

MA207 Sample Exam I

Name: KEY

Directions: You have one and a half hours to complete this exam. You may use your calculators, but no other electronic device. You may use the Ma207 Formula Sheet, t -distribution table, and standard normal distribution table provided, but no other notes or handouts. You may not use your book. Please turn off the ringer of your cellphone and store. Also, you must work alone.

I, _____, am fully aware of and have abided by the BSC Honor Code in completing this exam.

Please work in the space provided and *show all of your work*.

1. In Hobbiton, the average height of male hobbits is 58 inches with a standard deviation of 4 inches, while the average height of female hobbits is 54 inches with a standard deviation of 3 inches. Both distributions are approximately symmetric and bell-shaped.

- (a) (2 pts) What height range would you expect about 68% of all *male hobbits* to fall into?
Be sure to include units.

54 inches to 62 inches.

- (b) (2 pts) What height range would you expect about 95% of all *female hobbits* to fall into?
Be sure to include units.

48 inches to 60 inches

- (c) (3 pts) Approximately what percentage of *females* are shorter than 48 inches?

2.5%

- (d) (3 pts) Rosie is at the 95th percentile in height, how tall is she?

From SNT $P(X < Z) = .95 \Rightarrow Z = 1.645$

$$1.645 = \frac{\text{obs} - 54}{3}$$

$$\text{obs} = \boxed{59 \text{ inches}}$$

- (e) (3 pts) Sam is only at the 45th percentile in height, how tall is he?

From SNT $P(X < Z) = .45 \Rightarrow Z = \cancel{1.645} - .13$

$$-.13 = \frac{\text{obs} - 58}{4}$$

$$\text{obs} = \boxed{57 \frac{1}{2} \text{ inches}}$$

- (f) (2 pts) If Sam and Rosie get married, who would be taller in the wedding photo?

Rosie

- (g) (3 pts) Frodo's percentile is 15 percentage points more than Sam's percentile. How tall is Frodo (Frodo is a male hobbit)?

Frodo is at 60th percentile $\Rightarrow Z = .25$
From SNT

$$.25 = \frac{\text{obs} - 58}{4}$$

$$\Rightarrow \boxed{\text{Frodo is 59 inches tall.}}$$

2. Theoden, King of Rohan, has divided the 2000 Horsemen of the Mark in preparation for the assault by the Urak-hai Orcs of Saruman. While all are natural horse riders, the warriors have preference for sword versus bow as indicated in the table below.

Division	Number of Swordsmen	Number of Archers	
1	134	278	412
2	176	296	472
3	220	324	544
4	352	220	572
	<u>882</u>	<u>1118</u>	<u>2000</u>

- (a) (1 pt each) Describe each variable as Quantitative or Categorical by circling the appropriate choice.

- i. Division - Quantitative / Categorical
 ii. Number of Swordsmen - Quantitative / Categorical
 iii. Number of Archers - Quantitative / Categorical

- (b) (2+1 pts) Find the proportion of the 2000 Horsemen of the Mark who prefer the bow (i.e. are archers). What type of proportion is this (marginal, conditional or joint)?

$$\frac{1118}{2000}$$

marginal

- (c) (2+1 pts) What proportion of the first division are swordsmen? What type of proportion is this (marginal, conditional or joint)?

$$\frac{134}{412}$$

conditional

- (d) (2+1 pts) What proportion of the total are archers in the third division? What type of proportion is this (marginal, conditional or joint)?

$$\frac{324}{2000}$$

joint

- (e) (4 pts) For each of the four divisions, calculate the proportion of warriors who prefer the sword. Is there an association between division and preferring a sword? How do you know?

Division	prop of swordsmen
1	.325
2	.373
3	.404
4	.615

yes
 b/c the numbers are different

(This problem continues on the next page)

(This problem began on the previous page)

- (f) (4 pts) Find the ratio of the proportion of swordsmen in the fourth division to the proportion of swordsmen in the first division. What does this ratio mean in practical terms when wondering about the chances that a randomly selected warrior from one of these two divisions prefers the sword?

$$\frac{\text{Swords in 4th}}{\text{Swords in 1st}} = \frac{.615}{.325} = 1.89$$

A randomly selected ~~swordsmen~~ warrior is almost twice as likely to be a Swordsman from 4th than from 1st.

- (g) (4 pts) One of the divisions will perform a flanking maneuver to use bows from a distance on the Orcs. Which of the four divisions is most suitable and why? (There are a couple of good answers, you will be graded on the quality of the written explanation.)

Either 1st because the proportion of Archers is higher or 3rd because the number of Archers is higher.

- (h) (4 pts) Another of the divisions will blunt the Orc's charge with close combat using swords. Which of the four division is most suitable and why? (Really only one answer makes sense here, you will be graded on the explanation.)

the 4th division as both the number of swordsmen + the proportion of swordsmen is higher.

3. After a battle an inquisitive hobbit surveys a selection of elf warriors asking about their number of kills. The hobbit collects the following data:

12, 8, 10, 17, 9, 10, 15, 4, 7, 12, 8, 9, & 45.

The last elf asked was Legalos.

- (a) (2 pts) Find the mean number of kills for the collection of elves.

$$\bar{x} = \frac{166}{13} = 12.77$$

- (b) (2 pts) Find the median number of kills for the collection of elves.

$$4, 7, 8, 8, 9, 9, \underset{\uparrow}{10}, 10, 12, 12, 15, 17, 45$$

$$Q_2 = 10$$

- (c) (3 pts) What is the five number summary for this data?

$$4 - 8 - 10 - 13.5 - 45$$

- (d) (3 pts) Find the interquartile range (IQR) for this data.

$$IQR = 5.5$$

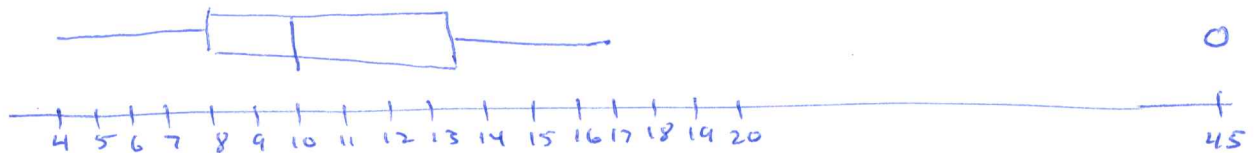
- (e) (3 pts) Are there any outliers according to the 1.5 IQR criterion?

Yes!

$$\begin{aligned} \text{upper} &= Q_3 + 1.5 IQR \\ &= 13.5 + 5.5 = 19 \end{aligned}$$

- (f) (6 pts) In the space below, construct a box-n-whisker plot for the data.

So 45 is
an outlier



- (g) (3 pts) Is the mean or the median a better measure of center for this data? Why or why not?

median, b/c of the outlier.

4. (4 pts - short answer) The correlation between variables A and B is $r = .45$ while the correlation between A and C is $r = -.87$. Which linear association is stronger? A with B , or A with C ? Why?

Between A and C as $|r| = 0.87$ is largest.

5. (4 pts - short answer) For a particular set of data, you know $\bar{x} = 5.23$ while $Q2 = 3.5$. What can you say about the shape of the data?

The data is skewed to the right as the mean is larger than the median.

6. In a group project, you identify weight as explanatory and height as response. You then calculate the correlation, $r = .8$.

- (a) (4 pts) Your partner thinks you should switch the two variables. How does this affect the correlation?

It doesn't. The correlation remains $r = .8$ because correlation does not "see" the labels on the variables.

- (b) (4 pts) What percentage of variation in height is due to the linear association with weight?

$$r^2 = (.8)^2 = .64$$

So 64% of the variation in height is due to the association with weight.

7. The sales price of a home is often associated to the size of the home measured in square feet. Suppose that for a certain neighborhood the regression line for sale price (in thousands) as a function of size (in square feet) is given by:

$$\hat{y} = 130 + .023x$$

- (a) (4 pts) Estimate the sales price of a 2000 ft^2 home. Include units.

$$\begin{aligned}\hat{y} &= 130 + .023(2000) \\ &= 176\end{aligned}$$

$$\text{So } \boxed{\$176,000}$$

- (b) (4 pts) Suppose a home in this neighborhood sold for \$166,800. Estimate the size. Include units.

$$166.8 = 130 + .023x$$

$$\frac{36.8}{.023} = x$$

$$\boxed{1600 \text{ ft}^2}$$

- (c) (2 pts) Is the association between sales price and size positive or negative? Why?

Positive, b/c the slope is positive.

- (d) (2 pts) A particular 1000 ft^2 home sold for \$160,000. Find the residual.

$$\hat{y} = 130 + .023(1000) = 153$$

$$160,000 - 153,000 = \boxed{\$7,000} \leftarrow \text{Residual}$$

- (e) (4 pts) Consider the same house as in part (d). Would you predict that this home was particularly nice for the neighborhood, or a bit dated and run-down? Why?

Nice, it sold for over the predicted price.