8.2 Properties of the Normal Distribution
Most continuous distributions are distributed symmetrically and unimodally about the mean, forming a distinctive bell-shape.

This bell-shaped distribution is the 'normal' distribution.
This symmetry makes both the \textit{mode and median} equal to the mean. This allows one to use the mean, $\mu$, and the standard deviation, $\sigma$, to describe the data.

\[
\bar{X} \pm 1 \sigma \quad \text{for a sample}
\]

Some distinctive patterns can be found within the normal distribution.
The mean, μ, lies at the halfway point in the distribution, so there is an equal number of data values of X above and below the mean. 

- Approximately 68% of the data values of X will lie within the range μ + σ and μ - σ.
- Approximately 95% of the data values of X will lie within the range μ + 2σ and μ - 2σ.
- Approximately 99.7% of the data values of X will lie within the range μ + 3σ and μ - 3σ.
Properties and Predictions

Ex/ The weights of 10,000 students are normally distributed. The mean of the weights is 55 kg and the standard deviation is 5 kg.

$\mu = 55$ $\sigma = 5$

a) Find the number of students weighing between 50 and 60 kg.
b) Find the number of students weighing less than 45 kg.
c) Find a range of weights that would include 99% of the measures.

\[
\begin{align*}
\Pr(50 \leq X \leq 60) &= 68\% \\
&= 0.68 \times 10000 \\
&= 6800 \\
\Pr(X \geq 45) &= 50\% + \frac{0.25}{2} \\
&= 47.5\% \\
\Pr(40 \leq X \leq 70) &= 99\% \text{ of data - within 3 } \sigma \\
\end{align*}
\]

To be more precise

\[
\frac{1.96}{2} = 0.98 \%
\]

\[
= 0.005
\]

The $z$-score for 0.005 = -2.575

\[
\mu - 2.575 \sigma < X < \mu + 2.575 \sigma
\]

42.125 < X < 67.875 kg

\[
\sigma = 5
\]

\[
\text{Ex/ The weights of 10,000 students are normally distributed.}
\]

\[
\mu = 55
\]

\[
\sigma = 5
\]

\[
\Pr(50 \leq X \leq 60) = 68\% \times 10000 = 6800
\]

\[
\Pr(X \geq 45) = 50\% + \frac{0.25}{2} = 47.5\%
\]

\[
\Pr(40 \leq X \leq 70) = 99\% \text{ of data - within 3 } \sigma
\]

To be more precise

\[
\frac{1.96}{2} = 0.98 \%
\]

\[
= 0.005
\]

The $z$-score for 0.005 = -2.575

\[
\mu - 2.575 \sigma < X < \mu + 2.575 \sigma
\]

42.125 < X < 67.875 kg
- The area under the curve can be found by using Z-scores, where \( z = \frac{x - \mu}{\sigma} \) because these scores also have a normal distribution with mean 0 and standard deviation 1.

- The z-score refers to the number of standard deviations away from the mean a value is.

- To determine a probability using a z-score:
  1) Calculate the z-score using the formula: \( z = \frac{x - \mu}{\sigma} \)
  2) Using a z-score table, find the probability value for \( P(X \leq x) = P(Z \leq a) \)

- A positive z-score means the value lies above the mean, a negative value is below the mean.
The times that people wait in a grocery checkout line is normally distributed. The mean waiting time is 10 minutes, with a 2.5 minute standard deviation. What is the probability that Jack will be through the line in less than 5 minutes.

\[ P(X < 5) \]

where \( X = \text{wait time} \)

\[ z = \frac{x - \mu}{\sigma} = \frac{5 - 10}{2.5} = -2 \]

\[ P(X < 5) = P(Z < -2) \]

\[ = 0.0228 \]

\[ \therefore \text{only 2.28\% of people wait less than 5 minutes} \]
Ex.
A region receives an average of 5100 mm of rain annually. If the annual rainfall is assumed to be normally distributed, with a standard deviation of 150 mm, find the probability that

a) More than 5500 mm of rain will fall in a year.

b) Less than 5000 mm of rain will fall in a year.

\[ Z = \frac{X - \mu}{\sigma} = \frac{5500 - 5100}{150} = \frac{400}{150} = 2.667 \]

Look up 2.667 in the z-table

\[ Z = 2.66 \quad \Rightarrow 0.9961 \]

\[ Z = 2.67 \quad \Rightarrow 0.9962 \]

\[ 99.62\% \text{ of the rainfall is below } 5500 \text{ mm} \]

\[ 100\% - 99.62\% = 0.38\% \]

There is only a 0.38\% probability that more than 5500 mm of rain will fall in a year.

\[ Z = \frac{5000 - 5100}{150} = \frac{-100}{150} = -0.667 \]

\[ \Rightarrow \text{from table } = 0.2514 \text{ or } 25.14\% \text{ is to the left of } 5000 \text{ mm} \]

There is a 25.14\% chance of having less than 5000 mm of rain.
Using Percentiles

A percentile refers to the set of values contained within a range of the normal distribution.

i.e. The 90% percentile of a distribution means that a given value is higher than or equal to 90% of the values in the data set.

You are looking for the value that satisfies the equation

\[ P(X \leq x) = 90\% \]
Ex.

Resting pulse rates of humans are normally distributed. The mean pulse rate is 72 beats per minute and the standard deviation is 10.

a) What is the percentile rank of an athlete with a resting pulse rate of 80?
b) How many, of 1 million people, would have a rate of less than 60 bpm?

\[ z = \frac{x - \mu}{\sigma} = \frac{80 - 72}{10} = 0.8 \]

\[ P(X < 80) = P(Z < 0.8) = 0.7881 \]

Therefore, 78th percentile

\[ z = \frac{x - \mu}{\sigma} = \frac{60 - 72}{10} = -1.2 \]

\[ P(X < 60) = P(Z < -1.2) = 0.1151 \]

\[ 0.1151(1000000) = 115100 \]

---

After exercise

\[ z = \frac{128 - 72}{10} = 5.6 \]

Too high

to compare with

resting heart rate

for 68 bpm

\[ z = \frac{68 - 72}{10} = -0.4 \]

\[ 0.3446 \]

34.5% of people have

a lower heart rate

34.5% of people have a lower heart rate
Cutoff Values

Ex/ What heart rate would the top 10% of the population have?

- Look up 90% in the z-score table...

\[ 0.90 \approx 0.9015 \text{ which equates to } z\text{-score of 1.29} \]

\[
z = \frac{x - \mu}{\sigma}
\]

\[
1.29 = \frac{x - 72}{10}
\]

\[
x = 72 + 1.29 \times 10
\]

\[ x = 84.9 \]

The top 10% of people have a resting heartrate of 84.9 or higher.

How about the top 10% of athletes?

Practice Questions:

Pg 430

#1 - 4, 7 - 10, 12