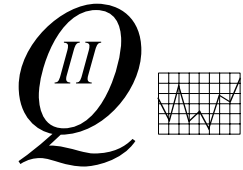


Exam 2, Mar. 29, 1995, Quant II



The exam questions and answer sheet are both to be turned in to your Discussion Section instructor at the end of the exam. Please mark all of your answers on the answer sheet using a #2 pencil. **Code your name, ID number, and Section number on the answer sheet.**

1. If an experimental treatment is deemed to have harmful side effects, it is called a placebo.
A) True B) False
2. If data have a very strong upward but curved relationship, it is possible to get a correlation coefficient that is larger than +1.
A) True B) False
3. In a factorial experiment all treatments combinations are run in one replication.
A) True B) False
4. A class has asked their instructor to “grade on the curve.” With this system the instructor is required to give preselected percentages of the various possible grades. In particular, the lowest 10% of the class **must receive** F’s. If exam scores are normally distributed with mean 83.0 and standard deviation 2.34, what exam score corresponds to the highest F grade?
A) 50
B) 60
C) 70
D) 78
E) 80
5. Correlations based on pairs of averages or totals of smaller units of study are usually larger in magnitude than if they were based on individual measurements.
A) True B) False
6. Autocorrelation documents the correlation of a time series with its own past history.
A) True B) False
7. In general, the larger the residual standard deviation, the better the regression model.
A) True B) False
8. In multiple regression modeling, the most basic residual plot is the plot of residuals versus the corresponding fitted values.
A) True B) False

9. A straight-line has been fit to model the relationship between Salary and Experience for 20 people. The residual plot is shown below. Points plotted for males use M as a plotting symbol. Females are plotted with an F.



Which of the following statements best describes this plot?

- A) The randomness shown in the plot indicates a good model.
 - B) The plot shows that normality is a reasonable assumption.
 - C) The plot shows that a quadratic model should be considered.
 - D) The plot shows that least squares is an excellent criterion for fitting the line.
 - E) The plot indicates that a parallel-lines model would explain the relationship better.
10. We have 17 pairs of x-y data and are fitting a quadratic regression curve by least squares. How many degrees of freedom do the residuals have?
- A) 14
 - B) 15
 - C) 16
 - D) 32
 - E) 33
11. A time series consists of the values 5, 9, 7, 4, 5 in that order. What is the value of the lag 1 autocorrelation coefficient?
- A) 0
 - B) 0.1
 - C) 0.2
 - D) 0.3
 - E) 0.4

12. In general, the smaller the adjusted R^2 value the better the regression model.
A) True B) False
13. The model $GP100M = -1.3 + 0.0025\text{Weight} + 0.7\text{Trans}$ is obtained from fitting Gallons per 100 miles to Weight and Transmission type for a number of cars. Here Weight is in pounds and Transmission type is an indicator variable with $\text{Trans} = 1$ for automatic transmissions and $= 0$ for manuals. A new car, the Ford Contour, comes on the market. It weighs 2910 pounds, has a 4 cylinder engine and a 4 speed automatic transmission. What does the regression model predict for the Ford Contour in usual terms of Miles per Gallon (*not* gallons per 100 miles)?
A) 6
B) 7
C) 15
D) 16
E) 17
14. In an experiment, the control group receives the new treatment.
A) True B) False
15. We have 27 triples of x-y-z data (x and y are continuous variables and z is an indicator variable) and are fitting a parallel-lines model by least squares. How many degrees of freedom do the residuals have?
A) 24
B) 25
C) 26
D) 27
E) None of the above.
16. A lurking variable is a variable that plausibly explains a relationship between two other variables.
A) True B) False
17. In multiple regression modeling, residual plots containing nonrandom patterns indicate that the model can be improved.
A) True B) False

18. The following table shows the responses of ten students to questions about their gender and the difficulty of an exam.

Student ID	Gender [†]	Difficulty [‡]		Student ID	Gender	Difficulty
1	0	2		6	0	2
2	1	3		7	0	1
3	1	1		8	1	3
4	1	2		9	0	1
5	0	1		10	1	2

[†] 1 = female, 0 = male

[‡] 1 = easy, 2 = moderate, 3 = hard

What percent of students thought the exam was easy?

- A) 10%
- B) 20%
- C) 30%
- D) 40%
- E) 50%

19. Referring to the table in question 18: What percent of females thought the exam was hard?

- A) 20%
- B) 40%
- C) 60%
- D) 80%
- E) 100%

20. Referring to the table in question 18: What percent of those who thought the exam was easy were males?

- A) 50%
- B) 60%
- C) 65%
- D) 70%
- E) 75%

21. The table below shows four data pairs together with some partial results on fitted values and residuals for two possible models—one linear and one quadratic. *These models were not necessarily found using least squares.*

Data			Curve I $\hat{y} = 0.7x$		Curve II $\hat{y} = 1 - 0.5x + 0.25x^2$	
y	x	x^2	FITTED	RESIDUAL	FITTED	RESIDUAL
1	1	1	0.7	0.3	0.75	0.25
1	2	4	1.4	-0.4		0.00
2	3	9	2.1	-0.1	1.75	0.25
3	4	16	2.8		3.00	

Which curve fits the data better in the sense of least squares?

- A) Curve I fits better since its residuals add to zero.
 - B) Curve I fits better since its sum of squared residuals is smaller than for curve II.
 - C) Curve I fits better since it is the least squares regression line for these data.
 - D) Curve II fits better since its sum of squared residuals is smaller than for curve I.
 - E) Curve II fits better since one of its residuals is zero.
22. Each year the General Social Survey, GSS, selects 1500 households throughout the U.S. using complicated randomization methods. People in the selected households are interviewed extensively to gather social science data. Such a study is best classified as
- A) a designed experiment with blinding of the subjects
 - B) an observational study
 - C) a designed study
 - D) a double blind, designed experiment
 - E) a designed experiment with neither blinding nor double blinding
23. A factorial experiment has two factors. The first factor has two levels and the second factor has three levels. If the experiment is replicated twice, how many observations will be recorded?
- A) 2
 - B) 3
 - C) 5
 - D) 6
 - E) 12

24. In the Physicians' Health Study 22,000 male doctors were used as subjects to measure the effect of a simple drug in the prevention of heart attacks. In this study aspirin was used as a placebo.
- A) True B) False
25. The Physicians' Health Study used double blinding.
- A) True B) False
26. The Physicians' Health Study failed to use randomization to decide which subjects received the real drug and which received the placebo.
- A) True B) False
27. A prediction based on a regression line is likely to be very precise when
- I. the residual sum of squares is small
 - II. the residual standard deviation is small
 - III. the value of R^2 is close to 100%
- A) I only
B) II only
C) III only
D) I, II, and III
E) None of the above
28. What is the area under the standard normal curve between the numbers -1.23 and $+1.19$?
- A) 0.1093
B) 0.7737
C) 0.8830
D) 0.9923
E) None of the above.

29. The table below displays counts of people by gender and marital status.

Gender	Marital Status		
	Single	Married	Other
female	200	50	50
male	100	60	40

What percent of the married people are female? (Round to the nearest whole percent.)

- A) 17%
- B) 33%
- C) 45%
- D) 50%
- E) 60%

30. Referring to the table in Question 29, what percent of the people are single? (Round to the nearest whole percent.)

- A) 17%
- B) 33%
- C) 45%
- D) 60%
- E) 67%

31. Consider normal scores calculated for a sample of 28 highly skewed observations. An analyst (mistakenly?) displays a stem-and-leaf diagram for the 28 normal scores values. Which one of the following describes a correct reaction to the shape of the stem-and-leaf diagram?

- A) The analyst is *not* surprised that the stem-and-leaf diagram shows the skewness in the data.
- B) The analyst *is* surprised that the stem-and-leaf diagram shows the skewness in the data.
- C) The analyst is *not* surprised that the stem-and-leaf diagram looks very normal.
- D) The analyst *is* surprised that the stem-and-leaf diagram looks very normal.
- E) The analyst *is* surprised that normal scores can be calculated for skewed data.

32. A truck carries 40 standard-sized containers in a load. The weights of the containers vary according to many factors but may be described by a distribution with mean 230 pounds and standard deviation 40.8 pounds. Over many loads, what percent of loads will exceed the legal load limit of 10,000 pounds?
- A) 1%
 - B) 5%
 - C) 10%
 - D) 20%
 - E) None of the above.
33. The times necessary to complete service for a class of bank customers is described by a normal distribution with mean 15 minutes and standard deviation 2.1 minutes. Service times are considered excessive if they exceed 20 minutes. Over the long run, what percent of customers will experience excessive service times?
- A) 0.87%
 - B) 2.38%
 - C) 97.62%
 - D) 99.13%
 - E) None of the above.
34. Hawkeye Supply Company has randomly selected 100 steel bolts from a large shipment. Suppose that the bolt lengths in the shipment may be described by a distribution with mean 3 inches and standard deviation 0.1 inches. Let \bar{y} denote the average bolt length for the sample of 100 bolts. If the sampling were repeated many times, what fraction of the averages, \bar{y} , would be less than 2.98 inches?
- A) 0.0227
 - B) 0.4207
 - C) 0.5793
 - D) 0.9773
 - E) None of the above.
35. What is the 90th percentile of a normal distribution with mean 75 and standard deviation 8.6?
- A) 1.28
 - B) 10
 - C) 86
 - D) 90
 - E) 92

36. The text describes the relationship between inflation rate, y , and an index of independence, x , of central banks in 16 countries. Calculation produced the values: $\bar{x} = 2.359$, $s_x = 0.837$, $\bar{y} = 5.588$, $s_y = 1.629$, and $r = -0.843$. Use the least squares regression line to answer the following question. If the laws of a country were changed so that the index of independence of central bank were to increase by one unit, what would we predict for the **change** in the inflation rate for that country?
- A) Since the slope of the regression line is -1.64 , we predict that the inflation rate would go **down** by 1.64 percentage points.
 - B) Since the slope of the regression line is -1.64 , we predict that the inflation rate would go **up** by 1.64 percentage points.
 - C) Since the constant term in the regression line is 9.46, we predict that the inflation rate would go **down** by 9.46 percentage points.
 - D) Since the constant term in the regression line is 9.46, we predict that the inflation rate would go **up** by 9.46 percentage points.
 - E) Since we are not told whether the Index is changing from 1.5 to 2.5 or 2 to 3, we cannot make a sensible prediction.

37. What is the correlation coefficient between the x - y pairs shown below?

x	y						
1,000,000	2						
1,000,001	1						
1,000,002	3						

- A) 0.00
 - B) 0.33
 - C) 0.50
 - D) 0.75
 - E) 1.00
38. Outliers have a strong impact on regression analysis.
- A) True
 - B) False

39. The sequence plot of a time series of 48 monthly values is shown below. The lag 1 autocorrelation coefficient for this series is best described as: (Hint: What would the plot of the series versus the lag 1 of the series look like?)



- A) Strongly negative
- B) Moderately negative
- C) Near zero
- D) Moderately positive
- E) Strongly positive

40. The normal probability plot for the Iowa city sizes data is shown below.



This plot shows that:

- A) normality is strongly supported.
- B) normality is *not* supported since the curvature in the plot indicates that the distribution is skewed towards the high values.
- C) normality is *not* supported since the curvature in the plot indicates that the distribution is skewed towards the low values.
- D) normal scores plots cannot be used with skewed data.
- E) None of the above.

Defective Question Report

Name: _____

Section: _____

ID: _____

If you believe that a test question is defective in some way, please list your concern here. Your concerns will be considered in our interpretation of the test results.

Remove this last page from the exam questions and turn it in with your exam questions and answers to one of the instructors in the course.

Question number: