Probability Models.S3
Continuous Random Variables

A continuous random variable is one that is measured on a continuous scale. Examples are measurements of time, distance and other phenomena that can be determined with arbitrary accuracy. This section reviews the general concepts of probability density functions and presents a variety of named distributions often used to model continuous random variables[1].

Probability Distributions

Computing Probabilities

$$P(a < X < b) = \int_a^b f_X(y)\, dy \,.$$

The *probability density function* (p.d.f.) is a function, $f_X(y)$, which defines the probability density for each value of a continuous random variable. Integrating the probability density function between any two values gives the probability that the random variable falls in the range of integration.

Triangular Distribution

Parameter: $0 \le c \le 1$

$$f(x) = \begin{cases} \dfrac{2x}{c} & \text{for } 0 < x \le c \\ \dfrac{2(1-x)}{(1-c)} & \text{for } c \le x < 1 \\ 0 & \text{elsewhere} \end{cases}$$

$$\mu = \frac{1 + c}{3}$$

$$\sigma^2 = \frac{1 - 2c + 2c^2 - c^3}{18(1 - c)}$$

The meaning of p.d.f. is different for discrete and continuous random variables. For discrete variables p.d.f. is the *probability distribution function* and it provides the actual probabilities of the values associated with the discrete random variable. For continuous variables, p.d.f. is the *probability density function*, and probabilities are determined by integrating this function. The probability of any specific value of the random variable is zero.

The example for this introductory section is the triangular distribution illustrated in Fig. 6. The functional form for this density has a single parameter *c* that determines the location of the highest point, or *mode*, of the function. The random variable ranges from 0 to 1. No values of the random variable can be observed outside this range where the density function has the value 0.

[1] The Add-ins allow the specification of any of the distributions considered in this section computes probabilities and moments for each.

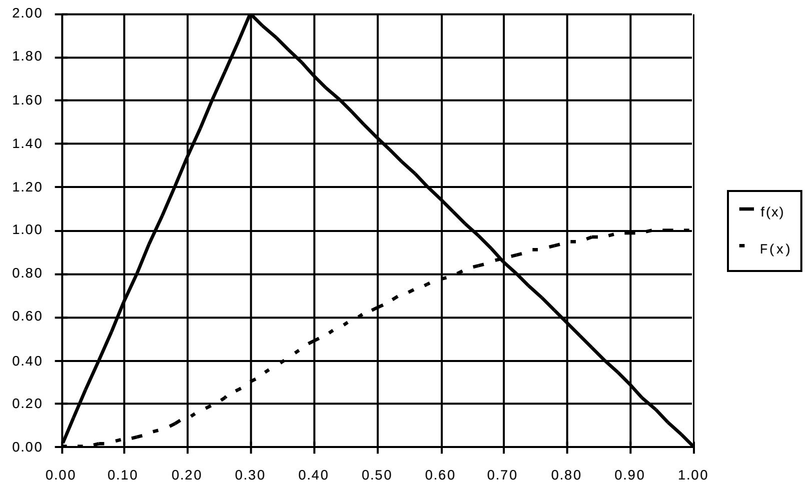


Figure 6. The Triangular Distribution with c = 0.3.

Cumulative Distribution Function

$$F_X(x) = P(X < x) = \int_{-\infty}^{x} f_X(y)\, dy .$$

$$f_X(x) = \frac{dF_X(x)}{dx} .$$

The cumulative distribution function (c.d.f.), $F_X(x)$, of the random variable is the probability that the random variable is less than the argument x. The p.d.f. is the derivative of the c.d.f. We drop the subscript on both f_X and F_X when there is no loss of clarity. The c.d.f. has the same meaning for discrete and continuous random variables, however values of the probability distribution function are summed for the discrete case while integration of the density function is required for the continuous case.

As x goes to infinity, $F_X(x)$ must go to 1. The area under an acceptable p.d.f. must, therefore, equal 1. Since negative probabilities are impossible, the p.d.f. must remain nonnegative for all x.

The c.d.f. of the Triangular Distribution

$$F(x) = \begin{cases} 0 \text{ for } x \le 0 \\ \dfrac{x^2}{c} \text{ for } 0 < x \le c \\ \dfrac{x(2-x)c}{(1-c)} \text{ for } c \le x < 1 \\ 1 \text{ for } x > 1 \end{cases} .$$

Integrating the density function for the triangular distribution results in the c.d.f. also shown in Fig. 6. The example illustrates the characteristics of every c.d.f.. It is zero for the values of x below the lower limit of the range. Within the range the function increases to the value of 1. It remains at 1 for all x values greater than the upper limit of the range. The c.d.f. never decreases and remains constant only when the p.d.f. is zero.

$$P\{a < X < b\} = \int_{a}^{b} f_X(y)\, dy = F(b) - F(a).$$

The probability that the random variable falls between two given values is the integral of the density function between these two values. It doesn't matter whether the equalities are included in the probability statements, since specific values of the random variable have zero probability.

To illustrate, consider again a random variable with the triangular distribution with *c* equal to 0.3. Since we have the closed form representation of the c.d.f., probabilities are easily determined using the c.d.f. rather than by integration. For example,

$$P\{0 < X < 0.2\} = F(0.2) - F(0) = 0.1333 - 0 = 0.1333,$$

$$P\{0.2 < X < 0.5\} = F(0.5) - F(0.2) = 0.6429 – 0.1333 = 0.5095,$$

$$P\{X > 0.5\} = 1 – P\{X < 0.5\} = 1 – F(0.5) = 1 - 0.6429 = 0.3571.$$

As for discrete distributions, descriptive measures include the mean, variance, standard deviation, skewness and kurtosis of continuous distributions. The general definitions of these quantities are given in Table 10. In addition we identify the *mode* of the distribution, that is the value of the variable for which the density function has its greatest probability, and the *median*, the value of the variable that has the c.d.f equal to 0.5.

Table 10. Descriptive Measures

Measure	**General Definition**	**Triangular Distribution (c = 0.3)**
Mean	$\mu = E\{X\} = \int_{-\infty}^{+\infty} x f(x) dx$.	$\mu = \frac{1 + c}{3} = 0.433$
Variance	$\sigma^2 = \int_{-\infty}^{\infty} (x - \mu)^2 f(x) dx$	$\sigma^2 = \frac{1 - 2c + 2c^2 - c^3}{18(1 - c)} = 0.0439$
Standard Deviation	$\sigma = \sqrt{\sigma^2}$	$\sigma = \sqrt{0.0439} = 0.2095.$
Skewness	$\beta_1 = \frac{(\mu_3)^2}{\sigma^6}$ $\mu_3 = \int_{-\infty}^{\infty} (x - \mu)^3 f(x) dx$	
Kurtosis	$\beta_2 = \frac{\mu_4}{\sigma^4}$ $\mu_4 = \int_{-\infty}^{\infty} (x - \mu)^4 f(x) dx$	

Named Continuous Distributions

Models involving random variables, require the specification of a probability distribution for each random variable. To aid in the selection, a number of named distributions have been identified. We consider several in this section that are particularly useful for modeling random variables that arise in operations research studies.

Logical considerations may suggest appropriate choices for a distribution. Obviously a time variable cannot be negative, and perhaps upper and lower limits caused by physical limitations may be identified. All of the distributions described in this section are based on logical assumptions. If one abstracts the system under study to obey the same assumptions, the appropriate distribution is apparent. For example, the queueing analyst determines that the customers of a queueing system are independently calling on the system. This is exactly the assumption that leads to the exponential distribution for time between arrivals. In another case, a variable is determined to be the sum of independent random variables with exponential distributions. This is the assumption that leads to the Gamma distribution. If the number of variables in the sum is moderately large, an appropriate distribution may be the Normal.

Very often, it is not necessary to determine the exact distribution for a study. Solutions may not be sensitive to distribution form as long as the mean and variance are approximately correct. The important requirement is to represent explicitly the variability inherent in the situation.

In every case the named distribution is specified by the mathematical statement of the probability density function. Each has one or more parameters that determine the shape and location of the distribution. Cumulative distributions may be expressed as mathematical functions; or, for cases when integration is impossible, extensive tables are available for evaluation of the c.d.f.. The moments of the distributions have already been derived, aiding the analyst in selecting one that reasonably represents his or her situation.

The Normal Distribution

Ex. 11. Consider the problem of scheduling an operating room. From past experience we have observed that the expected time required for a single operation is 2.5 hours with a standard deviation of 1 hour. We decide to schedule the time for an operation so that there is a 90% chance that the operation will be finished in the allotted time. The question is how much time to allow? We assume the time required is a random variable with a Normal distribution.

The Normal Distribution

Parameters: μ , σ

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

The probability density function of the Normal distribution has the familiar "bell shape". It has two parameters μ and σ, the mean and standard deviation of the distribution. The Normal distribution has applications in many practical contexts. It is often used, with theoretical justification, as the experimental variability associated with physical measurements. Many other distributions can be approximated by the Normal distribution using suitable parameters. This distribution reaches its maximum value at μ, and it is symmetric about the mean, so the mode and the median also have the value μ.. Fig. 7 shows the distribution associated with the example.

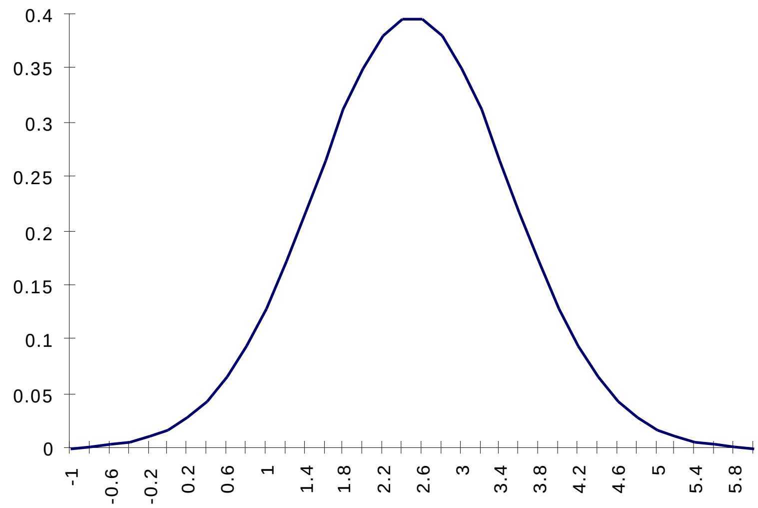


Figure 7. The Normal distribution

It is impossible to symbolically integrate the density function of a Normal distribution, so there is no closed form representation of the cumulative distribution function. Probabilities are computed using tables that appear in many text books or using numerical methods implemented with computer programs. Certain well known probabilities associated with the Normal distribution are in Table 11.

Table 11. Ranges of the Normal Distribution

within one standard deviation of the mean	$P(\mu-\sigma < X < \mu+\sigma)$	0.6826
within two standard deviations of the mean	$P(\mu-2\sigma < X < \mu+2\sigma)$	0.9545
within three standard deviations of the mean	$P(\mu-3\sigma < X < \mu+3\sigma)$	0.9973

Considering the example given at the beginning of this section, we note that the assumption of Normality cannot be entirely appropriate, because there is a fairly large probability that the time is less than zero. We compute

$$P(X < 0) = 0.0062$$

Perhaps this is a small enough value to be neglected.

The original requirement that we reserve the operating room such that the reserved time will be sufficient 90% of the time, asks us to find a value z such that

$$P(X < z) = 0.9.$$

For this purpose we identify the *inverse probability function* $F^{-1}(p)$ defined as

$$z = F^{-1}(p) \text{ such that } F(z) = P(x < z) = p.$$

For the Normal distribution, we compute these values numerically using a computer program. For the example

$$z = F^{-1}(0.9) = 3.782 \text{ hours}.$$

Sums of Independent Random Variables

Consider again the operating room problem. Now we will assign the operating room to a single doctor for three consecutive operations. Each operation has a mean time of 2.5 hours and a standard deviation of 1 hour. The doctor is assigned the operating room for an entire 8 hour day. What is the probability that the 8 hours will be sufficient to complete the three operations?

There are important results concerning the sum of independent random variables. Consider the sum of n independent random variables, each with mean μ and standard deviation σ,

$$Y = X_1 + X_2 + \ldots + X_n.$$

Four results will be useful for many situations.

1. Y has mean and variance

   $$\mu_Y = n\mu \text{ and } \sigma_Y^2 = n\sigma^2.$$

2. If the X_i individually have Normal distributions, Y will also have a Normal distribution.
3. If the X_i have identical but not Normal distributions, probabilities associated with Y can be computed using a Normal distribution with acceptable approximation as n becomes large.
4. The mean value of the n random variables is called the sample mean $M = Y/n$. The mean and variance of the sample mean is

   $$\mu_M = \mu \text{ and } \sigma_M^2 = \sigma^2/n.$$

   If the distribution of the X_i are Normal, the distribution of the sample mean is also Normal. If the distribution of the X_i are not Normal, the distribution of the sample mean approaches the Normal as n becomes large. This is called the *central limit theorem.*

Results 1 and 2 are true regardless of the distribution of the individual X_i. When the X_i are not normally distributed, the accuracy of results 3 and 4 depend on the size of *n*, and the shape of the distribution.

For the example, let X_1, X_2 and X_3 be the times required for the three operations. Since the operations are done sequentially, the total time the room is in use is, *Y*, where

$$Y = X_1 + X_2 + X_3.$$

Since we approximated the time for the individual operation as Normal, then *Y* has a Normal distribution with

$$\mu_Y = 3\mu = 7.5, \sigma_Y^2 = 3\sigma^2 = 3 \text{ and } \sigma_Y = \sqrt{3}\,\sigma = 1.732.$$

The required probability is

$$P\{Y < 8\} = 0.6136.$$

There is a 61% chance that the three operations will be complete in eight hours.

Lognormal Distribution

Ex. 12. Consider the material output from a rock crushing machine. Measurements have determined that the particles produced have a mean size of 2" with a standard deviation of 1". We plan to use a screen with 1" holes to filter out all particles smaller than 1". After shaking the screen repeatedly, what proportion of the particles will be passed through the screen? For analysis purposes we assume, the size of particles has a Lognormal distribution.

Lognormal Distribution

Parameters: α, β

$$f(x) = \frac{1}{\beta\sqrt{2\pi}}\, exp\{\frac{(\ln(x)-\alpha)^2}{2\beta^2}\} \text{ for } x > 0.$$

$$\mu = exp(\alpha + \beta^2/2)$$

$$\sigma^2 = exp[2\alpha + \beta^2][exp(\beta^2) - 1]$$

$$x_{\text{mode}} = exp[\alpha - \beta].$$

The Lognormal distribution exhibits a variety of shapes as illustrated in Fig. 8. The random variable is restricted to positive values. Depending on the parameters, the distribution rapidly rises to its mode, and then declines slowly to become asymptotic to zero.

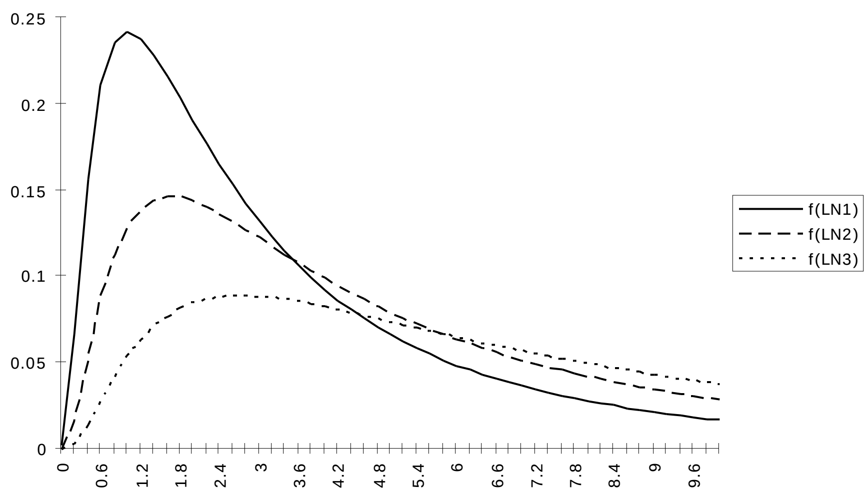


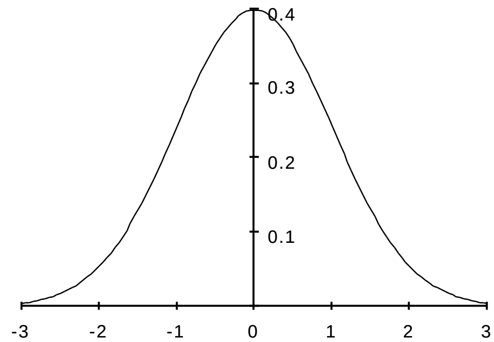
Figure 8. Lognormal Distribution

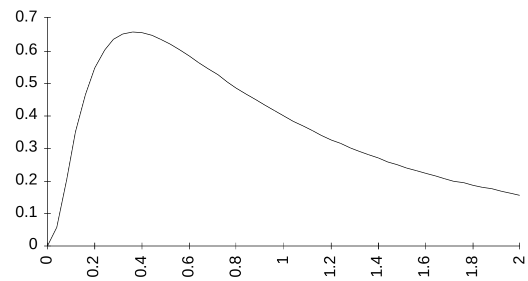
The Lognormal and the Normal distributions are closely related as shown in Fig. 9. When some random variable X has a Lognormal distribution, the variable Z has a Normal distribution when

$$Z = ln(X).$$

Alternatively, when the random variable Z has a Normal distribution, the random variable X has a Lognormal distribution when

$$X = exp(Z).$$

<table>
<tr><td>Random Variable Z
Normal distribution
Mean μ_Z and variance σ_Z^2</td><td>Random Variable X
Lognormal distribution
Mean μ_X and variance σ_X^2</td></tr>
<tr><td>


</td><td>


</td></tr>
<tr><td>$\mu_Z = 0$ and $\sigma_Z^2 = 1$ for the case shown.</td><td>$\mu_X = 1.6487$ and $\sigma_X^2 = 4.6708$.</td></tr>
<tr><td>$Z = ln(X)$</td><td>$X = exp(Z)$</td></tr>
</table>

Figure 9. Relation between the Normal and Lognormal Distribution

The Lognormal has two parameters that are the moments of the related Normal distribution.

$$\alpha = \mu_Z \text{ and } \beta = \sigma_Z.$$

The formula for the p.d.f. for the Lognormal distribution is given in terms of α and β. This expression is useful for plotting the p.d.f., however, it is not helpful for computing probabilities since it cannot be integrated in the closed form. The parameters of the three cases of Fig. 8 are shown in the Table 12.

Table 12. Distribution Parameters for Fig. 8

Case	$\alpha = \mu_Z$	$\beta = \sigma_Z$	μ_X	σ_X	x_{mode}
C1	1	1	4.48	5.87	1
C2	1.5	1	7.39	9.69	1.65
C3	2	1	12.18	15.97	2.72

For some cases, one might be given the mean and variance of the random variable *X* and would like to find the corresponding parameters of the distribution. Solving for α and β in terms of the parameters of the distribution underlying Normal distribution

$$\beta^2 = ln[(\sigma_X^2/\mu_X^2) + 1], \text{ and } \alpha = ln[\mu_X] - \beta^2/2.$$

The Lognormal distribution is a flexible distribution for modeling random variables that can assume only nonnegative values. Any positive mean and variance can be obtained by selecting appropriate parameters.

For the example given at the beginning of this section, the given data provides estimates of the parameters of the distribution of *X*,

$$\mu_X = 2 \text{ and } \sigma_X^2 = 1.$$

From this information, we compute the values of the parameters of the Lognormal distribution.

$$\beta^2 = 0.2231 \text{ or } \beta = 0.4723, \text{ and } \alpha = 0.6683.$$

$$P\{X < 1\} = 0.0786.$$

Approximately 8% of the particles will pass through the screen.

Exponential Distribution

Ex. 13. Telephone calls arrive at a switch board with an average time between arrivals of 30 seconds. Since the callers are independent, calls arrive at random. What is the probability that the time between one call and the next is greater than one minute?

Exponential Distribution
Parameter: λ

$$f(x) = \lambda \, exp(-\lambda x) \text{ for } x \geq 0.$$

$$F(x) = 1 - exp(-\lambda x) \text{ for } x \geq 0.$$

$$\mu = \sigma = 1/\lambda \text{ and } \beta_1 = 4.$$

The exponential distribution is often used to model situations involving the random variable of time between arrivals. When we say that the average arrival rate is λ, but the arrivals occur independently, then the time between arrivals has an exponential distribution. The general form has the single positive parameter λ. The density has the shape shown in Fig. 10.

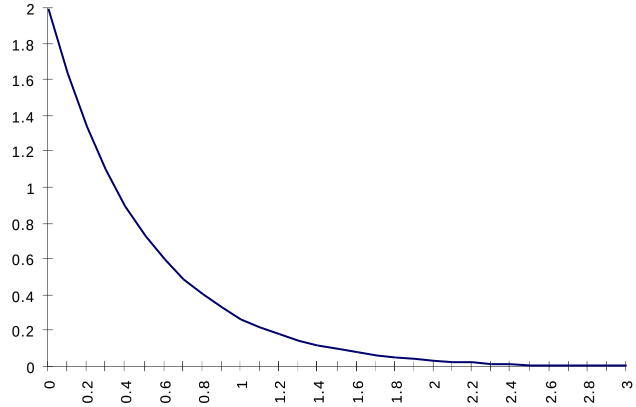


Figure 10. The Exponential Distribution

This density function is integrated to yield the closed form expression for the c.d.f. The distribution has equal mean and standard deviation, and the skewness measure is always 4. Because it is skewed to the right, the mean is always greater than the median. The probability that the random variable is less than its mean is independent of the value of λ: $P(x < \mu) = 0.6321$.

For the example problem, we are given that the mean of the time between arrivals is 30 seconds. The arrival rate is therefore 2 per minute. The required probability is then

$$P(x > 1) = 1 - F(1) = exp(-2) = 0.1353.$$

The exponential distribution is the only memoryless distribution. If an event with an exponential distribution with parameter λ has not occurred prior to some time T, the distribution of the time until it does occur has an exponential distribution with the same parameter.

The exponential distribution is intimately linked with the discrete Poisson distribution. When some process has the time between arrivals governed by the exponential distribution with rate λ, the number of arrivals in some fixed interval T is governed by the Poisson distribution with mean value equal to λT. Such a process is called a Poisson process.

Gamma Distribution

Ex. 14. Consider a simple manufacturing operation whose completion time has a mean equal to ten minutes. We ask, what is the probability that the completion time will exceed 11 minutes? We first investigate this question assuming the time has an exponential distribution (a special case of the Gamma), then we divide the operation into several parts, each with an exponential distribution.

Gamma Distribution

Parameters: $\lambda > 0$ and $r > 0$

$$f(x) = \frac{\lambda}{\Gamma(r)}(\lambda x)^{r-1} exp(-\lambda x) \quad \text{for } x \geq 0.$$

$\Gamma(r)$ is the Gamma function.

$$\Gamma(r) = \int_0^\infty x^{r-1} e^{-x}\, dx .$$

When r is an integer,

$$\Gamma(r) = (r-1)!$$

$$\mu = \frac{r}{\lambda},\ \sigma^2 = \frac{r}{\lambda^2},\ x_{\text{mode}} = \frac{r-1}{\lambda} .$$

The Gamma distribution models a random variable that is restricted to nonnegative values. The general form has two positive parameters r and λ determining the density function. We restrict attention to integer values of r although the Gamma distribution is defined for noninteger values as well.

Fig. 11 shows several Gamma distributions for different parameter values. The distribution allows only positive values and is skewed to the right. There is no upper limit on the value of the random variable. The parameter r has the greatest affect on the shape of the distribution. With r equal to 1, the distribution is the exponential distribution. As r increases, the mode moves away from the origin, and the distribution becomes more peaked and symmetrical. As *r* increases in the limit, the distribution approaches the Normal distribution.

The Gamma distribution is used extensively to model the time required to perform some operation. The parameter λ primarily affects the location of the distribution. The special case of the exponential distribution is important in queueing theory where it represents the time between entirely random arrivals. When *r* assumes integer values, the distribution is often called the Erlang distribution. This is the distribution of the sum of *r* exponentially distributed random variables each with the mean $1/\lambda$. All the distributions in Fig. 11 are Erlang distributions. This distribution is often used to model a service operation comprised of a series of individual steps. When the time for each step has an exponential distribution, the average time for all steps are equal, and the step times are independent random variables, the total time for the operation has an Erlang distribution.

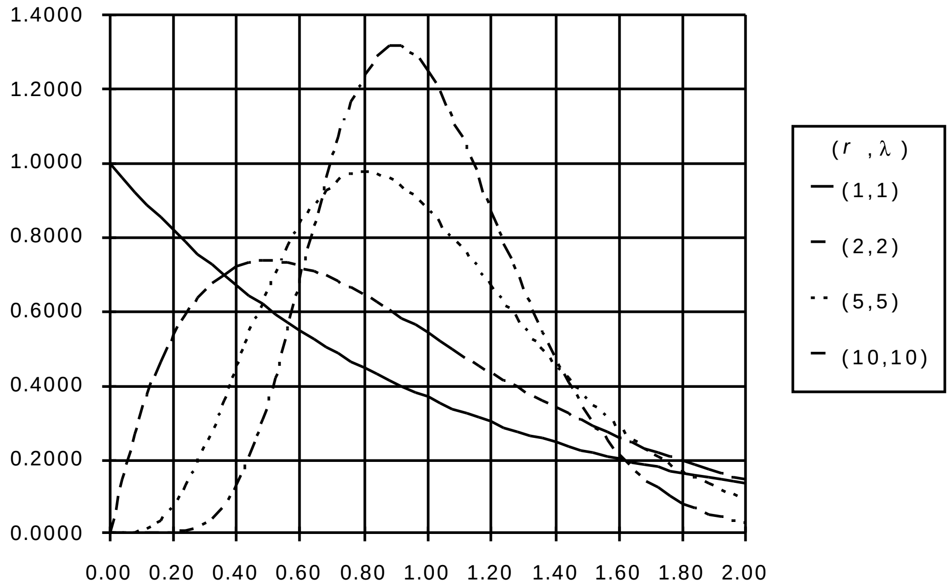


Figure. 11. The Gamma Distribution

Cumulative Distribution of the Gamma

$$P(X \le x) = F(x) = 1 - \sum_{k=0}^{r-1} \frac{(\lambda x)^k \exp(-\lambda x)}{k!} \text{ for } x \ge 0.$$

With integer *r*, a closed form expression for the cumulative distribution of the Gamma allows easy computation of probabilities. The summation expression is the cumulative distribution of the discrete Poisson distribution with parameter λx.

In the example problem for this section, we have assumed that the mean processing time for a manufacturing operation is 10 minutes and ask for the probability that the time is less than 11 minutes. First assume that the time has an exponential distribution. The exponential distribution is a special case of the

Gamma distribution with r equal to 1. The given mean, 10, provides us with the information necessary to compute the parameter λ.

$$\mu = \frac{1}{\lambda} \text{ or } \lambda = \frac{1}{\mu} = 0.1.$$

We compute ($r = 1$, $\lambda = 0.1$): $\mu = 10$, $\sigma = 10$, $P(X < 11) = 0.6671$

Now assume that we can divide the operation into two parts such that each part has a mean time for completion of 5 minutes and the parts are done in sequence. In this situation, the total completion time has a Gamma (or Erlang) distribution with $r = 2$. The parameter λ is computed from the mean completion time of each part, $\mu = 5$.

$$\lambda = \frac{1}{\mu} = 0.2.$$

We compute the results from the Gamma distribution.

For ($r = 2$, $\lambda = 0.2$): $\mu = 10$, $\sigma = 7.071$, $P(X < 11) = 0.6454$.

Continuing in this fashion for different values of r with the parameters such that $\lambda = r/10$.

For ($r = 10$, $\lambda = 1$): $\mu = 10$, $\sigma = 3.162$, $P(X < 11) = 0.6594$.

For ($r = 20$, $\lambda = 2$): $\mu = 10$, $\sigma = 2.236$, $P(X < 11) = 0.6940$.

For ($r = 40$, $\lambda = 4$): $\mu = 10$, $\sigma = 1.581$, $P(X < 11) = 0.7469$.

As r increases, the distribution has less variance and becomes less skewed.

One might ask whether the Normal distribution is a suitable approximation as r assumes higher values. We have computed the probabilities for $r = 20$ and $r = 40$ using a Normal approximation with the same mean and variance. The results are below.

For X Normal with ($\mu = 10$, $\sigma = 2.236$): $P(X < 11) = 0.6726$.

For X Normal with ($\mu = 10$, $\sigma = 1.581$): $P(X < 11) = 0.7364$.

It is apparent that the approximation becomes more accurate as r increases.

Beta Distribution

Ex. 15. Consider the following hypothetical situation. Grade data indicates that on the average 27% of the students in senior engineering classes have received A grades. There is variation among classes, however, and the proportion must be considered a random variable. From past data we have measured a standard deviation of 15%. We would like to model the proportion of A grades with a Beta distribution.

Beta Distribution

Parameters: $\alpha > 0$ and $\beta > 0$.

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

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

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

$$x_{\text{mode}} = \frac{\alpha - 1}{\alpha + \beta - 2}$$

The Beta distribution is important because it has a finite range, from 0 to 1, making it useful for modeling phenomena that cannot be above or below given values. The distribution has two parameters, α and β, that determine its shape. When α and β are equal, the distribution is symmetric. Increasing the values of the parameters decreases the variance. The symmetric case is illustrated in Fig. 12.

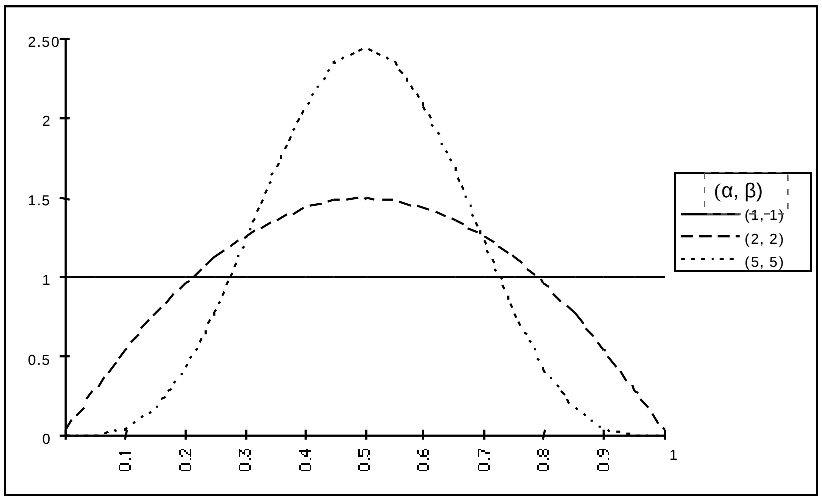


Figure 12. The Symmetric Case for the Beta Distribution.

When α is less than β the distribution is skewed to the right as shown in Fig. 13. The distribution function is symmetric with respect to α and β, so when α is greater than β, the distribution is skewed to the left.

The Uniform distribution is the special case of the Beta distribution with α and β both equal to 1. The density function has the constant value of 1 over the 0 to 1 range. When $\alpha = 1$ and $\beta = 2$, the resulting Beta distribution is a triangular distribution with the mode at 0.

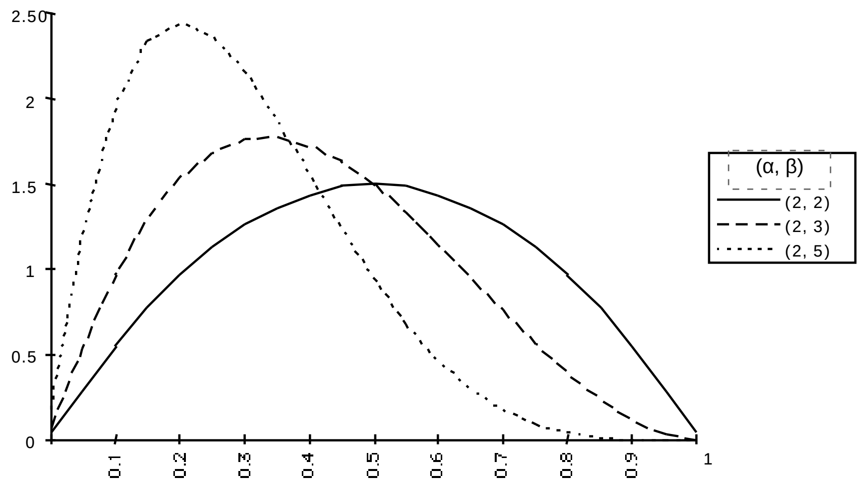


Figure 13. The Beta Distribution as β increases for constant α.

A linear transformation of the Beta variable provides a random variable with an arbitrary range. The distribution is often used when an expert provides a lower bound, *a*, upper bound, *b*, and most likely value, *m*, for the time to accomplish a task. A transformed Beta variable could be used to represent the task time in a model.

Generalized Beta Distribution

Parameters: $\alpha > 0$, $\beta > 0$, a, b

$a < b$

$$\mu = a + (b-a)\left(\frac{\alpha}{\alpha+\beta}\right)$$

$$\sigma^2 = \frac{(b-a)^2\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}$$

$$mode = a + (b-a)\left\{\frac{\alpha-1}{\alpha+\beta-2}\right\}$$

With a linear transformation we change the domain of the Beta distribution. Assume *X* has the Beta distribution, and let

$$Y = a + (b-a)X.$$

The transformed distribution called, the generalized Beta, has the range *a* to *b*. The mean and mode for the distribution are shifted accordingly. The mode of *Y* is the most likely value, *m*, and relates to the mode of *X* as

$$x_{\text{mode}} = \frac{\alpha-1}{\alpha+\beta-2} = \frac{m-a}{b-a}.$$

Given values of *a*, *b* and *m*, one obtains a relation between α and β. Selecting a value for one parameter determines the other. Probabilities are calculated from the Beta distribution.

Uniform Distribution
Parameters: $a < b$
$f(x) = \dfrac{1}{b-a}$ for $a \le x \le b$
$\mu = \dfrac{a+b}{2}$
$\sigma^2 = \dfrac{(b-a)^2}{12}$

$$P\{Y \le y\} = P\{X \le \frac{x-a}{b-a}\}.$$

The generalized Uniform distribution has an arbitrary range from *a* to *b*. Its moments are determined by specifying α and β equal to 1.

We use the Beta distribution to model the proportions of the example problem, since they are restricted to values between zero and 1. The example gives the average proportion of A's as 0.27. The standard deviation is measured as 0.15. To use this data as estimates of μ and σ, we first solve for β as a function of μ and α.

$$\beta = \frac{\alpha(1-\mu)}{\mu}.$$

Substituting $\mu = 0.27$ into this expression, we obtain: $\beta = 2.7037\alpha$.

There are a variety of combinations of the parameters that yield the required mean. We use the given standard deviation to make the selection. Table 13 shows the value of β and the standard deviation for several integer values of α.

Table 13. Standard Deviation of the Beta Distribution

α	1	2	3	4	5	6	7	8	9	10
β	2.704	5.41	8.11	10.81	13.52	16.22	18.93	21.63	24.33	27.04
σ	0.204	0.15	0.13	0.11	0.10	0.09	0.09	0.08	0.08	0.07

The standard deviation associated with $\alpha = 2$ and $\beta = 5.41$ approximates the data. With the selected model, we now ask for the probability that a class has more than 50% A grades. A spreadsheet function provides the values of the c.d.f. for the given parameters $(\alpha, \beta) = (2, 5.4074)$. We find: $P\{X > 0.5\} = 0.0873$.

Weibull Distribution

Ex. 16. A truck tire has a mean life of 20,000 miles. A conservative owner decides to replace the tire at 15,000 miles rather than risk failure. What is the probability that the tire will fail before it is replaced? Assume the life of the tire has a Weibull distribution with β equal to 2.

Weibull Distribution

Parameters: $\alpha > 0$, $\beta > 0$

$$f(x) = \alpha\beta x^{\beta-1} exp\{-\alpha x^{\beta}\} \text{ for } x \geq 0.$$

$$F(x) = 1 - exp\{-\alpha x^{\beta}\} \text{ for } x \geq 0.$$

$$\mu = \alpha^{-1/\beta}\,\Gamma(1 + \frac{1}{\beta}).$$

$$\sigma^2 = \alpha^{-2/\beta}\,\{\Gamma(1 + \frac{2}{\beta}) - [\Gamma(1 + \frac{1}{\beta})]^2\}.$$

$$x_{\text{mode}} = \sqrt[\beta]{\frac{\beta - 1}{\alpha\beta}}.$$

This distribution has special meaning to reliability experts, however, it can be used to model other phenomena as well. As illustrated in Fig. 14, the distribution is defined only for nonnegative variables and is skewed to the right. It has two parameters α and β.

The parameter β affects the form of the distribution, while for a given β, the parameter α affects the location of the distribution. The cases of Fig. 14 have α adjusted so that each distribution has the mean of 1. When β is 1, the distribution is an exponential. As β increases and the mean is held constant, the variance decreases and the distribution becomes more symmetric.

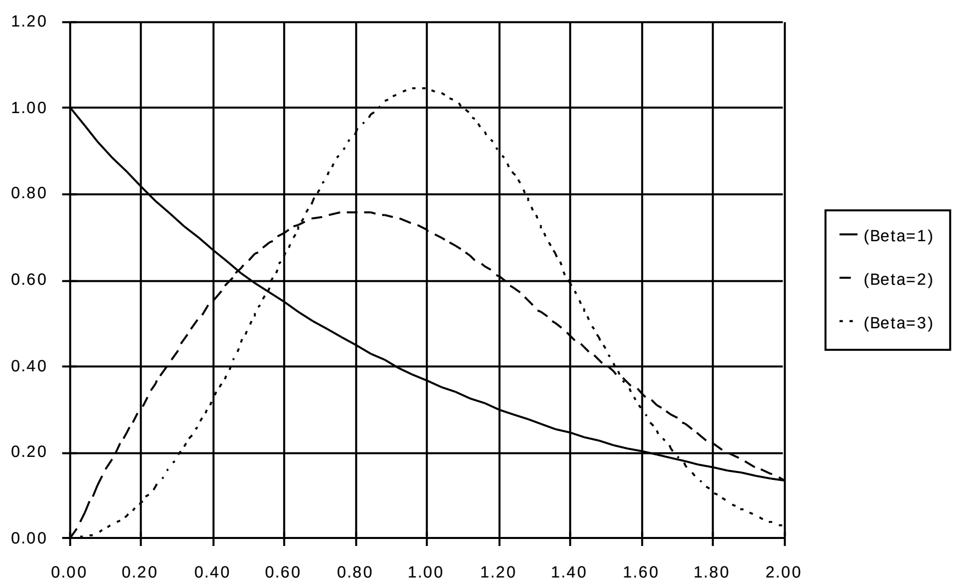


Figure 14. The Weibull Distribution

Conveniently, the cumulative distribution has a closed form expression.

In reliability modeling, the random variable x is the time to failure of a component. We define the hazard function as

$$\lambda(x) = \frac{f(x)}{1 - F(x)}.$$

For a small time interval Δ, the quantity $\lambda(x)\Delta$ is the probability that the component will fail in the time interval $\{x, x + \Delta\}$, given it did not fail prior to time x. The hazard function can be viewed as the *failure rate* as a function of time. For the Weibull distribution

$$\lambda(x) = \alpha\beta x^{\beta-1}.$$

For β equal to 1, the hazard function is constant, and we say that the component has a constant failure rate. The distribution for time to failure is the exponential distribution. This is often the assumption used for electronic components.

For β equal to 2, the hazard rate is increasing linearly with time. The probability of failure in a small interval of time, given the component has not failed previously, is growing with time. This is the characteristic of wear out. As the component gets older, it begins to wear out and the likelihood of failure increases. For larger values of β the hazard function increases at a greater than linear rate, indicating accelerating wear out.

Alternatively for β less than 1, the hazard function is decreasing with time. This models the characteristic of infant mortality. The component has a high failure rate during its early life; but when it survives that period, it becomes less likely to fail.

For the example, we assume the truck tire has a mean life of 20,000 miles, and we assume the life has a Weibull distribution with β equal to 2. The owner decides to replace the tire at 15,000 miles rather than risk failure. We ask for the probability that the tire will fail before it is replaced.

To answer this question, we see that the mean of the distribution is given as 20 (measured in thousands of miles). Since β is also specified, we must compute α for the distribution. The expression for the mean is solved for the parameter α to obtain

$$\alpha = \left[\frac{\Gamma(1+\frac{1}{\beta})}{\mu}\right]^{\beta}.$$

Evaluating this expression for the given information we obtain

$$\alpha = \left[\frac{\Gamma(1.5)}{20}\right]^2 = 0.0019635.$$

Now to compute the required probability with these parameters

$$P\{x < 15\} = F\{15\} = F(x) = 1 - exp\{-\alpha(15)^{\beta}\} = 0.3571.$$

This result looks a little risky for our conservative driver. He asks, how soon must the tire be discarded so the chance of failure is less than 0.1?

To answer this question, we must solve for the value of x that yields a particular value of $F(x)$ ($z = F^{-1}(0.1)$) . Using the c.d.f. this is easily accomplished.

$$z = \sqrt[\beta]{\frac{-ln[1 - F(x)]}{\alpha}} = \sqrt[2]{\frac{-ln[0.9]}{0.001964}} = 7.325. \qquad (21)$$

To get this reliability, the owner must discard his tires in a little more than 7000 miles.

Translations of the Random Variable

Ex. 17. An investment in an income producing property is \$1,000,000. In return for that investment we expect a revenue for each year of a 5 year period. The revenue is uncertain however. We estimate that the annual revenues are independent random variables, each with a Normal distribution having a mean of \$400,000 and a standard deviation of \$250,000. Negative values of the random variable indicate a loss. We evaluate the investment by computing the net present worth of the cash flows using a minimum acceptable rate of return (MARR) of 10%.

Say the amount of the income in year t is $R(t)$ and the MARR is i. Assuming discrete compounding and end of the year payments, the present worth for the cash flow at year t is

$$\text{PW}(t) = \frac{R(t)}{(1+i)^t}.$$

With the initial investment in the amount I the Net Present Worth for a property that lasts n years is

$$\text{NPW} = -I + \sum_{t=1}^{n} \frac{R(t)}{(1+i)^t}.$$

A NPW greater than 0 indicates that the rate of return of the investment in the property is greater than the MARR and it is an acceptable investment. The problem here is that the revenues are independent random variables, so the NPW is also a random variable. The best we can do is compute the probability that the NPW will be greater than 0. Although the annual revenues all have the same distributions they are linearly translated by the factor $1/(1 + i)^t$, a factor that is different for each year.

We use the following general results to deal with this situation. Say we have the probability distribution of the random variable X described by $f_X(x)$ and $F_X(x)$. We know the associated values of the mean (μ_X), variance (σ_X^2) and the mode (x_{mode}) of the distribution. We are interested, however, in a translated variable

$$Y = a + bX,$$

where a is any real number and b is positive. We want to compute probabilities about the new random variable and state its mean, variance and mode.

Probabilities concerning Y are easily computed from the cumulative distribution of X.

$$F_Y(y) = P\{Y \le y\} = P\{X \le \frac{y-a}{b}\} = F_X(\frac{y-a}{b}).$$

The mean, variance and mode of Y are respectively

$$\mu_Y = E\{Y\} = E\{a + bX\} = a + b\mu_X.$$

$$\sigma_Y^2 = E\{(Y - \mu_Y)^2\} = b^2\sigma_X^2 \text{ or } \sigma_Y = b\sigma_X.$$

$$y_{\text{mode}} = a + bx_{\text{mode}}$$

To use these results for the example problem, we note that the present worth factors are a linear translation with

$$a = 0 \text{ and } b = 1/(1 + i)^t.$$

Thus the contribution of year t to the net present worth is a random variable with: $\mu = 400/(1 + i)^t$ and $\sigma = 250/(1 + i)^t$. We use three facts to continue: the sum of independent random variables with Normal distributions is also Normal, the mean of the sum is the sum of the means, and the variance of the sum is the sum of the variances. We conclude that the NPW has a Normal distribution with

$$\mu = -1000 + \sum_{t=1}^{5} 400/(1+i)^t = 516.3$$

$$\text{and } \sigma = 250\sqrt{\sum_{t=1}^{5}\left(\frac{1}{(1+i)^t}\right)^2} = 427.6.$$

Based on these parameters and using $i = 10\%$: $P(\text{NPW} < 0) = 11.4\%$. It is up to the decision maker to decide whether the risk is acceptable.

Modeling

To model a continuous random variable for analytical or simulation studies, some distribution must be selected and parameters determined. Logical considerations and statistical analysis of similar systems can guide the selection.

When independent random variables are summed, the Normal distribution is the choice when the individuals have Normal distributions, and the Erlang distribution (a special case of the Gamma) is the choice when the individuals have exponential distributions. When a large number of random variables are summed that are of approximately the same magnitude, the Normal distribution is often a good approximation.

A disadvantage of the Normal distribution is that it allows negative values. Where the mean is not much larger than the standard deviation this might be a problem when the random variable is obviously nonnegative. Of course, time is a prime example where negativity is a logical impossibility. Here we have the Lognormal, Gamma and Weibull distributions that restrict the random variable to positive values. These distributions also exhibit the asymmetry that might be present and important in practice. The Weibull distribution is important to reliability theory when failure rate information is available.

When the physical situation suggests both upper and lower bounds, the uniform, triangular, and generalized Beta distributions are available. The uniform distribution has only the bounds as parameters, and might be used where there is no other information. The triangular distribution has a third parameter, the mode, and is used for cases when one estimates the most likely value for the random variable in addition to its range. The generalized Beta is very flexible with are large variety of shapes possible with a suitable selection of parameters.