (X = 4) + 90P (X = 6).
- What is the standard deviation of X?
- **Solution:** Let X be a Poisson variate with parameter m. Then the probability mass function of X is
- Thus P (X = 2) = 9P (X = 4) + 90P (X = 6)
- Þ f(2) = 9 f(4) + 90 f(6)
- e–m .m2 (m2 + 4) (m2 – 1) = 0
- Þ m2 – 1 = 0 (as e–m > 0 m > 0 and m2 + 4 ¹ 0)
- Þ m =1 (as m > 0, m ¹ –1)
- Thus the standard deviation of X is sq root of 1 = 1

- **Fitting a Poisson distribution**
- Since Poisson distribution is uniparametric, we equate m, the parameter of Poisson distribution, to the arithmetic mean of the observed distribution and get the estimate of m.
- i.e. $\hat{m} = \bar{x}$
- The fitted Poisson distribution is then given by
- $\hat{f}(x) = e^{-\hat{m}} \cdot (\hat{m})^x / x!$ For x= 0,1,2,3,…… infinity

Recap

- Types of PD
- Industry & Business applications
- Random variable
- Discrete & continuous Probability distribution
- PMF & PDF
- Features of Binomial & Poisson distribution – mean, SD, mode,
- Solved examples – analysis

THANK YOU