Lectures 16 - Continuous Probability Distributions

Text: WMM, Chapter 5. Sections 6.1-6.4, 6.6-6.8, 7.1-7.2

In the previous section, we introduce some of the common probability distribution functions (PDFs) for discrete sample spaces. In this chapter, we introduce some of the common PDFs for continuous sample spaces. The goal of this section is to become familiar with these probability distributions and, when given a word problem, know which PDF is appropriate.

Continuous Uniform Distribution (p. 143)

The continuous uniform distribution describes a sample space where every point within the space is equally (uniforml7) likely.

The density function of the continuous uniform random variable X on the interval [A,B] is

$$f(x;A,B) = \begin{cases} \dfrac{1}{B-A} & \text{for } A \le x \le B \\ 0 & \text{otherwise} \end{cases} \tag{16.1}$$

The mean is (from equation (7.2))

$$\mu = \frac{A+B}{2} \tag{16.2}$$

and the variance is (from equation (7.4))

$$\sigma^2 = \frac{(B-A)^2}{12} \tag{16.3}$$

Example 16.1: The default random number generators for most programming languages provides a uniform distribution on the interval [0,1]. What is the probability that a FORTRAN random number generator yields (a) X=0.25 (b) X<0.25, (c) $x \le 0.25$, (d) X>0.25, (e) 0.1<X<0.25?

(a) The probability of gettting exactly any number in a continuous sample space is zero.

(b) By equation (4.3), the definition of a contiuous PDF:

$$P(X < 0.25) = P(0 < X < 0.25) = \int_0^{0.25} f(x)dx = \int_0^{0.25} \frac{1}{1-0} dx = 0.25$$

(c)

$$P(X \le 0.25) = P(X < 0.25) + P(X = 0.25) = P(X < 0.25) = 0.25$$

(d)

$$P(X > 0.25) = 1 - P(X < 0.25) = 1 - 0.25 = 0.75$$

(e)

$$P(0.1 < X < 0.25) = \int_{0.1}^{0.25} f(x)dx = \int_{0.1}^{0.25} \frac{1}{1-0} dx = 0.15$$

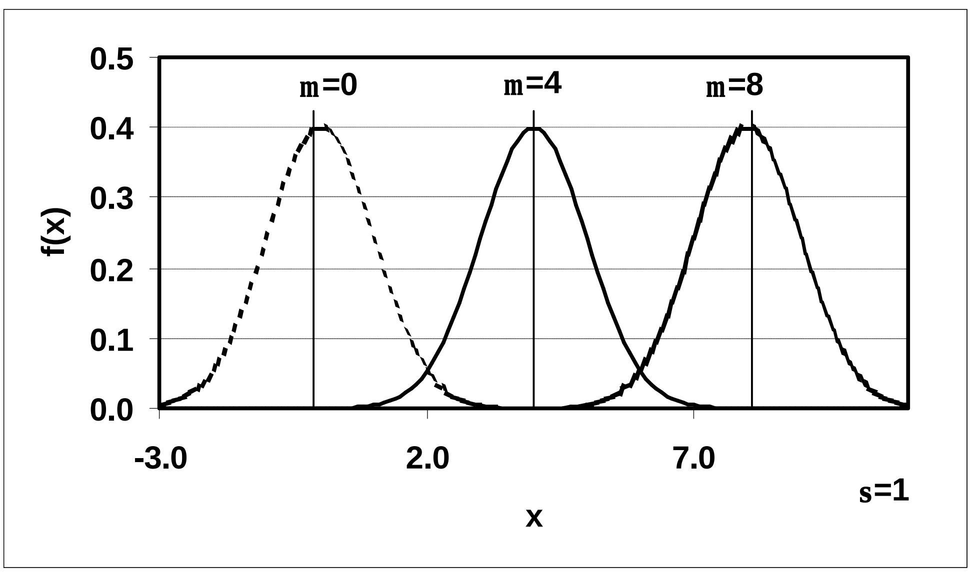
Normal Distribution (p. 145)

The normal distribution is the most important continuous PDF in statistics and in science. It is also known as the Gaussian distribution. It's shape has given rise to the nickname "the bell curve".

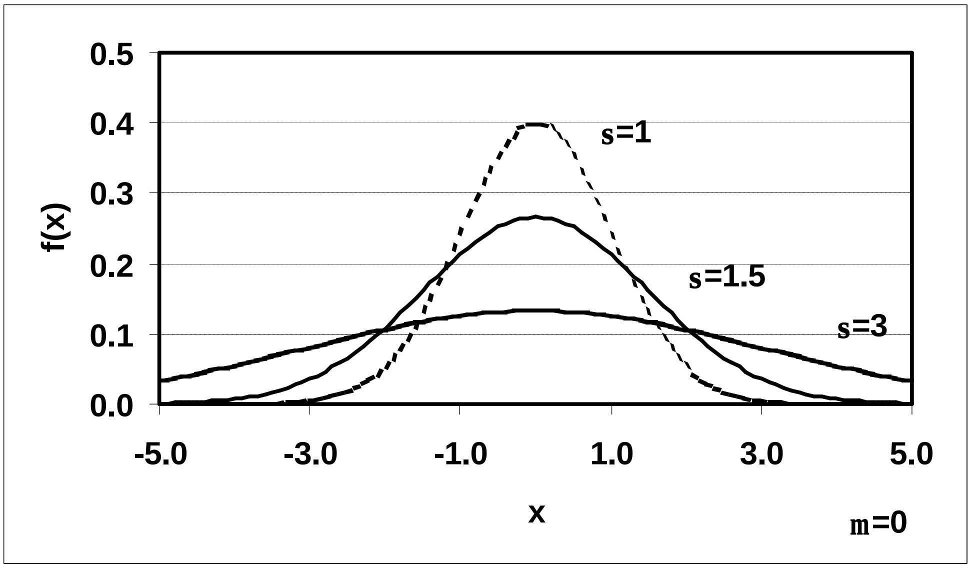
The density function of the normal random variable X, with mean μ and variance σ^2, is normal distribution

$$f(x;\mu,\sigma) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} \quad (16.4)$$

You will notice immediately that this distribution is different than the PDF's we've studied previously. The normal PDF is defined by the mean and variance, μ and σ^2. In our previous cases, the PDF defined μ and σ^2. The effect of the mean on the normal PDF is shown below.



The effect of the variance on the normal PDF is shown below



Some characteristics of the Normal PDF:

- The mode, (the most probable value occurs at the mean).
- The curve is symmetric about the mean.
- The curve has inflection point at $x = \mu \pm \sigma$. (Remember: inflection point are where the second derivative is zero, where the curve changes from concave up to concave down.
- The normal curve approaches the x-axis as we move from the mean.
- The toral area under the curve and above the x-axis is one (as it is for all PDF's.)
- The probability $P(a < X < b) = \int_a^b f(x)dx$ is the area under the normal curve between a and b.
- The Normal PDF with $\mu = 0$ and $\sigma = 1$ is called the **Standard Normal PDF**.
- Any Normal PDF, $f(x;\mu,\sigma)$,can be converted to a standard normal PDF, $f(z;0,1)$, with the change of variable $z = \frac{x-\mu}{\sigma}$
- The Normal function in equation (6.4) cannot be analytically evaluated. It can either be numerically evaluated or you can convert your normal PDF to the Standard Normal PDF and use the tabulated values in Table A.3 of WMM.
- The values in Table A.3 are listed by z-value and mean $P(-\infty < X < z)$, the probability that a value is less than z.
- The Normal PDF $f(x;\mu,\sigma)$ is equal to the binomial PDF

$b(x;n,p) = \binom{n}{x} p^x q^{n-x}$ with $\mu = np$ and $\sigma = npq$ when $n \to \infty$

Example 6.2: (example 6.7 WMM, p. 153) A type of battery lasts on the average 3.0 years with a standard deviation of 0.5 years. What is the probability that the battery will last (a) less than 2.3 years? (b) between 2.5 and 3.5 years? (c) more than five years?

(a) First step: convert the non-standard Normal PDF to the standard Normal PDF application of normal distribution

$$z_{2.3} = \frac{x-\mu}{\sigma} = \frac{2.3-3.0}{.5} = -1.4$$

$$P(X \le 2.3) = P(Z \le -1.4) = 0.0808$$

That last number comes from Table A.3 of WMM.

(b)

$$z_{2.5} = \frac{x-\mu}{\sigma} = \frac{2.5-3.0}{.5} = -1.0,\ z_{3.5} = \frac{x-\mu}{\sigma} = \frac{3.5-3.0}{.5} = 1.0$$

$$P(2.5 \le X \le 3.5) = P(-1.0 \le Z \le 1.0) = P(Z \le 1.0) - P(Z \le -1.0)$$
$$= 0.8413 - 0.1587 = 0.6826$$

(c)

$$z_{4.0} = \frac{x-\mu}{\sigma} = \frac{4.0-3.0}{.5} = 2.0$$

$$P(4.0 \ge X) = P(2.0 \ge Z) = 1 - P(2.0 \ge Z)$$
$$= 1 - 0.9772 = 0.0228$$

Gamma and Exponential Distribution (WMM, p. 166)

Two other types of PDF's commonly used in engineering are the Gamma and Exponential Functions.

The gamma function is the basis of one the Gamma Distribution. The Gamma Function has the following definition (Definition 6.2, p. 167)

$$\Gamma(\alpha) = \int_0^\infty x^{\alpha-1} e^{-x} dx \quad \text{for } \alpha > 0 \tag{16.5}$$

The Gamma function has some special values: $\Gamma(1) = 1$, $\Gamma(n) = (n-1)!$ where n is a positive integer, and $\Gamma(1/2) = \pi$.

The Gamma distribution is defined for the continuous random variable X with parameters α and β as

$$f_\Gamma(x;\alpha,\beta) = \begin{cases} \dfrac{1}{\beta^\alpha \Gamma(\alpha)} x^{\alpha-1} e^{-x/\beta} & \text{for } x > 0 \\ 0 & \text{elsewhere} \end{cases} \tag{16.6}$$

When a = 1, the Gamma distribution is called the **exponential distribution**

$$f_e(x;\beta) = \begin{cases} \dfrac{1}{\beta} e^{-x/\beta} & \text{for } x > 0 \\ 0 & \text{elsewhere} \end{cases} \tag{16.7}$$

The mean and variance of the Gamma distribution are

$$\mu = \alpha\beta \qquad \text{and} \qquad \sigma^2 = \alpha\beta^2 \tag{16.8}$$

Example 16.3:

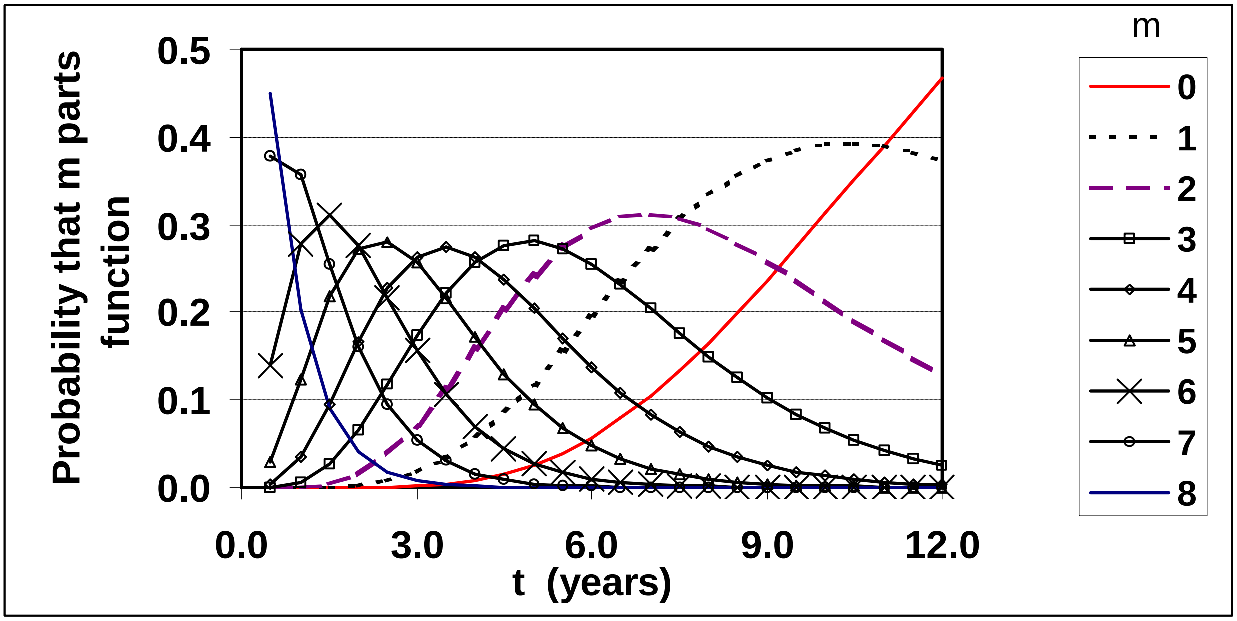
The lifetime of sparkplugs is measure in time, t, and is modelled by the exponential distribution with an average time to failure of 5 years, $\beta = 5$. If new sparkplugs are installed in an 8-cylinder engine and never replaced, what is the probability that m spark plugs are still alive at the end of t_i years where $0 \le m \le 8$ and t_i =1, 2.5, 5, 7.5, 10. The sparkplugs are independent of each other.

The probability that a single independent spark plug is still alive at the end of t_i years is given by:

$$P(t_i < t) = \int_{t_i}^{\infty} f_e(t;\beta)dt = \int_{t_i}^{\infty} \frac{1}{\beta} e^{-t/\beta} dt = e^{-t_i/\beta}$$

We can compute this for any desired value of t_i. Now, we need that probability that m of n=8 sparkplugs are still functioning at t_i, given the probability above. This is precisely the function of the binomial distribution, b(x;n,p) where x = m, the number of functioning sparkplugs, n=8, the number of total sparkplugs; and p = $P(t_i < t)$, the probability that a single sparkplug makes it to time t_i. We can calculate b(m,8, $P(t_i < t)$) for all values of m, (namely $0 \le m \le 8$) and for several values of t_i.

This is done in the graph below. Let's check that plot out. The plot shows the probability that m spark plugs function at time t. At any time t, the sum of the probabilities is 1. At time near 0, it is most probable that all 8 sparkplugs still function. At six years, it is most probable that only 2 sparkplugs still function, followed by 3, 1, 4, 0, 5, 6, 7, 8. At twelve years, it is most probable that no sparkplugs function anymore.



Chi-Squared Distribution (WMM, p. 172)

An additional special case of the Gamma Distribution is obtained when $\alpha = v/2$ and $\beta = 2$, where v is called the "degrees of freedom" and is a positive integer. The random variable X has a Chi-Squared Distribution with v degrees of freedom if

$$f_{\chi^2}(x;v) = \begin{cases} \frac{1}{2^{v/2}\Gamma(v/2)} x^{v/2-1} e^{-x/2} & \text{for } x > 0 \\ 0 & \text{elsewhere} \end{cases} \tag{16.9}$$

The mean and variance of the Chi-squared distribution are

$$\mu = v \quad \text{and} \quad \sigma^2 = 2v \tag{16.10}$$

We will encounter applications of the Chi-squared distribution in the following section. In particular, the Chi-squared distribution is important component of statistical hypothesis testing and estimation.

Functions of Random Variables

Any PDF of a variable can be transformed into a PDF of a function of that variable. In general, $f(x)$ can be transformed to g(y(x)).

Example 16.4:

When we changed the variable of a normal distribution, $f(x;\mu,\sigma)$, to the standard normal distribution, $f(z;\mu = 0,\sigma = 1)$, we used the transformation,

$$z(x) = \frac{x-\mu}{\sigma}$$

Example 16.5:
Let X be a continuous random variable with PDF

$$f(x) = \begin{cases} \frac{x}{12} & \text{for } 1 < x < 5 \\ 0 & \text{elsewhere} \end{cases}$$

then the probability, P(x<X) is given by

$$P(x > X) = \int_1^x f(x')dx' = \int_1^x \frac{x'}{12}dx' = \frac{x^2-1}{24}$$

Find the probability distribution for y, where y(x) = 2x+2.
Rearranging for x yields x=(y-2)/2 and dx=dy/2 and g(y)=f(x)=f((y-2)/2) with limits y(x=1)=4 and y(x=5)=12.

So,

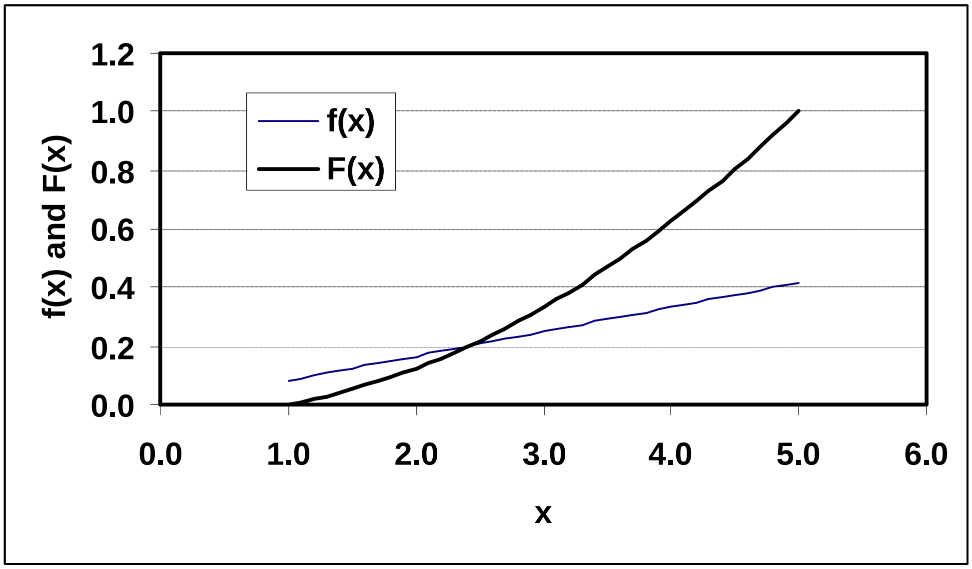
$$g(y) = f((y-2)/2) = \frac{(y-2)/2}{12} = \frac{y-2}{24}$$

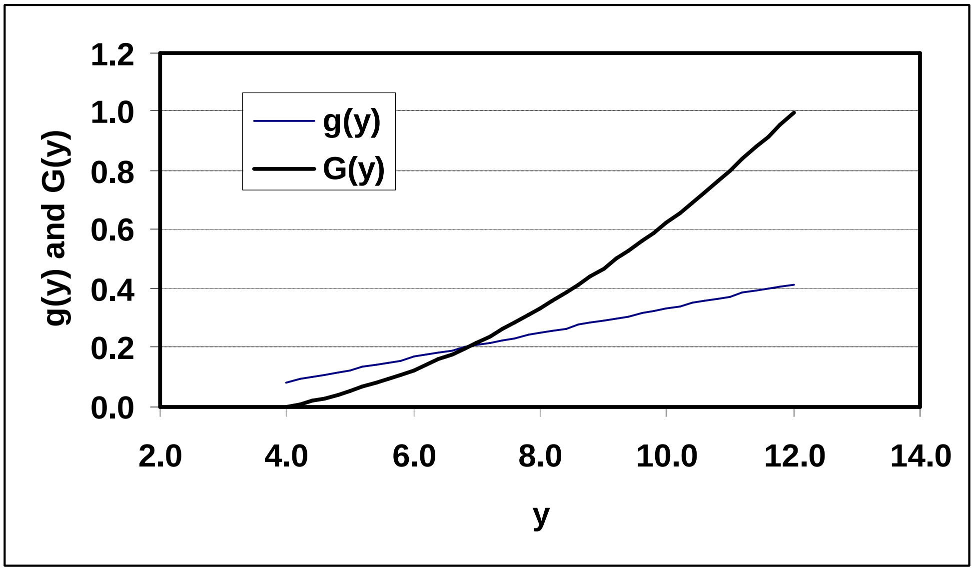
and

$$P(x > X) = P(y(x) > Y(X)) = \int_1^x f(x')dx' = \int_4^y g(y')\frac{dy'}{2}$$

$$P(y(x) > Y(X)) = \int_4^y \frac{y'-2}{24}\frac{dy'}{2} = \left(\frac{y'^2}{96} - \frac{4y'}{96}\right)\Bigg|_4^y = \frac{y^2-4y}{96}$$

Below are two plots which show f(x) and F(x)=P(x>X), the cumulative PDF, as a function of x. And g(y) and G(y)=P(y>Y), the cumulative PDF, as a function of y.


1.2
1.0
0.8
0.6
0.4
0.2
0.0
f(x) and F(x)
f(x)
F(x)
0.0
1.0
2.0
3.0
4.0
5.0
6.0
x


1.2
1.0
0.8
0.6
0.4
0.2
0.0
g(y) and G(y)
g(y)
G(y)
2.0
4.0
6.0
8.0
10.0
12.0
14.0
y