Homework #1

Solutions
Math 1011
21 July, 2014

Chapter 2

1 (a) what about population size?
(Minnesota has twice the population of Michigan so the crime rate in Michigan is actually lower.)

(b) Again, think about population. The population of the US increased from 1991 to 2001, so actually the crime rate did decrease, but even more than the investigator seems to suspect.

4 (a) They were controlling for age and gender as possible confounders.

(b) This is the wrong conclusion. This is an observational study, because to become an ex-smoker you must choose to quit smoking. Many people choose to quit smoking because they are sick, and this creates bias.

5 No! Trust the double-blind study. The results of the other study must have a bias.

11 (a) treatment group = those who finished boot camp

control group = all others (that is, prisoners who did not complete the program or did not volunteer.

(b) Observational. This is because the prisoners decide if they want to volunteer and later if they want to drop out.

(c) FALSE.

(It is very likely that the prisoners who were very serious about being rehabilitated and not returning to prison are exactly those who would sign up for boot camp and not drop out.)

12 FALSE. This is an example of "Simpson's Paradox".

Example:

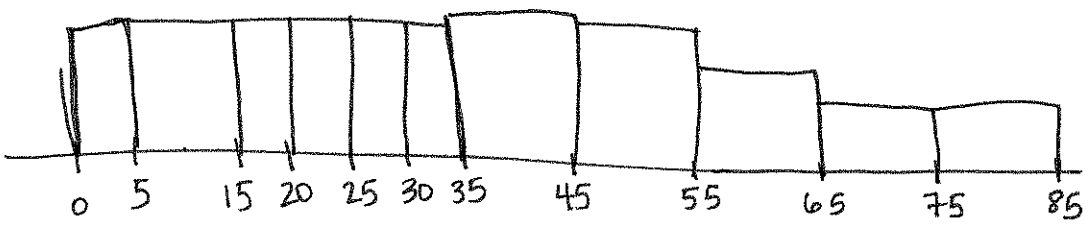
	DEM		REP	
	Total number	Number voting	total number	Number voting
Ward A	1000	100	100	5
Ward B	100	60	1000	500
Total	1100	160	1100	505

Notice a higher percent of democrats voted in each ward but a higher percent of Republicans voted in total.

Chapter 3

1 66 inches, 72 inches.

2 Histogram:



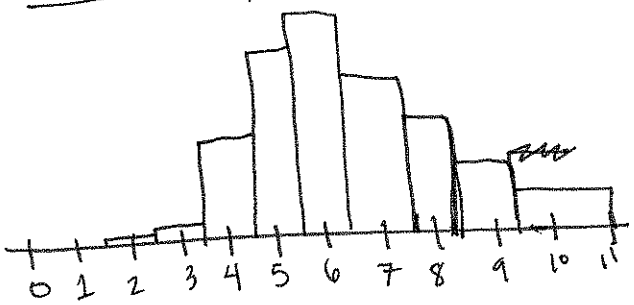
(a) More children age 1.

(b) 21 year olds

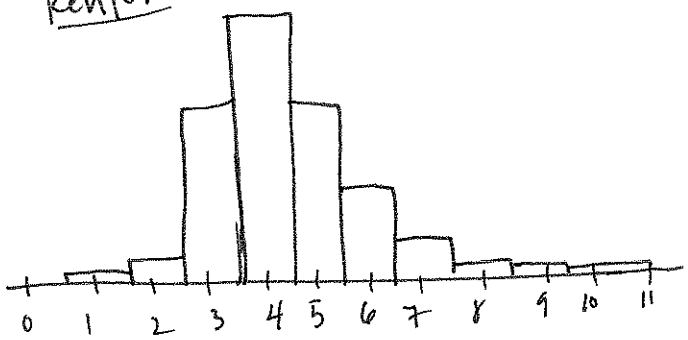
(c) 0 - 4

(d) 50%

3 Owner Occupied



Renter:



(a) Rounding

(b) No. Taking percentages adjusts for the different total numbers. In general, rented units are just smaller.

(c) Owner - Occupied.

5 $100,000 - $90,000 = $10,000 (width of interval we are interested in)

Area of block = width × height

$= \$10{,}000 \times 1\ \%/\$1000 = \boxed{10\%}$

Chapter 4:

1 (a) $\text{Ave} = \dfrac{41+48+50+50+54+37}{6}$

$= \dfrac{300}{6} = 50$

So the deviations from average are

$-9, -2, 0, 0, 4, 7$

Thus SD is

$SD = \sqrt{\dfrac{(-9)^2+(-2)^2+0^2+0^2+4^2+7^2}{6}}$

$= \sqrt{\dfrac{150}{6}} = \sqrt{25} = 5$

$\boxed{\text{Ave} = 50,\ SD = 5}$

(b) 48, 50, 50 are within 0.5 of an SD

48, 50, 50, 54, 57 are within 1.5 SD.

2 (a) List (ii) has the smaller SD because it has more values at the average.

(b) List (i) has smaller SD because list (ii) has two outliers stretching the SD.

6 (a) (i) 60, (ii) 50, (iii) 40

(b) (i) median bigger than average

(ii) about equal

(iii) median less than average

(c) 15. (Most of the area is within 50 of average, so it is too big but very little area is within 5 of average, so 5 is too small.)

(d) False. They are almost mirror images, so they should have about the same SD.

(reflecting the histogram does not change how "spread out" it is.)

9 (a) Yes.

The average goes up by

$$\frac{\$986{,}000 - \$98{,}600}{1000} = \boxed{\$887.40}$$

(b) No. The median is not changed by outliers.

12 The data provides weak support for the theory.

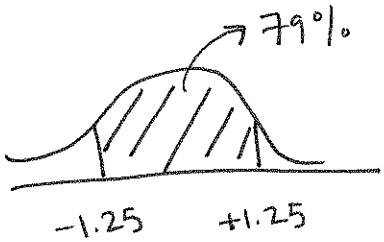
The problem is that this data is cross-sectional, so it could be true that each year many people escape from poverty and are replaced by new people who were not in poverty the previous year.

(In fact, longitudinal data shows most spells of poverty are short.)

Chapter 5

1 (a) From page A-104

z	Area
1.25	78.87



About 79% of the scores should be within 1.25 SDs. $79\% \cdot (25 \text{ scores}) = \boxed{20}$

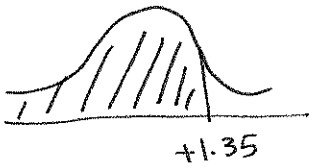
(b) 1.25 SDs = [redacted] 12.5

There are $\boxed{18 \text{ scores}}$ between 37.5 and 62.5

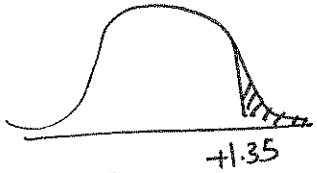
4 (a) Ave = 538
SD = 120

$\frac{700-538}{120} = +\frac{162}{120} \approx 1.35$ ← a score of 700 in standard units (for men)

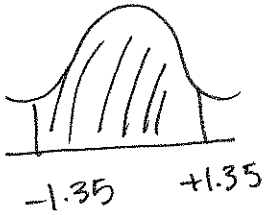
So we want to compute:



to find

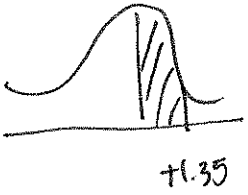


From page A-104 we know



= 82.30

So



$= \frac{82.30}{2} = 41.15$

Thus

$= 50 + 41.15 = 91.15$

So about 91.15% of students score less than 700.

This means $100 - 91.15 = 8.85\%$ score over 700.

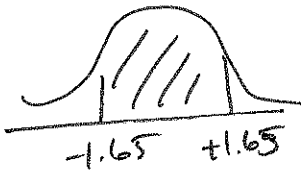
so | Approximately 9% score over 700

(b) First convert 700 to standard units

Ave = 504
SD = 120

$$\frac{700-504}{120} = \frac{196}{120} \approx 1.65$$

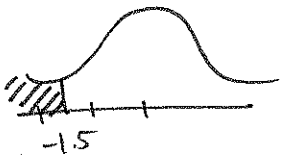
According to A104, $z = 1.65$ corresponds to an area of 90.11 ≈ 90.

So  = 90

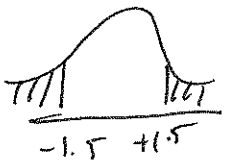
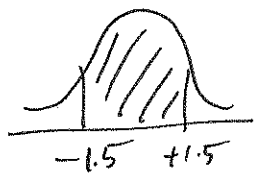
Thus  = 95 so  = 5.

Therefore, about 5% score over 700.

7 (a) 400 is 1.5 SD below average $\left(\frac{400-550}{100} = 1.5\right)$

so we are looking for 

Notice

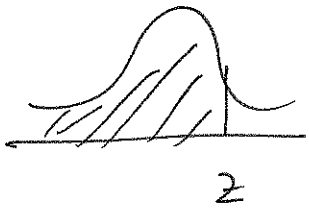
 = $\frac{1}{2}$ ·  = $\frac{1}{2}\left(100 - \right.$  $\left.\right)$

By page A104  = 86.64 ≈ 87.

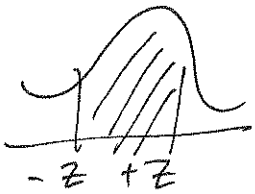
Thus the answer is

$$\frac{1}{2}(100-87) = \frac{1}{2}(13) = \boxed{6.5}$$

(b) We want to find the z such that



$= 75$

which would mean 

$= 50$

so by page A104 $z \approx 0.7$

So we must find the score which is +0.7 SD above average

$$550 + 0.7(100) = \boxed{620}$$

8 (a) T (b) F (does not change deviation from average)

(c) T (d) T (all deviations from average are doubled)

(e) T (f) F

11 Histogram (i)

Since people cannot take a negative number of courses the data is all positive. Since the Ave is [redacted] smaller than the SD we know the graph must stretch out to the right to create a large SD (there is no room to the left)

Chapter 6

[1] FALSE! There is always chance error. An experienced scientist will make many measurements.

[3] (a) False. Chance errors are sometimes positive and sometimes negative (so they mostly cancel out in the average) but bias pushes in one direction.

(b) False. Same reason ↑

(c) True. (although scientists work hard to keep bias small.)

Chapter 8

[1] [d] with (a) the average is too low.
with (b) the SD is too low
with (c) the SD is too high and the correlation is too high.

[2] (a) Negative; older cars get fewer miles per gallon.

(b) People with more money own newer cars (and maintain them better.)

[7]

0.62	−1.00
−0.85	0.97
0.06	−0.38

(to match up with where the images are on the page)