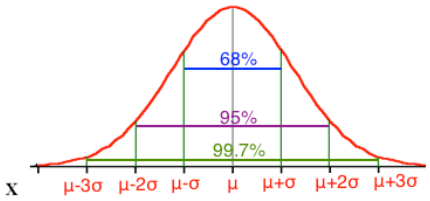
Section 4.2
The Normal Distribution

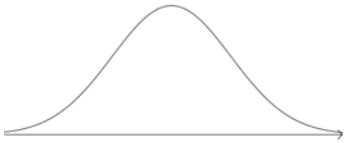
A density curve that is symmetric, single peaked and bell shaped is called a **normal distribution**.

The Empirical Rule

The Empirical Rule states if a distribution has a normal distribution,
1. Approximately 68% of all observations fall within one standard deviation of the mean.
2. Approximately 95% of all observations fall within two standard deviations of the mean.
3. Approximately 99.7% of all observations fall within three standard deviations of the mean.



Example: The length of time needed to complete a certain test is normally distributed with mean 60 minutes and standard deviation 10 minutes. First sketch the distribution.



a. What is the probability that someone will take between 40 and 80 minutes to complete the test?

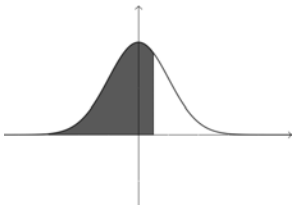
b. Find the interval that contains the middle 68% of completion times for all people taking the test.

c. What percent of people take more than 80 minutes to complete the test?

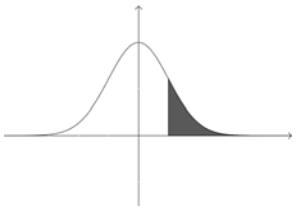
What if our values are not exactly within one, two or three standard deviations from the mean? Probabilities for these can still be found a number of ways, one of which we will explore in the next section.

Using R, the probability $P(X < x)$ can be found with the command pnorm(x, μ, σ).
*Note: The command in R only gives the probability that **X is less than a given value**. If we need to find the probability that X is greater than the given value, we will need to subtract the answer from 1. See the pictures below.*

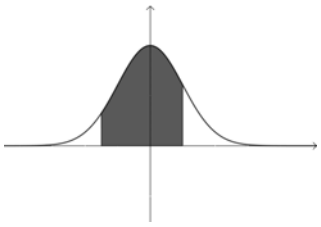
R will only give us area to the left of a value, i.e. $P(Z < z)$. Command: pnorm(z, μ, σ)



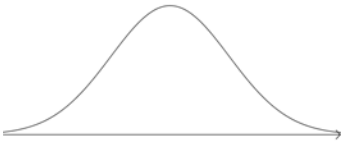
To calculate area to the right of a number, $P(Z > z)$, command: 1 – pnorm(z, μ, σ)



To calculate area between two numbers, $P(a < Z < b)$, command:
pnorm(a, μ, σ) – pnorm(b, μ, σ)

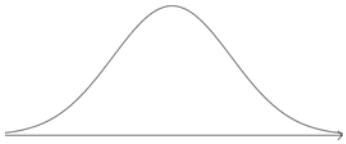


Continuing the example,
d. What is the probability that someone will take less than 45 minutes to complete the test?



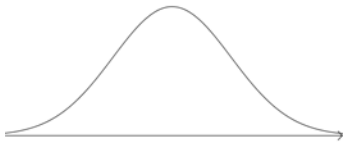
Command: Answer:

e. What is the probability that someone will take more than 30 minutes to complete the test?



Command: Answer:

f. What is the probability that someone will take between 30 and 45 minutes to complete the test?

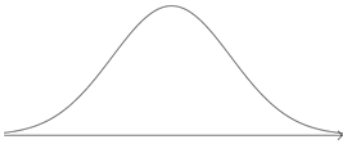


Command: Answer:

Now let's say that we know a probability and want to find the *x* value such that:

- $P(X < c) = p$, command: qnorm(p, μ, σ)
- $P(X > c) = p$, command: qnorm($1 - p, \mu, \sigma$)
- $P(-c < X < c) = p$, command: qnorm($(p+1)/2, \mu, \sigma$)

g. How long would it take someone to finish the test if they are in the top10% of the times? Sketch first the area:



Command: Answer: