Chapter Five
Continuous Distributions

The Chi-Square Distribution:

A random variable *X* of the continuous type that has the *pdf*

$$f(x) = \frac{1}{\Gamma(n/2)2^{n/2}} x^{n/2-1} e^{-x/2} \qquad , \quad 0 < x < \infty$$

where

Γ(n) = (n - 1)(n - 2)…(3)(2)(1)Γ(1) = (n - 1)!

and

Γ(1) = 1

and the moment generating function:

$$M(t) = (1-2t)^{-n/2}, \qquad t < \frac{1}{2}$$

is said to have a chi - square distribution. The mean and the variance of the chi - square distribution are $\mu = n$ and $\sigma^2 = 2n$, respectively. For no obvious reason, we call the parameter *n* the number of degrees of freedom of the chi - square distribution. Because the chi - square distribution has an important role in statistics and occurs so frequently , we will use the notation $X \sim \chi_n^2$ to signify that *X* has a chi - square distribution with *n* degrees of freedom.

Example 1: If *X* has the *pdf*

$$f(x) = \begin{cases} \frac{1}{4} x e^{-x/2}, & 0 < x < \infty \\ 0 & elsewhere \end{cases}$$

Then *X* is $\chi^2(4)$. Hence $\mu = 4$, $\sigma^2 = 8$ and

$$M(t) = (1-2t)^{-2}, \qquad t < \frac{1}{2}$$

Example 2: If *X* has the moment generating function

$$M(t) = (1-2t)^{-8}, \qquad t < \frac{1}{2}$$

then *X* is $\chi^2(16)$.

The Exponential Distribution:

A continuous random variable *X* whose probability density function is given, for some λ > 0, by

$$f(x) = \begin{cases} \lambda e^{-\lambda x} & x \geq 0 \\ 0 & x < 0 \end{cases}$$

is said to have an exponential distribution with parameter λ. The cumulative distribution function *F(x)* is given by

$$\begin{aligned} F(x) &= P(X \leq x) \\ &= \int_0^x \lambda e^{-\lambda t} dt \\ &= 1 - e^{-\lambda x} \quad , x \geq 0 \end{aligned}$$

The exponential distribution often arises, in practice, as being the distribution of the amount of time until some specific event occurs. For instance, the amount of time (starting from now) until an earthquake occurs, or until a new war breaks out, or until a telephone call you receive turns out to be a wrong number are all random variables that tend in practice to have exponential distributions.

The moment generating function of the exponential is given by:

$$\begin{aligned} M(t) = E[e^{tx}] &= \int_0^\infty e^{tx} \lambda e^{-\lambda x} dx \\ &= \lambda \int_0^\infty e^{-(\lambda - t)x} dx \\ &= \frac{\lambda}{\lambda - t} \qquad t < \lambda \end{aligned}$$

Differentiation yields

$$M'(t) = \frac{\lambda}{(\lambda - t)^2}$$

$$M''(t) = \frac{2\lambda}{(\lambda - t)^3}$$

And so

$$E[X] = M'(0) = \frac{1}{\lambda}$$

$$V(X) = E[X^2] - [E(X)]^2$$

$$= M''(0) - \frac{1}{\lambda^2} = \frac{1}{\lambda^2}$$

Hence λ is the reciprocal of the mean.

Example 1:

Let *X* have an exponential distribution with a mean of 100. Find the probability that *X* is less than 90.

$$f(x) = \frac{1}{100} e^{-x/100}, \qquad 0 \le x < \infty$$

$$p(X < 90) = \int_0^{90} \frac{1}{100} e^{-x/100} dx = 1 - e^{-90/100} = 0.593$$

<u>Normal Distribution</u>:

The normal (or Gaussian) distribution is the best known continuous distribution and occupies a central position in statistical theory and practice. It describes many random phenomena that occur in every day life, including test scores, weights heights, and many others. Errors in measurements are assumed to have a normal distribution.

The random variable *X* has a normal distribution if its *pdf* is defined by:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp[-\frac{(x-\mu)^2}{2\sigma^2}] \qquad -\infty < x < \infty$$

where μ and σ are parameters satisfying $-\infty < \mu < \infty$, $\sigma > 0$.
Briefly, we say that X is $N(\mu, \sigma^2)$
$E(X) = \mu$ location parameter
$Var(X) = \sigma^2$ shape parameter

Properties of the normal distribution:
1. The curve of the normal distribution is symmetrical about the mean.
2.The mean, median and mode are identical.
3.The two tails of the normal curve are asymptotic, they approach closer to the X-axis and will never touch it.
4. There are two points of inflection on the curve. They occur at a distance of one standard deviation to the left and to the right of the mean.
5. The total area under the normal curve represents *100%* of the individual items of the distribution. The bulk of the individual items lie very close to the mean $\mu \pm \sigma = 68\%$ of the individual items. $\mu \pm 2\sigma$ will include *95%* of the frequencies and $\mu \pm 3\sigma$ will include *99.73%* of the items of the distribution.

The *CDF* of the normal distribution is given by

$$F(X) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}\sigma} \exp[-\frac{(y-\mu)^2}{2\sigma^2}]dy$$

The expression cannot be evaluated in closed form. Very exact tables are available for normal distribution. These tables are based on the following standard normal *pdf*:
The normal distribution with $\mu = 0$ and $\sigma = 1$ is referred to as the standard normal distribution. A random variable Z has a standard normal distribution if the *pdf* of Z is

$$f(Z) = \frac{1}{\sqrt{2\pi}} \exp(-\frac{z^2}{2}) \qquad -\infty < z < \infty$$

The cumulative distribution function of *Z* is:

$$\Phi(Z) = p(Z \le z) = \int_{-\infty}^{z} \frac{1}{\sqrt{2\pi}} \exp(-\frac{y^2}{2}) dy$$

Because of the symmetry of the standard normal *pdf* ,we have that:

$$\Phi(-Z) = 1 - \Phi(Z)$$

When *X* is $N(\mu, \sigma^2)$, probabilities involving *X* are computed by standardizing .The standardized variable $Z = \frac{X-\mu}{\sigma}$ is *N(0, 1)*.
Since any normally distributed random variable can be transformed to the standard normal probabilities can be evaluated simply by having a table of standard normal integrals available. In particular

$$p(X \le x) = p(Z \le \frac{x-\mu}{\sigma}) = \Phi(\frac{X-\mu}{\sigma}) = \Phi(Z)$$

Values of the integral are given for *Z* between *0* and *3.09.*

Example 1:

If *Z* is a random variable that is *N(0, 1)*, find:

$P(-0.5 \le Z \le 1.5) = p(0 \le Z \le 1.5) + p(-0.5 \le Z \le 0)$
$= A_1 + A_2 = 0.4332 + 0.1915$
$= 0.6247$

Example 2:

a. $p(Z \le 1) = p(Z \le 0) + p(0 \le Z \le 1)$
$= 0.5 + 0.3413 = 0.8413$

b. $p(Z > 1) = 0.5 - p(0 \le Z \le 1)$
$= 0.5 - 0.3413 = 0.1587$

Or $p(Z > 1) = 1 - p(Z \le 1) = 1 - 0.8413$
$= 0.1587$

c. $p(Z < -1.5) = p(Z > 1.5)$
$= 0.5 - p(0 \le Z \le 1.5)$
$= 0.5 - 0.4332 = 0.0668$

d. $p(-1.5 \le Z \le 0.5) = p(-1.5 \le Z \le 0) + p(0 \le Z \le 0.5)$

$= 0.4332 + 0.1915 = 0.6247$

e. Find the value of Z_0 such that $p(0 \leq Z \leq Z_0) = 0.49$

We look for the probability *0.49* on the area side of the table. The closest we can come is at *0.4901* which corresponds to $Z_0 = 2.33$.

Example 3:

If *X* is *N(4.35, 0.3481)*, find

$$p(X \geq 5.2) = p(\frac{X-\mu}{\sigma} \geq \frac{5.2-4.53}{0.59}) = p(Z \geq 1.44)$$

$= 0.5 - 0.4251 = 0.0749$

Example 4:

If *X* is *N(16, 1)*, find

$$p(X > 17) = p(\frac{X-\mu}{\sigma} > \frac{17-16}{1}) = p(Z > 1) = 0.1587$$