Chapter 2 _B: Basic Statistical Concept

Random Variable, Probability Mass Function and Probability Density Function , Mathematical Expectation

Samir K Mahajan, Ph.D. , UGC NET

A. Random Variable or Random Quantity/ or Stochastic Variable or Aleatory Variable

- **Described informally as a variable whose values depends on outcome of a random phenomena.**
- **Formal mathematical treatment of random variables is a topic in probability theory.**

A. Random Variable contd.

- A random variable X is a real-valued function whose domain is a set of possible outcomes of a random experiment, and range is a sub-set of the set of real numbers and has the following properties:

i) Each particular value of the random variable can be assigned some probability i.e.

$$0 \leq p(X=x) \leq 1$$

ii) Uniting all the probabilities associated with all the different values of the
random variable gives the value 1(unity).

$$\sum p(X = x) = 1$$

A. Random Variable : Discrete and Continuous

- **A discrete random variable takes finite or countable number of distinct value and takes integer/ whole number such as 0, 1, 2, 3,.....**

- **Continuous variables are numeric variables that have an infinite number of possible values between any two values. It takes integer as well as fractional.**

  - **Hence, continuous random variable is not defined at specific values.**

  - **The probability of observing any single value is equal to zero**

A.2 Random Variable and Probability Distribution

- **PROBABILITY MASS FUNCTION i.e. Distribution of Discrete random variable**
- **PROBABILITY DENSITY FUNCTION i.e. Distribution of Continues Random Variable**

A.2.1 PROBABILITY MASS FUNCTION (pmf)

Let X be a r.v. which takes the values $x_1, x_2, \ldots$ and let $P[X = x_i] = p(x_i)$. This function $p(x_i)$, $i = 1, 2, \ldots$ defined for the values $x_1, x_2, \ldots$ assumed by X is called probability mass function of X satisfying $p(x_i) \geq 0$ and $\sum_i p(x_i) = 1$.

The set $\{(x_1, p(x_1)), (x_2, p(x_2)), \ldots\}$ specifies the probability distribution of a discrete r.v. X. Probability distribution of r.v. X can also be exhibited in the following manner:

X	x_1 x_2 $x_3 \ldots$
$p(x)$	$p(x_1)$ $p(x_2)$ $p(x_3) \ldots$

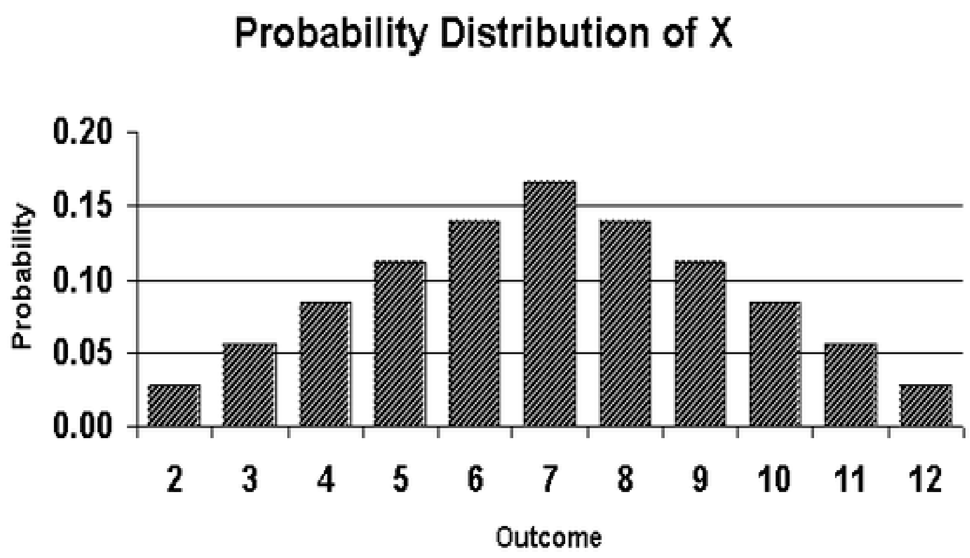
A.2.1 PROBABILITY MASS FUNCTION: Example

Example: Let X represent the sum of two dice.

Then the probability distribution of X is as follows:

X	2	3	4	5	6	7	8	9	10	11	12
$P(X)$	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{3}{36}$	$\frac{4}{36}$	$\frac{5}{36}$	$\frac{6}{36}$	$\frac{5}{36}$	$\frac{4}{36}$	$\frac{3}{36}$	$\frac{2}{36}$	$\frac{1}{36}$

To graph the probability distribution of a discrete random variable, construct a **probability histogram**.



A.2.1 PROBABILITY Density FUNCTION (pdf)

To determine the distribution of a discrete random variable we can either provide its PMF or CDF. For continuous random variables, the CDF is well-defined so we can provide the CDF. However, the PMF does not work for continuous random variables, because for a continuous random variable $P(X = x) = 0$ for all $x \in \mathbb{R}$. Instead, we can usually define the **probability density function (PDF)**. The PDF is the **density** of probability rather than the probability mass.

A.2.1.2 PROBABILITY Density FUNCTION (pdf) ctd.

Probability Density Function

Let $f(x)$ be a continuous function of x. Suppose the shaded region ABCD shown in the following figure represents the area bounded by $y = f(x)$, x-axis and the ordinates at the points x and $x + \delta x$, where δx is the length of the interval $(x, x + \delta x)$.

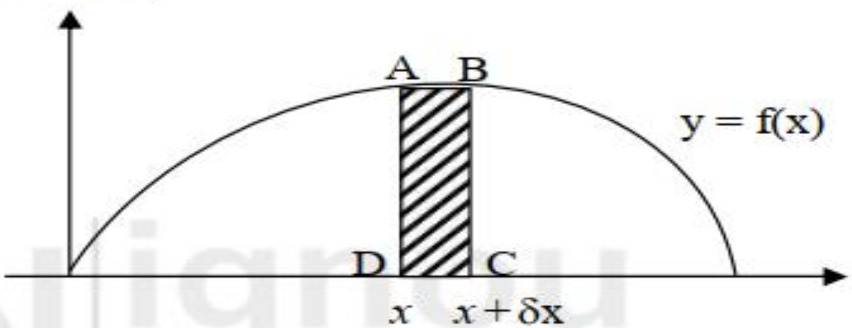


Fig. 5.1

Now, if δx is very-very small, then the curve AB will act as a line and hence the shaded region will be a rectangle whose area will be $AD \times DC$ i.e. $f(x)\,\delta x$ [$\because$ AD = the value of y at x i.e. $f(x)$, DC = length δx of the interval $(x, x+\delta x)$]

Also, this area = probability that X lies in the interval $(x, x+\delta x)$

$$= P[x \le X \le x+\delta x]$$

A.2.1.2 PROBABILITY Density FUNCTION (pdf) ctd.

Hence,

$$P[x \le X \le x + \delta x] = f(x)\delta x$$

$$\Rightarrow \frac{P[x \le X \le x + \delta x]}{\delta x} = f(x), \text{ where } \delta x \text{ is very-very small}$$

$$\Rightarrow \lim_{\delta x \to 0} \frac{P[x \le X \le x + \delta x]}{\delta x} = f(x).$$

$f(x)$, so defined, is called probability density function.

Probability density function has the same properties as that of probability mass function. So, $f(x) \ge 0$ and sum of the probabilities of all possible values that the random variable can take, has to be 1. But, here, as X is a continuous random variable, the summation is made possible through 'integration' and hence

$$\int_R f(x)\,dx = 1,$$

where integral has been taken over the entire range R of values of X.

A.2.1.2 PROBABILITY Density FUNCTION (pdf) ctd.

probability that it takes a value in specified intervals is non-zero and is calculable as a definite integral of the probability density function of the random variable and hence the probability that a continuous r.v. X will lie between two values a and b is given by

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

B. Characteristics of Distribution of Random Variables : Mathematical Expectation

Two widely used characteristics of Probability Distribution are

- **Mean /Expected Value of Random Variable**
- **Variance of a Random Variable**
- **Covariance of Random Variables**
- **Correlation**

B. 1. Expected Values/ Mean of Random Variables

❑ For Discrete Random Variables

The expected value of a discrete rv X, denoted by $E(X)$, is defined as follows:

$$E(X) = \sum_x x f(x)$$

where $\sum_x$ means the sum over all values of X and where $f(x)$ is the (discrete) PDF of X.

❑ For Continuous Random Variables

The expected value of a continuous rv is defined as

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

The only difference between this case and the expected value of a discrete rv is that we replace the summation symbol by the integral symbol.

B. 1.1. Properties of Expected Values

1. The expected value of a constant is the constant itself. Thus, if b is a constant, $E(b) = b$.

2. If a and b are constants,

$$E(aX + b) = aE(X) + b$$

This can be generalized. If $X_1, X_2, \ldots, X_N$ are N random variables and $a_1, a_2, \ldots, a_N$ and b are constants, then

$$E(a_1X_1 + a_2X_2 + \cdots + a_NX_N + b) = a_1E(X_1) + a_2E(X_2) + \cdots + a_NE(X_N) + b$$

B. 1.2. Properties of Expected Values contd.

3. If X and Y are *independent* random variables, then

$$E(XY) = E(X)E(Y)$$

That is, the expectation of the product XY is the product of the (individual) expectations of X and Y.

B. 2. Variance of Random Variables

Let X be a random variable and let $E(X) = \mu$. The distribution, or spread, of the X values around the expected value can be measured by the variance, which is defined as

$$\operatorname{var}(X) = \sigma_X^2 = E(X - \mu)^2$$

The positive square root of σ_X^2, σ_X, is defined as the **standard deviation** of

For computational convenience, the variance formula given above can also be expressed as

$$\begin{aligned}\operatorname{var}(X) = \sigma_x^2 &= E(X - \mu)^2 \\ &= E(X^2) - \mu^2 \\ &= E(X^2) - [E(X)]^2\end{aligned}$$

B. 2.1. Properties of Variance

1. $E(X-\mu)^2 = E(X^2) - \mu^2$, as noted before.

2. The variance of a constant is zero.

3. If a and b are constants, then

$$\operatorname{var}(aX+b) = a^2 \operatorname{var}(X)$$

4. If X and Y are *independent* random variables, then

$$\operatorname{var}(X+Y) = \operatorname{var}(X) + \operatorname{var}(Y)$$

$$\operatorname{var}(X-Y) = \operatorname{var}(X) + \operatorname{var}(Y)$$

This can be generalized to more than two independent variables.

5. If X and Y are *independent* rv's and a and b are constants, then

$$\operatorname{var}(aX+bY) = a^2 \operatorname{var}(X) + b^2 \operatorname{var}(Y)$$

B. 3. Covariance of random variables

Let X and Y be two rv's with means μ_x and μ_y, respectively. Then the **covariance** between the two variables is defined as

$$\operatorname{cov}(X, Y) = E\{(X - \mu_x)(Y - \mu_y)\} = E(XY) - \mu_x \mu_y$$

It can be readily seen that the variance of a variable is the covariance of that variable with itself.

B. 3.1 Properties of Covariance of Random Variables

Properties of Covariance

1. If X and Y are independent, their covariance is zero, for

$$\begin{aligned}\operatorname{cov}(X, Y) &= E(XY) - \mu_x\mu_y \\ &= \mu_x\mu_y - \mu_x\mu_y \qquad \text{since } E(XY) = E(X)E(Y) = \mu_x\mu_y \text{ when } X \text{ and } Y \text{ are independent} \\ &= 0\end{aligned}$$

2.

$$\operatorname{cov}(a + bX, c + dY) = bd\operatorname{cov}(X, Y)$$

where a, b, c, and d are constants.

B.4 Correlation

Correlation Coefficient

The (population) correlation coefficient ρ (rho) is defined as

$$\rho = \frac{\operatorname{cov}(X, Y)}{\sqrt{\{\operatorname{var}(X)\operatorname{var}(Y)\}}} = \frac{\operatorname{cov}(X, Y)}{\sigma_x \sigma_y}$$

Thus defined, ρ is a measure of *linear* association between two variables and lies between -1 and $+1$, -1 indicating perfect negative association and $+1$ indicating perfect positive association.

From the preceding formula, it can be seen that

$$\operatorname{cov}(X, Y) = \rho \sigma_x \sigma_y$$

B.4 Correlation cotd.

Variances of Correlated Variables. Let X and Y be two rv's. Then

$$\begin{aligned}\operatorname{var}(X+Y) &= \operatorname{var}(X)+\operatorname{var}(Y)+2\operatorname{cov}(X,Y)\\ &= \operatorname{var}(X)+\operatorname{var}(Y)+2\rho\sigma_x\sigma_y\\ \operatorname{var}(X-Y) &= \operatorname{var}(X)+\operatorname{var}(Y)-2\operatorname{cov}(X,Y)\\ &= \operatorname{var}(X)+\operatorname{var}(Y)-2\rho\sigma_x\sigma_y\end{aligned}$$

If, however, X and Y are independent, $\operatorname{cov}(X, Y)$ is zero, in which case the $\operatorname{var}(X+Y)$ and $\operatorname{var}(X-Y)$ are both equal to $\operatorname{var}(X)+\operatorname{var}(Y)$, as noted previously.

B.5 Conditional Expectation

Conditional Expectation. Note that $E(X \mid Y)$ is a random variable because it is a function of the conditioning variable Y. However, $E(X \mid Y = y)$, where y is a specific value of Y, is a constant.

Conditional Variance. The conditional variance of X given $Y = y$ is defined as

$$\operatorname{var}(X \mid Y = y) = E\{[X - E(X \mid Y = y)]^2 \mid Y = y\}$$

Reference :

https://www.investopedia.com/terms/p/population.asp

IGNOU Books

Inputs in these slides are exclusively collected from above sources.