Ch. 4. Continuous Random variables & Their Probability Distributions

The type of r.v. that takes on any value in an interval is called continuous.

The p.d. for a continuous r.v., unlike the p.d. for a discrete r.v., can not be obtained by assigning nonzero probabilities to all the points on a line interval and at the same time satisfy the requirement that the probabilities of the distinct possible values sum to one.

2. Probability Density Function (p.d.f.) for a Continuous R.V.

Distribution Function

Let Y be any r.v.. The *distribution function* (or *cumulative distribution function,c.d.f.*) of Y,

denoted by $F(y)$ is:

$$F(y) = P(Y \le y), \ -\infty < y < \infty.$$

Properties of F(y)

1. $$\lim_{y \to -\infty} F(y) = F(-\infty) = 0.$$

2. $$\lim_{y \to \infty} F(y) = F(\infty) = 1.$$

3. If y_1 and y_2 are any values such that $y_1 < y_2$, then $F(y_1) \le F(y_2)$. Thus, $F(y)$ is a

non decreasing function of y.

4. The c.d.f. of a discrete r.v. Y increases in jumps or steps at each of the possible values of Y and stayed flat between the possible values of Y.

5. The c.d.f. $F(y)$ of a continuous r.v. Y is continuous for $-\infty < y < \infty$.

6. The continuous r.v. have zero probability at discrete points, $P(Y = y) = 0$.

Probability Density Function

$f(y)$, called the $p.d.f.$ for the continuous r.v. Y, is given by

$$f(y) = \frac{dF(y)}{dy} = F'(y).$$

Thus, it implies that

$$F(y) = \int_{-\infty}^{y} f(t)dt.$$

Properties of a p.d.f.

1. $f(y) \geq 0$, for any value of y.

2. $\int_{-\infty}^{\infty} f(y)dy = 1$.

3. In general the $c.d.f.$ for a continuous r.v. must be continuous, but the $d.f.$ need not to be everywhere continuous.

4. If the continuous r.v. Y has a p.d.f. $f(y)$ and $a \leq b,$ then the probability that Y falls in the interval $[a, b]$ is

$$P(a \leq Y \leq b) = \int_{a}^{b} f(y)dy.$$

3. Expected Value of a Continuous R.V.

Let Y be a continuous r.v. with the p.d.f. $f(y)$. Then the expected value of Y:

$$E(Y) = \int_{-\infty}^{\infty} yf(y)dy.$$

Expected Value of a function of a R.V.

Let $g(Y)$ be a function of a continuous r.v. Y with p.d.f. $f(y)$. Then

$$E[g(Y)] = \int_{-\infty}^{\infty} g(y)f(y)dy.$$

Variance of a Continuous R.V.

$$\sigma^2 = E[(Y-\mu)^2] = \int_{-\infty}^{\infty} (Y-\mu)^2 f(y)dy.$$

and standard deviation $= +\sqrt{V(Y).}$

More Results on Expected Value

If c is a constant, then

$$E(c) = c.$$

$$E[c\, g(Y)] = c\, E[g(Y)]$$

$$E[g_1(Y) + g_2(Y) + ..] = E[g_1(Y)] + E[g_2(Y)] + ..$$

4. Uniform Probability Distribution

A r.v. Y is said to have a continuous uniform probability distribution on the interval (θ_1, θ_2) iff the p.d.f. of Y is

$$f(y) = \frac{1}{\theta_2 - \theta_1}, \theta_1 \leq y \leq \theta_2.$$

The constants that determine the specific form of a p.d.f. are called *parameters* of the p.d.f..

a. The quantities θ_1 and θ_2 are parameters of the uniform p.d.f..

b. Let Y be a uniform distribution with parameters θ_1 and θ_2. Then

$$\mu = E(Y) = \frac{\theta_1 + \theta_2}{2}; \ \sigma^2 = V(Y) = \frac{(\theta_1 - \theta_1)^2}{12}.$$

5. Normal Probability Distribution

Measurements on many continuous r.v's appear to have been generated from population frequency distributions that are bell-shaped and are closely approximated by a normal distribution.

A continuous r.v. Y is said to have a normal p.d. iff the p.d.f. of Y is

$$f(y) = \frac{e^{-\frac{(y-\mu)^2}{2\sigma^2}}}{\sigma\sqrt{2\pi}}, \sigma > 0, \ -\infty < \mu, y < \infty.$$

1. The normal p.d.f. has two parameters mean $= \mu$ and standard deviation $= \sigma$.

2. The normal p.d.f. is symmetric around mean μ and its mean $=$ median $=$ mode.

3. The standard normal variable (s.n.v.) Z is a normal p.d. with mean 0 and standard deviation 1.

4. z-score is the distance from the mean of a normal r.v. measured in units of standard deviation of the original normal r.v. given

by $z = \frac{y-\mu}{\sigma}$.

6. Gamma Probability Distribution

Some r.v's. are always non-negative and have distributions of data which are skewed (non-symmetric) to the right (meaning most of the area under the p.d.f. is located near the origin).

Examples: The lengths of time between malfunctions for aircraft engines; the lengths of time between arrivals at a super-store checkout queue; the lengths of time to complete a maintenance checkup for automobile or aircraft engine.

A random variable Y is said to have a gamma p.d. with parameters α and β iff the p.d.f. of Y is

$$f(y) = \frac{y^{\alpha-1}e^{-y/\beta}}{\beta^{\alpha}\Gamma(\alpha)}, \ \alpha, \beta > 0; \ 0 \le y < \infty,$$

where $\Gamma(\alpha) = \int_0^{\infty} y^{\alpha-1}e^{-y}dy,$ called the gamma function.

1. $\Gamma(\alpha) = (\alpha - 1)\Gamma(\alpha - 1),\ \alpha \geq 1$ and $\Gamma(n) = (n-1)!.$ $(\Gamma(0) = 1)$.

2. α is called the shape parameter and β is the scale parameter.

3. $\mu = E(Y) = \alpha\beta,\ \sigma^2 = V(Y) = \alpha\beta^2$.

4. Let υ be a positive integer. A random variable Y is said to have a $chi - square$ p.d. with υ degrees of freedom iff

$$Y \sim G(\frac{\upsilon}{2}, 2).$$

5. If Y is a chi-square r.v. with υ degrees of freedom then

$$\mu = E(Y) = \upsilon,\ \sigma^2 = V(Y) = 2\upsilon.$$

6. The gamma p.d.f. for $\alpha = 1$ is called the exponential p.d.f..

7. A random variable Y is said to have an exponential p.d. with parameter β iff the p.d.f. of Y is

$$f(y) = \frac{e^{-y/\beta}}{\beta},\ \beta > 0;\ 0 \le y < \infty.$$

8. The exponential p.d.f. is often useful for modeling the length of life of electronic components.

9. If Y is an exponential r.v. with parameter β, then

$$\mu = E(Y) = \beta,\ \sigma^2 = V(Y) = \beta^2.$$

7. Beta Probability Distribution

The beta function is a two-parameter d.f., defined over the closed interval $0 \le y \le 1$, and provides a good model for proportions such as the proportion of impurities in a chemical product or the proportion of time a machine is under repair.

A random variable Y is said to have a beta p.d. with parameters α and β iff the p.d.f. of Y is

$$f(y) = \frac{y^{\alpha-1}(1-y)^{\beta-1}}{B(\alpha,\beta)},\ \alpha,\beta > 0;\ 0 \le y \le 1,$$

where

$$B(\alpha,\beta) = \int_0^1 y^{\alpha-1}\ (1-y)^{\beta-1} dy = \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)}.$$

1. If $Y \sim B(\alpha,\beta)$, then

$$\mu = \frac{\alpha}{\alpha+\beta},\quad \sigma^2 = \frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}.$$

2. The c.d.f. for the beta r.v. is called the *incomplete beta function* and is

$$F(y) = \int_0^y \frac{t^{\alpha-1}(1-t)^{\beta-1}}{B(\alpha,\beta)} dt = I_y(\alpha,\beta).$$

3. The beta d.f. can be applied to a r.v. defined on the interval $c \le y \le d$ by transforming to a new beta r.v. as

$$y^* = \frac{y-c}{d-c}.$$

4. For integral values of α and β ,
$I_y(\alpha,\beta)I_y(\alpha,\beta)$ is related to the binomial p.f..
For $y = p$,

$$F(p) = \int_0^p \frac{t^{\alpha-1}(1-t)^{\beta-1}}{B(\alpha,\beta)}dt$$

$$= \sum_{i=\alpha}^{n} \binom{n}{i} p^i(1-p)^{n-i},$$

where 0<p<1 and n$=\alpha+\beta-1$.

9. Expected Values of a Continuous R. V.

1. The $i-th\ moment$ of a continuous r.v. Y taken about the *origin* is defined to be

$$\mu_i^{'} = E(Y^i) = \int y^i f(y)dy, \text{ i=1,2,...}$$

2. The $i-th\ moment$ of a continuous r.v. Y taken about its *mean*, or the $i-th$ *central moment* of Y, is defined to be

$$\mu_i = E[(Y-\mu)^i] = \int (y-\mu)^i f(y)dy, \text{ i=1,2,...}$$

3. The moment-generating function $m(t)$ for a continuous r.v. Y is defined to be

$$m(t) = E(e^{ty}) = \int e^{ty} f(y)dy.$$

4. Let $g(Y)$ be a single-valued function of a continuous r.v. Y with p.d.f. $f(y)$. Then

$$E[e^{tg(y)}] = \int e^{tg(y)} f(y)dy.$$

11. Tchebysheff's Theorem

The same as in case of the discrete r.v..