Chapter 5 Review: Statistics

Textbook p.210-282
Summary: p.238-239, p.278-279
Practice Questions p.240, p.280-282
Z-Score Table p.592

Key Concepts: Central Tendency, Standard Deviation, Graphing, Normal Distribution, Z-Score, Confidence

Central Tendency

1) Mean = The calculated <u>average</u>. Add up the numbers and divide by how many there are.
2) Median = The <u>middle</u> number (might be half way between two numbers in the list)
3) Mode = The number that occurs <u>most often</u>
4) **Standard Deviation** = The average <u>distance</u> of each number from the <u>mean</u>

FORMULA FOR STANDARD DEVIATION:

$$\sigma = \sqrt{\frac{\text{sum of the squares of the differences from the mean}}{\text{number of values}}}$$

Example: Find all measures of central tendency for the following list of numbers:
6,4,9,4,8,5

a) Mean $\dfrac{6+4+9+4+8+5}{6} = \dfrac{36}{6} = \underline{6}$

b) Median $4,4,5,\underset{\uparrow}{6},8,9 \qquad \dfrac{5+6}{2} = \underline{5.5}$

c) Mode $= 4$

d) Standard Deviation

$6 \rightarrow 6 = 0^2 = 0$
$4 \rightarrow 6 = 2^2 = 4$
$9 \rightarrow 6 = 3^2 = 9$
$4 \rightarrow 6 = 2^2 = 4$
$8 \rightarrow 6 = 2^2 = 4$
$5 \rightarrow 6 = 1^2 = \underline{1}$
$\qquad\qquad\qquad 22$

$\sigma = \sqrt{\dfrac{22}{6}}$

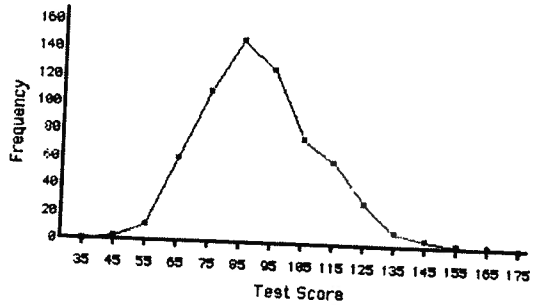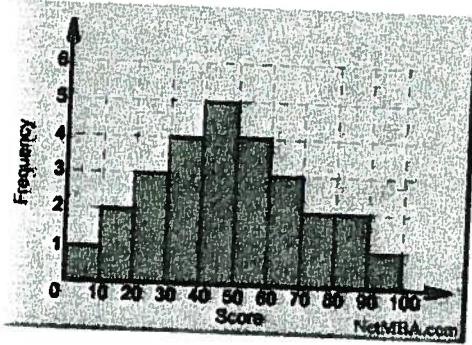$\sigma = \underline{1.915}$

Graphing

For larger sets of data, we break it into chunks to find the measures of central tendency. We also drew graphs to represent the data:

Histogram = _Bars_   Frequency Polygon = _Line_

L_1 L_2

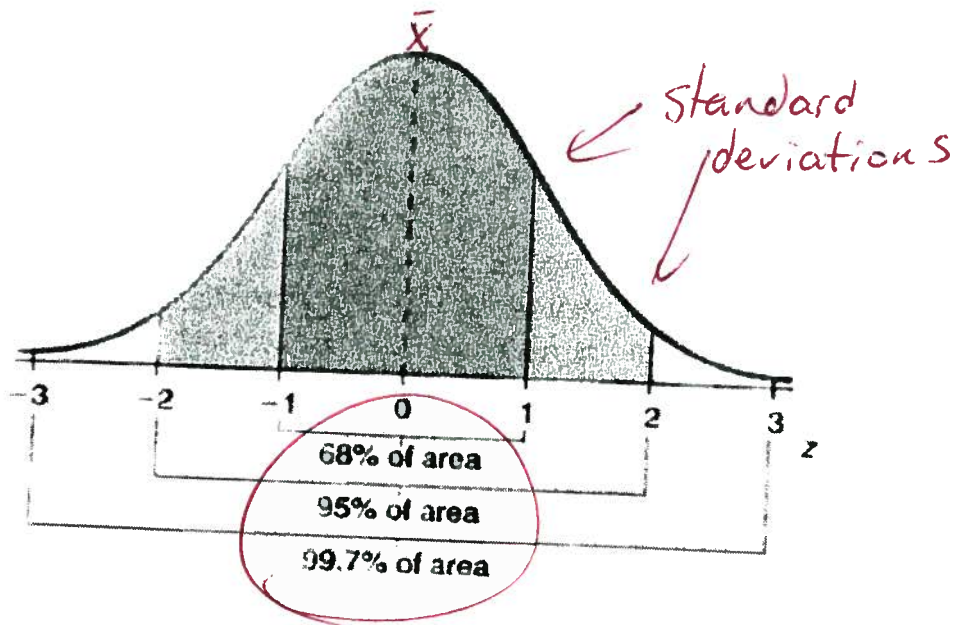| L_1	L_2
5	1
15	2
25	3
.	.
.	.

Examples:

For large data like this, you can input the numbers into your calculator to find the measures of central tendency. Remember to use the _middle_ number for each category.

Step #1: _STAT, Edit_

Step #2: _STAT, CALC, 1-Var Stats_

$\bar{x}$ = average

σx = standard deviation

Med = median

Normal Distribution

Standard deviations

68% of area
95% of area
99.7% of area

Z-Scores

The z-score is the distance from a point to the _mean_, in terms of
Standard deviations.

*If I am 1.5 standard deviations above the mean, then my z-score is_1.5_*
*If I am 0.6 standard deviations below the mean, then my z-score is_-0.6_*

Formula: $Z = \dfrac{x - \mu}{\sigma}$

$x = point$
$\mu = mean$
$\sigma = standard\ deviation$

Confidence

For interpreting a survey

1) **Confidence interval** = a specific interval estimate of the whole population by using information obtained from a sample. (example: between 9.5 and 10.2 million people are likely to vote for the Liberal party)

2) **Margin of error** = the possible difference between your estimate based on the survey and real-life. (example: 40% will vote Conservative with a margin of error of ±3%)

3) **Confidence level** = the probability that the answer in real-life matches your estimate from the survey. (example: this survey is correct 19 times out of 20)

Key Example: The average score on an normally distributed exam is 64% with a standard deviation of 7%.

a) Label the normal curve

60
73

43   50   57   64   71   78   85
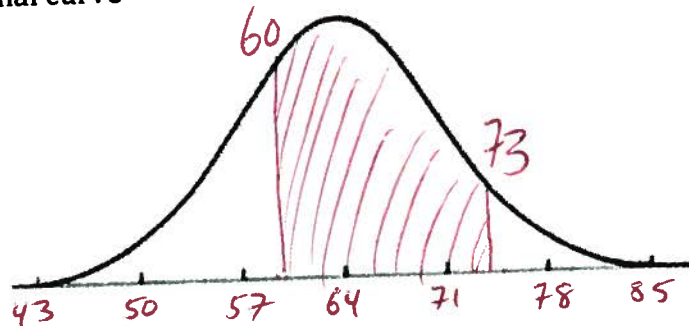
b) Find the percentage of people who score a C (between 60% and 73%)

Z-Score for 60%

$\dfrac{60-64}{7} = -0.57$

CHART → 0.2843

Z-Score for 73%

$\dfrac{73-64}{7} = 1.29$

CHART → 0.9015

$0.9015 - 0.2843 = 0.6172$ OR $\boxed{61.7\%}$

CALCULATOR SHORTCUT

[DISTR]   normalcdf $(60, 73, 64, 7) = .617$ $\boxed{61.7\%}$

Chapter 5 Review: Statistics

Practice #1: Find the measures of central tendency for the marks on sample of 10 quizzes: 14,12,17,3,8,12,15,8,10,11

a) Mode(s) $12, 8$

b) Median $3, 8, 8, 10, 11, 12, 12, 14, 15, 17$ $\frac{11+12}{2} = 11.5$

c) Mean $\frac{110}{10} = 11$

d) Standard Deviation $3^2 + 1^2 + 6^2 + 8^2 + 3^2 + 1^2 + 4^2 + 3^2 + 1^2 = \sqrt{\frac{146}{10}} = 3.82$

Practice #2: Students recorded their heights, in inches, when they graduated from kindergarten in 1999 and again when they graduated from high school in 2011.
  1999: 39 41 41 43 45 46 47 46 48 47 44 38 41 39 43 46 44
  2011: 60 74 76 62 64 61 66 68 71 76 74 73 72 69 64 63 60

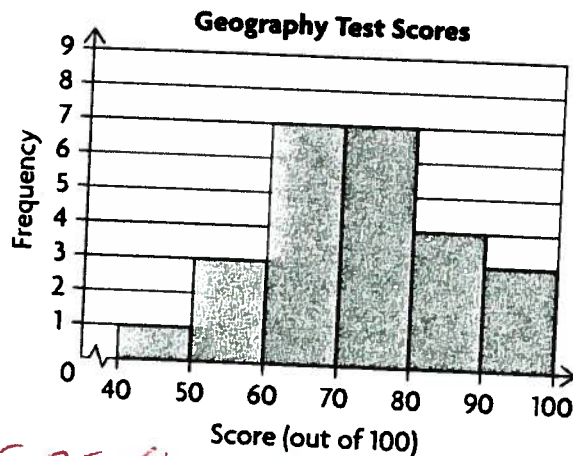a. Determine the mean and standard deviation for each year.

STAT
CALC
1-Var Stats

1999: $\bar{x} = 43.4$ $\sigma x = 3.03$
2011: $\bar{x} = 67.8$ $\sigma x = 5.55$

b. In which year are the heights most consistent?

1999 (standard deviation is lower)

Practice #3: Find the measure of central tendency for the following histogram:

Geography Test Scores

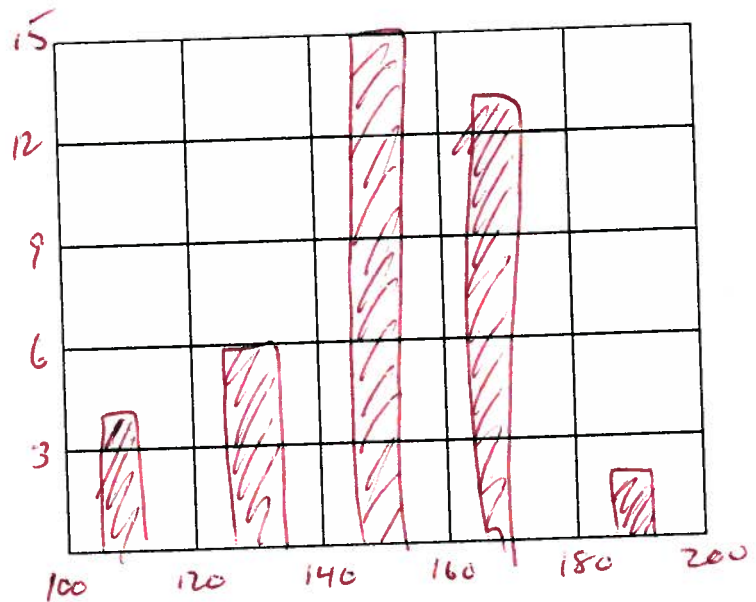| L_1	L_2
45	1
55	3
65	7
75	7
85	4
95	3

a) Mode(s) $65, 75$ (two tallest bars)

b) Median $Med = 75$

c) Mean $\bar{x} = 72.6$

d) Standard Deviation $\sigma x = 13.05$

**Practice #4:** Four groups of students recorded their pulse rates after a 2 km run.

Group 1	126	168	158	192	146	166	104	164	116	138
Group 2	158	132	156	160	108	150	178	136	172	140
Group 3	136	174	156	176	150	166	142	156	130	182
Group 4	144	150	142	152	174	176	118	152	178	164

a) Make a frequency table with five intervals to organize the pulse rates.
b) Construct a histogram of the data.

Interval	Frequency
1) 100 – 119	IIII
2) 120 – 139	IIII I
3) 140 – 159	IIII IIII IIII
4) 160 – 179	IIII IIII III
5) 180 – 200	II

Practice #5: *Suppose scores on an IQ test are normally distributed. If the test has a mean of 100 and a standard deviation of 10, what is the probability that a person who takes the test will score between 90 and 110?*

68%
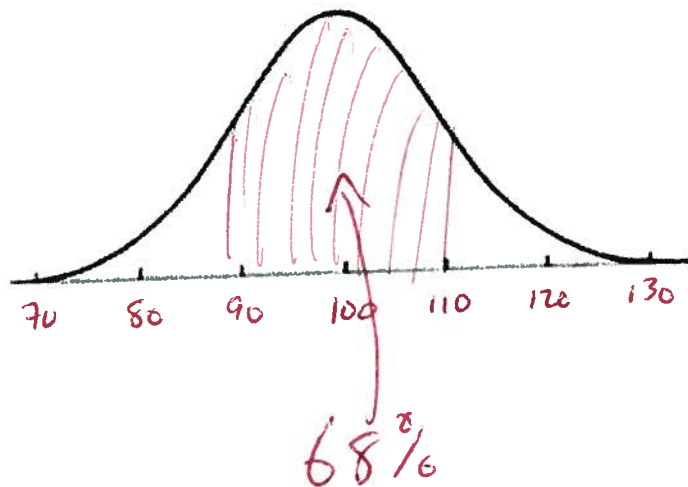
Practice #6: *A machine is used to fill soda bottles. The amount of soda dispensed into each bottle varies slightly. Suppose the amount of soda dispensed into the bottles is normally distributed. If at least 99.7% of the bottles must have between 585 and 595 milliliters of soda, find the greatest standard deviation, to the nearest hundredth, that can be allowed.*

$$\frac{10}{6} = \boxed{1.67} = \sigma$$

585   590   595
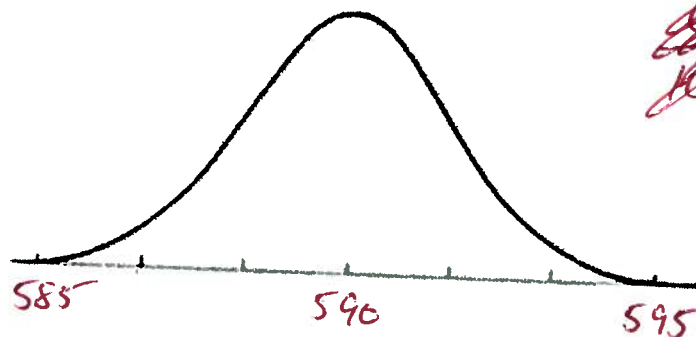
Practice #7: Determine the z-score for the given value.
$\mu = 120, \sigma = 10, x = 125$

$$\frac{125 - 120}{10} = \frac{5}{10} = 0.5$$

Practice #8:
Determine the percent of data to the left of the z-score: z = –1.50.
CHART   0.0668   6.68%
Determine the percent of data to the right of the z-score: z = 2.26.
CHART   0.9881   98.81%

Practice #9: What is the probability of getting a z-score of...
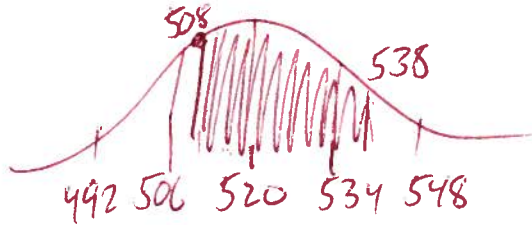a) Less than 1.61?   0.9463   94.63%
b) Less than -0.55?   0.2912   29.12%

Practice #10: An IQ test has a mean of 100 with a standard deviation of 15. What is the probability of getting less than 80 on the test?

$$z = \frac{80 - 100}{15} = -1.33$$

CHART = 0.0918   or   9.18%

Practice #11: Yumi always waits until her gas tank is nearly empty before refuelling. She keeps track of the distance she drives on each tank of gas. The distance varies depending on the weather and the amount she drives on the highway. The distance has a mean of 520 km and a standard deviation of 14 km. What percent of the time does Yumi drive between 508 km and 538 km on a tank of gas?

normalcdf (508, 538, 520, 14)

= .7050 70.5%

492 506 520 534 548

Z-Score

$\frac{508-520}{14}$ = -.857

CHART .1949

Z-Score

$\frac{538-520}{14}$ = 1.28

.9015

.9015 - .1949 = .7066

70.66%

Practice #13: A poll was conducted to ask voters the following question: If an election were held today, whom would you vote for? The results indicated that 53% would vote for Smith and 47% would vote for Jones. The results were stated as being accurate within 3.8 percentage points, 19 times out of 20.

a) What is the confidence interval (percentage)? Smith : 49.2 – 56.8

Jones : 43.2 – 50.8

b) If the number of people likely to vote is 300 000, how many people will vote for Smith, and how many will vote for Jones?

Smith : 147600 – 170400

Jones : 129600 – 152400

c) What is the confidence level?

$\frac{19}{20}$ OR 95%

d) If the polling company conducted this same survey using the same sample size, but used a confidence level of 99%, what would happen to the margin of error? larger

e) Who will win the election?

probably Smith, but not certain