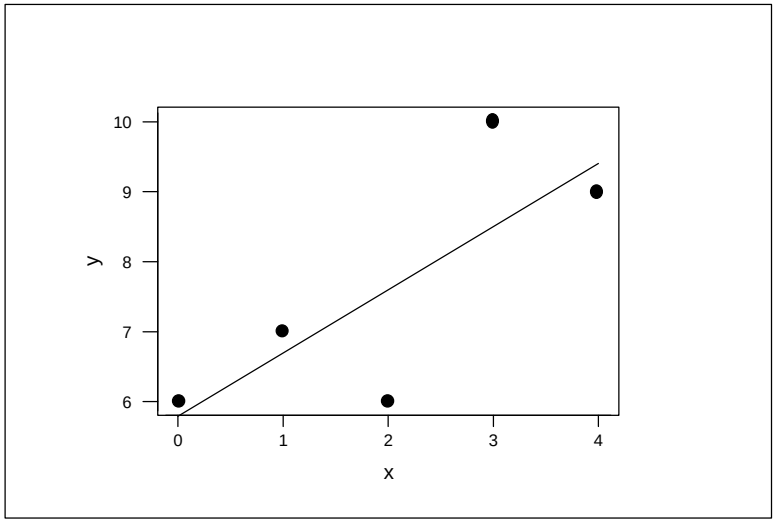
Exam 2, Nov. 2, 1995, Quant II

The exam questions and answer sheet are both to be turned in to your Discussion Section instructor at the end of the exam. **Be sure to code your Section number under OPTIONAL CODES in positions L M N.**

1. 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 top 10% of the class ***must receive*** A's. If exam scores are normally distributed with mean 82.0 and standard deviation 2.34, what exam score corresponds to the lowest A grade?

   A) 70
   B) 75
   C) 80
   D) 85
   E) 90

2. A scatterplot and least squares regression line are shown below for five data points. If the point with coordinates $(x,y)=(4,9)$ were changed to $(x,y)=(4,6)$ what can we say about the new fitted regression line?



   A) the *y*-intercept would increase and the slope would increase
   B) the *y*-intercept would decrease and the slope would increase
   C) the *y*-intercept would increase and the slope would decrease
   D) the *y*-intercept would decrease and the slope would decrease
   E) None of the above.

3. Data **design quality** refers to assessment of whether the data collected are relevant to the problem we wish to solve.

   A) True B) False

4. The least squares regression line is that line which makes the sum of squared vertical distances between the observation pairs and the fitted line as small as possible.

   A) True B) False

5. A set of data pairs has the following summary statistics: $\bar{x} = 10$, $\bar{y} = 50$, $s_x = 2$, $s_y = 3$, and $r = 0.5$. What is the equation for the least squares regression line in ***original terms*** of x and y?

   A) $\hat{y} = 0.5x$

   B) $\hat{y} = 0.75x$

   C) $\hat{y} = 45 + 0.5x$

   D) $\hat{y} = 42.5 + 0.75x$

   E) None of the above.

6. What is the area under the standard normal curve between the z-scores of −1.3 and +0.91?

   A) 0.0968
   B) 0.7218
   C) 0.8186
   D) 0.9154
   E) None of the above.

7. Twenty-five data pairs are under consideration. The first pair is $x = 2$ with $y = 6$. What is the **fitted value** for this data pair for the quadratic regression curve with equation $\hat{y} = 1 + x^2$?

   A) 1
   B) 2
   C) 5
   D) 6
   E) None of the above.

8. The **major** difference between simple straight-line regression models and multiple regression models is:

   A) Multiple regression models are curved.

   B) Regression coefficients in multiple regression models are not found by least squares.

   C) Multiple regression models have several response variables.

   D) Multiple regression models have an associated residual standard deviation.

   E) Multiple regression models have several predictor variables.

9. Suppose y is a response variable of annual salary, x is a continuous predictor variable of years experience, and z is a binary indicator variable indicating gender (z=1 means female). Consider the regression model with equation: $\hat{y} = b_0 + b_1 x + b_2 z$. What coefficient (or combination of coefficients) represents the predicted *starting* annual salary (experience=0) for a female?

   A) b_0

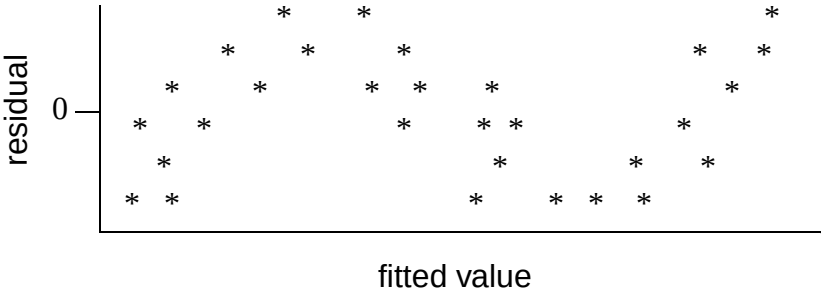
   B) b_1

   C) b_2

   D) b_0+b_1

   E) b_0+b_2

10. The residual plot from a least-squares multiple regression model fit is shown below. Which of the following statements best describes the implication of the 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 parallel-lines model should be considered.

   D) The plot shows that least squares is an excellent criterion for fitting the line.

   E) None of the above.

11. A factorial experiment has two factors. The first factor has three levels and the second factor has two levels. If the experiment has three replications, how many observations will be recorded?

    A) 2
    B) 3
    C) 6
    D) 12
    E) 18

12. An elevator carries 9 people in a load. The weights of people vary according to many factors but may be described by a distribution with mean 150 pounds and standard deviation 30 pounds. Over many loads each of 9 people, about what percentage of loads will exceed the safe load limit of 1700 pounds?

    A) 0.01%
    B) 0.1%
    C) 1%
    D) 5%
    E) None of the above.

13. In general, the larger the value of the adjusted R^2, the better the regression model.

    A) True B) False

14. Hawkeye Supply Inc. has randomly selected 100 steel bolts from a large shipment. Suppose that the individual 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 larger than 2.98 inches?

    A) 0.0228
    B) 0.4207
    C) 0.5793
    D) 0.9772
    E) None of the above.

15. Is the following expression true or false for a multiple regression model?
    Residual = Observed Response – Fitted Value

    A) True B) False

16. What is the 75^{th} percentile of a normal distribution with mean 65 and standard deviation 7.463?

    A) 67
    B) 68
    C) 69
    D) 70
    E) 91

17. Twenty data pairs are under consideration. The second pair has $x = 2$ with $y = 3$. What is the value of the **residual** for this data pair for the line with equation $\hat{y} = 1 + x$?

    A) 0
    B) 1
    C) 2
    D) 3
    E) 4

18. Alligators have weights that are distributed approximately according to a normal distribution with mean 150 kilograms and standard deviation 56 kilograms. Only alligators that weigh at least 200 kilograms may be legally caught. What percentage of all alligators may be legally caught? (to the nearest **whole percent**)

    A) 11%
    B) 19%
    C) 81%
    D) 89%
    E) None of the above.

19. Stanford-Binet IQ scores are approximately normally distributed with a mean of 100. If the 84th percentile of the scores is 116, what is the standard deviation of the scores?

    A) 1
    B) 16
    C) 84
    D) 100
    E) 116

20. In multiple regression modeling, residual plots that are random indicate that the model can be improved.

    A) True B) False

21. **Replication** is important in experimentation since process results vary even under identical experimental conditions.

    A) True B) False

22. In the Physicians' Health Study 22,000 male doctors were used as subjects to measure the effect of asprin in the prevention of heart attacks. In this study the physicians receiving aspirin are said to be in the **control group**.

    A) True B) False

23. In the Physicians' Health Study 22,000 male doctors were used as subjects to measure the effect of asprin in the prevention of heart attacks. By saying that the experiment was **double blinded** we mean that the physicians were given a choice as to whether they would receive aspirin or not.

    A) True B) False

24. In a taste test of two brands of beer, the brands were not revealed to the tasters. This an example of :

    A) randomization
    B) replication
    C) blinding
    D) stratification
    E) placebo

25. Control charts are designed to look at both the cross-sectional and longitudinal aspects of processes.

    A) True B) False

26. We have 27 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) 24
    B) 25
    C) 26
    D) 27
    E) None of the above.

27. The purpose of a mean control chart is to detect changes in variation within subgroups.

    A) True B) False

28. In multiple regression modeling, the most basic residual plot is the plot of residuals versus the corresponding response values.

    A) True B) False

29. 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 <u>least</u> 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	0.2	3.00	0.00

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 two of its residuals are zero.

30. A standard deviation control chart displays the changes in variation within subgroups over time.

A) True B) False

31. In a normal probablility plot, the closer the plot is to some straight line, the more support we have for a normal distribution for the data.

A) True B) False

32. **Common causes** of variation in a measured variable that are due to chance and remain in the system unless the process is fundamentally altered.

A) True B) False

33. Suppose that a change in supplier of raw materials has the effect of increasing variability of a measured variable but not its mean level. Which of the following statements is true?

A) A mean chart will detect the change but a standard deviation chart will be unaffected.

B) A standard deviation chart will detect the change but a mean chart will be unaffected.

C) Both mean charts and standard deviations charts will detect the change.

D) No control chart can be expected to effectively deal with this situation.

E) None of the above.

34. A portion of an ANOVA (analysis-of-variance) table from a multiple regression calculation is shown below. What is the value of *s* (the residual standard deviation) for this regression?

Source	SS	df
Regression	510.90	3
Error	127.33	6
Total		

A) 4.61
B) 8.42
C) 21.22
D) 70.91
E) None of the above.

35. Referring to the ANOVA table in problem 33, what is the value of R^2 for this regression?

A) 0.80%
B) 19.95%
C) 70.01%
D) 80.05%
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:
Please also tell us which response you chose.