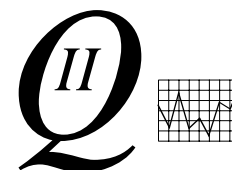


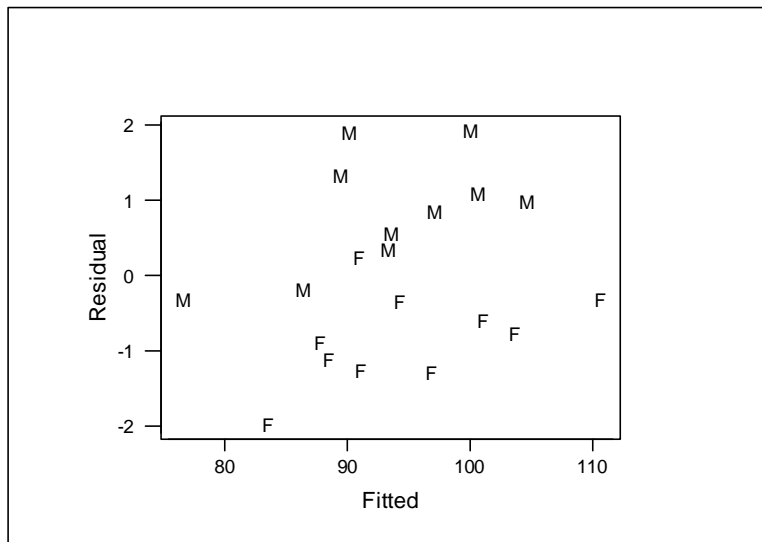
Final Exam, May 12, 1995, Quant II



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 (Optional codes under LMN) on the answer sheet.**

1. A universe consists of 8 elements. How many samples of size 3 are possible?
 - A) 8
 - B) 24
 - C) 6720
 - D) 40,320
 - E) None of the above.
2. A **statistic** is a numerical characteristic of a population.
 - A) True
 - B) False
3. 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 80.0 and standard deviation 7.8, what exam score corresponds to the highest F grade?
 - A) 50
 - B) 55
 - C) 60
 - D) 65
 - E) 70
4. Longitudinal data were collected on a continuous variable from a stable process. Unfortunately, they were given to us only in numerical order from smallest value to largest value rather than in time order. Which of the following could we reasonably produce from the data as given?
 - I. the one-sample t -statistic for testing $H_0: \mu = 55$
 - II. the standard error of the mean
 - III. the p -value for testing $H_0: \mu = 100$
 - A) I and II only
 - B) II and III only
 - C) I and III only
 - D) I, II, and III
 - E) None of the above.

5. 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.
6. 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 *gallons per 100 miles*? (Rounded to the nearest gallon.)
- A) 6
 - B) 7
 - C) 15
 - D) 16
 - E) 17

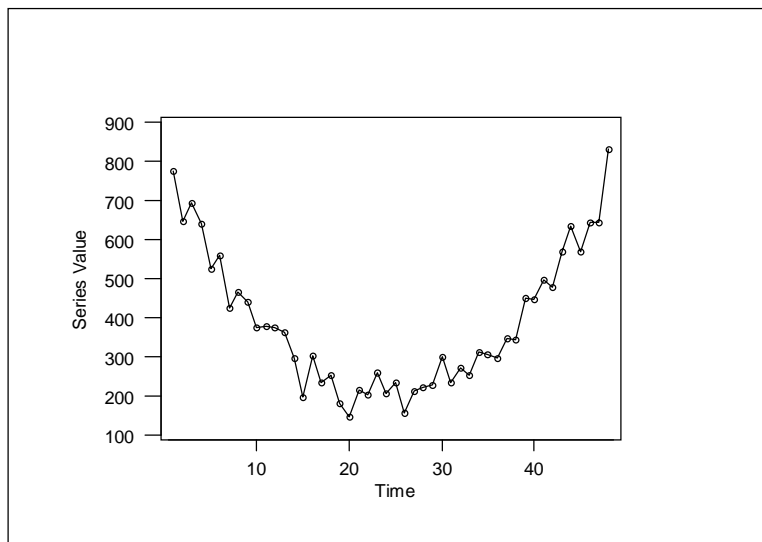
7. The table below shows four data pairs together with some partial results on fitted values and residuals for two possible models—both quadratic. *These models were not necessarily found using least squares.*

Data			Curve I $\hat{y} = 1.25 - 0.55x + 0.25x^2$		Curve II $\hat{y} = 1 - 0.5x + 0.25x^2$	
y	x	x^2	FITTED	RESIDUAL	FITTED	RESIDUAL
1	1	1	0.95	0.05	0.75	0.25
1	2	4		-0.15		0.00
2	3	9	1.85	0.15	1.75	0.25
3	4	16	3.05		3.00	

Which curve fits the data better in the sense of least squares? (You may assume that all values given in the table are correct.)

- 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 II fits better since its sum of squared residuals is smaller than for curve I.
 D) Curve II fits better since one of its residuals is zero.
 E) None of the above.
8. 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) 0.1%
 B) 5%
 C) 10%
 D) 20%
 E) None of the above.

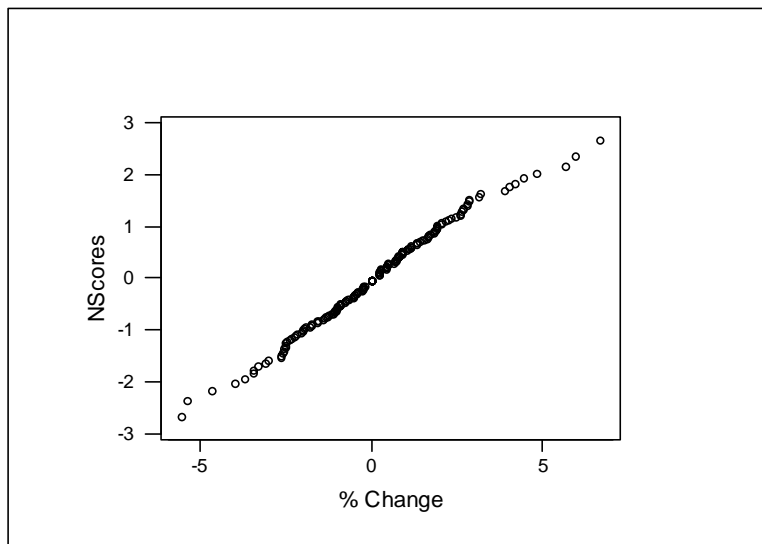
9. Hawkeye Supply Company has randomly selected 100 steel bolts from a large shipment. Suppose that the bolt lengths in the shipment may be described by a distribution with mean 3 inches and standard deviation 0.1 inches. Let \bar{y} denote the average bolt length for the sample of 100 bolts. If the sampling were repeated many times, what fraction of the averages, \bar{y} , would be less than 2.98 inches?
- A) 0.0228
 B) 0.4207
 C) 0.5793
 D) 0.9773
 E) None of the above.
10. The use of the *finite population correction* factor (*fpc*), decreases the standard error of the mean to account for the effects of sampling without replacement.
- A) True B) False
11. 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

12. Consider a Binomial distribution with $n = 3$ and $\pi = 0.3$. What is the chance of *at most one* success?
- A) 0.147
 B) 0.343
 C) 0.441
 D) 0.784
 E) None of the above.

13. The normal probability plot for the percentage changes of the daily Chrysler stock price data is shown below:



This plot shows that:

- A) normality of the percentage changes is strongly supported.
 B) normality of the percentage changes is *not* supported since the curvature in the plot indicates that the distribution is strongly skewed towards the high values.
 C) normality of the percentage changes is *not* supported since the curvature in the plot indicates that the distribution is strongly skewed towards the low values.
 D) normal scores plots cannot be used with percentage change data.
 E) normal scores should have been calculated for the original stock prices.
14. The Central Limit Effect applied to a Binomial variable y from n trials of a Bernoulli process says that for large n the distribution of y is approximately normal with mean π and standard deviation $\sqrt{\frac{\pi(1-\pi)}{n}}$.
- A) True B) False

15. A Binomial distribution has its *smallest variability* when $\pi = 0.5$.
- A) True B) False
16. In a **probability survey**, each element has a known probability of being chosen for the sample.
- A) True B) False
17. After an exam has been proofread 4 times, typos still appear at a rate of about 1/10 of a typo *per page*. Using a Bernoulli model (with each page representing a trial and assuming no more than one typo per page), what is the mean for the number of typos on an entire 10 page exam?
- A) 2/10
B) 1
C) 2
D) 4
E) 10
18. **Sampling error** is the difference between the value of a sample estimate and the corresponding equal complete coverage value that is due only to the sampling process.
- A) True B) False
19. As discussed in lecture, “A leading pollster has repudiated a widely publicized survey that seemed to show that a third of Americans were open to the possibility that the Holocaust never happened. ‘The so-called Holocaust question was flawed,’ said Burns Roper...”
- This is an example of:
- A) a coverage error
B) a sampling error
C) a survey instrument error
D) a frame error
E) interviewer bias
20. The article “A Pollster’s Worst Nightmare: Declining Response Rates” discussed the importance of
- A) coverage errors
B) survey instrument errors
C) sampling errors
D) avoiding interviewer bias
E) having a large sample size
21. A **parameter** is a numerical characteristic of a sample.
- A) True B) False

22. Multiple call-backs to nonrespondents are **not** possible in
- A) a probability survey
 - B) a nonprobability survey
 - C) a telephone survey
 - D) a confidential mail survey
 - E) an anonymous mail survey
23. A survey is to be carried out using a simple random sample. Assume that the universe is very large. We need to be 95% confident that our estimate of a certain proportion has a margin of error of ± 0.02 (that is, ± 2 percentage points). How large a sample do we need?
- A) 50
 - B) 95
 - C) 836
 - D) 1190
 - E) 2501
24. We have observed 9 values of a stable process and obtained a sample mean of 100.5 and sample standard deviation of 0.23. We need to estimate the mean of the process. What is the value of the upper endpoint of a 95% confidence interval for the process mean?
- A) $100.5 + 2.306 \times (0.23)$
 - B) $100.5 + 1.860 \times (0.23)$
 - C) $100.5 + 1.860 \times (0.23 / (\sqrt{9}))$
 - D) $100.5 + 2 \times (0.23 / (\sqrt{8}))$
 - E) $100.5 + 2.306 \times (0.23 / (\sqrt{9}))$
25. A result is said to be **statistically significant** when the observed value of a test statistic is close to the value specified by the null hypothesis.
- A) True
 - B) False
26. We need to find the critical value, $t_{1-\alpha/2}$, for a t -test with 11 degrees of freedom. The alternative is two-sided. What is the value of $t_{1-\alpha/2}$ if the significance level is 5%?
- A) -1.796
 - B) -2.201
 - C) 1.796
 - D) 2.201
 - E) None of these

27. In a simple random sample, increasing the sample size will decrease the margin of sampling error (other things being equal).
- A) True B) False
28. **Type I error** is failing to reject a false null hypothesis.
- A) True B) False
29. Suppose that we have two independent estimates, \bar{y}_1 and \bar{y}_2 , of μ_1 and μ_2 . Both means are based on 10 observations. The standard error of \bar{y}_1 is 2.4 and of \bar{y}_2 is 2.1. What is the standard error of $\bar{y}_1 - \bar{y}_2$?
- A) 0.30
B) 1.16
C) 1.35
D) 3.19
E) 10.17
30. A simple random sample of size 100 from a large universe yielded a sample proportion of 0.40. Find the upper endpoint of the 99.7% confidence interval for the population proportion π .
- A) 0.05
B) 0.15
C) 0.41
D) 0.55
E) 0.95
31. A new marketing strategy is tested against the one in present use. The results reported that the p -value was 0.02. Which **one** of the following statements is correct? (Here the null hypothesis is that the two strategies are no different.)
- A) H_0 should be rejected at the 1% significance level.
B) H_0 should be rejected at both the 1% and 5% significance levels.
C) H_0 should not be rejected at the 5% significance level.
D) H_0 should not be rejected at the 10% significance level.
E) H_0 should not be rejected at any significance level less than 2%.
32. The **frame** is the list of elements from which a sample is drawn.
- A) True B) False

33. Suppose that 30 samples each of size 50 are randomly selected from a process. A binary variable is measured and gives a grand total of 225 successes. If the next sample of size 50 produces 17 successes, which one of the following is true?
- A) The process is out of control since $17/50$ exceeds 0.30.
 - B) The process is in control since $17/50$ is less than 0.35.
 - C) The process is out of control since $17/50$ exceeds 0.16.
 - D) The process is in control since $17/50$ is less than 0.997.
 - E) The process is in control since $17/50$ exceeds 0.003.
34. **Nonsampling error** is the difference between the value of a sample estimate and the corresponding equal complete coverage value that is due to some cause other than the sampling process.
- A) True B) False
35. Which *one* of the following statements is *false* regarding control charts?
- A) Control charts can help avoid tampering with a stable process.
 - B) Control charts can help managers decide appropriate action to take concerning their processes.
 - C) Control charts help managers detect special causes of variation in their processes.
 - D) Control charts can be applied to service businesses as well as manufacturing.
 - E) Control charts were invented by the Japanese within the last 20 years.
36. A **probability survey** is a survey conducted with the intention of obtaining information from every element in the universe.
- A) True B) False
37. The **equal complete coverage** is the set of measurements that would be obtained if the survey procedures were used on every element in the universe.
- A) True B) False
38. A simple random sample of size 50 is selected from a universe of 200 accounts. The sample mean of the account size is \$250 and the sample standard deviation is \$40. Find the upper end-point for the 99.7% confidence interval for the population mean account size. Notice that the sampling fraction is 25% so that the *fpc* is needed. (Round to the nearest dollar. Do *not* use the *t*-distribution.)
- A) \$260
 - B) \$261
 - C) \$265
 - D) \$267
 - E) \$371

39. In an experiment a class of students were asked to guess the age of their professor. 190 of them were given no additional information but 189 were told the ages of the professor's three children. Here are the results:

Group	n	Mean Guess of Age (years)	Standard Deviation	SE Mean
No information	190	48.14	5.557	0.403
Given Ages of Kids	189	53.97	3.201	0.233

What is the value of the t -statistic for testing the null hypothesis of no difference between the two group means? Here $H_0: \mu_{\text{info}} - \mu_{\text{noinfo}} = 0$.

- A) 0.05
 B) 5.83
 C) 12.5
 D) 302
 E) Cannot be determined from the information given.
40. Referring to the previous problem: What is the critical t value for testing against a one-sided alternative (that the group with more information would have a higher mean, $H_1: \mu_{\text{info}} - \mu_{\text{noinfo}} > 0$) at the 5% significance level? (We would reject the null hypothesis if our observed two-sample t -statistic is greater than this value.)
- A) 1.282
 B) 1.645
 C) 1.960
 D) 2.326
 E) 2.576
41. Which one of the following displays the cross-sectional characteristics of a process?
- A) Deming's PDCA wheel
 B) standard deviation chart
 C) flow diagram
 D) cause-and-effect diagram
 E) density histogram
42. In a probability survey, increasing the sample size will decrease nonsampling errors.
- A) True B) False

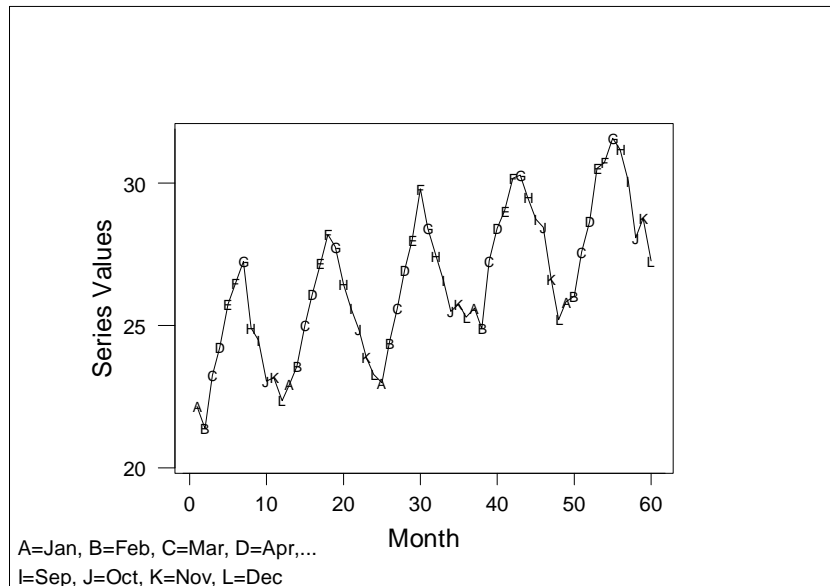
43. Which of the following tools would be useful for a longitudinal analysis of a process?

- I. Control charts for means
- II. p -charts
- III. Sequence plots

- A) I only
- B) II only
- C) III only
- D) I, II, and III
- E) None of the above.

44. A sequence plot of 5 years of monthly data is shown below. Which of the following best describes the plot?

- A) random
- B) meandering
- C) seasonal
- D) straight-line trend
- E) straight-line trend plus seasonal



45. Five percent of the products produced by a stable process are defective. If a sample of 4 items is taken, what is the probability that the sample contains *at least one* defective? (to the nearest hundredth)

- A) 0.17
- B) 0.19
- C) 0.81
- D) 0.83
- E) 0.95

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.

For each question, please report which answer you choose.

Question number: