

STAT303: Secs 506—508

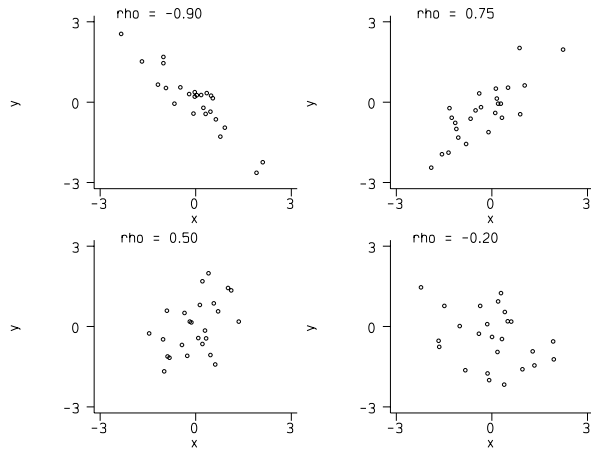
Fall 1998

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 section number and your test form (A, B, C or D) on the scantron!
3. Sign your name where indicated on your scantron and write your section number and computer number beside it. Also, you must place your scantron in the correct 'Return To' section stack (for next Thursday).
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. 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.
7. This exam is worth 100 points, and will constitute 25% of your final grade.
8. Good luck!



1. Labelling the plots above A, B, C and D left to right, top row then bottom row, which plot has the strongest correlation?

- A. A
- B. B
- C. C
- D. D
- E. You can't tell just from the scatterplot.

2. Find z^* such that $P(Z > z^* \text{ or } Z < -z^*) = 0.30$?

- A. $z^* = 0.52$
- B. $z^* = 0.6179$
- C. $z^* = 0.5596$
- D. $z^* = -0.52$
- E. $z^* = 1.03$

3. Suppose the heights of the males in this class are normally distributed (heights really are normal). If one of the students has a z -score of 0 for his height, this means

- A. His height is 5'10" since that's the true mean height of all men.
- B. His height is 6' even.
- C. All of the heights of the males in the class are the same.
- D. He is the average height of the males in the class.
- E. He is the average male in the class.

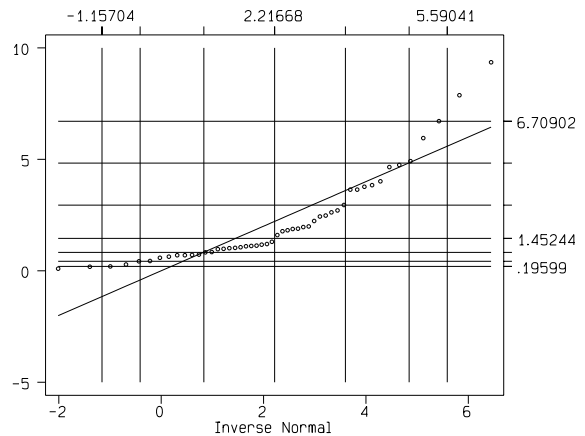
4. Let $Z \sim N(0, 1^2)$. What is $P(Z > 3.01)$?

- A. 0.0010
- B. 0.0013
- C. 1.88
- D. 0.9990
- E. 0.9987

5. Suppose we are calculating the average January temperature for the state of Alaska. All daily temperatures are below 0 degrees C, *i.e.*, they are all negative. If we came up with an IQR of -18 degrees, then

- A. the distribution of these temperatures is skewed negatively (to the left).
- B. the mean is less than the median.
- C. we made a mistake.
- D. Both A and B are true.
- E. None of the above are true.

(Grid lines are 5, 10, 25, 50, 75, 90, and 95 percentiles)

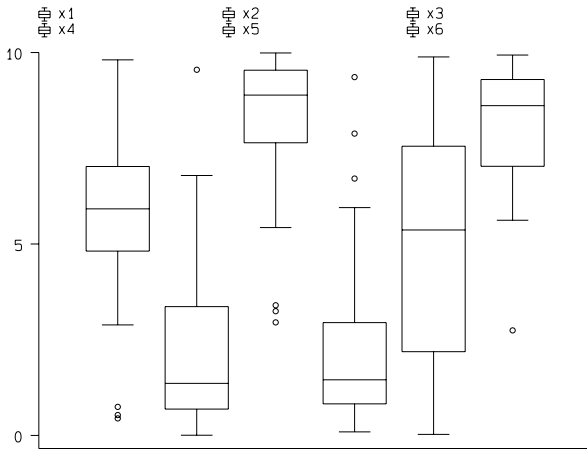


6. Assuming that the mean, \bar{x} and standard deviation, s_x are both about 2, which of the following statements is true?

- A. Approximately 95% of the observations are between -2 and +2.
- B. Approximately 95% of the observations are between -2 and +6.
- C. At least 75% of the observations are between -2 and +2.
- D. At least 94% of the observations are between -6 and +10.
- E. Approximately 68% of the observations are between -2 and +2.

7. Let X be a continuous uniform random variable like problem 6.12 on your homework, so the range of X is 0 to 10 and the mean is 5. What is $P(X = 5)$?

- A. 0 because 5 is the mean
- B. 0.5 because 5 is the mean
- C. 0.5 because 5 is the median
- D. 0.5 because X is continuous
- E. 0 because X is continuous



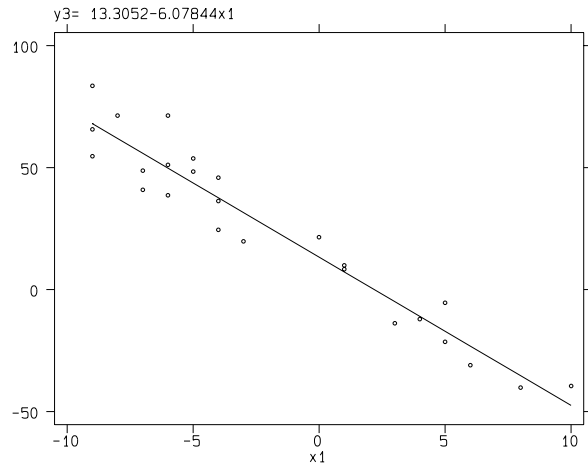
8. Which of the boxplots above has the median, \hat{x} , less than the mean, \bar{x} ?
- A. 2 and 4
 - B. 3 only
 - C. 3 and 6 only
 - D. 1, 3 and 6
 - E. 1, 3, 5 and 6

9. Again referring to the boxplots above, which one has an IQR of about 3?
- A. You can't tell just looking at the boxplots.
 - B. 1
 - C. 2
 - D. 3
 - E. 4

	Percentiles	Smallest
	1%	2.74105
	5%	5.73632
	10%	6.13405
	25%	7.028718
	50%	8.613952
	75%	9.294461
	90%	9.675049
	95%	9.726476
	99%	9.934505
		Largest
		9.709767
		9.726476
		9.79606
		9.934505

10. The output above describes which boxplot?
- A. 2
 - B. 3
 - C. 4
 - D. 5
 - E. 6

11. The standard deviation can be thought of as how far an observation is from the mean on average. Using this idea, which boxplot has the largest standard deviation, *i.e.*, which has the biggest spread for most of the points?
- A. 1
 - B. 2
 - C. 3
 - D. 4
 - E. 5



12. Assuming (0,25) is the point (\bar{x}, \bar{y}) , which of the following points will have the most effect on the correlation coefficient, r ?
- A. (0,25)
 - B. (0,50)
 - C. (50,0)
 - D. (0,10)
 - E. B and C would have about the same effect.
13. The regression equation for the data above is approximately $\hat{y}_3 = 13.3 - 6.1 * x_1$. What is the residual for the point (25,0)?
- A. 0 since that point is the mean of the data
 - B. 139.2
 - C. -139.2
 - D. 11.7
 - E. We can't guarantee that the equation is still valid at (25,0), so we can't figure a residual.
14. Again, let $X \sim N(6, 4^2)$. If $Y = (1/2)X + 2$, what are μ_Y and σ_Y ?
- A. $\mu_Y = 6, \sigma_Y = 4$
 - B. $\mu_Y = 5, \sigma_Y = 2$
 - C. $\mu_Y = 5, \sigma_Y = 4$
 - D. $\mu_Y = 5, \sigma_Y = 6$
 - E. $\mu_Y = 3, \sigma_Y = 2$

MasterCard			
Visa	Yes	No	Total
Yes	5000	2000	7000
No	1000	2000	3000
Total	6000	4000	10000

15. The table above shows the number of students (out of a total of 10,000) with a MasterCard, a Visa card or both. Which of the following are true?
- Having a MasterCard and a Visa card are mutually exclusive events since you don't need but one or the other.
 - Having a MasterCard and a Visa card are independent events since it doesn't matter which one you have.
 - Having a MasterCard and a Visa card are not independent events since you are more likely to have a MasterCard if you have a Visa card than you are if you don't have a Visa card.
 - Having a MasterCard and a Visa card are mutually exclusive events since it doesn't matter which one you have.
 - Having a MasterCard and a Visa card are independent events since you are just as likely to have or NOT have a Visa card if you DON'T have a MasterCard.
16. Let $X \sim N(6, 4^2)$, *i.e.*, X is normally distributed with a mean $\mu_X = 6$ and a standard deviation, $\sigma_X = 4$. What is $P(X < 8)$?
- 0.5
 - 0.6915
 - 0.3085
 - 0.5
 - 0.9772
17. For the same $X \sim N(6, 4^2)$, what is the 75th percentile, *i.e.*, what is x^* such that $P(X < x^*) = 0.75$?
- 0.75
 - 0.675
 - 8.7
 - 3.3
 - 2.7
- | X | -1 | 0 | 1 | 3 |
|--------|-----|-----|-----|-----|
| $p(X)$ | 0.5 | 0.2 | 0.2 | 0.1 |
18. What are the mean, μ_X and standard deviation, σ_X , for the random variable, X , whose distribution is as above?
- $\mu_X = 0.5, \sigma_X = 1.6$
 - $\mu_X = 1.6, \sigma_X = 1.26$
 - $\mu_X = 0, \sigma_X = 1.6$
 - $\mu_X = 0, \sigma_X = 1.26$
 - $\mu_X = -0.5, \sigma_X = 1.6$
19. Referring to the random variable, X , above, which of the following is true?
- The distribution is skewed positively (to the right), and the median, $\tilde{X} = -0.5$ is less than the mean, μ_X .
 - The distribution is skewed positively (to the right), and the median, $\tilde{X} = 0.5$ is less than the mean, μ_X .
 - The distribution is skewed positively (to the right), and the median, $\tilde{X} = 0.5$ is greater than the mean, μ_X .
 - The distribution is skewed negatively (to the left), and the median, $\tilde{X} = -0.5$ is less than the mean, μ_X .
 - The distribution is skewed negatively (to the left), and the median, $\tilde{X} = 0.5$ is greater than the mean, μ_X .
20. If we took a sample of 25 from the distribution above, multiplied each observation by 5, which of the following would be true?
- The mean would be approximately 0.
 - The mean would be approximately 5.
 - The mean would be approximately equal to the median.
 - The mean would be less than the median.
 - Exactly two of the above are true.
21. Which of the following statements about bivariate data is/are true?
- If Spearman's and Pearson's correlation coefficients are both about -0.9, then we can assume that the data strongly linear.
 - Adding 5 to all of the y values will cause the slope to change but not the intercept in the regression equation.
 - Multiplying all of the x values by 10 will cause the intercept to change but not the slope in the regression equation.
 - All of the above are true statements.
 - None of the above are true statements.

1A,2E,3D,4B,5C,6D,7E,8A,0C,10E,11E,
12C,13E,14B,15C,16B,17C,18D,19A,20A,21A