University of Toronto Scarborough
STAB22 Final Examination

December 2008

For this examination, you are allowed two handwritten letter-sized sheets of notes (both sides) prepared by you, a non-programmable, non-communicating calculator, and writing implements.

This question paper has 17 numbered pages. Before you start, check to see that you have all the pages. In addition, there are statistical tables at the back.

This examination is multiple choice. Each question has equal weight, and there is no penalty for guessing. On the Scantron answer sheet, ensure that you enter your last name, first name (as much of it as fits), and student number (in "Identification").

Mark in each case the best answer out of the alternatives given (which means the numerically closest answer if the answer is a number and the answer you obtained is not given.)

If you need paper for rough work, use the back of the sheets of this question paper. The question paper will be collected at the end of the examination, but any writing on it will not be read or marked.

Before you begin, two more things:

- Check that the colour printed on your Scantron sheet matches the colour of your question paper. If it does not, get a new Scantron from an invigilator.
- Complete the signature sheet, but *sign it only when the invigilator collects it.* The signature sheet shows that you were present at the exam.

1. An industrial process produces items of which 80% conform to specifications. An inspector will take a random sample of 100 items. In the sample, the probability that x or more items will conform to specifications is 0.30. What, using a suitable approximation, is x?

   (a) 78
   (b) 90
   (c) * 82
   (d) 70
   (e) 55

2. A local country club operates facilities that include a golf course and tennis courts. A survey of the members indicates that 72% regularly use the golf course, 42% regularly use the tennis courts and 8% use neither of these facilities regularly. (Some members play both golf and tennis.) Use this information for this question and the next one.

   The club has 600 members. How many members regularly use at least one of the golf or tennis facilities?

   (a) *552
   (b) 276
   (c) 424
   (d) 578
   (e) 244

3. Suppose that we randomly selected one of the members of the country club mentioned in Question 2. The probability that this member regularly uses exactly one facility is

   (a) 0.79
   (b) 0.50
   (c) 0.17
   (d) 0.29
   (e) *0.70

4. A population mean is known to be 20. A simple random sample of 80 individuals is drawn from this population. What can you say about what the sample mean might be?

   (a) it will be exactly 20
   (b) it could be anything
   (c) * it could be anything, but will probably be close to 20

5. A fraud investigator is investigating a group of physicians accused of forging their bills. The physicians work at three different hospitals (labelled A, B and C). The investigator's goal is to estimate the overall proportion of forged bills. It is known that bills from hospital A are very hard to forge. The investigator only has resources to sample 10% of the relevant bills. What sampling method might the investigator use?

   (a) Multi-stage sample
   (b) * Stratified sample
   (c) Simple random sample
   (d) Block design
   (e) Systematic sample

6. Travel times for buses on a certain journey have a mean of 22 minutes and a standard deviation of 3 minutes. The travel times have a normal distribution. Use this information for this question and the next one.

   What is the probability that on a randomly chosen day, the journey will take more than 24 minutes?

   (a) 0.75
   (b) 0.50
   (c) * 0.25
   (d) 0.35
   (e) 0.10

7. Refer to Question 6. The journey described there is made on 10 randomly chosen days. What is the probability that the sample mean travel time on those days is greater than 24 minutes?

   (a) 0.98
   (b) 0.30
   (c) 0.70
   (d) * 0.02
   (e) Same answer as Question 6

8. A bottling company needs to produce bottles that will hold 12 ounces of liquid for a local brewery. The company receives a complaint that their bottles are not holding enough liquid. To assess the evidence for this complaint, the bottling company randomly samples 64 bottles and finds that the average amount of liquid held by the 64 bottles is 11.915 ounces with a sample standard deviation of 0.40 ounces. Suppose that the P-value of this test turned out to be 0.0446. The proper conclusion from this test is:

   (a) We fail to reject the null hypothesis at $\alpha = 0.05$
   (b) * There is evidence that the company bottles are not holding enough liquid at $\alpha = 0.05$.
   (c) There is evidence that the company bottles are not holding enough liquid at $\alpha = 0.025$.
   (d) There is not enough information to make any conclusion.
   (e) We reject the null hypothesis at $\alpha = 0.01$.

9. We want to examine and compare the effects on digestion of taking varying amounts of a certain drug: 0, 25, 50, 75 mg. What kind of experimental design might we use?

   (a) Block design
   (b) Matched pairs design
   (c) Systematic sample
   (d) * Completely randomized design

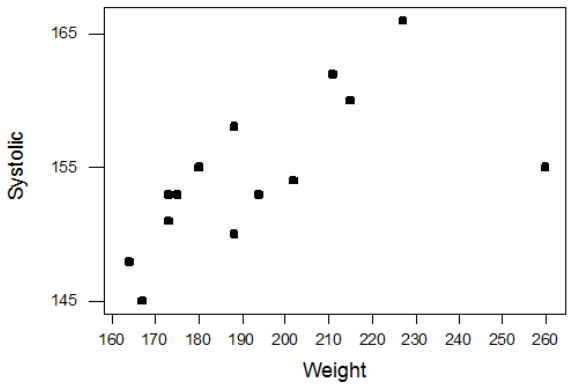
10. The capacity of a motel is 360 rooms. To minimize losses due to last minute cancellation the motel accept reservations in excess of its capacity. The records of the motel show that, on average, 18% of their prospective guests will cancel their reservation last minute and will not show up. If the motel accepts 420 reservations what is the probability that all the guests who arrive to claim a room will receive one i.e. the number of guests who will show up will not exceed the motel capacity?

    (a) * 0.9761
    (b) 1.00
    (c) 0.9612

(d) 0.9126

(e) 0.9272

11. Systolic blood pressure readings of individuals are thought to be related to age and weight. For a sample of 14 men, the following scatterplot is obtained:



Three statements are made about the plot above:

I: Based on the graph above, there is, on the whole, a linear relationship between systolic blood pressure and weight.

II: Systolic blood pressure is the response variable while weight is the explanatory variable.

III: There is one observation that is a suspected outlier with respect to systolic blood pressure.

Which of these three statements are true?

(a) Only statements I and III are true.

(b) Fewer than two statements are true.

(c) All of the statements are true.

(d) * Only statements I and II are true.

(e) Only statements II and III are true.

12. A recent study on teenage pregnancy indicates that 4% of the female population will get pregnant sometime in their teenage years. The study stated that teenage pregnancies occur independently of one another. Suppose 15 teenage women were randomly selected. What is the probability that between 3 and 5 inclusive of the sampled women will get pregnant sometime during their teenage years?

(a) 0.001

(b) 0.105

(c) 0.018

(d) 0.099

(e) * 0.020

13. Test marks for the mid-term test in each of the three large Statistics service courses are normally distributed and have an average of 72 with the same standard deviation. In STA100, tutorials have 18 students each; in STA200 there are 36 students per tutorial, and in STA250, there are 26 per tutorial. Each course has the same large number of tutorial sections. For each tutorial, the test average is recorded. In which course, do you expect to observe more tutorials with a test average below 69? Assume that tutorials are formed from students in a completely random manner.

(a) I'd expect to see about the same number of such tutorials in each course

(b) * STA100

(c) Insufficient information to answer

(d) STA200

(e) STA250

14. A sample of size 16 is taken from a population with mean 25 and SD 3. The population has an approximately normal shape. What is the probability that the sample mean will be between 23.5 and 26.5?

(a) * 0.95

(b) 0.69

(c) 0.98

(d) cannot say anything about the probability without more information about the population

(e) 0.38

15. 452 mothers were classified according to their alcohol intake (ounces per day) prior to pregnancy recognition and their nicotine intake during pregnancy (milligrams per day). The data are summarized below, where 'alcnone' denotes no alcohol, 'alc001' denotes 0.01-0.10 ounces, 'alc011' denotes 0.11-0.99 ounces, and 'alc100' denotes 1.00 or more ounces. 'nic01' denotes 1-15 mg., 'nic16' denotes 16 or more mg., and 'nicnone' denotes no nicotine. Consider this to be a simple random sample from some particular population.

Rows: Alcohol Columns: Nicotine

	nicnone	nic01	nic16	All
alcnone	105	7	11	123
alc001	58	5	13	76
alc011	84	37	42	163
alc100	57	16	17	90
All	304	65	83	452

Cell Contents --
Count

Use this information for this question and the next one.

Here are two statements about the above data:

I: The estimate of the proportion of mothers who are either heavy-drinkers (alc100) or heavy smokers (nic16) is equal to 156/452.

II: Suppose we want to estimate the proportion of the non-smokers mothers (nicnone) among the non-drinking mothers (alcnone). This proportion can be estimated using the above information. Its standard deviation may also be estimated from the data and equals 0.03 (2 decimal places).

(a) Only statement I is true.

(b) Neither statement is true.

(c) Only statement II is true.

(d) There is insufficient information to answer

(e) * Both statements are true.

16. Suppose that we randomly select one person from the 452 mothers mentioned in Question 15. Three statements are made about this person:

    I: The probability that the selected person is a heavy-drinker (alc100) is equal to 0.199 (3 decimal places).
    II: The probability that the selected person is a heavy-drinker (alc100) and a non-smoker (nicnone) equals 0.013.
    III: We know that the selected mother is one of the heavy smokers. The probability that she is also a drinker (alc001 or alc011 or alc100) equals 0.867.

    Which of these statements are true?

    (a) Only one statement is true
    (b) Statements I and II are true
    (c) * Statements I and III are true
    (d) All the statements are true
    (e) Statements II and III are true

17. Suppose that STAB22 students' ages follow a skewed distribution with a mean of 24 years old and standard deviation of 3 years. Suppose we randomly select 200 students. The following statements about the sampling distribution of the sample mean age can be made:

    I: The mean of the sampling distribution is approximately 9 years.
    II: The shape of the sampling distribution is approximately normal.
    III: The standard deviation of the sampling distribution is equal to 0.21 years.

    Which of these three statements are true?

    (a) All the statements are true
    (b) Only statements I and II are true
    (c) Only statements I and III are true
    (d) Fewer than two statements are true
    (e) * Only statements II and III are true

18. A list of enumeration areas in Ontario is made. From the list, we pick every 10th one, after a random start. For the selected areas, we obtain maps. For each map, we number the blocks, from 1 to N (n is the number of blocks in that area). Using a random number table, we select two distinct numbers between 1 and N (inclusive), and include the corresponding blocks in our sample. On each selected block, we start at the northeast courner, and walk around the block, selecting every 5th household into our sample (from a random start). Which types of sampling methods are used here (in no particular order)?

    (a) * multi-stage; simple random sample; systematic
    (b) systematic; multi-stage; stratified
    (c) multi-stage; simple random sample; stratified
    (d) stratified; simple random sample; systematic
    (e) systematic; simple random sample

19. A human gene carries a certain disease from the mother to the child with a probability rate of 60%. That is, there is a 60% chance that the child becomes infected with the disease. Suppose a female carrier of the gene has four children and that the infections of the four children are independent of one another. The probability that at least one of the children does not get the disease from their mother is:

    (a) 0.1272

(b) *0.8704

(c) 0.9744

(d) 0.0256

(e) 0.7226

20. About 25% of students in this course will end up with some kind of A grade (A+, A or A-). For this question, "an A grade" means one of the grades A+, A or A-. If you take a random sample of 50 students from this course, what is the approximate probability that 10 or fewer of them will end up with an A grade?

    (a) cannot calculate using the methods learned in this course

    (b) 0.79

    (c) * 0.21

    (d) 0.50

    (e) 0.38

21. Suppose you are surveying customers shopping in a department store about their choice of products. What kind of sampling procedure might you use?

    (a) * Simple random sample

    (b) Matched pairs design

    (c) Multi-stage sample

    (d) Stratified sample

    (e) Systematic sample

22. A local eat-in pizza restaurant wants to investigate the possibility of starting to deliver pizzas. The owner of the store has determined that home delivery will be successful if the average delivery time does not exceed 28 minutes. To investigate that she has randomly selected 26 costumers and has delivered pizzas to their homes. She recorded the delivery times. The owner of the store views the delivery possibility as a method of expanding her business. Three statements are made about this situation:

    I: The appropriate hypotheses that the store owner should test are: H_0: $\mu = 28$ and H_a: $\mu < 28$.

    II: If she decide to expand her business when in fact the mean delivery time is 28 minutes, she makes a type I error.

    III: If she decides to expand her business when in fact the mean delivery time is less than 28 minutes, she makes a type II error.

    Which of these three statements are true?

    (a) * Only statements I and II are true

    (b) Only statements II and III are true

    (c) All of the statements are true

    (d) Only statements I and III are true

    (e) None of the statements are true

23. The amount of corn chips an "18-ounce" bag contains follows a normal distribution with mean 18.5 ounces and standard deviation of 0.2 ounce. Suppose that 100 bags of chips were randomly selected. The probability that the sample mean weight of these 100 bags is less than 18.6 ounces is

    (a) 0.6915

    (b) 0.7272

(c) * close to 1

(d) 0.3085

(e) 0.1915

24. A telephone survey of opinion (of Torontonians) about Canadian participation in the Kyoto greenhouse gas emission accord was conducted by an ambitious young undergraduate student, named Bob, who chose households for his survey by randomly selecting telephone numbers from the current Toronto telephone directory. Results were needed quickly, so Bob's family members and friends volunteered to help do the interviewing. The selected phone numbers were all called between 10am and 4pm on Wednesday, Nov 20, after which results were tabulated. If no one answered, another phone number was selected at random from the phone book. All of those successfully contacted by phone answered every question in the survey. The following three problems may have occurred in this survey:

    I: Nonresponse
    II: Undercoverage
    III: Interviewer bias

    Which problems occurred?

    (a) * All of them

    (b) I and II

    (c) Fewer than two.

    (d) II and III

    (e) I and III

25. Two drugs A and B, used in the treatment of glaucoma, were tested for effectiveness on 10 diseased dogs. Drug A was administered to one eye of each dog and drug B to the other eye. Pressure measurements were taken 1 hour later on both eyeballs of each dog (the smaller the measurement, the less serious the eye disease). Use this information to answer this question and the next one.

    Which of the following best describes this experimental design?

    (a) This is a matched pairs design but not a randomized block design.

    (b) * This is both a matched pairs design and a randomized block design.

    (c) This is a randomized block design but not a matched pairs design.

    (d) This is a completely randomized design.

26. In the study described in Question 25, which of the following is the most important?

    (a) * We need to randomize the assignment of drugs to eyes.

    (b) We need to pair the dogs based on some relevant criteria related to the response.

    (c) We need to stratify the dogs before assigning the drugs.

    (d) We need to select the dogs randomly from a bigger population.

    (e) We need to randomize the assignment of dogs to drugs.

27. A coin is biased so that it has probability 0.6 of coming up heads. Use this information for this question and the next one.

    If the coin is tossed 7 times (independently), what is the probability that exactly 4 of the tosses are heads?

    (a) * 0.29

    (b) 0.60

    (c) cannot answer from the information available

(d) 0.13

(e) 0.19

28. In the situation of Question 27, what is the probability of obtaining 5 or more heads?

    (a) 0.29

    (b) cannot answer from the information available

    (c) 0.51

    (d) 0.26

    (e) * 0.42

29. Suppose a large labour union wishes to estimate the mean number of hours per month a union member is absent from work. The union decides to sample 440 of its members at random and monitor their working time for one month. At the end of the month, the total number of hours absent from work is recorded for each employee. Which of the following assumptions is necessary in order for a confidence interval (for the mean number of hours absent) to be valid?

    (a) The population sampled from has an approximate Uniform distribution.

    (b) The mean of the sample equals the mean of the population.

    (c) The population sampled from has an approximate Normal distribution.

    (d) * None of these assumptions are necessary.

    (e) The variance of the sample equals the variance of the population.

30. Nine subjects are available for an experiment. The experiment has two treatment groups and a control group. Three of the nine subjects will receive the first treatment. The subjects are:

1	Chin
2	Chowdhury
3	Evans
4	Fernandez
5	Johnson
6	Patel
7	Schmidt
8	Yao
9	Zheng

    The following extract from a random digit table is given:

```
4 0 7 4 7 0 5 3 1 2 7 7 4 1 3 5 8 6 9
```

    Using the subjects as numbered above, which subjects will receive the first treatment?

    (a) 4, 7 and 5

    (b) * Fernandez, Johnson and Schmidt

    (c) Fernandez, Zheng and Schmidt

    (d) Chin, Chowdhury and Evans

    (e) Fernandez, Schmidt and Fernandez again

31. According to a survey, 72% of the students at a certain college own a cell phone. We select a random sample of 720 students from this college. Use this information to answer this question and the next one.

    The mean number of students in the sample who own a cell phone is 518.4. What is the standard deviation of the number of students in the sample who own a cell phone?

(a) 0.02

(b) 12.58

(c) cannot answer from the information available

(d) * 12.05

(e) 101.79

32. Using the information in question 31, what is the probability that at least half of the students in the sample own a cell phone?

(a) close to 0

(b) 0.2636

(c) 0.0008

(d) 0.4272

(e) * close to 1

33. Low temperatures in Toronto for the month of February follow a uniform distribution over the interval -13°C to 8°C. What is the probability that a randomly selected February day has a low temperature that exceeds 0°C?

(a) 0.782

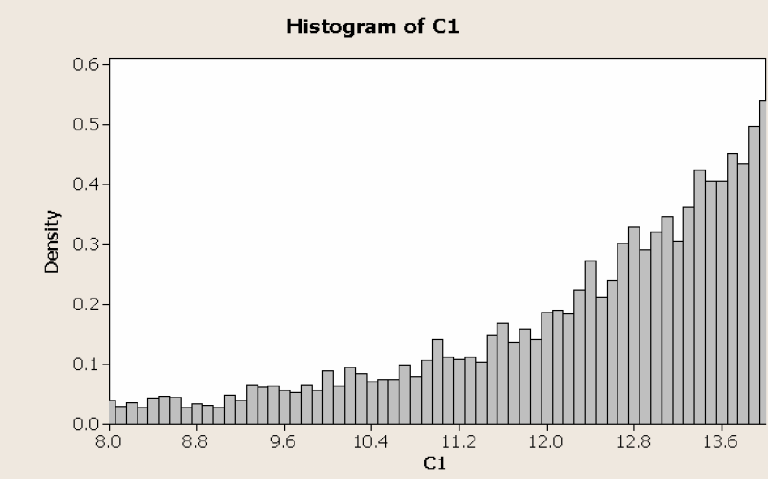
(b) 0.459

(c) 0.724

(d) * 0.381

(e) 0.217

34. Consider the histogram of 4570 measurements ranging from 8.0 to 14.0 shown below.



Here are three statements about the data shown in this histogram:

I: The standard deviation is approximately 3.

II: If we standardize each of these measurements, the distribution of these standardized measurements will be closer to a normal distribution.

III: Fewer than half of these measurements fall below the sample mean.

Which of the three statements are true?

(a) I only

(b) II only

(c) None of the statements are correct.

(d) More than one statement is correct.

(e) * III only

35. For simple random samples, a larger sample size is better than a smaller sample size. Why is this true?

(a) the sampling distribution of the statistic has smaller bias

(b) the sampling distribution of the parameter has smaller variability

(c) * the sampling distribution of the statistic has smaller variability

(d) the statistic is certain to be closer to the parameter

36. Data on salary and years of experience for a sample of social workers were examined, producing the output shown below, where salary was predicted from years of experience. Use the output and to answer this question and the next one.

Regression Analysis

The regression equation is
Salary = 11341 + 2152 Experience

Predictor	Coef	StDev	T	P
Constant	11341	2899	3.91	0.000
Experien	2151.7	147.6	14.58	0.000

S = 7929 R-Sq = 81.6% R-Sq(adj) = 81.2%

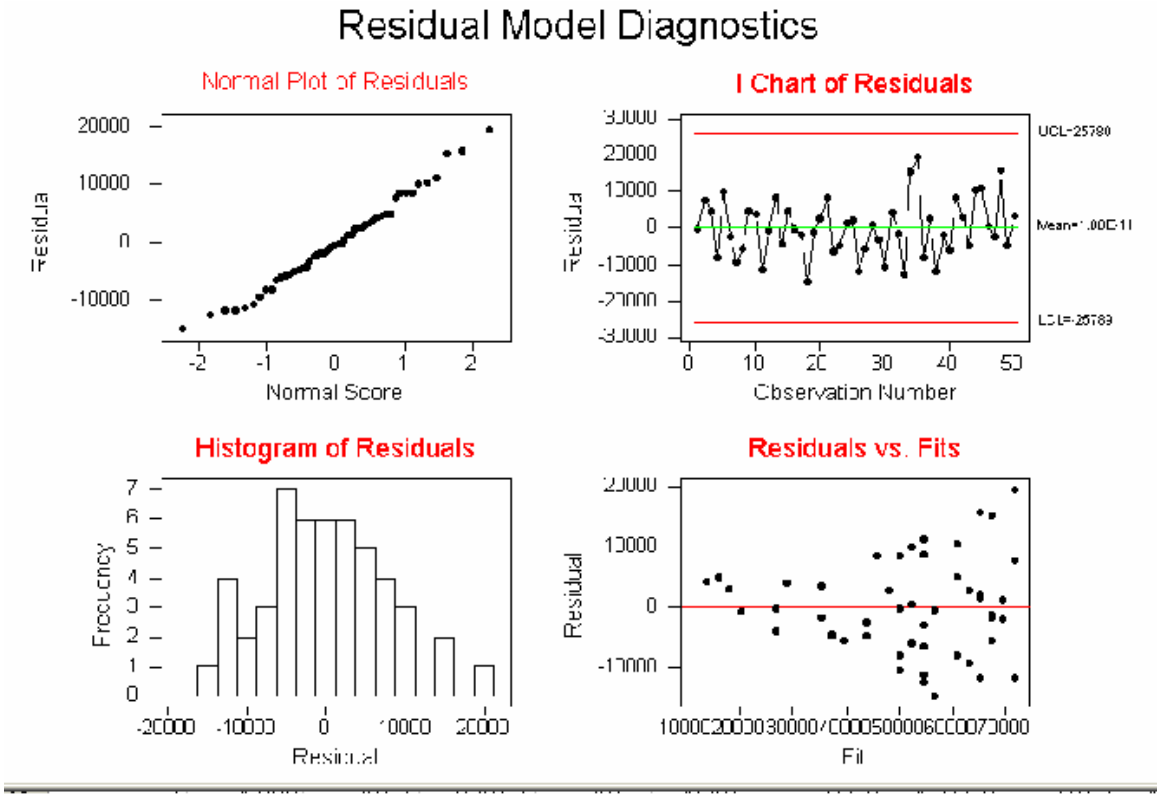
Analysis of Variance

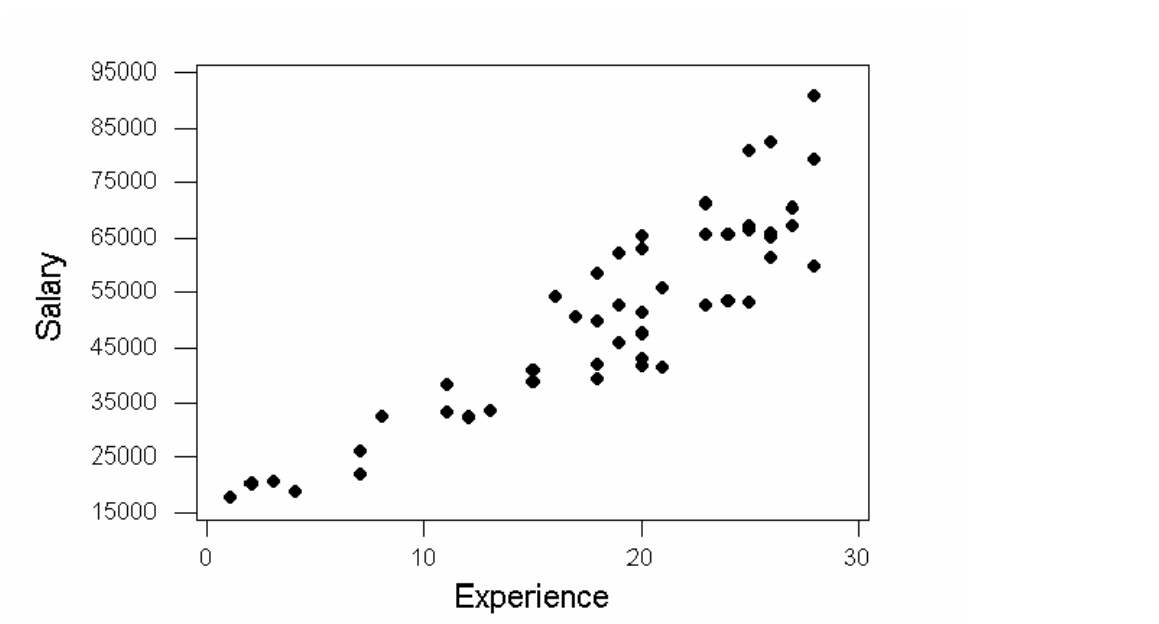
Source	DF	SS	MS	F	P
Regression	1	13367733925	13367733925	212.65	0.000
Residual Error	48	3017437076	62863272		
Total	49	16385171001			

Unusual Observations

Obs	Experien	Salary	Fit	StDev Fit	Residual	St Resid
31	1.0	17833	13493	2764	4340	0.58 X
35	28.0	91139	71589	1839	19550	2.53R
48	25.0	80931	65134	1513	15797	2.03R

R denotes an observation with a large standardized residual
X denotes an observation whose X value gives it large influence.





Here are some statements about the regression:

I: About 82% of the variation in salaries has been accounted for by the linear relationship with experience.

II: On average, salaries go up by \$10,800 for every additional 5 years of experience (to the nearest \$100).

III: The salary of someone with 10 years of experience is estimated to be \$32,900 (to the nearest \$100)

IV: The biggest residual belongs to an individual with 28 years experience.

Which of the statements are true?

(a) All the statements are true except III

(b) All the statements are true except I

(c) All the statements are true except II

(d) *All the statements are true

(e) All the statements are true except IV

37. Review the information given in Question 36 above. In assessing the adequacy of a fitted model, the following possible problems are typically checked for. Which of the following conditions appears to be the biggest problem with our fit above?

(a) None of the other statements are true for this regression.

(b) * The variation in the response is changing.

(c) The residuals have a clearly non-normal distribution.

(d) A non-linear (curved) type of model would be more appropriate for these data than a straight line model.

(e) An inspection of the scatterplot shows that a least squares line will not well represent the general relationship, due to the influence of a small number of points or the existence of distinct clusters.

38. One sample of 15 observations has mean 82 and standard deviation 5. A second sample of 10 observations from the same process but taken by a different scientist, has mean 88 and standard deviation 7. We are interested in seeing whether this is evidence that the two scientists would obtain different results if they each took a very large number of measurements of this process. Use this information for this question and the next one.

For a suitable test of significance, what is the P-value?

(a) between 0.025 and 0.04

(b) between 0.02 and 0.025

(c) * between 0.04 and 0.05

(d) less than 0.02

(e) greater than 0.05

39. Using the information in Question 38, what would you conclude, using $\alpha = 0.05$?

(a) There is evidence that the two scientists give the same results

(b) * There is evidence that the two scientists give different results

(c) There is no evidence that the two scientists give different results

(d) There is no evidence that the two scientists give the same results

40. You want to rent an unfurnished one-bedroom apartment for next semester. The mean monthly rent for a random sample of 13 apartments advertised in the local newspaper is \$570. Assume that the standard deviation of all monthly rents is \$91. Use this information for this question and the next one.

What is the margin of error of the 95% confidence interval for the population mean monthly rent?

(a) * \$49.50

(b) \$619.50

(c) \$520.50

(d) \$55.00

(e) $625.00

41. Using the information in Question 40, how large a sample of one-bedroom apartments would be needed in order to estimate the mean monthly rent μ to within ±$18 with 95% confidence?

    (a) * 99
    (b) 102
    (c) 117
    (d) 128
    (e) 35

42. We are interested in testing the hypothesis H_0: $\mu = 42$ versus H_a: $\mu \neq 42$. It is known that the standard deviation of the population is 7.2. In a sample of n = 101 observations we obtained a sample mean of 41. The P-value of this test is

    (a) * 0.16
    (b) 0.08
    (c) 0.32
    (d) 0.005
    (e) 0.04

43. A marketing researcher is running a study to estimate the average amount spent on electronics last year. He conducted a simple random sample of 1000 adults from Toronto and obtained a 90% confidence interval for this mean. The same study is run in Manhattan by using a simple random sample of 4000 adults to obtain again a 90% confidence interval. Assume the population in Toronto to be 4.8 million, and in Manhattan to be 1.2 million. Also, assume that the true standard deviation of the amount spent is the same for the two different cities. The confidence interval for Toronto is approximately:

    (a) 2 times smaller than the Manhattan confidence interval
    (b) 4 times smaller than the Manhattan confidence interval
    (c) The same width as the Manhattan confidence interval
    (d) 4 times larger than the Manhattan confidence interval
    (e) * 2 times larger than the Manhattan confidence interval

44. In a factory, a machine produces a certain type of part. The machine is such that 95% of parts produced meet specifications. A simple random sample of 100 parts is taken. What is the approximate probability that 90 or fewer of these parts will meet specifications?

    (a) 0.10
    (b) 0.50
    (c) 0.01
    (d) 0.05
    (e) * cannot obtain a reasonable approximate probability in this case

45. Assume that the annual salaries of elementary school teachers in Canada are normally distributed with a mean of $36,000 and standard deviation of $4000. In a school with 100 teachers, approximately how many teachers make more than $42,000 a year?

    (a) * 7
    (b) 72

(c) 8

(d) 18

(e) 42

46. A population has mean 40, standard deviation 20, and a strongly right-skewed shape. A sample of 10 items is taken from this population. What can you say about the sampling distribution of the sample mean?

    (a) it will have a strongly right-skewed shape

    (b) * it will have a right-skewed shape, but not as much so as the population.

    (c) we cannot say anything about its shape.

    (d) it will have a normal shape

47. In the situation of Question 46, what can you say about the sample mean itself?

    (a) it could be above 60

    (b) * it will probably be somewhere near 40

    (c) we need a larger sample to be able to say anything

    (d) it will be exactly 40

48. Which part of the hypothesis testing procedure determines the value of the P-value?

    (a) * All of the other alternatives.

    (b) The sampling distribution of the test statistic.

    (c) The test statistic

    (d) The alternative hypothesis.

49. The increasing cost of health care is an important issue today. Suppose that a random sample of 22 small companies that offer paid health insurance as a benefit was selected. The mean health insurance cost per worker per month in the sample was \$156. The sample standard deviation of the health cost per worker per month was \$36. The upper limit of a 99% confidence interval for the mean health cost per worker per month for all small companies is:

    (a) \$171.04

    (b) \$175.80

    (c) \$134.30

    (d) \$171.90

    (e) * \$177.70

50. In a large Statistics class (not this one), the probability that a randomly-selected student was born in India is 0.07. 10 students are independently and randomly selected. What is the probability that at least one of them was born in India?

    (a) 0.48

    (b) * 0.52

    (c) 0.30

    (d) 0.07

    (e) 0.93

51. A manufacturing company claims that its new floodlight will last 1200 hours. After collecting a simple random sample of size 20, you determine that a 98% confidence interval for the true mean number of hours that the floodlights will last, μ, is from 972 to 1101.

From this information, three statements are made concerning tests of significance about μ. Assume all the tests mentioned are two-sided.

I: We have sufficient evidence to reject the null hypothesis that $\mu = 1200$ against $H_a : \mu \neq 1200$ at both the 2% and the 5% significance level.

II: If a 99% confidence interval for the mean number of hours the floodlights will last were computed, the numerical value 992 would lie in this interval.

III: If we were interested in testing the null hypothesis $H_0 : \mu = 991$ against $H_a : \mu \neq 991$, we know that the P-value must be less than 0.01.

Which of these statements are true?

(a) Only II and III are true

(b) Only one statement is true

(c) * Only I and II are true

(d) Only I and III are true

(e) All the statements are true

52. In Ontario as a whole, over the last year, 56% of all babies born were boys. A government agency selects 100 birth records from this period, and finds that 53% of these babies were boys. We are interested only in the province of Ontario. Are the values 56% and 53% parameters or statistics?

(a) 56% is a statistic and 53% is a parameter.

(b) 56% and 53% are both parameters.

(c) 56% and 53% are both statistics.

(d) * 56% is a parameter and 53% is a statistic.

53. The stemplot below shows salaries for a company:

```
Stem-and-leaf of salary     N   = 20
Leaf Unit = 1000

     1     1 8
     3     2 39
     8     3 25578
    (5)    4 04567
     7     5 144
     4     6 2
     3     7 88
     1     8 8
```

Here are three statements about the data shown in the stemplot:

I: The median equals $44,500.

II: The inter-quartile range equals $19,000.

III: The 40% trimmed mean equals $43,750.

Which of these statements are true?

(a) I and III are true.

(b) * All of the statements are true.

(c) Fewer than two statements are true.

(d) II and III are true.

(e) I and II are true.

54. A special roulette wheel has seven equally likely outcomes: 0, 1, 2, 3, 4, 5, 6. If you bet that an odd number comes up, you win or lose \$10 according to whether or not that event occurs. If X denotes your net gain for one play, $X = 10$ if we see 1, 3, or 5, and $X = -10$ otherwise. (That is, 0 counts as an even number.) Suppose that you play this game 400 times. Let Y be your total net gain after these 400 plays. The mean (expectation) and standard deviation of Y are, respectively (accurate to the number of figures shown):

(a) * -571; 198

(b) -571; 988

(c) -270; 198

(d) none of the other answers displayed

(e) -270; 988

55. Three of the following statements are valid conclusions that may be properly justified by the Central Limit Theorem, as presented in this course.

I: The mean weight of apples in a simple random sample of 72 (standard size) boxes of apples from a large population will be approximately normally distributed.

II: 2000 students write an exam. The IQR of the test scores for a simple random sample of 100 students selected from all 2000 students has an approximately normal distribution.

III: The proportion of part time students at U of T calculated from a simple random sample of 800 students is approximately normally distributed.

IV: A simple random sample of families is selected from all Ontario families and family income for the year 2006 is recorded for each family. In Ontario, family incomes possess a moderately skewed distribution. If we select a simple random sample of 400 families and record each family's income, the histogram of these 400 incomes will be approximately normal.

V: The number of people who owned a cellular phone in 2007, in a simple random sample of 550 people, selected from all adults residing in Toronto, will be approximately normally distributed.

Which three statements are correct?

(a) I, II, and III

(b) II, IV, and V

(c) I, IV and V

(d) II, III and V

(e) * I, III and V

56. Mamma Luisa bakes six pies a day that cost \$3 each to produce. On 26% of the days she sells only two pies. On 57% of the days, she sells 4 pies, and on the remaining 17% of the days, she sells all six pies. If Mamma Luisa sells her pies for \$6 each, what is her expected net profit for a day?

(a) * \$4.92

(b) \$3.82

(c) -\$7.20

(d) \$7.92

(e) -\$8.10