6| The Normal Distribution

Chapter 6 Table of Contents

Adapted from: Claude Laflamme, Business Statistics -- BSTA 200 -- Humber College -- Version 2016 Revision A. OpenStax CNX. Jan 10, 2020 http://cnx.org/contents/0bec2053-5fc5-49c9-a97a-c33b7e0a095c@12.342.

Introduction

Figure 6.1 If you ask enough people about their shoe size, you will find that your graphed data is shaped like a bell curve and can be described as normally distributed. (credit: Ömer Ünlü)

The normal, a continuous distribution, is the most important of all the distributions. It is widely used and even more widely abused. Its graph is bell-shaped. You see the bell curve in almost all disciplines. Some of these include psychology, business, economics, the sciences, nursing, and, of course, mathematics. Some of your instructors may use the normal distribution to help determine your grade. Most IQ scores are normally distributed. Often real-estate prices fit a normal distribution. The normal distribution is extremely important, but it cannot be applied to everything in the real world.

In this chapter, you will study the normal distribution, the standard normal distribution, and applications associated with them.

The normal distribution has two parameters (two numerical descriptive measures), the mean (μ) and the standard deviation (σ). If X is a quantity to be measured that has a normal distribution with mean (μ) and standard deviation (σ), we designate this by writing

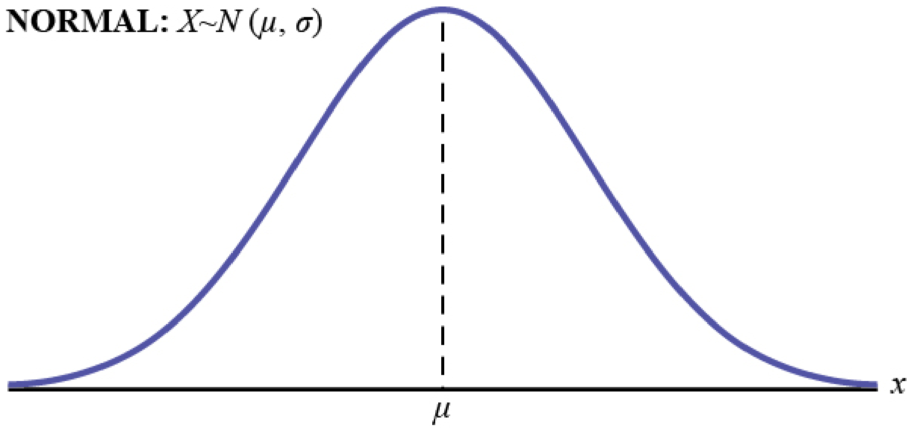


Figure 6.2

The probability density function is a rather complicated function. **Do not memorize it**. It is not necessary.

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

The cumulative distribution function is $P(X < x)$. It is calculated either by a calculator or a computer, or it is looked up in a table.

The curve is symmetrical about a vertical line drawn through the mean, μ. The mean is the same as the median, which is the same as the mode, because the graph is symmetric about μ. As the notation indicates, the normal distribution depends only on the mean and the standard deviation. Since the area under the curve must equal one, a change in the standard deviation, σ, causes a change in the shape of the curve; the curve becomes fatter and wider or skinnier and taller depending on σ. A change in μ causes the graph to shift to the left or right. This means there are an infinite number of normal probability distributions. One of special interest is called the **standard normal distribution**.

6.1 | The Standard Normal Distribution

The **standard normal distribution** is a normal distribution of **standardized values called *z*-scores**. **A z-score is measured in units of the standard deviation.** For example, if the mean of a normal distribution is five and the standard deviation is two, the value 11 is three standard deviations above (or to the right of) the mean. The calculation is as follows:

$$z = \frac{x - \mu}{\sigma} = \frac{11 - 5}{2} = 3$$

The mean for the standard normal distribution is zero, and the standard deviation is one. That is, $Z \sim N(0, 1)$. The value x comes from a normal distribution with mean μ and standard deviation σ.

z-Scores

If X is a normally distributed random variable and $X \sim N(\mu, \sigma)$, then the z-score is:

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

The *z*-score tells you how many standard deviations the value *x* is above (to the right of) or below (to the left of) the mean, *μ*. Values of x that are larger than the mean have positive z-scores, and values of x that are smaller than the mean have negative z-scores. If x equals the mean, then x has a z-score of zero.

Example 6-1

Suppose $X \sim N(5, 6)$. This means that x is a normally distributed random variable with mean $\mu = 5$ and standard deviation $\sigma = 6$. Suppose $x = 17$. Find the z-score for $x = 17$ and $x = 1$ and interpret.

Solution 6-1

For $\boldsymbol{x = 17}$:

$$z = \frac{x - \mu}{\sigma} = \frac{17 - 5}{6} = 2$$

This means that **$x = 17$ is two standard deviations (2σ) above (or to the right) of the mean $\mu = 5$**. The standard deviation is $\sigma = 6$.

Now suppose $x = 1$. Then:

$$z = \frac{1 - 5}{6} \approx -0.67$$

This means that $x = 1$ is 0.67 standard deviations below (or to the left) of the mean $\mu = 5$.

Summarizing, when z is positive, x is above or to the right of μ. And when z is negative, x is to the left of or below μ. Or, when z is positive, x is greater than μ, and when z is negative x is less than μ.

Try It 6-1

What is the z-score of x, when x = 7.5 and $X \sim N(12,3)$?

Chapter 6 Try It Solutions

Example 6-2

Some doctors believe that a person can lose five pounds, on the average, in a month by reducing his or her fat intake and by exercising consistently. Suppose weight loss has a normal distribution. Let X = the amount of weight lost(in pounds) by a person in a month. Use a standard deviation of two pounds. $X \sim N(5, 2)$. Fill in the blanks.

a) Suppose a person **lost** ten pounds in a month. The z-score when x = 10 pounds is z = ________ . This z-score tells you that x = 10 is _________ standard deviations to the _________ (right or left) of the mean which is _______ .

b) Suppose a person **gained** three pounds (a negative weight loss). Then z = ___________. This z-score tells you that x = −3 is _________ standard deviations to the ___________ (right or left) of the mean.

Solution 6-2

a) $z = \frac{10-5}{2} = 2.5$

This z-score tells you that x = 10 is **2.5** standard deviations to the **right** of the mean which is **five**.

b) $z = \frac{-3-5}{2} = -4$

This z-score tells you that x = −3 is four standard deviations to the **left** of the mean.

Suppose the random variables X_1 and X_2 have the following normal distributions: $X_1 \sim N(5, 6)$ and $X_2 \sim N(2, 1)$.

If $x_1 = 17,\ z = \frac{x_1-\mu}{\sigma} = \frac{17-5}{6} = 2$ (μ=5, σ=6)

If $x_2 = 4,\ z = \frac{x_2-\mu}{\sigma} = \frac{4-2}{1} = 2$ (μ=2, σ=1)

Therefore, x_1 = 17 and x_2 = 4 are both two (of **their own**) standard deviations to the right of **their** respective means.

The *z*-score allows us to compare data that are scaled differently. To understand the concept, suppose X_1 ~ *N*(5, 6) represents weight gains for one group of people who are trying to gain weight and X_1 ~ *N*(2, 1) measures the weight gain for a second group of people. A negative weight gain would be a weight loss. Since x_2 = 17 and x_2 = 4 are each two standard deviations to the right of their means, they represent the same standardized weight gain **relative to their means**.

Try It 6-2

Fill in the blanks.

Jerome averages 16 points a game with a standard deviation of four points. *X* ~ *N*(16,4). Suppose Jerome scores ten points in a game. The *z*–score when *x* = 10 is ______. This score tells you that *x* = 10 is ______ standard deviations to the ______(right or left) of the mean which is ______.

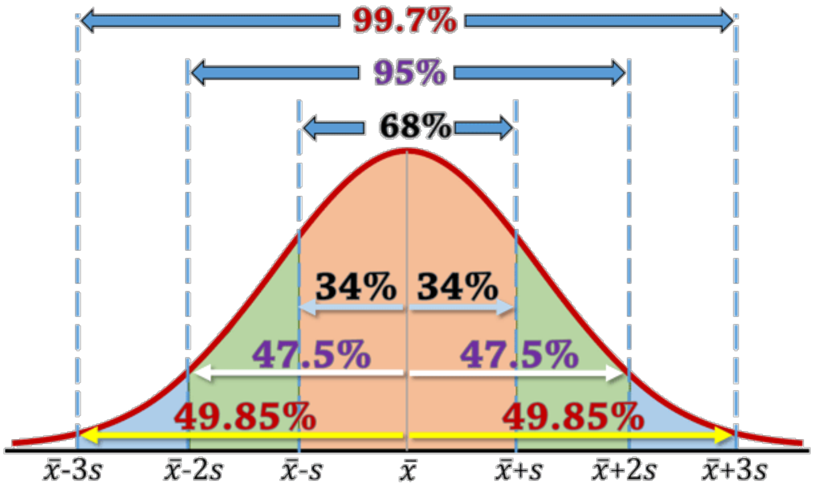
Chapter 6 Try It Solutions

The Empirical Rule

If *X* is a random variable and has a normal distribution with mean *μ* and standard deviation *σ*, then the **Empirical Rule** says the following:

- About 68% of the *x* values lie between –1*σ* and +1*σ* of the mean *μ* (within one standard deviation of the mean).
- About 95% of the *x* values lie between –2*σ* and +2*σ* of the mean *μ* (within two standard deviations of the mean).
- About 99.7% of the *x* values lie between –3*σ* and +3*σ* of the mean *μ* (within three standard deviations of the mean). Notice that almost all the *x* values lie within three standard deviations of the mean.
- The *z*-scores for +1*σ* and –1*σ* are +1 and –1, respectively.
- The *z*-scores for +2*σ* and –2*σ* are +2 and –2, respectively.
- The *z*-scores for +3*σ* and –3*σ* are +3 and –3 respectively.

The empirical rule is also known as the 68-95-99.7 rule. See video



Example 6-3

The mean height of 15 to 18-year-old males from Chile from 2009 to 2010 was 170 cm with a standard deviation of 6.28 cm. Male heights are known to follow a normal distribution. Let *X* = the height of a 15 to 18-year-old male from Chile in 2009 to 2010. Then *X* ~ *N*(170, 6.28).
Suppose a 15 to 18-year-old male from Chile was 168 cm tall from 2009 to 2010. The *z*-score when *x* = 168 cm is *z* = ________. This *z*-score tells you that *x* = 168 is _________ standard deviations to the _________ (right or left) of the mean which is ______ .

Solution 6-3

–0.32, 0.32, left, 170

$$z = \frac{x - \mu}{\sigma} = \frac{168 - 170}{6.28} = -0.32$$

Try It 6-3

In 2012, 1,664,479 students took the SAT exam. The distribution of scores in the verbal section of the SAT had a mean μ = 496 and a standard deviation σ = 114. Let *X* = a SAT exam verbal section score in 2012. Then *X* ~ *N*(496, 114).

Find the *z*-scores for x_1 = 325 and x_2 = 366. Interpret each *z*-score. What can you say about x_1 = 325 and x_2 = 366?

Chapter 6 Try It Solutions

Example 6-4

Suppose *x* has a normal distribution with mean 50 and standard deviation 6.

a) About 68% of *x* values lie between what two values?

b) About 95% of *x* values lie between what two values?

c) About 99.7% of *x* values lie between what two values?

Solution 6-4

a) About 68% of the *x* values lie between $-1\sigma = (-1)(6) = -6$ and $1\sigma = (1)(6) = 6$ of the mean (50). The values $50 - 6 = 44$ and $50 + 6 = 56$ are within one standard deviation of the mean (50). The *z*-scores are −1 and +1 for 44 and 56, respectively.

b) About 95% of the *x* values lie between $-2\sigma = (-2)(6) = -12$ and $2\sigma = (2)(6) = 12$ of the mean (50). The values $50 - 12 = 38$ and $50 + 12 = 62$ are within two standard deviations of the mean (50). The *z*-scores are −2 and +2 for 38 and 62, respectively.

c) About 99.7% of the *x* values lie between $-3\sigma = (-3)(6) = -18$ and $3\sigma = (3)(6) = 18$ of the mean (50). The values $50 - 18 = 32$ and $50 + 18 = 68$ are within three standard deviations of the mean (50). The *z*-scores are −3 and +3 for 32 and 68, respectively.

Try It 6-4

Suppose *X* has a normal distribution with mean 25 and standard deviation five. Between what values of *x* do 68% of the values lie?

Chapter 6 Try It Solutions

Example 6-5

From 1984 to 1985, the mean height of 15 to 18-year-old males from Chile was 172.36 cm, and the standard deviation was 6.34 cm. Let *X* = the height of 15 to 18-year-old males in 1984 to 1985. Then $X \sim N(172.36, 6.34)$.

a) About 68% of the *x* values lie between _____________ and _____________. The *z*-scores are _______ and _______, respectively.

b) About 95% of the *x* values lie between _____________ and _____________. The *z*-scores are _______ and _______, respectively.

c) About 99.7% of the *x* values lie between ____________ and ____________. The *z*-scores are _______ and _______, respectively.

Solution 6-5

a) About 68% of the values lie between 166.02 and 178.7. The *z*-scores are −1 and 1.

b) About 95% of the values lie between 159.68 and 185.04. The *z*-scores are −2 and 2.

c) About 99.7% of the values lie between 153.34 and 191.38. The *z*-scores are −3 and 3.

Try It 6-5

The scores on a college entrance exam have an approximate normal distribution with mean, μ = 52 points and a standard deviation, σ = 11 points.

a) About 68% of the scores lie between ____________ and ____________. The *z*-scores are _______ and _______, respectively.

b) About 95% of the scores lie between ____________ and ____________. The *z*-scores are _______ and _______, respectively.

c) About 99.7% of the scores lie between ____________ and ____________. The *z*-scores are _______ and _______, respectively.

Chapter 6 Try It Solutions

Example 6-6

Suppose a normal distribution has a mean score of 15 and a standard deviation of 3. That is X ~ (15, 3).

a) Approximately what percent of the scores are Between 9 and 21?

b) Approximately what percent of the scores are less than 6?

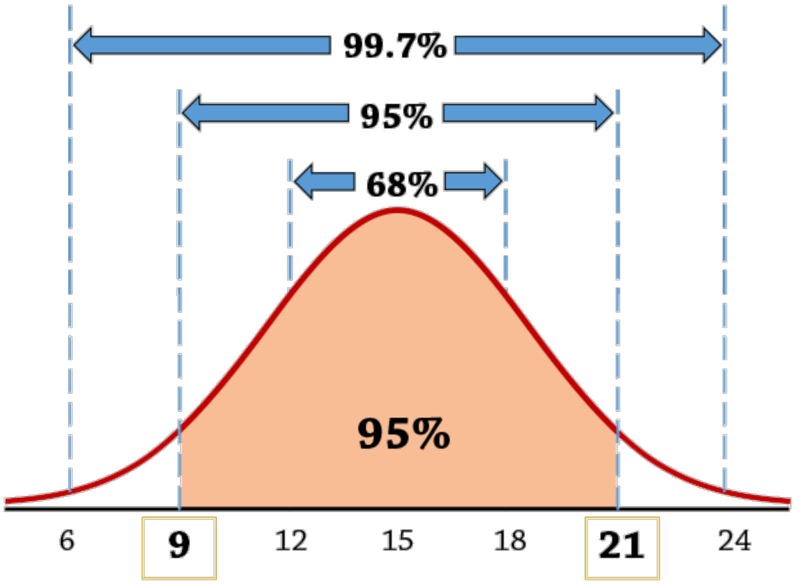
c) Approximately what percent of the scores are greater than 18?

d) Approximately what percent of the scores are below 21?

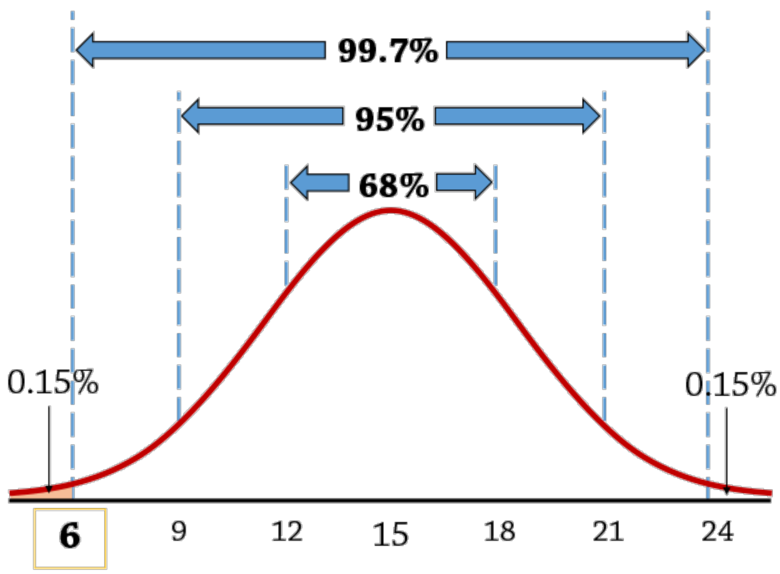
e) Approximately what percent of the scores are between 12 and 24?

Solution 6-6

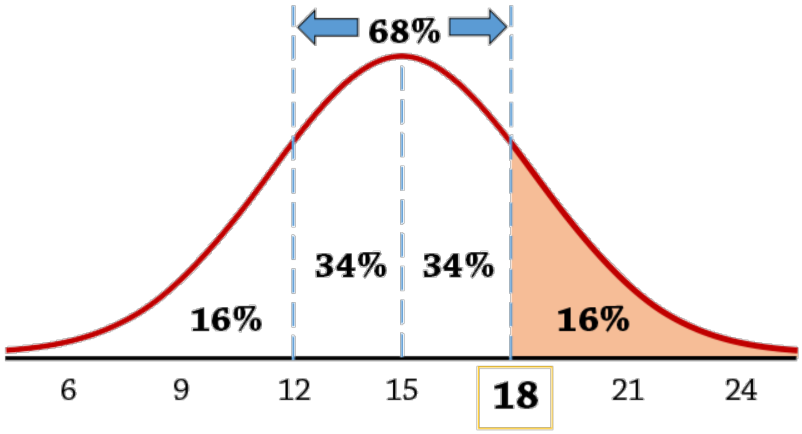
a) Approximately 95%



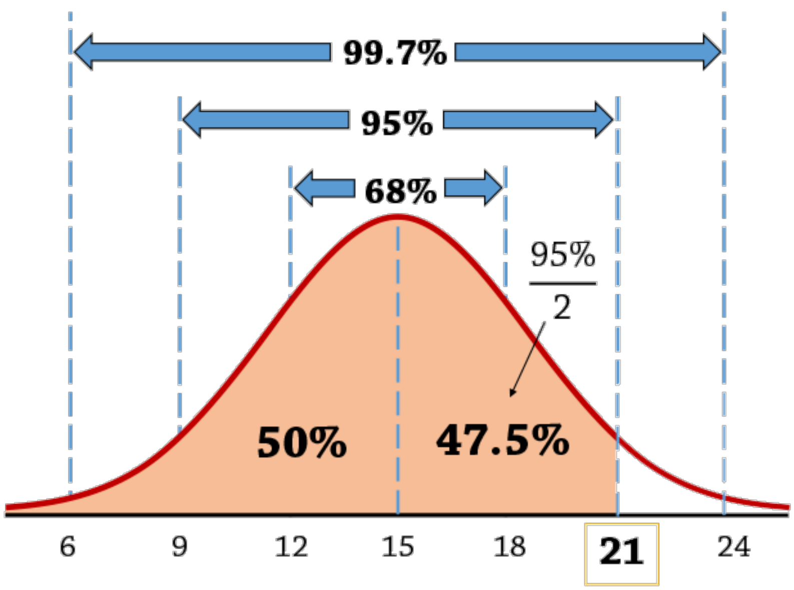
b) Approximately 0.15%



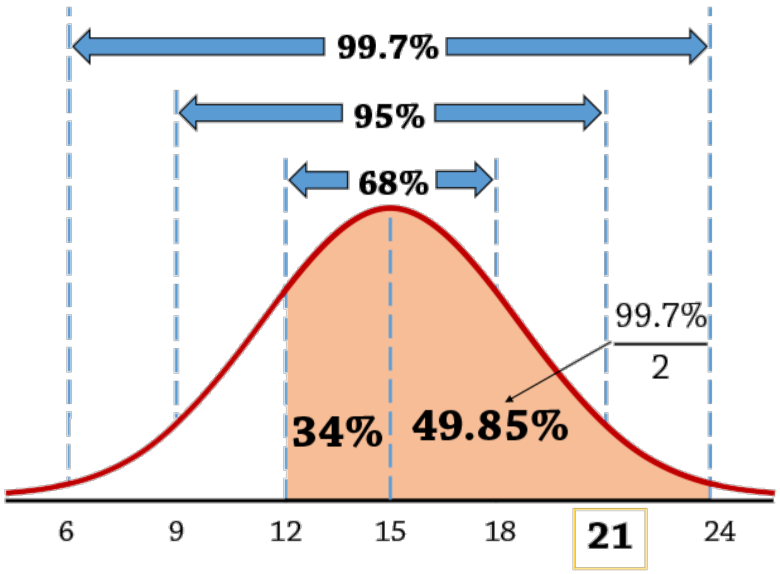
c) Approximately 16%



d) Approximately 97.5%



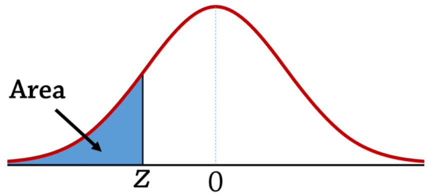
e) Approximately 83.85%



6.2 | Using the Normal Distribution

The Standard Normal Distribution Tables (shown below) provide the probability that *Z*, the Standard Normal Variable, is **less than** a certain value *z*. *z* values (values in the left column and on the top row) are points on the horizontal scale while areas or probabilities (values in the body of the table) are the regions bounded by the normal curve and the horizontal scale.

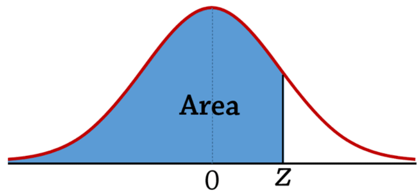
Table of Standard Normal Probabilities for Negative z-scores



z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
-3.4	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0002
-3.3	0.0005	0.0005	0.0005	0.0004	0.0004	0.0004	0.0004	0.0004	0.0004	0.0003
-3.2	0.0007	0.0007	0.0006	0.0006	0.0006	0.0006	0.0006	0.0005	0.0005	0.0005
-3.1	0.0010	0.0009	0.0009	0.0009	0.0008	0.0008	0.0008	0.0008	0.0007	0.0007
-3.0	0.0013	0.0013	0.0013	0.0012	0.0012	0.0011	0.0011	0.0011	0.0010	0.0010
-2.9	0.0019	0.0018	0.0018	0.0017	0.0016	0.0016	0.0015	0.0015	0.0014	0.0014
-2.8	0.0026	0.0025	0.0024	0.0023	0.0023	0.0022	0.0021	0.0021	0.0020	0.0019
-2.7	0.0035	0.0034	0.0033	0.0032	0.0031	0.0030	0.0029	0.0028	0.0027	0.0026
-2.6	0.0047	0.0045	0.0044	0.0043	0.0041	0.0040	0.0039	0.0038	0.0037	0.0036
-2.5	0.0062	0.0060	0.0059	0.0057	0.0055	0.0054	0.0052	0.0051	0.0049	0.0048
-2.4	0.0082	0.0080	0.0078	0.0075	0.0073	0.0071	0.0069	0.0068	0.0066	0.0064
-2.3	0.0107	0.0104	0.0102	0.0099	0.0096	0.0094	0.0091	0.0089	0.0087	0.0084
-2.2	0.0139	0.0136	0.0132	0.0129	0.0125	0.0122	0.0119	0.0116	0.0113	0.0110
-2.1	0.0179	0.0174	0.0170	0.0166	0.0162	0.0158	0.0154	0.0150	0.0146	0.0143
-2.0	0.0228	0.0222	0.0217	0.0212	0.0207	0.0202	0.0197	0.0192	0.0188	0.0183
-1.9	0.0287	0.0281	0.0274	0.0268	0.0262	0.0256	0.0250	0.0244	0.0239	0.0233
-1.8	0.0359	0.0351	0.0344	0.0336	0.0329	0.0322	0.0314	0.0307	0.0301	0.0294
-1.7	0.0446	0.0436	0.0427	0.0418	0.0409	0.0401	0.0392	0.0384	0.0375	0.0367
-1.6	0.0548	0.0537	0.0526	0.0516	0.0505	0.0495	0.0485	0.0475	0.0465	0.0455
-1.5	0.0668	0.0655	0.0643	0.0630	0.0618	0.0606	0.0594	0.0582	0.0571	0.0559
-1.4	0.0808	0.0793	0.0778	0.0764	0.0749	0.0735	0.0721	0.0708	0.0694	0.0681
-1.3	0.0968	0.0951	0.0934	0.0918	0.0901	0.0885	0.0869	0.0853	0.0838	0.0823
-1.2	0.1151	0.1131	0.1112	0.1093	0.1075	0.1056	0.1038	0.1020	0.1003	0.0985
-1.1	0.1357	0.1335	0.1314	0.1292	0.1271	0.1251	0.1230	0.1210	0.1190	0.1170
-1.0	0.1587	0.1562	0.1539	0.1515	0.1492	0.1469	0.1446	0.1423	0.1401	0.1379
-0.9	0.1841	0.1814	0.1788	0.1762	0.1736	0.1711	0.1685	0.1660	0.1635	0.1611
-0.8	0.2119	0.2090	0.2061	0.2033	0.2005	0.1977	0.1949	0.1922	0.1894	0.1867
-0.7	0.2420	0.2389	0.2358	0.2327	0.2296	0.2266	0.2236	0.2206	0.2177	0.2148
-0.6	0.2743	0.2709	0.2676	0.2643	0.2611	0.2578	0.2546	0.2514	0.2483	0.2451
-0.5	0.3085	0.3050	0.3015	0.2981	0.2946	0.2912	0.2877	0.2843	0.2810	0.2776
-0.4	0.3446	0.3409	0.3372	0.3336	0.3300	0.3264	0.3228	0.3192	0.3156	0.3121
-0.3	0.3821	0.3783	0.3745	0.3707	0.3669	0.3632	0.3594	0.3557	0.3520	0.3483
-0.2	0.4207	0.4168	0.4129	0.4090	0.4052	0.4013	0.3974	0.3936	0.3897	0.3859
-0.1	0.4602	0.4562	0.4522	0.4483	0.4443	0.4404	0.4364	0.4325	0.4286	0.4247
-0.0	0.5000	0.4960	0.4920	0.4880	0.4840	0.4801	0.4761	0.4721	0.4681	0.4641

Figure 6.3 Table of Standard Normal Probabilities for Negative z-scores. Table entries are "less than" areas.

Table of Standard Normal Probabilities for Positive z-scores



z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.7	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7852
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
1.0	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810	0.8830
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997	0.9015
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9279	0.9292	0.9306	0.9319
1.5	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394	0.9406	0.9418	0.9429	0.9441
1.6	0.9452	0.9463	0.9474	0.9484	0.9495	0.9505	0.9515	0.9525	0.9535	0.9545
1.7	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	0.9608	0.9616	0.9625	0.9633
1.8	0.9641	0.9649	0.9656	0.9664	0.9671	0.9678	0.9686	0.9693	0.9699	0.9706
1.9	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.9750	0.9756	0.9761	0.9767
2.0	0.9772	0.9778	0.9783	0.9788	0.9793	0.9798	0.9803	0.9808	0.9812	0.9817
2.1	0.9821	0.9826	0.9830	0.9834	0.9838	0.9842	0.9846	0.9850	0.9854	0.9857
2.2	0.9861	0.9864	0.9868	0.9871	0.9875	0.9878	0.9881	0.9884	0.9887	0.9890
2.3	0.9893	0.9896	0.9898	0.9901	0.9904	0.9906	0.9909	0.9911	0.9913	0.9916
2.4	0.9918	0.9920	0.9922	0.9925	0.9927	0.9929	0.9931	0.9932	0.9934	0.9936
2.5	0.9938	0.9940	0.9941	0.9943	0.9945	0.9946	0.9948	0.9949	0.9951	0.9952
2.6	0.9953	0.9955	0.9956	0.9957	0.9959	0.9960	0.9961	0.9962	0.9963	0.9964
2.7	0.9965	0.9966	0.9967	0.9968	0.9969	0.9970	0.9971	0.9972	0.9973	0.9974
2.8	0.9974	0.9975	0.9976	0.9977	0.9977	0.9978	0.9979	0.9979	0.9980	0.9981
2.9	0.9981	0.9982	0.9982	0.9983	0.9984	0.9984	0.9985	0.9985	0.9986	0.9986
3.0	0.9987	0.9987	0.9987	0.9988	0.9988	0.9989	0.9989	0.9989	0.9990	0.9990
3.1	0.9990	0.9991	0.9991	0.9991	0.9992	0.9992	0.9992	0.9992	0.9993	0.9993
3.2	0.9993	0.9993	0.9994	0.9994	0.9994	0.9994	0.9994	0.9995	0.9995	0.9995
3.3	0.9995	0.9995	0.9995	0.9996	0.9996	0.9996	0.9996	0.9996	0.9996	0.9997
3.4	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9998

Figure 6.4 Table of Standard Normal Probabilities for Positive z-scores. Table entries are "less than" areas.

The shaded area in the following graph indicates the area to the left of z_1. This area is represented by the probability $P(z < z_1)$.

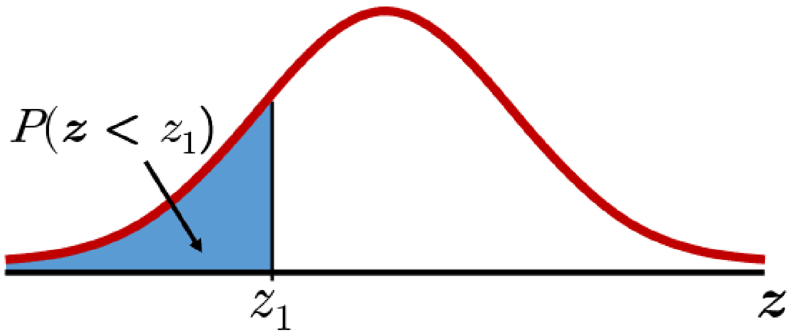


Figure 6.5

$P(z < z_1)$ = **Area to the left** of the vertical line through z_1.
$P(z > z_1) = 1 - P(z < z_1)$ = **Area to the right** of the vertical line through z_1.
$P(z < z_1)$ is the same as $P(z \leq z_1)$ and $P(z > z_1)$ is the same as $P(z \geq z_1)$ for continuous distributions.

To locate the area less than z = 1.43 on the table, for example, locate the row for 1.4 and the column for 0.03. The intersection of the row and column gives 0.9236. Hence $P(z < 1.43) = 0.9236$. Consequently, $P(z > 1.43) = 1 - P(z < 1.43) = 1 - 0.9236 = 0.0764$.

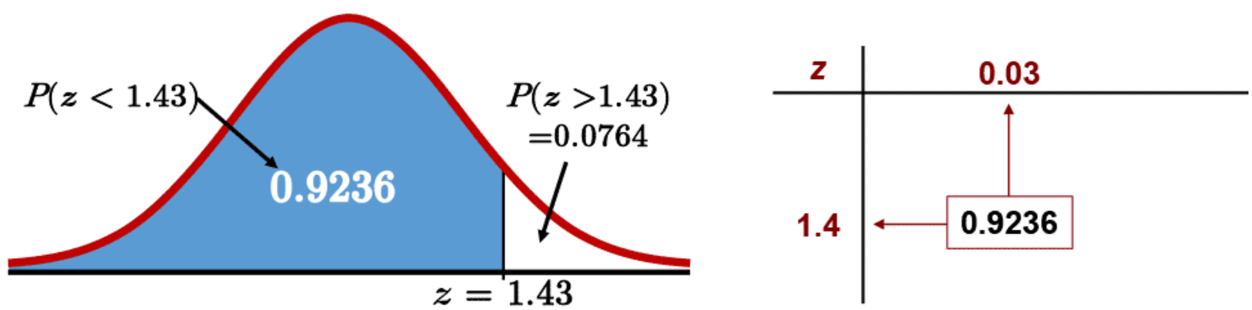


Figure 6.6

Calculation of Probabilities

Example 6-7

Use the standard normal tables to find the following areas/probabilities:

a) Area to the left of $z = 1.36$: $P(z<1.36)$

b) Area to the left of $z = -0.60$: $P(z<-0.60)$

c) Area to the right of $z = 1.47$: $P(z>1.47)$

d) Area to the right of *z* = -0.33: P(z>−0.33).

e) Area between *z* = -2.16 and *z* = 0.45: P(−2.16<z<−0.33)

f) Area between *z* = 1.13 and *z* = 3.10: P(1.13<z<3.10)

Solution 6-7

a) $P(z < 1.36) = 0.9131$

b) $P(z < -0.60) = 0.2743$

c) $P(z > 1.47) = 1 - P(z < 1.47) = 1 - 0.9292 = 0.0708$

d) $P(z > -0.33) = 1 - P(z < -0.33) = 1 - 0.3707 = 0.6293$

e) $P(-2.16 < z < -0.33) = P(z < 0.45) - P(z < -2.16) = 0.6736 - 0.0154 = 0.6582$

f) $P(1.13 < z < 3.10) = P(z < 3.10) - P(z < 1.13) = 0.9990 - 0.8708 = 0.1282$

Example 6-8

The final exam scores in a statistics class were normally distributed with a mean of 63 and a standard deviation of 5.

a) Find the probability that a randomly selected student scored less than 65 on the exam.

b) Find the probability that a randomly selected student scored more than 57 on the exam.

c) Find the probability that a randomly selected student scored between 55 and 70.

Solution 6-8

a) Let *X* = score on the final exam. *X* ~ *N*(63, 5) where μ = 63 and σ = 5. We begin by calculating the z-score for 65, using the formula

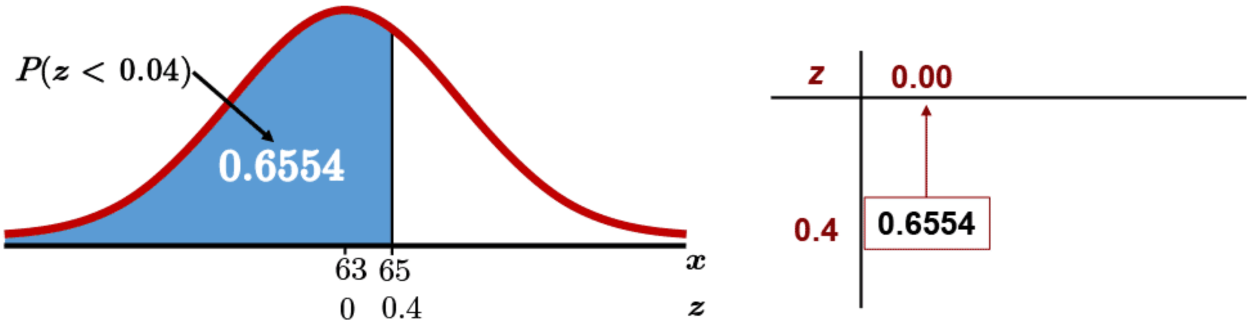
$$z = \frac{x - \mu}{\sigma}$$

For x = 65, the corresponding z-score is

$$z = \frac{65 - 63}{5} = 0.4$$

From the *z*-tables, the area corresponding to *z* = 0.40 is 0.6554. That is,

$$P(x < 65) = P(z < 0.40) = 0.6554$$



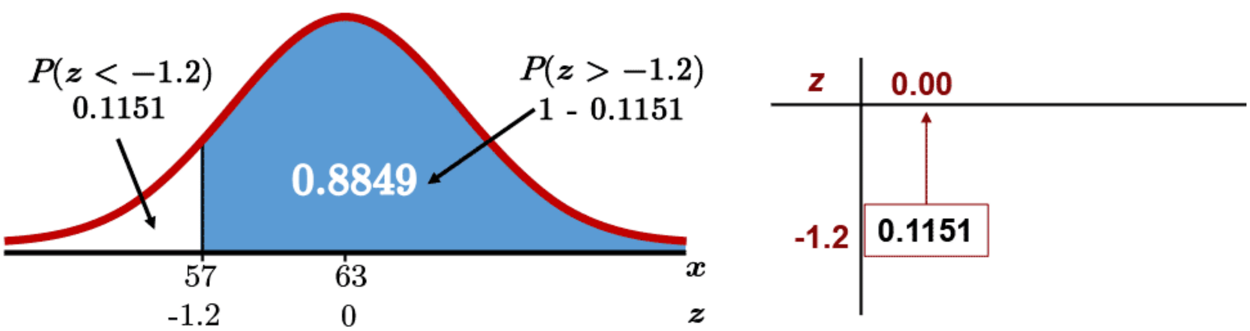
The probability that any student selected at random scores less than 65 is 0.6554.
Interpretation: About 65.54% of students who took the test scored below 65.

b) For x = 57 (μ = 63 and σ = 5), the corresponding z-score is

$$z = \frac{x - \mu}{\sigma} = \frac{57 - 63}{5} = -1.20$$

From the *z*-tables, the area below *z* = -1.20 is 0.1151. Therefore,

$$P(x > 57) = P(z > -1.20) = 1 - P(z < -1.20) = 1 - 0.1151 = 0.8849$$



The probability that any student selected at random scored more than 57 is 0.8849.
Interpretation: 88.49% of students scored above 57 on this test.

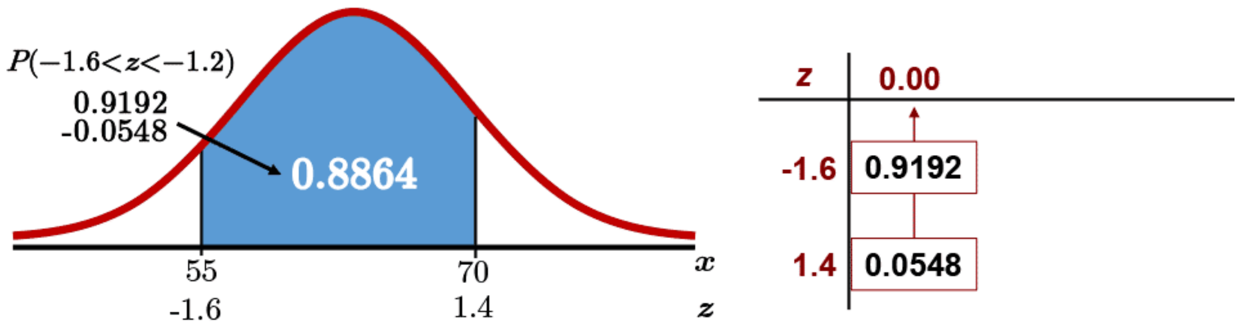
c) The z-scores for x = 55 and x = 70 are calculated as follows:

$$x = 55; \ z = \frac{x - \mu}{\sigma} = \frac{55 - 63}{5} = -1.60$$

$$x = 70; \ z = \frac{x - \mu}{\sigma} = \frac{70 - 63}{5} = 1.40$$

From the *z*-tables, the areas less than *z* = -1.60 and *z* = 1.40 are 0.0548 and 0.9192 respectively. Therefore,

$$P(55 < x < 70) = P(-1.60 < z < 1.40) = 0.9192 - 0.0548 = 0.8644$$



Interpretation: 86.44% of the students who took the exam scored between 50 and 70.

Try It 6-6

The golf scores for a school team were normally distributed with a mean of 68 and a standard deviation of three. Find the probability that a randomly selected golfer scored more than 63.

Chapter 6 Try It Solutions

Example 6-9

A personal computer is used for office work at home, research, communication, personal finances, education, entertainment, social networking, and a myriad of other things. Suppose that the average number of hours a household personal computer is used for entertainment is two hours per day. Assume the times for entertainment are normally distributed and the standard deviation for the times is half an hour.

a) Find the probability that a household personal computer is used for entertainment between 1.31 and 1.83 hours per day.

b) What percent of household personal computers are used for entertainment for less than 1.31 hours or for more than 1.83 hours per day.

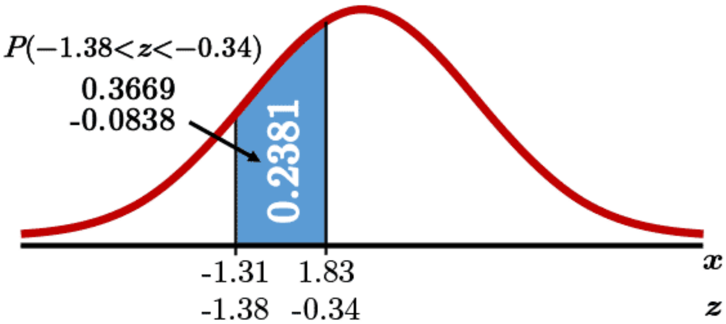
Solution 6-9

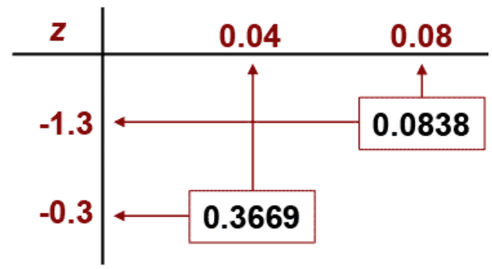
a) Let X = the amount of time (in hours) a household personal computer is used for entertainment. X ~ N(2, 0.5) where $\mu = 2$ and $\sigma = 0.5$. We need to find $P(1.31 < x < 1.83)$. The z-scores for x = 1.31 and x = 1.83 are calculated as follows:

$$x = 1.31; \; z = \frac{x-\mu}{\sigma} = \frac{1.31-2}{0.5} = -1.38$$

$$x = 1.83; \; z = \frac{x-\mu}{\sigma} = \frac{1.83-2}{0.5} = -0.34$$

$$P(1.31 < x < 1.83) = P(-1.38 < z < -0.34) = 0.3669 - 0.0838 = 0.2381$$

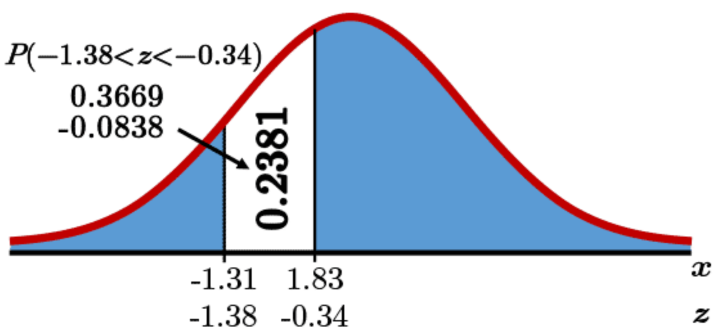


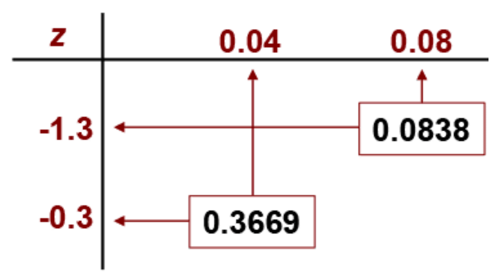


Interpretation: A household personal computer is used for entertainment between 1.31 and 1.83 hours per day 23.81% of the time.

b) The probability that a household personal computer is used for entertainment for "less than 1.31 hours or for more than 1.83 hours" per day is the **complement (or opposite)** of "between 1.31 and 1.83 hours". From part a), $P(1.31 < x < 1.83) = 0.2381$. Therefore,

$$P(x < 1.31 \; or \; x > 1.83) = 1 - P(1.31 < x < 1.83) = 1 - 0.2381 = 0.7619$$





Interpretation: 76.19% of household personal computers are used for entertainment for less than 1.31 hours or for more than 1.83 hours per day.

Try It 6-7

The golf scores for a school team were normally distributed with a mean of 68 and a standard deviation of three. Find the probability that a golfer scored between 64 and 70.

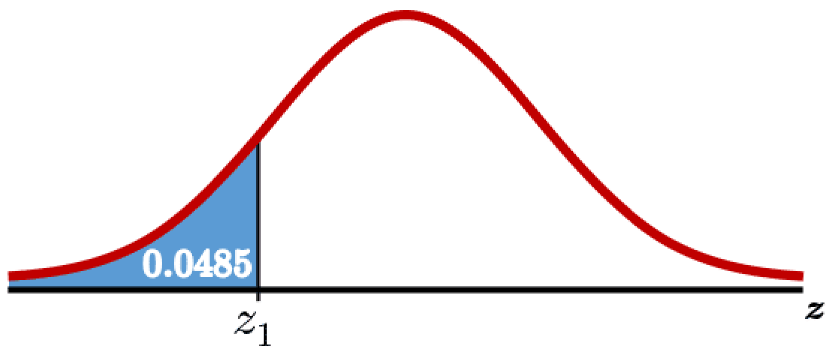
Chapter 6 Try It Solutions

Calculating Scores

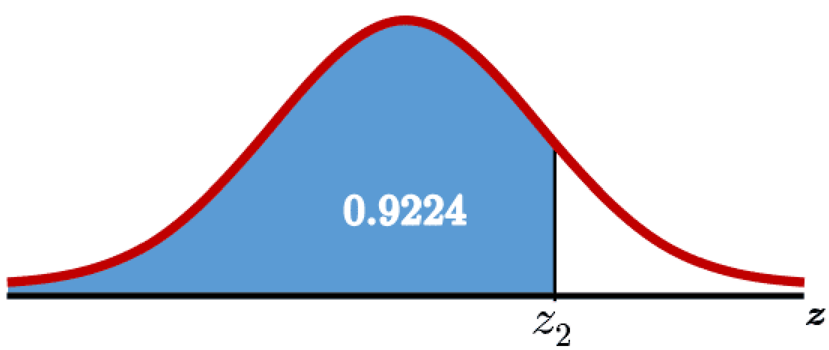
Example 6-10

We are sometimes faced with the problem of finding a score (z or x), rather than probability (or area). That is, we are provided with the probability (or area) in a region and are required to find the corresponding z-score and/or x value. This reverse lookup process is shown in the following examples.

a) Find the value of z_1 in the figure below. The area to the left of z_1 is 0.0485 or $P(z < z_1) = 0.0485$.



b) Find the value of z2 in the figure below.

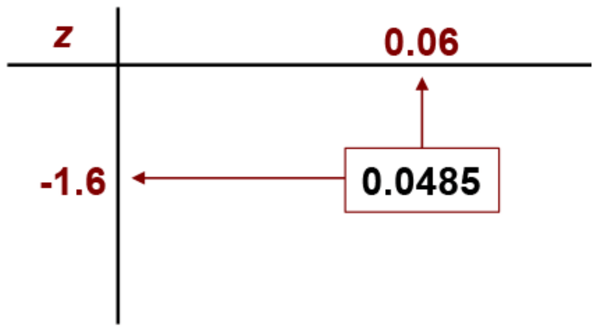


c) Find the z-score above which lies 20% of scores in the normal distribution.

d) Find b such that P(z > b) = 0. 9121

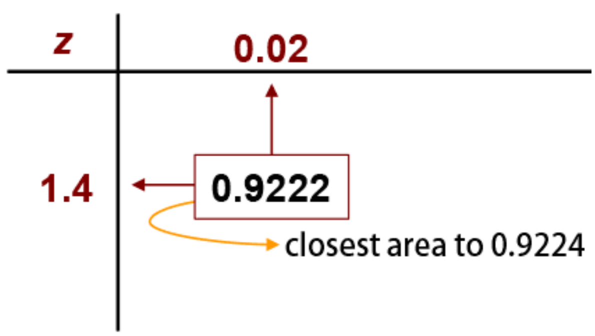
Solution 6-10

a) The area less than z_1 is 0.0485. Since this area is less than 0.5, it can be found on the negative z-score table. Looking 0.0485 up in the **area** section of the z-table, we see that the corresponding *z*-score is **-1.66**.



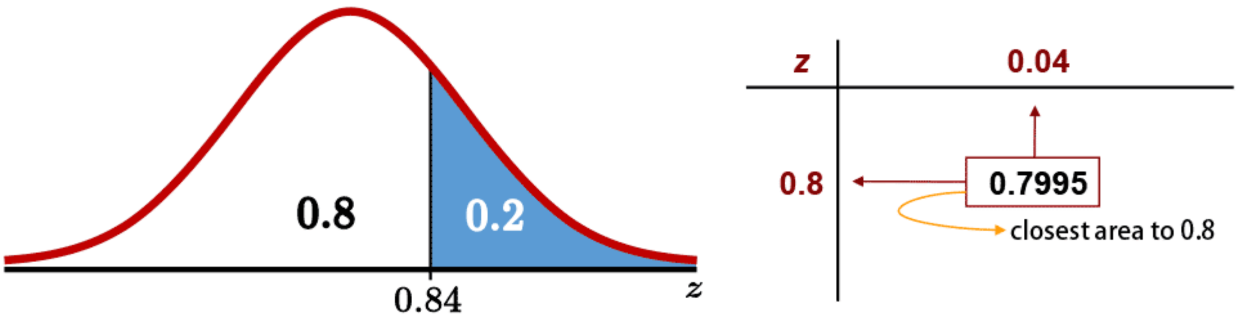
$$z_1 = -1.66; \quad 0.0485 = P(z < -1.66)$$

b) The area less than z_2 is 0.9224. Since this area is greater than 0.5, it can be found on the positive z-score table. Looking up 0.9224 in the **area** section of the z-table, we find that the closest value to 0.9224 is 0.9222, corresponding to a *z*-score of **1.42**.



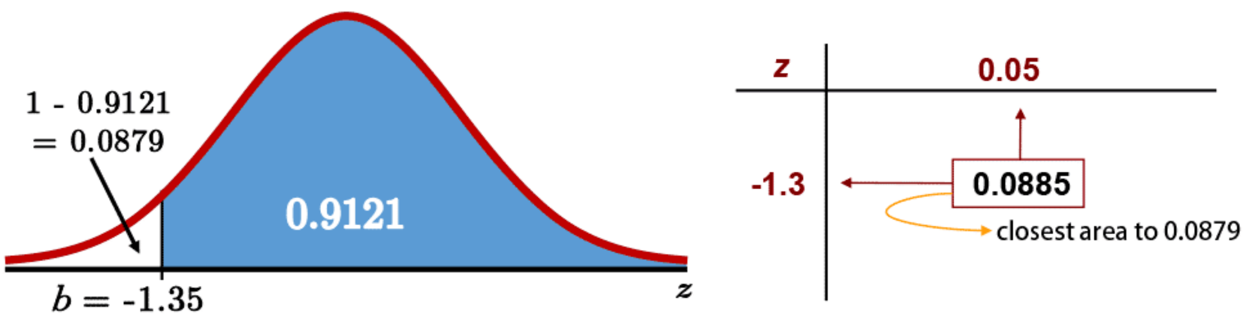
$$z_2 = 1.42; \quad 0.9224 \approx P(z < 1.42)$$

c) The right tail area is 20%, but our z-tables only contain left tail (less than) areas. We therefore subtract 20% from 100% to find the left tail area of 80% or 0.8. From the body of the z-table, we find the closest value to 0.8 to be 0.7995, corresponding to a z-score of **0.84**.



$$P(z > 0.84) \approx 0.2$$

d) P(z > *b*) = 0. 9121 implies that the area to the right of *b* is 0.9121. Since this area is on the right, and the normal tables we have only provide areas on the left, we subtract 0.9121 from 1, so that the "less than area" is 0.0879. The closest area in the table is 0.0885, corresponding to a *z*-score is -1.35.



$$b = -1.35; \quad P(z > -1.35) \approx 0.9121$$

Calculating x-values

Starting with the z-formula, we can solve for x as follows:

$$z = \frac{x - \mu}{\sigma} \quad \Rightarrow \quad z\sigma = x - \mu \quad \Rightarrow \quad \mu + z\sigma = x$$

$$\boldsymbol{x = \mu + z\sigma}$$

Example 6-11

Suppose the highway fuel consumption of cars sold in a city follows a normal distribution with a mean of 8.7 L/100km and a standard deviation of 2.5 L/100km.

a) What fuel consumption rate represents the third quartile?

b) Determine the fuel consumption rate above which 90% of the cars will fall.

c) What is the fuel consumption rate of the least efficient 20% of the cars?

d) What is the minimum and the maximum fuel consumption rate of the middle 95% of the cars?

Solution 6-11

a) Let X = car fuel consumption. $X \sim N(8.7, 2.5)$ where $\mu = 8.7$ and $\sigma = 2.5$. The area below the 3rd quartile (75th percentile) is 0.75. Since we already have μ and σ, we simply need to find the z-score corresponding to the lower area of 0.75 in the table. The closest area to 0.75 in the standard normal table is 0.7486 and this corresponds to $z = 0.67$. Therefore,

$$z = \frac{x - \mu}{\sigma} \quad becomes \quad 0.67 = \frac{x - 8.7}{2.5}$$

$$0.67(2.5) = x - 8.7 \quad \Rightarrow \quad x = 8.7 + 0.67(2.5) = 10.375$$

Or we can simply use the x transformation formula obtained earlier:

$$x = \mu + z\sigma = 8.7 + 0.67(2.5) = 10.375$$

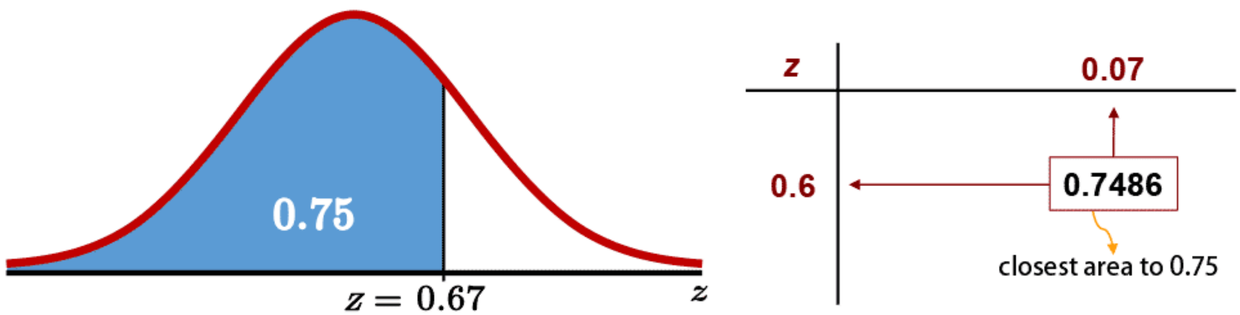


Figure 6.7

The fuel consumption rate representing the third quartile is 10.375 L/100km.

b) If the fuel consumption rate has 90% higher than it, then it has 10% lower. We then search for the closest z-score to 0.1. The closest area to 0.1 is 0.1003, corresponding to a z-score of -1.28. Therefore,

$$x = \mu + z\sigma = 8.7 - 1.28(2.5) = 5.5$$

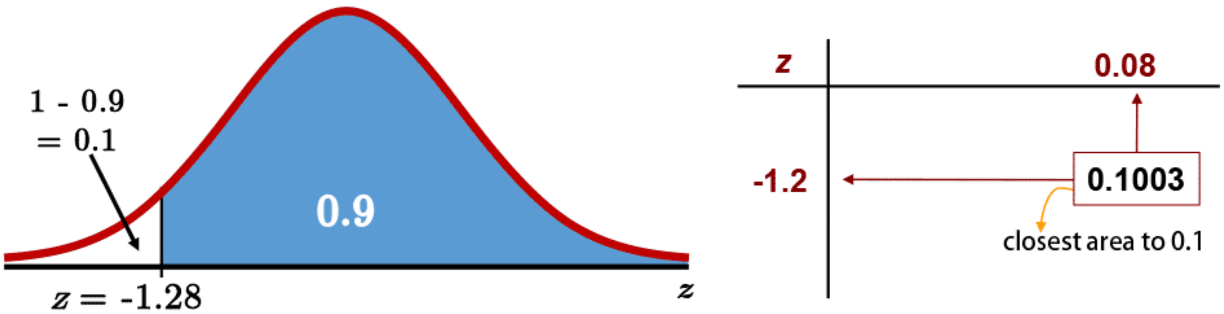


Figure 6.8

The fuel consumption rate above which 90% of the cars will fall is 5.5L/100km. This is the 10th percentile rate. That is, only 10% of the cars have consumption rate of 5.5L/km or better.

c) The least efficient 20% of the cars will have the worst fuel consumption rates, and thus the highest 20% L/100km values. As a result, the required rate will separate the top 20% of the rates from bottom 80%. Searching for the closest area to 0.8 we find 0.7995 which corresponds to a z-score of 0.84. The required consumption rate is

$$x = \mu + z\sigma = 8.7 + 0.84(2.5) = 10.8$$

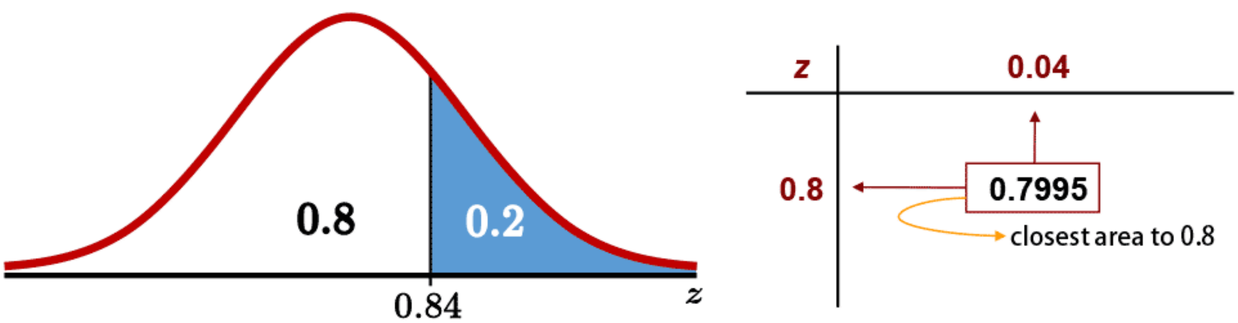


Figure 6.9

The fuel consumption rate of the least efficient 20% of the cars is 10.8 L/100km.

d) For the middle 95% of the rates, we have 5% of the rates divided equally into the two tails (2.5% each). See Figure 6.10 below. We will refer to the minimum and maximum z-scores as z_1 and z_2 respectively, and their corresponding consumption rates as x_1 and x_2.

The area to the left of z_1 is 0.025. So by reverse lookup, we see that

z_1 = -1.96.
The area to the left of z_2 is 0.975 (0.95+0.025). And by reverse lookup, we see that z_2 = 1.96 (It is expected that z_1 and z_2 will only differ in sign because the standard normal distribution is symmetric about 0).

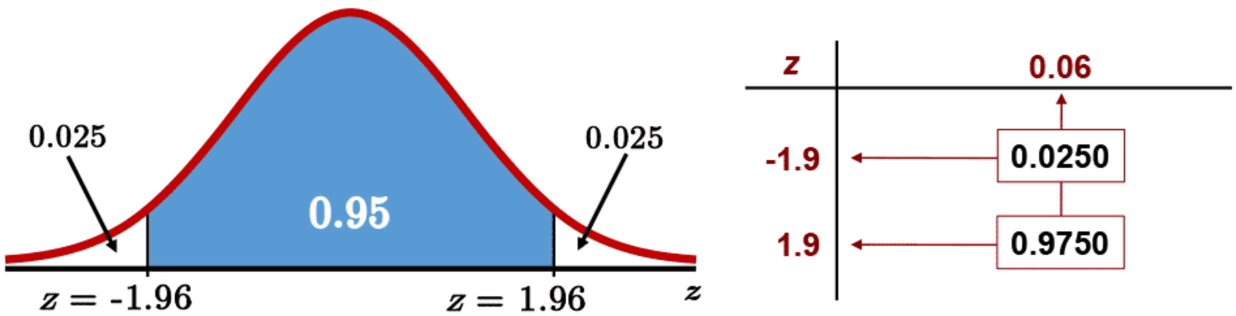


Figure 6.10

$$Minimum:\ x_1 = \mu + z_1\sigma = 8.7 - 1.96(2.5) = 3.8$$

$$Maximum:\ x_2 = \mu + z_2\sigma = 8.7 + 1.96(2.5) = 13.6$$

The consumption rates of the middle 95% of the cars are between 3.8 L/100km and 13.6 L/100km. That is, only 2.5% of the cars have fuel consumption rates less than 3.8 L/100km and only 2.5% have rates above 13.6 L/100km.

Try It 6-8

The scores on an exam have an approximate normal distribution with a mean μ = 81 points and standard deviation σ = 15 points.

a) Calculate the minimum score of the top 15% students on this exam.

b) Calculate the first- and third-quartile scores for this exam.

c) The middle 50% of the exam scores are between what two values?

Chapter 6 Try It Solutions

Using Excel for Normal Probabilities

Excel can be used to obtain probabilities and z-values by using the following built-in functions.

1. `NORM.DIST(x, mean, standard_dev, cumulative)` or `NORM.DIST(x, μ, σ, cumulative)`

This function is used to calculate the probability of obtaining values less than x, given μ and σ. We will always set cumulative to **TRUE** to obtain cumulative probabilities.

Suppose X ~ N(8, 1.5) where μ = 8 and σ = 1.5.
P(x < 6.5) can be found using `=NORM.DIST(6.5, 8, 1.5, TRUE)` which will result in 0.1587.
P(x > 5.6) can also be found using `=1-NORM.DIST(5.6, 8, 1.5, TRUE)` which gives 0.9452.

Results obtained using Excel may be slightly different from those from tables because of rounding.

2. **`NORM.S.DIST(z, cumulative)`**
This function is used to find the probability of obtaining values less than a standard normal value z. Again, `cumulative = TRUE`.

We can find the probability that z is less than 0.54, and that z is greater than 1.28 respectively as follows:
P(z < 0.54) `=NORM.S.DIST(0.54, TRUE)` which gives 0.7054.
P(z > 1.28) `=1-NORM.S.DIST(1.28, TRUE)` which equals 0.1003.

3. **`NORM.INV(probability, mean, standard_dev)`**
This function is used to obtain the value x, given the "less than" probability, μ, and σ.

Suppose X ~ N(8, 1.5) where μ = 8 and σ = 1.5.
Then the score corresponding to the 10th percentile can be obtained by `=NORM.INV(0.1, 8, 1.5)` = 6.08.
The minimum score of the top 30% of the scores can also be obtained by `=NORM.INV(1-0.3, 8, 1.5)` = 8.79.

4. **`NORM.S.INV(probability)`**
This function is used to obtain a standard normal score z corresponding to a specified "less than" probability.

The z-score corresponding to the 37th percentile can be obtained by `=NORM.S.INV(0.37)` = -0.33
The z-score that has 7.5% of the area above it can be obtained by `=NORM.S.INV(1-0.075)` = 1.44

Chapter 6 Try It Solutions

Try It 6-1

$z = \frac{7.5-12}{3} = -1.5$

Try It 6-2

-1.5, 1.5, left, 16

$z = \frac{x-\mu}{\sigma} = \frac{10-16}{4} = -1.5$

Try It 6-3

The z-score for $\boldsymbol{x}_1$ = 325 is $\boldsymbol{z}_1$ = −1.50.

The z-score for $\boldsymbol{x}_2$ = 366 is $\boldsymbol{z}_2$ = −1.14.

Student 2 scored closer to the mean than Student 1. Although they both had negative z-scores, Student 2 had the better score.

Try It 6-4

between 20 and 30.

Try It 6-5

a) About 68% of the values lie between the values 41 and 63. The z-scores are −1 and 1, respectively.

b) About 95% of the values lie between the values 30 and 74. The z-scores are −2 and 2, respectively.

c) About 99.7% of the values lie between the values 19 and 85. The z-scores are −3 and 3, respectively.

Try It 6-6

For x = 63 (μ = 68 and σ = 3), the corresponding z-score is

$$z = \frac{x-\mu}{\sigma} = \frac{63-68}{3} = -1.67$$

From the z-tables, the area below z = 1.67 is 0.9525. Therefore,

$$P(x > 63) = P(z > -1.67) = 1 - P(z < -1.67) = 1 - 0.0475 = 0.9525$$

The probability that a randomly selected golfer scored more than 63 is 0.9525.

Try It 6-7

For x = 64 and x = 70 (μ = 68 and σ = 3), the corresponding z-scores are

$$x = 64; \; z = \frac{x - \mu}{\sigma} = \frac{64 - 68}{3} = -1.33$$

$$x = 70; \; z = \frac{x - \mu}{\sigma} = \frac{70 - 68}{3} = 1.67$$

$$P(64 < x < 70) = P(-1.33 < z < 1.67) = 0.7486 - 0.0918 = 0.6568$$

The probability that a randomly selected golfer scored between 64 and 70 is 0.6568.

Try It 6-8

a) The minimum of the top 15% scores has 85% of the scores below it. The area to closest to 0.85 from the normal tables is 0.8508, corresponding to a z-score of 1.04. As a result, the minimum top 15% score is

$$x = \mu + z\sigma = 81 + 1.04(15) = 96.6$$

b) Since 25% of the scores are less than the first quartile, Q_1, the area to the left of Q_1 is 0.25 and that corresponds to a z-score of -0.67 in the tables. Similarly, 25% of the scores are greater than the third quartile, Q_3, so the z-score will be positive 0.67 (due to symmetry). The first and third quartiles therefore calculated as follows.

$$Q_1: x = \mu + z\sigma = 81 - 0.67(15) = 70.95$$

$$Q_3: x = \mu + z\sigma = 81 + 0.67(15) = 91.05$$

c) Since the middle 50% of all scores are between Q_1 and Q_3, the middle 50% of the scores are between 70.95 and 91.05.

CHAPTER REVIEW

6.1 | The Standard Normal Distribution

A z-score is a standardized value. Its distribution is the standard normal, $Z \sim N(0, 1)$. The mean of the z-scores is zero and the standard deviation is one. If z is the z-score for a value x from the normal distribution $N(\mu, \sigma)$ then z tells you how many standard deviations x is above (greater than) or below (less than) μ.

6.2 | Using the Normal Distribution

The normal distribution, which is continuous, is the most important of all the probability distributions. Its graph is bell-shaped. This bell-shaped curve is used in almost all disciplines. Since it is a continuous distribution, the total area under the curve is one. The parameters of the normal are the mean μ and the standard deviation σ. A special normal distribution, called the standard normal distribution is the distribution of z-scores. Its mean is zero, and its standard deviation is one.

REFERENCES

6.1 | The Standard Normal Distribution

"Blood Pressure of Males and Females." StatCruch, 2013. Available online at http://www.statcrunch.com/5.0/viewreport.php?reportid=11960 (accessed Sep 2, 2016).

"The Use of Epidemiological Tools in Conflict-affected populations: Open-access educational resources for policy-makers: Calculation of z-scores." London School of Hygiene and Tropical Medicine, 2009. Available online at http://conflict.lshtm.ac.uk/page_125.htm (accessed Sep 2, 2016).

"2012 College-Bound Seniors Total Group Profile Report." CollegeBoard, 2012. Available online at http://media.collegeboard.com/digitalServices/pdf/research/TotalGroup-2012.pdf (accessed Sep 2, 2016).

"Digest of Education Statistics: ACT score average and standard deviations by sex and race/ethnicity and percentage of ACT test takers, by selected composite score ranges and planned fields of study: Selected years, 1995 through 2009." National Center for Education Statistics. Available online at http://nces.ed.gov/programs/digest/d09/tables/dt09_147.asp (accessed Sep 2, 2016).

"List of stadiums by capacity." Wikipedia. Available online at https://en.wikipedia.org/wiki/List_of_stadiums_by_capacity (accessed Sep 2, 2016).

Data from the National Basketball Association. Available online at www.nba.com (accessed May 14, 2013).

6.2 | Using the Normal Distribution

"Naegele's rule." Wikipedia. Available online at http://en.wikipedia.org/wiki/Naegele's_rule (accessed June 23, 2016).

"403: NUMMI." Chicago Public Media & Ira Glass, 2013. Available online at http://www.thisamericanlife.org/radio-archives/episode/403/nummi (accessed June 23, 2016).

"Facebook Statistics." Statistics Brain. Available online at http://www.statisticbrain.com/facebook-statistics/(June 23, 2016).