Probability Models

To describe of the outcome of an uncertain event, we often speak of the probability of its occurrence. Meteorologists tell us the probability of rain, electrical engineers predict the reliability (or probability of success) of a computer system, quality control technicians measure the probability of a defect in a manufacturing process, gamblers estimate their chances (or probability) of winning, doctors tell their patients the risk (or probability of failure) of a medical procedure, and economists try to forecast the likelihood of a recession. For many, risk and uncertainty are unpleasant aspects of daily life; while for others, they are the essence of adventure. To the analyst it is the inescapable result of almost every activity.

In the first part of the book, we mostly neglect the effects of uncertainty and risk and assume that the results of our decisions are predictable and deterministic. This abstraction allows large and complex problems to be modeled and solved using the powerful methods of mathematical programming. In the remainder of the book, we examine models that explicitly include some aspects of uncertainty, and describe analytic techniques for dealing with decision problems in this environment.

In this chapter, we establish the language of probability that is used to model situations whose outcomes cannot be predicted with certainty. In important instances, reasonable assumptions allow single random variables to be described using one of the named discrete or continuous random variables. The chapter summarizes the formulas for obtaining probabilities and moments for the distributions. When combinations of several random variables affect the outcome, simulation is often used to learn about a system. Here we provide an introduction to the subject.

21.1 Variability, Uncertainty and Complexity

If we observe the price of a security on the stock market from hour to hour, the time it takes a worker to perform some operation, the number of calories a person consumes at breakfast each day, the number of telephone calls arriving at a receptionist's desk, or the interest rate on government bonds we see that these quantities are variable. They change over time. Moreover, the change is uncertain or random. In other cases, the quantity of interest may change with time but in a known or predictable way. This would be true for a printed circuit board assembly shop where the output was a function of the product being assembly. In general, variability complicates the problems of planning and decision making, and often reduces the effectiveness of the systems we design.

Variability is the source of many of our difficulties so decision making would be much simpler if exact values of the variable quantities could be predicted with some degree of accuracy. Unfortunately, variability in the future is almost always associated with uncertainty. We are backward looking beings. We can measure the variability of the past by taking data and analyzing it using the methods of statistics. Statistics tell us what *did* happen. As designers and planners, operations research professionals are concerned about answering the questions: what *will* happen and what *should* be done about it? In cases where the dual problems of variability and uncertainty are present, the deterministic methods of the first part of this book are usually insufficient.

Even if there is both variability and uncertainty, we can sometimes deal with a problem if it is "simple." Most often, however, the real world is characterized by complexity. For example, the production planner is faced by multiple sources of variability and uncertainty. Not only is future demand unknown, but also the availability and lead times for raw materials, the production times for the sequential stations of the manufacturing process, the possibility of machine failures, the availability of the work force, and many other factors relating to marketing and sales. In a complex system, variable quantities often interact in unexpected ways. In the face of such randomness, it is surprising that those charged with decision making are at all successful.

This section begins the study of the characteristics of variability, uncertainty and complexity. We will find that the models available are much more limited than those we have described for deterministic systems. While in many cases deterministic models can be mathematically solved to obtain optimal solutions for decision variables, this is true for only very simple models that explicitly consider uncertainty. For more complex models, the methods for dealing with variability and uncertainty provide the analyst with information that is useful in many contexts. Optimization, however, is usually accomplished by trial and error.

Variability

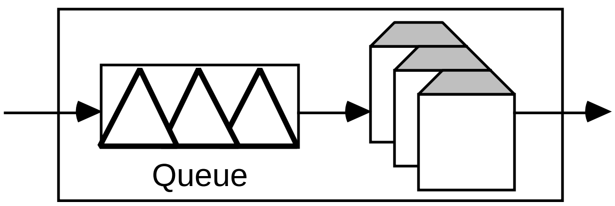


Figure 1. Machine station with queue

To illustrate variability, consider a machine station implementing some operation in a manufacturing process as shown in the Fig. 1. There are three machines, each capable of processing the product that queues up on the left. After the necessary operations are performed, the product leaves to the right for subsequent processing at another station or delivery to a finished goods inventory.

There are many reasons for variability in both the arrival of items to the station and the time to perform the operation. Perhaps the rate of arrival to the workstation is affected by customer orders. Occasionally, the demand rate be high, while at other times it may be low. The customers are probably acting independently, so the time between adjacent orders are extremely variable. Even if the customer demand is constant or effectively buffered by a finished goods inventory, there are many sources of variability within the production system. Scheduling of several products through a shop typically results in variations in the arrival pattern at a particular machine. The removal of defects in a production line causes randomness in arrivals to stations downstream from the point of removal. Machine failures cause periods of downtime for components of the system resulting in disruptions and variability in product flow.

The processing time is also a source of variability. Items of a product may not be exactly alike, requiring more or less processing at a particular operation. Human activities introduce significant variability in operation times. Even a robot may require several tries at performing its activity, leading to variability. A particular machine may be required to manufacture different items, introducing variability through different processing requirements.

One effect of arrival and processing variability is the formation of a queue of items waiting for service. Even if the machines have sufficient capacity to satisfy the average rate of input, a queue eventually appears at the workstation. The queue will vary in length over time, sometimes growing, sometimes shrinking and sometimes disappearing. If the queue and its effects are important to the design, operation and evaluation of the system, we must provide ways to measure and model variability. A deterministic model using average values for the arrival and processing rates does not indicate that a queue will exist, much less provide estimates of its characteristics.

We should note that the queue will form even if we know all about the variability. If given the number of items in the queue at the start of some observation period, a table of arrival times for products and a second table of processing times for each item, we can predict quite accurately how the length of queue will vary over time.

The queue is the result of variability, not uncertainty. It is true that we could propose a plan to handle the variability if we had prior knowledge of the variable quantities. This planning problem is often difficult, but it can be approached with the optimization methods in the first part of this book. It is the combined effects of variability and uncertainty that makes planning difficult and suggests the methods in the latter chapters.

Uncertainty

We can observe the variability of the past but it is impossible to predict with 100 percent accuracy the variability of the future. Even so, all of us must make decisions in the face of this uncertainty. The stock investor plots the market price of a particular stock in the hope that an investment now will yield a profit in the future; the production planner purchases raw materials and initiates a production process with the idea of fulfilling future demand; parents buy insurance to protect their family against catastrophic events; the government spends huge amounts of money on water storage facilities to protect against the variability in future rainfall; and the gambler lays down a bet that the next card will be the one he wants. Although, we cannot know the future, we must respond to its possibilities. We need methods for dealing with uncertainty.

It is just about impossible to enumerate the aspects of a particular situation that we might *not* know with certainty. One could easily argue that we know nothing with certainty about the future. To propose a model we will abstract reality by supposing that some aspects are known while others are not. By doing so, we are able to construct the models and methods presented in the second part of the book.

The first issue that we must come to grips with is whether an uncertain quantity is entirely unknown or can be assigned a probability distribution. This is an important distinction because most of the methods associated with models of random systems assume that one can specify a probability distribution for unknown quantities. For example, a card player may be interested in the probability of drawing an ace from a deck of cards. If, she knows that cards come from a standard bridge deck (4 aces in 52 cards) and that the cards are thoroughly mixed, she can logically say that the probability of an ace is 4/52 or 1/13. Most people understand the idea of "chance" and feel comfortable with this definition of probability. Another way of saying this is that if the player performed the experiment of drawing a card over and over again, the expected number of aces would be 1/13 of the total number of experiments. Any person that tries this experiment will not observe 1/13 of the cards to be aces, but that should not be a concern. When events are governed by chance, one expects variability in the results of experiments.

The information about probability can be quite useful. Perhaps the player is about to place a bet on the result of the next draw. Although she cannot know the result with certainty, knowing the probability of drawing an ace will affect the odds she will require for the bet.

The situation would be quite changed if the player did not know about the number of aces or the total number of cards in the deck, or suspected that the person mixing the deck could manipulate the order of the cards or the card selected. Here both the identity of the next card and the probability that the next card is an ace are unknown. Any decisions based on this situation are clearly different than the case with known probabilities. The subject of *game theory* provides techniques that deal with limited cases of this type.

In more formal discussions, the phrase *decision making under risk* is used to describe situations where probabilities are assumed known, and *decision making under uncertainty* is used to describe situations where we do not know probabilities. In the text, we concentrate on problems involving probability; however, we use the more comfortable term *uncertainty* to describe them.

Complexity

When the situation is very simple and probability distributions for uncertain variables are assumed known, simple mathematical models can be described. Some examples are observed in the study of inventory systems. Here the expected cost of operating the inventory system is expressed as a nonlinear function of its parameters. The methods of nonlinear programming can be used to determine the optimal parameter values.

Simple queueing system models in Chapter 16 are manipulated to yield closed form expressions for such quantities as the expected length and expected time in the queue. Given cost information regarding the system, a discrete search procedure is used in the section on Markov queueing systems to find the optimal design characteristics.

Only a slight increase in system complexity, however, makes analytical results difficult to obtain. In most cases, the task of computing the response of the system to a single setting of parameters is so difficult that we concentrate on this analysis problem. In Chapters 12 through most of 17 we are satisfied to compute the probabilities for the states of stochastic processes. The goal of finding closed form or algorithmic approaches to optimal solutions is beyond our capabilities.

When a situation involving the interactions of several random variables is considered, analytical approaches usually fail to give information about the system response. Simulation, discussed in this chapter, may be necessary. Here a run of a simulation provides a single observation of the system response. When the response is a random variable, it is necessary to run the simulation many times to find good estimates of the expected values. Optimization is reduced to trial and error, a very expensive computational procedure for complex systems.

Decision Making

There is no question that decision making in the face of variability, uncertainty and complexity is an important topic. It is one of our prime activities, arising in every field of purposeful work and in most of the games that we invent for our amusement.

In the second part of the book, we introduce various special cases of problems that are amenable to quantitative analysis. We propose methods for dealing with these models and arriving at solutions for the problems stated. Because of the difficulty of dealing with uncertainty, these models are far less complex and the solution methods are far less powerful than those proposed for deterministic systems. We do not suggest that these results handle all problems or solve unequivocally any real problem. We believe, however, that the models and methods do yield insights and guidance to experts and decision makers operating in this difficult area. The formulation of an effective model, interpretation of the results of a computational procedure and the implementation of a policy depend for success on the experience of the analyst, or the team charged with the analysis, and the person responsible for the decision.

21.2 Examples of Random Outcomes

For problems in which randomness places a role, the decision maker acts but is unable to predict with certainty the response of the system. A principal tool for quantitative modeling and analysis in this context is probability theory. Even though a system response is uncertain, specifying the probability distribution over the various possible outcomes can still be very useful to the decision making.

Example 1: The Game of Craps

This is your first trip to Las Vegas and you are thinking about trying out the gaming tables. As a student, you don't have much money, so you're a little afraid of diving right in. After all, the minimum bet is $5, and it won't take many losses before your gambling budget of $100 is exhausted. You are looking forward to a good time but the uncertainty and risk involved is unsettling.

The hotel has a special channel on the television that helps you learn the rules. You turn it on, and they are discussing the game of craps. In this game, the player rolls a pair of dice and sums the numbers showing. A sum of 7 or 11 wins for the player, and a sum of 2, 3 or 12 loses. Any other number is called the point. The player then rolls the dice again. If she rolls the point number she wins, if she throws a 7 she loses. Any other number requires another roll. The game continues until the gambler rolls a 7 or her point.

Before going downstairs to the tables, you decide to do a little analysis. Perhaps that probability course you took in college will finally bear fruit. With a little thought you determine the probabilities of the possible outcomes of the roll of a pair of dice and construct Table 1.

Table 1. Probabilities for a throw of two dice

Sum	2	3	4	5	6	7	8	9	10	11	12
Probability	1/36	2/36	3/36	4/36	5/36	6/36	5/36	4/36	3/36	2/36	1/36
Cumulative	1/36	3/36	6/36	10/36	15/36	21/36	26/36	30/36	33/36	35/36	36/36

In technical terms, the activity associated with an uncertain outcome is a random *experiment*, a process that can be repeated and which has two or more outcomes (sample points) that are affected by chance. A sample space (S) is the set of all possible outcomes of the experiment. For most random experiments, the interest is not on the sample points but on some value derived from the attribute being measured. This typically takes the form of an observed or computed value known as a *random variable*. The realization of a random variable is a random variate.

The current experiment is throwing a pair of dice, the sample space is all possible combinations of the two faces, and the random variable as defined here is the sum of the numbers facing up. Table 1, enumerating all possible values of the random variable and their probabilities, constitutes the model of the game. Games of chance are interesting examples because logical arguments or repeated trials easily verify the probabilities describing the model.

The table row labeled "sum" provides the possible observations of the random variable. The sum is a *discrete random variable*. The set of possible values is

$$\boldsymbol{X} = \{2, 3, \ldots, 12\}.$$

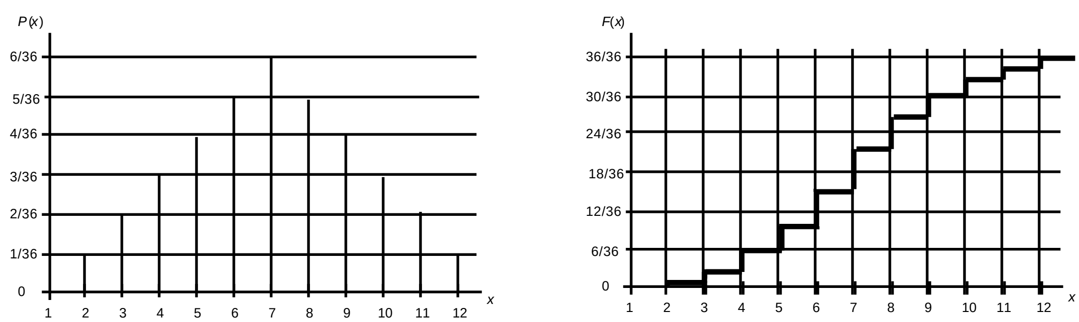
The row labeled probability shows the chance that a single throw results in each of the various possible values of the random variable. The collection of probabilities is *probability distribution function* (PDF) of the random variable. Assign the notation $P_X(k)$ to represent the probability that the random variable X takes on the value k. Probabilities are always nonnegative, and their sum over all possible values of the random variable must be 1.

The *cumulative distribution function* (CDF) describes the probability that the random variable is less than or equal to a specified value. The value of the cumulative distribution function at b is

$$F_X(b) = \sum_{k \le b} P_X(k)$$

We drop the subscript x on both P_X and F_X when there is no loss of clarity.

The distributions for this situation are graphed in Fig. 2. The PDF. has the value 0 at $x = 1$, rises linearly to a peak at $x = 7$, and falls linearly to the value of 0 at $x = 13$. The function is called a *triangular distribution.* The CDF shown in Fig. 2 has the typical pattern of all functions of this type, starting at 0 and rising until it reaches the value of 1. The function remains at 1 for all values greater than 12. For the discrete random variable, the PDF has nonzero values only at the integers, while the CDF is defined for all real values.



Probability distribution function

Cumulative distribution function

Figure 2. Distribution functions for total on dice (triangular distribution)

Computing Probabilities of Events

With the PDF in hand, you easily compute the probability that you will win or lose on the first throw of the dice. Because the possible values are mutually exclusive, you simply add the probabilities to find the probability of win or lose.

$$P(\text{win}) = P(7) + P(11) = 6/36 + 2/36 = 0.222$$

$$P(\text{lose}) = P(2) + P(3) + P(12) = 1/36 + 2/36 + 1/36 = 0.111$$

The odds look good, with the probability of a win equal to twice that of a loss. You confidently head toward the tables. Unfortunately, the game does not go well, and you lose a good deal of your stake. Perhaps your analysis stopped a little early. We will continue this example later to find a better estimate of your chances of winning this game.

An *event* is a subset of the outcomes of an experiment. Event probabilities are computed by summing over the values of the random variable that make up the event. In the case of the dice, we identified the event of a "win" as the outcomes 7 and 11. The probability of a win is the sum of the probabilities for these two values of the random variable.

In many cases events are expressed as ranges in the random variable. In general, the probability that the discrete random variable falls between a and b is

$$P(a \le x \le b) = \sum_{k=a}^{b} P_x(k) = \sum_{k=0}^{b} P_X(k) - \sum_{k=0}^{a-1} P_X(k)$$

$$= F_X(b) - F_X(a-1) \text{ for } a \le b.$$

Range event probabilities are computed by summing over the values of the random variable that make up the event or differencing the cumulative distribution for the values defining the range.

For discrete random variables, the relation ≤ is different than the relation < since a nonzero probability is assigned to a specific value. Another useful expression comes from the fact that the total probability equals 1.

$$P(x \ge a) = 1 - F_X(a-1).$$

To illustrate, several events for the dice example appear in Table 2. The random variable, x, is the sum of the two dice. The examples show how different phrases are used to describe ranges. The values of the cumulative distribution come from Table 1.

Table 2. Simple range events

Event	Probability
Sum on the dice is less than 7	$P(x < 7) = P(x \le 6) = F_X(6) = 15/36$
Sum is between 3 and 10 inclusive	$P(3 \le x \le 10) = F_X(10) - F_X(2) = 32/36$
Sum is more than 7	$P(x > 7) = 1 - P(x \le 7) = 1 - F_X(7) = 15/36$
Sum is at least 7	$P(x \ge 7) = 1 - F_X(6) = 21/36$
Sum is no more than 7	$P(x \le 7) = F_X(7) = 21/36$

Combinations of Events

Based on the results of the first night's play, you feel that you were a little rash considering only the probability of winning or losing on the first roll of the dice. In fact, most of your loses occurred when didn't roll a winning or losing number and were forced to roll for a point. You must roll for a point if the first roll is between 4 and 6 or between 8 and 10. You define the event $\boldsymbol{A}$ as the outcomes $\{4 \le x \le 6\}$ and the event $\boldsymbol{B}$ as the outcomes $\{8 \le x \le 10\}$. Using the rules of probability you determine that

$$P(\boldsymbol{A}) = F(6) - F(3) = 15/36 - 3/36 = 1/3,$$

$$P(\boldsymbol{B}) = F(10) - F(7) = 33/36 - 21/36 = 1/3.$$

The probability that you must roll a point is the probability that either event $\boldsymbol{A}$ or event $\boldsymbol{B}$ occurs. This is the union of these two events (written $\boldsymbol{A} \cup \boldsymbol{B}$). Again from your probability course you recall that since the events are mutually exclusive, the probability of the union of these two events is the sum of their probabilities.

$$P(\text{throw a point }) = P(\boldsymbol{A} \cup \boldsymbol{B}) = 1/3 + 1/3 = 2/3.$$

Together win, lose and throw a point represent all possibilities of the first roll, so it is not surprising that

$$P(\text{win}) + P(\text{lose}) + P(\text{throw a point}) = 0.222 + 0.111 + 0.667 = 1.$$

Events are sets of outcomes so we use set notation such as $\boldsymbol{A}$ and $\boldsymbol{B}$ to identify them. The union is two events, that is the outcomes in either $\boldsymbol{A}$ *or* $\boldsymbol{B}$, is written $\boldsymbol{A} \cup \boldsymbol{B}$. The intersection of two events, that is the set of outcomes in both events $\boldsymbol{A}$ *and* $\boldsymbol{B}$, is written $\boldsymbol{A} \cap \boldsymbol{B}$.

Rule 1: If $\boldsymbol{A}$ and $\boldsymbol{B}$ are mutually exclusive, $P(\boldsymbol{A} \cup \boldsymbol{B}) = P(\boldsymbol{A}) + P(\boldsymbol{B})$.

When two events are mutually exclusive they have no outcomes in common, and the probability of their union is the sum of the two event probabilities as illustrated above.

Rule 2: Probability of the union of any two events is

$$P(\boldsymbol{A} \cup \boldsymbol{B}) = P(\boldsymbol{A}) + P(\boldsymbol{B}) - P(\boldsymbol{A} \cap \boldsymbol{B}).$$

When the events are not mutually exclusive, we have the more general expression where $P(\boldsymbol{A} \cap \boldsymbol{B})$, is the probability that both events occur. For example, let $\boldsymbol{A} = \{\text{roll less than } 8\}$ and $\boldsymbol{B} = \{\text{roll more than } 6\}$. We compute

$$P(\boldsymbol{A}) = F(7) = 21/36,\ P(\boldsymbol{B}) = 1 - F(6) = 1 - 15/36 = 21/36.$$

To compute the probability of the union, we must first find the probability of the intersection event $P(\boldsymbol{A} \cap \boldsymbol{B}) = P(7) = 6/36$. Using the general expression for the union we find

$$P(\boldsymbol{A}) + P(\boldsymbol{B}) - P(\boldsymbol{A} \cap \boldsymbol{B}) = 21/36 + 21/36 - 6/36 = 1.$$

Indeed the union of these two events includes all possible outcomes.

Rule 3: The probabilities of the intersection of independent events is

$$P(\boldsymbol{A} \cap \boldsymbol{B}) = P(\boldsymbol{A})\, P(\boldsymbol{B}).$$

Rule 4: The probability of the union of independent events is

$$P(\boldsymbol{A} \cup \boldsymbol{B}) = P(\boldsymbol{A}) + P(\boldsymbol{B}) - P(\boldsymbol{A})\, P(\boldsymbol{B}).$$

Two events are independent if the occurrence of one does not affect the probabilities associated with the other. As an example, we might be interested in the probability that two sixes are thrown for a pair of dice. The results of the two throws are independent, and we compute

$$P(\{\text{six on first}\} \cap \{\text{six on second}\}) = P(\text{six on first})P(\text{six on second}) = (1/6)(1/6) = 1/36.$$

When events are not independent, they are related by conditional probabilities. Consider the events $\boldsymbol{A}$ = {roll a 4 on the first play} and $\boldsymbol{B}$ = {win the game}. These events are not independent because the probability of winning depends on whether or not we roll a 4. We can, however, define the conditional probability of winning given that a 4 is rolled on the first play, written $P(\boldsymbol{B} \mid \boldsymbol{A})$. To compute this probability, we assume we have rolled a point of 4. On the second roll, the probability that we win, given that the game ends is

$$P(\text{roll a 4}) / [P(\text{roll a 4}) + P(\text{roll a 7})] = 1/3.$$

Rule 5: The general expression for the intersection probability is

$$P(\boldsymbol{A} \cap \boldsymbol{B}) = P(\boldsymbol{A})\, P(\boldsymbol{B} \mid \boldsymbol{A}) .$$

We can't say how many rolls it will take, but this probability remains the same throughout the remainder of the game, so we conclude that

Rule 6: For independent events, $P(\boldsymbol{B} \mid \boldsymbol{A}) = P(\boldsymbol{B})$ and $P(\boldsymbol{A} \mid \boldsymbol{B}) = P(\boldsymbol{A})$.

As a consequence, we have $P(\boldsymbol{B} \mid \boldsymbol{A}) = P$(win the game / roll a 4 on the first play) = 1/3. Also, using the general formula for computing the intersection probability, gives

$$P(\boldsymbol{A} \cap \boldsymbol{B}) = (3/36)(1/3) = 1/36.$$

You now have enough tools to complete the analysis of the game. Consider an experiment that is a single complete play of the game. You define the following events:

$\boldsymbol{W}$: Win the game

$\boldsymbol{L}$: Lose the game

$\boldsymbol{W}_1$: Throw a point of k on the first roll, $k = 4, 5, 6, 8, 9, 10$.

The probabilities associated with the events are $P(\boldsymbol{P}_4) = 3/36$, $P(\boldsymbol{P}_5) = 4/36$, $P(\boldsymbol{P}_6) = 5/36$, $P(\boldsymbol{P}_8) = 5/36$, $P(\boldsymbol{P}_9) = 4/36$, $P(\boldsymbol{P}_{10}) = 3/36$.

The conditional probabilities, $P(\boldsymbol{W} \mid \boldsymbol{P}_k)$, of a win given a point of k on the first roll are $P(\boldsymbol{W} \mid \boldsymbol{P}_4) = 1/3$, $P(\boldsymbol{W} \mid \boldsymbol{P}_5) = 4/10$, $P(\boldsymbol{W} \mid \boldsymbol{P}_6) = 5/11$, $P(\boldsymbol{W} \mid \boldsymbol{P}_8) = 5/11$, $P(\boldsymbol{W} \mid \boldsymbol{P}_9) = 4/10$, $P(\boldsymbol{W} \mid \boldsymbol{P}_{10}) = 1/3$.

The event of winning the game combines all possible winning combinations of events.

$$\boldsymbol{W} = \boldsymbol{W}_1 \quad (\boldsymbol{P}_4 \cap \boldsymbol{W}_4) \quad (\boldsymbol{P}_5 \cap \boldsymbol{W}_5) \quad (\boldsymbol{P}_6 \cap \boldsymbol{W}_6) \quad (\boldsymbol{P}_8 \cap \boldsymbol{W}_8)$$
$$(\boldsymbol{P}_9 \cap \boldsymbol{W}_9) \quad (\boldsymbol{P}_{10} \cap \boldsymbol{W}_{10})$$

$$P(\boldsymbol{W}) = P(\boldsymbol{W}_1) + P(\boldsymbol{P}_4)\, P(\boldsymbol{W}|\boldsymbol{P}_4) + P(\boldsymbol{P}_5)\, P(\boldsymbol{W}|\boldsymbol{P}_5) + P(\boldsymbol{P}_6)\, P(\boldsymbol{W}|\boldsymbol{P}_6)$$
$$+ P(\boldsymbol{P}_8)\, P(\boldsymbol{W}|\boldsymbol{P}_8) + P(\boldsymbol{P}_9)\, P(\boldsymbol{W}|\boldsymbol{P}_9) + P(\boldsymbol{P}_{10})\, P(\boldsymbol{W}|\boldsymbol{P}_{10})$$
$$= 0.493.$$

If you don't win you lose, so $P(\boldsymbol{L}) = 1 - P(\boldsymbol{W}) = 0.507$.

Your odds of winning on any one play are slightly less than 50 - 50. You now know that the casino has a better chance of winning the game than you do, so if you play long enough you will lose all your money. If this were an economic decision, your best bet is not to play the game. If you do decide to play, you go to the tables armed with the knowledge of your odds of winning and losing. The probability model has changed the situation to one of risk rather than uncertainty. You can use the probability model to answer many other questions about this game of chance.

Example 2: Two-Machine Workstation

You are an machine operator and have recently been promoted to be in charge of two identical machines in a production line. You just received your first performance review from your supervisor. The report is not good and it ends with the comment,

> "Your station fails in both efficiency and work-in-process. You are inconsistent in your work habits. At times you sit idle and at other times the backlog is overflowing from your station. A work study has determined that you have sufficient capacity to process all the jobs assigned to you. It's hard to imagine how you can fail to meet your schedules, have idle time and excess work-in-process. If you don't improve, we'll have to reevaluate your promotion."

You throw down the review in anger, blaming the workstation that comes before yours as not providing jobs on a regular basis. You also blame the maintenance department for not keeping your machines in good shape. They often require frequent adjustments and no two jobs take the same time. You also complain about the production manager who takes work away from you and gives it to another station whenever the number waiting exceeds 6 jobs. After you cool down you pledge to yourself to try to improve. You really need this position and the extra pay it provides.

You visit the industrial engineering department and ask for some ideas to improve your efficiency. The engineer listens to your problem and points out that perhaps the situation is not your fault. He explains that variability in the arrival process and machine operation times can result in both queues and idleness. Using data from the work study he determines that you are assigned an average of 50 jobs per week and that each job takes an average of 1.4 hours for a total work load of 70 hours per week. With two machines working 40 hours per week, you have a production capacity of 80 hours, sufficient time to do the work assigned. As a first attempt at modeling your station, the engineer uses a queueing analysis[1] program and

[1] Here we take these probabilities as given. Chapter16 provides formulas to compute them.

finds the following probability distribution for the number of jobs that are at your station.

No. in system	0	1	2	3	4	5	6	7	8
Probability	0.098	0.172	0.15	0.131	0.115	0.101	0.088	0.077	0.067
Cumulative	0.098	0.27	0.42	0.552	0.667	0.767	0.855	0.933	1.0

The table does give some interesting insights. The model predicts that you are idle almost 10% of the time (when the number in the system is 0). About 7% of the jobs entering the station will find all six queue positions full (when the number in the system is 8). Since these jobs are sent to another operator, the number of jobs actually processed in your station is 50(1 – 0.067) = 46.63. Between three and four jobs per week are taken away from you because of congestion at your station. The remaining jobs require an average of 46.63 × 1.4 = 65.3 hours per week. Since your machines have 80 hours of capacity, their efficiency is 65.3/80 = 82%. The work-in-process (WIP) is the expected number of jobs in the station. Using probability theory you calculate the expected value for the number of jobs as

$$\mu = \sum_{k=0}^{8} kP_x(k) = 0(0.098) + 1(0.172) + 2(0.15) + \cdots + 8(0.067) = 3.438.$$

You are not very encouraged by these results. They clearly support your supervisor's opinion that you are simultaneously inefficient, lazy, and generally not doing your job. The kindly engineer consoles you that the results are the inevitable cause of variability. You can do nothing about the situation unless you can reduce the variability of job arrivals to your station, the variability of the processing times, or both. If these factors are beyond your control, you better try to convince your supervisor that the problem is not your fault and that he should look elsewhere for solutions.

Both the gambling and manufacturing examples illustrate how probability models can be used to provide insight about a current situation. The gambling example illustrates a case where the underlying causes of variability are well known and the analysis yields accurate estimates about the probabilities of winning and losing. The model for the manufacturing example rests on arguable assumptions so the results may not be accurate. The model is valuable however, in that it illustrates an aspect of the situation often not apparent to decision makers. The analysis shows that the inherent randomness in the process is the source of the difficulty, and should stimulate the search for new solutions.

21.3 Discrete Random Variables

Results of an experiment involving uncertainty are described by one or more random variables. This section considers the discrete random variable, while the continuous case is the subject of the next section. Here the probability distribution is specified by a nonzero probability assigned to each possible value of the random variable.

For a particular decision situation, the analyst must assign a distribution to each random variable. One method is to perform repeated replications of the experiment. Statistical analysis can then be used to estimate the probability of each possible occurrence. Another and often more practical method is to identify the distribution to be one of the *named distributions*. It is much easier to estimate the parameters of such a distribution, rather than to estimate the entire set of probabilities. A catalog of some of the important named distributions and examples of their use is provided below

Also presented are formulas for the moments and probabilities for the general discrete distribution as well as for the more prominent named distributions. Fortunately, when working with real problems, it is not always necessary to know the formulas or even how to use them. Computer programs are readily available to do the analysis for given parameter values[2]. The formulas, however, provide a higher level of knowledge regarding the distributions and are frequently very helpful for estimating parameters.

Describing Probability Distributions

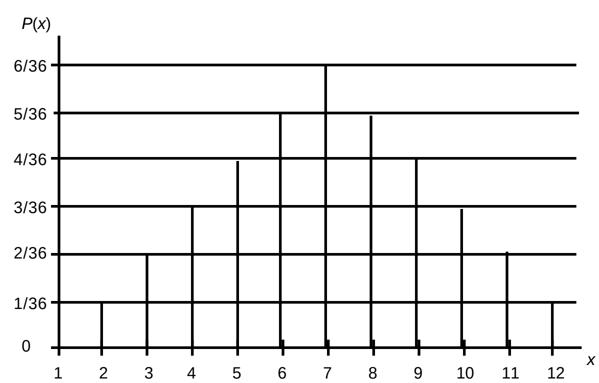


Figure 3. Distribution functions for total on dice (Triangular Distribution)

A discrete probability distribution function is completely described by the set of possible values the random variable can take and by the probabilities assigned to each point in the sample space. The notation $P_X(k)$ for k = 0, 1,…, indicates that the random variable takes on the value of any nonnegative integer. Providing values of $P_X(k)$ for each k completely specifies the distribution. The triangular distribution of Fig. 3 is used as an illustration.

Certain quantities can be computed from the PDF that describe simple characteristics of the distribution. These are called moments. The most common is the mean, μ, the first moment about the origin. We give the general definition for the mean along with its computation for the triangular distribution in Table 3. This and following definitions assume that the random variable is defined over all nonnegative integers. The definitions can be easily expanded to cover other domains. The mean

[2] The Probability Add-in allows the user to define random variables and assign them named distributions. Functions are available to compute moments and probabilities. Random variables may be simulated to model more complex situations.

provides a measure of the *center* of the distribution. For a symmetric distribution, the mean is the middle value of the range of the random variable. Another term often used for the mean is the *expected value of X*, written $E[X]$.

Table 3. Descriptive measures of distributions

Measure	General formula	Sum of the dice
Mean	$\mu = \sum_{k=0}^{\infty} k P_X(k)$	$\mu = 2(1/36) + 3(2/36) + \cdots + 12(1/36) = 7$
Variance	$\sigma^2 = \sum_{k=0}^{\infty} (k-\mu)^2 P_X(k)$	$\sigma^2 = (2-7)^2(1/36) + (3-7)^2(2/36) + \cdots + (12-7)^2(1/36) = 5.833$
Standard deviation	$\sigma = \sqrt{\sigma^2}$	$\sigma = \sqrt{5.833} = 2.415$
Skewness	$\beta_1 = \frac{(\mu_3)^2}{\sigma^6}$ $\mu_3 = \sum_{k=0}^{\infty} (k-\mu)^3 P_X(k)$	$\beta_1 = 0$
Kurtosis	$\beta_2 = \frac{\mu_4}{\sigma^4}$ $\mu_4 = \sum_{k=0}^{\infty} (k-\mu)^4 P_X(k)$	$\beta_2 = 2.365$

The variance, σ^2, is a measure of the *spread* of the distribution about the mean. It is the second moment about the mean. Where the random variable has high probabilities near the mean, the variance is small, while large probabilities away from the mean imply that the variance is large. The standard deviation, σ, is simply the positive square root of the variance and also describes the spread of the distribution. Discussions more often use the standard deviation because it has the same dimension as the mean.

Two additional measures that describe features of a distribution are the skewness and kurtosis with general measures β_1 and β_2 given in Table 3. A positive skewness indicates that the distribution is concentrated to the with a long thin tail pointing to the right, and a negative skewness has the concentration to the right and the tail pointed to the left. Kurtosis measures the peakedness of the distribution. We illustrate the skewness measure in the several examples of this section. The example in Fig. 3 has a skewness of 0 because the distribution is symmetric.

Named Discrete Distributions

It is useful for modeling purposes to know about the named discrete distributions. When an experiment on which a random variable is based satisfies the logical conditions associated with a named distribution, the distribution for the random variable is immediately determined. Then we can use the distribution without extensive experimentation to answer decision questions about the situation.

Bernoulli Distribution

Example 3: Consider again the Craps game. If on the first roll of the dice you throw a number other than 2, 3, 7, 11, or 12, the number you do throw is your point. The rules say you must roll the dice again and continue to roll until you throw your point and win, or a 7, and lose. Say your point is 4. Based on your probability model you determine that on any given roll following the first:

$P(\text{win}) = P(x = 4) = 3/36.$

$P(\text{lose}) = P(x = 7) = 6/36.$

$P(\text{roll again}) = 1 - P(\text{win}) - P(\text{lose}) = 27/36 = 3/4.$

For each roll, the game either terminates with probability 1/4, or you must roll again with probability 3/4.

Bernoulli Distribution
Parameter: $0 < p < 1$
$P(1) = p$ and $P(0) = 1 - p$
$\mu = p,\ \sigma^2 = p(1 - p)$

An experiment that has two outcomes is called a *Bernoulli trial*. For the example we take the two outcomes as “roll again” and “terminate”, and arbitrarily assign the value 0 to the roll again outcome and value 1 to the terminate outcome. The parameter associated with the probability distribution is the probability that the variable assumes the value 1 indicated by p. Given the value of p, the entire distribution is specified. For the example

$$P(\text{terminate}) = P(1) = 1/4 \text{ and } P(\text{roll again}) = P(0) = 3/4.$$

The simple *Bernoulli* distribution illustrated with this example is the first of several named distributions presented in this chapter. These distributions are useful because they model a variety of situations.

Geometric Distribution

Example 4: If you don’t win or lose on the first roll, you might wonder how long the game will last. Assume you roll a point of 4. Now you begin the second phase of the game and define the random variable as the number of rolls prior to the last roll. That number may be 0, 1, 2,… which is random variable described by the geometric distribution.

This random variable has an infinity of possible values in that there is no upper limit to the number of rolls conceivably required. There is only one way a particular value, k, of the random variable can occur. There must be k *roll again* outcomes followed by one termination. The probability of this occurrence describes the probability distribution function.

Geometric Distribution
Parameter: $0 < p < 1$
$P(k) = p\,(1-p)^k$ for $k = 0, 1, 2, \ldots$
$\mu = \dfrac{1-p}{p}, \quad \sigma^2 = \dfrac{1-p}{p^2}$

The geometric distribution has the single parameter p, with the requirement that ($0 < p < 1$). The logical condition for this distribution is that the separate trials be independent, that is, the outcome of one trial does not affect the probability of a later trial. This is certainly true for sequential throws of the dice.

For the example the parameter is $p = P(\text{terminate}) = 1/4$. Table 4 shows the probability distribution for this case, and Fig. 4 shows a plot of the distribution. The game may take quite a few rolls to complete with a greater than 10% chance that more than 7 are required. The moments for the example are:

$$\mu = 3,\ \sigma^2 = 12,\ \sigma = 3.46,\ \beta_1 = 4.08,\ \beta_2 = 9.08$$

The positive skewness is clearly indicated in the plot.

Table 4. Geometric distribution

Number, k	0	1	2	3	4	5	6	7
Probability	0.25	0.188	0.141	0.105	0.079	0.059	0.044	0.033
Cumulative	0.25	0.438	0.578	0.684	0.763	0.822	0.867	0.9

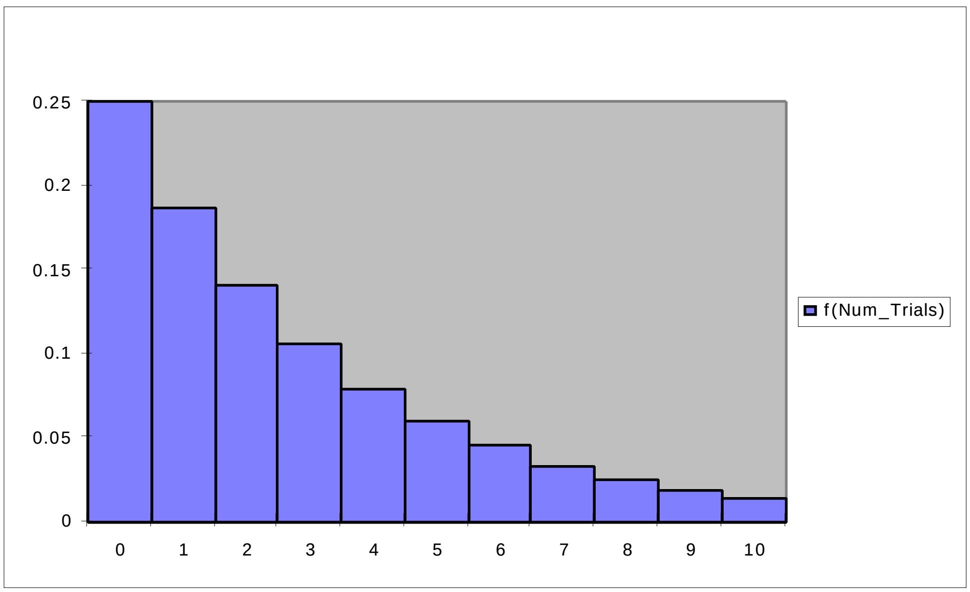


Figure 4. Plot of the geometric distribution

Using this probability model, it is easy to compute the mean and variance of the number of rolls to termination for each of the point numbers as given below.

Point	4	5	6	8	9	10
p	0.25	0.278	0.306	0.306	0.278	0.25
μ	3	2.6	2.273	2.273	2.6	3
σ^2	12	9.36	7.438	7.438	9.36	12

Negative Binomial Distribution

Example 5: In the game of craps, you decide to play until you lose 5 games. You wonder how many games you will play with this termination rule. Recall that the probability of losing any one game is 0.5071. The games are a series of independent Bernoulli trials, and the random variable is the number of wins until the fifth loss. This is a situation described by the negative binomial distribution.

Negative Binomial Distribution

Parameters: $0 < p < 1$, $r \geq 1$ and integer

$$P_X(k) = \binom{r+k-1}{r-1} p^r (1-p)^k$$

for $k = 0, 1, 2\ldots$

$$\mu = \frac{r(1-p)}{p}, \quad \sigma^2 = \frac{r(1-p)}{p^2}$$

For this distribution we first identify the result of success. In this case, we perversely defined success as a “loss” with p the probability of a success equal to 0.5071 for the example. The random variable is the number of trials that result in 0 before the r th 1 is observed. For this case, $r = 5$.

The distribution for the example is shown in the Table 5. It is important to remember that the random variable is not the total number of trials but the number of failed trials before the rth success. In Table 5, the entry for 0 describes the probability that the first five plays were losses and there were no wins. The geometric distribution is a special case when $r = 1$.

Table 5. Negative binomial distribution

Number, k	0	1	2	3	4	5	6	7
Probability	0.034	0.083	0.122	0.141	0.139	0.123	0.101	0.078
Cumulative	0.034	0.116	0.238	0.379	0.517	0.64	0.741	0.82

Binomial Distribution

Example 6: The reliability of a computer is defined as the probability of successful operation throughout a particular mission. A study determines that the reliability for a given mission as 0.9. Because the mission is very important and computer failure is extremely serious, we provide five identical computers for this mission. The computers operate independently and the failure or success of one does not affect the probability of failure or success of the others. Our job is to compute the probability of mission success, or system reliability, under the following three operating rules:

a. All five computers must work for mission success

b. At least three out of five must work for mission success

c. At least one computer must work for mission success

Consider an experiment that involves n independent Bernoulli trials. Associate with each outcome, the random variable that is the sum of the of the n Bernoulli random variables. This is called the binomial random variable. The variable has n+1 possible values ranging from 0 to n. Its PDF is the binomial distribution which has two parameters, p and n.

> *Binomial Distribution*
>
> Parameters: $0 < p < 1$, $n \geq 1$ and integer
>
> $$P_X(k) = \binom{n}{k} p^k (1 - p)^{n-k} \quad \text{for } k = 0,1,2,\ldots,n$$
>
> $$\binom{n}{k} = \frac{n!}{(n - k)!\ k!}$$
>
> $$\mu = np,\ \sigma^2 = np(1 - p)$$

In the case given, the success or failure of each computer is a Bernoulli random variable with 1 representing success and 0 representing failure. The probability of success is p, and we assume that the computers are independent with respect to failure.

The number of working computers, x, is the random variable of interest, and the binomial distribution, with parameters n = 5 and p = 0.9, is the appropriate PDF. With these parameters the probability of k successful computers is computed and the results are shown in the table. The moments of the distribution are:

$$\mu = 4.5,\ \sigma^2 = 0.45,\ \sigma = 0.67,\ \beta_1 = -1.42,\ \beta_2 = 4.02$$

The negative skewness is illustrated in the Fig. 5.

Number, k	0	1	2	3	4	5
Probability	0.00001	0.00045	0.0081	0.0729	0.32805	0.59049
Cumulative	0.00001	0.00046	0.00856	0.08146	0.40951	1

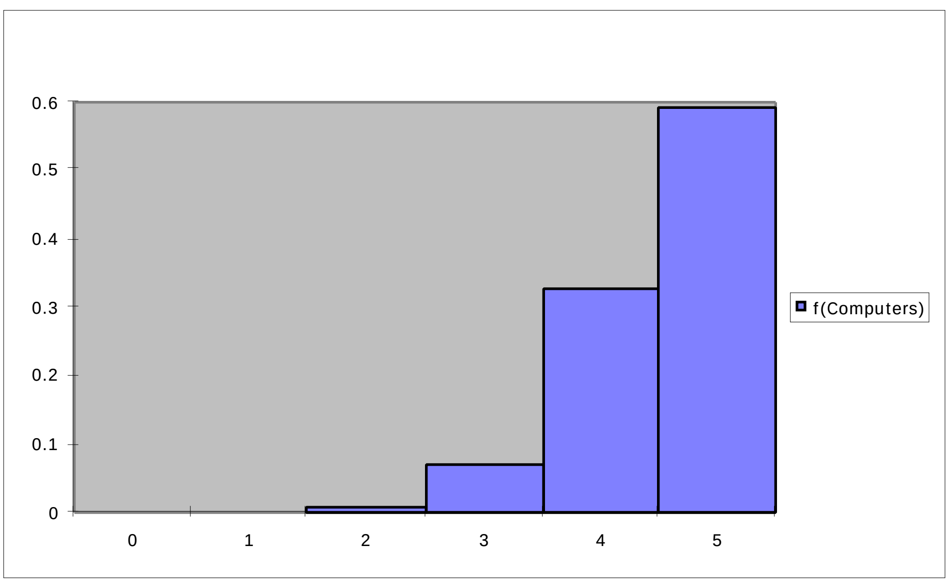


Figure 5. Plot of the binomial distribution

The probabilities of mission success under the three conditions listed earlier are

a. $P(x = 5) = 0.59049$.

b. $P(x \geq 3) = P(3) + P(4) + P(5) = 0.99144$.

c. $P(x \geq 1) = 1 - P(0) = 0.99999$.

These results show the value of redundancy for increasing reliability. In case *a*, none of the computers is redundant since all are required for successful operation. In case *b*, we say that two computers are redundant since only three are required. In case *c*, four are redundant. The reliability of the system obviously increases as redundancy is increased.

Poisson Distribution

Example 7: A traffic engineer is interested in the traffic intensity at a particular street corner during the 1 – 2 a.m. time period. Using a mechanical counting device, the number of vehicles passing the corner is recorded during the one hour interval for several days of the week. Although the numbers observed are highly variable, the average is 50 vehicles. The engineer wants a probability model to answer a variety of questions regarding the traffic.

Poisson Distribution
Parameter: $\theta > 0$
$P_X(k) = \dfrac{e^{-\theta}(\theta)^k}{k!}$ for $k \geq 0$
$\mu = \theta,\ \sigma^2 = \theta$

This situation fits the logical requirements of the *Poisson Distribution.* Consider arrivals that occur randomly but independently in time. Let the average arrival rate be equal to λ per unit of time and let the time interval be t. Then one would expect the number of arrivals during the interval to be $\theta = \lambda t$. The actual number of arrivals occurring in the interval is a random variable governed by the Poisson distribution.

The parameter of the distribution is the dimensionless quantity θ, which is the mean number of arrivals in the interval. We call an arrival process that gives rise to this kind of distribution a *Poisson process.* The distribution is very important in queueing theory and is discussed more fully in Chapter 16.

To use the distribution for the example, we must only assume that vehicles arrive independently and that the average arrival rate is constant. This does not mean that the vehicles pass in a steady stream with a fixed interval between cars. Rather, with the assumption of randomness, vehicle arrivals are extremely variable; the rate of 50 per hour is an average. Using the distribution, we can model the probabilities for any interval of time, however consider a one minute period. The random variable is the number of vehicles passing during the one minute period. The parameter of the distribution is

$$\theta = (50/\text{hour})\ (1/60\ \text{hour}) = 5/6 = 0.833.$$

The probability distribution for this example is computed with the formulas given for the Poisson distribution and shown in Table 6. Descriptive measures are

$$\mu = 0.833,\ \sigma^2 = 0.833,\ \sigma = 0.913,\ \beta_1 = 1.2,\ \beta_2 = 4.2$$

The mean and variance of the Poisson distribution are always equal.

Table 6. The Poisson Distribution with $\theta = 0.833$

Number, k	0	1	2	3	4	5
Probability	0.4346	0.3622	0.1509	0.0419	0.0087	0.0015
Cumulative	0.4346	0.7968	0.9477	0.9896	0.9983	0.9998

From the distribution, various probability statements can be made.

- P(no cars pass the corner) = P(0) = 0.4346
- P(at least two cars pass) = P(2) + P(3) + ⋯

This expression can be evaluated if enough terms are added together however an easier, more accurate way is

- P(at least two cars pass) = 1 – P(no more than one car passes)
  = 1 – P(0) – P(1) = 1 – F(1) = 1 – 0.7968 = 0.2032

Hypergeometric Distribution

Example 8. You are dealt a hand of 5 cards from a standard deck of 52 cards. Before looking at the hand you wonder about the number of aces among the five cards. This is a case for the hypergeometric distribution.

Hypergeometric Distribution

Parameters: N, a, n

all positive and integer

$a \le N,\ n \le N$

$$P(k) = \frac{\binom{a}{k}\binom{N-a}{n-k}}{\binom{N}{n}} \quad \text{for } k = 0, 1, \ldots, n$$

$$\mu = \frac{na}{N},\quad \sigma^2 = n\left(\frac{a}{N}\right)\left(\frac{N-a}{N}\right)\left(\frac{N-n}{N-1}\right)$$

The situation involves N items. a of the items are assigned the label 1 and $N - a$ are assigned the label 0. Select at random n items from the N available without replacement. The number of items labeled 1 is the random variable of interest. For the example, there is a fixed number from which to draw (N = 52 cards), the number of aces in the deck (a = 4 aces) and the sample is 5 cards (n = 5).

$P(k)$ is the probability that k of the items selected have the label 1. There are three parameters a, N and n. All parameters are integers. $0 \le a \le N$. $1 \le n \le N$. A combination expression is zero if the bottom number is greater than the top number. The probability distribution for the number of aces is in Table 7.

Table 7. Hypergeometric distribution with N = 52, n = 5, a = 4

Number, k	0	1	2	3	4	5
Probability	0.6588	0.2995	0.0399	0.0017	0.00001	0
Cumulative	0.6588	0.9583	0.9982	1	1	1

Triangular Distribution

Example 9: A computer is shipped with a multiple number of some part. A company that assembles computers is interested in the distribution of the number of parts required. No statistics are available, however, design of the computer assures that every computer requires at least 1 part and the most that can be installed is six parts. A production supervisor estimates that most computers require two parts. The only information we have is the range of the random variable and its mode (the most likely number). A reasonable estimate for the distribution of parts is the triangular distribution.

Triangular Distribution
Parameters: a, b, m all integer
$a \le m \le b$
$P(k) = d(k - a + 1)$ for $a \le k \le m$,
and $P(k) = e(b - k + 1)$ for $m \le k \le b$
Use the general formulas to compute moments

In general, identify a as the lower limit to the range, b as the upper limit, and m as the mode. Construct the distribution with the constants d and e that satisfy the requirements:

$$\sum_{k=a}^{b} P(k) = 1 \text{ and } d(m - a + 1) = e(b - m + 1).$$

For the example situation, d = 0.143 and e = 0.057, and the distribution is presented in Table 8.

Table 8. Triangular distribution with a = 1, b = 6, m = 2

Number, k	1	2	3	4	5	6
Probability	0.143	0.286	0.229	0.171	0.114	0.057
Cumulative	0.143	0.429	0.657	0.829	0.943	1

Uniform Distribution

Example 10: We continue the example used for the triangular distribution, but now assume the production manager has no idea how many parts will be used in a computer. He knows only that the design limits the range of the random variable between 1 and 6. In this situation of complete uncertainty, the uniform distribution might be used.

Uniform Distribution
Parameters: a, b all integer
$P(k) = 1/(b - a + 1)$ for $a \leq k \leq b$

$$\mu = \frac{a+b}{2}$$

Use general formula for σ^2

This distribution assigns equal probabilities to all possible values of the random variable within the range a to b, inclusive. For the example situation: $P(k) = 1/6$ for $1 \leq k \leq 6$.

Modeling

When attempting to use the results of this section to model a problem, the first step is to determine if the random variable is discrete. Whenever the question concerns counting, only the integers are relevant so the variable is obviously discrete. To assign a probability distribution to the random variable, review the special cases to see if any are appropriate. In Chapter 18, several goodness-of-fit tests are provided to aid this process.

If the experiment only has two outcomes, the Bernoulli random variable is the clear choice. If the experiment involves a sequence of independent observations each with two outcomes, then the binomial, geometric or negative binomial distributions may fit. A fixed number in the sequence suggests the binomial, while a question related to the first occurrence of one of the two outcomes suggests the geometric. The negative binomial is a generalization of the geometric where the random variable is the number of unsuccessful trials before the rth success.

The binomial and hypergeometric are appropriate when we are selecting a fixed number of items from a population. The binomial models the number of successes when the population is infinite or when the population is finite and the items are replaced after each trial. When items are not replaced the hypergeometric is the correct distribution.

The Poisson distribution is used when the question relates to counting the number of occurrences of some event in an interval of time (or some other measure). Key phrases that suggest the appropriateness of the Poisson are that the arrivals are "independent" or "random".

Triangular and uniform distributions are often used when very little is known concerning the situation. The uniform requires only the range, while the triangular needs the additional knowledge of the mode. The distributions may also be logical consequences of the features of an experiment. For example, the uniform distribution models the number on the face of a single die, while the triangular models the sum of two dice.

If any of the special cases can logically be applied, then the parameters of the distribution must be determined. In practice, they may be determined by the logic of a situation or by estimation using statistics from historical data. If none of the special cases fit, the random variable still has a probability distribution; however, some other rational must be used to assign probabilities to possible outcomes.

21.4 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.

Probability Distributions

Computing Probabilities

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

The *probability density function* (pdf) 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.

The meaning of pdf is different for discrete and continuous random variables. For discrete variables PDF is the *probability distribution function* and it provides the actual probabilities of the values associated with the discrete random variable. For continuous variables, pdf 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.

Triangular Distribution

Parameter: $0 \le c \le 1$

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

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

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

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.

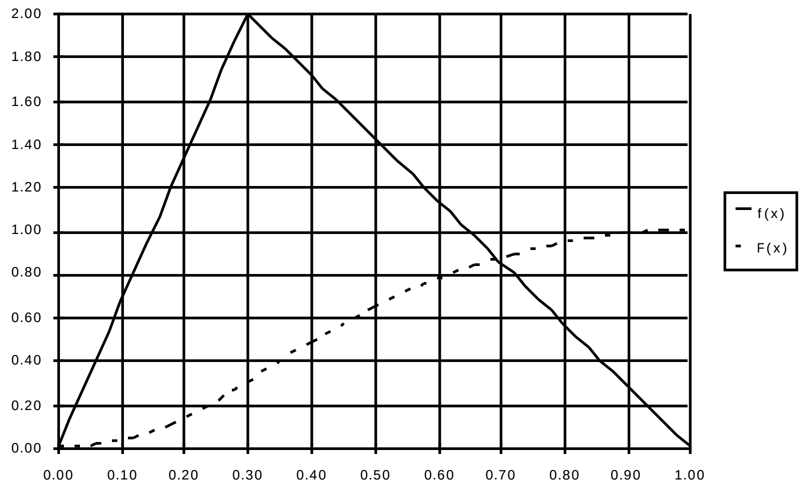


Figure 6. Triangular distribution with $c = 0.3$

Cumulative Distribution Function

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

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

The cumulative distribution function (CDF), $F_X(x)$, of the random variable is the probability that the random variable is less than or equal to the argument x. This is the same as the probability that the random variable is strictly less than the argument x because the probability that a continuous random variable takes on any particular value is zero; that is, $P(X = x) = 0$. Consequently, $P(X \le x) = P(X < x)$.

The pdf is the derivative of the CDF We drop the subscript on both f_X and F_X when there is no loss of clarity. The CDF 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 a legitimate pdf must therefore equal 1. Because negative probabilities are impossible, the pdf must remain nonnegative for all x.

CDF of the Triangular Distribution

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

Integrating the density function for the triangular distribution results in the CDF also shown in Fig. 6. The example illustrates the characteristics of every CDF. 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 CDF never decreases and remains constant only when the pdf 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 CDF, probabilities are easily determined using the CDF rather than by integration. For example,

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

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

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

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 CDF equal to 0.5.

Table 10. Descriptive measures

Measure	General formula	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	$\mu_3 = \int_{-\infty}^{\infty} (x-\mu)^3 f(x)\,dx$ $\beta_1 = \frac{(\mu_3)^2}{\sigma^6}$	
Kurtosis	$\mu_4 = \int_{-\infty}^{\infty} (x-\mu)^4 f(x)\,dx$ $\beta_2 = \frac{\mu_4}{\sigma^4}$	

Named Continuous Distributions

Models involving random occurrences 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 phenomena 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 bounds due to physical limitations may be identified. All of the distributions described below 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 telephone support line 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, the normal distribution may be appropriate.

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 CDF. The moments of the distributions have already been derived, aiding the analyst in selecting one that reasonably represents his or her situation.

Normal Distribution

Example 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.

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. 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. Figure 7 plots the distribution associated with the example.

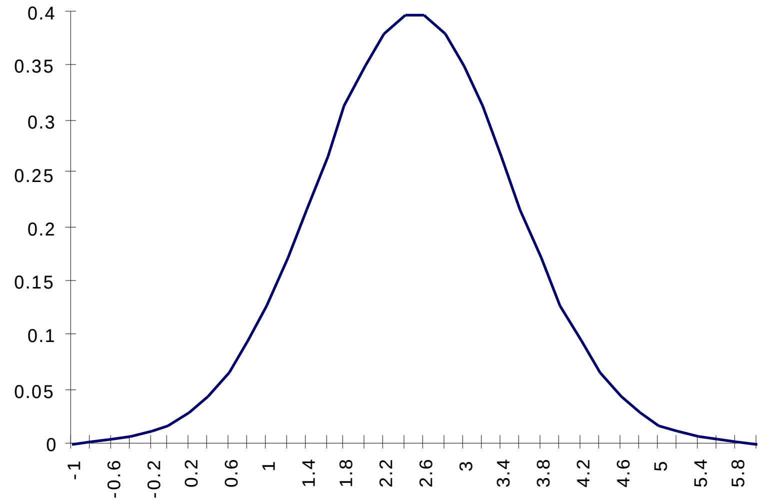


Figure 7. Normal distribution

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

Range	Formula	Probability
Within one standard deviation of the mean	$P(\mu-\sigma \le X \le \mu+\sigma)$	0.6826
Within two standard deviations of the mean	$P(\mu-2\sigma \le X \le \mu+2\sigma)$	0.9545
within three standard deviations of the mean	$P(\mu-3\sigma \le X \le \mu+3\sigma)$	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

$$P(X \le 0) = 0.0062$$

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

$$P(X \le z) = 0.9.$$

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

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

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,

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

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

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

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

Lognormal Distribution

Example 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

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

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

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

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

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.

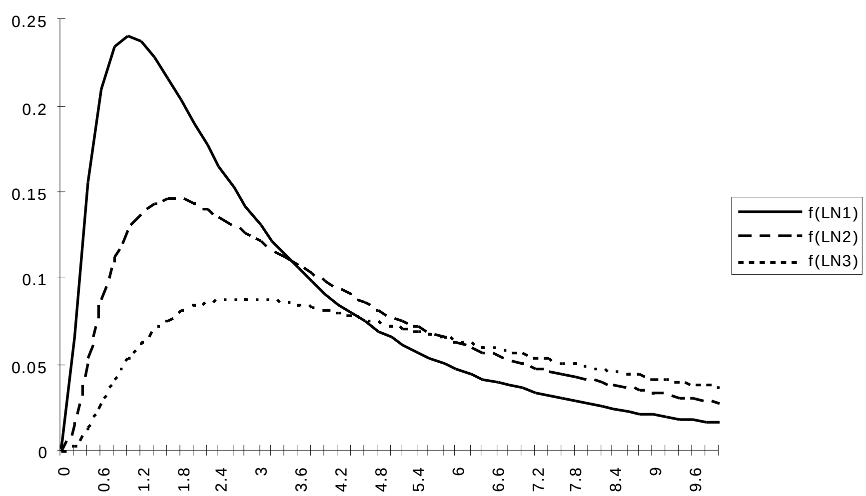


Figure 8. Lognormal Distribution

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

$$X = exp(Z).$$

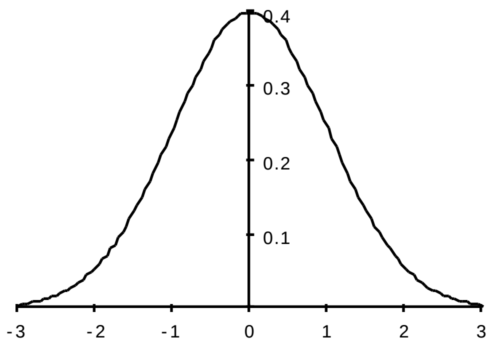
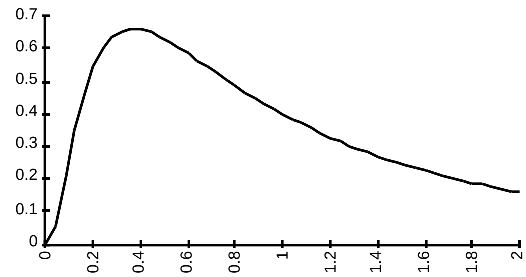
<table>
<tr><td>Random variable Z
with normal distribution;
mean μ_Z and variance σ_Z^2</td><td>Random variable X
with lognormal distribution;
mean μ_X and variance σ_X^2</td></tr>
<tr><td>
</td><td>
</td></tr>
<tr><td>$\mu_Z = 0$ and $\sigma_Z^2 = 1$ for the case shown</td><td>$\mu_X = 1.6487$ and $\sigma_X^2 = 4.6708$</td></tr>
<tr><td>$Z = ln(X)$</td><td>$X = exp(Z)$</td></tr>
</table>

Figure 9. Relation between normal and lognormal distributions

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 pdf of the lognormal distribution is given in terms of α and β. This expression is useful for plotting the pdf, but not for computing probabilities since it cannot be integrated in closed form. The parameters of the three cases given in 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 gives

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

The lognormal 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 described 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.

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

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

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

Exponential Distribution

Example 13. Telephone calls arrive at a switch board every 30 seconds on average. Because each is initiated independently, they arrive at random. We would like to know the probability that the time between one call and the next is greater than 1 minute?

Exponential Distribution

Parameter: λ

$f(x) = \lambda\, exp(-\lambda x)$ for $x \geq 0$

$F(x) = 1 - exp(-\lambda x)$ for $x \geq 0$

$\mu = \sigma = 1/\lambda$ 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 which is characterized by the single positive parameter λ. The density has the shape shown in Fig. 10.

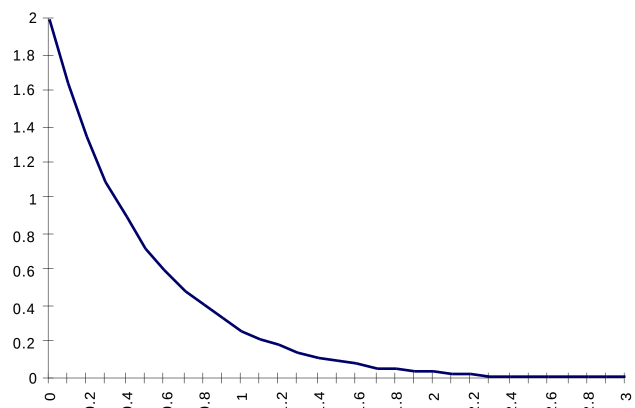


Figure 10. Exponential pdf

Integrating the pdf yields a closed form expression for the CDF. 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, we know that the mean time between arrivals is 30 seconds. The average 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 the time between arrivals of some process is governed by the exponential distribution with rate λ, the number of arrivals in a fixed interval T is governed by the Poisson distribution with mean value λT. Such a process is called a Poisson process.

Gamma Distribution

Example 14: Consider a simple manufacturing operation whose completion time has a mean equal to 10 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, $\Gamma(r) = (r-1)!$

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

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.

Figure 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 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.

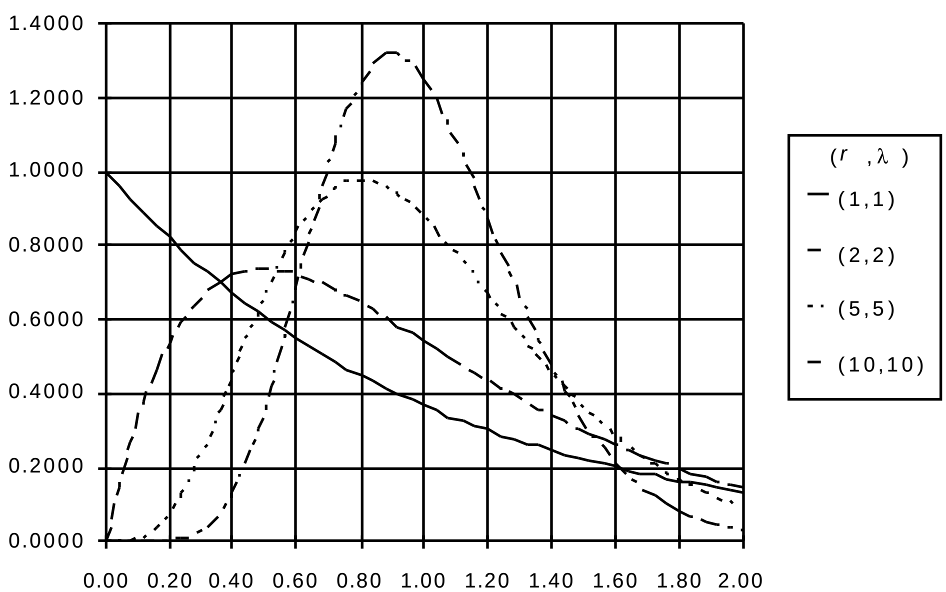


Figure. 11. Several gamma density functions

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 is the CDF of the Poisson with parameter λx.

In the example 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 -- a special case of the gamma with r equal to 1. The given mean, 10, provides us with the information necessary to compute the parameter λ.

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

For $r = 1$, $\lambda = 0.1$ we get $\mu = 10$, $\sigma = 10$, and $P(X < 11) = 0.6671$

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

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

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 and Uniform Distributions

Example 15. Consider the following hypothetical situation. Grades in a senior engineering class indicate that on average 27% of the students received an A. 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 $\beta > 0$

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

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

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

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.

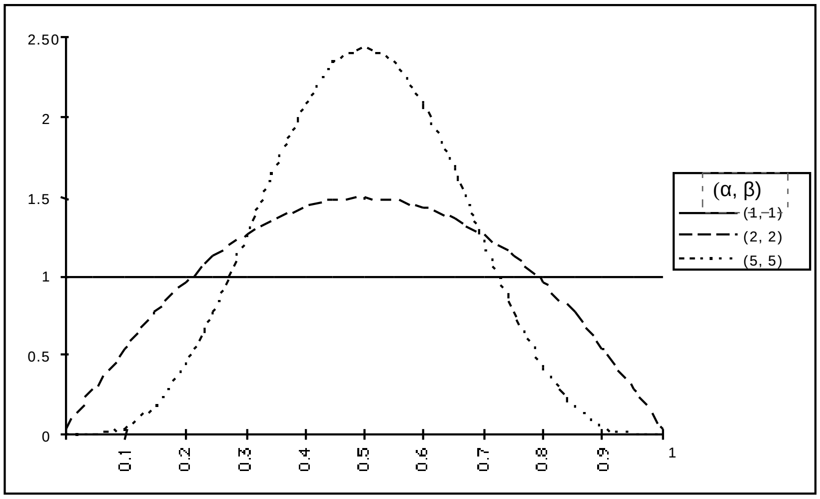


Figure 12. The symmetric case for the beta Distribution

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.

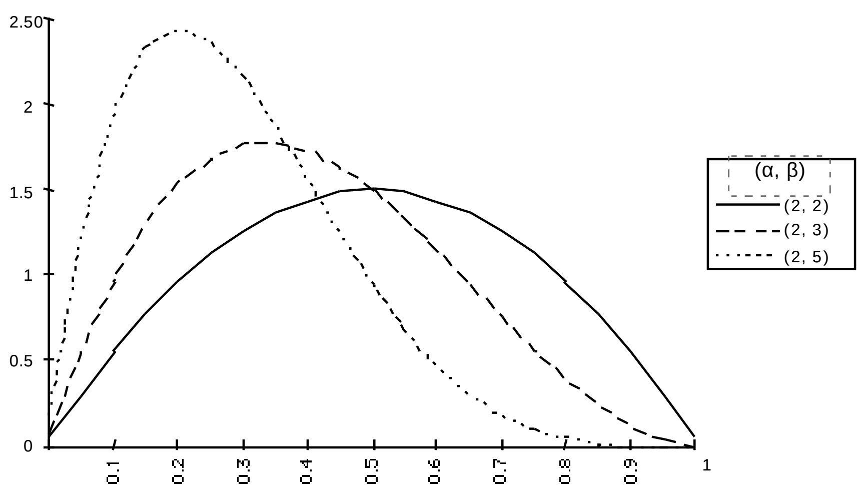


Figure 13. Beta distribution as β increases for constant α

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

Generalized Beta Distribution
Parameters: $\alpha > 0$, $\beta > 0$, a, b where $a < b$
$\mu = a + (b - a)\left(\frac{\alpha}{\alpha + \beta}\right)$
$\sigma^2 = \frac{(b-a)^2\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}$
$mode = a + (b-a)\left\{\frac{\alpha - 1}{\alpha + \beta - 2}\right\}$

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

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

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 0 and 1. The example gives the average proportion of A's as 0.27. The standard deviation is specified 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 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 the grades in a class re more than 50% A. An Excel spreadsheet function provides the values of the CDF for the given parameters $(\alpha, \beta) = (2, 5.4074)$. The results indicate that $P\{X > 0.5\} = 0.0873$.

Weibull Distribution

Example 16: A truck tire has a mean life of 20,000 miles. A conservative owner decides to always replace a 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 parameter $\beta = 2$.

Weibull Distribution

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

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

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

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

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

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

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.

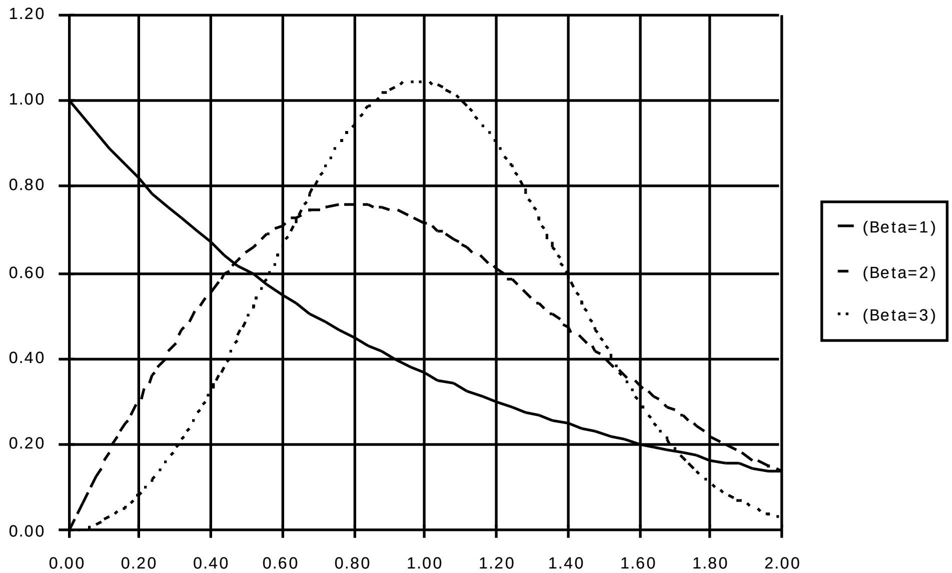


Figure 14. Weibull pdf

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

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

when $X = x$. 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

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

For β =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 $\beta = 2$, 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 $\beta < 1$, 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). Because only 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

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

Computing the required probability with these parameters gives

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

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 CDF this is easily accomplished.

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

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

Translations of a Random Variable

Example 17: An investment in an income producing property is $1,000,000. In return for that investment we expect revenue for each of the next 5 years. 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

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

With the initial investment in the amount I the net present worth (NPW) for a property that lasts n years is

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

A NPW greater than 0 indicates that the rate of return of the investment is greater than the MARR and is therefore acceptable. 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.

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

The mean, variance and mode of Y are, respectively,

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

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

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

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: (i) the sum of independent random variables with normal distributions is also normal, (ii) the mean of the sum is the sum of the means, and (iii) 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%, we have P(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, a distribution must be selected and its 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 preferred 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 logically restricted to be nonnegative. Time is a prime example of this situation. The lognormal, gamma and Weibull distributions are some examples that might be appropriate to model time. These distributions also exhibit asymmetry that might be important in practice. The Weibull distribution is important in 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 can be tried. The uniform distribution has only the bounds as parameters, and might be used when 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 a large variety of shapes possible with a suitable selection of parameters.

21.5 Simulating Random Variables

Depending on the application, we often create probability models of certain aspects of systems. For simple cases, we compute probabilities of events using probability theory. There are many situations, though, that are affected by more than one random variable. Because of their complexity, many problems will not yield to easy analysis. Rather, it is necessary to *simulate* the several random variables that describe the situation, and observe and analyze their effects with the help of statistics. The approach is widely used in operations research. We provide an elementary discussion of simulation in this section and add depth in Chapter 18 . In addition to its power as a tool for analyzing complex systems, simulation provides a medium for illustrating the ideas of variability, uncertainty and complexity. We construct a simulated reality and perform experiments on it.

A Production Process

Example 18: Consider an assembly process with three stations. Each station produces a part, and the three parts are assembled to produce a finished product. The stations have been balanced in terms of work load so that each has an average daily production equal to 10 units. The production of each station is variable, however, and we have established that the daily rate has a discrete uniform distribution ranging from 8 to 12. Work in process cannot be kept from one day to the next, so the output of the line is equal to the minimum production rate of the three stations. Let X_1, X_2 and X_3 be random variables associated with the production at the three stations, respectively. The daily production of the line is a random variable, Y, such that $Y = \text{Min}\{X_1, X_2, X_3\}$. Our interest is learning about the distribution of the random variable Y.

The dedicated student may be able to derive the PDF of Y given this information; however, we will take the conceptually easier approach of simulation. We observe in Table 15 ten days of simulated operation. The production in each station is simulated[3] using the methods described later in this chapter. With these observations we compute the simulated production of the line as

$$y = \text{Min}\{x_1, x_2, x_3\}.$$

Table 15. Simulated observations of production line

Random variate	Day 1	2	3	4	5	6	7	8	9	10
x_1	11	10	8	11	9	12	8	11	12	9
x_2	12	10	10	12	8	9	8	9	10	9
x_3	9	8	8	9	9	12	12	9	10	10
y	9	8	8	9	8	9	8	9	10	9

The simulation provides information about the system output. The statistical mean, variance, and standard deviation are

$$\bar{y} = 8.7,\ s^2 = 0.4556,\ s = 0.6749.$$

[3]The observations were simulated by x_i = ROUND(7.5 + 5r). This provides a discrete uniform distribution with integer values between 8 and 12.

We conclude that the average daily production of the system is well below the average of the individual stations. The cause of this reduction in capacity is variability in station production.

Because the statistics depend on the specific simulated observations, the analyst usually runs replications of the experiment to learn about the variability of the simulated results. The table below shows six replications of this experiment. Although the results vary, they stay within a narrow range. Our conclusion about the reduction in the system capacity is certainly justified.

Replication	1	2	3	4	5	6
$\bar{y}$	8.7	8.6	9.1	9	8.8	8.8

Sample Statistics

mean: $\bar{y} = \sum_{i=1}^{n} y_i \Big/ n$

sample variance:

$$s^2 = \sum_{i=1}^{n} (y_i - \bar{y})^2 \Big/ (n-1)$$

standard deviation: $s = \sqrt{s^2}$

The grand average and the standard deviation of the six observations is 8.833 and 0.186, respectively. Since each of the replication averages are determined by the sum of 10 numbers, we can conclude from the central limit theorem that the average observations are approximately normally distributed. This allows a number of additional statistical tests on the results of the simulation replications.

Simulation is a very powerful tool for the analysis of complex systems involving probabilities and random variables. Virtually any system can be simulated to obtain numerical estimates of its parameters when its logic is understood and the distributions of its random variables are known. The approach used extensively in the latter chapters of the book to illustrate the theoretical concepts, as well as to provide solutions when the theoretical results are either not available or too difficult to apply.

Simulating with the Reverse Transformation Method

The conceptually simplest way to simulate a random variable X with density function $f_X(x)$, is called the reverse transformation method. The cumulative distribution function is $F_X(x)$.

Let R be a random variable taken from a uniform distribution ranging from 0 to 1, and let r be a specific observation in that range. Because R is from a uniform distribution

$$P\{R \le r\} = r.$$

Let the simulated value x_S be that value for which the CDF, $F(x_S)$, equals r.

$$r = F(x_S) \text{ or } x_S = F^{-1}(r).$$

The random observation is the reverse transformation of the CDF, thus providing the name of the method. The process is illustrated for the example in Fig. 14 where we show the CDF for the production at a station in Example 18. Say we select the random number 0.65 from a continuous uniform distribution. We locate 0.65 on the $F(x)$ axis, project to the cumulative distribution function (the dotted line), and find the value of the random variable x_s for which $F(x_s) = 0.65$. For this instance, $x_s = 11$. Table 16 shows the random numbers that provided the observations of Table 15.

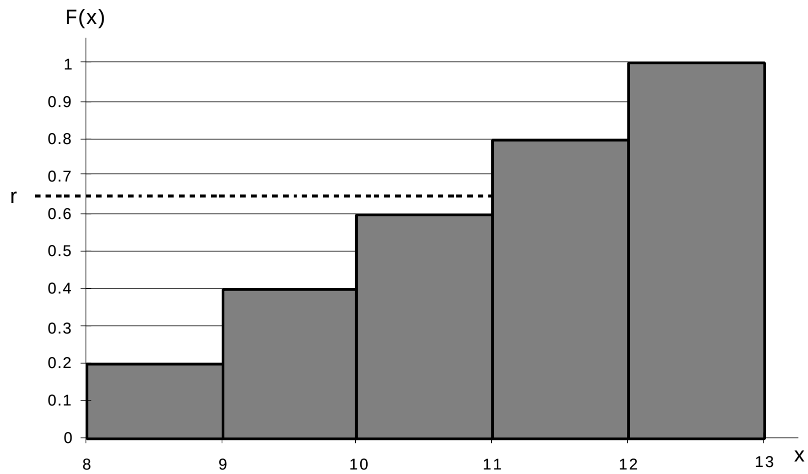


Figure 14. Reverse transformation method for a discrete distribution

Table 16. Random numbers for 10 days

Random variate	Day 1	2	3	4	5	6	7	8	9	10
x_1	0.6151	0.5569	0.18	0.7291	0.3046	0.869	0.1611	0.6633	0.9535	0.3298
x_2	0.861	0.5541	0.4883	0.983	0.1391	0.251	0.1389	0.3261	0.5123	0.3793
x_3	0.3108	0.0923	0.141	0.3124	0.2558	0.9785	0.9638	0.3351	0.5778	0.4323

Discrete Random Variables

For a second illustration of the reverse transformation method for a discrete random variable, consider the binomial distribution with $n = 3$ and $p = 0.9$. The PDF is

$$P_X(k) = \binom{3}{k}(0.9)^{k}(0.1)^{3-k} \text{ for } k = 0,1,2,3.$$

Table 17 shows the probabilities associated with this distribution.

Next we define intervals on the real number line from 0 to 1 for the possible values of the random variable. The intervals are computed using the cumulative distribution. In general, when r is a real number and x is a finite discrete random variable, the intervals are

$$I(k) = \{r \mid F_X(k-1) < r < F_X(k)\} \quad k = 0,1,2,\ldots,n.$$

For the example case the intervals are shown in Table 17.

Table 17. Simulation intervals for binomial distribution with $n = 3$, $p = 0.9$.

k	0	1	2	3
$P(k)$	0.001	0.027	0.243	0.729
$F(k)$	0.001	0.028	0.271	1.000
$I(k)$	0 - 0.001	0.001 – 0.028	0.028 – 0.271	0.271 – 1.0

To simulate an observation from the distribution, we select a random number and determine the simulation interval in which the random number falls. The corresponding value of k is the simulated observation. For instance, using the random number 0.0969, we find that 0.0969 is in the range $I(2)$. The simulated observation is then 2. In general, let r_i be a random number drawn from a uniform distribution in the range (0, 1). The simulated observation is

$$x_i = \{k \mid F_X(k-1) < r_i < F_X(k)\}.$$

Six simulated values are shown in Table 18.

Table 18. Simulated values for binomial distribution

Random number	0.0969	0.2052	0.0013	0.2637	0.6032	0.5552
Simulated observation	2	2	1	2	3	3

Based on a sample of 30 observations, the sample mean, variance and standard deviation are computed as shown in Table 19. We also estimate the probabilities of each value of the random variable from the proportion of times that value appears in the data. The sample results are compared to the population values in the table.

Table 19. Distribution information estimated from simulation

Parameter	Estimated from sample	Population value
Mean	2.667	2.7
Variance	0.2989	0.27
Standard deviation	0.5467	0.5196
$P(0)$	0	0.001
$P(1)$	0.0333	0.027
$P(2)$	0.2667	0.243
$P(3)$	0.7	0.729

The statistics obtained from the sample are estimates of the population parameters. If we simulate again using a new set of random numbers we will surely obtain different estimates. Comparing the statistics to the population values we note that the statistical mean and variance are reasonably close to the parameters they estimate. We did not observe any values of 0 from the sample. This is not surprising as the probability of such an occurrence is one in a thousand, and there were only 30 observations.

Continuous Random Variables

Example 19: A particular job consists of three tasks. Tasks A and B are to be done simultaneously. Task C can begin only when both tasks A and B are complete. The times required for the tasks are T_A, T_B, and T_C respectively, and all times are random variables. T_A has an exponential distribution with a mean of 10 hours, T_B has a uniform distribution that ranges between 6 and 14 hours, and T_C has a normal distribution with a mean of 10 hours and a standard deviation of 3 hours. The time to complete the project, Y, is a random variable that depends on the task times as follows.

$$Y = \mathrm{Max}\{ T_A, T_B \} + T_C$$

What is the probability that the promised completion time of 20 hours is met?

An analytic solution to this problem would be difficult because the completion time is a complex combination of random variables. Simulation provides statistical estimates concerning the completion time.

Table 20 shows the results for 10 simulations. The rows labeled r provide the random numbers used to simulate the random variables and the rows labeled t_A, t_B and t_c show the associated observations. The row labeled y is computed using the system equation and the observations in each column. The process of producing the simulated observations is described presently.

Table 20. Simulated results for example

Random	Run number									
variate	1	2	3	4	5	6	7	8	9	10
r	0.663	0.979	0.256	0.312	0.141	0.092	0.311	0.251	0.139	0.983
t_A	10.89	38.41	2.954	3.745	1.52	0.968	3.722	2.89	1.497	40.76
r	0.953	0.33	0.139	0.326	0.512	0.379	0.964	0.335	0.578	0.432
t_B	13.63	8.638	7.111	8.609	10.1	9.034	13.71	8.681	10.62	9.459
r	0.731	0.615	0.557	0.18	0.729	0.305	0.869	0.861	0.554	0.488
t_C	11.85	10.88	10.43	7.253	11.83	8.466	13.36	13.25	10.41	9.912
y	25.48	49.28	17.54	15.86	21.93	17.5	27.07	21.94	21.03	50.67

It is clear from the 10 completion times in Table 20, that there is quite a bit of variability in an estimate of the time to complete the project. There are 3 numbers below 20, so an initial estimate of the probability of a completion time less than 20 would be 0.3. A much larger sample would be necessary for an accurate determination.

Closed Form Simulation

The reverse transformation method can be used to simulate continuous random variables but the probabilities do not occur at discrete values. Rather we have the CDF that indicates the probability that the random variable falls with a specified range.

$$P\{X \le x\} = F(x)$$

The process for generating random variates follows.

- Select a random number r from the uniform distribution.
- Find the value of x for which $r = F(x)$ or $x = F^{-1}(r)$.
- The resulting x is the simulated value.

For some named distributions we have a closed form functional representation of the CDF, while for others we must use a numerical procedure to find the inverse probability function. When we have a closed form expression that can be solved for x given r, the simulation procedure is easily performed. For example, consider the exponential distribution for task A. The mean of 10 hours yields the parameter $\lambda = 0.1$. The general form of the pdf is

$$f(x) = \lambda e^{-\lambda x} \text{ and } F(x) = 1 - e^{-\lambda x} \text{ for } x \ge 0.$$

Setting the cumulative distribution equal to r and solving we obtain

$$r = 1 - e^{-\lambda x} \text{ or } x = -(1/\lambda)\ln(1 - r).$$

The first line labeled r in Table 20 is used to simulate T_A, shown in the following line. For example, when

$$r = 0.663,\ t_A = -10\ln(1 - 0.663) = 10.89.$$

The uniform distribution associated with T_B is also simulated with a closed form expression. The CDF for a uniform distribution ranging from a to b is

$$F(x) = (x - a) / (b - a) \text{ for } a \le x \le b.$$

Setting r equal to the CDF and solving form x yields the simulated value

$$x = a + r\,(b - a).$$

To illustrate, we compute the first simulated value of T_B using the random number $r = 0.953$.

$$t_B = 8 + 0.953(14 - 8) = 13.63.$$

The Weibull distribution also allows a closed form simulation. We set the expression for the CDF equal to the random number

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

Solving for x, we obtain the expression for the simulated random variable

$$x = \sqrt[\beta]{\frac{-ln[1 - r]}{\alpha}}.$$

We recognize that when r is a random number from the range 0 to 1, $(1 - r)$ is also a random number from the same range. This simplifies the expression for the simulated observation to

$$x = \sqrt[\beta]{\frac{-ln(r)}{\alpha}}.$$

Simulation using Inverse Probability Functions

When a closed form expression is not available for the cumulative distribution, numerical methods may be used to compute the inverse probability function. For the example we used the inverse normal distribution function available with the Excel spreadsheet program to simulate the observations from T_C. First we must find

$$t_C = F^{-1}(r)$$

where $F^{-1}(r)$ is the inverse probability function for the normal distribution with mean 10 and standard deviation 3 evaluated for the value r. Figure 15 illustrates the reverse transformation in the case where a random number equal to 0.731 yields t_C equal to 11.85.

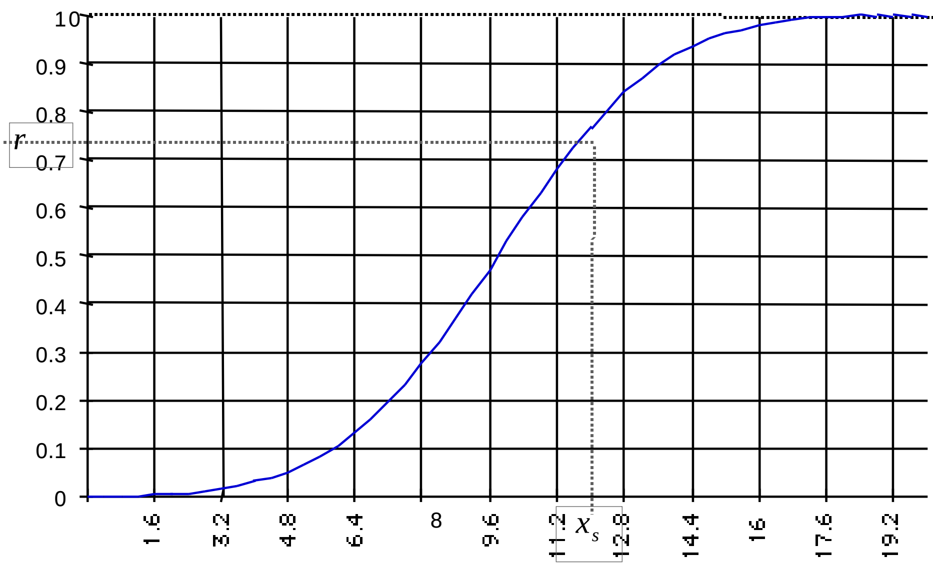


Figure 15. Reverse transformation method for a normal distribution

We complete the first column in Table 20 by computing y with the equation defining the interaction between the random variables.

$$y = \text{Max}\{10.89, 13.63\} + 11.85 = 25.48.$$

Using Simulation to Solve Problems

One of the main applications of simulation is to analyze problems involving combinations of random variables that are too complex to be conveniently analyzed by probability theory. It is a very powerful tool because virtually any system can be simulated to obtain numerical estimates of its parameters when its logic is understood and the distributions of its random variables are known. The approach is used extensively in the practice of operations research.

Simulation does have its drawbacks, however. Except for very simple systems, simulation analysis is computationally expensive. Inverse probability functions that do not have a closed form may be difficult to compute and complex systems with many random variables will burn up computer time. Simulation results are statistical. One selection of random variables yields a single observation. Because most situations involve variability, it is necessary to make a large number of observations and use statistical techniques to draw conclusions. Simulation results are sometimes difficult to interpret. This is partly due to the complexity of many simulated situations and partly due to their inherent variability. Effects may be observed, but causes are difficult to assign.

21.6 Exercises

1. From the daily newspaper identify five quantities that are variable in time and uncertain for the future. Provide data or charts that describe the historical variability. Describe decisions that are affected by the values of the quantities.

2. A group of six students draws tickets to the football game. The tickets are for six adjacent seats all on the same row. One of the guys and one of the gals have a secret affection for each other. If the tickets are passed out at random, what is the probability that they will sit next to each other?

3. A repeatable experiment involves observing two numbers a and b. The possible number pairs (a, b) include

$$\boldsymbol{S} = \{(1,2), (1,3), (1,4), (2,1), (2,2), (2,3), (2,4), (3,1), (3,2), (3,4), (4,1), (4,2)\}.$$

   Each of the 12 elements in the set $\boldsymbol{S}$ are equally likely. For the following events

$$\boldsymbol{W} = \{(a, b) | a = 1\}, \boldsymbol{X} = \{(a, b) | b < 2\}, \boldsymbol{Y} = \{(a, b) | a + b \le 4\}, \boldsymbol{Z} = \{(a, b) | b = 2\}$$

   find the probabilities of $\boldsymbol{W}$, $\boldsymbol{X}$, $\boldsymbol{Y}$, $\boldsymbol{Z}$, $\boldsymbol{W} \quad \boldsymbol{Y}$, $\boldsymbol{W} \cap \boldsymbol{Z}$, $\boldsymbol{W} \quad \boldsymbol{X} \quad \boldsymbol{Y} \quad \boldsymbol{Z}$.

4. This problem uses the events $\boldsymbol{W}$, $\boldsymbol{X}$, $\boldsymbol{Y}$ and $\boldsymbol{Z}$ defined in previous exercise. Again all outcomes are equally likely. Consider in turn each of the six pairs that can be formed from these events. For each pair below, indicate whether the events are mutually exclusive or independent.

   Pairs: $\boldsymbol{W}$ and $\boldsymbol{X}$, $\boldsymbol{W}$ and $\boldsymbol{Y}$, $\boldsymbol{W}$ and $\boldsymbol{Z}$, $\boldsymbol{X}$ and $\boldsymbol{Y}$, $\boldsymbol{X}$ and $\boldsymbol{Z}$, $\boldsymbol{Y}$ and $\boldsymbol{Z}$.

5. Five events -- **A**, **B**, **C**, **D**, **E**-- are defined on a sample space of a random experiment. Events **A**, **B** and **C** are mutually exclusive, and you are given the following probabilities.

$$P(\mathbf{A}) = 0.1,\ P(\mathbf{B}) = 0.4,\ P(\mathbf{C}) = 0.5,\ P(\mathbf{D}) = 0.16,\ P(\mathbf{D} | \mathbf{A}) = 0.8,$$
$$P(\mathbf{D} | \mathbf{B}) = 0.2,\ P(\mathbf{D} | \mathbf{C}) = 0,\ P(\mathbf{E} | \mathbf{A}) = 0,\ P(\mathbf{E} | \mathbf{B}) = 0.9,\ P(\mathbf{E} | \mathbf{C}) = 0.1$$

   Compute $P(\mathbf{D} \cap \mathbf{A})$, $P(\mathbf{B} | \mathbf{D})$, $P(\mathbf{D} \quad \mathbf{B})$, $P(\mathbf{E})$, $P(\mathbf{B} | \mathbf{E})$.

6. A quality control test involves removing a sample of three products from the production line in each hour. They are then is tested to determine whether or not they meet specifications. The random variable for this situation is the number of units that meet specifications. Define the following events.

   $\boldsymbol{A}$: at most 2 meet the specification

   $\boldsymbol{B}$: exactly 3 meet the specifications

   $\boldsymbol{C}$: at least 2 meet the specifications

   a. Which events are mutually exclusive and which are not?

   b. In terms of the random variable, what are the events $\boldsymbol{A} \quad \boldsymbol{B}$, $\boldsymbol{A} \quad \boldsymbol{C}$, and $\boldsymbol{A} \cap \boldsymbol{C}$.

c. Say each unit meets specifications with probability 0.95 and that the units are independent. Assign a probability to each event (and the combinations of events) considered in parts *a* and *b*.

7. The random variable X is the number of cars entering the campus of Big State University from 1:00 AM to 1:05 AM. Assign probabilities according to the formula:

$$P(X = k) = \frac{e^{-5}(5^k)}{k!} \text{ for } k = 0, 1, 2, \ldots$$

a. For the events

$\boldsymbol{A}$: More than 5 cars enter

$\boldsymbol{B}$: Fewer than 3 cars enter

$\boldsymbol{C}$: Between 4 and 8 cars enter

find $P(\boldsymbol{A})$, $P(\boldsymbol{B})$, $P(\boldsymbol{C})$, $P(\boldsymbol{A} \quad \boldsymbol{B})$, $P(\boldsymbol{A} \quad \boldsymbol{C})$, $P(\boldsymbol{B} \quad \boldsymbol{C})$, $P(\boldsymbol{A} \quad \boldsymbol{B} \quad \boldsymbol{C})$, $P(\boldsymbol{A} \cap \boldsymbol{B})$,

$P(\boldsymbol{A} \cap \boldsymbol{C})$, $P(\boldsymbol{B} \cap \boldsymbol{C})$, $P(\boldsymbol{A} \cap \boldsymbol{B} \cap \boldsymbol{C})$, $P(\boldsymbol{A} \mid \boldsymbol{B})$, $P(\boldsymbol{C} \mid \boldsymbol{A})$, $P(\boldsymbol{A} \mid \boldsymbol{C})$.

b. Compute the CDF associated with the random variable.

8. Assign probabilities to an event $\boldsymbol{E} = \{t \mid t_1 < t < t_2\}$ with the function

$$P(\boldsymbol{E}) = exp(-t_1) - exp(-t_2)$$

Define $\boldsymbol{A} = \{t \mid 0 < t < 1\}$, $\boldsymbol{B} = \{t \mid 1.5 < t < \infty\}$ and $\boldsymbol{B} = \{t \mid 0.5 < t < 1.5\}$

Find $P(\boldsymbol{A})$, $P(\boldsymbol{B})$, $P(\boldsymbol{C})$, $P(\boldsymbol{A} \quad \boldsymbol{B})$, $P(\boldsymbol{A} \quad \boldsymbol{C})$, $P(\boldsymbol{B} \quad \boldsymbol{C})$, $P(\boldsymbol{A} \quad \boldsymbol{B} \quad \boldsymbol{C})$, $P(\boldsymbol{A} \cap \boldsymbol{B})$,

$P(\boldsymbol{A} \cap \boldsymbol{C})$, $P(\boldsymbol{B} \cap \boldsymbol{C})$, $P(\boldsymbol{A} \cap \boldsymbol{B} \cap \boldsymbol{C})$, $P(\boldsymbol{A} \mid \boldsymbol{B})$, $P(\boldsymbol{C} \mid \boldsymbol{A})$, $P(\boldsymbol{A} \mid \boldsymbol{C})$.

9. Consider the situation of throwing two dice (each die has sides numbered 1 through 6). For each of the following random variables find $P(3 \le X \le 5)$.

a. X = sum of the two dice

b. X = number on the first die

c. X = average of the two dice

d. X = largest of the two numbers on the dice

10. Hillary makes $15 each Saturday night that she babysits. Based on past experience, she computes that there is an 80% chance that she will be called for a job in one week. The weeks are independent. Hillary wants to save $90 for a new dress. Assuming that she saves all the money she makes, what is the probability that more than eight weeks are required to save the $90?

11. Computations indicate that a missile fired at a target will destroy the target 35% of the time. Assume all missiles fired are independent and have the same probability of

destruction. How many missiles must be fired to assure that the target will be destroyed with a 0.90 probability?

12. The kicker on the football squad has a 60% chance of successfully kicking a field goal on each try. If he is given five opportunities in a game, what is the probability that he will complete at least four field goals in the game?

13. Ten students in a class of 30 are girls. On the first midterm, girls score four of the top five grades. What is the probability that this result could have occurred purely by chance?

14. During the preregistration period, students arrive at random to a professor's office at an average rate of 5 per hour. What is the probability that fewer than two students will arrive during a 10- minute period?

15. We select an item from the assembly line and note that it is either operating or failed. The associated probabilities are P(operating) = 0.8 and P(failed) = 0.2. For each situation below identify the probability distribution, specify the parameters, and answer the question.

    a. We draw 10 such items from the line. What is the probability that a majority of them will work?

    b. What is the probability that at least 10 items must be taken from the line before we find five that are working?

    c. What is the probability that we will draw 5 working items before the first failed one is encountered?

16. In a standard deck of 52 cards there are 13 different card types labeled as follows: ace, 2, 3, 4, 5, 6, 7, 8, 9, 10, J, Q, K. Each card type appears in 4 suits: clubs, diamonds, hearts and spades.

    a. You are about to be dealt a hand of five cards from the pack of 52 (this is selection without replacement). What is the probability that at least three of the cards will be aces?

    b. The dealer will show you five cards. After each card is drawn, it is returned to the deck (this is selection with replacement). What is the probability that at least three of the cards shown are aces?

    c. The dealer will show you a series of cards off the top of the deck. After each card is displayed, it is returned to the deck (this is selection with replacement). The dealer will continue showing you cards until 3 aces have appeared. What is the probability that five or fewer cards must be shown?

17. An OR professor always gives tests made up of 20 short answer questions. Five points are given for each correct answer so a maximum total score of 100 is possible. She grades the questions as either right or wrong with no partial credit. Based on your current knowledge of the subject, you judge that you have a 0.7 probability of answering any one question correctly. Assume the questions are independent in

terms of chance of success. If a score of 80 or more earns at least a B, what is the probability that you will earn at least a B for this exam?

18. An accounting professor always gives tests made up of 20 short answer questions. Five points are given for each correct answer so a maximum total score of 100 is possible. No question may receive partial credit – it is either right or wrong. Based on your understanding of the material in the course, you judge that your mark on each question is a random variable with mean equal to 3 and a standard deviation equal to 1. Assume that the questions are independent in terms of chance of success. Before you take the test what is your estimate of the mean and standard deviation of your final score? If a score of 80 or more earns at least a B, what is the probability that you will earn a B for this exam? You will have to make an approximation to answer this question.

19. A major manufacturing company needs at least three design engineers for an upcoming project, but will take more. The manager in charge of hiring will conduct a series of interviews in hopes of finding at least three "good" applicants who will accept a job offer. The probability that an applicant is "good" is 0.4, and the probability that a "good" applicant will accept the job is 0.6. She makes offers to every "good" applicant. In answering the following questions, assume that the applicants are independent.

    a. If the manager conducts 10 interviews, what distribution describes the number of engineers she will hire? What is the probability that at least three will be hired?

    b. If the manager conducts as many interviews as necessary to find the three engineers, what distribution describes the number of interviews required? What is the probability that at least 10 interviews will be required?

20. A robotic insertion device must place 4 rivets on a printed circuit board. Rivets are presented to it at random, but 10% of the supply are faulty. When a faulty rivet is encountered it must be discarded, and another drawn. We want to compute probabilities regarding the number of draws required before 4 good rivets are found. In particular, what is the probability that exactly 10 draws are required before we find 4 good rivets.

21. The Department of Industrial Engineering at a local university has six full professors, four associate professors, and five assistant professors. A committee with five members is chosen at random from all the professors. What is the probability that a majority of the committee will be full professors.

22. A box contains three red dice, two blue dice and one green die. For the following experiments describe the probability distribution function for the random variable.

    a. Three dice are drawn from the box with replacement. The random variable is the number of red dice drawn. What is the probability that at least 2 red dice are drawn?

    b. Three dice are drawn without replacement. The random variable is the number of blue dice drawn. What is the probability that at exactly 2 blue dice are drawn?

c. Draw a die out of the box. Observe its color. Return the die to the box. Continue drawing dice in this manner until the green die is found or five dice have been drawn. The random variable is the number of red or blue dice drawn before the experiment stops. What is the probability that the process stops in three or fewer draws.

23. The game of Russian Roulette uses a gun which has a single live bullet placed in one of the six chambers in the cylinder. Player *A* spins the cylinder, places the gun to his head and pulls the trigger. If player *A* survives, player *B* repeats the process. The game continues with the players taking turns until one finds the bullet.

a. Compute the probability distribution of the number of times the trigger will be pulled until a player is killed. The last pull kills one of the players.

b. Find the probability that the game lasts longer than six rounds.

c. If the players pledge to stop after six rounds, what is the chance that each player will be killed?

24. A line forms in front of a ticket booth with one window. Observing the number of persons in the line (including the one being served), you estimate the probability for *k* persons in line, p_k, to be the values given below. You never observe more than four persons in the line.

k	0	1	2	3	4
p_k	0.3278	0.2458	0.1844	0.1383	0.1037

a. Compute the mean value and standard deviation of the number of persons in the line.

b. What is the probability that the server is busy?

25. Find the values of the mean and standard deviations for the following distributions:

a. $f(x) = 1/4$ for $x = 1, 2, 3, 4$.

b. The random variable X has a binomial distribution with $n = 10$ and $p = 0.6$.

c. $f(x) = \dfrac{e^{-4}(4)^x}{x!}$ for $x = 0, 1, 2, \ldots$

26. A continuous random variable has the pdf

$$f(x) = \begin{cases} 1 \text{ for } 0 \le x \le 1 \\ 0 \text{ elsewhere} \end{cases}.$$

Derive the formulas for the CDF, mean, variance, and median.

27. Consider a continuous random variable with the pdf

$$f(x) = \begin{cases} 0.5 & \text{for } 0 \le x < 0.5 \\ k & \text{for } 0.5 \le x \le 1 \\ 0 & \text{elsewhere} \end{cases}$$

What is the value of k that makes this a legitimate pdf? Find an expression for the CDF for this random variable.

28. Consider a continuous random variable Y with the pdf

$$f(y) = \begin{cases} k(y-10) & \text{for } 10 \le x \le 20 \\ 0 & \text{elsewhere} \end{cases}.$$

a. Find the value of k that makes this a pdf. Find an expression that describes the CDF, $F(y)$, for all y.

b. Compute the mean, standard deviation, median, mode of the distribution.

c. From the CDF find $P\{Y < 12\}$, $P\{12 < Y < 18\}$.

29. Consider the previous exercise by now express Y as a linear transformation of a random variable X (i.e., $Y = a + bX$), where X has a triangular distribution that ranges from 0 to 1.

a. What is c for the distribution of X?

b. What are a and b for the transformation?

c. Compute the mean, standard deviation, median, mode of Y from the corresponding values for X.

d. From the CDF of X find $P\{Y< 12\}$, $P\{12 < Y < 18\}$.

30. For the accompanying probability distribution, find

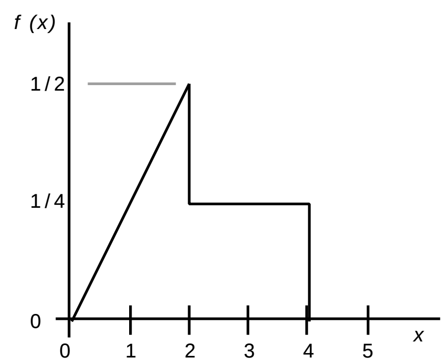
a. $P(1 \le X \le 3)$

b. the mean

c. the variance

d. the median

e. the mode



31. A manufacturer of men's bathrobes makes only one size which is designed to fit any man with a chest size more than 39 inches and less than 43 inches. Using data gathered by the U.S. Dept. of Statistics on Interesting Phenomena, the company determines that the mean chest size of men in the United States is 41.5 inches. The standard deviation is 1.75 inches. Assuming that chest size has a normal distribution, what proportion of the population *will not* fit into the robe?

32. Consider the information given in the previous exercise with the additional caveat that the robe will not fit men who are less than 68 inches tall or more than 75 inches tall. Assume that the height of men has a normal distribution with mean 70 inches and a standard deviation of 3 inches. For simplicity assume also that height and chest size are independent random variables. What proportion of the men *will* fit into the robe considering both restrictions?

33. A grocery store clerk must pack n items in a bag. The average time to pack an item is 1 second with a standard deviation of 0.5 seconds. What is an appropriate distribution for modeling the total time required to pack the bag (some assumptions are needed).

34. Uncharacteristically, Martha Stewart is having difficulty cutting a pie into six even pieces. Starting from a given point she can will cut a slice of pie with a central angle α. The central angle is a normal distributed random variable with a mean of 60 degrees and a standard deviation of 5 degrees. The first five pieces are cut with this accuracy. The last piece is what remains.

    a. If the Stewart family judges as acceptable a piece with an angle between 55 and 65 degrees what is the probability that each of the first five pieces is acceptable?

    b. What is the probability that the final piece is acceptable?

    c. What is the probability that all six pieces are acceptable?

35. Consider a restaurant open only for the lunch hour. The number of persons, Z, seeking service has a normal distribution with mean μ_Z and standard deviation σ_Z. Because of congestion it has been found that the time for the restaurant to clear after the lunch hour, X, is an exponential function of the number of customers. That is

    $$X = a \cdot e^{bZ}, \text{ where } a \text{ and } b \text{ are constants.}$$

    What is the probability distribution for the random variable X? In terms of μ_Z, σ_Z, a and b, what are the mean and standard deviation of X?

36. A manufacturing operation has a mean time of 10 minutes. We have the option of dividing the operation into 1, 2, 5 or 10 components. The components must be processed sequentially and the entire operation is complete only when all the components are finished. When the number of components is n, the mean time to complete a component is $10/n$. The time to complete each component has an exponential distribution. Compute the probability that the total completion time, Y, is between 9 and 11 minutes. Hint: make use of the gamma distribution.

37. Use the standard normal distribution to approximate the case of 10 components in the previous exercise. Compute the $P(9 \le Y \le 11)$ and compare it with the value computed with the gamma distribution.

38. The chief technology officier estimates that a project will take at least four months and at most eight months. He would like to associate a generalized beta distribution with

these estimates. For the several possible values of the most likely times given in the table, fill in the missing values of α and β.

Most likely, *m*	4	5	6	7	8
α		2	2		2
β	2			2	

39. It is estimated that a software development project will take at least four months and at most eight months. Find the probability that the completion time will be less than six months when we use the following beta distribution parameters to model the time to completion, Y.
    a. $\alpha = 2$, $\beta = 2$
    b. $\alpha = 2$, $\beta = 5$
    c. $\alpha = 5$, $\beta = 2$

40. Ten students join hands. The reach of one student (the distance from the left to the right hand) is a random variable with a normal distribution having $\mu = 38$ inches and $\sigma = 4$ inches. What is the probability distribution of the combined reach of all ten students?

41. Tillie enters a bank and finds a single teller at work with a line forming in front of the window. From past experience she estimates a mean time of 2 minutes for each customer. For the service time distributions specified in each part below, what is the probability that she will begin her service within the next 15 minutes if there is one person ahead of her (the person is already in service)? Answer this question if she finds two persons waiting or in service when she enters the bank. Answer the question for 5 persons.
    a. The time for service of each customer has an exponential distribution.
    b. The time for service of each customer has a normal distribution with a mean of 2 minutes and a standard deviation of 1 minute. Accept the possibility of negative times.
    c. The time for service has a lognormal distribution with a mean of 2 minutes and a standard deviation of 1 minute.

42. A submarine must remain submerged for a period of three months. A particularly important subsystem of the submarine has an exponential failure time distribution with a mean of 5 months.
    a. What is the reliability of the subsystem? The reliability is defined as the probability that the subsystem will operate without any failures for the full three month mission.
    b. The submarine stores one redundant subsystem (so there are two in all). The redundant subsystem has no failure rate while in storage. What is the reliability of the subsystem with redundancy.

c. The submarine stores four redundant subsystems (five in all). The redundant subsystems have no failure rate while not operating. What is the reliability of the subsystem with redundancy.

d. Comment on the effect of redundancy on the reliability of the system. What are the critical assumptions that make the above analysis valid?

e. Say the cost of buying and holding a component for a mission is \$1000. The cost of provisioning the submarine with new components if all on board fail is \$200,000. What is the optimal number of components to bring on the mission?

f. Change the situation of this problem so that all components have the same failure rate whether they are operating or not. The sub still requires one good component for a successful mission. What is the optimal number of components to bring?

g. Change the situation of this problem so that the time to failure of the operating component has a Weibull distribution with a mean of 5 months and parameter β = 2. Assume that non-operating components do not fail. Use simulation to estimate the reliability of the system with 1, 2 and 3 components. You will have to simulate to answer this problem for 2 and 3 components. Use 500 observations.

Exercises 43 – 45. Use the random numbers in the table below to answer these questions.

Random Numbers

	1	2	3	4	5	6
1	0.3410	0.4921	0.5907	0.9658	0.8632	0.6327
2	0.8228	0.4294	0.0809	0.3400	0.4116	0.8931
3	0.6066	0.3841	0.8915	0.6096	0.7013	0.4947
4	0.3729	0.7230	0.1621	0.6537	0.0011	0.3888
5	0.1241	0.9864	0.7059	0.3734	0.9895	0.0768

43. Use row 1 to generate 6 observations from a:

a. Bernoulli distribution p = 0.4. Use the range ($0 \le r \le 0.6$) for 0 and ($0.6 < r \le 1$) for 1.

b. Binomial distribution with n = 5 and p = 0.4.

c. Geometric distribution with p = 0.4.

d. Poisson distribution with θ = 2.

44. Use row 1 to generate 6 observations from a:

a. Standard normal distribution.

b. Normal distribution with μ = 100 and σ = 20.

c. Lognormal distribution with a standard normal underlying distribution.

45. Use all six numbers from each row for the following problems.

a. Simulate five observations of a Binomial random variable with $n = 6$ and $p = 0.4$ by simulating Bernoulli trials with $p = 0.4$.

b. Simulate five replications of the sum of six observations by summing six simulated observations from the standard normal distribution.

c. Simulate five observations from a gamma distribution with $n = 6$ and $\lambda = 2$ by summing simulated exponential random variables.

46. When evaluating investments with cash flows at different points in time, it is common to compute their net present worth (NPW). Consider an investment of the amount P with annual returns I_k for $k = 1,\ldots,n$, where n is the life of the investment and I_k is the return in year k. The NPW is defined in terms of i and represents the investor's minimum acceptable rate of return (MARR). If NPW ≥ 0, the investment provides a return at least as great as the MARR. If NPW < 0, the investment fails to provide the MARR. Your investment advisor tells you a particular investment of $1000 will provide a return of $500 per year for three years. Your MARR is 20% or 0.2.

a. Is the investment acceptable according to the NPW criterion.

b. The advisor adds the information that the life of the investment is uncertain. Rather there is a uniform distribution on the life with $p_n = 0.2$ for a life of $n = 1$, 2, 3, 4 and 5. Use the random numbers below to simulate ten replications of the life and compute the NPW in each case.

0.5758	0.4998	0.3290	0.3854	0.6784	0.8579
0.5095	0.6824	0.3762	0.1351	0.2555	0.9773

Based on your simulation, what is the probability that the investment will yield the MARR?

c. The advisor now tells you that the annual revenue per year is uncertain. Although it will be the same each year, the revenue is normally distributed with a mean of $500 and a standard deviation of $200. Use the random numbers below to simulate ten observations of the annual revenue.

0.5153	0.3297	0.6807	0.0935	0.9872	0.6339
0.0858	0.3229	0.5285	0.4451	0.3177	0.1562

Combine the lives simulated with the revenues of this section to determine the probability that the investment will yield the MARR.

Bibliography

Barnes, J. W., *Statistical Analysis for Engineers and Scientists*, McGraw Hill, New York 1994.

Bowden, R., “The Spectrum of Simulation Software,” *IIE Solutions,* Vol. 30, No. 5, pp. 44-46, 1998.

Hayter, A.J., *Probability and Statistics for Engineers and Scientists*, PWS, Boston, 1996.

Johnson, R.A., *Miller & Freund’s Probability and Statistics for Engineers*, Sixth Edition, Prentice Hall, Upper Saddle Ridge, NJ, 2000.

Law, A.M., *ExpertFit*, Averill M. Law & Associates, Tucson, AZ, 1998.

Montgomery, D.C., G.C. Runger and N.F. Hubele, *Engineering Statistics*, John Wiley & Sons, New York, 1998.

Ross, S., *A First Course in Probability*, Fifth Edition, Prentice Hall, Upper Saddle Ridge, NJ, 1998.

Rubinstein, R.Y., B. Melamed and A. Shapiro, *Modern Simulation and Modeling*, John Wiley & Sons, New York, 1998.

Walpole, R.E., R.H. Meyers, S.L. Myers and K. Ye, *Probability & Statistics for Engineers & Scientists*, Seventh Edition, Prentice Hall, Upper Saddle Ridge, NJ, 2002.