4.3 Areas under a Normal Curve

Like the density curve in Section 3.4, we can use the normal curve to approximate areas (probabilities) between different values of $Y$ that follow a normal distribution $Y \sim N(\mu, \sigma)$

What we have to do is standardize $Y$. We do this by using a transformation which we discussed in Section 2.7

**Standardized Normally Distributed Variable $Z$**
The standardized version of a normally distributed variable $Y$,

$$Z = \frac{y - \mu}{\sigma}$$

Has the standard normal distribution.
Standard Normal distribution: Standard Normal Curve
A normally distributed variable having mean 0 and standard deviation 1 is said to have the standard normal distribution. Its associated normal curve is called the standard normal curve.

Basic Properties of the Standard Normal Curve

Property 1: The total area under the standard normal curve is 1.

Property 2: The standard normal curve extends indefinitely in both directions, approaching but never touching, the horizontal axis as it does so.

Property 3: The standard normal curve is symmetric about 0; that is, the part of the curve to the left of the dashed line is the mirror image of the part of the curve to the right of it.

Property 4: Almost all the area under the standard normal curve lies between -3 and 3.

For a normally distributed variable, we can find the percentage or the probability of all possible observations that lie within any specified range.
Procedure 1: Expressing the range in terms of z-scores and finding the corresponding area.

Step 1: Draw a standard normal curve.
Step 2: Label the z-score(s) on the curve.
Step 3: Shade in the region of interest.
Step 4: Determining the corresponding area under the standard normal curve using Table III.

Example 1

1. Find the area to the left of specified z-scores
   a. -0.87
   b. 2.56
   c. 5.12

2. Find the area to the right of specified z-scores
   a. 2.02
   b. -0.56
   c. -4

3. Determine the area under the standard normal curve that lies between
   a. -0.88 and 2.24

4. Find the area under the standard normal curve that lies
   a. Either to the left of -1 or to the right of 2
The Zα Notation

The symbol $Z_\alpha$ is used to denote the z-score that has an area of $\alpha$ (alpha) to its right under the standard normal curve. It is read as “$Z_\alpha$” as “z sub $\alpha$” or more simply as “$z \alpha$”

Example 2 from Elementary Statistics by Neil Weiss

6.18 Green Sea Urchins. From the paper “Effects of Chronic Nitrate Exposure on Gonad Growth in Green Sea Urchin Strongylocentrotus Droebachiensis” (Aquaculture, Vol. 242, No. 1–4, pp. 357–363) by S. Siikavuopio et al., we found that weights of adult green sea urchins are normally distributed with mean 52.0 g and standard deviation 17.2 g. Let $x$ denote weight of adult green sea urchins.

a. Sketch the distribution of the variable $x$.
b. Obtain the standardized version, $z$, of $x$.
c. Identify and sketch the distribution of $z$.
d. The percentage of adult green sea urchins with weights between 50 g and 60 g is equal to the area under the standard normal curve between _____ and _____.

e. The percentage of adult green sea urchins with weights above 40 g is equal to the area under the standard normal curve that lies to the _____ of _____.

Example 3

Obtain the following z-scores

a. $z_{0.20}$
b. $z_{0.05}$
Procedure 2: To Determine a Percentage or Probability for a Normally Distributed Variable

**Step 1:** Sketch the normal curve associated with the variable

**Step 2:** Shade the region of interest and mark its delimiting $y$-value(s).

**Step 3:** Find the $z$-score(s) for the delimiting $y$-value(s) found in Step 2.

**Step 4:** Use Table III to find the area under the standard normal curve delimited by the $z$-score(s) found in Step 3.

---

**Example 4**

A variable is normally distributed with mean 68 and standard deviation 10. Find the percentage of all possible values of the variable that

- d. lie between 73 and 80
- e. are at least 75
- f. are at most 90
The 68.26-95.44-99.74 Rule (Empirical Rule)

Any normally distributed variable has the following properties

Property 1: 68.26% of all possible observations lie within one standard deviation to either side of the mean, that is, between \( \mu - \sigma \) and \( \mu + \sigma \)

Property 2: 95.44% of all possible observations lie within two standard deviation to either side of the mean, that is, between \( \mu - 2\sigma \) and \( \mu + 2\sigma \)

Property 3: 99.74% of all possible observations lie within three standard deviations to either side of the mean, that is, between \( \mu - 3\sigma \) and \( \mu + 3\sigma \)

Procedure 3: To Determine the Observations Corresponding to a Specified Percentage or Probability for a Normally Distributed Variable

Step 1: Sketch the normal curve associated with the variable

Step 2: Shade the region of interest.

Step 3: Use Table III to determine the \( z \)-score(s) for the delimiting region found in Step 2.

Step 4: Find the \( y \)-value(s) having the \( z \)-scores(s) found in Step 3.
Example 5
A variable is normally distributed with mean 68 and standard deviation 10.
   a. Determine and interpret the quartiles of the variable
   b. Obtain and interpret the 99th percentile
   c. Find the value that 85% of all possible values of the variable exceed.
   d. Find the two values that divide the area under the corresponding normal curve into a middle area of 0.90 and two outside areas of 0.05. Interpret your answer.

Example 6
According to the *National Health and Nutritional Examination Survey*, published by the national Center for Health Statistics, the serum (noncellular portion of blood) total cholesterol level of U.S. females 20 years old or older is normally distributed with a mean of 206 mg/dL (milligrams per deciliter) and a standard deviation of 44.7 mg/dL.
   a. Determine the percentage of U.S. females 20 years old or older who have a serum total cholesterol level between 150 mg/dL and 250 mg/dL.
   b. Determine the percentage of U.S. females 20 years old or older who have a serum total cholesterol level below 220 mg/dL.
   c. Pr(Y < 170)
   d. Pr(Y > 230)
   e. Pr(140 < Y < 240)
   f. Obtain and interpret the quartiles for serum total cholesterol level of U.S. females 20 years old or older.
   g. Find and interpret the 90th percentile for serum total cholesterol level of U.S. females 20 years old or older.