

# Statistics 101.002, Fall 2004, Practice Exam A for Midterm 1, 2004

Name (Please print clearly) \_\_\_\_\_

Lab Time (circle one): Thurs. 9:10 – 10:00; Thurs. 10:30 – 11:20; Fri. 9:10 – 10:00; Fri. 10:30 – 11:20

Honor Pledge: I have not given or received assistance on this exam.

Signature: \_\_\_\_\_

Directions:

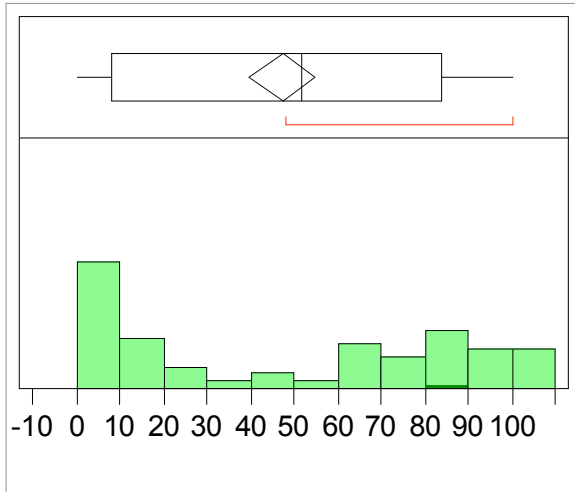
- 1) Print clearly on this exam. Only correct solutions that can be read will be given credit.
- 2) You may use a calculator and 2-pages (with both sides) as crib sheets.
- 3) Show your method of solution on problems requiring calculations. Only answers with supporting work will be given credit.
- 4) Carry out all calculations to 2 decimal places of accuracy. You can leave answers in fractions.
- 5) When you are asked for estimate, we allow a range of responses as correct ones. Estimate as well as you can, making sure to budget your time too.
- 6) As a rough guideline, allot yourself about 10 minutes per page. If you get stuck, move on.

**The data for the problems on pages 2 – 4 of this exam pertain to the impeachment proceedings against President Clinton in 1996 in the Senate.**

<i>Page</i>	<i>Points Possible</i>	Points Scored
2	11	
3	8	
4	9	
5	14	
6	8	

Figure 1 displays ratings of each Senator's degree of conservatism, as determined by the American Conservative Union. The scale goes from 0 to 100, with 0 being extremely liberal and 100 being extremely conservative. (Jesse Helms scored a 100, and Ted Kennedy scored a 4). No scores in the data set go below 0, and no scores go above 100.

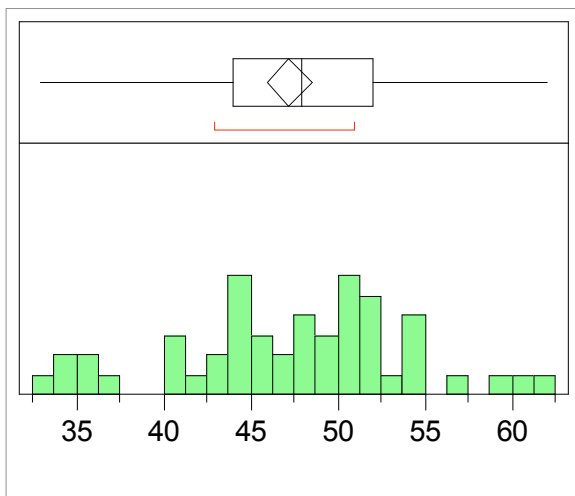
**Figure 1: Conservatism Ratings**



- 1) T F The median is less than the mean.
- 2) Estimate the standard deviation: \_\_\_\_\_
- 3) Estimate the percentage of senators whose rating is less than 10. \_\_\_\_\_
- 4) Estimate the percentage of senators whose rating equals 100. \_\_\_\_\_
- 5) T F A “normal quantile plot” in JMP, referred to as a normal probability plot in the class slides, would show points falling on a straight line.
- 6) Suppose we change the ratings of all the senators with ratings in between 30 and 60. Specifically, we set them equal the average rating in Figure 1. Which one of the following is a consequence of these changes (circle one):
  - a) The standard deviation will increase.
  - b) The standard deviation will decrease.
  - c) The standard deviation will not change.

Figure 2 displays the percentages of votes Clinton received in the 50 states in the 1996 Presidential election. There are 100 data points in this histogram corresponding to the 100 senators. As a reminder, there are two senators per state.

**Figure 2: % of vote in state going to Clinton in election**



- 7) Estimate the average of the state percentages. \_\_\_\_\_
- 8) Estimate the median of the state percentages. \_\_\_\_\_
- 9) Estimate the standard deviation of the state percentages. \_\_\_\_\_
- 10) Estimate the percentage of states that had less than 39% of votes for Clinton. Be very precise with your answer. \_\_\_\_\_
- 11) There are two peaks in the histogram: one around 44 and one around 52. What is the most likely explanation of this fact of the data? You may consider variables not shown in these figures in your answer.

\_\_\_\_\_

**Figure 3: Oneway Analysis of Conservatism By Perjury (1=yes, 0 =no)**

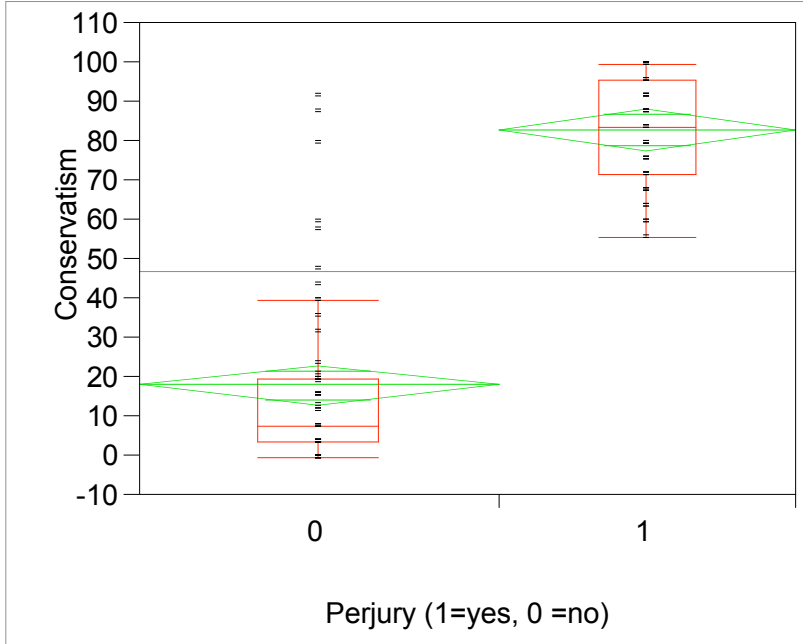
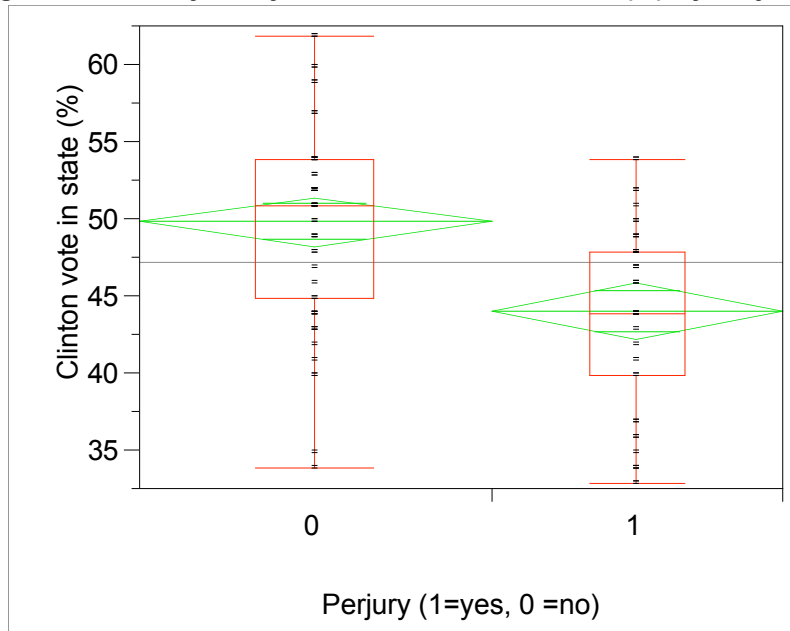


Figure 3 shows plots of conservatism versus vote on whether Clinton was guilty of perjury.

- 12) T F A normal is appropriate for describing the conservatism rates for those who voted no on perjury.
- 13) Of the following choices, circle the letter of the one that is true:
- The standard deviation (SD) of conservatism for those who voted yes is larger than the SD for those who voted no.
  - The SD of conservatism for those who voted yes is the same as the SD for those who voted no.
  - The SD of conservatism for those who voted yes is smaller than the SD for those who voted no.
- 14) T F For senators who voted no, there are more people with conservative ratings above 20 than there are below 10.
- 15) T F This display indicates that senators with high conservative ratings voted yes on perjury more frequently than did senators with low conservative ratings.

Figure 4 shows plots of percentage of people voting Clinton versus perjury vote.

**Figure 4: Oneway Analysis of Clinton vote in state (%) By Perjury (1=yes, 0 =no)**



16) Estimate the difference in average Clinton vote percentage between senators who said yes and senators who said no (yes group – no group). \_\_\_\_\_

17) Check all that apply for Figure 4:

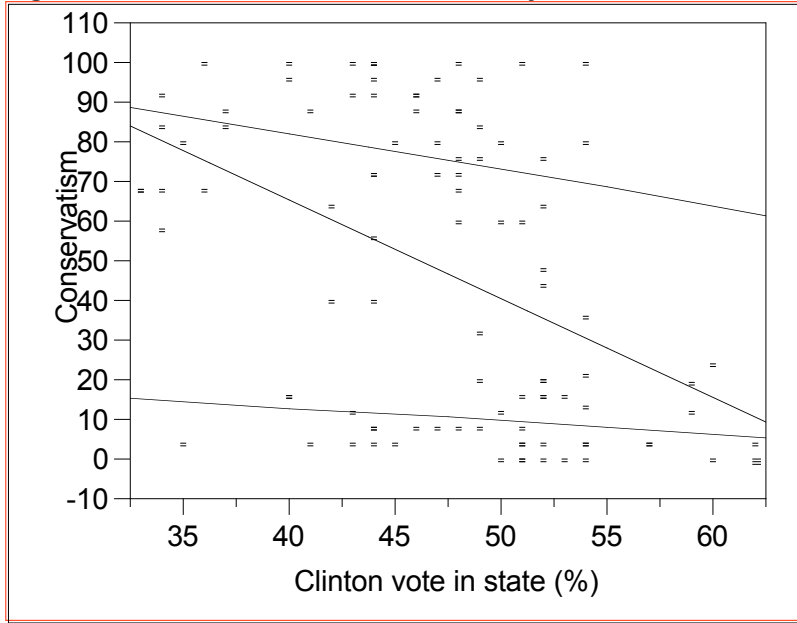
- The ratio of the larger to smaller SD is less than 1.5.
- The ratio of the larger to smaller SD is at least 1.5.
- The percentages are typically larger for the no group than the yes group.
- The percentages are typically smaller for the no group than the yes group.

18) Which of the two variables, % of Clinton votes or conservatism rating, is a better predictor of perjury vote?

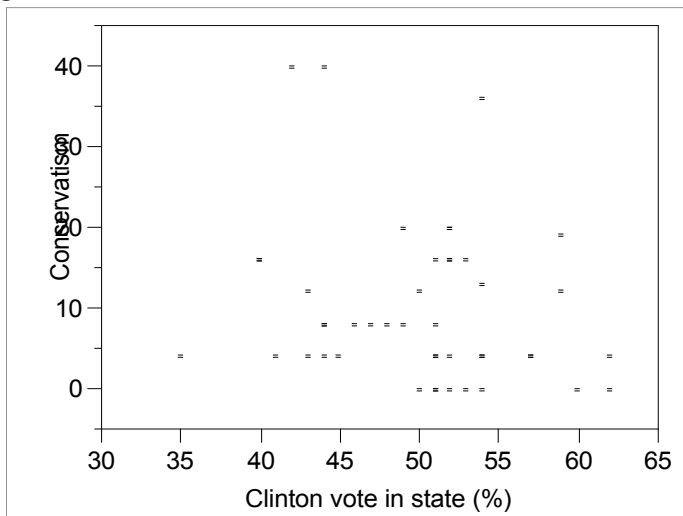
\_\_\_\_\_

Figure 5 shows the relationship between conservatism and percentage in state who vote for Clinton. Three regression lines are drawn to the data. One of the lines corresponds to the regression of conservatism on all Clinton vote for all 100 senators. One line corresponds to the regression only for Democratic senators. One line corresponds to the regression only for Republican senators.

**Figure 5: Bivariate Fit of Conservatism By Clinton vote in state**



**Figure 6: Conservatism versus Clinton vote for Democrats only**



19) You can use any answers or graphs from previous problems, as well as both Figure 5 and Figure 6 for these problems. Additionally, the average and SD of conservatism rating for Democrats equal 10 and 10, and the average and SD of conservatism rating for Republicans equal 78 and 20.

- On the graph to the left, identify each line as one of the following: “Democrats”, “Republicans”, “All senators”. Write the appropriate label over each line.
- Estimate the correlation between conservatism and Clinton vote based on all 100 senators \_\_\_\_\_
- T F States with senators that rate high on the conservative scale tended to have high percentages of people voting for Clinton.
- Estimate the average and standard deviation for Clinton vote percentage for Democratic senators only.  
Avg: \_\_\_\_\_ SD: \_\_\_\_\_
- Estimate the correlation between conservatism and Clinton vote percentage for the Democratic senators. \_\_\_\_\_
- Estimate the slope of the regression line for conservatism (y) on Clinton vote percentage (x) for Democrats only as accurately as possible. \_\_\_\_\_
- Estimate the intercept of the regression line for conservatism (y) on Clinton vote percentage (x) for Democrats only as accurately as possible. \_\_\_\_\_
- T F Because the Democratic senators are a sample of all senators, their average conservatism rating of 10 is subject to chance error.

20) For each of the following statements, decide whether the statement is true or false. If you think the statement is true, justify why it is true. If you think the statement is false, justify why it is false. Write at most three sentences for your answer (graders will read only up to the first three sentences).

a) (3 points) Scores on the SAT exam are designed to follow a normal curve with mean of 1000 and standard deviation of 100. Using this distribution, the percentage of scores above 1140 equals 3.5%. Justify your answer with calculations.

b) (3 points) For the SAT exam described in part (a), what score equals the 90<sup>th</sup> percentile?

c) (2 points) If the test is changed so that the standard deviation of scores decreases but the average score remains the same, the percentage of scores above 1140 increases. (Be sure to read correctly what is increasing and what is decreasing!)

d) (3 points) You throw two fair, six-sided dice. What is the chance that their sum will be less than five and even?

e) (3 points) M&M's candies come in colors, and 30% of all M&M's produced at the Mars Candy factory are brown. Bags of M&M's are filled randomly with M&M's, so that colors are well-mixed. The chance that at least one of the first ten M&Ms in a bag will be brown equals 0.3. (Calculate the probability as justification for your answer.)

- 21) For each problem below, you are given a method for collecting data and a corresponding conclusion. If the conclusions are justifiable from the method of data collection, explain why you think the method of data collection is effective. If the conclusions are not justifiable from the method of data collection, say why you think the method of data collection is ineffective. Don't assume that all the study designs below are ineffective; some may be perfectly reasonable.

For the surveys, assume there is no ambiguity in questions. For the causal studies, assume the definition of any treatments and response is specific. In other words, don't worry about the method of stating the question, just focus on the data collection. Assume sample sizes are sufficiently large.

Write at most three sentences for your answer (graders only will read up to the first three sentences).

a) (4 points) A headline in *USA Today* in 1998 announced that “Prayer can lower blood pressure”. The opening sentence of the news story stated, “Attending religious services lowers blood pressure more than tuning in to religious TV or radio, a new study says.” The news story was based on a journal article (Davis, 1998) describing a study done by the National Institute of Health. The study followed 2391 people aged 65 or older for six years. One of the principal findings in the article was that, “People who attended a religious service once a week and prayed or studied the Bible once a day were 40% less likely to have high blood pressure than those who don't go to church every week and prayed and studied the Bible less” (Davis, 1998). Comment on the whether the news paper is correct; that is, comment on whether this study says that prayer reduces blood pressure.

b) (4 points) A researcher conducted an experiment to see if her meditation technique helps people relax. First, she interviewed every person in the study to record their background characteristics. Then, she randomly assigned the subjects to two groups. The researcher taught one group her meditation technique for relaxation, and simply told the other group to relax the way they usually relax. After each person used his or her assigned relaxation technique, the researcher re-interviewed the person and rated his or her relaxation level. Based on the ratings, the group that used meditation was significantly more relaxed than the group that did not use meditation. The researcher attributes the difference to the effect of her meditation technique.

