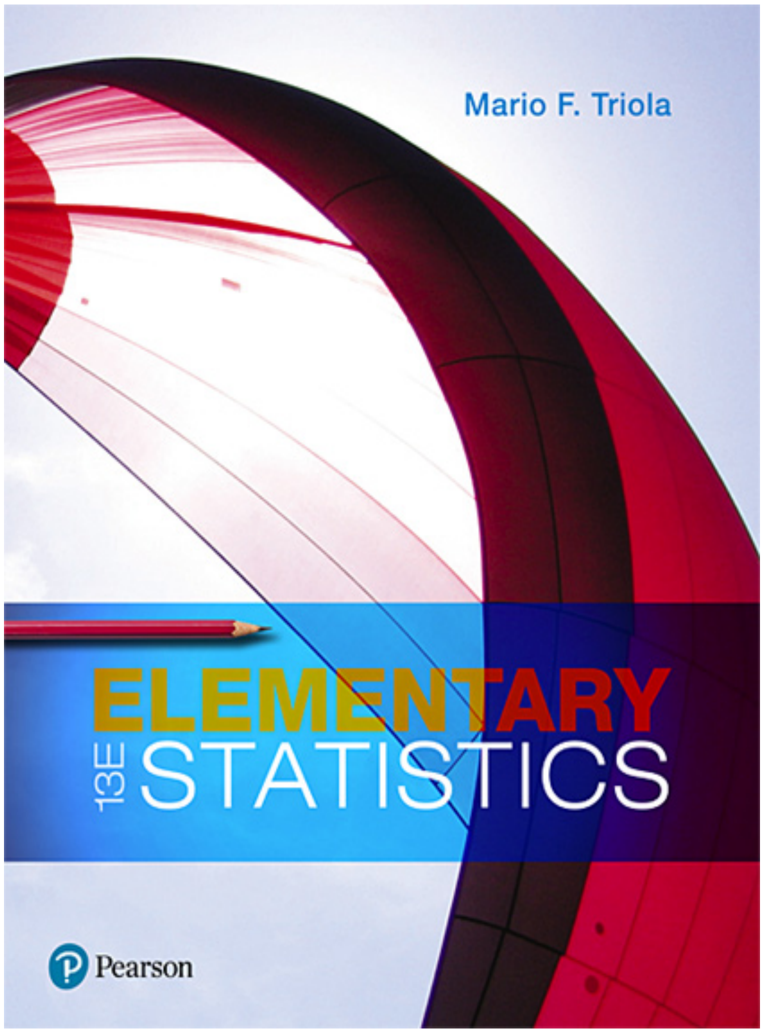
Elementary Statistics

Thirteenth Edition



Chapter 6

Normal Probability Distributions

Normal Probability Distributions

6-1 The Standard Normal Distribution

6-2 Real Applications of Normal Distributions

6-3 Sampling Distributions and Estimators

6-4 The Central Limit Theorem

6-5 Assessing Normality

6-6 Normal as Approximation to Binomial

Key Concept

In this section we present the **standard normal distribution,** which is a specific normal distribution having the following three properties:

1. Bell-shaped: The graph of the standard normal distribution is bell-shaped.
2. $\mu = 0$: The standard normal distribution has a mean equal to 0.
3. $\sigma = 1$: The standard normal distribution has a standard deviation equal to 1.

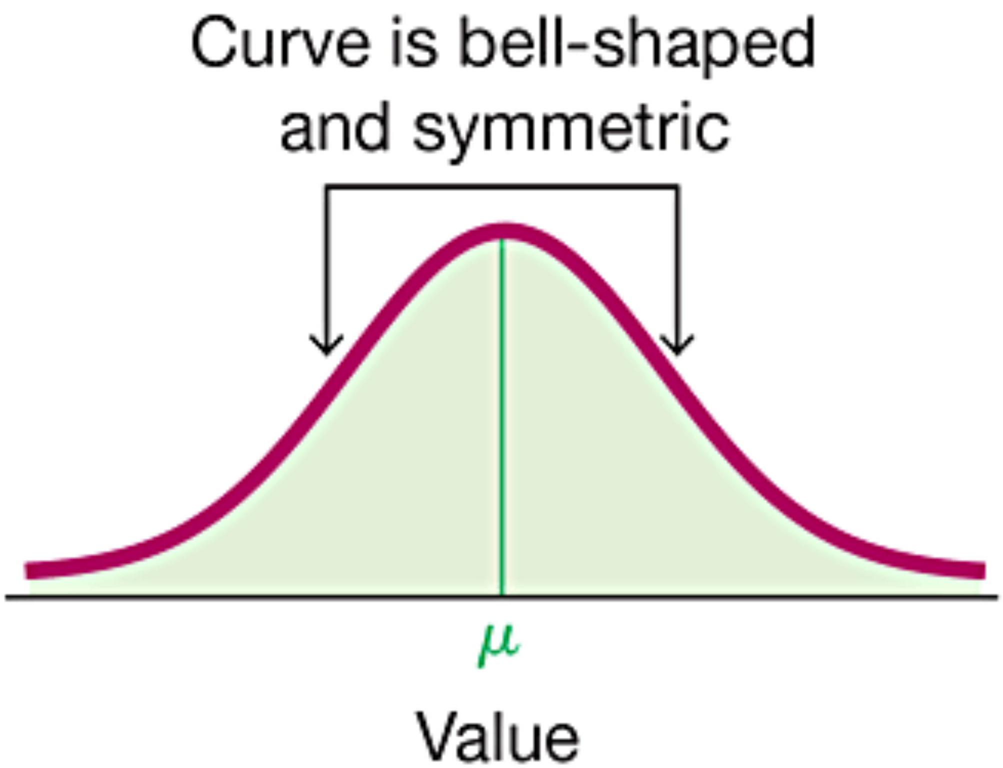
In this section we develop the skill to find areas (or probabilities or relative frequencies) corresponding to various regions under the graph of the standard normal distribution. In addition, we find *z* scores that correspond to areas under the graph.

Normal Distribution (1 of 2)

- Normal Distribution
  - If a continuous random variable has a distribution with a graph that is symmetric and bell-shaped, we say that it has a **normal distribution.**

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

Normal Distribution (2 of 2)



Uniform Distribution (1 of 2)

Properties of uniform distribution:

1. The area under the graph of a continuous probability distribution is equal to 1.
2. There is a correspondence between area and probability, so probabilities can be found by identifying the corresponding areas in the graph using this formula for the area of a rectangle:
   Area = height × width

Uniform Distribution (2 of 2)

- Uniform Distribution
  - A continuous random variable has a **uniform distribution** if its values are spread **evenly** over the range of possibilities. The graph of a uniform distribution results in a rectangular shape.

Density Curve

- Density Curve
  - The graph of any continuous probability distribution is called a **density curve,** and any density curve must satisfy the requirement that the total area under the curve is exactly 1.

Because the total area under any density curve is equal to 1, there is a correspondence between area and probability.

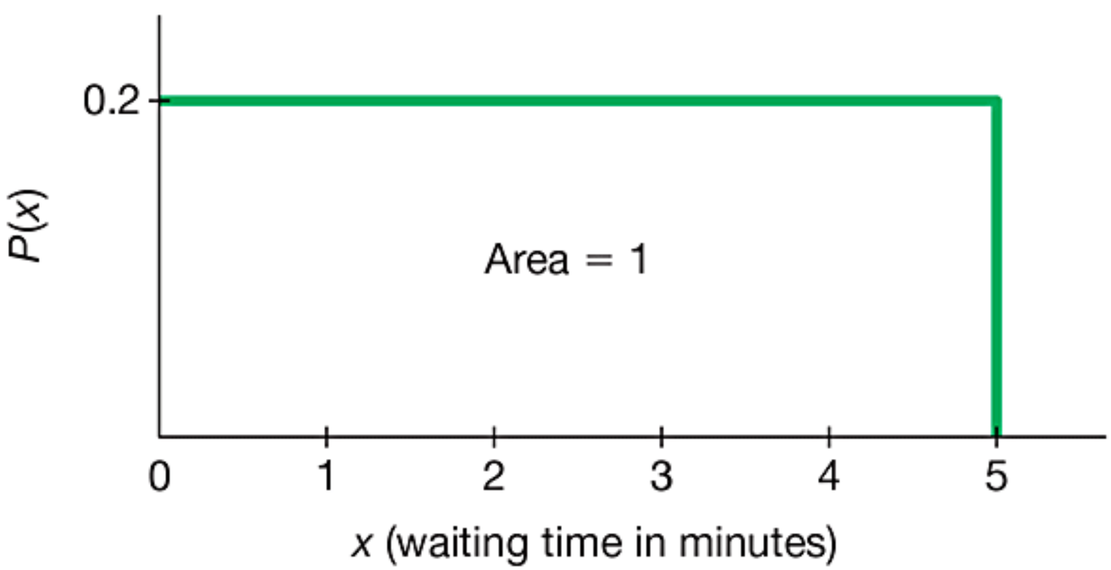
Example: Waiting Times for Airport Security (1 of 7)

During certain time periods at JFK airport in New York City, passengers arriving at the security checkpoint have waiting times that are uniformly distributed between 0 minutes and 5 minutes, as illustrated in the figure on the next page.

Example: Waiting Times for Airport Security (2 of 7)

Refer to the figure to see these properties:

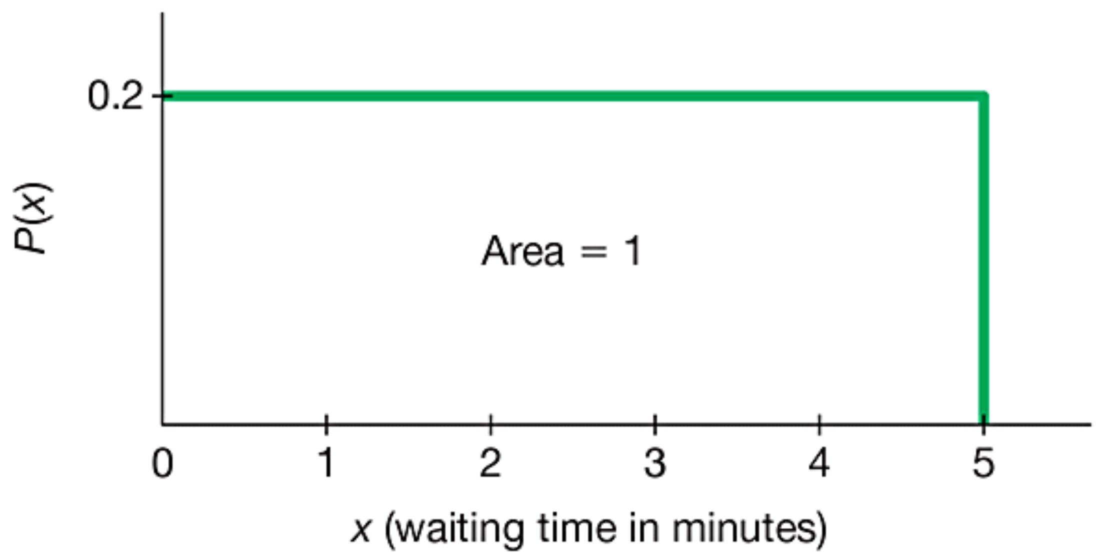
- All of the different possible waiting times are **equally likely.**



Example: Waiting Times for Airport Security (3 of 7)

Refer to the figure to see these properties:

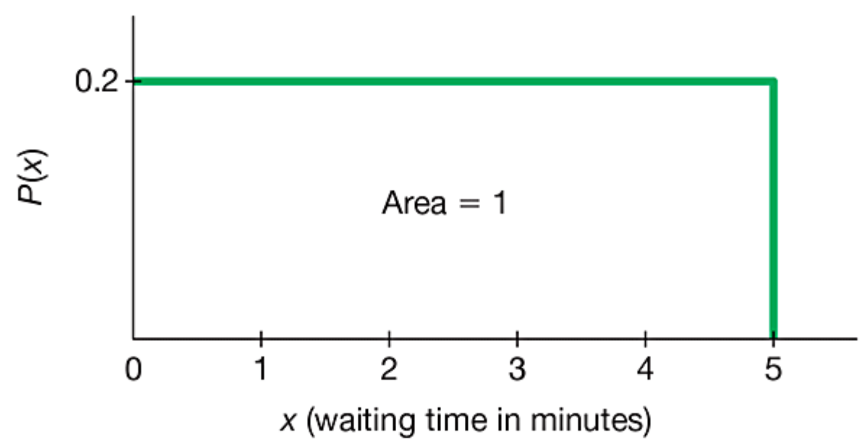
- Waiting times can be **any** value between 0 min and 5 min, so it is possible to have a waiting time of 1.234567 min.



Example: Waiting Times for Airport Security (4 of 7)

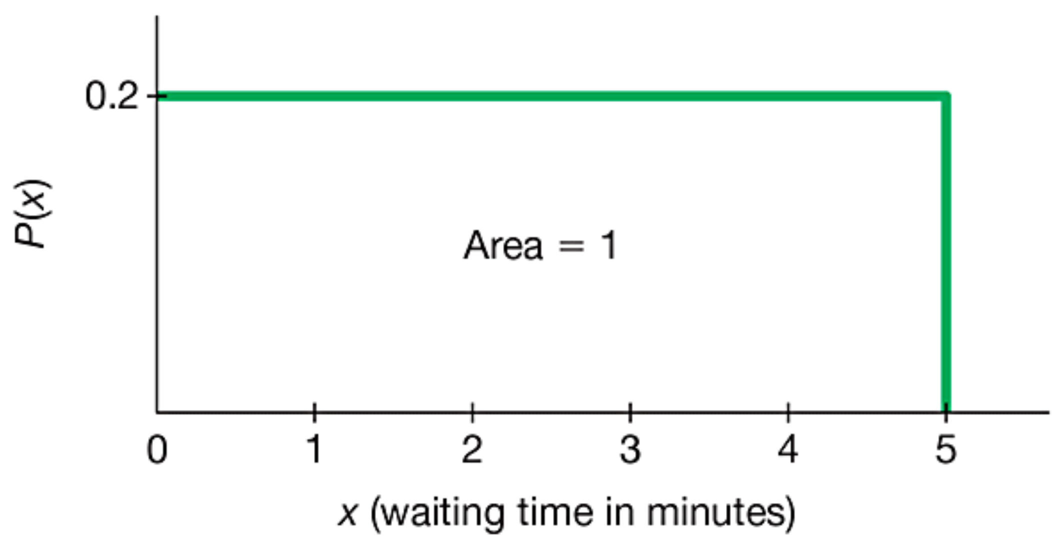
Refer to the figure to see these properties:

- By assigning the probability of 0.2 to the height of the vertical line in the figure, the **enclosed area is exactly 1.**
  (In general, we should make the height of the vertical line in a uniform distribution equal to $\frac{1}{\text{range}}$.)



Example: Waiting Times for Airport Security (5 of 7)

Given the uniform distribution illustrated in the figure, find the probability that a randomly selected passenger has a waiting time of at least 2 minutes.



Example: Waiting Times for Airport Security (6 of 7)

Solution

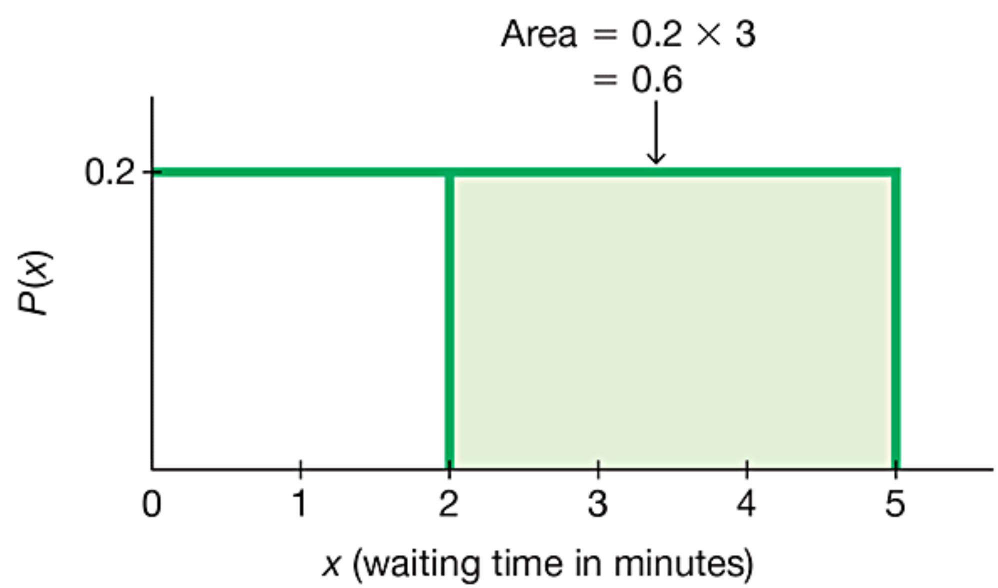
The shaded area represents waiting times of at least 2 minutes. Because the total area under the density curve is equal to 1, there is a correspondence between area and probability. We can easily find the desired **probability** by using **areas.**

Example: Waiting Times for Airport Security (7 of 7)

Solution

P(wait time of at least 2 min) = height × width of shaded area in the figure = 0.2 × 3 = 0.6

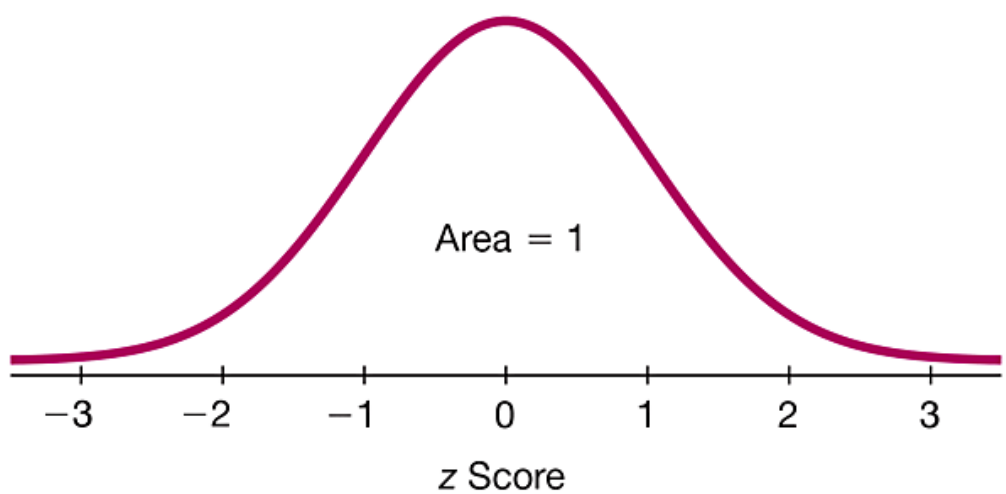
The probability of randomly selecting a passenger with a waiting time of at least 2 minutes is 0.6.



Pearson

Standard Normal Distribution

- Standard Normal Distribution
  - The **standard normal distribution** is a normal distribution with the parameters of $\mu = 0$ and $\sigma = 1$. The total area under its density curve is equal to 1.



Finding Probabilities When Given z Scores (1 of 3)

- We can find areas (probabilities) for different regions under a normal model using technology or Table A-2.
- Technology is strongly recommended.

Because calculators and software generally give more accurate results than Table A-2, we **strongly** recommend using technology.

Finding Probabilities When Given z Scores (2 of 3)

If using Table A-2, it is essential to understand these points:

1. Table A-2 is designed only for the **standard** normal distribution, which is a normal distribution with a mean of 0 and a standard deviation of 1.
2. Table A-2 is on two pages, with the left page for **negative** *z* scores and the right page for **positive** *z* scores.
3. Each value in the body of the table is a **cumulative area from the left** up to a vertical boundary above a specific *z* score.

Finding Probabilities When Given z Scores (3 of 3)

4. When working with a graph, avoid confusion between *z* scores and areas.

***z* score: Distance along the horizontal scale of the standard normal distribution (corresponding to the number of standard deviations above or below the mean); refer to the leftmost column and top row of Table A-2.**

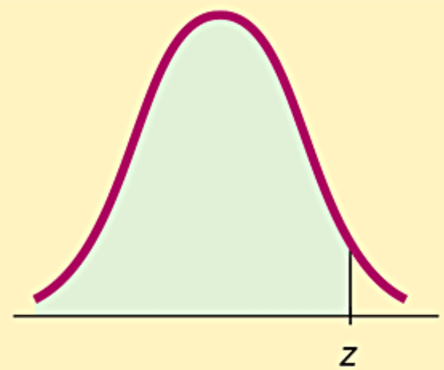
Area: Region under the curve; refer to the values in the body of Table A-2.

5. The part of the *z* score denoting hundredths is found across the top row of Table A-2.

Formats Used for Finding Normal Distribution Areas

Cumulative Area from the Left
The following provide the *cumulative area from the left* up to a vertical line above a specific value of z:

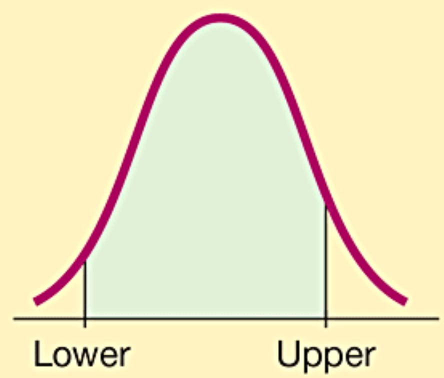
- **Table A-2**
- **Statdisk**
- **Minitab**
- **Excel**
- **StatCrunch**



Cumulative Left Region

Area Between Two Boundaries
The following provide the area bounded on the left and bounded on the right by vertical lines above specific values.

- **TI-83/84 Plus calculator**
- **StatCrunch**



Area Between Two Boundaries

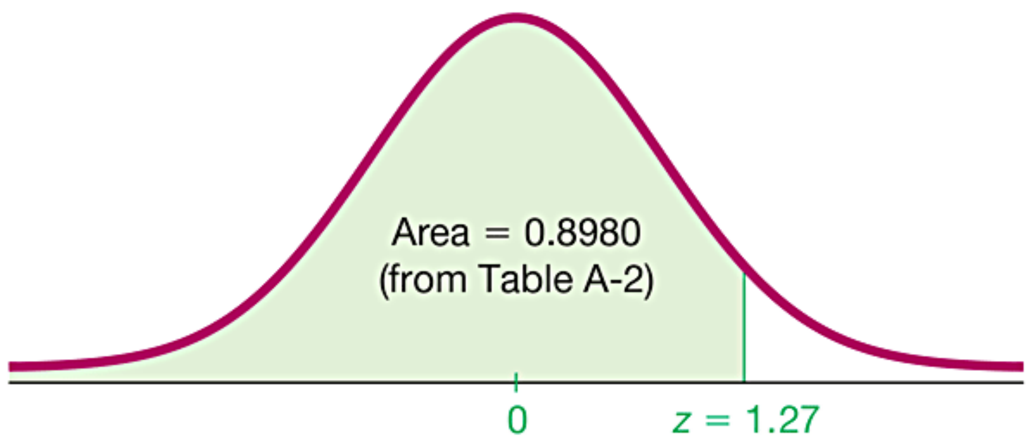
Example: Bone Density Test (1 of 7)

A bone mineral density test can be helpful in identifying the presence or likelihood of osteoporosis. The result of a bone density test is commonly measured as a *z* score. The population of *z* scores is normally distributed with a mean of 0 and a standard deviation of 1, so these test results meet the requirements of a standard normal distribution.

Example: Bone Density Test (2 of 7)

The graph of the bone density test scores is as shown in the figure.

A randomly selected adult undergoes a bone density test. Find the probability that this person has a bone density test score less than 1.27.

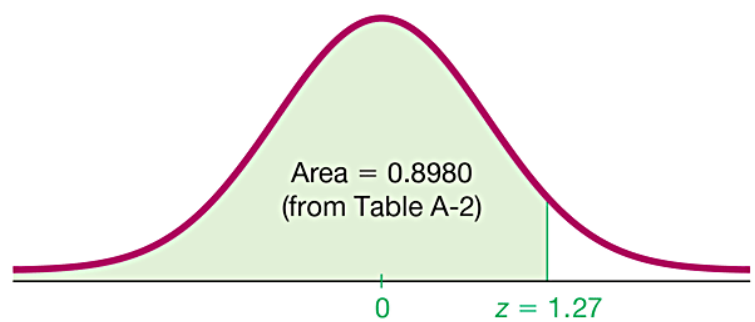


Example: Bone Density Test (3 of 7)

Solution

Note that the following are the **same** (because of the aforementioned correspondence between probability and area):

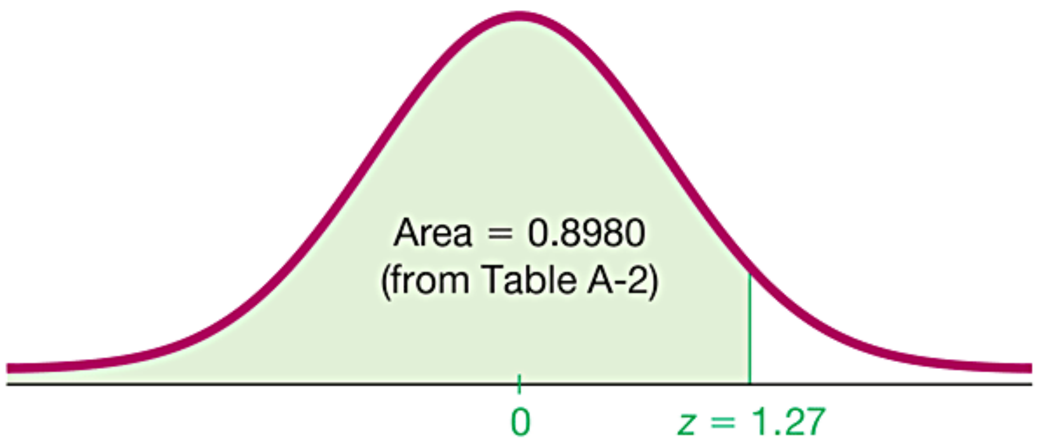
- **Probability** that the bone density test score is less than 1.27
- Shaded **area** shown in the figure



Example: Bone Density Test (4 of 7)

Solution

So we need to find the area in the figure to the left of $z = 1.27$.



Example: Bone Density Test (5 of 7)

Solution

Using Table A-2, begin with the *z* score of 1.27 by locating 1.2 in the left column; next find the value in the adjoining row of probabilities that is directly below 0.07, as shown:

z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
0.0	.5000	.5040	.5080	.5120	.5160	.5199	.5239	.5279	.5319	.5359
0.1	.5398	.5438	.5478	.5517	.5557	.5596	.5636	.5675	.5714	.5753
0.2	.5793	.5832	.5871	.5910	.5948	.5987	.6026	.6064	.6103	.6141
1.0	.8413	.8438	.8461	.8485	.8508	.8531	.8554	.8577	.8599	.8621
1.1	.8643	.8665	.8686	.8708	.8729	.8749	.8770	.8790	.8810	.8830
1.2	.8849	.8869	.8888	.8907	.8925	.8944	.8962	.8980	.8997	.9015
1.3	.9032	.9049	.9066	.9082	.9099	.9115	.9131	.9147	.9162	.9177
1.4	.9192	.9207	.9222	.9236	.9251	.9265	.9279	.9292	.9306	.9319

Example: Bone Density Test (6 of 7)

Solution

Table A-2 shows that there is an area of 0.8980 corresponding to $z = 1.27$. We want the area **below** 1.27, and Table A-2 gives the cumulative area from the left, so the desired area is 0.8980.

Because of the correspondence between area and probability, we know that the probability of a z score below 1.27 is 0.8980.

Example: Bone Density Test (7 of 7)

Interpretation

The **probability** that a randomly selected person has a bone density test result below 1.27 is 0.8980, shown as the shaded region. Another way to interpret this result is to conclude that 89.80% of people have bone density levels below 1.27.

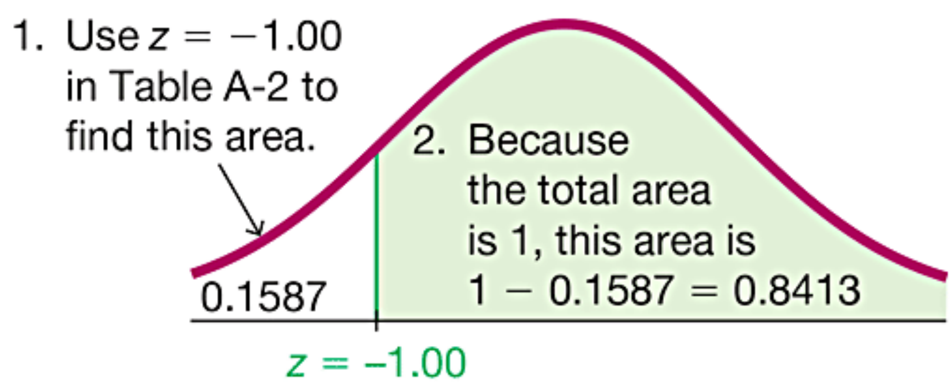
Example: Bone Density Test: Finding the Area to the Right of a Value (1 of 4)

Using the same bone density test, find the probability that a randomly selected person has a result above −1.00 (which is considered to be in the "normal" range of bone density readings).

Example: Bone Density Test: Finding the Area to the Right of a Value (2 of 4)

Solution

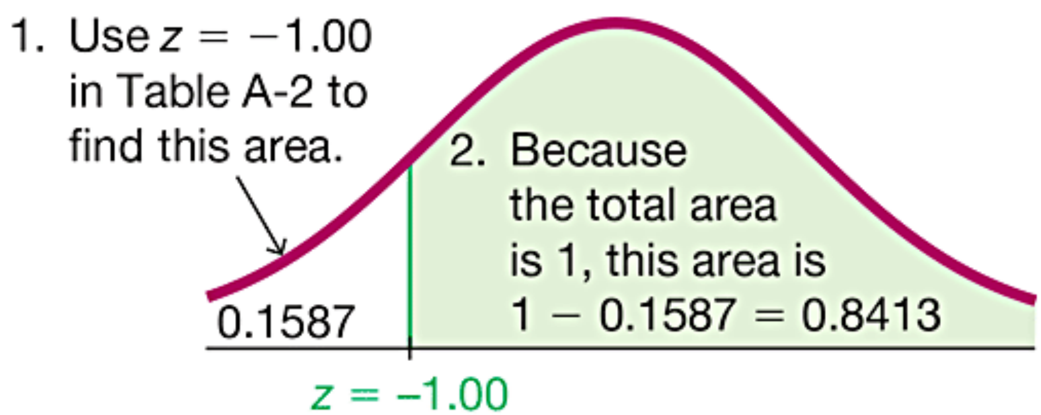
If we use Table A-2, we should know that it is designed to apply only to cumulative areas from the **left.** Referring to the page with **negative** z scores, we find that the cumulative area from the left up to $z = -1.00$ is 0.1587, as shown in the figure.



Example: Bone Density Test: Finding the Area to the Right of a Value (3 of 4)

Solution

Because the total area under the curve is 1, we can find the shaded area by subtracting 0.1587 from 1. The result is 0.8413.



Example: Bone Density Test: Finding the Area to the Right of a Value (4 of 4)

Interpretation

Because of the correspondence between probability and area, we conclude that the **probability** of randomly selecting someone with a bone density reading above −1 is 0.8413 (which is the **area** to the right of $z = -1.00$). We could also say that 84.13% of people have bone density levels above −1.00.

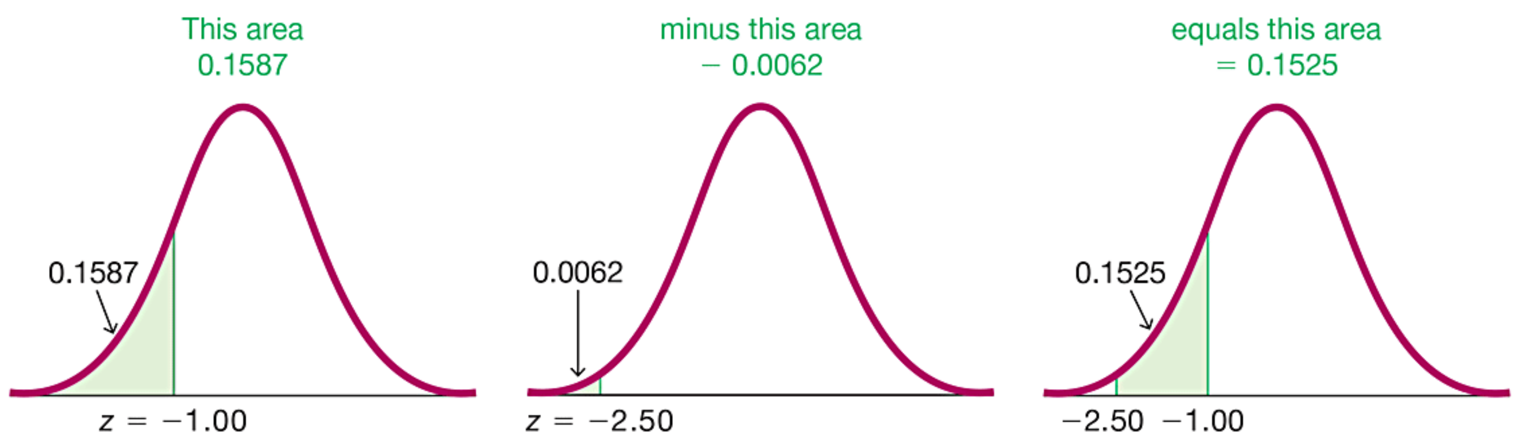
Example: Bone Density Test: Finding the Area Between Two Values (1 of 3)

A bone density reading between −1.00 and −2.50 indicates the subject has osteopenia, which is some bone loss. Find the probability that a randomly selected subject has a reading between −1.00 and −2.50.

Example: Bone Density Test: Finding the Area Between Two Values (2 of 3)

Solution

1. The area to the left of $z = -1.00$ is 0.1587.
2. The area to the left of $z = -2.50$ is 0.0062.
3. The area between $z = -1.00$ and $z = -2.50$ is the difference between the areas found above.



Example: Bone Density Test: Finding the Area Between Two Values (3 of 3)

Interpretation

Using the correspondence between probability and area, we conclude that there is a probability of 0.1525 that a randomly selected subject has a bone density reading between −1.00 and −2.50.

Another way to interpret this result is to state that 15.25% of people have osteopenia, with bone density readings between −1.00 and −2.50.

Generalized Rule

The area corresponding to the region **between** two z scores can be found by finding the difference between the two areas found in Table A-2.

Don’t try to memorize a rule or formula for this case. Focus on **understanding** by using a graph. Draw a graph, shade the desired area, and then get creative to think of a way to find the desired area by working with cumulative areas from the left.

Notation

- $P(a < z < b)$ denotes the probability that the z score is between a and b.
- $P(z > a)$ denotes the probability that the z score is greater than a.
- $P(z < a)$ denotes the probability that the z score is less than a.

Finding *z* Scores from Known Areas

1. Draw a bell-shaped curve and identify the region under the curve that corresponds to the given probability. If that region is not a cumulative region from the left, work instead with a known region that is a cumulative region from the left.
2. Use technology or Table A-2 to find the z score. With Table A-2, use the cumulative area from the left, locate the closest probability in the **body** of the table, and identify the corresponding z score.

Critical Value (1 of 2)

- Critical Value
  - For the standard normal distribution, a **critical value** is a *z* score on the borderline separating those *z* scores that are **significantly low** or **significantly high.**

Critical Value (2 of 2)

Notation

The expression z_α denotes the z score with an area of α to its right.

Example: Finding the Critical Value

z_{α} (1 of 3)

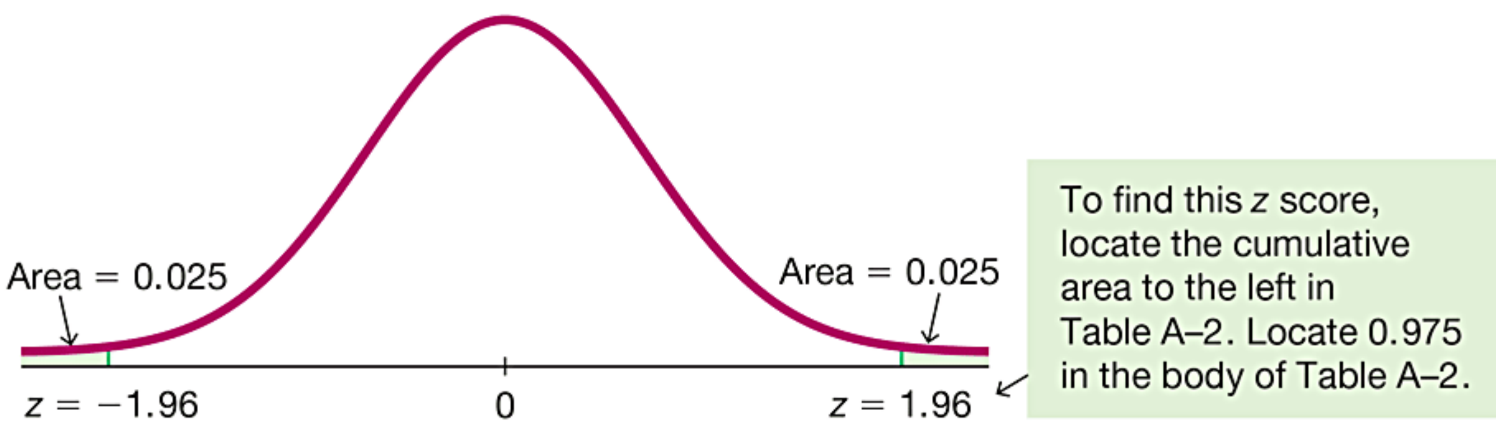
Find the value of $z_{0.025}$. (Let $\alpha = 0.025$ in the expression z_{α}.)

Example: Finding the Critical Value

z_α (2 of 3)

Solution

The notation of $z_{0.025}$ is used to represent the z score with an area of 0.025 to its **right.** Refer to the figure and note that the value of z = 1.96 has an area of 0.025 to its right, so $z_{0.025}$ = 1.96. Note that $z_{0.025}$ corresponds to a cumulative left area of 0.975.



Example: Finding the Critical Value

z_α (3 of 3)

CAUTION

When finding a value of z_α for a particular value of α, note that α is the area to the **right** of z_α, but Table A-2 and some technologies give cumulative areas to the **left** of a given z score.

To find the value of z_α, resolve that conflict by using the value of $1 - \alpha$. For example, to find $z_{0.1}$, refer to the z score with an area of 0.9 to its left.