Probability Concepts and Distribution

Introduction

Managers often base their decisions on an analysis of uncertainties such as:

1. What are the chances that the sales will decrease if we increase prices?
2. How likely is it that the project will be finished on time?
3. What is the chance that a new commercial real estate loan will go into default.

Probability is a numerical measure of the likelihood that an event will occur.

Probability values are always assigned on a scale from 0 to 1:

- A probability near zero indicates an event is quite unlikely to occur.
- A probability near one indicates an event is almost certain to occur.
- Other probabilities between 0 and 1 represent degrees of likelihood that an event will occur.

Random Experiment and Sample Space

A **Random experiment** is a process that generates *experimental outcomes*, and possesses the following properties:

1. The experimental outcomes are well-defined and may also be listed prior to conducting the experiment.
2. On any single repetition (*trial*) of the experiment, one and only one of the possible experimental outcomes will occur.
3. The experimental outcome that occurs on any trial is determined solely by chance.

An experimental outcome is also called a **sample point**. The **sample space**, *S*, for an experiment is the set of all experimental outcomes.

Example: consider the random experiment of tossing a coin.

- The experimental outcomes (sample points) are Head and Tail.
- The sample space is: $S = \{\text{Head}, \text{Tail}\}$

Counting Rule for Multiple-Step Experiments

A **multiple-step experiment** can be described as a sequence of *k* steps with n_1 possible outcomes on the first step, n_2 possible outcomes on the second step, and so on.

The total number of experimental outcomes is given by $(\boldsymbol{n_1})(\boldsymbol{n_2})\ldots(\boldsymbol{n_k})$.

As an example, consider the random experiment of tossing two coins, one at a time.

The first tossed coin has $n_1 = 2$ outcomes, and the second tossed coin has also $n_2 = 2$ outcomes.

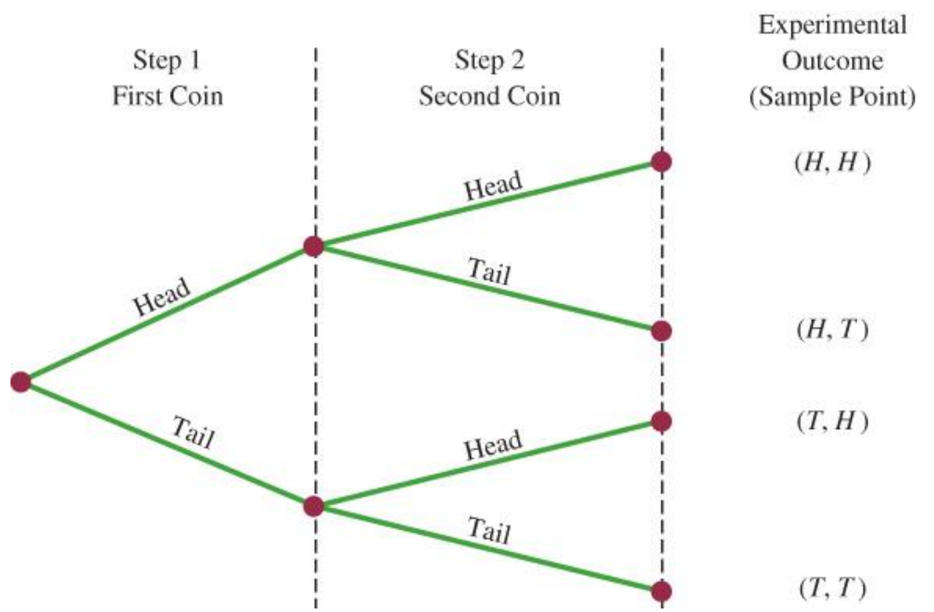
From the counting rule, we have:

$(n_1)(n_2) = (2)(2) = 4$ experimental outcomes

The sample space is :

$S = \{(\mathrm{H}, \mathrm{H}), (\mathrm{H}, \mathrm{T}), (\mathrm{T}, \mathrm{H}), (\mathrm{T}, \mathrm{T})\}$

The **tree diagram** for the two-coin toss shown to the right is a graphical representation that helps visualizing a multiple-step experiment.



The KP&L Project: a Multiple-Step Experiment

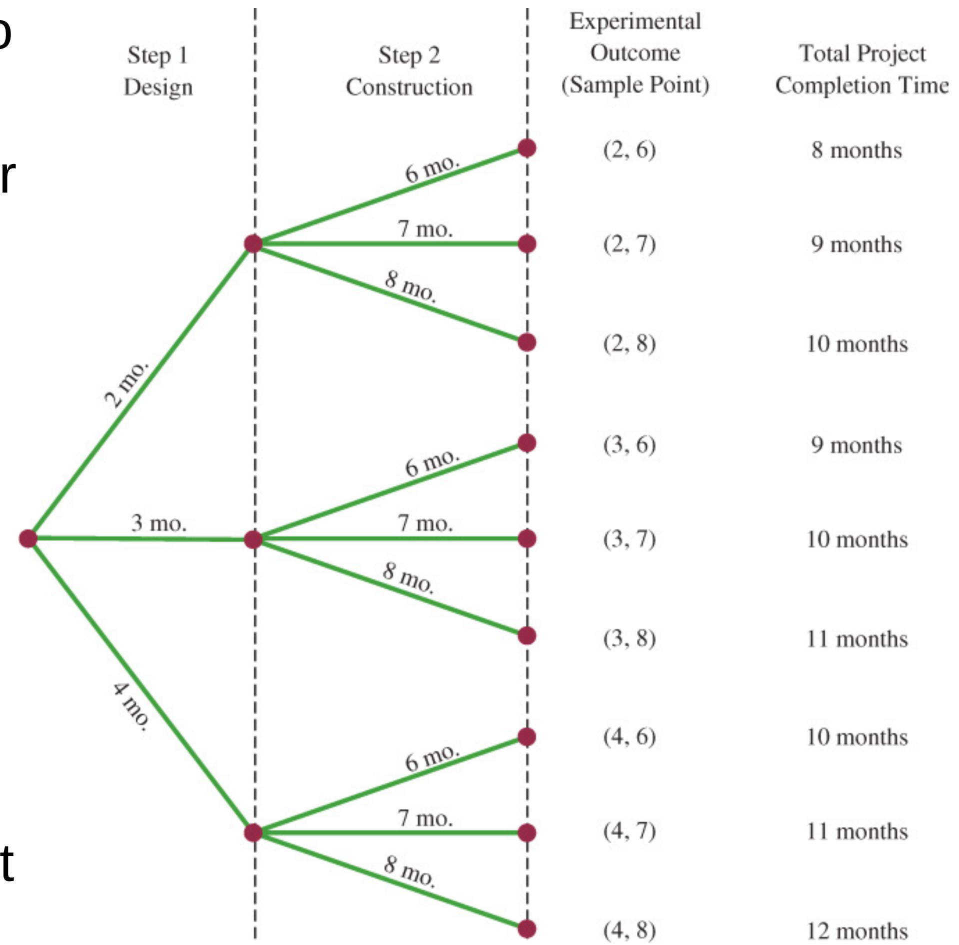
The KP&L capacity expansion project consists of two sequential stages: design and construction.

Management expects the design stage to last 2, 3, or 4 months, the construction stage 6, 7, or 8 months, with an overall completion goal of 10 months.

Application of the counting rule for multiple-step experiments reveals a total of (3)(3) = 9 experimental outcomes.

The tree diagram to the right shows how the 9 outcomes (sample points) occur and the project completion time for each outcome.

We need to assign probabilities to each of the 9 outcomes before we can compute the probability that the project is completed on time.



Assigning Probabilities

The basic requirements for assigning probabilities are

1. The probability assigned to each experimental outcome must be between 0 and 1, inclusively.

   $$\mathbf{0 \leq P(E_i) \leq 1}$$

   where E_i is the ith experimental outcome and $P(E_i)$ its probability.

2. The sum of the probabilities for all experimental outcomes must equal 1.

   $$\boldsymbol{P(E_1) + P(E_2) + \cdots P(E_n) = 1}$$

   where n is the number of experimental outcomes.

There are three methods of assigning probabilities:

- The **classical method**, is based on the assumption of equally likely outcomes.
- The **empirical method**, based on relative frequencies from experimental or historical data.
- The **subjective method**, based on experience or intuition when no data are available.

Assigning Probabilities for the KP&L Project

Using experience and judgment, KP&L management concluded that not all the outcomes were equally likely. Thus, the classical method could not be used.

A study on the completion times of 40 similar projects undertaken over the past three years revealed the frequency distribution shown.

Using the empirical method, we can assign the probabilities to each of the nine outcomes.

For example, sample point (2, 6) on the top row, with an expected completion time of 8 months, has probability:

$$P(2,6) = 6/40 = 0.15$$

Completion Time (months)					
Stage 1 Design	Stage 2 Design	Sample Point	Number of Past Projects	Project Completion Time (months)	Probability of Sample Point
2	6	(2, 6)	6	8	*P*(2, 6) = 6/40 = 0.15
2	7	(2, 7)	6	9	*P*(2, 7) = 6/40 = 0.15
2	8	(2, 8)	2	10	*P*(2, 8) = 2/40 = 0.05
3	6	(3, 6)	4	9	*P*(3, 6) = 4/40 = 0.10
3	7	(3, 7)	8	10	*P*(3, 7) = 8/40 = 0.20
3	8	(3, 8)	2	11	*P*(3, 8) = 2/40 = 0.05
4	6	(4, 6)	2	10	*P*(4, 6) = 2/40 = 0.05
4	7	(4, 7)	4	11	*P*(4, 7) = 4/40 = 0.10
4	8	(4, 8)	6	12	*P*(4, 8) = 6/40 = 0.15
			40		

Events

An **event** is a collection of sample points.

If we can identify all the sample points of an experiment and assign a probability to each, we can compute the probability of an event.

In the KP&L example, management was interested in the probability that the project is completed in 10 months or less.

From the table in the previous slide, we see that there are six sample points that provide a project completion time of 10 months or less.

Let *C* denote the event that the project is completed in 10 months or less. Thus, we have:

$$C = \{(2, 6), (2, 7), (2, 8), (3, 6), (3, 7), (4, 6)\}$$

Other events of interest may be:

L, the event that the project will complete in less than 10 months: $L = \{(2, 6), (2, 7), (3, 6)\}$

M, the event that the project will complete in more than 10 months: $M = \{(3, 8), (4, 7), (4, 8)\}$

Probability of an Event

The **probability of an event** is equal to the sum of the probabilities of the sample points in the event.

We can calculate the probability that the KP&L project is completed in 10 months or less by adding the probabilities of the six sample points belonging to event *C*.

$$P(C) = P(2,6) + P(2,7) + P(2,8) + P(3,6) + P(3,7) + P(4,6)$$

If we refer to the sample point probabilities listed in the previous table, we have

$$P(C) = 0.15 + 0.15 + 0.05 + 0.10 + 0.20 + 0.05 = 0.7$$

Similarly, the probability that the KP&L project is completed in less than 10 months is

$$P(L) = P(2,6) + P(2,7) + P(3,6) = 0.15 + 0.15 + 0.10 = 0.4$$

And the probability that the KP&L project is completed in more than 10 months is

$$P(M) = P(3,8) + P(4,7) + P(4,8) = 0.05 + 0.10 + 0.15 = 0.3$$

Some Basic Relationships of Probability

In an experiment with a large number of sample points, the identification of all the sample points, as well as the determination of their associated probabilities, can become an extremely cumbersome task.

In this section, we present some basic probability relationships that can be used to compute the probability of an event without knowledge of all the sample point probabilities.

These basic relationships of probability are:

- The complement of an event
- The union of two events
- The intersection of two events
- The addition law
- The multiplication law

Complement of an Event

The **complement of event $\boldsymbol{A}$** is defined as the event consisting of all sample points that are *not* in A.

The complement of A is denoted by A^C.

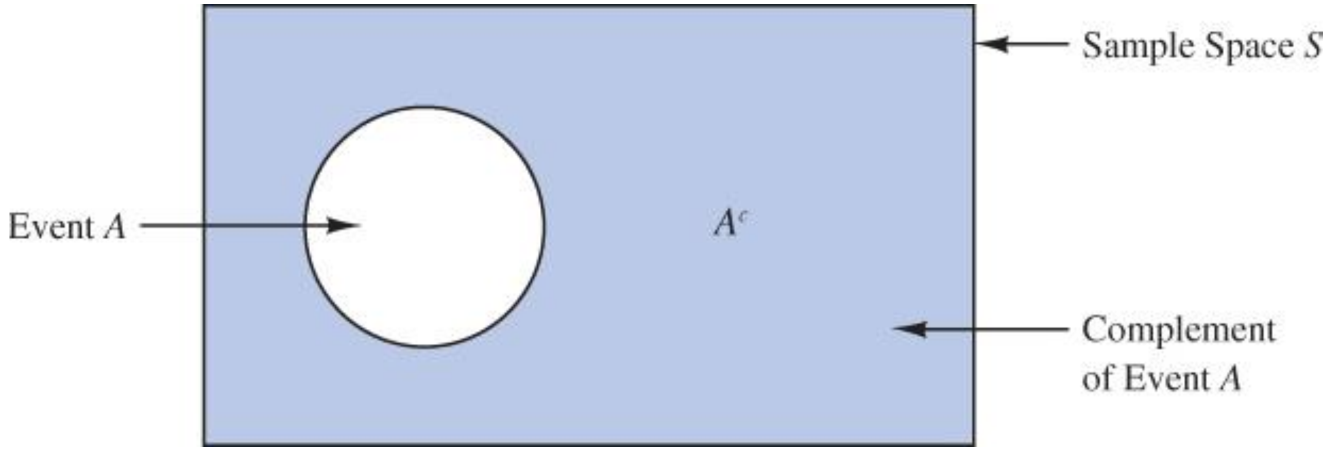
In any probability application, either event A or its complement A^C must occur, as shown in the Venn diagram to the right.

Therefore, we have

$$P(A) + P(A^C) = 1$$

Solving for $P(A)$, we obtain the formula for the probability of the complement of an event

$$\boldsymbol{P(A) = 1 - P\left(A^C\right)}$$

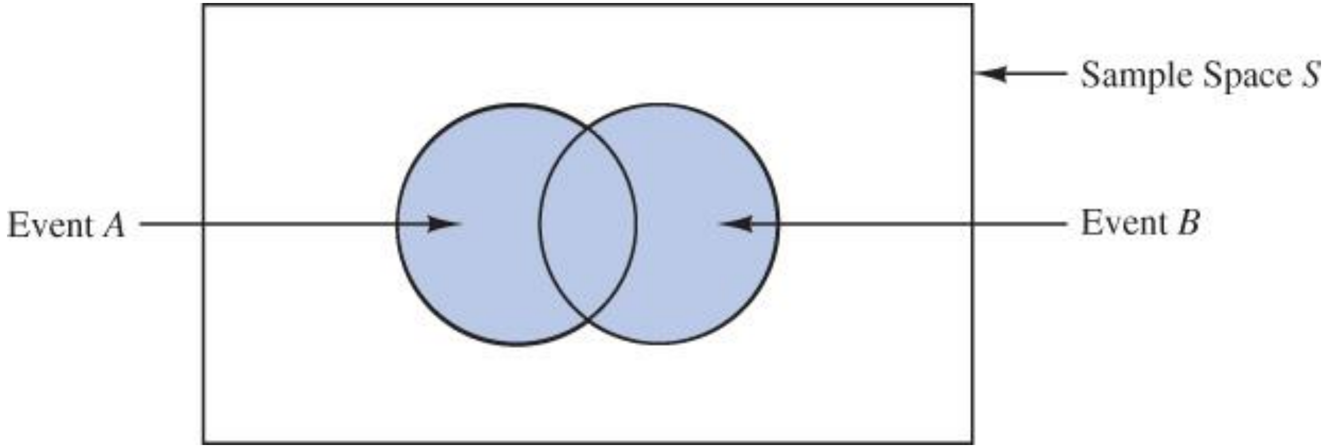


Union of Two Events

The **union of events $\boldsymbol{A}$ and $\boldsymbol{B}$** is the event containing all sample points that are in A and B or both.

The union of events A and B is denoted by $A \cup B$.

Note that the two circles in the Venn diagram overlap, indicating that some sample points are contained in both A and B.



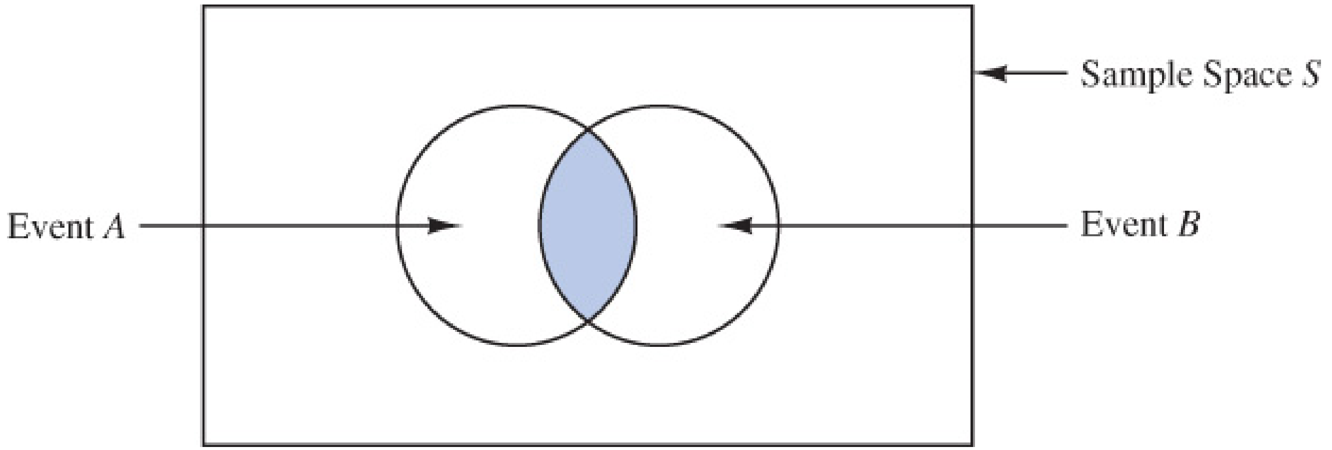
Intersection of Two Events

The **intersection of events A and B** is the event containing all sample points belonging to both A and B.

The intersection of events A and B is denoted by $A \cap B$.

Note that the area in the Venn diagram where the two circles overlap is the intersection.

The intersection contains all the sample points that are in both A and B.



Addition Law

The **addition law** provides a way to compute the probability of the union of two events:

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

To understand the addition law intuitively, consider the following:

- The sample points in the intersection $A \cap B$ are in both A and B.
- When we compute $P(A) + P(B)$, we are in effect counting each of the sample points in $A \cap B$ twice.
- We correct for this overcounting by subtracting $P(A \cap B)$.

In the next slide, we consider an example as an application of the addition law.

An Application of the Addition Law

Consider a group of 50 software engineers who work in online banking.

At the end of an evaluation period, a bank manager found that

- 5 engineers completed work late (event *L*)
- 6 engineers produced code that contains errors (event *E*)
- 2 engineers completed work late and produced code that contains errors (event $L \cap E$)

Thus, we have:

$$P(L) = 5/50 = 0.10$$

$$P(E) = 6/50 = 0.12$$

$$P(L \cap E) = 2/50 = 0.04$$

The probability that an engineer completed work late or produced code that contains errors (event $L \cup E$) can be calculated using the addition law as

$$\mathrm{P}(L \cup E) = P(L) + P(E) - P(L \cap E) = 0.10 + 0.12 - 0.04 = 0.18$$

Mutually Exclusive Events

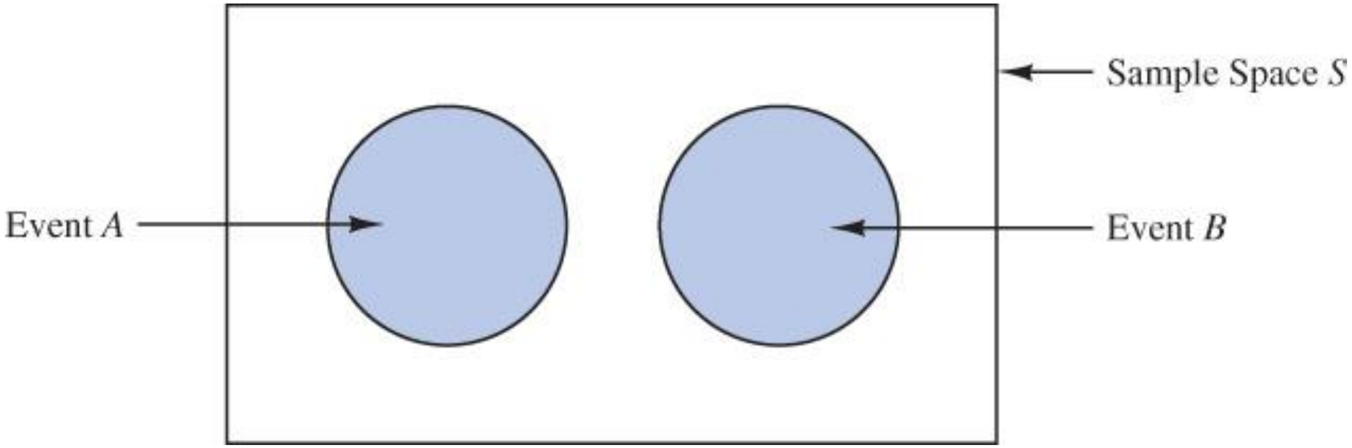
Two events are said to be **mutually exclusive** when they have no sample points in common.

In other words, events A and B are mutually exclusive if, when one event occurs, the other cannot occur.

Thus, the intersection of A and B does not contain any points, and $P(A \cap B) = 0$.

When two events are mutually exclusive, the addition law simplifies to

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

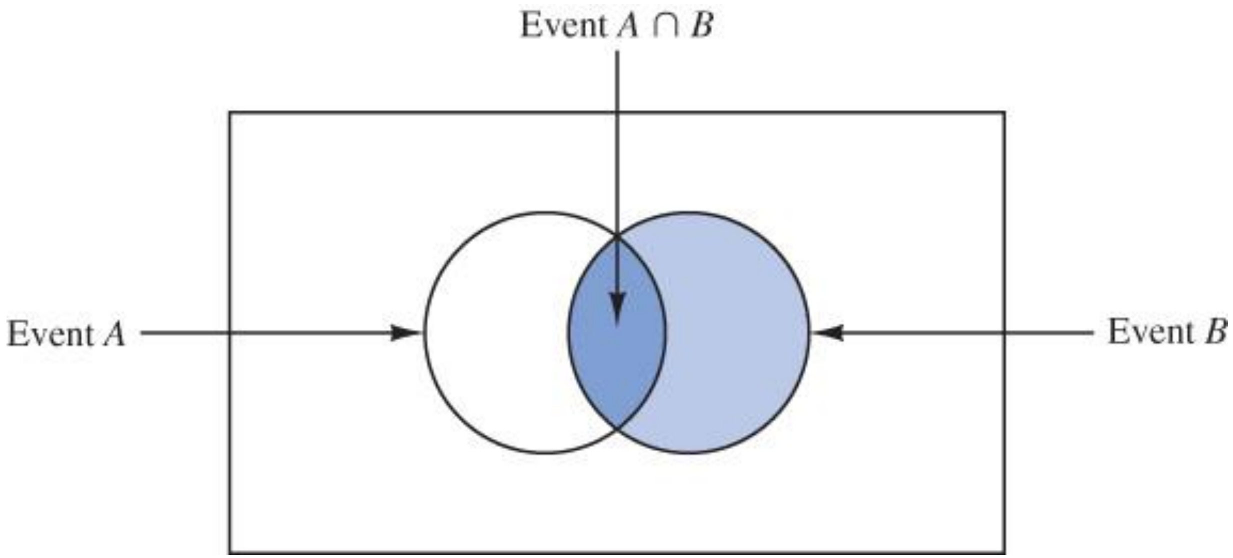


Conditional Probability

We use the notation $P(A \mid B)$ to denote the **conditional probability** of event A, given that event B has occurred.

We calculate the probability of event A given B as

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



We calculate the probability of event *B* given *A* as

$$P(B \mid A) = \frac{P(A \cap B)}{P(A)}$$

Note that

$$P(B \mid A) \neq P(A \mid B)$$

Joint Probability Table

Dividing the data values in the crosstabulation by the total of 1,200 officers enables us to summarize the available information on promotions with a **joint probability table**.

In the body of the table, we have the **joint probabilities**, which are the intersection probabilities of the events of whether a police officer was promoted (A) or not (A^c) with whether the police officer is male (M) or female (F).

The joint probability that an officer was promoted (A) and is a male (M)

$$P(A \cap M) = 288/1200 = 0.24$$

On the margins of the table, we have the **marginal probabilities**, providing the probabilities of each separate event.

The probability that an officer is male (M)

$$P(M) = 960/1200 = 0.80$$

	Male (M)	Female (F)	Total
Promoted (A)	0.24	0.03	0.27
Not Promoted (A^c)	0.56	0.17	0.73
Total	0.80	0.20	1.00

Conditional Probability Analysis

We note that the marginal probabilities are found by summing the joint probabilities in the corresponding row or column of the joint probability table. For example

$$P(A) = P(A \cap M) + P(A \cap F) = 0.24 + 0.03 = 0.27$$

The conditional probability that an officer was promoted (*A*) given that the officer is male (*M*)

$$P(A \mid M) = \frac{P(A \cap M)}{P(M)} = \frac{288/1200}{960/1200} = \frac{288}{960} = 0.30$$

Note that this could have also been calculated directly from the corresponding values in the crosstabulation using the conditional distribution (column) for M as $P(A \mid M) = 288/960 = 0.30$.

The conditional probability that an officer was promoted (*A*) given that the officer is female (*F*)

$$P(A \mid F) = \frac{P(A \cap F)}{P(F)} = \frac{36/1200}{240/1200} = \frac{36}{240} = 0.15$$

Thus, because $P(A \mid M) > P(A \mid F)$, the analysis supports the female officers' argument.

Independent Events

Two events *A* and *B* are **independent** if

$$P(A \mid B) = P(A) \quad \text{or} \quad P(B \mid A) = P(B)$$

Otherwise, the events are *dependent* if

$$P(A \mid B) \neq P(A) \quad \text{or} \quad P(B \mid A) \neq P(B)$$

In the analysis of the promotion of police officers, we found

$$P(A) = 0.27,\ P(A \mid M) = 0.30, \text{ and } P(A \mid F) = 0.15$$

Because $P(A \mid M) \neq P(A)$, that is, the probability of whether an officer is promoted (A) is altered or affected by knowing whether the officer is male (M), we conclude that the events A and M are dependent.

Similarly, because $P(A \mid F) \neq P(A)$, we would say that events A and F are also dependent

Multiplication Law

The **multiplication law** is based on the definition of conditional probability, and it provides a way to compute the probability of the intersection of two events

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

or

$$P(B \cap A) = P(B \mid A)P(A)$$

In the case of independent events, $P(A \mid B) = P(A)$ or $P(B \mid A) = P(B)$, and the **multiplication law for independent events** becomes

$$P(A \cap B) = P(A)P(B)$$

or

$$P(B \cap A) = P(B)P(A)$$

Therefore, two events *A* and *B* are dependent when $P(A \cap B) \neq P(A)P(B)$.

Bayes' Theorem

The **Bayes' theorem (two-event case)** states that

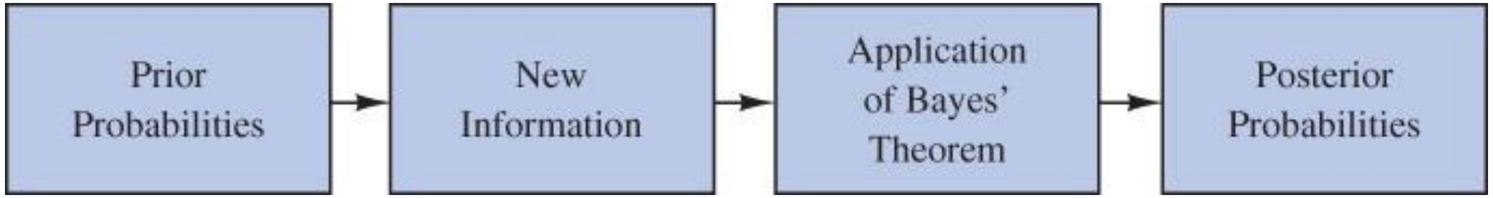
$$P(A_i \mid B) = \frac{P(A_i)P(B \mid A_i)}{P(A_1)P(B \mid A_1) + P(A_2)P(B \mid A_2)} \quad \text{with } i = 1,2$$

Where:

$P(A_1)$ and $P(A_2)$ are the initial or **prior probabilities**, available at the beginning.

$P(B \mid A_1)$ and $P(B \mid A_2)$ are conditional probabilities calculated from a sample, special report, or product test.

Bayes' theorem provides a mean to calculate the revised or **posterior probabilities** $P(A_1 \mid B)$ and $P(A_2 \mid B)$.



An Application of Bayes' Theorem

Consider a manufacturing firm that purchases 65% of its parts from supplier 1 and 35% from supplier 2.

If we let A_1 denote that event that a part is from supplier 1, and A_2 the event that the part is from supplier 2, we can write the *prior probabilities* as

$$P(A_1) = 0.65 \qquad P(A_2) = 0.35$$

Historical data for the quality ratings of the two suppliers are shown in the table below.

Thus, we can write the *conditional probabilities* of receiving a good (G) or bad (B) part from either supplier as:

$$P(G \mid A_1) = 0.98 \qquad P(B \mid A_1) = 0.02$$

$$P(G \mid A_2) = 0.95 \qquad P(B \mid A_2) = 0.05$$

	Percentage Good Parts	Percentage Bad Parts
Supplier 1	98	2
Supplier 2	95	5

Posterior Probabilities

Using Bayes' theorem and the prior and conditional probabilities, we can compute the *posterior probabilities* that either supplier ships a bad part

$$P(A_1 \mid B) = \frac{P(A_1)P(B \mid A_1)}{P(A_1)P(B \mid A_1) + P(A_2)P(B \mid A_2)} = \frac{0.65(0.02)}{0.65(0.02) + 0.35(0.05)} = \frac{0.130}{0.130 + 0.175} = 0.426$$

$$P(A_2 \mid B) = \frac{P(A_2)P(B \mid A_2)}{P(A_1)P(B \mid A_1) + P(A_2)P(B \mid A_2)} = \frac{0.35(0.05)}{0.65(0.02) + 0.35(0.05)} = \frac{0.175}{0.130 + 0.175} = 0.574$$

Note that, for supplier 1, we began with an initial probability of $P(A_1) = 0.65$. However, given the information that the part was bad, the revised probability drops to $P(A_1 \mid B) = 0.426$.

Conversely, for supplier 2, we began with an initial probability of $P(A_2) = 0.35$ and when we include the information that the part was bad, the revised probability grows to $P(A_2 \mid B) = 0.574$.

In conclusion, despite the fact more parts come from supplier 1 than supplier 2, there is more than a 50-50 chance that a bad part comes from supplier 2.

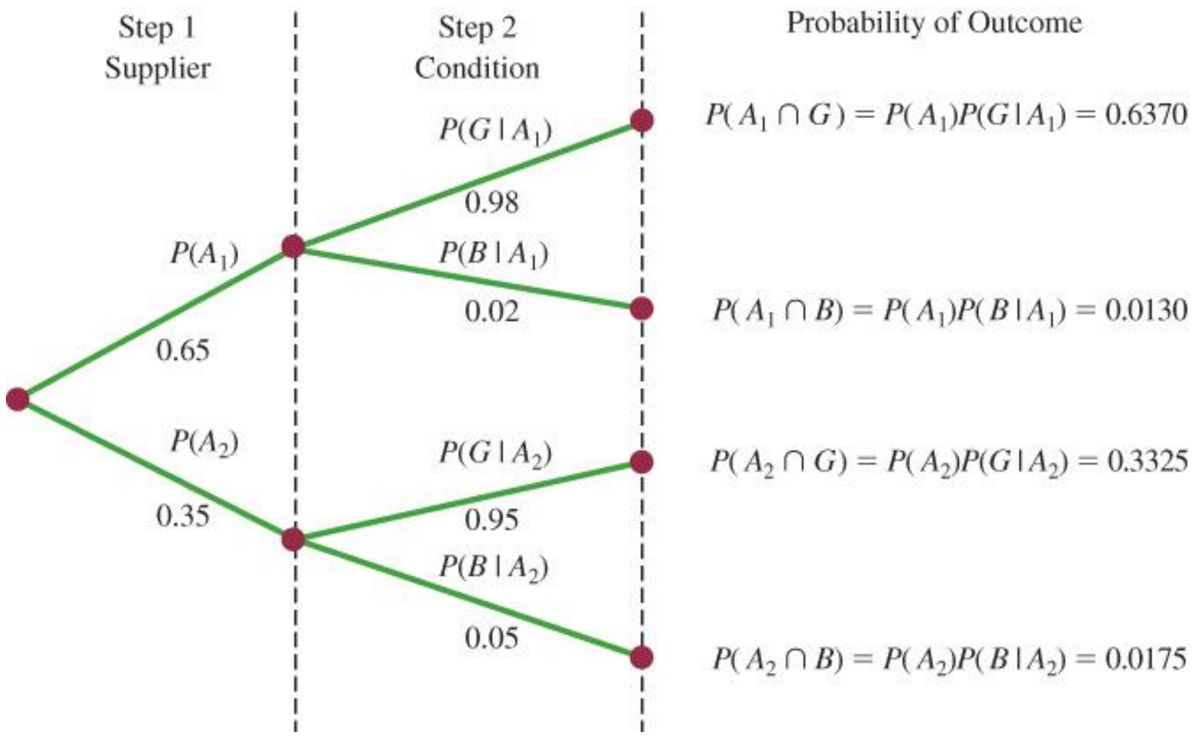
Bayes' Theorem with a Probability Tree

Bayes' theorem can be visualized with a **probability tree**. From left to right:

Step 1 includes the two branches for the prior probabilities of the two suppliers.

Step 2 branches out the four conditional probabilities of receiving a good (G) or bad (B) part from either supplier 1 (A_1) or 2 (A_2).

The resulting four outcomes represent the possible intersections of supplier (A_1) or (A_2) with either a good (G) or bad (B) part.



Because $P(B) = P(A_1 \cap B) + P(A_2 \cap B)$, the posterior probabilities in Bayes' theorem can be calculated from the joint probabilities as

$$P(A_i \mid B) = \frac{P(A_i \cap B)}{P(B)} = \frac{P(A_i \cap B)}{P(A_1 \cap B) + P(A_2 \cap B)} = \frac{P(A_i)P(B \mid A_i)}{P(A_1)P(B \mid A_1) + P(A_2)P(B \mid A_2)} \quad \text{with } i = 1,2$$

Tabular Approach to Bayes' Theorem

A tabular approach is helpful in conducting Bayes' theorem calculations (see table below.)

Step 1: Create three columns: (1) with the two events A_1 and A_2 for which posterior probabilities are desired, (2) with the prior probabilities $P(A_1)$ and $P(A_2)$, and (3) with the conditional probabilities from the available data for event B.

Step 2: In column (4), for each of the two rows compute the joint probabilities $P(A_i \cap B)$ by multiplying the values in columns (1) and (2).

Step 3: Sum the joint probabilities in column (4) to obtain $P(B)$.

Step 4: In column (5), compute the posterior probabilities using the values from column (4):

$$P(A_1 \mid B) = \frac{P(A_1 \cap B)}{P(B)}$$

$$P(A_2 \mid B) = \frac{P(A_2 \cap B)}{P(B)}$$

(1) Events A_i	(2) Prior Probabilities $P(A_i)$	(3) Conditional Probabilities $P(B\|A_i)$	(4) Joint Probabilities $P(A_i \cap B)$	(5) Posterior Probabilities $P(A_i\|B)$
A_1	0.65	0.02	0.0130	0.0130 / 0.0305 = 0.4262
A_2	0.35	0.05	0.0175	0.0175 / 0.0305 = 0.5738
	1.00		0.0305	1.0000

Discrete Probability Distributions

Examples of Discrete Random Variables

A **random variable** is a numerical description of the outcome of an experiment.

A **discrete random variable** may assume either a finite number of values or an infinite sequence of values.

Random Experiment	**Random Variable (x)**	**Possible Values for the Random Variable**
Flip a coin	Face of coin showing	1 if heads; 0 if tails
Roll a die	Number of dots showing on top of die	1, 2, 3, 4, 5, 6
Contact five customers	Number of customers who place an order	0, 1, 2, 3, 4, 5
Operate a health care clinic for one day	Number of patients who arrive	0, 1, 2, 3, …
Offer a customer the choice of two products	Product chosen by customer	0 if none; 1 if choose product A; 2 if choose product B

Examples of Continuous Random Variables

A **continuous random variable** may assume any numerical value in an interval or collection of intervals (*see notes.)

Random Experiment	Random Variable (x)	Possible Values for the Random Variable
Customer visits a web page	Time customer spends on web page in minutes	$x \geq 0$
Fill a soft drink can (max capacity = 12.1 ounces)	Number of ounces	$0 \leq x \leq 12.1$
Test a new chemical process	Temperature when the desired reaction takes place (min temperature = 150°F; max temperature = 212°F)	$150 \leq x \leq 212$
Invest $10,000 in the stock market	Value of investment after one year	$x \geq 0$

Discrete Probability Distributions

The **probability distribution** for a random variable describes how probabilities are distributed over the values of the random variable.

We can describe a discrete probability distribution with a table, graph, or formula.

The Two Types of Discrete Probability Distributions

1. Uses the rules of assigning probabilities to experimental outcomes to determine probabilities for each value of the random variable.
2. Uses a special mathematical formula to compute the probabilities for each value of the random variable.

Probability Function

The **probability function**, denoted by *f*(*x*), defines the probability distribution by providing the probability for each value of the random variable.

The required conditions for a discrete probability function are:

$$\boldsymbol{f(x) \geq 0} \quad \text{and} \quad \sum \boldsymbol{f(x) = 1}$$

Empirical Discrete Distribution

There are three methods for assigning probabilities to random variables: classical method, subjective method, and relative frequency method.

The use of the relative frequency method to develop discrete probability distributions leads to an **empirical discrete distribution**.

Example: Number of automobiles sold during a day at DiCarlo Motors.

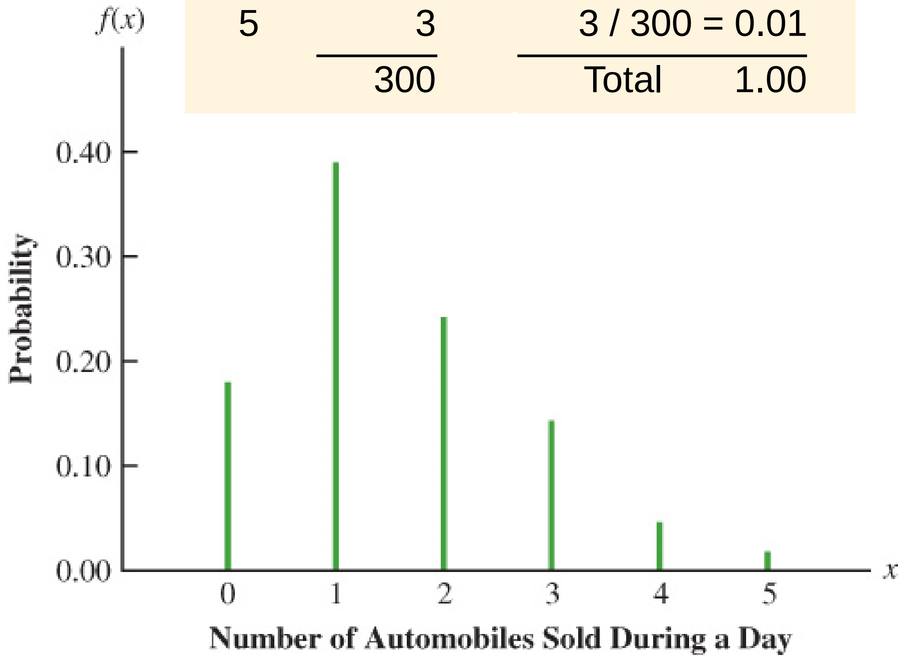
The table to the right shows the number of automobiles sold at DiCarlo over the past 300 days.

We define the random variable of interest as:

x = the number of automobiles sold during a day.

We can use the relative frequencies to represent graphically the empirical discrete distribution for x.

x	**# of days**	***f*(*x*)**
0	54	54 / 300 = 0.18
1	117	117 / 300 = 0.39
2	72	72 / 300 = 0.24
3	42	42 / 300 = 0.14
4	12	12 / 300 = 0.04
5	3	3 / 300 = 0.01
	300	Total 1.00



Probability Distribution Given by a Formula

In addition to tables and graphs, a formula that gives the probability function, $f(x)$, for every value of x is often used to describe the probability distributions.

Typical discrete probability distributions specified by formulas are the discrete-uniform, binomial, Poisson, and hypergeometric distributions.

The **discrete uniform probability distribution** is the simplest example of a discrete probability distribution given by a formula.

The discrete uniform probability function is

$$f(x) = \frac{1}{n}$$

Where: n = the number of values the random variable may assume

An example of a discrete uniform probability distribution is provided by the experiment of rolling a die, in which $n = 6$ equally likely outcomes, the random variable x is represented by the number of dots on the upward face of the die, and $f(x) = 1/n = 1/6$.

Expected Value

The **expected value**, or mean, of a random variable is a measure of its central location.

The expected value of a discrete random variable is calculated as a weighted average of the values the random variable may assume. The weights are the probabilities.

$$\boldsymbol{E(x) = \sum xf(x)}$$

The expected value for the DiCarlo Motors example is

$E(x) = \sum xf(x) = 1.50$ automobiles (*see notes)

Although sales from 0 to 5 automobiles are possible on any one day, over time DiCarlo can anticipate selling an average of 1.50 automobiles per day, or 30(1.50) = 45 automobiles per month.

x	***f*(*x*)**	***xf*(*x*)**
0	0.18	0(0.18) = 0.00
1	0.39	1(0.39) = 0.39
2	0.24	2(0.24) = 0.48
3	0.14	3(0.14) = 0.42
4	0.04	4(0.04) = 0.16
5	0.01	5(0.01) = 0.05
		1.50

Variance and Standard Deviation

The variance summarizes the variability in the values of a random variable.

The variance of a discrete random variable is calculated as

$$\boldsymbol{Var(x) = \sigma^2 = \sum (x-\mu)^2 f(x)}$$

For the DiCarlo Motors example, variance and standard deviation are calculated as

$$Var(x) = \sum (x-\mu)^2 f(x) = 1.25$$

$$\boldsymbol{\sigma} = \sqrt{\boldsymbol{Var(x)}} = \sqrt{1.25} = 1.118 \text{ automobiles}$$

Because the standard deviation is measured in the same units as the random variable, it is often preferred in describing the variability of a random variable.

x	$x-\mu$	$(x-\mu)^2$	$f(x)$	$(x-\mu)^2 f(x)$
0	0 − 1.5 = −1.50	2.25	0.18	2.25(0.18) = 0.4050
1	1 − 1.5 = −0.50	0.25	0.39	0.25(0.39) = 0.0975
2	2 − 1.5 = 0.50	0.25	0.24	0.25(0.24) = 0.0600
3	3 − 1.5 = 1.50	2.25	0.14	2.25(0.14) = 0.3150
4	4 − 1.5 = 2.50	6.25	0.04	6.25(0.04) = 0.2500
5	5 − 1.5 = 3.50	12.25	0.01	12.25(0.01) = 0.1225
				1.2500

Continuous Probability Distributions

Uniform Probability Density Function

A random variable is described by a **uniform probability distribution** whenever the probability is proportional to the interval's length.

The *uniform probability density function* is:

$$\boldsymbol{f(x)} = \begin{cases} \dfrac{\mathbf{1}}{\boldsymbol{b-a}} & \textbf{for } \boldsymbol{a \le x \le b} \\ \mathbf{0} & \textbf{elsewhere} \end{cases}$$

Where

a = smallest value the variable x can assume

b = largest value the variable x can assume

A uniform continuous probability distribution has expected value and variance.

$$\boldsymbol{E(x) = \frac{a+b}{2}} \qquad \boldsymbol{Var(x) = \frac{(b-a)^2}{12}}$$

Area as a Measure of Probability

The flight time of an airplane traveling from Chicago to New York can be described by a continuous random variable that can assume any value in the interval from 120 to 140 minutes.

Let us assume that sufficient actual flight data are available to conclude that the probability of a flight time within any 1-minute interval is the same from 120 to 140 minutes.

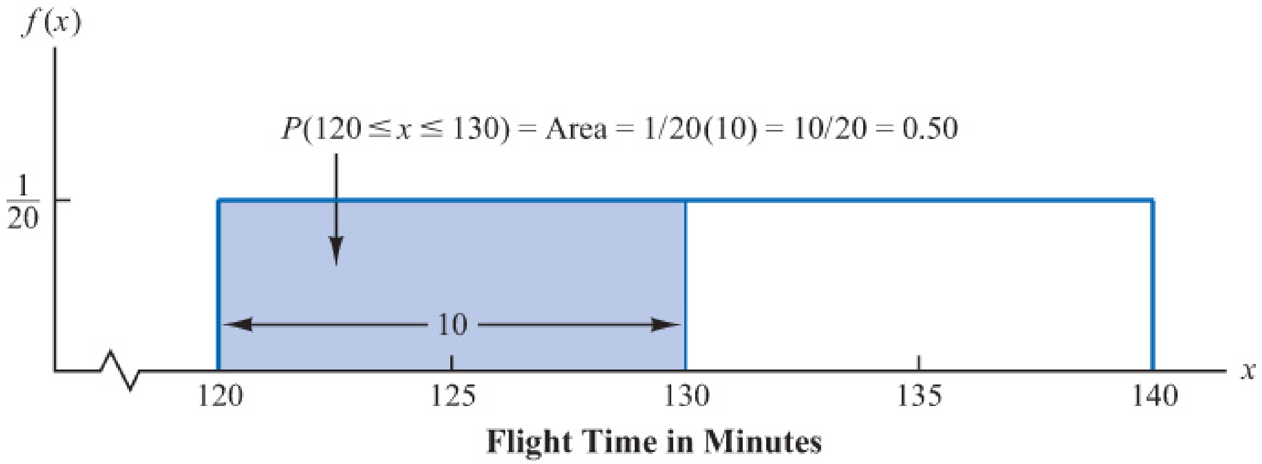
Thus, the random variable x is said to have the following uniform probability distribution

$$f(x) = \begin{cases} 1/20 & \text{for } 120 \leq x \leq 140 \\ 0 & \text{elsewhere} \end{cases}$$

What is the probability that the flight time is between 120 and 130 minutes?

As shown in the graph, the area is:

$$P(120 \leq x \leq 130) = 1/20(10) = 0.50$$



Properties of a Uniform Probability Distribution

All continuous probability distributions have the following properties:

1. $f(x) \geq 0$ for any value of x
2. the area under the graph of $f(x)$ is equal to 1

For example, in the airplane traveling time problem seen in the previous slide

$$P(120 \leq x \leq 140) = 1/20(140 - 120) = 1/20(20) = 1$$

Applying the formulas for expected value and standard deviation (as the square root of the variance) to the uniform distribution for flight times from Chicago to New York, we obtain

$$E(x) = \frac{a+b}{2} = \frac{120+140}{2} = \frac{260}{2} = 130 \text{ minutes}$$

$$\sigma = \sqrt{Var(x)} = \sqrt{\frac{(b-a)^2}{12}} = \sqrt{\frac{(140-120)^2}{12}} = \sqrt{\frac{400}{12}} = \sqrt{33.33} = 5.77 \text{ minutes}$$

Normal Probability Density Function

The **normal probability distribution** is the most common probability continuous distribution.

The normal distribution, used in a wide variety of practical applications, is widely applied to statistical inference, which is the major topic of the remainder of this book.

The normal distribution assumes the characteristic bell-shaped curve and provides a description of the likely results obtained through sampling.

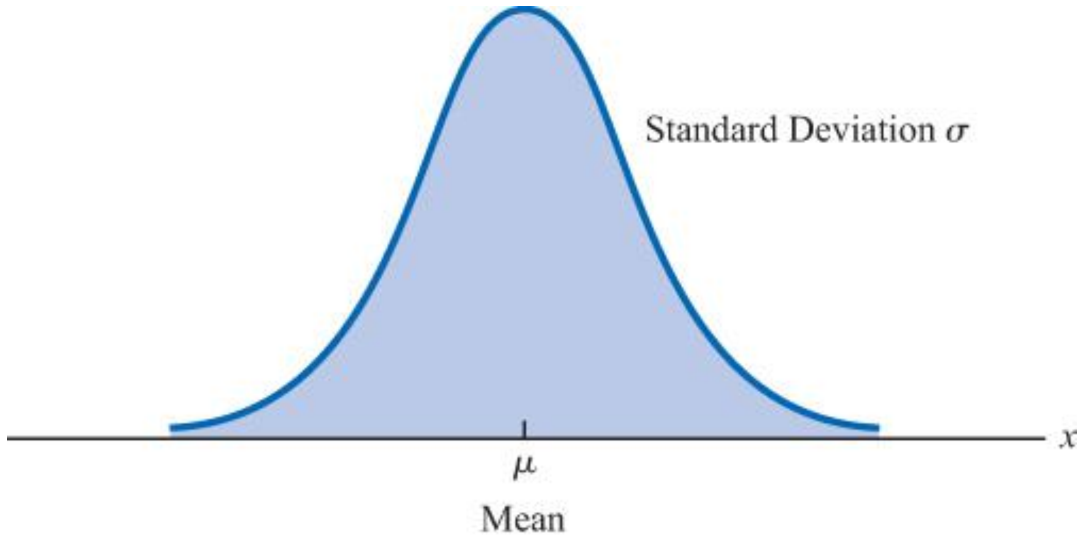
The *normal probability density function* is described by two parameters: the mean μ and the standard deviation σ.

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

Where

$\pi = 3.14159$ pi

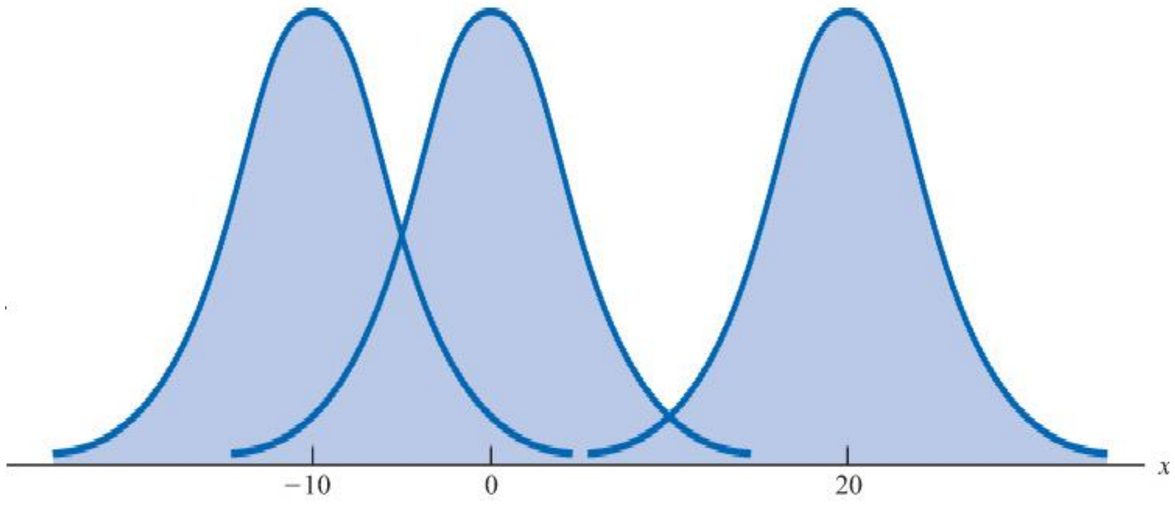
$e = 2.71828$ Euler's number



The Mean of the Normal Distribution

We make the following observations about the mean μ of the normal distribution:

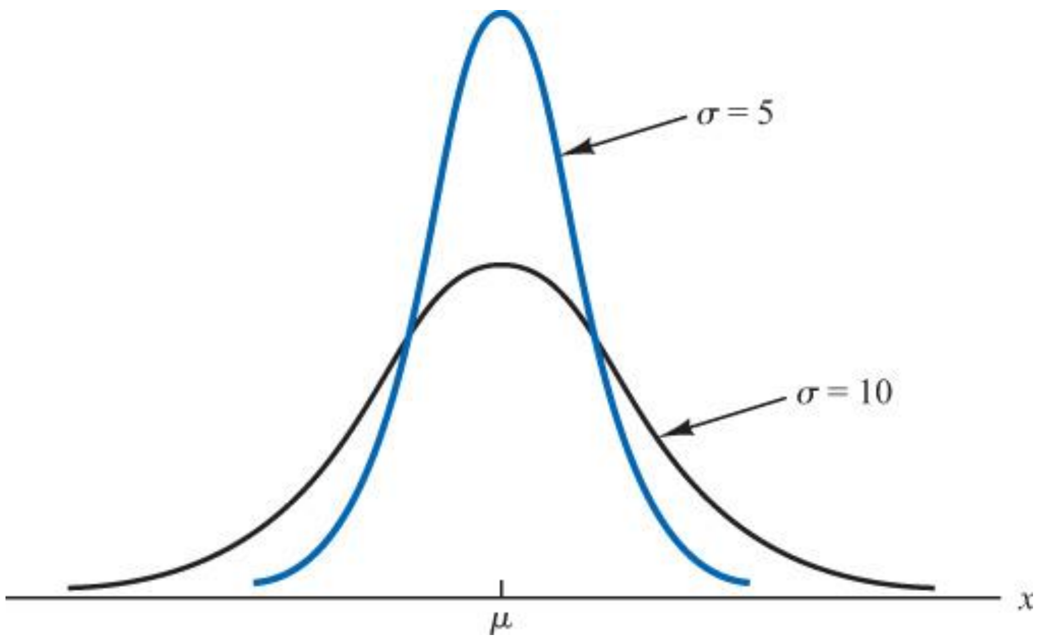
- The mean μ is the highest point on the normal curve.
- The mean μ is also the median and mode of the distribution.
- The mean of the distribution can be any numerical value: negative, zero, or positive. The figure shows three normal distributions with the same σ and $\mu = -10, 0,$ and 20.
- The normal distribution is symmetric.
- The area under the curve to the left of the mean and the area under the curve to the right of the mean are both 0.50.
- The tails of the normal curve extend to infinity in both directions and theoretically never touch the horizontal axis.



The Spread of the Normal Distribution

We make the following observations about the standard deviation σ of the normal distribution:

- The standard deviation determines how flat and wide the normal curve is.
  - Larger values of σ result in wider, flatter curves, showing more variability in the data.
  - The figure shows two normal distributions with the same mean but $\sigma = 5$ and $\sigma = 10$.
- The percentage of values in some typical intervals are:
  - 68.3% of the values of a normal random variable are within $\pm 1\ \sigma$ of μ.
  - 95.4% of the values of a normal random variable are within $\pm 2\ \sigma$ of μ.
  - 99.7% of the values of a normal random variable are within $\pm 3\ \sigma$ of μ.



Standard Normal Probability Distribution

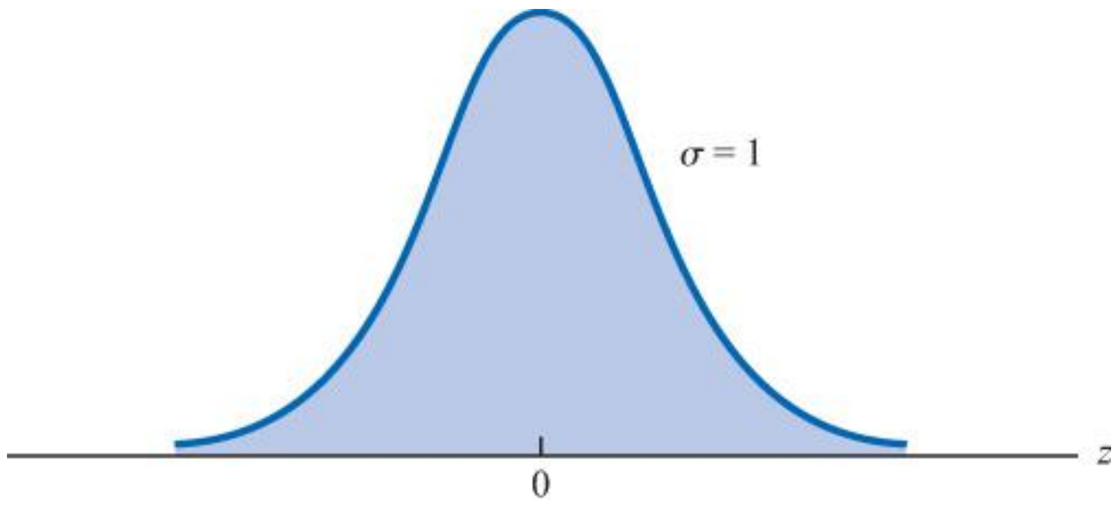
A random variable having a normal distribution with a mean of 0 and a standard deviation of 1 is said to have a **standard normal probability distribution**.

The letter z is commonly used to designate the standard normal random variable.

The standard normal distribution has the same general appearance as other normal distributions, but with the special properties of having $\mu = 0$ and $\sigma = 1$.

Because of $\mu = 0$ and $\sigma = 1$, the formula for the standard normal probability density function is a simpler version of the equation shown earlier for the general normal probability distribution:

$$\boldsymbol{f(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}}}$$



Probability Calculations with the Normal Distribution

As with other continuous random variables, probability calculations with a normal distribution are made by computing the area under the graph of the normal curve over a given interval.

Excel can be used to calculate the probabilities associated with the

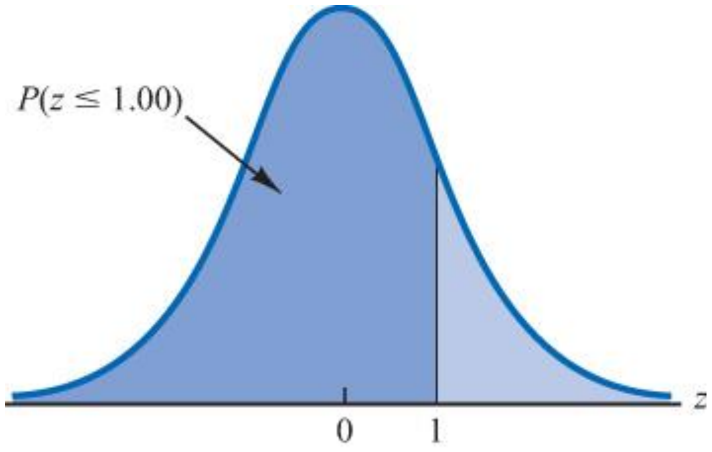
- normal probability distributions using the function $=\text{NORM.DIST}(x, \mu, \sigma, \text{TRUE})$
- the standard normal probability distribution using the function $=\text{NORM.S.DIST}(z, \text{TRUE})$

We will now show how to use the ‘Z table’ to calculate the following cumulative probabilities associated with a standard normal probability distribution

- the probability that the standard normal random variable z will be less than or equal to a given value Z: $P(z \leq Z)$
- the probability that z will be between two given values, Z_1 and Z_2: $P(Z_1 \leq z \leq Z_2)$
- the probability that z will be greater than or equal to a given value Z: $P(z \geq Z)$

Cumulative Probability for $P(z \leq 1.00)$

The cumulative probability that z is less than or equal to 1.00, that is, $P(z \leq 1.00)$, is the area under the standard normal curve to the left of $z = 1.00$.



To find the cumulative probability corresponding to $z \leq 1.00$ in the ‘Z table’, we identify the value located at the intersection of the row labeled 1.0 and the column labeled 0.00.

1. we find 1.0 in the left column of the table
2. we find 0.00 in the top row of the table.

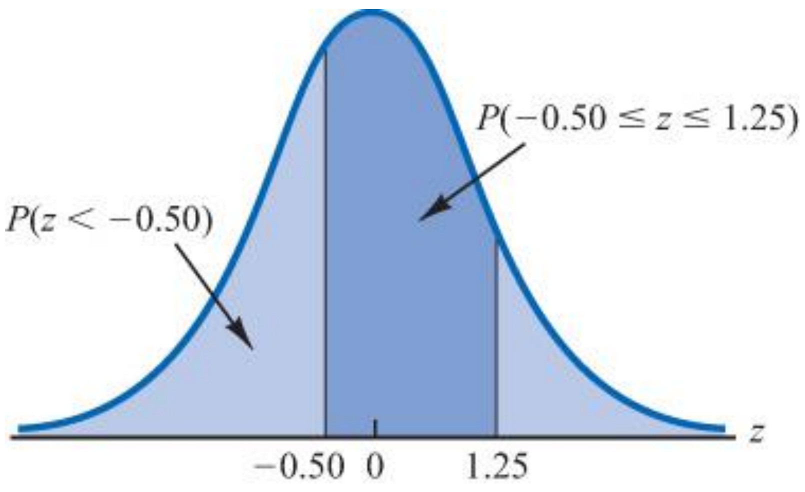
z	**0.00**	**0.01**	**0.02**	**...**	**0.09**
⋮	⋮	⋮	⋮		⋮
0.9	0.8159	0.8186	0.8212	...	0.8389
1.0	$\boxed{0.8413}$	0.8438	0.8461	...	0.8621
1.1	0.8643	0.8665	0.8686	...	0.8830
⋮	⋮	⋮	⋮		⋮

By looking in the body of the table, we find that the 1.0 row and the 0.00 column intersect at the value of 0.8413. Thus

$$P(z \leq 1.00) = 0.8413$$

Cumulative Probability for $P(-0.50 \leq z \leq 1.25)$

The cumulative probability that z is in the interval between –0.50 and 1,25, that is, $P(-0.50 \leq z \leq 1.25)$, is the area depicted in the figure to the right.



Three steps are required to compute this probability using Table 1 of Appendix B.

1. we calculate $P(z \leq -0.50)$ as the intersection of the –0.5 row and the 0.00 column:

   $$P(z \leq -0.50) = 0.3085$$

2. we calculate $P(z \leq 1.25)$ as the intersection of the 1.2 row and the 0.05 column:

   $$P(z \leq 1.25) = 0.8944$$

3. we calculate the difference of the cumulative probabilities to the left of $z = 1.25$ and $z = -0.50$ that we calculated in steps 1 and 2, so that:

   $$P(-0.50 \leq z \leq 1.25) = P(z \leq 1.25) - P(z \leq -0.50) = 0.8944 - 0.3085 = 0.5859$$

Cumulative Probability for $P(z \geq 1.58)$

The cumulative probability that z is greater than or equal to 1.58, that is, $P(z \geq 1.58)$, is the area under the standard normal curve to the right of $z = 1.58$.

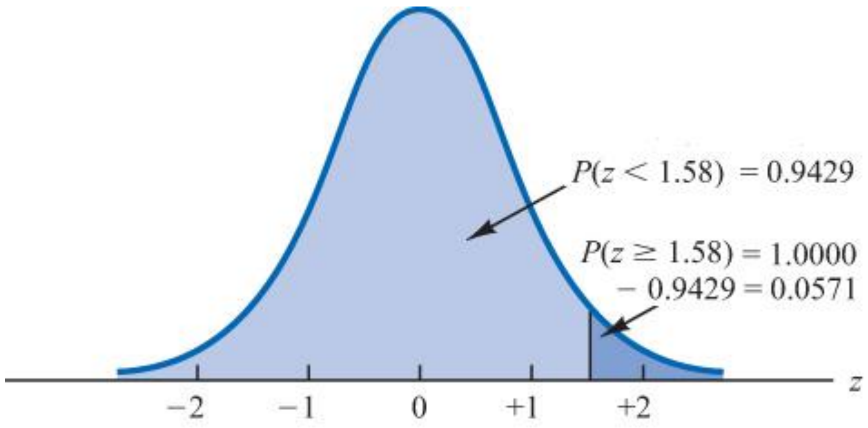
To find the probability to the right of $z = 1.58$, we need to first find the probability corresponding to the left of $z = 1.58$, and then calculate its complement.

To calculate the area to the left of $z = 1.58$, identify the value located at the intersection of the row labeled 1.5 and the column labeled 0.08. Thus

$$P(z \leq 1.58) = 0.9429$$

However, because the total area under the normal curve is 1, we have

$$P(z \geq 1.58) = 1 - 0.9429 = 0.0571$$



z	0.00	0.01	...	0.08	0.09
⋮	⋮	⋮		⋮	⋮
1.4	0.9192	0.9207	...	0.9306	0.9319
1.5	0.9332	0.9345	...	$\boxed{0.9429}$	0.9441
1.6	0.9452	0.9463	...	0.9535	0.9545
⋮	⋮	⋮		⋮	⋮

Find Z Such That $P(z \geq z^*) = 0.10$

We want to find a z^* value such that the probability of obtaining a larger z value is 0.10. That is

$$P(z \geq z^*) = 0.10$$

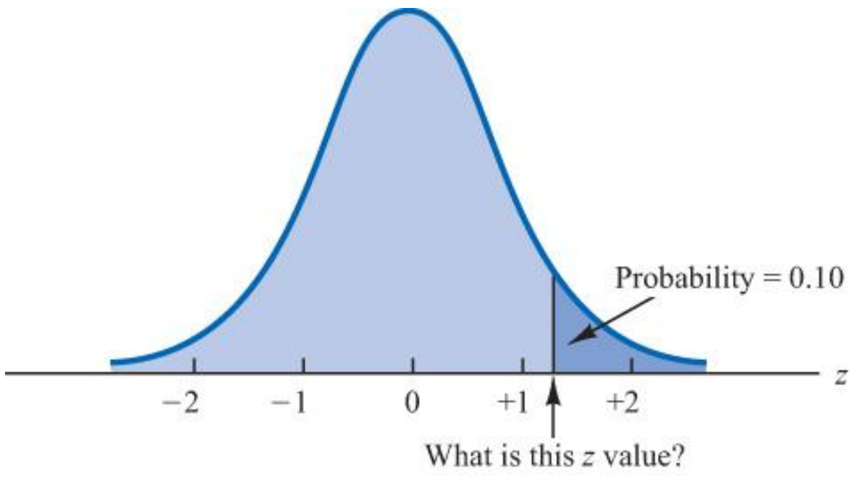
In this inverse problem, we need to find the z value that corresponds to a given probability.

However, the standard normal probability table gives the area under the curve to the *left* of a particular z value.

Thus, we use again the complement rule to find a z^* value such that

$$P(z \leq z^*) = 1 - P(z \geq z^*) = 1 - 0.10 = 0.90$$

'Z table' shows the closest probability to 0.9000 to be 0.8997, corresponding to the 1.2 row and the 0.08 column. That is, $z^* = 1.28$.



z	0.00	0.01	...	0.08	0.09
⋮	⋮	⋮		⋮	⋮
1.1	0.8643	0.8665	...	0.8810	0.8830
1.2	0.8849	0.8869	...	$\boxed{0.8997}$	0.9015
1.3	0.9032	0.9049	...	0.9162	0.9177
⋮	⋮	⋮		⋮	⋮

Computing Probabilities for Any Normal Curve

We use the following formula to convert any normal random variable x with mean μ and standard deviation σ to the standard normal random variable z (*see notes)

$$\boldsymbol{z = \frac{x - \mu}{\sigma}}$$

We see that a value of x equal to its mean μ corresponds to $z = 0$, and that a value of x that is one standard deviation above its mean μ, that is, $x = \mu + \sigma$, has z value

$$z = [(\mu + \sigma) - \mu]/\sigma = \sigma/\sigma = 1.$$

Thus, we can interpret z as the number of standard deviations that x is from its mean μ.

Example: find $P(10 \leq x \leq 14)$ for a normal random variable x described by $\mu = 10$ and $\sigma = 2$.

$$P(10 \leq x \leq 14) = P\left(\frac{10 - \mu}{\sigma} \leq z \leq \frac{14 - \mu}{\sigma}\right) = P\left(\frac{10 - 10}{2} \leq z \leq \frac{14 - 10}{2}\right)$$

$$= P(0 \leq z \leq 2) = P(z \leq 2) - P(z \leq 0) = 0.9772 - 0.5000 = 0.4772$$

Grear Tire Company Problem

Grear Tire Company has developed a new steel-belted radial tire to be sold through a national chain of discount stores. From tire road tests, Grear's engineering group estimated

mean tire mileage: $\mu = 36{,}500$ miles

standard deviation: $\sigma = 5{,}000$ miles

In addition, the data collected indicate that a normal distribution is a reasonable assumption.

Thus, we define a normal random variable as

x = number of miles the tire will last

Before finalizing the tire mileage guarantee policy, Grear's managers want an answer to the following two questions:

1. What is the probability that the tire mileage will exceed 40,000 miles?
2. What should the guaranteed mileage be if Grear wants no more than 10% of the tires to be eligible for a discount?

Probability Tire Mileage Exceeds 40,000 Miles

As depicted by the darker area in the figure, we need to find $P(x \geq 40{,}000)$.

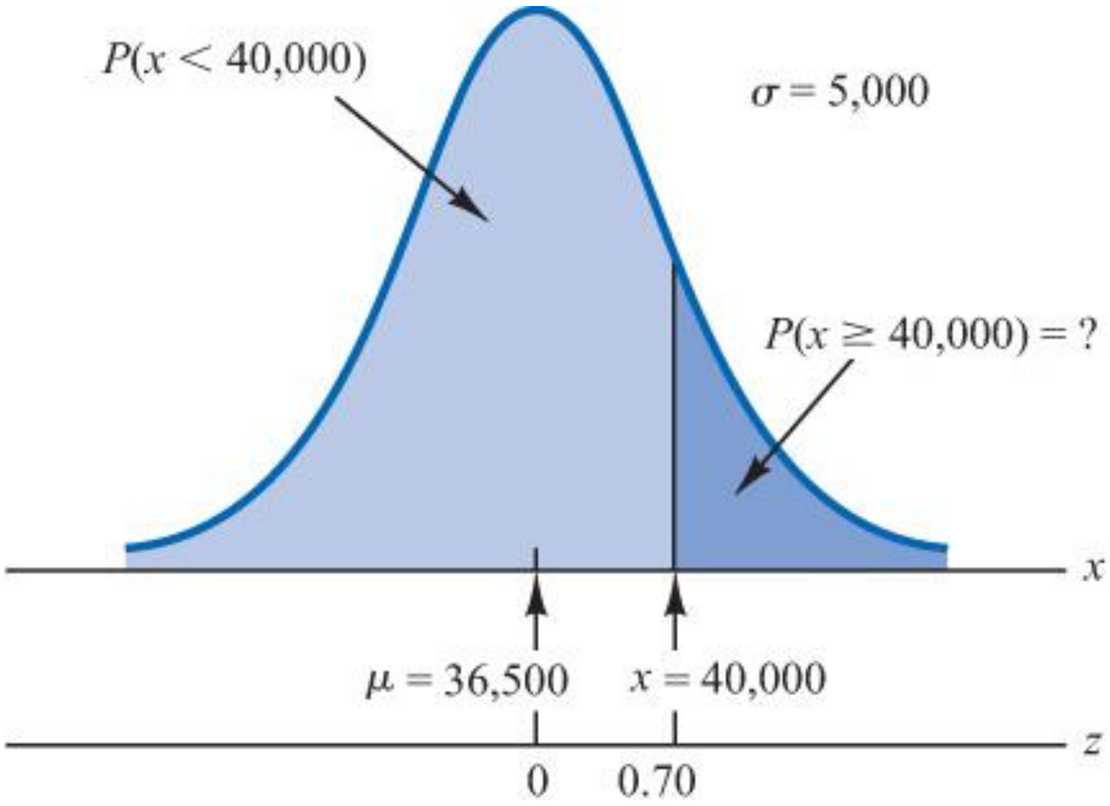
We can convert the normal random variable x to the standard normal random variable z, as

$$z = \frac{x - \mu}{\sigma} = \frac{40{,}000 - 36{,}500}{5{,}000} = \frac{3{,}500}{5{,}000} = 0.70$$

The probability that z is greater than or equal to 0.70 is the complement to the probability that z is less than or equal to 0.70. Thus,

$$P(x \geq 40{,}000) = P(z \geq 0.70)$$
$$= 1 - \mathrm{P}(\mathrm{z} < 0.70) = 1 - 0.7580 = 0.2420$$

We can conclude that about 24.2% of the tires will exceed 40,000 in mileage.



Grear Tire Guaranteed Mileage

In this inverse problem, we need to find a value of x^* such that: $P(x \leq x^*) = 0.10$

'Z table' shows the closest probability to 0.1000 to be 0.1003, corresponding to the –1.2 row and the 0.08 column. That is, $z^* = -1.28$.

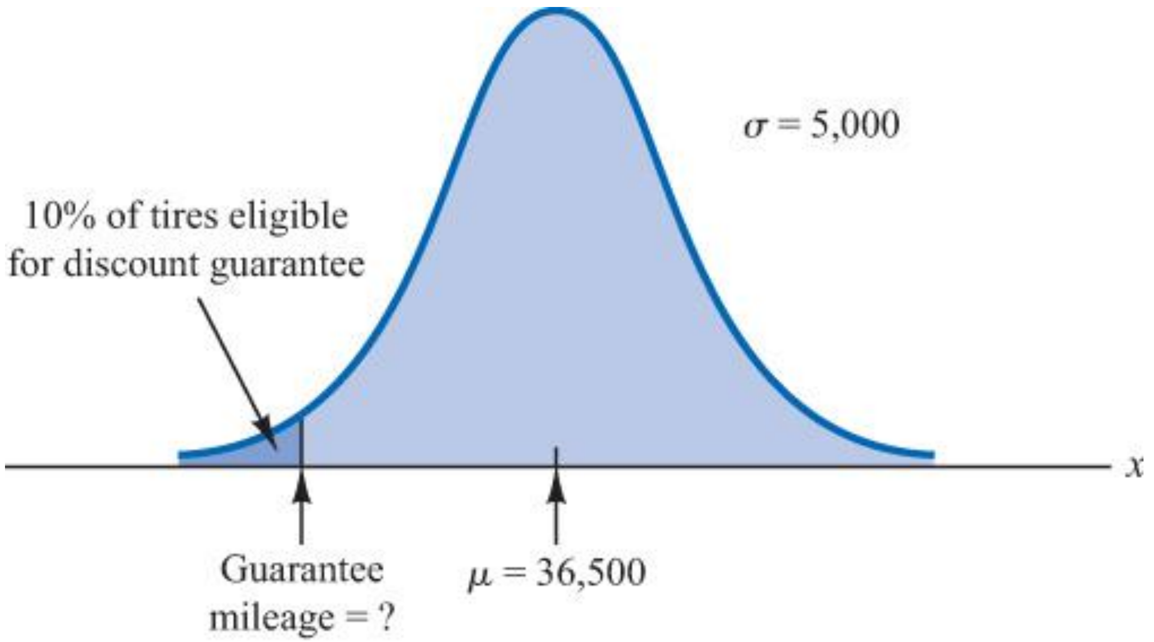
Thus, we have

$$z^* = \frac{x^* - \mu}{\sigma} = -1.28$$

Solving for x^*, we find

$$x^* = \mu - 1.28\sigma = 36{,}500 - 1.28(5{,}000) = 30{,}100 \approx 30{,}000 \text{ miles}$$

Thus, a guarantee of 30,000 miles will meet the requirement that no more than 10% of the tires will be eligible for the discount.



Normal Approximation of Binomial Probabilities

When the number of trials, n, in a binomial experiment becomes large, evaluating the binomial probability function by hand or with a calculator can become cumbersome.

In cases where $np \geq 5$, and $n(1-p) \geq 5$, where p represents the probability of success in a trial, the normal distribution provides an easy-to-use approximation of binomial probabilities.

When using the normal approximation to the binomial, we set a normal curve defined as

$$\boldsymbol{\mu = np}$$

$$\boldsymbol{\sigma = \sqrt{np(1-p)}}$$

Let us illustrate the normal approximation to the binomial distribution by supposing that a particular company has a history of making errors in 10% of its invoices ($p = 0.10$.)

Using a sample of $n = 100$ invoices, we want to compute the following probabilities:

1. The probability that $x = 12$ of the 100 invoices contain errors.
2. The probability that 13 or fewer ($x \leq 13$) of the 100 invoices contain errors.

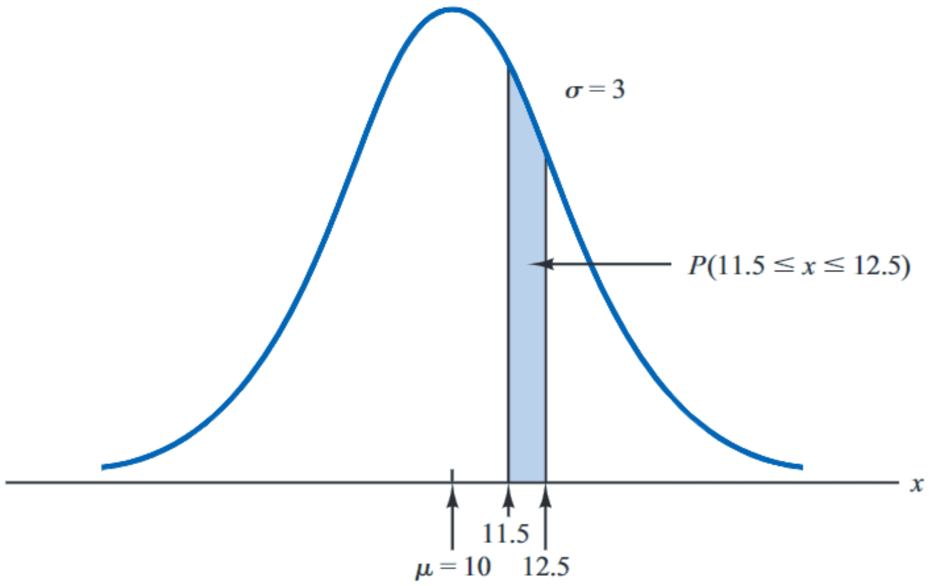
Probability that 12 of 100 Invoices Contain Errors

To find the binomial probability of $x = 12$ successes in $n = 100$ trials, we can apply the normal approximation of the binomial distribution because $np = 100(0.10) = 10 \geq 5$, and $n(1-p) = 100(1 - 0.10) = 90 \geq 5$.

Thus, mean and standard deviation are

$$\mu = np = 100(0.10) = 10$$

$$\sigma = \sqrt{np(1-p)} = \sqrt{100(0.10)(0.90)} = \sqrt{9} = 3$$



To approximate the binomial probability of $x = 12$, we apply the **continuity correction factor** and compute the area under the corresponding normal curve between 11.5 and 12.5.

$$P(11.5 \leq x \leq 12.5) = P[(11.5 - 10)/3 \leq z \leq (12.5 - 10)/3]$$

$$= P(0.5 \leq z \leq 0.83) = P(z \leq 0.83) - P(z \leq 0.50) = 0.7967 - 0.6915 = 0.1052$$

Probability that 13 or Fewer Invoices Contain Errors

The figure shows the area under the normal curve that approximates the probability that 13 of fewer invoices out of a sample of 100 contain errors.

Because of the continuity correction factor introduced in the previous slide, we set $x = 13.5$.

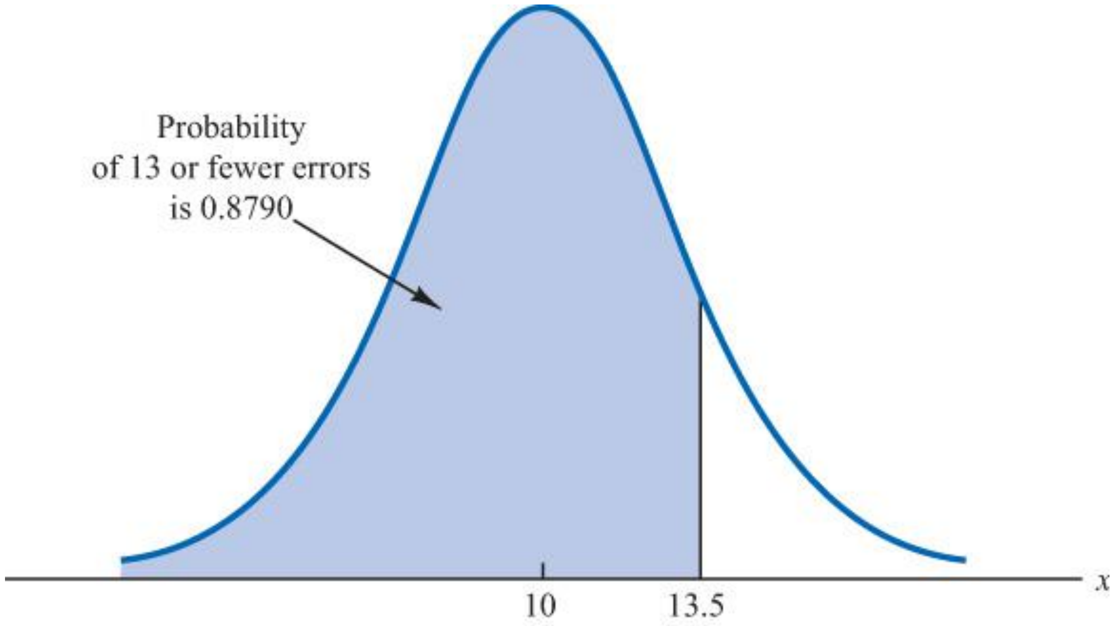
The z-score can be calculated as:

$$z = \frac{x - \mu}{\sigma} = \frac{13.5 - 10}{3} = \frac{3.5}{3} = 1.17$$

Thus, the probability to have 13 or fewer invoices is

$$P(x \leq 13.5) = P(z \leq 1.17) = 0.8790$$

We can conclude that about 88% of the invoices contain 13 or fewer errors.



Exponential Probability Distribution

The **exponential probability distribution** may be used for random variables that describe the length of the interval between occurrences, such as the time between arrivals at a hospital emergency room or the distance between major defects in a highway.

The exponential probability density function is

$$\boldsymbol{f(x) = \frac{1}{\mu} e^{-\frac{x}{\mu}}} \qquad \text{for } x \geq 0$$

Where μ is the expected value or mean.

For example, if the loading time x for a truck at the Schips loading dock has mean $\mu = 15$ minutes, and follows an exponential distribution, the probability density function is

$$f(x) = \frac{1}{15} e^{-\frac{x}{15}}$$

Let us calculate the probability that the loading time x is between 6 and 18 minutes.

Exponential Distribution: Cumulative Probability

To compute exponential probabilities, we use the formula for the cumulative probability of obtaining a value for the exponential random variable that is less than or equal to some specific value denoted by x_0.

$$\boldsymbol{P(x \leq x_0) = 1 - e^{-\frac{x_0}{\mu}}}$$

For the Schips loading dock example, with x = loading time in minutes and $\mu = 15$ minutes, we have

$$P(x \leq 6) = 1 - e^{-\frac{6}{15}} = 0.3297$$

$$P(x \leq 18) = 1 - e^{-\frac{18}{15}} = 0.6988$$

Thus,

$$P(6 \leq x \leq 18) = P(x \leq 18) - P(x \leq 6)$$

$$= 0.6988 - 0.3297 = 0.3691$$