PROBABILITY FUNCTION AND DISTRIBUTION FUNCTION

Presented
By
Pradip Panda
Asstt. Professor, Deptt. Of Statistics
Serampore College

CONTENTS:

- Probability Function
- Probability – mass function
- Probability - density function
- Joint probability – mass function
- Joint probability – density function
- Distribution function

PROBABILITY FUNCTION

- Probability – mass function (p.m.f.)
- Probability – density function (p.d.f.)

With the help of these functions , the probabilities of the random variables can be found for different values or different sets of values.

PROBABILITY – MASS FUNCTION

For a discrete random variable X, there exists a function f(x) such that f(x) = P(X = x), x being the general value of X.

A function of this type satisfying the conditions:

$$i)\ f(x) \geq 0, \text{ for all } x,$$

$$ii) \sum_{x} f(x) = 1,$$

is called the probability – mass function (p.m.f.) of X.

Mr. Pradip Panda, Asstt. Prof., Deptt. of Statistics, Serampore College

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

Is it a probability mass function?

Proof: $Here, \; f(x) \geq 0 \;\; for \; all \; x.$

$$and \sum_{x=0}^{\infty} f(x) = \sum_{x=0}^{\infty} \frac{e^{-\lambda}\lambda^x}{x!}$$

$$= e^{-\lambda}\left[\frac{\lambda^0}{0!} + \frac{\lambda}{1!} + \frac{\lambda^2}{2!} + \dots\dots\dots\right] = e^{-\lambda}e^{\lambda} \qquad = 1$$

So, f(x) is a p.m.f.

PROBABILITY - DENSITY FUNCTION

For a continuous random variable X, there may exist a function f(x) such that, for $a \leq b$,

$$\int_a^b f(x)\,dx = P(a \leq X \leq b).$$

A function of this type satisfying the conditions:

$$i)\ f(x) \geq 0, \forall x$$

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

is called the probability – density function (p.d.f.) of X.

Mr. Pradip Panda, Asstt. Prof., Deptt. of Statistics, Serampore College

Is the following a probability density function?

$$f(x) = 2x\,,\ 0 < x \le 1$$
$$= 4 - 2x\,, 1 < x \le 2$$
$$= 0\,,\ elsewhere$$

Proof: *Here,* $f(x) \ge 0\,,\ \forall x$ and

$$\int_{-\infty}^{\infty} f(x)\,dx = \int_0^1 2x\,dx + \int_1^2 (4-2x)\,dx$$

$$= 2.\frac{x^2}{2}\Big]_0^1 + \left(4x - 2.\frac{x^2}{2}\right)\Big]_1^2$$

$$= 1 + [(8-4)-(4-1)] = 2 \ne 1$$

Mr. Pradip Panda, Asstt. Prof., Deptt. of Statistics, Serampore College

So, f(x) is not a probability density function.

JOINT PROBABILITY - MASS FUNCTION

For a pair of discrete random variables X and Y , a function f(x,y) , called the joint probability – mass function of X and Y, such that

f(x,y) = P(X = x, Y = y) , x and y being the general values of X and Y, and f(x,y) satisfies the conditions :

$$i)\ f(x,y) \geq 0, \forall\ x, y$$

$$and\ ii) \sum_{x}\sum_{y} f(x,y) = 1$$

Mr. Pradip Panda, Asstt. Prof., Deptt. of Statistics,
Serampore College

JOINT PROBABILITY – DENSITY FUNCTION

For two continuous random variables X and Y , the joint probability – density function, f(x,y), is such that :

$$\int_c^d \int_a^b f(x,y)\,dx\,dy = P(a \le X \le b, c \le y \le d)$$

and f(x,y) satisfies the conditions:

$$i)\, f(x,y) \ge 0,\ \forall x, y$$

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

MARGINAL AND CONDITIONAL PROBABILITIES

DISCRETE RANDOM VARIABLE:

If f(x,y) be the joint probability – mass function of X and Y, then the marginal distribution of X and Y are respectively given by :

$$g(x) = \sum_{y} f(x, y), \quad and \; h(y) = \sum_{x} f(x, y)$$

The conditional distribution or array distribution of X given Y= y, is given by

$$g(x/y) = \frac{f(x, y)}{h(y)}, \; provided \, h(y) > 0$$

The conditional distribution or array distribution of Y given X= x, is given by

$$h(y/x) = \frac{f(x, y)}{g(x)}, \; provided \, g(x) > 0$$

Continuous random variable :

If f(x,y) be the joint probability – density function of X and Y, then the marginal distribution of X and Y are respectively given by :

$$g(x) = \int_y f(x,y)\,dy \qquad and\; h(y) = \int_x f(x,y)\,dx$$

If f(x,y) = g(x).h(y) , for all x and y , then the variables X and Y are said to be independent.

$$Cov(X,Y) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} (x-\mu_X)(y-\mu_Y)\, f(x,y)\,dx\,dy$$

$$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xy\; f(x,y)\,dx\,dy - \mu_X\;\mu_Y$$

$$Where,\ \mu_X = \int_{-\infty}^{\infty} x\, g(x)\,dx \quad and\ \mu_Y = \int_{-\infty}^{\infty} y\, h(y)\,dy$$

Is the following a joint probability – density function? If yes, examine whether the variables are independent or not.

$$f(x,y)=\frac{2}{3}(x+1)e^{-y},\ 0<x<1,\ y>0$$
$$=0,\qquad elsewhere.$$

Solution: $Here,\ f(x,y)\geq 0,\ for\ all\ x,y$ and

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)\,dx\,dy=\int_0^{\infty}\int_0^1 \frac{2}{3}(x+1)e^{-y}dx\,dy$$
$$=\frac{2}{3}\int_0^{\infty} e^{-y}dy.\int_0^1 (x+1)\,dx\ =\frac{2}{3}\left[-e^{-y}\right]_0^{\infty}.\left[\frac{x^2}{2}+x\right]_0^1\ =\frac{2}{3}.1.\frac{3}{2}=1$$

So, f(x,y) represents a joint p.d.f.

Marginal distribution of X is given by $g(x) = \int_{-\infty}^{\infty} f(x, y)\, dy$

$$= \int_0^{\infty} \frac{2}{3}(x+1)e^{-y}\, dy \quad = \frac{2}{3}(x+1)\int_0^{\infty} e^{-y}\, dy = \frac{2}{3}(x+1).1 \quad = \frac{2}{3}(x+1)$$

Marginal distribution of Y is given by $h(y) = \int_{-\infty}^{\infty} f(x, y)\, dx$

$$= \int_0^{1} \frac{2}{3}(x+1)e^{-y}\, dx \quad = \frac{2}{3}e^{-y}\int_0^{1} (x+1)\, dx \quad = \frac{2}{3}e^{-y}.\frac{3}{2} \quad = e^{-y}$$

Thus, $f(x, y) = \frac{2}{3}(x+1)e^{-y} = g(x).h(y)$ *for all x and y.*

So, the variables X and Y are independent.

Mr. Pradip Panda, Asstt. Prof., Deptt. of Statistics,
Serampore College

If the joint p.d.f. is given as

$$f(x,y)=e^{-x-y},\ x>0,\ y>0$$
$$=0 \quad ,elsewhere$$

Find the covariance between X and Y?

Solution: Marginal distribution of X is $g(x)=\int_{-\infty}^{\infty} f(x,y)\,dy$

$$=e^{-x}\int_{0}^{\infty} e^{-y}\,dy = e^{-x},\ x>0$$

and g(x) = 0 , otherwise

Marginal distribution of Y is $h(y)=\int_{-\infty}^{\infty} f(x,y)\,dx$

$$=e^{-y}\int_{0}^{\infty} e^{-x}\,dx \;=e^{-y},\; y>0$$

and h(y) = 0 , otherwise

$$\therefore E(X)=\int_{-\infty}^{\infty} x.\,g(x).dx \;=\int_{0}^{\infty} x.e^{-x}dx \;=\int_{0}^{\infty} e^{-x}.x^{2-1}dx$$

$$=\Gamma 2=1$$

Similarly, E(Y) = 1

$$E(XY)=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xy\; f(x,y)\,dx\,dy \;=\int_{0}^{\infty}\int_{0}^{\infty} xy\,e^{-x-y}dx\,dy$$

$$=E(X).E(Y)=1$$

So, Cov(X, Y) = 0.

DISTRIBUTION FUNCTION

It is useful in finding the quantiles and is denoted by F(x).

$$F(x) = P(X \le x)$$

Properties:

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

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

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

$$iv)\ F(a) \le F(b)\ \ for\ a < b$$

$$and\ v)\ F(x)\ is\ right\ continuous.$$

For a continuous random variable X with p.d.f. $f(x)$,

$$Mean(\mu) = E(X) = \int_{-\infty}^{\infty} x.f(x)dx$$

$$Median(\mu_e) \text{ is such that } P(X \le \mu_e) = \int_{-\infty}^{\mu_e} f(x)dx = \frac{1}{2}.$$

$$Mode(\mu_0) \text{ is the value at which } f(x) \text{ is maximum}$$

$$Variance(\sigma^2) = E[X - E(X)]^2 = \int_{-\infty}^{\infty} (x-\mu)^2 f(x)dx$$

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

$$\text{Mean deviation about } 'a' = E(|X-a|) = \int_{-\infty}^{\infty} |x-a|.f(x)dx$$

$$\text{Moment of order } r \text{ about } 'a' = \mu_r'(a) = E(X-a)^r$$

$$= \int_{-\infty}^{\infty} (x-\mu)^r .f(x)dx$$

$$\text{The distribution function } F(x) \text{ is defined as}$$

$$F(x) = P(X \le x) = \int_{-\infty}^{x} f(y)\,dy$$

$$f(x) = \frac{d}{dx}[F(x)];\ F(x) \text{ is a differentiable function.}$$

$$For\ a\ discrete\ random\ variable\ X\ with\ p.m.f.\ f(x),$$

$$Mean(\mu) = E(X) = \sum_x x.P(X = x) = \sum_x x.f(x).$$

$$Median(\mu_e)\ is\ that\ value\ of\ the\ variable\ which\ is$$

$$such\ that\ P(X \prec \mu_e) \prec \frac{1}{2} \leq P(X \leq \mu_e).$$

$$Mode(\mu_0)\ is\ the\ value\ of\ the\ variable\ with\ the$$

$$highest\ probability.$$

$$Variance(\sigma^2) = E[X - E(X)]^2 = \sum_x (x - \mu)^2.f(x)$$

$$= \sum_x (x-\mu)^2 . f(x) = \sum_x x^2 . f(x) - \mu^2 .$$

$$Mean\ deviation\ about\ 'a' = E(|X-a|) = \sum_x |x-a| . f(x).$$

$$Moment\ of\ order\ r\ about\ 'a' = \mu_r^{'}(a) = E(X-a)^r$$

$$= \sum_x (x-a)^r . f(x)$$

$$The\ distribution\ function\ F(x)\ is\ defined\ as$$

$$F(x) = P(X \le x) = \sum_{y \le x} f(y).$$

Mr. Pradip Panda, Asstt. Prof., Deptt. of Statistics,
Serampore College

Thank You