Ph.D. Economics (Course Work)
Research Methodology: ECON6001

RANDOM VARIABLES

Shreedhar Satyakam
Department of Economics
Mahatma Gandhi Central University, Bihar

RANDOM VARIABLE

A variable whose value is determined by the outcome of a chance experiment is called a random variable. (Gujarati, 2006)

Intuitively by a random variable we mean a real number X associated with the outcome of a random experiment E. Thus to each outcome ω, there corresponds a real number X(ω). (Gupta and Kapoor, 2006)

If S is the sample space, i.e., outcomes of an event, B is the σ-field of subsets in S, and P is a probability function on B, then a random variable is a function X(ω) with domain S and range $(-\infty, \infty)$ such that for every real number a, the event $[\omega : X(\omega) \leq a] \in B$.

Some Theorems on Random Variables

- A function $X(\omega)$ from S to R $(-\infty, \infty)$ is a random variable if and only if for real a, $\omega : X(\omega) < a] \in B$.

- If X_1 and X_2 are random variables and C is a constant then CX_1, X_1+X_2, $X_1 X_2$ are also random variables.

- If X_1 and X_2 are random variables, then (i) max $[X_1, X_2]$, and (ii) min $[X_1, X_2]$ are also random variables.

- If X is a random variable and $f(.)$ is a continuous function, then $f(X)$ is a random variable.

- If X is a random variable and $f(.)$ is an increasing function, then $f(X)$ is a random variable.

Distribution Function

A distribution function is the cumulative distribution function associated with a particular random variable X.

It is defined for a random variable X as

$$F(x) = P\,(X \le x) = P\,\{\,\omega : X\,(\omega) \le x\,\},\ -\infty < x < \infty$$

F(x) is defined for all real x. The domain of the distribution function is $(-\infty, \infty)$ and its range is [0,1].

If F is the distribution function of the random variable X and if a < b, then $P\,(a < X \le b) = F\,(b) - F\,(a)$.

Discrete Random Variable

A real valued function defined on a discrete sample space is called a discrete random variable. (Gujarati, 2006)

A discrete random variable can assume only a countable number of real values and these values depend on chance. Thus, values of a discrete random variable are distinct and separated by finite distances. Discrete random variables are also called discrete stochastic variables or discrete chance variables.

The entire set of permissible values together with their respective probabilities is called the probability distribution of the random variable X or probability mass function. (Koutsoyiannis, 2001)

Continuous Random Variable

A random variable is said to be continuous when its different values cannot be put in 1-1 correspondence with a set of positive integers. (Gujarati, 2006)

Thus, a continuous random variable may assume an infinite number of values within any given interval. It can take all possible integral and fractional values between certain limits.

In the case of a continuous variable the probability of any particular value must be zero, i.e. $P(X=X^*) = 0$. However, the probability of of X assuming values within an interval (X_1 and X_1), no matter how small this interval might be, is a finite number can be computed. (Koutsoyiannis, 2001)

$P(X=X^*) = 0$ does not meant that the value X_i^* is impossible. It should be interpreted as the average probability of values very close in the neighbourhood of X_i^*. (Goldberger, 1964)

The probability distribution of a continuous variable is called probability density function.

Probability Mass Function or Discrete Probability Density Function

If X is a discrete random variable with distinct values $x_1, x_{2, \ldots,}$ $x_{n, \ldots,}$ then the probability mass function of the random variable X is defined as

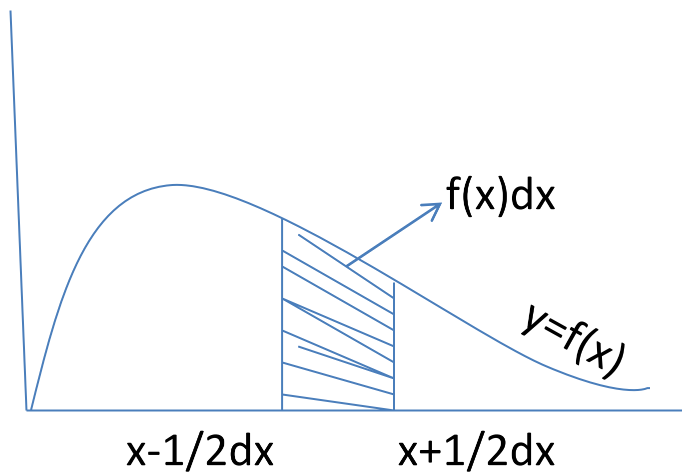
$$f(x) = P(X = x_i) \qquad \text{for } i = 1, 2, \ldots., n, \ldots$$

$$= 0 \qquad \text{for } x \neq x_i$$

Probability Density Function (of a Continuous Random Variable)

If X is a continuous random variable and there be a small interval of length dx around the point x. The probability of X falling in the infinitesimal interval (x, x+dx) is given by *f(x)dx*, where *f(x)* is a continuous function of X. Symbolically,

$$P\ (x \leq X \leq x+dx) = f_X(x)dx$$



The expression $f_X(x)dx$ is known as the probability differential and the curve y = f(x) is known as the probability density curve. $f_X(x)dx$ represents the area bounded by the curve y = f(x), x-axis and ordinates at the points x and x+dx.

$f_X(x)$ so defined is known as the probability density function and it is expressed as

$$f_X(x) = \lim_{\delta x \to 0} \frac{P\,(x \leq X \leq x+dx)}{\delta x}$$

Characteristics of Probability Distributions

The moments of the distribution characterise a probability distribution. Mean or expected value and variance are the two most widely used moments.

Mathematical Expectation or Expected Value of a random variable

Mathematical expectation or expected value refers to the average value of a random phenomenon.

The expected value of a discrete random variable is a weighted average of all possible values of the random variable, where the weights are the probabilities associated with the corresponding values.

If X is a discrete random variable, the expected value is defined as

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

Where $\sum_x$ is the sum over all value of X and where f(x) is the probability mass function of X.

In case of a continuous random variable X with probabilty density function f(x), the expected value is defined as

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

Let X is a random variable with probability density function (or probability mass function) F(x). If g(.) is a function such that g(X) is a random variable then E[g(X)] is defined as

$E[g(X)] = \sum_x g(x)f(x)$ (for discrete random variable)

$E[g(X)]= \int_{-\infty}^{\infty} g(x)f(x)dx$ (for continuous random variable)

Properties of Expectation

Addition Theorem of Expectation

If X and Y are random variables, then $E(X+Y) = E(X) + E(Y)$, provided all the expectations exist.

Multiplication Theorem Expectation

If X and Y are independent random variables, then

$$E(XY) = E(X) \, . \, E(Y)$$

Other Properties

- If X is a random variable and 'a' is a constant, then

  (i) $E[a\Psi(X)] = a\,E[\Psi(x)]$

  (ii) $E[\Psi(X)+a] = E[\Psi(x)]+a$

Where $\Psi(X)$ is a function of X and it is a random variable and all the expectations exist.

- If X is a random variable and ‘a’ and ‘b’ are constants, then

$$E\,(a X + b) = a\,E\,(X) + b$$

- Expectation of a Linear Combination of Random Variables

Let $X_1, X_2, \ldots, X_n$ are any n random variables and if $a_1, a_2, \ldots, a_n$ are any n constants, then

$$E\left(\sum_{i=1}^{n} a_i X_i\right) = \sum_{i=1}^{n} a_i\, E\,(X_i)$$

Provided all the expectations exist.

- If X and Y are two random variables such that $Y \leq X$, then $E\,(Y) \leq E\,(X)$, provided all the expectations exist.

- $|E\,(X)| \leq E\,|X|$, provided the expectations exist.

Variance of a Random Variable

Variance of a random variable measures the distribution of spread of the values of the random variable X around the expected value of X.

Variance of a random variable is defined as

(i) $Var\ (X) = \sum_x (x-\mu)^2 f(x)$ (for discrete random variable)

(ii) $Var\ (X) = \int_{-\infty}^{\infty} (x-\mu)^2 f(x)dx$ (for continuous random variable)

Variance of a random variable can also be expressed as

$$Var\ (X) = E(X^2) - [E(X)]^2$$

Properties of Variance

- If X is a random variable, then V (aX+b) = a^2 V(X), where a and b are constants.
- If b = 0, then V (aX) = a^2 V(X), i.e. variance is not independent of change of scale.
- If a = 0, then V (b) = 0, i.e. variance of a constant is zero.
- If a = 1, then V (X+b) = V(X), i.e. variance is independent of change of origin.

- If X and Y are independent random variables, then

$$V\ (X+Y) = V(X) + V(Y)$$

$$V\ (X - Y) = V(X) + V(Y)$$

- If X and Y are independent random variables and a and b are constants, then

$$V\ (aX + bY) = a^2\ V(X) + b^2\ V(Y)$$

Covariance of a Random Variable

If X and Y are two random variables with means μ_x and μ_y , respectively, then the covariance between the two variables is defined as

$$Cov\ (X, Y) = E\{(X- \mu_x)\ (Y- \mu_Y)\} = E(XY) - \mu_x\mu_Y$$

Properties of Covariance

- If X and Y are independent random variables, their covariance is zero.

- $Cov(a+bX, c+dY) = bd\ Cov\ (X, Y)$

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

Reading List

i. Gujarati, D.N. (2004). Basic Econoetrics, Tata Mcgraw Hill, New Delhi.

ii. Gupta, S.C. and V. K. Kapoor (2006). Fundamentals of Mathematical Statistics, Sultan Chand & Sons, New Delhi.

iii. Koutsoyiannis, A. (2007). Theory of Econometrics, Palgrave, New York.