6.3: Using R to Model Properties of a Normal Distribution

Given a mean and a standard deviation, we can use R's `dnorm()` function to plot the corresponding normal distribution

```
dnorm(x, mean, sd)
```

where `mean` is the value for μ, `sd` is the value for σ, and `x` is a vector of values that spans the range of x-axis values we want to plot.

```
# define the mean and the standard deviation
    mu = 12
    sigma = 2
# create vector for values of x that span a sufficient range of
# standard deviations on either side of the mean; here we use values
# for x that are four standard deviations on either side of the mean
    x = seq(4, 20, 0.01)
# use dnorm() to calculate probabilities for each x
    y = dnorm(x, mean = mu, sd = sigma)
# plot normal distribution curve
    plot(x, y, type = "l", lwd = 2, col = "blue", ylab = "probability", xlab = "x")
```

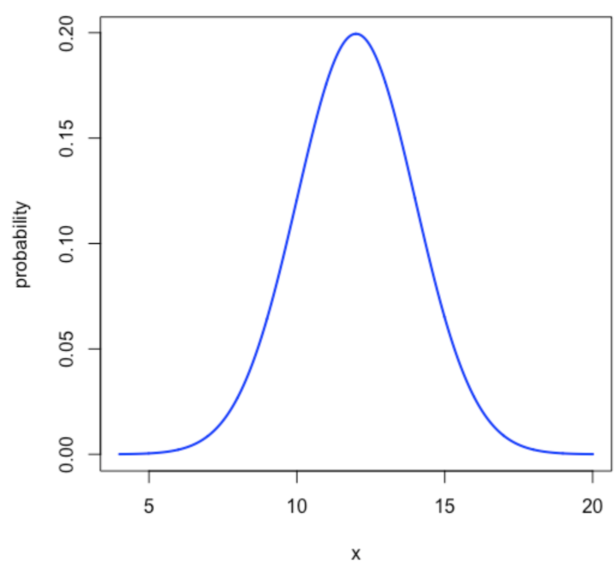

Figure 6.3.1: Plot showing the normal distribution curve for a population with $\mu = 12$ and $\sigma = 2$.

To annotate the normal distribution curve to show an area of interest to us, we use R's `polygon()` function, as illustrated here for the normal distribution curve in Figure $6.3.1$, showing the area that includes values between 8 and 15.

```
# define the mean and the standard deviation
    mu = 12
    sigma = 2
# create vector for values of x that span a sufficient range of
# standard deviations on either side of the mean; here we use values
```

```
# for x that are four standard deviations on either side of the mean
    x = seq(4, 20, 0.01)
# use dnorm() to calculate probabilities for each x
    y = dnorm(x, mean = mu, sd = sigma)
# plot normal distribution curve; the options xaxt = "i" and yaxt = "i"
# force the axes to begin and end at the limits of the data
    plot(x, y, type = "l", lwd = 2, col = "ivory4", ylab = "probability", xlab = "x",
    xaxs = "i", yaxs = "i")
# create vector for values of x between a lower limit of 8 and an upper limit of 15
lowlim = 8
uplim = 15
    dx = seq(lowlim, uplim, 0.01)
# use polygon to fill in area; x and y are vectors of x,y coordinates
# that define the shape that is then filled using the desired color
    polygon(x = c(lowlim, dx, uplim), y = c(0, dnorm(dx, mean = 12, sd = 2), 0),
    border = NA, col = "ivory4")
```

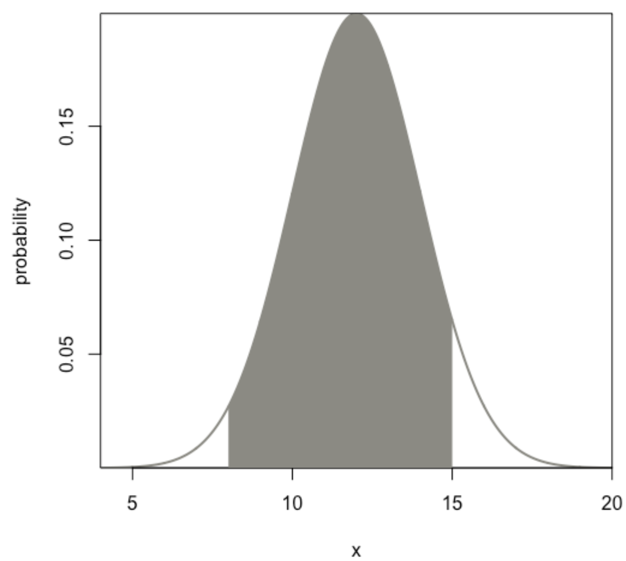

Figure 6.3.2: Plot showing the normal distribution curve for a population with $\mu = 12$ and $\sigma = 2$, and highlighting probability of obtaining a result between 8 and 15.

To find the probability of obtaining a value within the shaded are, we use R's `pnorm()` command

```
pnorm(q, mean, sd, lower.tail)
```

where `q` is a limit of interest, `mean` is the value for μ, `sd` is the value for σ, and `lower.tail` is a logical value that indicates whether we return the probability for values below the limit (`lower.tail = TRUE`) or for values above the limit (`lower.tail = FALSE`). For example, to find the probability of obtaining a result between 8 and 15, given $\mu = 12$ and $\sigma = 2$, we use the following lines of code.

```
# find probability of obtaining a result greater than 15
    prob_greater15 = pnorm(15, mean = 12, sd = 2, lower.tail = FALSE)
```

```
# find probability of obtaining a result less than 8
    prob_less8 = pnorm(8, mean = 12, sd = 2, lower.tail = TRUE)
# find probability of obtaining a result between 8 and 15
    prob_between = 1 - prob_greater15 - prob_less8 # display results
prob_greater15
    [1] 0.0668072
prob_less8
    [1] 0.02275013
prob_between
    [1] 0.9104427
```

Thus, 91.04% of values fall between the limits of 8 and 15.