Lecture

INSTITUTE OF ACTUARIAL & QUANTITATIVE STUDIES

Class: FY BSc

Subject : Probability and Statistics -1

Chapter: Unit 2 Chapter 2

Chapter Name: Theoretical Continuous Distributions

Today's Agenda

0. Introduction

1. Continuous Distributions
   1. Uniform Distribution
   2. Gamma Distribution
   3. Exponential Distribution
   4. Chi-square Distribution
   5. Beta Distribution
   6. Normal Distribution
   7. Standard Normal Distribution
   8. Lognormal Distribution
   9. T-Distribution
   10. F-Distribution

0 Introduction

In this chapter we shall study some of the probability densities that figure most prominently in statistical theory and in applications.

1.1 Uniform Distribution

A random variable X has a **uniform distribution** and it is referred to as a continuous uniform random variable if and only if its probability density is given by

$$f_x(x) = \frac{1}{\beta-\alpha} \quad \alpha < x < \beta$$

$$= 0 \quad \text{otherwise}$$

The parameters α and β of this probability density are real constants, with $\alpha < \beta$

The mean and the variance of the uniform distribution are given by:

$$\mu = \frac{\alpha+\beta}{2}$$

$$\sigma^2 = \frac{(\beta-\alpha)^2}{12}$$

1.2 Gamma Distribution (including exponential and chi-square)

The gamma family of distributions has 2 positive parameters and is a versatile family. The PDF can take different shapes depending on the values of the parameters. The range of the variable is $\{x: x > 0\}$.

First note that the gamma function Γ (α) is defined for α > 0 as follows:
$\Gamma(\alpha) = \int_0^\infty y^{\alpha-1} e^{-y} \, dy$

Note in particular that Γ (1) = 1

1.2 Gamma Distribution

The PDF of the gamma distribution with parameters α and λ is defined by:

$$f_x(x) = \frac{\lambda^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\lambda x} \quad \text{for } x > 0$$

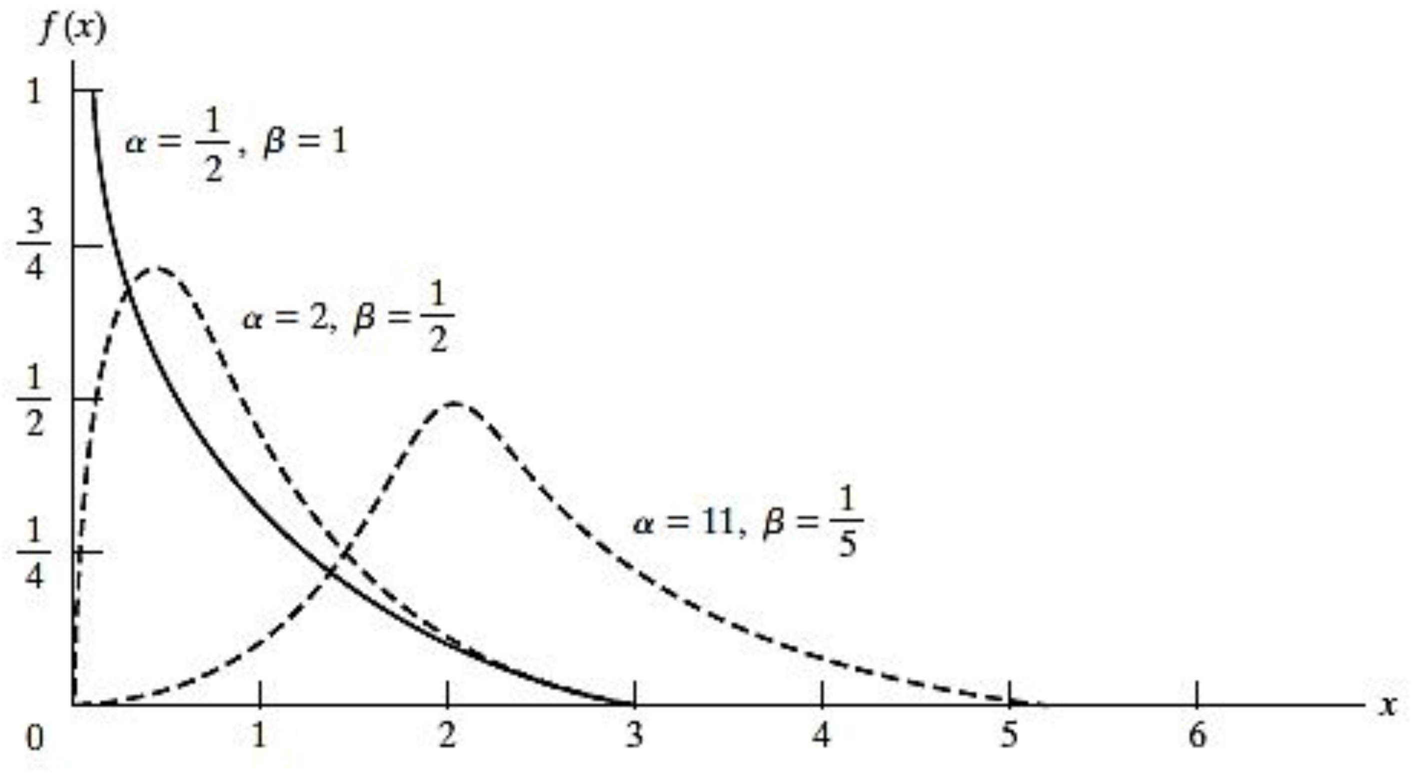
The moments of the distribution are:

$$\mu = \frac{\alpha}{\lambda}$$

$$\sigma^2 = \frac{\alpha}{\lambda^2}$$

1.2 Gamma Distribution

To get some idea about the shape of the graphs of gamma densities, those for several special values of α and β are shown in Figure. Some special cases of the gamma distribution play important roles in statistics;



1.3 Exponential Distribution (Gamma with α = 1)

A random variable X has an **exponential distribution** and it is referred to as an exponential random variable if its probability density is given by:

$f_x(x) = \lambda e^{-\lambda x}$ for $x > 0$

The moments of the distribution are:

$$\mu = \frac{1}{\lambda}$$

$$\sigma^2 = \frac{1}{\lambda^2}$$

The CDF is given as

$$F_X(x) = \int_0^x \lambda e^{-\lambda t}\, dt = 1 - e^{-\lambda x}$$

1.3 Exponential Distribution

The exponential distribution is used as a simple model for the lifetimes of certain types of equipment. Very importantly, it also gives the distribution of the waiting-time, *T*, from one event to the next in a Poisson process with rate λ.

P(T>t) = P(0 events in time t)
= P(X = 0) -> where X ~ Poisson (λt)
= $e^{-\lambda t}$

P(T<t) = 1 - $e^{-\lambda t}$

f_T(t) = $\lambda e^{-\lambda t}$

In fact the time from any specified starting point (not necessarily the time at which the last event occurred) to the next event occurring has this exponential distribution. This property can also be expressed as the "**memoryless" property**. So for the exponential distribution we can also show that:
P(X > x + n | X > n) = P(X > x)

1.4 Chi-square Distribution

Another **special case of the gamma distribution arises when $\alpha = \frac{\nu}{2}$ and λ = 1/2,** where ν is the lowercase Greek letter nu.

A random variable X has a **chi-square distribution** and it is referred to as a chi-square random variable if its probability density is given by:

$$f_x(x) = \frac{\left(\frac{1}{2}\right)^{\left(\frac{v}{2}\right)}}{\Gamma(v/2)} x^{\left(\frac{v}{2}\right)-1} e^{-\left(\frac{1}{2}\right)x} \quad \text{for } x > 0.$$

The moments of the distribution are:

$\mu = \nu$

$\sigma^2 = 2\nu$

The parameter ν is referred to as the **number of degrees of freedom**, or simply the **degrees of freedom**. The chi-square distribution plays a very important role in sampling theory.

1.4 Useful result

If *W Gamma* ~ (α, λ) , then 2λ*W* has a $\chi^2_{2\alpha}$ distribution (*ie* a chi-square distribution with 2α degrees of freedom).

This is an important result as it is the only practical way we can calculate probabilities for a gamma distribution in an exam.
Probability tables for the chi-square distribution can be found on pages 164-166 of the *Tables*

1.5 Beta Distribution

A random variable X has a **beta distribution** and it is referred to as a beta random variable if its probability density is given by:

$$f_X(x) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} x^{\alpha-1}(1-x)^{\beta-1} \qquad \text{for } 0<x<1.$$

The moments of the distribution are:

$$\mu = \frac{\alpha}{\alpha+\beta}$$

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

The (continuous) uniform distribution on (0,1) is a special case (with $\alpha=\beta=1$).

1.6 Normal Distribution

The **normal distribution**, is in many ways the cornerstone of modern statistical theory. It was investigated first in the eighteenth century when scientists observed an astonishing degree of regularity in errors of measurement. They found that the patterns (distributions) that they observed could be closely approximated by continuous curves, which they referred to as “normal curves of errors” and attributed to the laws of chance.

Many numerical populations have distributions that can be fit very closely by an appropriate normal curve.

1.6 Normal Distribution

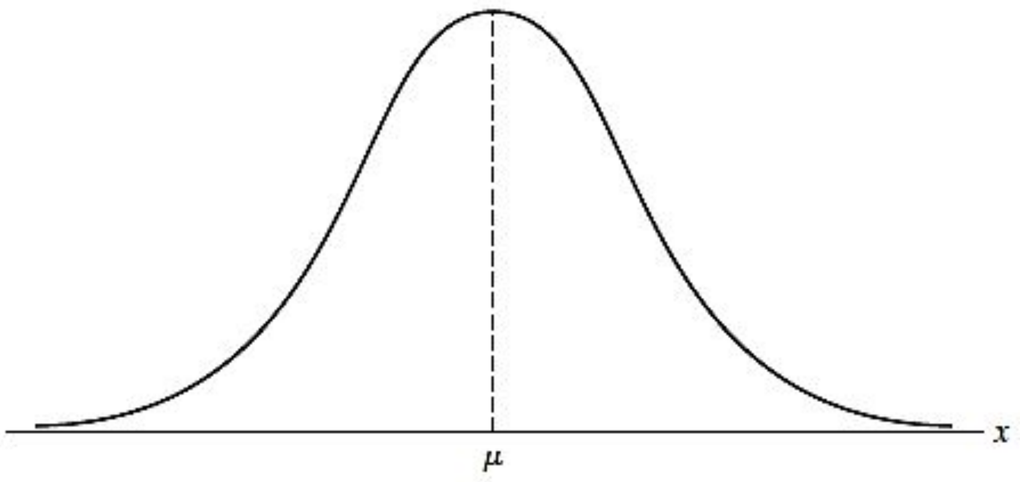
A random variable X has a **normal distribution** and it is referred to as a normal random variable with ts probability density given by

$$f_X(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} \quad \text{for } -\infty < x < \infty$$

The parameter μ is, in fact, $E(X)$ and that the parameter σ is, in fact, the square root of var(X), where X is a random variable having the normal distribution with these two parameters.

A linear function of a normal variable is also a normal variable, *ie* if X is normally distributed, so is $Y = aX + b$.

The graph of a normal distribution, shaped like a bell, is shown in Figure.



1.7 Standard Normal Distribution

The normal distribution with μ = 0 and σ = 1 is referred to as the standard normal distribution.

$$f_X(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2} \qquad \text{for } -\infty < x < \infty$$

Since the normal distribution plays a basic role in statistics and its density cannot be integrated directly, its areas have been tabulated for the special case where μ = 0 and σ = 1.

When X is a normal random variable with mean μ and standard deviation σ, then define Z as a standard normal variable, where

$$\mathbf{Z = \frac{X - \mu}{\sigma}}$$

The calculation of a probability for a normal variable is always done the same way – transform to standard normal via $Z = \frac{X - \mu}{\sigma}$ and look up in the tables.

Question

CT3 September 2007 Q8
Claim sizes in a certain insurance situation are modelled by a normal distribution with mean μ = £30,000 and standard deviation σ = £4,000 The insurer defines a claim to be a *large claim* if the claim size exceeds £35,000.

(i) Calculate the probabilities that the size of a claim exceeds
(a) £35,000, and
(b) £36,000

(ii) Calculate the probability that the size of a *large claim* (as defined by the insurer) exceeds £36,000.

(iii) Calculate the probability that a random sample of 5 claims includes 2 which exceed £35,000 and 3 which are less than £35,000.

Solution

$X \sim N$ with mean $\mu = 30$ and $\sigma = 4$ (working in units of £1000)

(i) (a) $P(X > 35) = P\left(Z > \frac{35-30}{4}\right) = P(Z > 1.25) = 1 - 0.89435 = 0.10565$

(b) $P(X > 36) = P\left(Z > \frac{36-30}{4}\right) = P(Z > 1.5) = 1 - 0.93319 = 0.06681$

(ii) $P(X > 36 \mid X > 35) = P(X > 36 \text{ and } X > 35) / P(X > 35)$

$= P(X > 36) / P(X > 35)$

$= P(Z > 1.5)/P(Z > 1.25) = 0.06681/0.10565 = 0.632$

(iii) $\binom{5}{2} \times 0.1056^2 \times 0.8944^3 = 0.0798$

1.8 Lognormal Distribution

If X represents, for example, claim size and Y = logX has a normal distribution, then X is said to have a lognormal distribution. (log X here refers to natural log)

$$f_X(x) = \frac{1}{x\,\sigma\,\sqrt{2\pi}}\, e^{-\frac{1}{2}\left(\frac{\log x - \mu}{\sigma}\right)^2} \quad \text{for } 0 < x < \infty$$

Notice that the lower limit for *x* is 0 and not -∞, as it was for the normal distribution.
This is because log *x* is not defined for values of *x* below zero.

The lognormal distribution is positively skewed and is therefore a good model for the distribution of claim sizes.

The moments of the lognormal distribution are *not* μ and σ^2, but are given by:

$$E[X] = e^{\mu + \frac{1}{2}\sigma^2}$$

$$Var[X] = e^{2\mu + \sigma^2}\,(e^{\sigma^2} - 1)$$

1.9 t Distribution

If the variable X has a χ_v^2 distribution and another independent variable Z has the standard normal distribution of the form $N(0,1)$ then the function:

$$\frac{Z}{\sqrt{X/v}}$$

is said to have a t-distribution with parameter "degrees of freedom" v .

The *t*-distribution, like the normal, is symmetrical about 0.

To calculate probabilities for t-distribution, we will look up probabilities using page 163 in the *Tables*.

This distribution is used to find confidence intervals and carry out hypothesis tests on the mean of a distribution.

1.10 F Distribution

Another distribution that plays an important role in connection with sampling from normal populations is the F distribution, named after Sir Ronald A. Fisher, one of the most prominent statisticians of the last century.

If two independent random variables, X and Y have χ^2 distributions with parameters $n_1$and n_2 respectively, then the function:

$$\frac{X/n_1}{Y/n_2}$$

is said to have an F distribution with parameters "degrees of freedom" n_1 and $n2$.

We find probabilities by using the F-tables given on pages 170-174 of the *Tables*.

This distribution is used to find confidence intervals and carry out hypothesis tests on the variances of two distributions.