# 4 The Normal Distribution

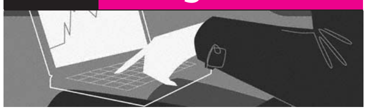

4–1 Using Statistics 147
4–2 Properties of the Normal Distribution 148
4–3 The Standard Normal Distribution 151
4–4 The Transformation of Normal Random Variables 156
4–5 The Inverse Transformation 162
4–6 The Template 166
4–7 Normal Approximation of Binomial Distributions 169
4–8 Using the Computer 171
4–9 Summary and Review of Terms 172
*Case 4* Acceptable Pins 177
*Case 5* Multicurrency Decision 177


## LEARNING OBJECTIVES

*After studying this chapter, you should be able to:*

- Identify when a random variable will be normally distributed.
- Use the properties of the normal distribution.
- Explain the significance of the standard normal distribution.
- Compute probabilities using normal distribution tables.
- Transform a normal distribution into a standard normal distribution.
- Convert a binomial distribution into an approximated normal distribution.
- Solve normal distribution problems using spreadsheet templates.

# 4–1 Using Statistics



The **normal distribution** is an important continuous distribution because a good number of random variables occurring in practice can be approximated to it. *If a random variable is affected by many independent causes, and the effect of each cause is not overwhelmingly large compared to other effects, then the random variable will closely follow a normal distribution.* The lengths of pins made by an automatic machine, the times taken by an assembly worker to complete the assigned task repeatedly, the weights of baseballs, the tensile strengths of a batch of bolts, and the volumes of soup in a particular brand of canned soup are good examples of normally distributed random variables. All of these are affected by several independent causes where the effect of each cause is small. For example, the length of a pin is affected by many independent causes such as vibrations, temperature, wear and tear on the machine, and raw material properties.

Additionally, in the next chapter, on sampling theory, we shall see that many of the sample statistics are normally distributed.

For a normal distribution with mean $\mu$ and standard deviation $\sigma$, the probability density function $f(x)$ is given by the complicated formula

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} \qquad -\infty < x < +\infty \tag{4–1}$$

In equation 4–1, $e$ is the natural base logarithm, equal to 2.71828 . . . By substituting desired values for $\mu$ and $\sigma$, we can get any desired density function. For example, a distribution with mean 100 and standard deviation 5 will have the density function

$$f(x) = \frac{1}{\sqrt{2\pi}5}e^{-\frac{1}{2}\left(\frac{x-100}{5}\right)^2} \qquad -\infty < x < +\infty \tag{4–2}$$

This function is plotted in Figure 4–1. This is the famous bell-shaped normal curve.

Over the years, many mathematicians have worked on the mathematics behind the normal distribution and have made many independent discoveries. The discovery

**FIGURE 4–1** **A Normal Distribution with Mean 100 and Standard Deviation 5**

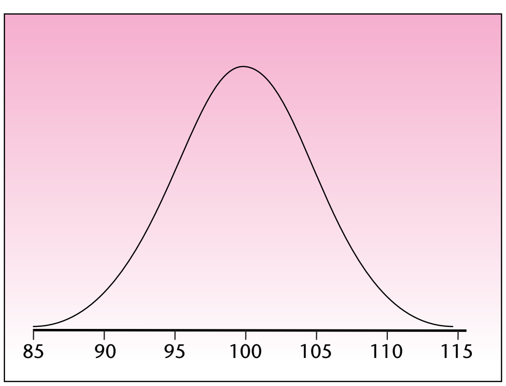
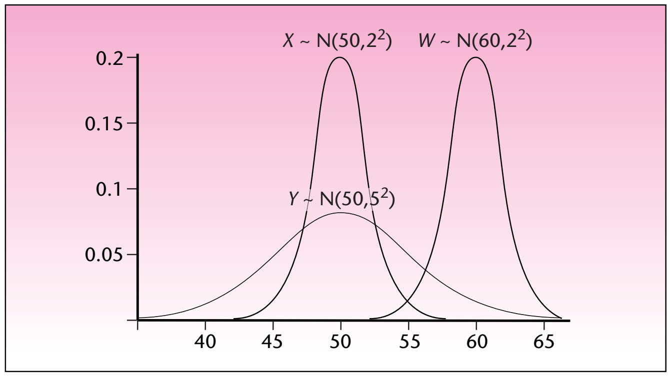


FIGURE 4–2 Three Normal Distributions




of equation 4–1 for the normal density function is attributed to Carl Friedrich Gauss (1777–1855), who did much work with the formula. In science books, this distribution is often called the *Gaussian distribution*. But the formula was first discovered by the French-born English mathematician Abraham De Moivre (1667–1754). Unfortunately for him, his discovery was not discovered until 1924.

As seen in Figure 4–1, the normal distribution is symmetric about its mean. It has a (relative) kurtosis of 0, which means it has average peakedness. The curve reaches its peak at the mean of 100, and therefore its mode is 100. Due to symmetry, its median is 100 too. In the figure the curve seems to touch the horizontal axis at 85 on the left and at 115 on the right; these points are 3 standard deviations away from the center on either side. Theoretically, the curve never touches the horizontal axis and extends to infinity on both sides.

If $X$ is normally distributed with mean $\mu$ and variance $\sigma^2$, we write $X \sim N(\mu, \sigma^2)$. If the mean is 100 and the variance is 9, we write $X \sim N(100, 3^2)$. Note how the variance is written. By writing 9 as $3^2$, we explicitly show that the standard deviation is 3. Figure 4–2 shows three normal distributions: $X \sim N(50, 2^2)$; $Y \sim N(50, 5^2)$; $W \sim N(60, 2^2)$. Note their shapes and positions.

## 4–2 Properties of the Normal Distribution

V S CHAPTER 5

There is a remarkable property possessed only by the normal distribution:

If several *independent* random variables are normally distributed, then their sum will also be normally distributed. The mean of the sum will be the sum of all the individual means, and by virtue of the independence, the variance of the sum will be the sum of all the individual variances.

We can write this in algebraic form as

If $X_1, X_2, \ldots, X_n$ are independent random variables that are normally distributed, then their sum $S$ will also be normally distributed with

$$E(S) = E(X_1) + E(X_2) + \cdots + E(X_n)$$

and

$$V(S) = V(X_1) + V(X_2) + \cdots + V(X_n)$$

Note that it is the *variances* that can be added as in the preceding box, and *not the standard deviations*. We will never have an occasion to add standard deviations.

We see intuitively that the sum of many normal random variables will also be normally distributed, because the sum is affected by many independent individual causes, namely, those causes that affect each of the original random variables.

Let us see the application of this result through a few examples.

**EXAMPLE 4–1**

Let $X_1$, $X_2$, and $X_3$ be independent random variables that are normally distributed with means and variances as follows:

| | Mean | Variance |
|---|---|---|
| $X_1$ | 10 | 1 |
| $X_2$ | 20 | 2 |
| $X_3$ | 30 | 3 |

Find the distribution of the sum $S = X_1 + X_2 + X_3$. Report the mean, variance, and standard deviation of $S$.

***Solution***

The sum $S$ will be normally distributed with mean $10 + 20 + 30 = 60$ and variance $1 + 2 + 3 = 6$. The standard deviation of $S = \sqrt{6} = 2.45$.

**EXAMPLE 4–2**

The weight of a module used in a spacecraft is to be closely controlled. Since the module uses a bolt-nut-washer assembly in numerous places, a study was conducted to find the distribution of the weights of these parts. It was found that the three weights, in grams, are normally distributed with the following means and variances:

| | Mean | Variance |
|---|---|---|
| Bolt | 312.8 | 2.67 |
| Nut | 53.2 | 0.85 |
| Washer | 17.5 | 0.21 |

Find the distribution of the weight of the assembly. Report the mean, variance, and standard deviation of the weight.

***Solution***

The weight of the assembly is the sum of the weights of the three component parts, which are three normal random variables. Furthermore, the individual weights are independent since the weight of any one component part does not influence the weight of the other two. Therefore, the weight of the assembly will be normally distributed.

The mean weight of the assembly will be the sum of the mean weights of the individual parts: $312.8 + 53.2 + 17.5 = 383.5$ grams.

The variance will be the sum of the individual variances: $2.67 + 0.85 + 0.21 = 3.73$ gram$^2$.

The standard deviation $= \sqrt{3.73} = 1.93$ grams.

Another interesting property of the normal distribution is that if $X$ is normally distributed, then $aX + b$ will also be normally distributed with mean $aE(X) + b$ and variance $a^2 V(X)$. For example, if $X$ is normally distributed with mean 10 and variance 3, then $4X + 5$ will be normally distributed with mean $4 * 10 + 5 = 45$ and variance $4^2 * 3 = 48$.

We can combine the above two properties and make the following statement:

If $X_1, X_2, \ldots, X_n$ are independent random variables that are normally distributed, then the random variable $Q$ defined as $Q = a_1X_1 + a_2X_2 + \cdots + a_nX_n + b$ will also be normally distributed with

$$E(Q) = a_1E(X_1) + a_2E(X_2) + \cdots + a_nE(X_n) + b$$

and

$$V(Q) = a_1^2V(X_1) + a_2^2V(X_2) + \cdots + a_n^2V(X_n)$$

The application of this result is illustrated in the following sample problems.

**EXAMPLE 4–3**

The four independent normal random variables $X_1$, $X_2$, $X_3$, and $X_4$ have the following means and variances:

| | Mean | Variance |
|---|---|---|
| $X_1$ | 12 | 4 |
| $X_2$ | −5 | 2 |
| $X_3$ | 8 | 5 |
| $X_4$ | 10 | 1 |

Find the mean and variance of $Q = X_1 - 2X_2 + 3X_3 - 4X_4 + 5$. Find also the standard deviation of $Q$.

***Solution***

$$E(Q) = 12 - 2(-5) + 3(8) - 4(10) + 5 = 12 + 10 + 24 - 40 + 5 = 11$$
$$V(Q) = 4 + (-2)^2(2) + 3^2(5) + (-4)^2(1) = 4 + 8 + 45 + 16 = 73$$
$$\text{SD}(Q) = \sqrt{73} = 8.544$$

**EXAMPLE 4–4**

A cost accountant needs to forecast the unit cost of a product for next year. He notes that each unit of the product requires 12 hours of labor and 5.8 pounds of raw material. In addition, each unit of the product is assigned an overhead cost of \$184.50. He estimates that the cost of an hour of labor next year will be normally distributed with an expected value of \$45.75 and a standard deviation of \$1.80; the cost of the raw material will be normally distributed with an expected value of \$62.35 and a standard deviation of \$2.52. Find the distribution of the unit cost of the product. Report its expected value, variance, and standard deviation.

***Solution***

Let $L$ be the cost of labor and $M$ be the cost of the raw material. Denote the unit cost of the product by $Q$. Then $Q = 12L + 5.8M + 184.50$. Since the cost of labor $L$ may not influence the cost of raw material $M$, we can assume that the two are independent. This makes the unit cost of the product $Q$ a normal random variable. Then

$$E(Q) = 12 \times 45.75 + 5.8 \times 62.35 + 184.50 = \$1095.13$$
$$V(Q) = 12^2 \times 1.80^2 + 5.8^2 \times 2.52^2 = 680.19$$
$$\text{SD}(Q) = \sqrt{680.19} = \$26.08$$

## 4–3 The Standard Normal Distribution

Since, as noted earlier, infinitely many normal random variables are possible, one is selected to serve as our *standard*. Probabilities associated with values of this standard normal random variable are tabulated. A special transformation then allows us to apply the tabulated probabilities to *any* normal random variable. The standard normal random variable has a special name, $Z$ (rather than the general name $X$ we use for other random variables).

We define the **standard normal random variable *Z*** as the normal random variable with mean $\mu = 0$ and standard deviation $\sigma = 1$.

In the notation established in the previous section, we say

$$Z \sim N(0, 1^2) \quad (4\text{–}3)$$

Since $1^2 = 1$, we may drop the superscript 2 as no confusion of the standard deviation and the variance is possible. A graph of the standard normal density function is given in Figure 4–3.

### *Finding Probabilities of the Standard Normal Distribution*

Probabilities of intervals are areas under the density $f(z)$ over the intervals in question. From the range of values in equation 4–1, $-\infty < x < \infty$, we see that any normal random variable is defined over the entire real line. Thus, the intervals in which we will be interested are sometimes *semi-infinite* intervals, such as $a$ to $\infty$ or $-\infty$ to $b$ (where $a$ and $b$ are numbers). While such intervals have infinite length, the probabilities associated with them are finite; they are, in fact, no greater than 1.00, as required of all probabilities. The reason for this is that the area in either of the "tails" of the distribution (the two narrow ends of the distribution, extending toward $-\infty$ and $+\infty$) becomes very small very quickly as we move away from the center of the distribution.

Tabulated areas under the standard normal density are probabilities of intervals extending from the mean $\mu = 0$ to points $z$ to its right. Table 2 in Appendix C gives areas under the standard normal curve between 0 and points $z > 0$. The total area under the normal curve is equal to 1.00, and since the curve is symmetric, the area from 0 to $-\infty$ is equal to 0.5. The *table area* associated with a point $z$ is thus equal to the value of the cumulative distribution function $F(z)$ minus 0.5.

We define the **table area** as

$$TA = F(z) - 0.5 \quad (4\text{–}4)$$

**FIGURE 4–3 The Standard Normal Density Function**

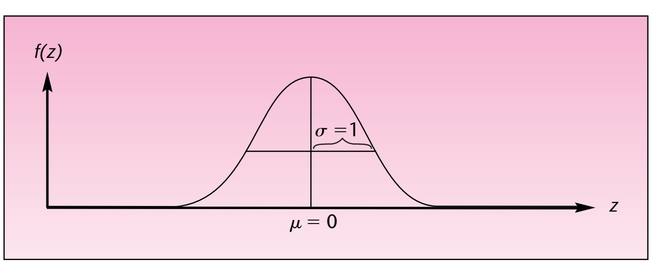


**FIGURE 4–4** **The Table Area TA for a Point *z* of the Standard Normal Distribution**

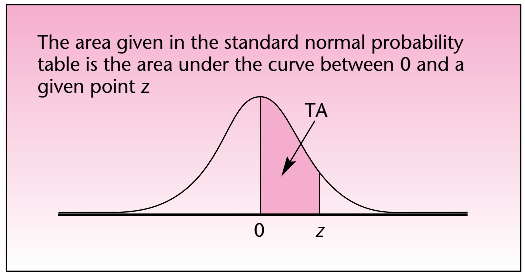
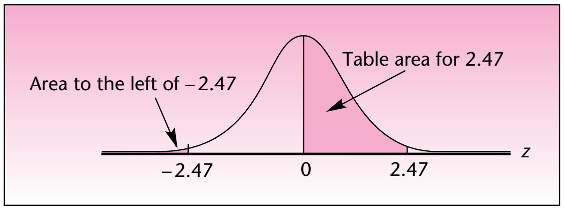



**TABLE 4–1** **Standard Normal Probabilities**

| z | .00 | .01 | .02 | .03 | .04 | .05 | .06 | .07 | .08 | .09 |
|---|---|---|---|---|---|---|---|---|---|---|
| 0.0 | .0000 | .0040 | .0080 | .0120 | .0160 | .0199 | .0239 | .0279 | .0319 | .0359 |
| 0.1 | .0398 | .0438 | .0478 | .0517 | .0557 | .0596 | .0636 | .0675 | .0714 | .0753 |
| 0.2 | .0793 | .0832 | .0871 | .0910 | .0948 | .0987 | .1026 | .1064 | .1103 | .1141 |
| 0.3 | .1179 | .1217 | .1255 | .1293 | .1331 | .1368 | .1406 | .1443 | .1480 | .1517 |
| 0.4 | .1554 | .1591 | .1628 | .1664 | .1700 | .1736 | .1772 | .1808 | .1844 | .1879 |
| 0.5 | .1915 | .1950 | .1985 | .2019 | .2054 | .2088 | .2123 | .2157 | .2190 | .2224 |
| 0.6 | .2257 | .2291 | .2324 | .2357 | .2389 | .2422 | .2454 | .2486 | .2517 | .2549 |
| 0.7 | .2580 | .2611 | .2642 | .2673 | .2704 | .2734 | .2764 | .2794 | .2823 | .2852 |
| 0.8 | .2881 | .2910 | .2939 | .2967 | .2995 | .3023 | .3051 | .3078 | .3106 | .3133 |
| 0.9 | .3159 | .3186 | .3212 | .3238 | .3264 | .3289 | .3315 | .3340 | .3365 | .3389 |
| 1.0 | .3413 | .3438 | .3461 | .3485 | .3508 | .3531 | .3554 | .3577 | .3599 | .3621 |
| 1.1 | .3643 | .3665 | .3686 | .3708 | .3729 | .3749 | .3770 | .3790 | .3810 | .3830 |
| 1.2 | .3849 | .3869 | .3888 | .3907 | .3925 | .3944 | .3962 | .3980 | .3997 | .4015 |
| 1.3 | .4032 | .4049 | .4066 | .4082 | .4099 | .4115 | .4131 | .4147 | .4162 | .4177 |
| 1.4 | .4192 | .4207 | .4222 | .4236 | .4251 | .4265 | .4279 | .4292 | .4306 | .4319 |
| 1.5 | .4332 | .4345 | .4357 | .4370 | .4382 | .4394 | .4406 | .4418 | .4429 | .4441 |
| 1.6 | .4452 | .4463 | .4474 | .4484 | .4495 | .4505 | .4515 | .4525 | .4535 | .4545 |
| 1.7 | .4554 | .4564 | .4573 | .4582 | .4591 | .4599 | .4608 | .4616 | .4625 | .4633 |
| 1.8 | .4641 | .4649 | .4656 | .4664 | .4671 | .4678 | .4686 | .4693 | .4699 | .4706 |
| 1.9 | .4713 | .4719 | .4726 | .4732 | .4738 | .4744 | .4750 | .4756 | .4761 | .4767 |
| 2.0 | .4772 | .4778 | .4783 | .4788 | .4793 | .4798 | .4803 | .4808 | .4812 | .4817 |
| 2.1 | .4821 | .4826 | .4830 | .4834 | .4838 | .4842 | .4846 | .4850 | .4854 | .4857 |
| 2.2 | .4861 | .4864 | .4868 | .4871 | .4875 | .4878 | .4881 | .4884 | .4887 | .4890 |
| 2.3 | .4893 | .4896 | .4898 | .4901 | .4904 | .4906 | .4909 | .4911 | .4913 | .4916 |
| 2.4 | .4918 | .4920 | .4922 | .4925 | .4927 | .4929 | .4931 | .4932 | .4934 | .4936 |
| 2.5 | .4938 | .4940 | .4941 | .4943 | .4945 | .4946 | .4948 | .4949 | .4951 | .4952 |
| 2.6 | .4953 | .4955 | .4956 | .4957 | .4959 | .4960 | .4961 | .4962 | .4963 | .4964 |
| 2.7 | .4965 | .4966 | .4967 | .4968 | .4969 | .4970 | .4971 | .4972 | .4973 | .4974 |
| 2.8 | .4974 | .4975 | .4976 | .4977 | .4977 | .4978 | .4979 | .4979 | .4980 | .4981 |
| 2.9 | .4981 | .4982 | .4982 | .4983 | .4984 | .4984 | .4985 | .4985 | .4986 | .4986 |
| 3.0 | .4987 | .4987 | .4987 | .4988 | .4988 | .4989 | .4989 | .4989 | .4990 | .4990 |

The table area TA is shown in Figure 4–4. Part of Table 2 is reproduced here as Table 4–1. Let us see how the table is used in obtaining probabilities for the standard normal random variable. In the following examples, refer to Figure 4–4 and Table 4–1.

**FIGURE 4–5** Finding the Probability That *Z* Is Less Than −2.47




1. Let us find the probability that the value of the standard normal random variable will be between 0 and 1.56. That is, we want $P(0 < Z < 1.56)$. In Figure 4–4, substitute 1.56 for the point $z$ on the graph. We are looking for the table area in the row labeled 1.5 and the column labeled 0.06. In the table, we find the probability 0.4406.
2. Let us find the probability that $Z$ will be less than −2.47. Figure 4–5 shows the required area for the probability $P(Z < -2.47)$. By the symmetry of the normal curve, the area to the left of −2.47 is exactly equal to the area to the right of 2.47. We find

$$P(Z < -2.47) = P(Z > 2.47) = 0.5000 - 0.4932 = 0.0068$$

3. Find $P(1 < Z < 2)$. The required probability is the area under the curve between the two points 1 and 2. This area is shown in Figure 4–6. The table gives us the area under the curve between 0 and 1, and the area under the curve between 0 and 2. Areas are additive; therefore, $P(1 < Z < 2) = \text{TA(for } 2.00) - \text{TA(for } 1.00) = 0.4772 - 0.3413 = 0.1359$.

In cases where we need probabilities based on values with greater than second-decimal accuracy, we may use a linear interpolation between two probabilities obtained from the table. For example, $P(0 \leq Z \leq 1.645)$ is found as the midpoint between the two probabilities $P(0 \leq Z \leq 1.64)$ and $P(0 \leq Z \leq 1.65)$. This is found, using the table, as the midpoint of 0.4495 and 0.4505, which is 0.45. If even greater accuracy is required, we may use computer programs designed to produce standard normal probabilities.

### *Finding Values of Z Given a Probability*

In many situations, instead of finding the probability that a standard normal random variable will be within a given interval, we may be interested in the reverse: finding an interval with a given probability. Consider the following examples.

**FIGURE 4–6** Finding the Probability That *Z* Is between 1 and 2

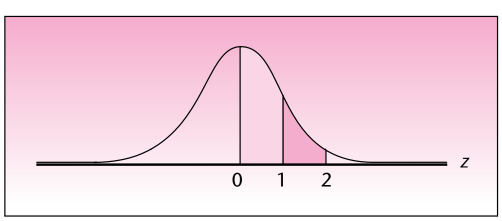


**FIGURE 4–7** **Using the Normal Table to Find a Value, Given a Probability**

| z | .00 | .01 | .02 | .03 | .04 | .05 | .06 | .07 | .08 | .09 |
|---|---|---|---|---|---|---|---|---|---|---|
| 0.0 | .0000 | .0040 | .0080 | .0120 | .0160 | .0199 | .0239 | .0279 | .0319 | .0359 |
| 0.1 | .0398 | .0438 | .0478 | .0517 | .0557 | .0596 | .0636 | .0675 | .0714 | .0753 |
| 0.2 | .0793 | .0832 | .0871 | .0910 | .0948 | .0987 | .1026 | .1064 | .1103 | .1141 |
| 0.3 | .1179 | .1217 | .1255 | .1293 | .1331 | .1368 | .1406 | .1443 | .1480 | .1517 |
| 0.4 | .1554 | .1591 | .1628 | .1664 | .1700 | .1736 | .1772 | .1808 | .1844 | .1879 |
| 0.5 | .1915 | .1950 | .1985 | .2019 | .2054 | .2088 | .2123 | .2157 | .2190 | .2224 |
| 0.6 | .2257 | .2291 | .2324 | .2357 | .2389 | .2422 | .2454 | .2486 | .2517 | .2549 |
| 0.7 | .2580 | .2611 | .2642 | .2673 | .2704 | .2734 | .2764 | .2794 | .2823 | .2852 |
| 0.8 | .2881 | .2910 | .2939 | .2967 | .2995 | .3023 | .3051 | .3078 | .3106 | .3133 |
| 0.9 | .3159 | .3186 | .3212 | .3238 | .3264 | .3289 | .3315 | .3340 | .3365 | .3389 |
| 1.0 | .3413 | .3438 | .3461 | .3485 | .3508 | .3531 | .3554 | .3577 | .3599 | .3621 |
| 1.1 | .3643 | .3665 | .3686 | .3708 | .3729 | .3749 | .3770 | .3790 | .3810 | .3830 |
| → 1.2 | .3849 | .3869 | .3888 | .3907 | .3925 | .3944 | .3962 | .3980 | .3997 | .4015 |
| 1.3 | .4032 | .4049 | .4066 | .4082 | .4099 | .4115 | .4131 | .4147 | .4162 | .4177 |
| 1.4 | .4192 | .4207 | .4222 | .4236 | .4251 | .4265 | .4279 | .4292 | .4306 | .4319 |
| 1.5 | .4332 | .4345 | .4357 | .4370 | .4382 | .4394 | .4406 | .4418 | .4429 | .4441 |

1. Find a value $z$ of the standard normal random variable such that the probability that the random variable will have a value between 0 and $z$ is 0.40. We look *inside* the table for the value closest to 0.40; we do this by searching through the values inside the table, noting that they increase from 0 to numbers close to 0.5000 as we go down ↓ the columns and across the rows. The closest value we find to 0.40 is the table area .3997. This value corresponds to 1.28 (row 1.2 and column .08). This is illustrated in Figure 4–7.
2. Find the value of the standard normal random variable that cuts off an area of 0.90 to its left. Here, we reason as follows: Since the area to the left of the given point $z$ is greater than 0.50, *$z$ must be on the right side of 0.* Furthermore, the area to the left of 0 all the way to $-\infty$ is equal to 0.5. Therefore, TA $= 0.9 - 0.5 = 0.4$. We need to find the point $z$ such that TA $= 0.4$. We know the answer from the preceding example: $z = 1.28$. This is shown in Figure 4–8.
3. Find a 0.99 probability interval, symmetric about 0, for the standard normal random variable. The required area between the two $z$ values that are equidistant from 0 on either side is 0.99. Therefore, the area under the curve between 0 and the positive $z$ value is TA $= 0.99/2 = 0.495$. We now look in our normal probability table for the area closest to 0.495. The area 0.495 lies exactly between the two areas 0.4949 and 0.4951, corresponding to $z = 2.57$ and $z = 2.58$. Therefore, a simple linear interpolation between the two values gives us $z = 2.575$. This is correct to within the accuracy of the linear interpolation. The answer, therefore, is $z = \pm 2.575$. This is shown in Figure 4–9.

**FIGURE 4–8** **Finding $z$ Such That $P(Z \leq z) = 0.9$**

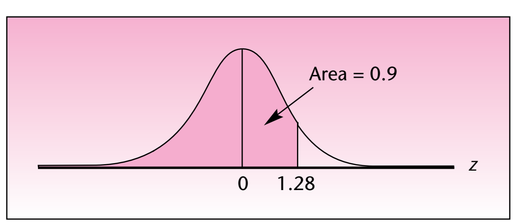
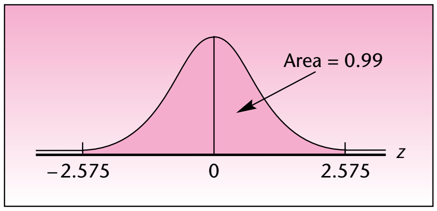


**FIGURE 4–9** A Symmetric 0.99 Probability Interval about 0 for a Standard Normal Random Variable




## PROBLEMS

**4–1.** Find the following probabilities: $P(-1 < Z < 1)$, $P(-1.96 < Z < 1.96)$, $P(-2.33 < Z < 2.33)$.

**4–2.** What is the probability that a standard normal random variable will be between the values $-2$ and 1?

**4–3.** Find the probability that a standard normal random variable will have a value between $-0.89$ and $-2.50$.

**4–4.** Find the probability that a standard normal random variable will have a value greater than 3.02.

**4–5.** Find the probability that a standard normal random variable will be between 2 and 3.

**4–6.** Find the probability that a standard normal random variable will have a value less than or equal to $-2.5$.

**4–7.** Find the probability that a standard normal random variable will be greater in value than $-2.33$.

**4–8.** Find the probability that a standard normal random variable will have a value between $-2$ and 300.

**4–9.** Find the probability that a standard normal variable will have a value less than $-10$.

**4–10.** Find the probability that a standard normal random variable will be between $-0.01$ and 0.05.

**4–11.** A sensitive measuring device is calibrated so that errors in the measurements it provides are normally distributed with mean 0 and variance 1.00. Find the probability that a given error will be between $-2$ and 2.

**4–12.** Find two values defining tails of the normal distribution with an area of 0.05 each.

**4–13.** Is it likely that a standard normal random variable will have a value less than $-4$? Explain.

**4–14.** Find a value such that the probability that the standard normal random variable will be above it is 0.85.

**4–15.** Find a value of the standard normal random variable cutting off an area of 0.685 to its left.

**4–16.** Find a value of the standard normal random variable cutting off an area of 0.50 to its right. (Do you need the table for this probability? Explain.)

**4–17.** Find $z$ such that $P(Z > z) = 0.12$.

**4–18.** Find two values, equidistant from 0 on either side, such that the probability that a standard normal random variable will be between them is 0.40.

**4–19.** Find two values of the standard normal random variable, $z$ and $-z$, such that $P(-z < Z < z) = 0.95$.

**4–20.** Find two values of the standard normal random variable, $z$ and $-z$, such that the two corresponding tail areas of the distribution (the area to the right of $z$ and the area to the left of $-z$) add to 0.01.

**4–21.** The deviation of a magnetic needle from the magnetic pole in a certain area in northern Canada is a normally distributed random variable with mean 0 and standard deviation 1.00. What is the probability that the absolute value of the deviation from the north pole at a given moment will be more than 2.4?

## 4–4 The Transformation of Normal Random Variables

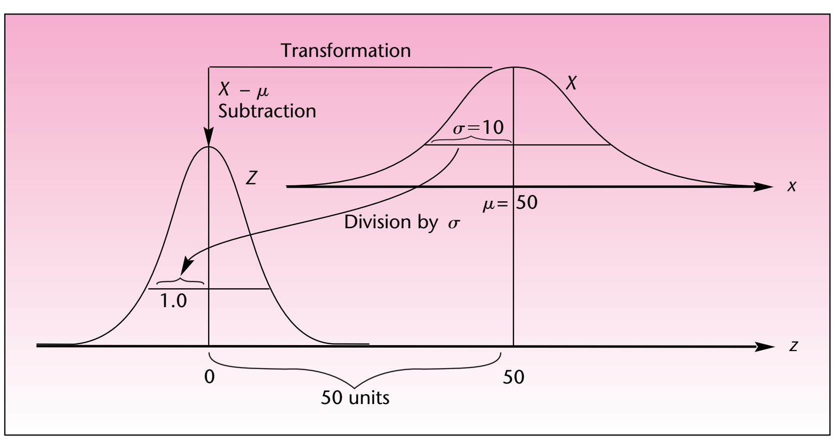
The importance of the standard normal distribution derives from the fact that any normal random variable may be transformed to the standard normal random variable. We want to transform $X$, where $X \sim N(\mu, \sigma^2)$, into the standard normal random variable $Z \sim N(0, 1^2)$. Look at Figure 4–10. Here we have a normal random variable $X$ with mean $\mu = 50$ and standard deviation $\sigma = 10$. We want to transform this random variable to a normal random variable with $\mu = 0$ and $\sigma = 1$. How can we do this?

We move the distribution from its center of 50 to a center of 0. This is done by *subtracting* 50 from all the values of $X$. Thus, we shift the distribution 50 units back so that its new center is 0. The second thing we need to do is to make the width of the distribution, its standard deviation, equal to 1. This is done by squeezing the width down from 10 to 1. Because the total probability under the curve must remain 1.00, the distribution must grow upward to maintain the same area. This is shown in Figure 4–10. Mathematically, squeezing the curve to make the width 1 is equivalent to dividing the random variable by its standard deviation. The area under the curve adjusts so that the total remains the same. *All probabilities* (areas under the curve) *adjust accordingly.* The mathematical transformation from $X$ to $Z$ is thus achieved by first subtracting $\mu$ from $X$ and then dividing the result by $\sigma$.

The transformation of $X$ to $Z$:

$$Z = \frac{X - \mu}{\sigma} \tag{4–5}$$

**FIGURE 4–10** **Transforming a Normal Random Variable with Mean 50 and Standard Deviation 10 into the Standard Normal Random Variable**



The transformation of equation 4–5 takes us from a random variable $X$ with mean $\mu$ and standard deviation $\sigma$ to the standard normal random variable. We also have an opposite, or *inverse,* transformation, which takes us from the standard normal random variable $Z$ to the random variable $X$ with mean $\mu$ and standard deviation $\sigma$. The inverse transformation is given by equation 4–6.

The inverse transformation of $Z$ to $X$:

$$X = \mu + Z\sigma \tag{4–6}$$

You can verify mathematically that equation 4–6 does the opposite of equation 4–5. Note that multiplying the random variable $Z$ by the number $\sigma$ increases the width of the curve from 1 to $\sigma$, thus making $\sigma$ the new standard deviation. Adding $\mu$ makes $\mu$ the new mean of the random variable. The actions of multiplying and then adding are the opposite of subtracting and then dividing. We note that the two transformations, one an inverse of the other, transform a *normal* random variable into a *normal* random variable. If this transformation is carried out on a random variable that is not normal, the result will not be a normal random variable.

### *Using the Normal Transformation*

Let us consider our random variable $X$ with mean 50 and standard deviation 10, $X \sim N(50, 10^2)$. Suppose we want the probability that $X$ is greater than 60. That is, we want to find $P(X > 60)$. We cannot evaluate this probability directly, but if we can transform $X$ to $Z$, we will be able to find the probability in the $Z$ table, Table 2 in Appendix C. Using equation 4–5, the required transformation is $Z = (X - \mu)/\sigma$. Let us carry out the transformation. In the probability statement $P(X > 60)$, we will substitute $Z$ for $X$. If, however, we carry out the transformation on one side of the probability inequality, we must also do it on the other side. In other words, transforming $X$ into $Z$ requires us also to transform the value 60 into the appropriate value of the standard normal distribution. We transform the value 60 into the value $(60 - \mu)/\sigma$. The new probability statement is

$$\begin{aligned} P(X > 60) &= P\left(\frac{X - \mu}{\sigma} > \frac{60 - \mu}{\sigma}\right) = P\left(Z > \frac{60 - \mu}{\sigma}\right) \\ &= P\left(Z > \frac{60 - 50}{10}\right) = P(Z > 1) \end{aligned}$$

Why does the inequality still hold? We subtracted a number from each side of an inequality; this does not change the inequality. In the next step we divide both sides of the inequality by the standard deviation $\sigma$. The inequality does not change because we can divide both sides of an inequality by a positive number, and a standard deviation is always a positive number. (Recall that dividing by 0 is not permissible; and dividing, or multiplying, by a negative value would reverse the direction of the inequality.) From the transformation, we find that the probability that a normal random variable with mean 50 and standard deviation 10 will have a value greater than 60 is exactly the probability that the standard normal random variable $Z$ will be greater than 1. The latter probability can be found using Table 2 in Appendix C. We find: $P(X > 60) = P(Z > 1) = 0.5000 - 0.3413 = 0.1587$. Let us now look at a few examples of the use of equation 4–5.

**EXAMPLE 4–5**

Suppose that the time it takes the electronic device in the car to respond to the signal from the toll plaza is normally distributed with mean 160 microseconds and standard deviation 30 microseconds. What is the probability that the device in the car will respond to a given signal within 100 to 180 microseconds?

***Solution***

Figure 4–11 shows the normal distribution for $X \sim N(160, 30^2)$ and the required area on the scale of the original problem and on the transformed $z$ scale. We have the following (where the probability statement inequality has three sides and we carry out the transformation of equation 4–5 on all three sides):

$$P(100 < X < 180) = P\left(\frac{100 - \mu}{\sigma} < \frac{X - \mu}{\sigma} < \frac{180 - \mu}{\sigma}\right)$$
$$= P\left(\frac{100 - 160}{30} < Z < \frac{180 - 160}{30}\right)$$
$$= P(-2 < Z < 0.6666) = 0.4772 + 0.2475 = 0.7247$$

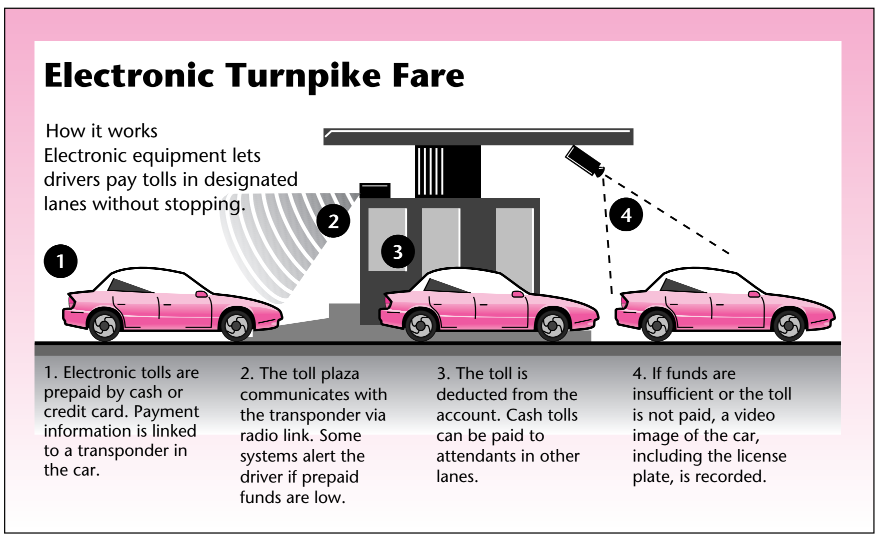
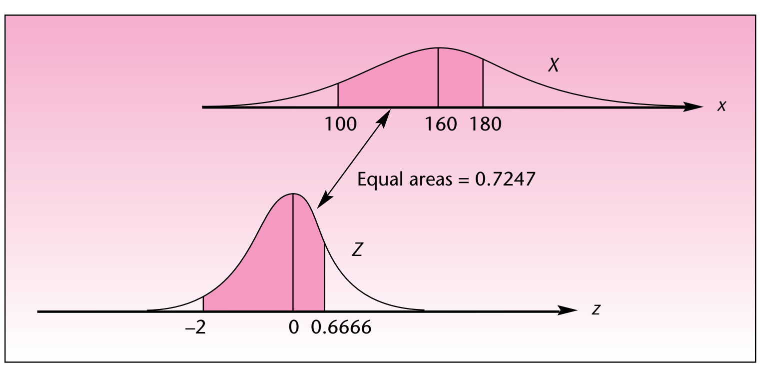





**FIGURE 4–11 Probability Computation for Example 4–5**



(Table area values were obtained by linear interpolation.) Thus, the chance that the device will respond within 100 to 180 microseconds is 0.7247.

**EXAMPLE 4–6**

The concentration of impurities in a semiconductor used in the production of microprocessors for computers is a normally distributed random variable with mean 127 parts per million and standard deviation 22. A semiconductor is acceptable only if its concentration of impurities is below 150 parts per million. What proportion of the semiconductors are acceptable for use?

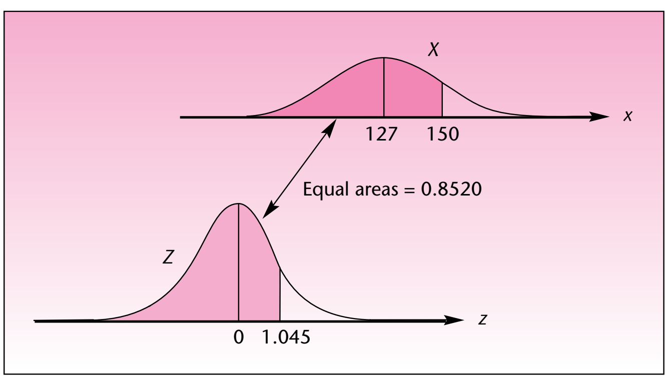
***Solution***

Now $X \sim N(127, 22^2)$, and we need $P(X < 150)$. Using equation 4–5, we have

$$P(X < 150) = P\left(\frac{X - \mu}{\sigma} < \frac{150 - \mu}{\sigma}\right) = P\left(Z < \frac{150 - 127}{22}\right)$$
$$= P(Z < 1.045) = 0.5 + 0.3520 = 0.8520$$

(The TA of 0.3520 was obtained by interpolation.) Thus, 85.2% of the semiconductors are acceptable for use. This also means that the probability that a randomly chosen semiconductor will be acceptable for use is 0.8520. The solution of this example is illustrated in Figure 4–12.

**FIGURE 4–12 Probability Computation for Example 4–6**



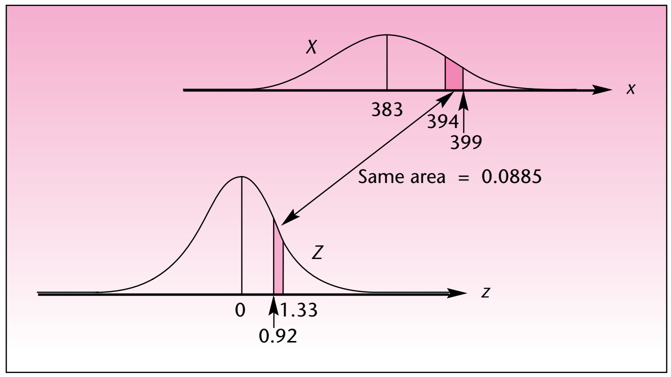

**EXAMPLE 4–7**

Fluctuations in the prices of precious metals such as gold have been empirically shown to be well approximated by a normal distribution when observed over short intervals of time. In May 1995, the daily price of gold (1 troy ounce) was believed to have a mean of \$383 and a standard deviation of \$12. A broker, working under these assumptions, wanted to find the probability that the price of gold the next day would be between \$394 and \$399 per troy ounce. In this eventuality, the broker had an order from a client to sell the gold in the client's portfolio. What is the probability that the client's gold will be sold the next day?

***Solution***

Figure 4–13 shows the setup for this problem and the transformation of $X$, where $X \sim N(383, 12^2)$, into the standard normal random variable $Z$. Also shown are the required areas under the $X$ curve and the transformed $Z$ curve. We have

**FIGURE 4–13** Probability Computation for Example 4–7




$$P(394 < X < 399) = P\left(\frac{394 - \mu}{\sigma} < \frac{X - \mu}{\sigma} < \frac{399 - \mu}{\sigma}\right)$$
$$= P\left(\frac{394 - 383}{12} < Z < \frac{399 - 383}{12}\right)$$
$$= P(0.9166 < Z < 1.3333) = 0.4088 - 0.3203 = 0.0885$$

(Both TA values were obtained by linear interpolation, although this is not necessary if less accuracy is acceptable.)

Let us summarize the transformation procedure used in computing probabilities of events associated with a normal random variable $X \sim N(\mu, \sigma^2)$.

Transformation formulas of $X$ to $Z$, where $a$ and $b$ are numbers:

$$P(X < a) = P\left(Z < \frac{a - \mu}{\sigma}\right)$$

$$P(X > b) = P\left(Z > \frac{b - \mu}{\sigma}\right)$$

$$P(a < X < b) = P\left(\frac{a - \mu}{\sigma} < Z < \frac{b - \mu}{\sigma}\right)$$

## PROBLEMS

**4–22.** For a normal random variable with mean 650 and standard deviation 40, find the probability that its value will be below 600.

**4–23.** Let $X$ be a normally distributed random variable with mean 410 and standard deviation 2. Find the probability that $X$ will be between 407 and 415.

**4–24.** If $X$ is normally distributed with mean 500 and standard deviation 20, find the probability that $X$ will be above 555.

**4–25.** For a normally distributed random variable with mean −44 and standard deviation 16, find the probability that the value of the random variable will be above 0.

**4–26.** A normal random variable has mean 0 and standard deviation 4. Find the probability that the random variable will be above 2.5.

**4–27.** Let $X$ be a normally distributed random variable with mean $\mu = 16$ and standard deviation $\sigma = 3$. Find $P(11 < X < 20)$. Also find $P(17 < X < 19)$ and $P(X > 15)$.

**4–28.** The time it takes an international telephone operator to place an overseas phone call is normally distributed with mean 45 seconds and standard deviation 10 seconds.

*a.* What is the probability that my call will go through in less than 1 minute?

*b.* What is the probability that I will get through in less than 40 seconds?

*c.* What is the probability that I will have to wait more than 70 seconds for my call to go through?

**4–29.** The number of votes cast in favor of a controversial proposition is believed to be approximately normally distributed with mean 8,000 and standard deviation 1,000. The proposition needs at least 9,322 votes in order to pass. What is the probability that the proposition will pass? (Assume numbers are on a continuous scale.)

**4–30.** Under the system of floating exchange rates, the rate of foreign money to the U.S. dollar is affected by many random factors, and this leads to the assumption of a normal distribution of small daily fluctuations. The rate of U.S. dollar per euro is believed in April 2007 to have a mean of 1.36 and a standard deviation of 0.03.[1] Find the following.

*a.* The probability that tomorrow's rate will be above 1.42.

*b.* The probability that tomorrow's rate will be below 1.35.

*c.* The probability that tomorrow's exchange rate will be between 1.16 and 1.23.

**4–31.** *Wine Spectator* rates wines on a point scale of 0 to 100. It can be inferred from the many ratings in this magazine that the average rating is 87 and the standard deviation is 3 points. Wine ratings seem to follow a normal distribution. In the May 15, 2007, issue of the magazine, the burgudy Domaine des Perdrix received a rating of 89.[2] What is the probability that a randomly chosen wine will score this high or higher?

**4–32.** The weights of domestic, adult cats are normally distributed with a mean of 10.42 pounds and a standard deviation of 0.87 pounds. A cat food manufacturer sells three types of foods for underweight, normal, and overweight cats. The manufacturer considers the bottom 5% of the cats underweight and the top 10% overweight. Compute what weight range must be specified for each of the three categories.

**4–33.** Daily fluctuations of the French CAC-40 stock index from March to June 1997 seem to follow a normal distribution with mean of 2,600 and standard deviation of 50. Find the probability that the CAC-40 will be between 2,520 and 2,670 on a random day in the period of study.

**4–34.** According to global analyst Olivier Lemaigre, the average price-to-earnings ratio for companies in emerging markets is 12.5.[3] Assume a normal distribution and a standard deviation of 2.5. If a company in emerging markets is randomly selected, what is the probability that its price-per-earnings ratio is above 17.5, which, according to Lemaigre, is the average for companies in the developed world?

**4–35.** Based on the research of Ibbotson Associates, a Chicago investment firm, and Prof. Jeremy Siegel of the Wharton School of the University of Pennsylvania, the

---

[1] This information is inferred from data on foreign exchange rates in *The New York Times,* April 20, 2007, p. C10.

[2] "The Ratings," *Wine Spectator,* May 15, 2007, p. 156.

[3] Mitchell Martin, "Stock Focus: Ride the Rocket," *Forbes,* April 26, 2004, p. 138.

average return on large-company stocks since 1920 has been 10.5% per year and the standard deviation has been 4.75%. Assuming a normal distribution for stock returns (and that the trend will continue this year), what is the probability that a large-company stock you've just bought will make in 1 year at least 12%? Will lose money? Will make at least 5%?

**4–36.** A manufacturing company regularly consumes a special type of glue purchased from a foreign supplier. Because the supplier is foreign, the time gap between placing an order and receiving the shipment against that order is long and uncertain. This time gap is called "lead time." From past experience, the materials manager notes that the company's demand for glue during the uncertain lead time is normally distributed with a mean of 187.6 gallons and a standard deviation of 12.4 gallons. The company follows a policy of placing an order when the glue stock falls to a predetermined value called the "reorder point." Note that if the reorder point is $x$ gallons and the demand during lead time exceeds $x$ gallons, the glue would go "stock-out" and the production process would have to stop. Stock-out conditions are therefore serious.

*a.* If the reorder point is kept at 187.6 gallons (equal to the mean demand during lead time) what is the probability that a stock-out condition would occur?

*b.* If the reorder point is kept at 200 gallons, what is the probability that a stock-out condition would occur?

*c.* If the company wants to be 95% confident that the stock-out condition will not occur, what should be the reorder point? The reorder point minus the mean demand during lead time is known as the "safety stock." What is the safety stock in this case?

*d.* If the company wants to be 99% confident that the stock-out condition will not occur, what should be the reorder point? What is the safety stock in this case?

**4–37.** The daily price of orange juice 30-day futures is normally distributed. In March through April 2007, the mean was 145.5 cents per pound, and standard deviation = 25.0 cents per pound.[4] Assuming the price is independent from day to day, find $P(x < 100)$ on the next day.

## 4–5 The Inverse Transformation

Let us look more closely at the relationship between $X$, a normal random variable with mean $\mu$ and standard deviation $\sigma$, and the standard normal random variable. The fact that the standard normal random variable has mean 0 and standard deviation 1 has some important implications. When we say that $Z$ is greater than 2, we are also saying that $Z$ is more than 2 *standard deviations above its mean.* This is so because the mean of $Z$ is 0 and the standard deviation is 1; hence, $Z > 2$ is the same event as $Z > [0 + 2(1)]$.

Now consider a normal random variable $X$ with mean 50 and standard deviation 10. Saying that $X$ is greater than 70 is exactly the same as saying that $X$ is 2 standard deviations above its mean. This is so because 70 is 20 units above the mean of 50, and 20 units = 2(10) units, or 2 standard deviations of $X$. Thus, the event $X > 70$ is the same as the event $X >$ (2 standard deviations above the mean). This event is identical to the event $Z > 2$. Indeed, this is what results when we carry out the transformation of equation 4–5:

$$P(X > 70) = P\left(\frac{X - \mu}{\sigma} > \frac{70 - \mu}{\sigma}\right) = P\left(Z > \frac{70 - 50}{10}\right) = P(Z > 2)$$

[4] "Futures," *The New York Times,* April 26, 2007, p. C9.

Normal random variables are related to one another by the fact that the probability that a normal random variable will be above (or below) its mean a certain number of standard deviations is exactly equal to the probability that any other normal random variable will be above (or below) its mean the same number of (its) standard deviations. In particular, this property holds for the standard normal random variable. The probability that a normal random variable will be greater than (or less than) $z$ standard-deviation units above its mean is the same as the probability that the standard normal random variable will be greater than (less than) $z$. The change from a *$z$ value* of the random variable $Z$ to *$z$ standard deviations* above the mean for a given normal random variable $X$ should suggest to us the inverse transformation, equation 4–6:

$$x = \mu + z\sigma$$

That is, the value of the random variable $X$ may be written in terms of the number $z$ of standard deviations $\sigma$ it is above or below the mean $\mu$. Three examples are useful here. We know from the standard normal probability table that the probability that $Z$ is greater than $-1$ and less than 1 is 0.6826 (show this). Similarly, we know that the probability that $Z$ is greater than $-2$ and less than 2 is 0.9544. Also, the probability that $Z$ is greater than $-3$ and less than 3 is 0.9974. These probabilities may be applied to *any* normal random variable as follows:[5]

1. The probability that a normal random variable will be within a distance of *1 standard deviation* from its mean (on either side) is 0.6826, or *approximately 0.68.*
2. The probability that a normal random variable will be within *2 standard deviations* of its mean is 0.9544, or *approximately 0.95.*
3. The probability that a normal random variable will be within 3 *standard deviations* of its mean is 0.9974.

We use the inverse transformation, equation 4–6, when we want to get from a given probability to the value or values of a normal random variable $X$. We illustrate the procedure with a few examples.

---

**EXAMPLE 4–8**

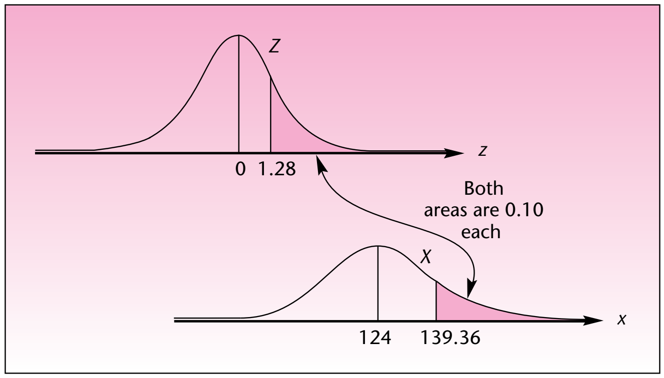
PALCO Industries, Inc., is a leading manufacturer of cutting and welding products. One of the company's products is an acetylene gas cylinder used in welding. The amount of nitrogen gas in a cylinder is a normally distributed random variable with mean 124 units of volume and standard deviation 12. We want to find the amount of nitrogen $x$ such that 10% of the cylinders contain more nitrogen than this amount.

***Solution***

We have $X \sim N(124, 12^2)$. We are looking for the value of the random variable $X$ such that $P(X > x) = 0.10$. In order to find it, we look for the value of the standard normal random variable $Z$ such that $P(Z > z) = 0.10$. Figure 4–14 illustrates how we find the value $z$ and transform it to $x$. If the area to the right of $z$ is equal to 0.10, the area between 0 and $z$ (the table area) is equal to $0.5 - 0.10 = 0.40$. We look inside the table for the $z$ value corresponding to TA = 0.40 and find $z = 1.28$ (actually, TA = 0.3997,

[5]This is the origin of the *empirical rule* (in Chapter 1) for mound-shaped data distributions. Mound-shaped data sets approximate the distribution of a normal random variable, and hence the proportions of observations within a given number of standard deviations away from the mean roughly equal those predicted by the normal distribution. Compare the empirical rule (section 1–7) with the numbers given here.

FIGURE 4–14 Solution of Example 4–8




which is close enough to 0.4). We need to find the appropriate $x$ value. Here we use equation 4–6:

$$x = \mu + z\sigma = 124 + (1.28)(12) = 139.36$$

Thus, 10% of the acetylene cylinders contain more than 139.36 units of nitrogen.

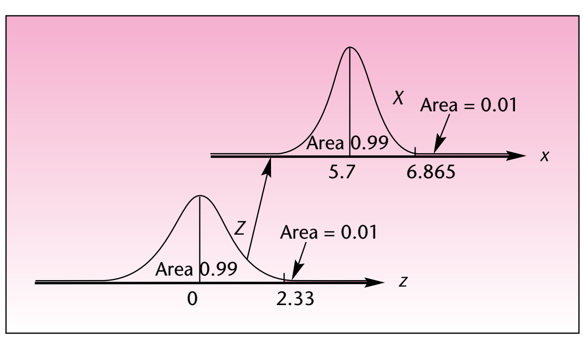
**EXAMPLE 4–9**

The amount of fuel consumed by the engines of a jetliner on a flight between two cities is a normally distributed random variable $X$ with mean $\mu = 5.7$ tons and standard deviation $\sigma = 0.5$. Carrying too much fuel is inefficient as it slows the plane. If, however, too little fuel is loaded on the plane, an emergency landing may be necessary. The airline would like to determine the amount of fuel to load so that there will be a 0.99 probability that the plane will arrive at its destination.

*Solution*

We have $X \sim N(5.7, 0.5^2)$. First, we must find the value $z$ such that $P(Z < z) = 0.99$. Following our methodology, we find that the required table area is TA $= 0.99 - 0.5 = 0.49$, and the corresponding $z$ value is 2.33. Transforming the $z$ value to an $x$ value, we get $x = \mu + z\sigma = 5.7 + (2.33)(0.5) = 6.865$. Thus, the plane should be loaded with 6.865 tons of fuel to give a 0.99 probability that the fuel will last throughout the flight. The transformation is shown in Figure 4–15.

FIGURE 4–15 Solution of Example 4–9



**EXAMPLE 4–10**

Weekly sales of Campbell's soup cans at a grocery store are believed to be approximately normally distributed with mean 2,450 and standard deviation 400. The store management wants to find two values, symmetrically on either side of the mean, such that there will be a 0.95 probability that sales of soup cans during the week will be between the two values. Such information is useful in determining levels of orders and stock.

*Solution*

Here $X \sim N(2{,}450, 400^2)$. From the section on the standard normal random variable, we know how to find two values of $Z$ such that the area under the curve between them is 0.95 (or any other area). We find that $z = 1.96$ and $z = -1.96$ are the required values. We now need to use equation 4–6. Since there are *two* values, one the negative of the other, we may combine them in a single transformation:

$$x = \mu \pm z\sigma \qquad (4\text{–}7)$$

Applying this special formula we get $x = 2{,}450 \pm (1.96)(400) = 1{,}666$ and $3{,}234$. Thus, management may be 95% sure that sales on any given week will be between 1,666 and 3,234 units.

The procedure of obtaining values of a normal random variable, given a probability, is summarized:

1. Draw a picture of the normal distribution in question and the standard normal distribution.
2. In the picture, shade in the area corresponding to the probability.
3. Use the table to find the $z$ value (or values) that gives the required probability.
4. Use the transformation from $Z$ to $X$ to get the appropriate value (or values) of the original normal random variable.

## PROBLEMS

**4–38.** If $X$ is a normally distributed random variable with mean 120 and standard deviation 44, find a value $x$ such that the probability that $X$ will be less than $x$ is 0.56.

**4–39.** For a normal random variable with mean 16.5 and standard deviation 0.8, find a point of the distribution such that there is a 0.85 probability that the value of the random variable will be above it.

**4–40.** For a normal random variable with mean 19,500 and standard deviation 400, find a point of the distribution such that the probability that the random variable will exceed this value is 0.02.

**4–41.** Find two values of the normal random variable with mean 88 and standard deviation 5 lying symmetrically on either side of the mean and covering an area of 0.98 between them.

**4–42.** For $X \sim N(32, 7^2)$, find two values $x_1$ and $x_2$, symmetrically lying on each side of the mean, with $P(x_1 < X < x_2) = 0.99$.

**4–43.** If $X$ is a normally distributed random variable with mean $-61$ and standard deviation 22, find the value such that the probability that the random variable will be above it is 0.25.

**4–44.** If $X$ is a normally distributed random variable with mean 97 and standard deviation 10, find $x_2$ such that $P(102 < X < x_2) = 0.05$.

**4–45.** Let $X$ be a normally distributed random variable with mean 600 and variance 10,000. Find two values $x_1$ and $x_2$ such that $P(X > x_1) = 0.01$ and $P(X < x_2) = 0.05$.

**4–46.** Pierre operates a currency exchange office at Orly Airport in Paris. His office is open at night when the airport bank is closed, and he makes most of his business on returning U.S. tourists who need to change their remaining euros back to U.S. dollars. From experience, Pierre knows that the demand for dollars on any given night during high season is approximately normally distributed with mean \$25,000 and standard deviation \$5,000. If Pierre carries too much cash in dollars overnight, he pays a penalty: interest on the cash. On the other hand, if he runs short of cash during the night, he needs to send a person downtown to an all-night financial agency to get the required cash. This, too, is costly to him. Therefore, Pierre would like to carry overnight an amount of money such that the demand on 85% of the nights will not exceed this amount. Can you help Pierre find the required amount of dollars to carry?

**4–47.** The demand for high-grade gasoline at a service station is normally distributed with mean 27,009 gallons per day and standard deviation 4,530. Find two values that will give a symmetric 0.95 probability interval for the amount of high-grade gasoline demanded daily.

**4–48.** The percentage of protein in a certain brand of dog food is a normally distributed random variable with mean 11.2% and standard deviation 0.6%. The manufacturer would like to state on the package that the product has a protein content of at least $x_1$% and no more than $x_2$%. It wants the statement to be true for 99% of the packages sold. Determine the values $x_1$ and $x_2$.

**4–49.** Private consumption as a share of GDP is a random quantity that follows a roughly normal distribution. According to an article in *BusinessWeek,* for the United States that was about 71%.[6] Assuming that this value is the mean of a normal distribution, and that the standard deviation of the distribution is 3%, what is the value of private consumption as share of GDP such that you are 90% sure that the actual value falls below it?

**4–50.** The daily price of coffee is approximately normally distributed over a period of 15 days with a mean in April 2007 of \$1.35 per pound (on the wholesale market) and standard deviation of \$0.15. Find a price such that the probability in the next 15 days that the price will go below it will be 0.90.

**4–51.** The daily price in dollars per metric ton of cocoa in 2007 was normally distributed with $\mu = \$2{,}014$ per metric ton and $\sigma = \$2.00$. Find a price such that the probability that the actual price will be above it is 0.80.

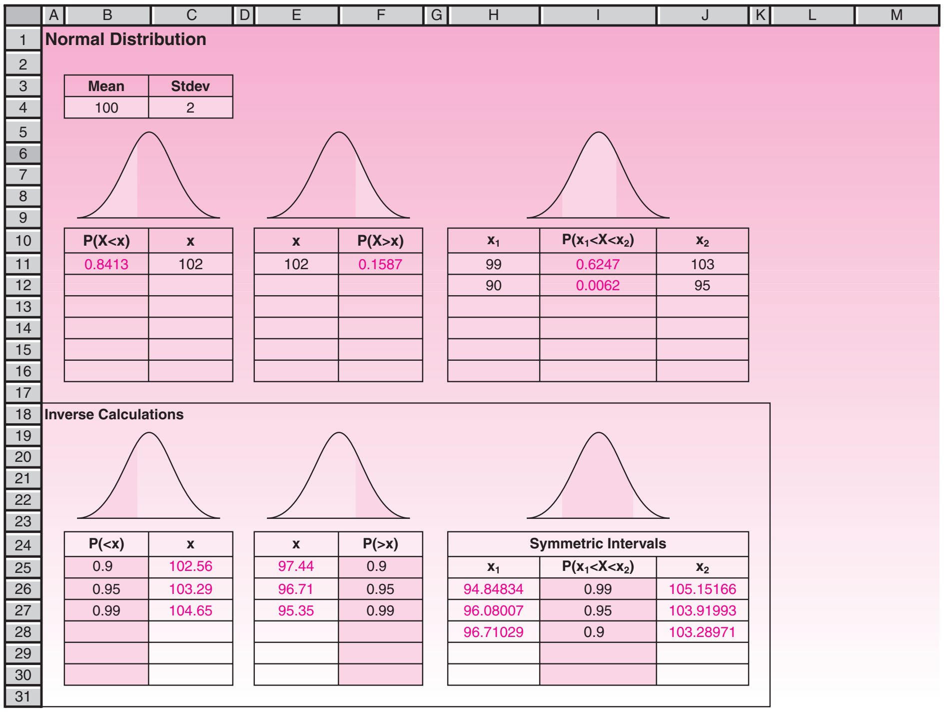
## 4–6 The Template

This normal distribution template is shown in Figure 4–16. As usual, it can be used in conjunction with the Goal Seek command and the Solver tool to solve many types of problems.

To use the template, make sure that the correct values are entered for the mean and the standard deviation in cells B4 and C4. Cell B11 gives the area to the left of the value entered in cell C11. The five cells below C11 can be similarly used. Cell F11 gives the area to the right of the value entered in cell E11. Cell I11 contains the area between the values entered in cells H11 and J11. In the area marked "Inverse Calculations," you can input areas (probabilities) and get $x$ values corresponding to those areas. For example, on entering 0.9 in cell B25, we get the $x$ value of 102.56 in cell C25. This implies that the area to the left of 102.56 is 0.9. Similarly, cell F25 has been used to get the $x$ value that has 0.9 area to its right.

[6]Dexter Roberts, "Slower to Spend," *BusinessWeek,* April 30, 2007, p. 34.

**FIGURE 4–16** Normal Distribution Template
[Normal Distribution.xls; Sheet: Normal]




Sometimes we are interested in getting the *narrowest* interval that contains a desired amount of area. A little thought reveals that the narrowest interval has to be symmetric about the mean, because the distribution is symmetric and it peaks at the mean. In later chapters, we will study *confidence intervals,* many of which are also the narrowest intervals that contain a desired amount of area. Naturally, these confidence intervals are symmetric about the mean. For this reason, we have the "Symmetric Intervals" area in the template. Once the desired area is entered in cell I26, the limits of the symmetric interval that contains that much area appear in cells H26 and J26. In the example shown in the Figure 4–16, the symmetric interval (94.85, 105.15) contains the desired area of 0.99.

### *Problem Solving with the Template*

Most questions about normal random variables can be answered using the template in Figure 4–16. We will see a few problem-solving strategies through examples.

**EXAMPLE 4–11**

Suppose $X \sim N(100, 2^2)$. Find $x_2$ such that $P(99 \leq X \leq x_2) = 60\%$.

***Solution***

Fill in cell B4 with the mean 100 and cell C4 with standard deviation 2. Fill in cell H11 with 99. Then on the Data tab, in the Data Tools group, click What If Analysis, and then click Goal Seek. In the dialog box, ask to set cell I11 to value 0.6 by changing cell J11. Click OK when the computer finds the answer. The required value of 102.66 for $x_2$ appears in cell J11.

**EXAMPLE 4–12** Suppose $X \sim N(\mu, 0.5^2)$; $P(X > 16.5) = 0.20$. What is $\mu$?

**Solution** Enter the $\sigma$ of 0.5 in cell C4. Since we do not know $\mu$, enter a *guessed* value of 15 in cell B4. Then enter 16.5 in cell F11. Now invoke the Goal Seek command to set cell F11 to value 0.20 by changing cell B4. The computer finds the value of $\mu$ in cell B4 to be 16.08.

The Goal Seek command can be used if there is only one unknown. With more than one unknown, the Solver tool has to be used. We shall illustrate the use of the Solver in the next example.

**EXAMPLE 4–13** Suppose $X \sim N(\mu, \sigma^2)$; $P(X > 28) = 0.80$; $P(X > 32) = 0.40$. What are $\mu$ and $\sigma$?

**Solution** *One way* to solve this problem is to use the Solver to find $\mu$ and $\sigma$ with the objective of making $P(X > 28) = 0.80$ subject to the constraint $P(X > 32) = 0.40$. The following detailed steps will do just that:

- Fill in cell B4 with 30 (which is a guessed value for $\mu$).
- Fill in cell C4 with 2 (which is a guessed value for $\sigma$).
- Fill in cell E11 with 28.
- Fill in cell E12 with 32.
- Under the Analysis group on the Data tab select the Solver.
- In the Set Cell box enter `F11`.
- In the To Value box enter `0.80` [which sets up the objective of $P(X > 28) = 0.80$].
- In the By Changing Cells box enter `B4:C4`.
- Click on the Constraints box and the Add button.
- In the dialog box on the left-hand side enter `F12`.
- Select the = sign in the middle drop down box.
- Enter `0.40` in the right-hand-side box [which sets up the constraint of $P(X > 32) = 0.40$].
- Click the OK button.
- In the Solver dialog box that reappears, click the Solve button.
- In the dialog box that appears at the end, select the Keep Solver Solution option.

The Solver finds the correct values for the cells B4 and C4 as $\mu = 31.08$ and $\sigma = 3.67$.

**EXAMPLE 4–14** A customer who has ordered 1-inch-diameter pins in bulk will buy only those pins with diameters in the interval $1 \pm 0.003$ inches. An automatic machine produces pins whose diameters are normally distributed with mean 1.002 inches and standard deviation 0.0011 inch.

1. What percentage of the pins made by the machine will be acceptable to the customer?
2. If the machine is adjusted so that the mean of the pins made by the machine is reset to 1.000 inch, what percentage of the pins will be acceptable to the customer?
3. Looking at the answer to parts 1 and 2, can we say that the machine must be reset?

**Solution**

1. Enter $\mu = 1.002$ and $\sigma = 0.0011$ into the template. From the template $P(0.997 < X < 1.003) = 0.8183$. Thus, 81.83% of the pins will be acceptable to the consumer.
2. Change $\mu$ to 1.000 in the template. Now, $P(0.997 < X < 1.003) = 0.9936$. Thus, 99.36% of the pins will be acceptable to the consumer.
3. Resetting the machine has considerably increased the percentage of pins acceptable to the consumer. Therefore, resetting the machine is highly desirable.

## 4–7 Normal Approximation of Binomial Distributions

When the number of trials $n$ in a binomial distribution is large (>1,000), the calculation of probabilities becomes difficult for the computer, because the calculation encounters some numbers that are too large and some that are too small to handle with needed accuracy. Fortunately, the binomial distribution approaches the normal distribution as $n$ increases and therefore we can approximate it as a normal distribution. Note that the mean is $np$ and the standard deviation is $\sqrt{np(1-p)}$. The template is shown in Figure 4–17. When the values for $n$ and $p$ of the binomial distribution are entered in cells B4 and C4, the mean and the standard deviation of the corresponding normal distribution are calculated in cells E4 and F4. The rest of the template is similar to the normal distribution template we already saw.

Whenever a binomial distribution is approximated as a normal distribution, a **continuity correction** is required because a binomial is discrete and a normal is continuous. Thus, a column in the histogram of a binomial distribution for, say, $X = 10$, covers, in the continuous sense, the interval [9.5, 10.5]. Similarly, if we include the columns for $X = 10$, 11, and 12, then in the continuous case, the bars occupy the interval [9.5, 12.5], as seen in Figure 4–18. Therefore, when we calculate

**FIGURE 4–17** **The Template for Normal Approximation of Binomial Distribution [Normal Distribution.xls; Sheet: Normal Approximation]**

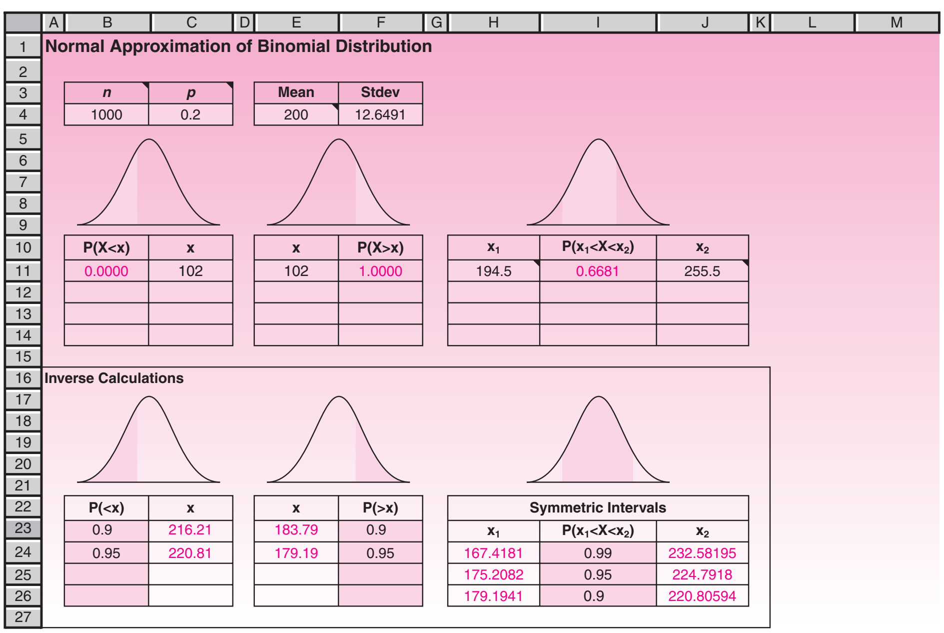
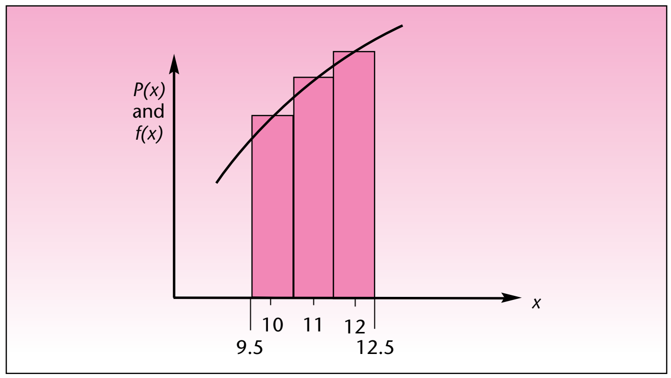


FIGURE 4–18 Continuity Correction




the binomial probability of an interval, say, $P(195 \leq X \leq 255)$, *we should subtract 0.5 from the left limit and add 0.5 to the right limit* to get the corresponding normal probability, namely, $P(194.5 < X < 255.5)$. Adding and subtracting 0.5 in this manner is known as the *continuity correction*. In Figure 4–17, this correction has been applied in cells H11 and J11. Cell I11 has the binomial probability of $P(195 \leq X \leq 255)$.

**EXAMPLE 4–15**

A total of 2,058 students take a difficult test. Each student has an independent 0.6205 probability of passing the test.

*a.* What is the probability that between 1,250 and 1,300 students, both numbers inclusive, will pass the test?
*b.* What is the probability that at least 1,300 students will pass the test?
*c.* If the probability of at least 1,300 students passing the test has to be at least 0.5, what is the minimum value for the probability of each student passing the test?

***Solution***

*a.* On the template for normal approximation, enter 2,058 for $n$ and 0.6205 for $p$. Enter 1,249.5 in cell H11 and 1,300.5 in cell J11. The answer 0.7514 appears in cell I11.
*b.* Enter 1,299.5 in cell E11. The answer 0.1533 appears in cell F11.
*c.* Use the Goal Seek command to set cell F11 to value 0.5 by changing cell C4. The computer finds the answer as $p = 0.6314$.

## PROBLEMS

In the following problems, use a normal distribution to compute the required probabilities. In each problem, also state the assumptions necessary for a binomial distribution, and indicate whether the assumptions are reasonable.

**4–52.** The manager of a restaurant knows from experience that 70% of the people who make reservations for the evening show up for dinner. The manager decides one evening to overbook and accept 20 reservations when only 15 tables are available. What is the probability that more than 15 parties will show up?

**4–53.** An advertising research study indicates that 40% of the viewers exposed to an advertisement try the product during the following four months. If 100 people are exposed to the ad, what is the probability that at least 20 of them will try the product in the following four months?

**4–54.** According to *The Economist,* 77.9% of Google stockholders have voting power.[7] If 2,000 stockholders are gathered in a meeting, what is the probability that at least 1,500 of them can vote?

**4–55.** Sixty percent of the managers who enroll in a special training program will successfully complete the program. If a large company sends 328 of its managers to enroll in the program, what is the probability that at least 200 of them will pass?

**4–56.** A large state university sends recruiters throughout the state to recruit graduating high school seniors to enroll in the university. University records show that 25% of the students who are interviewed by the recruiters actually enroll. If last spring the university recruiters interviewed 1,889 graduating seniors, what is the probability that at least 500 of them will enroll this fall?

**4–57.** According to *Fortune,* Missouri is within 500 miles of 44% of all U.S. manufacturing plants.[8] If a Missouri company needs parts manufactured in 122 different plants, what is the probability that at least half of them can be found within 500 miles of the state? (Assume independence of parts and of plants.)

**4–58.** According to *Money,* 59% of full-time workers believe that technology has lengthened their workday.[9] If 200 workers are randomly chosen, what is the probability that at least 120 of them believe that technology has lengthened their workday?

## 4–8 Using the Computer

### *Using Excel Functions for a Normal Distribution*

In addition to the templates discussed in this chapter, you can use the built-in functions of Excel to evaluate probabilities for normal random variables.

The **NORMDIST** function returns the normal distribution for the specified mean and standard deviation. In the formula `NORMDIST(x, mean, stdev, cumulative)`, *x* is the value for which you want the distribution, *mean* is the arithmetic mean of the distribution, *stdev* is the standard deviation of the distribution, and *cumulative* is a logical value that determines the form of the function. If cumulative is TRUE, NORMDIST returns the cumulative distribution function; if FALSE, it returns the probability density function. For example, `NORMDIST(102,100,2,TRUE)` will return the area to the left of 102 in a normal distribution with mean 100 and standard deviation 2. This value is 0.8413. `NORMDIST(102,100,2,FALSE)` will return the density function $f(x)$, which is not needed for most practical purposes.

`NORMSDIST(z)` returns the standard normal cumulative distribution function, which means the area to the left of $z$ in a standard normal distribution. You can use this function in place of a table of standard normal curve areas. For example `NORMSDIST(1)` will return the value 0.8413.

`NORMINV(probability, mean, stdev)` returns the inverse of the normal cumulative distribution for the specified mean and standard deviation. For example `NORMINV(0.8413, 100, 2)` will return the value of $x$ on the normal distribution with mean 100 and standard deviation 2 for which $P(X \leq x) = 0.8413$. The value of $x$ is 102.

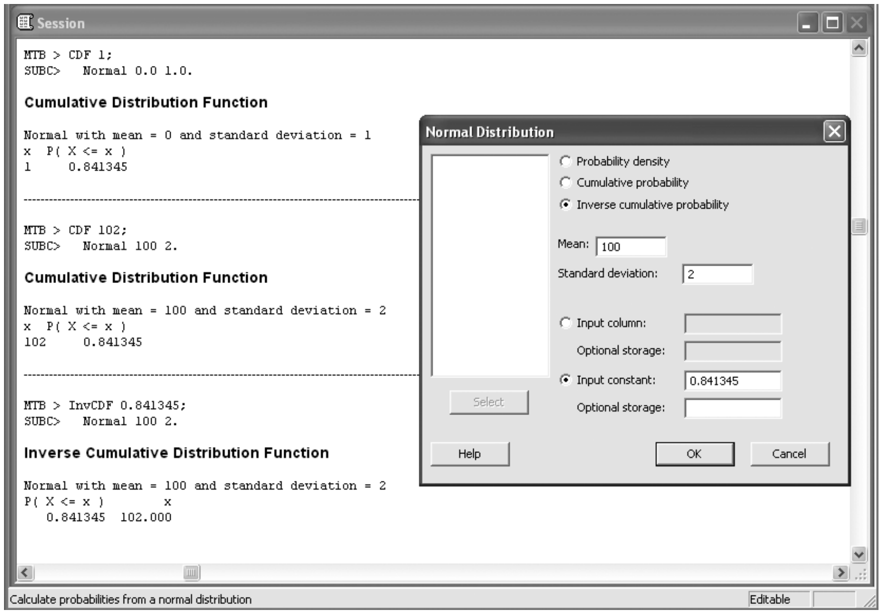
The function `NORMSINV(Probability)` returns the inverse of the standard normal cumulative distribution. For example, the formula `NORMSINV(0.8413)` will return the value 1, for which $P(Z \leq 1) = 0.8413$.

[7] "Our Company Right or Wrong," *The Economist,* March 17, 2007, p. 77.

[8] "Missouri," *Fortune,* March 19, 2007, p. 177.

[9] Jean Chatzky, "Confessions of an E-Mail Addict," *Money,* March 28, 2007, p. 28.

**FIGURE 4–19** Using MINITAB for Generating Cumulative and Inverse Cumulative Distribution Functions of a Normal Distribution




### *Using MINITAB for a Normal Distribution*

As in the previous chapter, choose Calc ▶ Probability Distributions ▶ Normal from the menu. The Normal Distribution dialog box will appear. Using the items available in the dialog box, you can choose to calculate probabilities, cumulative probabilities, or inverse cumulative probabilities for a normal distribution. You also need to specify the mean and standard deviation of the normal distribution. In the input section the values for which you aim to obtain probability densities, cumulative probabilities, or inverse cumulative probabilities are specified. These values can be a constant or a set of values that have been defined in a column. Then press OK to observe the obtained result in the Session window. Figure 4–19 shows the Session commands for obtaining the cumulative distribution in a standard normal distribution as well as a normal distribution with mean 100 and standard deviation 2. It also shows the dialog box and Session commands for obtaining inverse cumulative probabilities for a normal distribution with mean 100 and standard deviation 2.

## 4–9 Summary and Review of Terms

In this chapter, we discussed the **normal probability distribution,** the most important probability distribution in statistics. We defined the **standard normal random variable** as the normal random variable with mean 0 and standard deviation 1. We saw how to use a table of probabilities for the standard normal random variable and how to transform a normal random variable with any mean and any standard deviation to the standard normal random variable by using the **normal transformation.**

We also saw how the standard normal random variable may, in turn, be transformed into any other normal random variable with a specified mean and standard deviation, and how this allows us to find values of a normal random variable that conform with some probability statement. We discussed a method of determining the mean and/or the standard deviation of a normal random variable from probability statements about the random variable. We saw how the normal distribution is used as a model in many real-world situations, both as the true distribution (a continuous one) and as an approximation to discrete distributions. In particular, we illustrated the use of the normal distribution as an approximation to the binomial distribution.

In the following chapters, we will make much use of the material presented here. Most statistical theory relies on the normal distribution and on distributions that are derived from it.

## ADDITIONAL PROBLEMS

**4–59.** The time, in hours, that a copying machine may work without breaking down is a normally distributed random variable with mean 549 and standard deviation 68. Find the probability that the machine will work for at least 500 hours without breaking down.

**4–60.** The yield, in tons of ore per day, at a given coal mine is approximately normally distributed with mean 785 tons and standard deviation 60. Find the probability that at least 800 tons of ore will be mined on a given day. Find the proportion of working days in which anywhere from 750 to 850 tons is mined. Find the probability that on a given day, the yield will be below 665 tons.

**4–61.** Scores on a management aptitude examination are believed to be normally distributed with mean 650 (out of a total of 800 possible points) and standard deviation 50. What is the probability that a randomly chosen manager will achieve a score above 700? What is the probability that the score will be below 750?

**4–62.** The price of a share of Kraft stock is normally distributed with mean 33.30 and standard deviation 6.[10] What is the probability that on a randomly chosen day in the period for which our assumptions are made, the price of the stock will be more than $40 per share? Less than $30 per share?

**4–63.** The amount of oil pumped daily at Standard Oil's facilities in Prudhoe Bay is normally distributed with mean 800,000 barrels and standard deviation 10,000. In determining the amount of oil the company must report as its lower limit of daily production, the company wants to choose an amount such that for 80% of the days, at least the reported amount $x$ is produced. Determine the value of the lower limit $x$.

**4–64.** An analyst believes that the price of an IBM stock is a normally distributed random variable with mean $105 and variance 24. The analyst would like to determine a value such that there is a 0.90 probability that the price of the stock will be greater than that value.[11] Find the required value.

**4–65.** Weekly rates of return (on an annualized basis) for certain securities over a given period are believed to be normally distributed with mean 8.00% and variance 0.25. Give two values $x_1$ and $x_2$ such that you are 95% sure that annualized weekly returns will be between the two values.

**4–66.** The impact of a television commercial, measured in terms of excess sales volume over a given period, is believed to be approximately normally distributed with mean 50,000 and variance 9,000,000. Find 0.99 probability bounds on the volume of excess sales that would result from a given airing of the commercial.

**4–67.** A travel agency believes that the number of people who sign up for tours to Hawaii during the Christmas–New Year's holiday season is an approximately normally distributed random variable with mean 2,348 and standard deviation 762. For reservation purposes, the agency's management wants to find the number of people

[10] Inferred from data in "Business Day," *The New York Times*, April 4, 2007, p. C11.

[11] Inferred from data in "Business Day," *The New York Times*, March 14, 2007, p. C10.

such that the probability is 0.85 that at least that many people will sign up. It also needs 0.80 probability bounds on the number of people who will sign up for the trip.

**4–68.** A loans manager at a large bank believes that the percentage of her customers who default on their loans during each quarter is an approximately normally distributed random variable with mean 12.1% and standard deviation 2.5%. Give a lower bound $x$ with 0.75 probability that the percentage of people defaulting on their loans is at least $x$. Also give an upper bound $x'$ with 0.75 probability that the percentage of loan defaulters is below $x'$.

**4–69.** The power generated by a solar electric generator is normally distributed with mean 15.6 kilowatts and standard deviation of 4.1 kilowatts. We may be 95% sure that the generator will deliver at least how many kilowatts?

**4–70.** Short-term rates fluctuate daily. It may be assumed that the yield for 90-day Treasury bills in early 2007 was approximately normally distributed with mean 4.92% and standard deviation 0.3%.[12] Find a value such that 95% of the time during that period the yield of 90-day T-bills was below this value.

**4–71.** In quality-control projects, engineers use charts where item values are plotted and compared with 3-standard-deviation bounds above and below the mean for the process. When items are found to fall outside the bounds, they are considered nonconforming, and the process is stopped when "too many" items are out of bounds. Assuming a normal distribution of item values, what percentage of values would you expect to be out of bounds when the process is in control? Accordingly, how would you define "too many"? What do you think is the rationale for this practice?

**4–72.** Total annual textbook sales in a certain discipline are normally distributed. Forty-five percent of the time, sales are above 671,000 copies, and 10% of the time, sales are above 712,000 copies. Find the mean and the variance of annual sales.

**4–73.** Typing speed on a new kind of keyboard for people at a certain stage in their training program is approximately normally distributed. The probability that the speed of a given trainee will be greater than 65 words per minute is 0.45. The probability that the speed will be more than 70 words per minute is 0.15. Find the mean and the standard deviation of typing speed.

**4–74.** The number of people responding to a mailed information brochure on cruises of the Royal Viking Line through an agency in San Francisco is approximately normally distributed. The agency found that 10% of the time, over 1,000 people respond immediately after a mailing, and 50% of the time, at least 650 people respond right after the mailing. Find the mean and the standard deviation of the number of people who respond following a mailing.

**4–75.** The Tourist Delivery Program was developed by several European automakers. In this program, a tourist from outside Europe—most are from the United States—may purchase an automobile in Europe and drive it in Europe for as long as six months, after which the manufacturer will ship the car to the tourist's home destination at no additional cost. In addition to the time limitations imposed, some countries impose mileage restrictions so that tourists will not misuse the privileges of the program. In setting the limitation, some countries use a normal distribution assumption. It is believed that the number of kilometers driven by a tourist in the program is normally distributed with mean 4,500 and standard deviation 1,800. If a country wants to set the mileage limit at a point such that 80% of the tourists in the program will want to drive fewer kilometers, what should the limit be?

**4–76.** The number of newspapers demanded daily in a large metropolitan area is believed to be an approximately normally distributed random variable. If more newspapers are demanded than are printed, the paper suffers an opportunity loss,

[12] From "Business Day," *The New York Times*, March 14, 2007, p. C11.

in that it could have sold more papers, and a loss of public goodwill. On the other hand, if more papers are printed than will be demanded, the unsold papers are returned to the newspaper office at a loss. Suppose that management believes that guarding against the first type of error, unmet demand, is most important and would like to set the number of papers printed at a level such that 75% of the time, demand for newspapers will be lower than that point. How many papers should be printed daily if the average demand is 34,750 papers and the standard deviation of demand is 3,560?

**4–77.** The Federal Funds rate in spring 2007 was approximately normal with $\mu = 5.25\%$ and $\sigma = 0.05\%$. Find the probability that the rate on a given day will be less than 1.1%.[13]

**4–78.** Thirty-year fixed mortgage rates in April 2007 seemed normally distributed with mean 6.17%.[14] The standard deviation is believed to be 0.25%. Find a bound such that the probability that the actual rate obtained will be this number or below it is 90%.

**4–79.** A project consists of three phases to be completed one after the other. The duration of each phase, in days, is normally distributed as follows: Duration of Phase I $\sim N(84, 3^2)$; Duration of Phase II $\sim N(102, 4^2)$; Duration of Phase III $\sim N(62, 2^2)$. The durations are independent.

*a.* Find the distribution of the project duration. Report the mean and the standard deviation.

*b.* If the project duration exceeds 250 days, a penalty will be assessed. What is the probability that the project will be completed within 250 days?

*c.* If the project is completed within 240 days, a bonus will be earned. What is the probability that the project will be completed within 240 days?

**4–80.** The GMAT scores of students who are potential applicants to a university are normally distributed with a mean of 487 and a standard deviation of 98.

*a.* What percentage of students will have scores exceeding 500?

*b.* What percentage of students will have scores between 600 and 700?

*c.* If the university wants only the top 75% of the students to be eligible to apply, what should be the minimum GMAT score specified for eligibility?

*d.* Find the narrowest interval that will contain 75% of the students' scores.

*e.* Find $x$ such that the interval $[x, 2x]$ will contain 75% of the students' scores. (There are two answers. See if you can find them both.)

**4–81.** The profit (or loss) from an investment is normally distributed with a mean of $11,200 and a standard deviation of $8,250.

*a.* What is the probability that there will be a loss rather than a profit?

*b.* What is the probability that the profit will be between $10,000 and $20,000?

*c.* Find $x$ such that the probability that the profit will exceed $x$ is 25%.

*d.* If the loss exceeds $10,000 the company will have to borrow additional cash. What is the probability that the company will have to borrow additional cash?

*e.* Calculate the value at risk.

**4–82.** The weight of connecting rods used in an automobile engine is to be closely controlled to minimize vibrations. The specification is that each rod must be $974 \pm 1.2$ grams. The half-width of the specified interval, namely, 1.2 grams, is known as the *tolerance.* The manufacturing process at a plant produces rods whose weights are

[13] www.federalreserve.gov

[14] "Figures of the Week," *BusinessWeek,* April 30, 2007, p. 95.

normally distributed with a mean $\mu$ of 973.8 grams and a standard deviation $\sigma$ of 0.32 grams.

*a.* What proportion of the rods produced by this process will be acceptable according to the specification?

*b.* The *process capability index,* denoted by $C_p$, is given by the formula

$$C_p = \frac{\text{Tolerance}}{3 * \sigma}$$

Calculate $C_p$ for this process.

*c.* Would you say a larger value or a smaller value of $C_p$ is preferable?

*d.* The mean of the process is 973.8 grams, which does not coincide with the target value of 974 grams. The difference between the two is the *offset,* defined as the difference and therefore always positive. Clearly, as the offset increases, the chances of a part going outside the specification limits increase. To take into account the effect of the offset, another index, denoted by $C_{pk}$, is defined as

$$C_{pk} = C_p - \frac{\text{Offset}}{3 * \sigma}$$

Calculate $C_{pk}$ for this process.

*e.* Suppose the process is adjusted so that the offset is zero, and $\sigma$ remains at 0.32 gram. Now, what proportion of the parts made by the process will fall within specification limits?

*f.* A process has a $C_p$ of 1.2 and a $C_{pk}$ of 0.9. What proportion of the parts produced by the process will fall within specification limits? (*Hint:* One way to proceed is to assume that the target value is, say, 1,000, and $\sigma = 1$. Next, find the tolerance, the specification limits, and the offset. You should then be able to answer the question.)

**4–83.** A restaurant has three sources of revenue: eat-in orders, takeout orders, and the bar. The daily revenue from each source is normally distributed with mean and standard deviation shown in the table below.

| | Mean | Standard Deviation |
|---|---|---|
| Eat in | $5,780 | $142 |
| Takeout | 641 | 78 |
| Bar | 712 | 72 |

*a.* Will the total revenue on a day be normally distributed?

*b.* What are the mean and standard deviation of the total revenue on a particular day?

*c.* What is the probability that the revenue will exceed $7,000 on a particular day?

# CASE 4 Acceptable Pins

A company supplies pins in bulk to a customer. The company uses an automatic lathe to produce the pins. Due to many causes—vibration, temperature, wear and tear, and the like—the lengths of the pins made by the machine are normally distributed with a mean of 1.012 inches and a standard deviation of 0.018 inch. The customer will buy only those pins with lengths in the interval 1.00 $\pm$ 0.02 inch. In other words, the customer wants the length to be 1.00 inch but will accept up to 0.02 inch deviation on either side. This 0.02 inch is known as the *tolerance.*

1. What percentage of the pins will be acceptable to the consumer?

In order to improve percentage accepted, the production manager and the engineers discuss adjusting the population mean and standard deviation of the length of the pins.

2. If the lathe can be adjusted to have the mean of the lengths to any desired value, what should it be adjusted to? Why?
3. Suppose the mean cannot be adjusted, but the standard deviation can be reduced. What maximum value of the standard deviation would make 90% of the parts acceptable to the consumer? (Assume the mean to be 1.012.)
4. Repeat question 3, with 95% and 99% of the pins acceptable.
5. In practice, which one do you think is easier to adjust, the mean or the standard deviation? Why?

The production manager then considers the costs involved. The cost of resetting the machine to adjust the population mean involves the engineers' time and the cost of production time lost. The cost of reducing the population standard deviation involves, in addition to these costs, the cost of overhauling the machine and reengineering the process.

6. Assume it costs \$150 $x^2$ to decrease the standard deviation by $(x/1000)$ inch. Find the cost of reducing the standard deviation to the values found in questions 3 and 4.
7. Now assume that the mean has been adjusted to the best value found in question 2 at a cost of \$80. Calculate the reduction in standard deviation necessary to have 90%, 95%, and 99% of the parts acceptable. Calculate the respective costs, as in question 6.
8. Based on your answers to questions 6 and 7, what are your recommended mean and standard deviation?

# CASE 5 Multicurrency Decision

A company sells precision grinding machines to four customers in four different countries. It has just signed a contract to sell, two months from now, a batch of these machines to each customer. The following table shows the number of machines (batch quantity) to be delivered to the four customers. The selling price of the machine is fixed in the local currency, and the company plans to convert the local currency at the exchange rate prevailing at the time of delivery. As usual, there is uncertainty in the exchange rates. The sales department estimates the exchange rate for each currency and its standard deviation, expected at the time of delivery, as shown in the table. Assume that the exchange rates are normally distributed and independent.

| Customer | Batch Quantity | Selling Price | Exchange Rate: Mean | Exchange Rate: Standard Deviation |
|---|---|---|---|---|
| 1 | 12 | £ 57,810 | \$1.41/£ | \$0.041/£ |
| 2 | 8 | ¥ 8,640,540 | \$0.00904/¥ | \$0.00045/¥ |
| 3 | 5 | €97,800 | \$0.824/€ | \$0.0342/€ |
| 4 | 2 | R 4,015,000 | \$0.0211/R | \$0.00083/R |

1. Find the distribution of the uncertain revenue from the contract in U.S. dollars. Report the mean, the variance, and the standard deviation.
2. What is the probability that the revenue will exceed \$2,250,000?

3. What is the probability that the revenue will be less than $2,150,000?
4. To remove the uncertainty in the revenue amount, the sales manager of the company looks for someone who would assume the risk. An international bank offers to pay a sure sum of $2,150,000 in return for the revenue in local currencies. What useful facts can you tell the sales manager about the offer, without involving any of your personal judgment?
5. What is your recommendation to the sales manager, based on your personal judgment?
6. If the sales manager is willing to accept the bank's offer, but the CEO of the company is not, who is more risk-averse?
7. Suppose the company accepts the bank's offer. Now consider the bank's risk, assuming that the bank will convert all currencies into U.S. dollars at the prevailing exchange rates. What is the probability that the bank will incur a loss?
8. The bank defines its value at risk as the loss that occurs at the 5th percentile of the uncertain revenue. What is the bank's value at risk?
9. What is the bank's expected profit?
10. Express the value at risk as a percentage of the expected profit. Based on this percentage, what is your evaluation of the risk faced by the bank?
11. Suppose the bank does not plan to convert all currencies into U.S. dollars, but plans to spend or save them as local currency or convert them into some other needed currency. Will this increase or decrease the risk faced by the bank?
12. Based on the answer to part 11, is the assumption (made in parts 7 to 10) that the bank will convert all currencies into U.S. dollars a good assumption?