

STAT303: Secs 508 - 510

Spring 2005

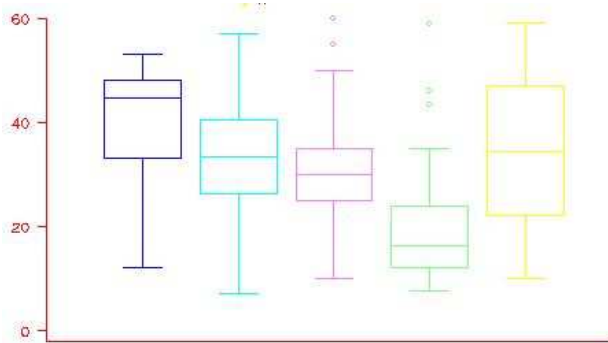
Exam #1

Form A

Instructor: Julie Hagen Carroll

1. **Don't EVEN open this until you are told to do so.**
2. Be sure to mark your **CORRECT** section number and your test form on the scantron!
3. Sign your name where indicated on your scantron. You may keep this exam.
4. There are 20 multiple-choice questions on this exam, each worth 5 points. There is partial credit. Please mark your answers **clearly** on the scantron. Multiple marks will be counted wrong.
5. You will have **60 minutes** to finish this exam.
6. This exam is worth 100 points, and will constitute 20% of your final grade.
7. If you are caught cheating or helping someone to cheat on this exam, you both will receive a grade of **zero** on the exam. You must work alone.
8. Good luck!

- We know that the distribution of our dataset is bell-shaped, the mean, $\bar{x} = 14$ and the standard deviation, $s = 6$. Which of the following statements would be true about our data?
 - The median, \tilde{x} , would also be 14.
 - The IQR would also be 6.
 - None of the observations would be negative.
 - All of the above are true.
 - Only two of the above are true.
- Which of the following graphs would have told us our data was normal (bell-shaped)?
 - A normal quantile plot in which the points fell along a bell curve.
 - A boxplot with the sections equally spaced to represent the 1, 2 and 3 standard deviations.
 - A stemplot with the middle row being the longest and the other rows progressively shorter as they got further from the middle.
 - A histogram with a peak, not flat.
 - All of the above would indicate the data is approximately normally distributed.
- In an experiment, an observed effect so large that it would rarely occur by chance is called
 - an outlier.
 - influential.
 - statistically significant.
 - a bias on the data.
 - Exactly two of the above are correct.



- Numbering from 1 to 5, left to right, which of the boxplots above most likely has the smallest standard deviation?
 - 1
 - 2
 - 3
 - 4

- 5
- Still referring to the boxplots, which has a mean about 40?
 - You can't tell the mean from a boxplot, only the median.
 - 1
 - 2
 - 3
 - 5
 - A simple random sample of size n is defined to be
 - a sample of size n chosen in such a way that every value in the population has the same (equal) chance of being chosen.
 - a sample of size n chosen in such a way that every unit (observations) in the population has the same (equal) chance of being chosen.
 - a sample of size n chosen in such a way that every value in the population has a known probability (chance) of being chosen.
 - a sample of size n chosen in such a way that every sample of size n units (observations) in the population has the same (equal) chance of being chosen.
 - All of the above. They are essentially the same.
 - If you scored at the 90th percentile on the SAT, this would mean
 - you scored 90% better than everyone else who took the exam.
 - you scored better than 90 another people who took the exam.
 - you got 90% correct on the exam.
 - 90% of the people who took the exam got the same amount correct as you did.
 - you scored better than 90% of the people took the exam.

	college	AGLS	grade			Total	
			A	B	C		
		10	8	2		20	
		ARCH	3	1	0		4
		EDUC	8	13	2		23
		ENGR	4	1	0		5
		GEOS	1	0	1		2
		LA	12	9	2		23
		SCI	6	2	0		8
		VMED	12	3	0		15
Total		56	37	7		100	

8. Which of the following best describes $12/23$ as a percentage of the data in the table above?

- of all the students in the College of Liberal Arts in the survey, the proportion who expect to get an A
- the proportion of A students in the College of Liberal Arts
- the proportion of all the students in the survey who are in the College of Liberal Arts and expect an A
- the proportion of all Liberal Arts students in the survey who expect an A
- of all students in the survey who expect an A, the proportion who are in the College of Liberal Arts

9. Does it appear that one's college and one's expected grade are related?

- Yes, more Education majors expect a B than any other major.
- No, an A is the 'most popular' expected grade in almost majors.
- No, A is the 'favorite' and C is the least 'favorite' in all majors.
- Yes, the number of students vary by college for each grade.
- Yes, the proportion of expected A's is not the same for each college.

10. Which of the following is not a statistic?

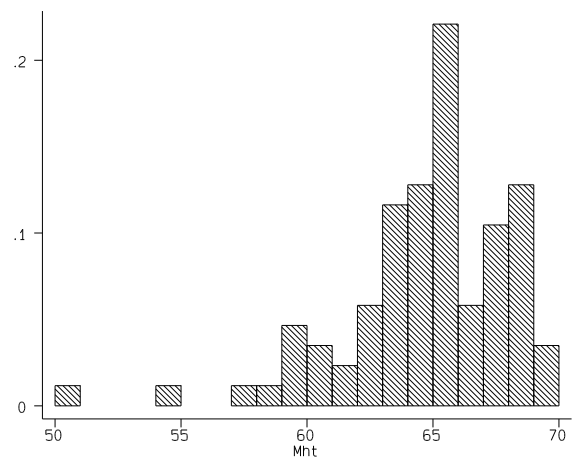
- the average height of this class
- the proportion of students in this class that are Education majors
- the correlation between the ages of the moms and dads for this class
- the average number of points A&M scores in a football game
- All of the above are statistics.

11. A study found a correlation coefficient, $r = 0.0$ between the first and second exam of a STAT30X class. You conclude that

- There is no relationship between the exams.
- The mean score of each exam is 0.0.
- The mean difference between the exams is 0.0.
- Using the first exam in a least squares equation is not useful for predicting the second exam.
- None of the above are true.

12. Suppose I decided to add 5 points to everyone's quiz grade. Which of the following would be true?

- The mean number of questions answered correctly, \bar{x} , would increase by 5.
- The mean grade on the quiz, \bar{Y} , would increase by 5.
- The standard deviation for the grade on the quiz, s_Y , would increase by 5.
- Everyone's z -score for the quiz would increase by 5.
- Exactly two of the above are true.

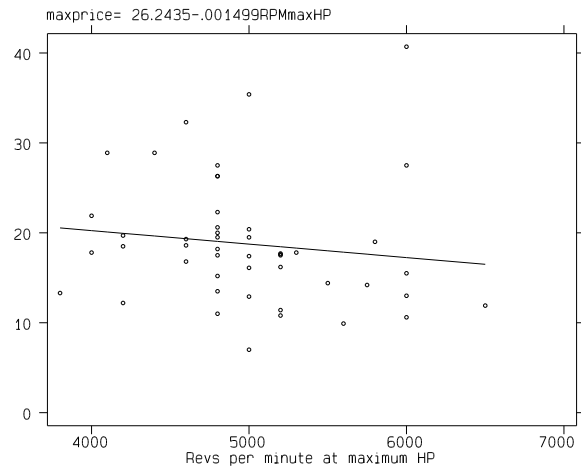


13. Which of the following is true about the data in the histogram above?

- The median, \tilde{x} , is about 60 and the mean, \bar{x} is something less than 60.
- The median, \tilde{x} , is about 65 and the mean, \bar{x} is something less than 65.
- The mean, \bar{x} , is about 60 and the median, \tilde{x} is something less than 60.
- The mean, \bar{x} , is about 65 and the median, \tilde{x} is something less than 65.
- We cannot conclude anything about the mean from a histogram.

14. Referring to the dataset in the last question, the histogram, what would happen if the smallest point, say 50, was changed to 55?
- Both the mean and median would increase by 5.
 - Both the mean and the standard deviation would increase by 5.
 - Only the mean would increase by 5.
 - Only the mean would increase.
 - None of the statements above are completely true.
15. The least squares regression line
- goes through all the data points.
 - is the line that minimizes the vertical distances of the data points to the line.
 - is the line that maximizes the vertical distances of the data points to the line.
 - is the line which best splits the data in half, with half of the points above the line and half below the line.
 - Exactly two of the above are correct statements.
16. What is a confounding variable?
- It is a variable that we should have measured but didn't.
 - It is a variable that so closely tied to the explanatory variable that we can't figure out which one caused the response.
 - It is a variable that so closely tied to the response variable that we can't prove causation.
 - It is another explanatory variable that we forgot to measure.
 - It is a variable that is only used in case control studies.
17. Consider a set a bivariate data, (x_i, y_i) , of size n . If we *add* 5 to each of the x 's, then which of the following would be true?
- The correlation coefficient for x and y would change.
 - The slope of the linear regression line would be smaller by a factor of 5.
 - The intercept of the linear regression line would be 5 points bigger.
 - All of the above will change.
 - None of the above will change.

18. Did you get cheated in your box of Valentine chocolates? The box claims to have 16 oz. of chocolate in it (so that's the true mean) and we're told the true standard deviation for these boxes is 0.5 oz. plus the distribution of the chocolate boxes is practically normal. Your box had only 14 oz. in it.
- Why YES, it's impossible to get a box that light if the numbers and distribution are correct.
 - Yes, there's less than a 0.15% chance of getting a box that light.
 - No, you just were unlucky, but someone should have gotten a box with 18 oz. in it.
 - No, since the weights are random, you could have easily have gotten a box that light.
 - Actually, someone ate a piece before you got it, so it really weighed 16 oz. originally.



19. What is the approximate correlation in the plot above?
- strongly negative
 - moderately negative
 - weak
 - moderately positive
 - strongly positive
20. If the point (6000, 40) were removed from the dataset in the last problem, what would change?
- The correlation would improve slightly.
 - The slope of the line would become positive.
 - The intercept would increase noticeably.
 - All of the above are happen.
 - Only two of the above would happen.

1A,2C,3C,4C,5B,6D,7E,8A,9E,10E,11D
12B,13B,14D,15B,16B,17E,18B,19C,20A