

Name:
Section:

Exam I

MATH 444
Show all your work

1. A president at a university with 12,000 students and 200 faculty members is considering changing the length of the winter intercession from 5 weeks to 2 weeks. The faculty senate is interested in the faculties' opinion concerning this change and will sample 50 faculty members.

a. (4 pts) What is the population of interest?

200 Faculty, or the 200 faculty opinions

b. (4 pts) What is the sample?

50 sampled faculty, or the opinions of the 50 sampled faculty

c. (4 pts) How could a systematic sample be selected?

Use a list of names and select every fourth name on the list. The first name should be selected at random from among the first 4 names.

d. (4 pts) If the sample is used to make inferences about faculties' opinions, why is a measure of reliability of the inferences important?

Because a sample is used, there is a risk of an incorrect inference. We should know something about our level of risk

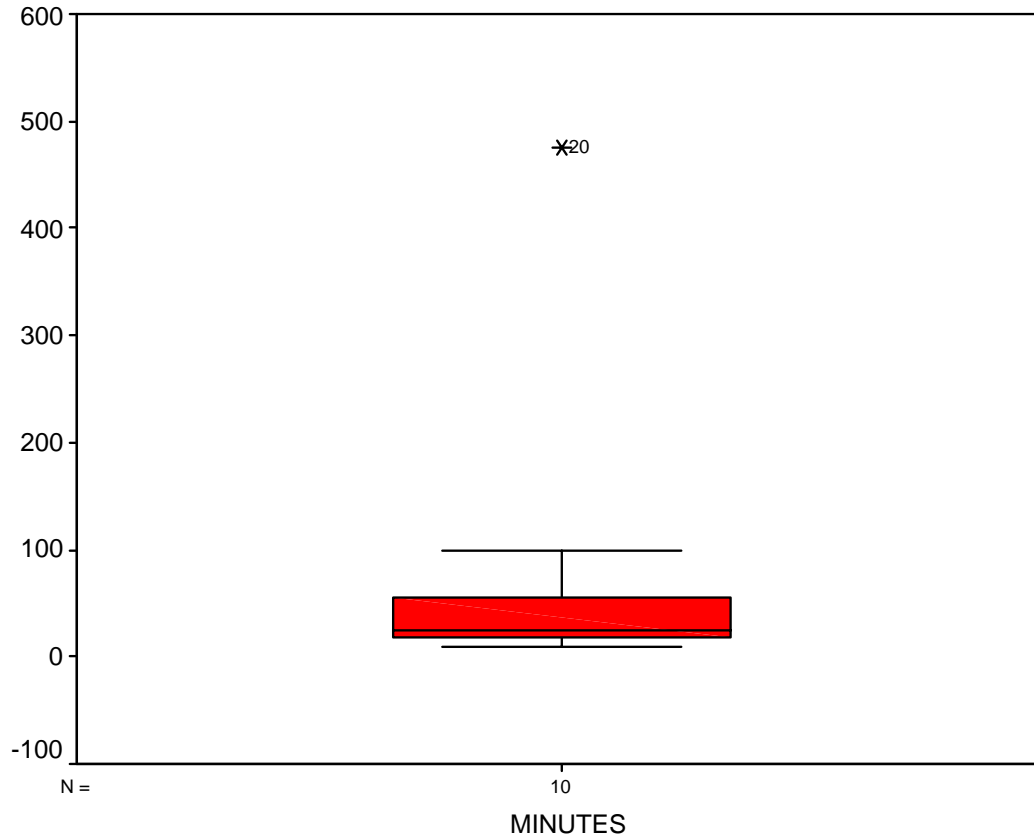
2. As discussed in class, a researcher taught 4 young chimpanzees 10 signs of the American sign language with the intent of determining whether some signs are easier to learn, and whether some chimps tended to learn more quickly than others. The data below are learning times (minutes) for all 4 of the chimps and 10 signs.

a. (4 pts) Find the median, first and third quartiles (Q_1 and Q_3), and IQR, of the learning times for Booe.

Word	Chimpanzee			
	Booe	Cindy	Bruno	Thelma
Listen	12	10	2	15
Drink	15	25	36	18
Shoe	14	18	60	20
Key	10	25	25	40
More	10	15	225	24
Food	80	55	14	190
Fruit	80	20	177	297
Hat	78	99	178	297
Look	115	54	345	420
String	129	476	287	372

$$Q_2 = \frac{15+78}{2} = 46.5; Q_1 = 12, Q_3 = 80, IQR = 68$$

b. (4 pts). Construct a boxplot of Booee's learning time next to the boxplot for Cindy's learning times. Be sure to use the y -axis given in the Figure when drawing the boxplot.



lower fence is located at 10, box is drawn at 12 and 80, the median is at 46.5, and the upper fence is at 129

c). (4 pts) Give numerical summaries of the center and variability for Booee's learning time distribution.

center is 46.5 (median), variability is measured by the $IQR = 168$. Mean and standard deviation are acceptable.

d). (4 pts) Using quantitative information, compare the learning time distribution for Cindy against that for Booee.

First quartiles and lower fences are approximately equal, but the upper fence and third quartiles are smaller form Booee's distribution

e). (4 pts) What other measure of center and variability (besides the ones that you used in d.) could be used?

Mean (or median) for center; std. dev. (or IQR) for variability. Range is a poor measure of spread

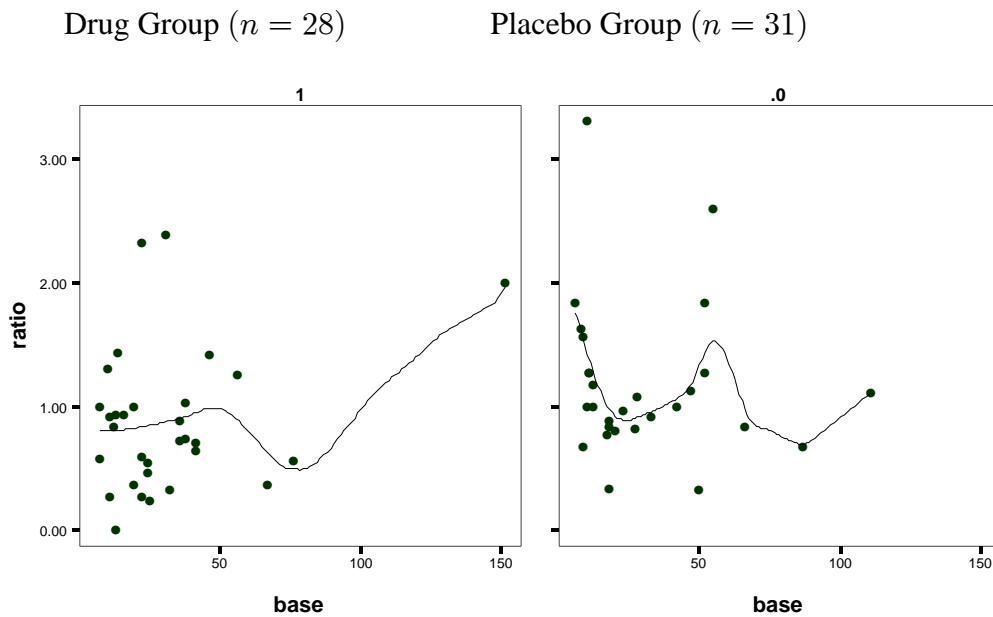
3. Case Study: Thall, P.F. and Vail, S.C. 1990. *Some covariance models for longitudinal count data with overdispersion. Biometrics*, **46**, 657-671. These are data from a study of the effectiveness of a drug for suppressing seizures in epileptics. 59 subjects were randomly assigned to treatment and placebo groups, and each subject reported the numbers of seizures experiences in each of 4 consecutive two-week periods. Also recorded was age and baseline seizure count. The baseline seizure count was the average number of seizures per two weeks, estimated by the patient and doctor.

To adjust for differences among subjects with respect to seizure rate, I have computed the variable $\text{ratio} = \text{total}/\text{baseline}$, and used this variable below.

a) (4 pts) Based on Figure 1, is there a association between the baseline count and the ratio variable? Why or why not?

There is no evidence of association because the points are generally random. Also, there is no real evidence suggesting a trend (up or down) with baseline

Figure 1. Scatter plots of ratio against baseline seizure count, by treatment group.



Using the summary statistics and boxplots on the last page of the Exam,

b) (10 pts) Describe the two sample distributions.

Centers are 1.00 and 0.73 (using the median) for the placebo and drug groups, respectively.

First Quartiles are 0.83 and 1.27 versus 0.50 and 1.01 for the placebo and drug groups, respectively. IQRs are 0.44 and 0.51, respectively

c) (4 pts) Is there evidence of differences between treatment groups with respect to the ratio? Why or why not?

Yes. The median of the drug group (1.00) is nearly equal to the third quartile (1.01) of the placebo group. More generally, there is a shift of the drug group downward relative to the placebo group. Admittedly, variability is rather large compared to the shift, so other interpretations may be valid (provided they are supported by quantitative statements)

Figure 2. Boxplots of ratio, by treatment group (trmt = 1 if the subject was in the drug group, and trmt=0 if the subject was in the placebo group) .

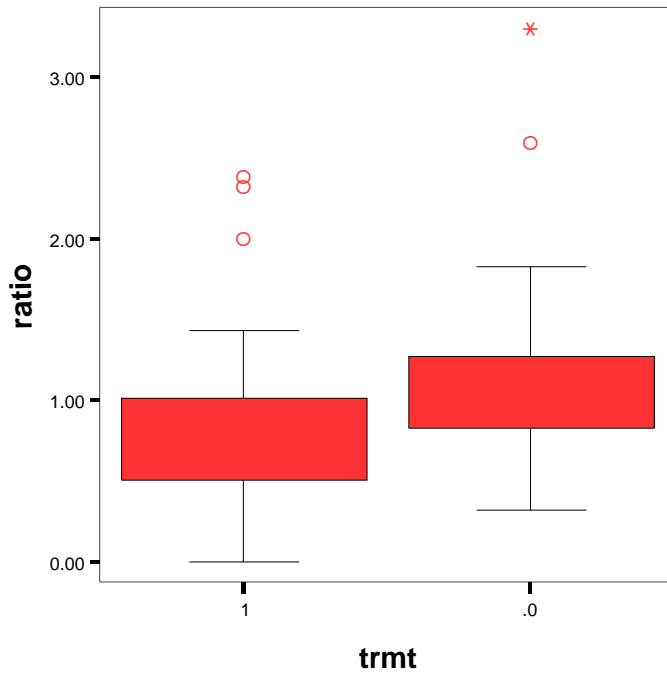


Table 1. Summary Statistics for ratio, by treatment group

	Placebo	Drug
Min:	0.3200000	0.0000000
1st Qu.:	0.8287037	0.5000000
Mean:	1.1724361	0.8695220
Median:	1.0000000	0.7368421
3rd Qu.:	1.2727273	1.0131579
Max:	3.3000000	2.3870968
Total N:	28.0000000	31.0000000
Std Dev.:	0.6297064	0.5784757