

Introduction to probability distribution

June 15, 2018

1 Starters

- When the random variable is discrete in nature, its probability distribution is characterized by **Probability Mass Function (PMF)**.

No. of fruits sold	no. of customers	PMF
3	30	30/60
5	20	20/60
7	10	10/60
	60	sum = 1

- A **Cumulative Distribution Function (CDF)** defines the less than, greater than or equal to argument of a function.

CDF is a monotonic increasing function.

For the above PMF, CDF of $P(X \leq x)$ could be calculated as -

No. of fruits sold	no. of customers	PMF	CDF
3	30	30/60	30/60
5	20	20/60	(30/60) + (20/60)
7	10	10/60	(30/60) + (20/60) + (10/60)
	60	sum = 1	last value itself becomes 1

- When the random variable is continuous in nature, its probability distribution is characterized by **Probability Density Function (PDF)**.

2 Discrete PD

Topics: Uniform — Binomial — Negative binomial — Poisson

2.1 Uniform

- A random variable which assumes equal probability for its outcomes, is termed as discrete uniform PD.
e.g. getting 5 in a throw of a dice. Same goes with other number on dice.

2.2 Binomial

- a binomial distribution is formed by multiplying the total number of ways an event can occur to the probability of one of the way that event can occur.
- PMF of Binomial can be written as -

$$PMF = C(n, x) * p^x q^{n-x}$$

where,

$C(n, x)$ determines the binomial coefficient and also the total number of ways the event can occur

n total number of trails

x total number of success

$n - x$ total number of failures

p probability of success

q probability of failure

- Example: consider a shop sells 4 kinds of fruits mango, kiwi, banana and apple.
Further, consider a customer purchases only one out of all four which brings equal probability to all four fruits.
Let us determine the probability of a customer buying an apple in upcoming 5 customers.
Moving with the binomial distribution we can arrive at following PMF -

$$PMF = C(4, 1) * 0.25^1 (1 - 0.25)^3 = 0.42$$

R code: `dbinom(x = 1, size = 4, prob = 0.25)`

2.3 Negative binomial

- In this distribution we need to pass the number of success not the number of trails and hence makes it opposite of binomial distribution and so does negative word comes.
- PMF of Negative Binomial can be written as -

$$PMF = C(x - 1, r - 1) * p^r q^{x-r}$$

where,

$x = r, r + 1, r + 2, \dots$

r number of success trails

- Example: consider the previous fruits example of a shop with four fruits with equally likely outcome. Let us try to find the probability that it takes exactly 7 trails to find 2 customers who purchases an apple.
here, $r = 2, x = 7, p = 0.25, q = (1 - 0.25)$

$$\text{PMF} = C(6, 1) * 0.25^2(1 - 0.25)^5 = 0.42$$

R code: `dnbinom(x = 5, size = 2, prob = 0.25)`

2.4 Poisson

- a poisson distribution is built upon three conditions -
 - given an interval of real number constituting of an outcome of interest/success, if this interval is broken down into various sub-intervals
 1. there occurs only one success per sub-interval i.e. probability of more than one success in a sub-interval is zero.
 2. the probability of success in all sub-intervals remain same and proportional to the length of whole interval.
 3. the count of success in each sub-interval is independent of each other
- PMF of poisson

$$\frac{e^{-\lambda} \lambda^x}{x!}$$

where

x number of success

λ expected number of successes

- Example: consider 15 customers arrive in an hour, what is the probability that exactly 20 customers can arrive in another hour.

$$\text{PMF} = \frac{e^{-20} 20^{15}}{15!} = 0.051$$

R code: `dpois(x = 15, lambda = 20)`

3 Continuous PD

Topics: uniform — normal — log normal — exponential — gamma — weibull — t — chi-squared — f

3.1 Uniform

- a continuous random variable X can consists of any values in range $[a, b]$ with PDF $\frac{1}{b-a}$

- Example: consider weight of a person can vary between 70.8kg to 110kg.

Here, the PDF will be $\frac{1}{110-70.8}$

R code: `dunif(x = 80, min = 70.8, max = 110)`

and CDF for having his weight 85kg is $\frac{85-70.8}{110-70.8}$

R code: `punif(x = 80, min = 70.8, max = 110)`