Section 2.2: The Normal Distributions

After this section, you should be able to…

- ✓ DESCRIBE and APPLY the 68-95-99.7 Rule
- ✓ DESCRIBE the standard Normal Distribution
- ✓ PERFORM Normal distribution calculations
- ✓ ASSESS Normality

Normal Distributions

A class of distributions whose density curves are symmetric, uni-modal, and bell-shaped.

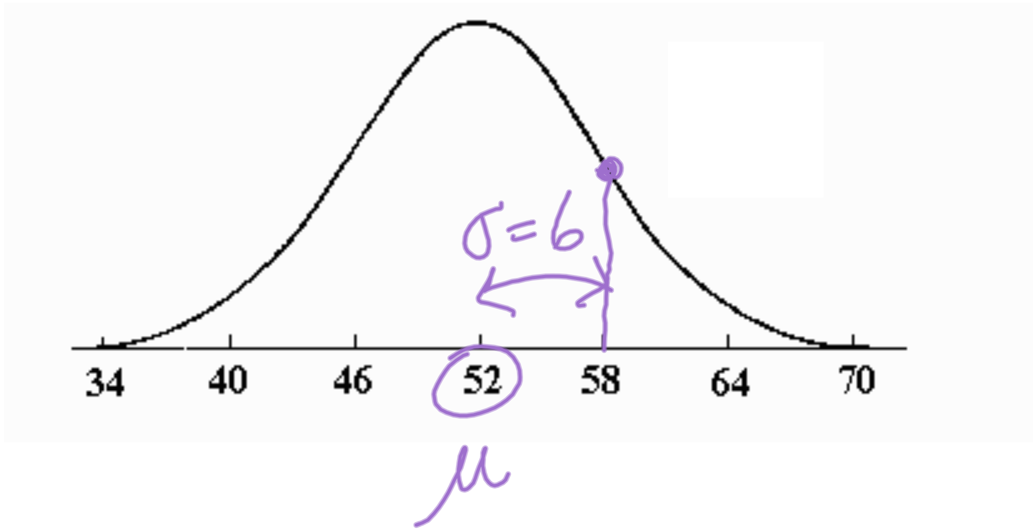
Normal distributions are VERY important in statistics.

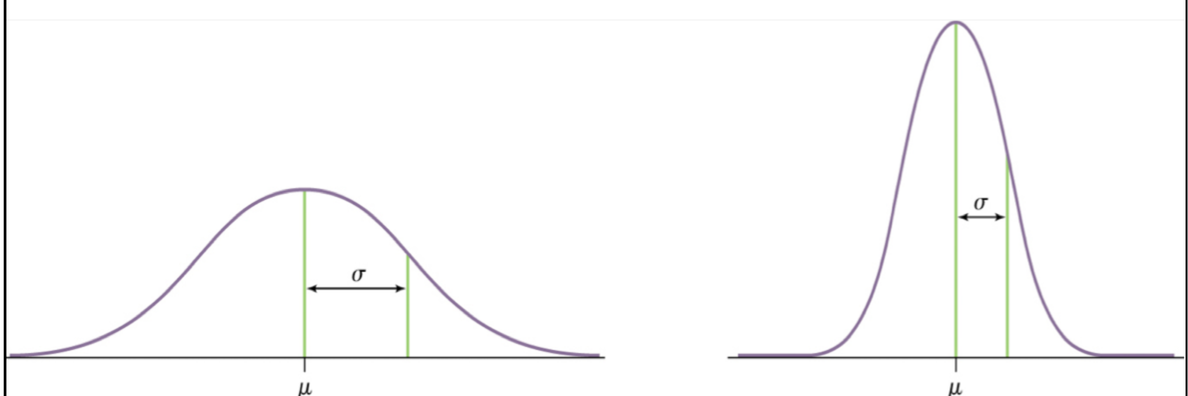
Which numerical summary would we use to describe the center and spread of a Normal distribution?

mean & st. dev ✓

Notation: $N(\mu, \sigma)$

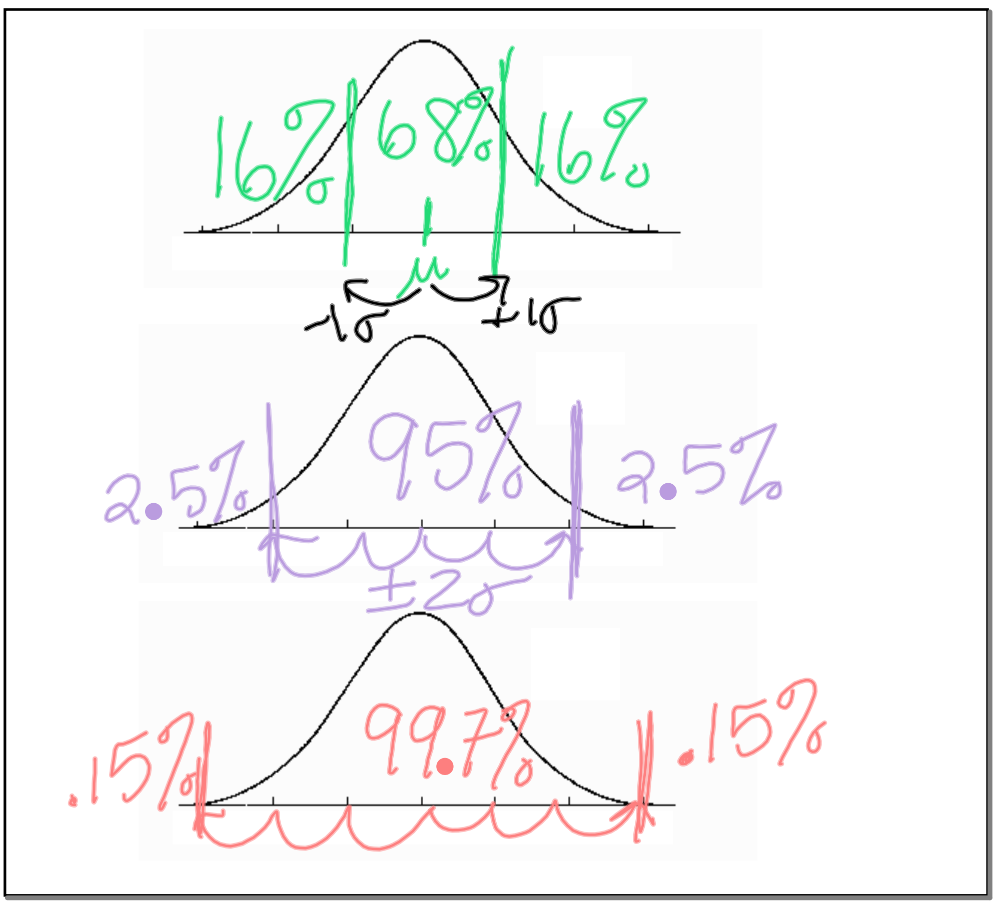
Calculating σ using the Normal density curve





The 68-95-99.7 Rule - In the Normal distribution with mean μ ad standard deviation σ:

- **68%** of all the observations fall within one standard deviation (σ) of the mean μ (in both directions)
- **95%** of all the observations fall within two standard deviations (2σ) of the mean μ (in both directions)
- **99.7%** of all the observations fall within three standard deviations (3σ) of the mean μ (in both directions)



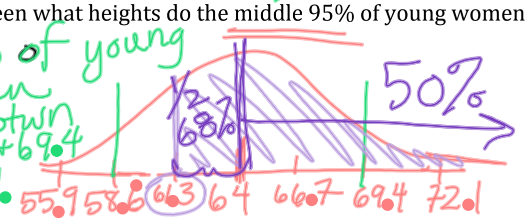
Section 2.2 Notes - Almost Done

The distribution of heights of women aged 20 to 29 is approximately Normal with mean 64 inches and standard deviation 2.7 inches. Use the 68-95-99.7 rule to answer the following questions.

$\pm 2\sigma$

(a) Between what heights do the middle 95% of young women fall?

95% of young women are btwn 58.6 + 69.4 in. tall.



(b) What percent of young women are taller than 61.3 inches?

84% of young women are taller than 61.3 in.

$<61.3 \rightarrow 16\%$

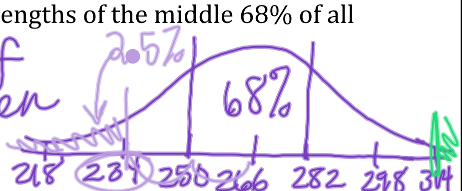
$100\% - 16\% \rightarrow 84\%$

$\frac{1}{2}(68\%) + 50\% \rightarrow 84\%$

You try:
The length of human pregnancies from conception to birth varies according to a distribution that is approximately Normal with mean 266 days and standard deviation 16 days. Use the 68-95-99.7 rule to answer the following questions.

(a) Between what values do the lengths of the middle 68% of all pregnancies fall?

The middle 68% of preg. last between 250 + 282 days.



(b) How short are the shortest 2.5% of all pregnancies?

The shortest 2.5% of all pregnancies are less than 234 days.

(c) What percent of pregnancies are longer than 314 days?

$+3\sigma \ldots 99.7\% \rightarrow \frac{100-99.7}{2}$

.15% of all preg. are longer than 314 days.

The standard Normal distribution

- Infinitely many Normal distributions
  One for every possible combination of means and standard deviations

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

- Standard Normal distribution - N(0, 1)

- We can *standardize* any value of a variable, *x*. This standardized value is called the *z*-score, or *z*. If we actually want to do calculations using this standardized score we need to know the distribution of the original variable. If the original variable is Normal then the *z*-score comes from a standard Normal distribution.

- **A *z*-score tells us how many standard deviations the original observation falls away from its mean AND in which direction.**

standardizing

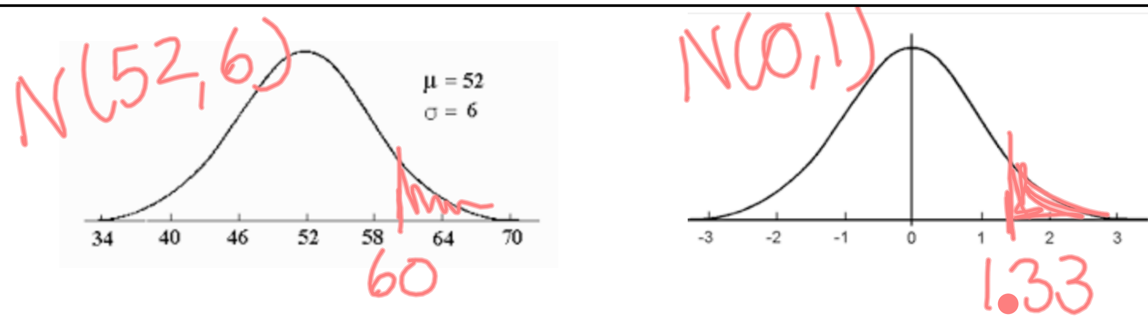
meas. of center • subtract μ (of original data)

meas. of center / meas. of spread • divide by σ (of original data)

orig. (N) $\xrightarrow{\text{standardizing}}$ Standard Normal

orig.		Standard Normal	
mean $\to \mu$	$\frac{\mu - \mu}{\sigma}$	mean	0
st. dev. $\to \sigma$	$\frac{\sigma}{\sigma}$	st. dev.	1

Section 2.2 Notes - Almost Done



$$z = \frac{60-52}{6} = 1.33$$

YOU TRY:
The heights of women aged 20 to 29 are approximately Normal with mean 64 inches and standard deviation 2.7 inches: $N(64, 2.7)$. Men the same age have mean height 69.3 inches with standard deviation 2.8 inches and follow an approximately Normal distribution: $N(69.3, 2.8)$. What are the z-scores for a woman 6 feet tall and a man 6 feet tall? Say in simple language what information the z-scores give the actual heights do not.

woman

$$z = \frac{72-64}{2.7} = 2.96$$

man

$$z = \frac{72-69.3}{2.8} = .96$$

A 6 ft tall woman (or taller) occurs about .15% of the time but for men it is more common (~16% of the time)

Using the Normal curve to determine proportions/percentiles (Beyond the 68-95-99.7 rule)

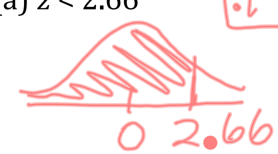
• The area under any Normal curve (or density curve for that matter) is equal to 1.

• If we want to know the proportion of observations that lie within a certain range of observation values we look for the area of the density curve between those two values (for ANY density curve - not just Normal)

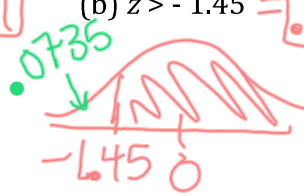
• We have a table that gives us these values for **ONLY** the standard Normal distribution.

Use table A to find the proportion of observations from a standard Normal distribution that satisfies each of the following statements. In each case, sketch a standard Normal curve and shade the area under the curve that is the answer to the question.

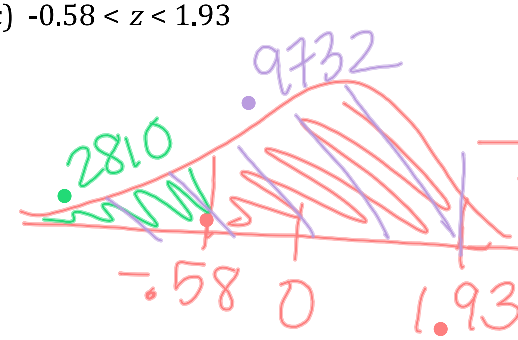
(a) $z < 2.66$ = $\boxed{.9961}$



(b) $z > -1.45$ = $1 - .0735$ = $\boxed{.9265}$



(c) $-0.58 < z < 1.93$



.9732
− .2810
$\boxed{.6022}$

Since we can standardize **ANY** Normal distribution we can use this table for **ANY** Normal distribution.

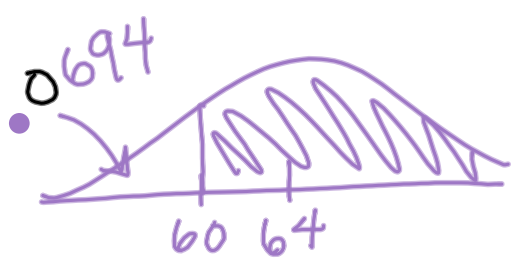
FOR EXAMPLE:
Suppose that the heights of young women have a Normal distribution, $N(64, 2.7)$. What proportion or percentage of all young women are less than 70 inches tall?

$$z = \frac{70-64}{2.7} = 2.22$$

$$\boxed{.9868}$$

98.68% of young women are shorter than 70 inches.

Using the same distribution from the last example, what proportion of women are greater than 60 inches tall?



$$z = \frac{60-64}{2.7} = -1.48$$

$$1 - .0694$$
$$= \boxed{.9306}$$

Using the same distribution from the last example, what proportion of women are between the heights of 62 and 68 inches tall?

$z=\frac{62-64}{2.7}=-.74$

$z=\frac{68-64}{2.7}=1.48$

YOU TRY:
Use table A to find the proportion of observations from a standard Normal distribution that satisfies each of the following statements. In each case, sketch a standard Normal curve and shade the area under the curve that is the answer to the question.

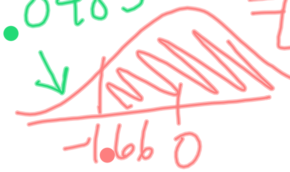
(a) $z < 2.85$ $= .9978$

(b) $z > 2.85$

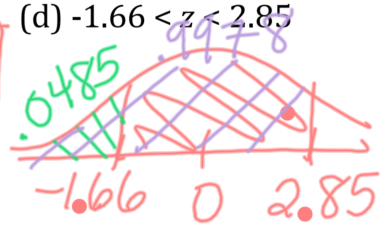
$z<-2.85$ $\quad 1-.9978 = \boxed{.0022}$

(c) $z > -1.66$ $= 1-.0485 = \boxed{.9515}$

.0485



(d) $-1.66 < z < 2.85$



$.9978-.0485=\boxed{.9493}$

Section 2.2 Notes - Almost Done

In the 2008 Wimbledon tennis tournament, Rafael Nadal averaged 115 miles per hour (mph) on his first serves. Assume that the distribution of his first serve speeds is Normal with a mean of 115 mph and a standard deviation of 6 mph.

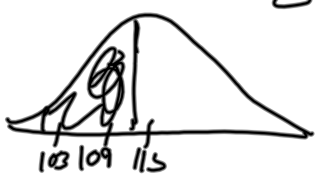
a) About what proportion of his first serves would you expect to exceed 120 mph?

$z = \frac{120-115}{6}$

$z = .83$ $\boxed{.2033}$

20.33% of his serves were faster than 120 mph

b) What percent of Rafael Nadal's first serves are between 100 and 110 mph?



$z = \frac{100-115}{6} = -2.5$ .0062

$z = \frac{110-115}{6} = -.83$ .2033

19.71% $.2033 - .0062 = .1971$

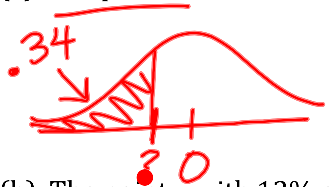
x $z = \frac{x-\mu}{\sigma}$ z Table A   %

μ
σ

Sometimes we are given a particular proportion of observations that lie above or below some observed value and we want to find that observed value.

FOR EXAMPLE:
Use table A to find the value of z of a standard Normal variable that satisfies each of the following conditions.

(a) The point z with 34% of the observations falling below it.

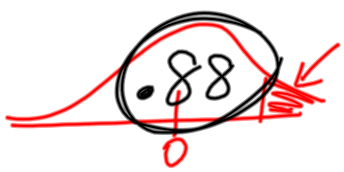


.01

$-0.4 \leftarrow$ .3409 .3372

$\boxed{z = -.41}$

(b) The point z with 12% of the observations falling above it.



.07 .08

$1.1 \leftarrow$ .8790 .8810

$\boxed{z = 1.17 \text{ or } 1.175 \text{ or } 1.18}$

FOR EXAMPLE:
Suppose that the heights of young women have a Normal distribution, $N(64, 2.7)$. What heights are 75% of young women less than?

$x = ?$ X = heights %$\rightarrow z \rightarrow x$

$z = .67$

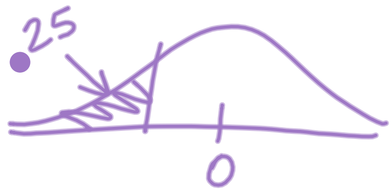
$$.67 = \frac{x - 64}{2.7}$$

$$1.809 = x - 64$$

$$\boxed{x = 65.81 \text{ in.}}$$

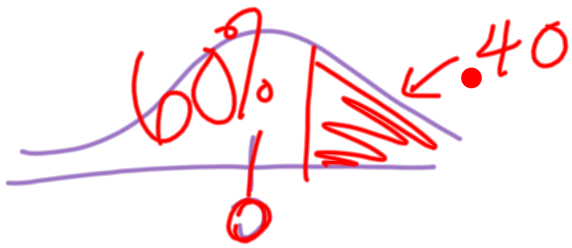
YOU TRY:
Use table A to find the value z of a standard Normal variable that satisfies each of the following conditions. (Use the value of z from Table A that comes closest to satisfying the condition.) In each case, sketch a standard Normal curve with your value of z marked on the axis.

(a) The point z with 25% of the observations falling below it.



$z = -.67$

(b) The point z with 40% of the observations falling above it.



$z = .25$

z-scores for middle 80%

$.10 \rightarrow \boxed{z = -1.28 \quad z = 1.28}$

Scores on the Wechsler Adult Intelligence Scale are approximately Normally distributed with $\mu = 100$ and $\sigma = 15$.

(a) What IQ scores fall in the lowest 25% of the distribution?

$$-.67 = \frac{x-100}{15}$$

$$-10.05 = x - 100$$

$x = 89.95$ or less

(b) How high an IQ score is needed to be in the highest 5%?

$$1.64 = \frac{x-100}{15}$$

$$x \geq 124.6$$

124.6 or higher

Normal Calculations using your Calculator

Taking an observation - x - and converting it to a percentile: %

1. Press 2nd VARS to get the DISTR menu
2. select option 2:normalcdf
3. enter the lower bound, upper bound, μ, σ
4. If you are interested in the values LESS than a certain x-value then use -10^99 as your lower bound
5. If you are interested in the values GREATER than a certain x-value then use 10^99 as your upper bound.
6. IF you have already converted your x-value to its corresponding z-score, don't enter the μ and σ

x or z $\longrightarrow$ % Normalcdf

Taking a percentile (percent to the left of a value) and converting it to a value of the original variable (x):

1. Press 2nd VARS to get the DISTR menu
2. select option invnorm
3. enter the percentile (percent to the left of a value), μ, σ
4. Remember, if you are given the percent GREATER than a value, subtract from 1 (100%) to get the percentile.
5. IF you just want the corresponding z-score (not x-value) don't enter μ and σ

left tail % → x or z invNorm

Assessing Normality

As we've seen, Normal models provide good models for some distributions of real data.

However, some common variables are usually skewed and therefore distinctly non-Normal.

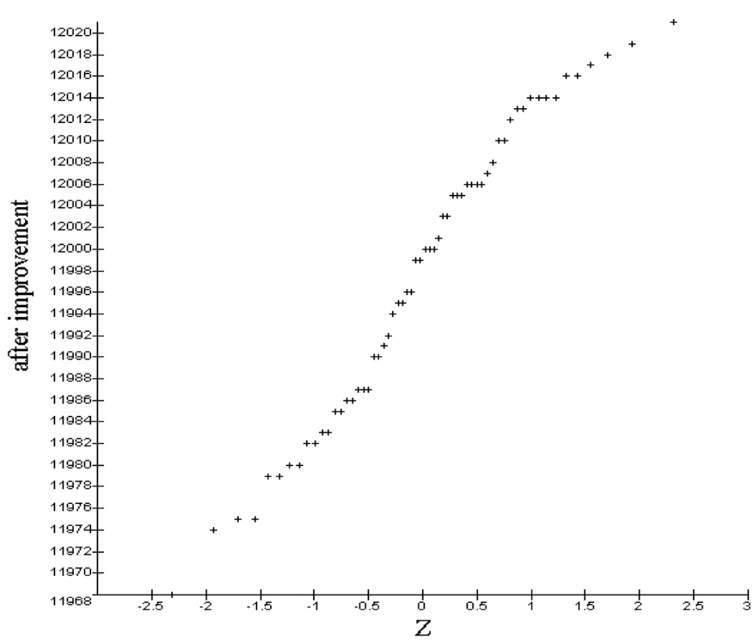
It is risky to assume that a distribution is Normal without inspecting the data or even if the data are uni-modal and roughly symmetric.

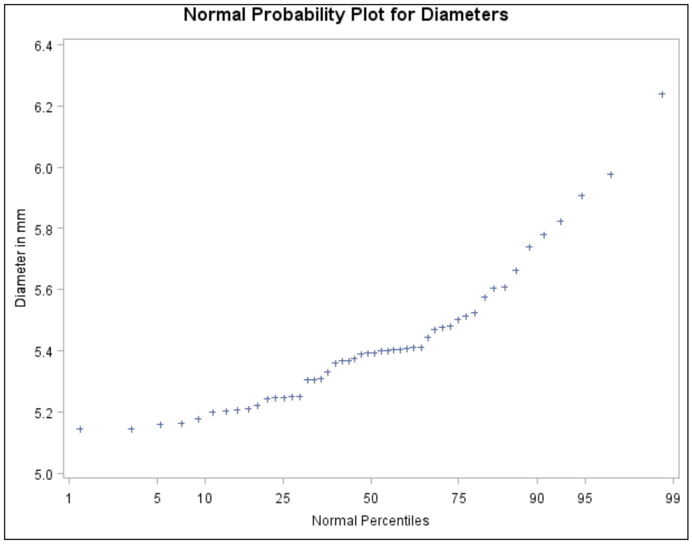
We an check to see if the distribution of the data follow the 68-95-99.7 rule.

We an also use a ***Normal Probability Plot*** - a plot of each observation against the corresponding z-score for the percentile it represents. If there is a strong linear pattern, the distribution is close to Normal.

Normal Probability Plots on the Calculator

1. Enter the data into a single list
2. Go to STATPLOT (2nd Y=)
3. Turn a single plot on
4. Select the LAST of the graphs - bottom right
5. Select the correct Data List
6. Select X as your Data Axis
7. Choose the mark you would like to see in your graph for the points
8. Go to your graph and use Zoom option 9:zoomstat





The measurements listed below describe the useable capacity (in cubic feet) of a sample of 36 side-by-side refrigerators. Are the data close to Normal?

12.9 13.7 14.1 14.2 14.5 14.5 14.6 14.7 15.1 15.2 15.3 15.3
15.3 15.3 15.5 15.6 15.6 15.8 16.0 16.0 16.2 16.2 16.3 16.4
16.5 16.6 16.6 16.6 16.8 17.0 17.0 17.2 17.4 17.4 17.9 18.4

Homework: p. 131 #s 41 - 59 odd, 63, 65, 66, 68, 69-74 all