University of Toronto Scarborough
STAB22 Final Examination

August 2007

This examination is multiple choice. Ensure that you have a Scantron answer sheet and a #2 pencil, and complete the Scantron sheet according to the instructions (otherwise your exam may not be marked).

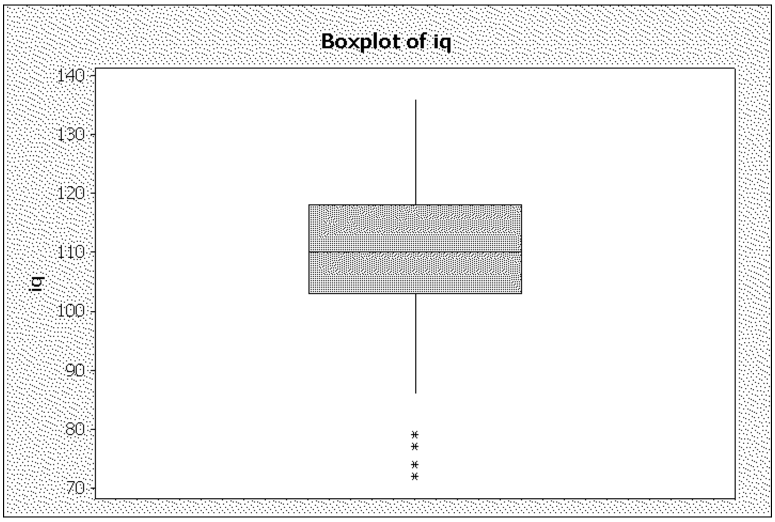
For this examination, you are allowed two letter-sized sheet of notes (both sides), handwritten and prepared by you, a non-programmable, non-communicating calculator, and writing implements.

Mark the answer that is most nearly correct from the alternatives 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.

This examination has xx numbered pages; before you start, check to see that you have all the pages.

1. Educational data was recorded for 78 grade 7 students. For the students' IQ, a boxplot was drawn, as shown below.



What is the interquartile range?

(a) 64
(b) 50
(c) * 15
(d) 7.5

2. For the same students as in Question 1, each student was assessed on the Piers-Harris Children's Self-Concept Scale, a psychological test. A histogram of the self-concept scores is shown below.

Describe the shape of this distribution of scores.

(a) approximately symmetric
(b) skewed to the right

(c) * skewed to the left

(d) cannot tell shape from a histogram

3. Look again at the histogram in Question 2. How do the mean and median compare for this data set?

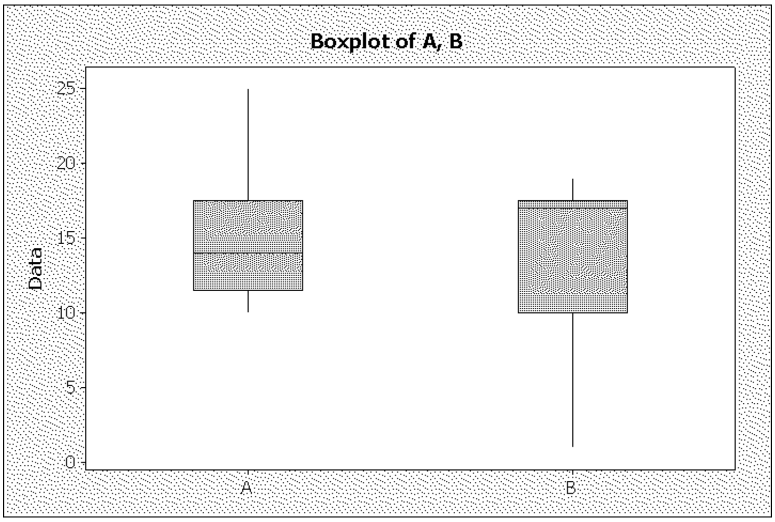
   (a) * The median is bigger than the mean.

   (b) The mean is bigger than the median.

   (c) The mean and median are about the same.

   (d) It is impossible to compare the mean and median using this histogram.

4. As part of an experiment, two groups of data were collected: for treatment A and treatment B. Boxplots for the two data distributions of the response variable are shown below.



The researcher wants to compare the mean response value for the two groups. Which of the statements below best describes the situation?

   (a) the means for the two groups are about the same.

   (b) * boxplots do not show means, so we cannot reasonably compare the groups.

   (c) the mean for group A is bigger.

   (d) the mean for group B is bigger.

5. Calorie contents of a number of different hot dogs were measured. Summary statistics for the calorie content of 20 beef hot dogs are as shown below:

```
Descriptive Statistics: Beef calories

Variable        N  N*    Mean  SE Mean  StDev  Minimum      Q1  Median      Q3
Beef calories  20   0  156.85     5.06  22.64   111.00  139.50  152.50  179.75
```

```
Variable        Maximum
Beef calories    190.00
```

Do a calculation to decide whether the highest and lowest values are outliers in this data set (that is, do two calculations). What do you conclude?

(a) The minimum value is an outlier, but the maximum is not.
(b) The maximum value is an outlier, but the minimum is not.
(c) * Neither minimum nor maximum value is an outlier.
(d) Both minimum and maximum values are outliers.

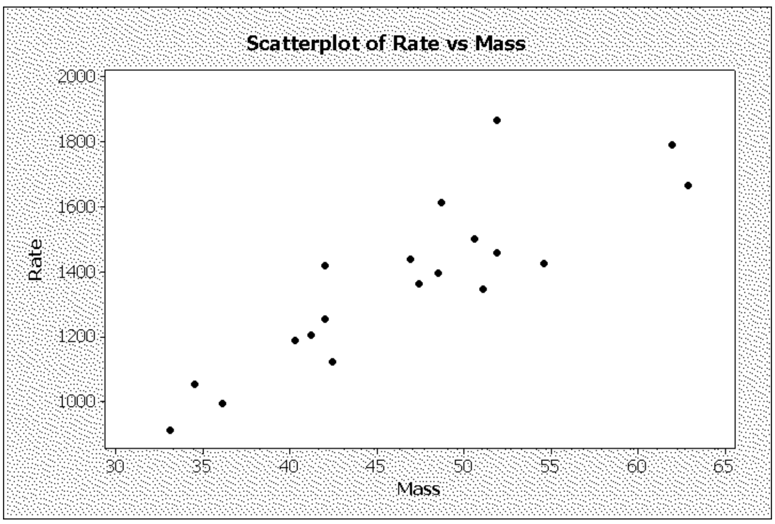
6. A physicist measures a number of temperatures in degrees Celsius. The mean is 5 and the standard deviation is 15. The physicist then converts the temperatures to Kelvin degrees by adding 273. Can you say, without looking at the original data, what the mean and SD of the converted temperatures will be?

   (a) The mean will be 5 and the SD will be 288.
   (b) * The mean will be 278 and the SD will be 15.
   (c) The mean will be 278 and the SD will be 288.
   (d) We cannot say without looking at the original data.

7. A normal distribution has mean 50 and standard deviation 4. Approximately what proportion of this normal distribution lies above 58?

   (a) * 0.025
   (b) 0.05
   (c) 0.95
   (d) 0.975

8. In a study of dieting, the body mass and resting metabolic rate were recorded for 19 people. A scatterplot is shown below.



What is the correlation between body mass and resting metabolic rate?

(a) close to 0 because there are outliers
(b) * 0.85
(c) −0.75
(d) 0.35

9. In the same study as in Question 8, some output is given from a regression predicting metabolic rate from body mass:

```
Regression Analysis: Rate versus Mass

The regression equation is
Rate = 113 + 26.9 Mass


Predictor     Coef  SE Coef     T      P
Constant     113.2    179.6  0.63  0.537
Mass        26.879    3.786  7.10  0.000


S = 133.075   R-Sq = 74.8%   R-Sq(adj) = 73.3%
```

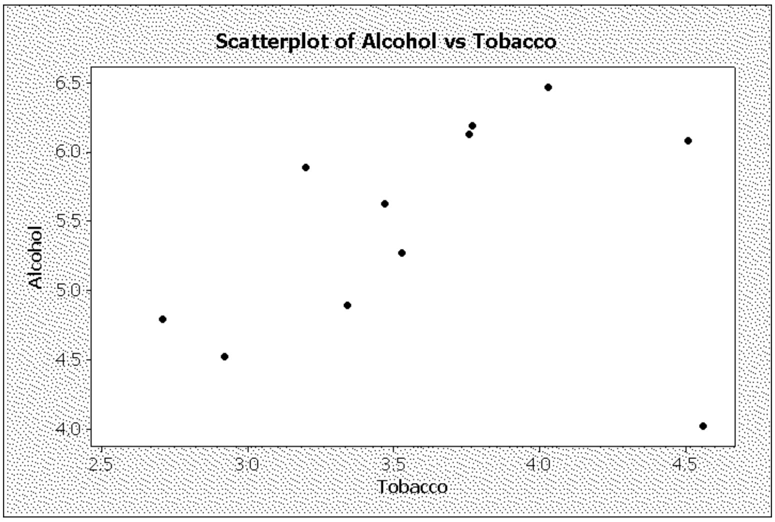
Which number from this output gives you the increase in metabolic rate associated with an increase in body mass of 1 unit?

(a) 113
(b) * 26.9
(c) 133.075
(d) 74.8%

10. Using the output in Question 9, what is the predicted metabolic rate for a person of mass 45?

(a) * 1320
(b) 1600
(c) 113
(d) 1000

11. The British government measures average average household weekly spending on alcohol and tobacco for each of 11 regions of the United Kingdom. A scatterplot for 1981 data is shown below (all figures are in pounds). In the survey, Northern Ireland had average household spending of 4.02 pounds on alcohol and 4.56 pounds on tobacco.



Suppose the correlation between alcohol and tobacco spending is calculated for two cases: (i) when Northern Ireland is included in the calculation, and (ii) when Northern Ireland is excluded. How do the correlations in cases (i) and (ii) compare?

(a) the correlations in cases (i) and (ii) are about the same.

(b) the correlation in case (i) is much closer to +1 than in case (ii).

(c) * the correlation in case (ii) is much closer to +1 than in case (i).

(d) the correlation in case (ii) is much closer to −1 than in case (i).

(e) the correlation in case (i) is much closer to −1 than in case (ii).

12. A piece of software is used to assess computer skills. One activity tests the user's skill with the mouse. A circle (of fixed size) is drawn in a random location on the screen, and the user has to click inside the circle. Two items are recorded: the amount of time (in milliseconds) that the user takes to click inside the circle, and the distance of the circle's centre from the cursor position. It is suspected that a user will take more time to click inside the circle if it is far away from the cursor (that is, the distance is greater).

One user has 20 trials with each hand (right and left). The regression line (for predicting time from distance) is calculated for all 40 trials, and the residuals are calculated. For the trials with the user's left hand, the residuals have mean 56.25 and SD 30; for the trials with the user's right hand, the residuals have mean −56.25 and SD 25.

What can you say about the regression?

(a) the residuals must be plotted against the original explanatory variable.

(b) the regression should be done with each hand separately because the user is obviously left-handed.

(c) * the regression should be done with each hand separately because the user is obviously right-handed.

(d) the residuals show no problem with the regression.

13. The National Halothane Study was a study of the safety of anesthetics used in surgery. Records of many surgeries were examined, the anesthetic used was noted, along with whether the patient survived. Death rates for the four different anesthetics used were as follows:

Anesthetic	A	B	C	D
Death rate	1.7%	1.7%	3.4%	1.9%

What can we conclude?

(a) This is a statistical experiment, so we can conclude that anesthetic C causes a higher death rate.

(b) * Anesthetic C might have been used on the most critical cases, so is not necessarily the most dangerous.

(c) This is anecdotal evidence, so we cannot conclude anything.

(d) This is an observational study, so anesthetic C is definitely not the most dangerous.

14. A study is to be made, comparing a promising new medical treatment with a control. How would you recommend that the researchers decide which subjects receive which treatment?

(a) The subjects whose cases are most serious have the greatest need for the new treatment, and so they should receive it.

(b) Each subject who receives the treatment should be matched up with a subject of the same age who is in the control group.

(c) Put all the males in one group and all the females in the other.

(d) * Use a method that does not rely on which subject is being considered, and does not rely on the researchers' judgement.

15. An experimenter claims that a meditation technique lowers the anxiety level. The experimenter conducts an experiment in which subjects' anxiety levels are measured at the beginning and end of a month, and during the month the subjects are taught how to meditate and given frequent practice at meditating. At the end of the month, the experimenter discovered that anxiety levels had been reduced. How do you react to this experiment?

(a) * The experimenter also needed to have a group of subjects that remained quiet without meditating.

(b) The experiment offers convincing evidence that meditation causes a reduction in anxiety.

(c) By comparing anxiety levels at the beginning and end of the month, the experimenter is properly allowing for the fact that some people are more anxious than others, and that teaching these people to meditate has an effect on their anxiety.

(d) This experiment is a load of nonsense because we all know that meditation has no effect.

16. A utility company is trying to learn whether providing households with electricity meters reduces their consumption of electricity. An executive from the utility company says "If we compare electricity consumption from last year, before the meters were introduced, with electricity consumption this year, that will tell us whether the meters are effective." Which of the statements below do you think is the best response?

    (a) Comparing this year with last year is a fair way to see whether the meters are reducing electricity consumption.

    (b) This study won't work because there is no control group.

    (c) * It would be better to choose some households at random, install meters there, and compare electricity consumption this year with the households that do not have meters installed.

    (d) This study suffers from lack of realism.

17. A community has 60% males and 40% females living in it. Which is the best way of taking a sample of 100 people from this community?

    (a) Take a simple random sample, because this will have about 60 men and about 40 women.

    (b) * Choose 60 people at random from the men, and 40 people at random from the women.

    (c) Take a simple random sample of phone numbers from the community.

    (d) Put out an appeal for volunteers, and sample the first 100 people who respond.

18. A systematic sample of size 10 is to be taken from a population of size 1000 by giving each experimental unit a number, choosing at random a unit numbered 1–100, and then sampling each 100th unit. Which statement below is true?

    (a) * Each experimental unit has an equal chance to be selected.

    (b) Each experimental unit has an equal chance to be selected, independently of the other experimental units that have been selected.

(c) If the experimental unit numbered 71 is in the sample, the experimental unit numbered 269 might be in the sample.

(d) If the experimental unit numbered 71 is in the sample, the experimental unit numbered 471 cannot be in the sample.

19. In an experiment to determine how caffeine affects our bodies, subjects were asked to push a button as quickly as they could after taking a caffeine pill and again after taking a placebo pill. (Some subjects had the caffeine pill first and some had the placebo pill first; the order was randomized.) The mean reaction time for the subjects taking the caffeine pill was 158 milliseconds, and the mean reaction time for the subjects taking the placebo pill was 197 milliseconds.

    Consider the numbers 158 and 197. Are they statistics or parameters?

    (a) They are both parameters.

    (b) * They are both statistics.

    (c) 158 is a statistic and 197 is a parameter.

    (d) 158 is a parameter and 197 is a statistic.

20. For a certain population, the sampling distribution of the sample mean, for a sample of size 50, is found to have mean 100 and SD 20. Which of the following calculations would use the sampling distribution (you do not need to do the calculations)?

    (a) * The probability that the mean of a sample of size 50 lies between 80 and 90.

    (b) The probability that the population mean is less than 95.

    (c) The probability that a value from this population would be bigger than 110.

    (d) The probability that any sample mean would lie between 105 and 115.

21. A coin is tossed three times, and the number of heads is counted. What is the sample space?

    (a) $S = \{HHH, HHT, HTH, HTT, THH, THT, TTH, TTT\}$.

    (b) * $S = \{0, 1, 2, 3\}$.

    (c) $S = \{\text{head}, \text{tail}\}$.

    (d) cannot say because we don't know whether the coin is fair.

22. Events A and B are such that $P(A) = 0.3$, $P(B) = 0.4$. If event A happens, then event B cannot happen. What is the probability of "either A or B or both"?

    (a) cannot say because the events are not disjoint

    (b) cannot say because the events may not be independent.

    (c) 0.3×0.4

(d) * $0.3 + 0.4$

23. A random variable X can take values 1, 2, 3 or 4 with probabilities 0.6, 0.2, 0.15 and 0.05 respectively. What is the mean of X?

    (a) 2.5
    (b) cannot say because we need the variance of X in order to calculate the mean
    (c) less than 1
    (d) * 1.65

24. A fair coin is tossed. Which of the following sequences of heads and tails is most likely (has the highest probability): (i) HTHHTT (ii) HTHH (iii) THHHHH?

    (a) (i)
    (b) (i) and (ii) are equally likely.
    (c) * (ii)
    (d) (iii)

25. Which is the best statement of the law of large numbers?

    (a) When a very large sample is drawn, the sample mean will be exactly equal to the population mean.
    (b) The sampling distribution of the sample mean will be approximately normal in shape.
    (c) * When a large sample is drawn, the sample mean is very likely to be close to the population mean.
    (d) Suppose we toss a fair coin, and get 14 heads in the first 20 tosses. Then, since the sample proportion of heads should be close to 0.5, we are more likely to get a tail on the next coin toss.

26. If U and V are random variables, which of the following can you say?

    (a) The standard deviation of $U + V$ is the sum of the standard deviations of U and V. It does not matter whether U and V are independent.
    (b) The standard deviation of $U + V$ is the sum of the standard deviations of U and V, if U and V are independent.
    (c) The variance of $U + V$ is the sum of the variances of U and V. It does not matter whether U and V are independent.
    (d) * The variance of $U + V$ is the sum of the variances of U and V, if U and V are independent.

27. A mail-order company claims to ship 95% of its orders on time. 4000 orders were received last week. You take a sample of 120 of these orders, and find that 108 were shipped on time. What is the sample proportion of orders shipped on time?

   (a) 3800
   (b) 108
   (c) 0.95
   (d) * 0.90

28. Suppose that X is the number of baby girls out of the next 50 children born at a local hospital. (Twins and other multiple births are not counted, and you can assume that any of these children are equally likely to be girls).

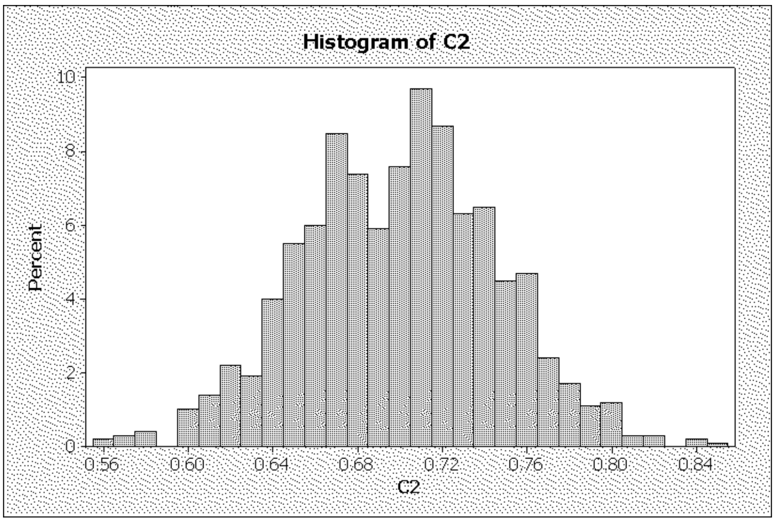
   Independently of the above, a couple decides to have children until their second boy is born; suppose that their total number of children is Y.

   What can you say about the applicability of the binomial distribution in these cases?

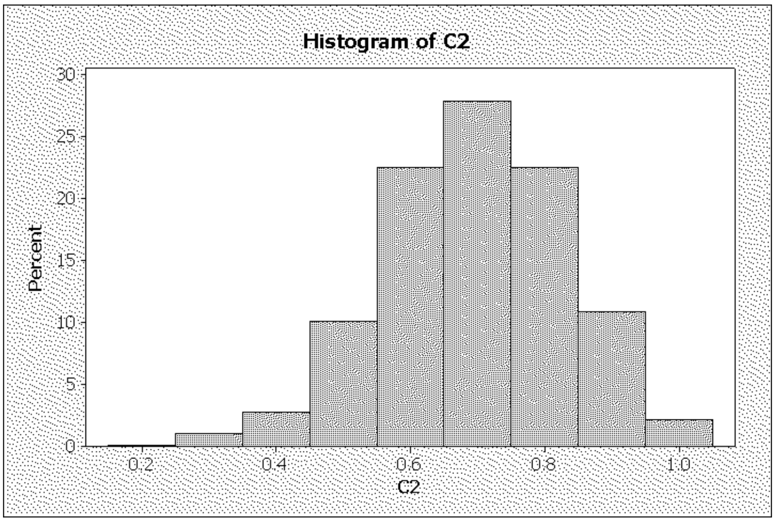
   (a) * X has a binomial distribution and Y does not.
   (b) Y has a binomial distribution and X does not.
   (c) X and Y both have binomial distributions.
   (d) Neither X nor Y has a binomial distribution.

29. A simple random sample of size 100 was drawn from a population containing a proportion 0.7 of successes. Of the five histograms below, which shows the simulated sampling distribution of the sample proportion?

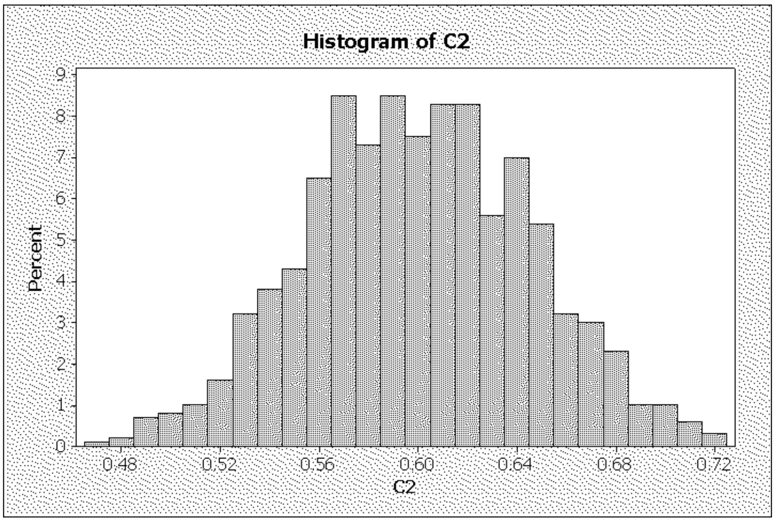
   Histogram 1:



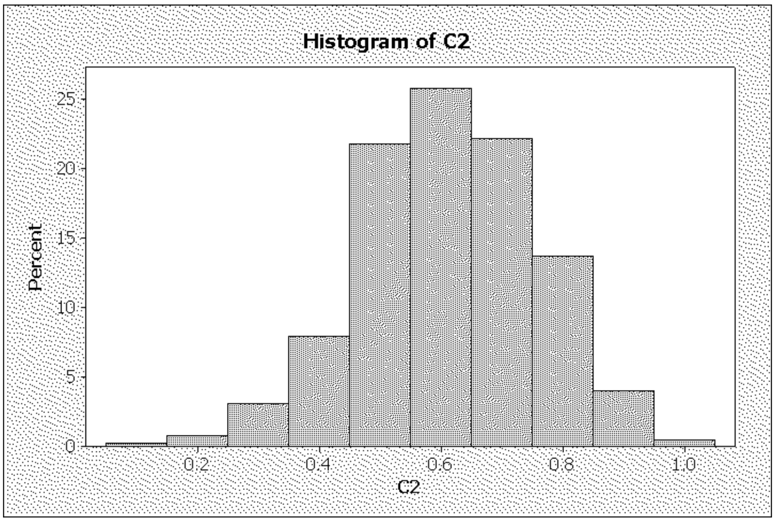
   Histogram 2:



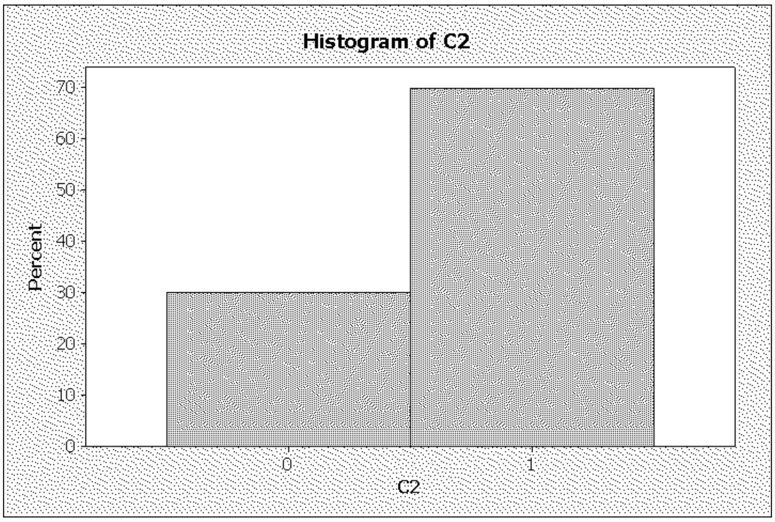
Histogram 3:



Histogram 4:



Histogram 5:



(a) * Histogram 1
(b) Histogram 2
(c) Histogram 3
(d) Histogram 4
(e) Histogram 5

30. Which is the best statement of the Central Limit Theorem?

(a) When a large sample is drawn, the sample mean is very likely to be close to the population mean.
(b) When a very large sample is drawn, the sample mean will be exactly equal to the population mean.
(c) Suppose we toss a fair coin, and get 14 heads in the first 20 tosses. Then, since the sample proportion of heads should be close to 0.5, we are more likely to get a tail on the next coin toss.
(d) * The sampling distribution of the sample mean will be approximately normal in shape.

31. A population with mean μ has a shape which is very much skewed to the right. A sample of size $n = 20$ is taken. Which of the following statements is the most accurate?

(a) The sampling distribution of the sample mean does not have mean μ because the population has a skewed shape.
(b) The Central Limit Theorem says that the sampling distribution of the sample mean must have a normal shape.
(c) * The sampling distribution of the sample mean might not be very normal in shape.
(d) The Central Limit Theorem does not apply to skewed populations.

32. The level of nitrogen oxides in the exhaust of a particular car has mean 1.1 grams per kilometre and SD 0.2 grams per kilometre. The shape of this distribution is slightly non-normal.

    A sample of 40 cars of this type is taken, and the nitrogen oxide level is measured for each car. What is the probability that the sample mean nitrogen oxide level exceeds 1.3 grams per kilometre?

    (a) 0.9992
    (b) 0.6915
    (c) 0.3085
    (d) * 0.0008
    (e) can't calculate because the population distribution is not normal.

33. A 95% confidence interval is an interval that:

    (a) has probability 0.95 of containing the population parameter for the one sample that you observe.
    (b) has probability 0.95 of containing the sample statistic.
    (c) * contains the population parameter value in 95% of all possible samples.
    (d) contains 95% of the population.

34. A hotel chain wanted to learn about the level of experience of its general managers. A random sample of 14 general managers was taken, and these managers had a mean of 11.72 years of experience. Suppose that the standard deviation of years of experience for all general managers in the chain is known to be 3.2 years.

    What is the lower limit of a 95% confidence interval for the mean experience of all general managers in this hotel chain?

    (a) * 10.04 years
    (b) 13.40 years
    (c) 5.45 years
    (d) 9.89 years

35. In Question 34, suppose that we had not known the standard deviation of years of experience for all general managers in the chain. What would you do to calculate the confidence interval?

    (a) Use the sample SD and z^* from the normal distribution.
    (b) It is impossible to calculate a confidence interval for the population mean if the population SD is not known.
    (c) * Use the sample SD and the t distribution.

(d) Do the same thing as in Question 34.

36. A taxation agency estimates the mean income of all residents of a certain city by taking a simple random sample of 200 residents. For each sampled resident, the income given on the most recent tax form is recorded. The resulting 95% confidence interval has margin of error $300, and the city's SD of incomes is known to be $2165.

    How large a sample would be needed to make the margin of error $150?

    (a) 200
    (b) 283
    (c) 400
    (d) * 800

37. Hypothesis tests are often done by comparing the calculated P-value with a quantity α. What does α represent?

    (a) The difference between the sample mean and the null hypothesis.
    (b) The probability that the null hypothesis is true.
    (c) The probability that the alternative hypothesis is false.
    (d) The probability of not rejecting the null hypothesis when it is false.
    (e) * The probability of rejecting the null hypothesis when it is true.

38. Experiments on animal learning often measure how long a mouse takes to get through a maze. For one particular maze, the mean time is 18 seconds. A researcher thinks that a loud noise will help the mice to complete the maze faster. She meaures the average time $\bar{x}$ required for 10 mice to complete the maze when the loud noise is played.

    Which is a suitable null hypothesis in this situation?

    (a) $\bar{x} = 18$
    (b) $\bar{x} < 18$
    (c) * $\mu = 18$
    (d) $\mu < 18$

39. Refer again to the situation described in Question 38. What would be a suitable alternative hypothesis?

    (a) $\bar{x} = 18$
    (b) $\bar{x} < 18$
    (c) $\mu = 18$
    (d) $\mu \neq 18$
    (e) * $\mu < 18$

40. A test of significance with $\alpha = 0.05$ results in a P-value of 0.1. What do you conclude?

  (a) The null hypothesis is rejected.
  (b) * The null hypothesis is not rejected.
  (c) The alternative hypothesis is rejected.
  (d) The alternative hypothesis is not rejected.
  (e) The null hypothesis is accepted.

41. Historically, the mean yield of corn in the United States has been 120 bushels per acre. A survey of 40 farmers this year gives a sample mean yield of 125 bushels per acre. In the past, the standard deviation of yields has been 12 bushels per acre.

    Let μ be the mean yield of corn nationally for this year. Supposing that the past standard deviation is still correct, what is your P-value for testing a null hypothesis of $\mu = 120$ against an alternative of $\mu \neq 120$?

  (a) 0.0041
  (b) * 0.0082
  (c) between 0.01 and 0.02
  (d) 2.64
  (e) 125

42. Radon is an odorless, slightly radioactive gas that can accumulate in tightly-sealed homes. There is therefore interest in developing radon detectors for use in homes. In a study of the accuracy of a certain model of detector, 12 detectors are exposed to a known concentration (105 picocuries per litre) of radon. Minitab output from a hypothesis test is shown below.

```
One-Sample Z: Radon reading

Test of mu = 105 vs not = 105
The assumed standard deviation = 9

Variable        N     Mean  StDev  SE Mean          95% CI             Z      P
Radon reading  12  104.133  9.397    2.598  (99.041, 109.225)  -0.33  0.739
```

    What do you conclude about the mean reading of all detectors of this type from the test?

  (a) * It could be 105.
  (b) It is equal to 105.
  (c) It is definitely not 105.

(d) It could be anywhere between 99 and 109.

43. A hypothesis test at $\alpha = 0.05$ for the null hypothesis that $\mu = 30$ against the alternative that $\mu \neq 30$ has a P-value of 0.02. What can you say about the 95% confidence interval for μ?

    (a) Since $\mu \neq 30$, 33 is inside the interval.
    (b) 30 is inside the interval.
    (c) * 30 is outside the interval.
    (d) There is no connection between the test and the confidence interval.

44. Which of the following questions does a hypothesis test (test of significance) answer?

    (a) Was the experiment properly designed, or the sample properly taken?
    (b) Is the observed effect important?
    (c) * Is the observed effect due to chance?

45. Scores on a certain standardized test have mean 150 and SD 10. An increase of 5 points does not make any difference in applying to a professional school. A sample of 100 people is given extra materials to use in studying for this test. The sample mean is 155. Suppose that μ is the mean score that all applicants would get if they used the extra materials. A test is done of the null hypothesis $\mu = 150$ against the alternative $\mu > 150$. The P-value is less than 0.0001. What do you conclude?

    (a) There must have been a mistake in calculating the P-value.
    (b) Because the P-value is so small, the extra materials are definitely worth using.
    (c) The sample size is too small to allow any useful conclusions.
    (d) * The extra materials are associated with a gain that is not helpful in practice, so the hypothesis test is pointless.

46. Every semester, a university administrator calculates the mean grade awarded by the university's professors in that semester's courses. There is concern that the mean grade score is changing over time, so the administrator does a hypothesis test that the mean grade score is 2.67 (B-minus) against the alternative hypothesis that the mean has changed. One semester, the result is statistically significant. What should the administrator conclude?

    (a) The administrator can conclude that the mean grade score has changed.
    (b) This test is the best way to see whether grade scores are increasing over time.
    (c) * Because many tests are being done, the "statistically significant" result may not be meaningful.
    (d) The administrator should use an α value of 0.01.

47. In a test of $\mu = 20$ against $\mu > 20$ using a sample size of 5, the null hypothesis is not rejected. What can the researcher conclude?

   (a) The mean is definitely not bigger than 20.
   (b) * The mean could still be bigger than 20.
   (c) Because the P-value is large, the power will also be large.
   (d) If a larger sample size is used, the null hypothesis will still not be rejected.

48. Consider the Minitab output below.

```
Power and Sample Size

1-Sample Z Test

Testing mean = null (versus not = null)
Calculating power for mean = null + difference
Alpha = 0.05  Assumed standard deviation = 8

            Sample
Difference    Size     Power
         5      30  0.928308
```

   What do you conclude?

   (a) If the null hypothesis is true, the probability of failing to reject it is 0.93.
   (b) If the null hypothesis is false in any way, the probability of rejecting it is 0.93.
   (c) * If the mean increases by 5 from its former value, the null hypothesis is likely to be rejected.
   (d) If the sample size is increased (and everything else stays the same), the power will decrease.

49. A hypothesis test is to be carried out at $\alpha = 0.05$, for a population whose SD is 12. Use the output below to decide how big a sample is needed to bring the type II error probability below 0.2.

```
Power and Sample Size

1-Sample Z Test

Testing mean = null (versus not = null)
Calculating power for mean = null + difference
Alpha = 0.05  Assumed standard deviation = 12
```

Difference	Sample Size	Target Power	Actual Power
5	29	0.6	0.611752
5	36	0.7	0.705418
5	46	0.8	0.806758
5	61	0.9	0.902220

(a) 29

(b) 36

(c) * 46

(d) 61

50. A test for the population mean, when the population SD is not known, is carried out based on a sample of size 12. The null hypothesis is that $\mu = 20$, and the alternative is that $\mu < 20$. The test statistic value is -2.45. Using a suitable table, what is the P-value?

    (a) 0.0071

    (b) large enough that the null hypothesis should not be rejected.

    (c) between 0.02 and 0.04

    (d) * between 0.01 and 0.02

51. A test is done under the same circumstances as Question 50, but this time the test statistic is 2.45. What is the P-value?

    (a) 0.9929

    (b) * large enough that the null hypothesis should not be rejected.

    (c) between 0.02 and 0.04

    (d) between 0.01 and 0.02

52. In a study of women's bone health, a random sample of 8 women was taken, and their daily calcium intake measured. The sample mean was 926 milligrams, and the sample standard deviation was 427.2 milligrams. What is the upper limit of a 99% confidence interval for the mean daily calcium intake of all women?

    (a) * 1454.6

    (b) 1432.7

    (c) 1315.7

    (d) 397.4

53. A study was carried out to see whether right-handed people find it easier to turn control knobs in a clockwise or counter-clockwise direction. 25 right-handed people took part in the study. Each person used both of the two kinds of control knob (the order was randomized). For each person and each control knob, the time to move the indicator a fixed distance was recorded. What hypothesis test would be appropriate for this kind of data?

   (a) * Matched pairs t-test
   (b) One-sample t-test
   (c) Two-sample t-test
   (d) One-sample z-test

54. A sample of 20 French teachers attended an intensive summer course in listening skills. Each teacher's listening skills were assessed before the course ("pretest") and after the course ("posttest"). Some Minitab output appears below.

```
Paired T-Test and CI: Posttest, Pretest

Paired T for Posttest - Pretest

             N     Mean    StDev  SE Mean
Posttest    20  28.7500   4.7448   1.0610
Pretest     20  27.3000   5.0378   1.1265
Difference  20  1.45000  3.20321  0.71626

T-Test of mean difference = 0 (vs > 0): T-Value = 2.02  P-Value = 0.029
```

What do you conclude?

   (a) A two-sample t-test should have been done instead.
   (b) There is a statistically significant difference in listening skills.
   (c) There is no statistically significant difference in listening skills.
   (d) * There is a statistically significant improvement in listening skills.

55. A two-sample confidence interval, based on the t distribution, is to be calculated. Suppose $\bar{x}_1$ and $\bar{x}_2$ are the two sample means, s_1 and s_2 the two sample SDs, and n_1 and n_2 are the two sample sizes. The formula is $\bar{x}_1 - \bar{x}_2 + (t^* \times w)$, where t^* is a number from the t table. What should w be?

   (a) $\frac{s_1}{\sqrt{n_1}} + \frac{s_2}{\sqrt{n_2}}$
   (b) * $\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$.
   (c) $\frac{s}{\sqrt{n}}$ where the SD and sample size come from the smaller sample.

(d) $\frac{s}{\sqrt{n}}$ where the SD and sample size come from the sample with the larger SD.

56. A comparative experiment is planned, with a treatment group and a control group. The experiment will use a two-sample t-test. A difference in means of 10 between the two populations is considered meaningful. It is desired to have power 0.8 to detect this difference. The standard deviation of each population is believed to be about 5. How would you interpret the Minitab output below?

```
Power and Sample Size

2-Sample t Test

Testing mean 1 = mean 2 (versus not =)
Calculating power for mean 1 = mean 2 + difference
Alpha = 0.05  Assumed standard deviation = 5

             Sample  Target
Difference     Size   Power  Actual Power
        10        6     0.8      0.876418
```

(a) We should collect 6 observations in total, 3 in each group.
(b) * We should collect 6 observations in each group.
(c) We should collect 6 observations in total, but it does not matter how we divide them between groups.
(d) Since the actual power is greater than 0.8, the sample size can be smaller than 6.

57. A study is comparing the taste of instant coffee with fresh-brewed coffee. 50 individuals (randomly sampled) take part in the study. Each individual tastes both kinds of coffee, and declares which kind he or she prefers. The order in which the two kinds of coffee are given to each individual is randomized.

32 individuals in the sample prefer the fresh-brewed coffee. Use a suitable normal approximation to find a 95% confidence interval for the proportion of all individuals who would prefer fresh-brewed coffee. What is the lower limit of this confidence interval?

(a) 0.77
(b) 0.64
(c) * 0.51
(d) 0.38
(e) 0.25

58. A basketball player made only 41% of her free throws last season. Over the summer, she worked with her coach on practicing her free throws. This season, she made 28 out of 50 free throws. Using the Minitab output below, what do you conclude? Use $\alpha = 0.05$.

```
Test and CI for One Proportion

Test of p = 0.41 vs p > 0.41

                                          95%
                                        Lower    Exact
Sample    X   N  Sample p     Bound  P-Value
1        28  50  0.560000  0.434282    0.023
```

(a) There is no evidence that her free-throw shooting has improved.

(b) * There is evidence that her free-throw shooting has improved.

(c) There is evidence of a change in her free-throw shooting percentage.

(d) A hypothesis test does not help us assess whether her free-throw shooting has improved.

59. Large trees can cause problems when they grow near power lines – when there is a storm, the tree can fall on the power lines and cut off power. One treatment is to apply a chemical that will slow the tree growth; however, the chemical will sometimes cause the tree to die. In an experiment, the chemical killed 38 of 112 trees. Using the Minitab output below, what can you say about the 99% confidence interval for the proportion of all trees that would be killed by the chemical?

```
Test and CI for One Proportion

Test of p = 0.25 vs p not = 0.25
                                      Exact
Sample    X    N  Sample p    P-Value
1        38  112  0.339286      0.037

Test of p = 0.4 vs p not = 0.4
                                      Exact
Sample    X    N  Sample p    P-Value
1        38  112  0.339286      0.210
```

(a) 0.25 and 0.4 are both outside the confidence interval.

(b) * 0.25 and 0.4 are both inside the confidence interval.

(c) 0.4 is inside the confidence interval, but 0.25 is outside.

(d) 0.25 is inside the confidence interval, but 0.4 is outside.

60. A university records the number of male and female applicants to a certain advanced degree program in two different years, as follows:

	Male	Female
1994	64	25
1995	49	16

To test whether the proportion of female applicants (to this program at all universities) is different in 1994 and 1995, a chisquare test will be done. Assuming that these applicants are a random sample from all possible applicants, what would be the expected frequency for female students in 1995?

(a) 16
(b) * 17.3
(c) 23.7
(d) 25

61. It is believed that caring for a pet is a significant factor in survival times after coronary heart disease. For a sample of people with coronary heart disease, it was recorded whether they survived at least one year ("alive") or not ("dead"), and whether they owned a pet. Some Minitab output from the analysis is below.

```
Chi-Square Test: no pet, pet

Expected counts are printed below observed counts
Chi-Square contributions are printed below expected counts

       no pet    pet  Total
alive      28     50     78
        33.07  44.93
        0.776  0.571

 dead      11      3     14
         5.93   8.07
        4.323  3.181

Total      39     53     92

Chi-Sq = 8.851, DF = 1, P-Value = 0.003
```

What can you conclude from this analysis?

(a) there is no evidence of any difference in survival rates for people who do and do not own a pet
(b) the test only allows us to assess whether the proportions of pet owners differ between the people who survived and those who did not.
(c) people are less likely to survive a year if they have a pet
(d) * people are more likely to survive a year if they have a pet

62. On wet pavement, the distance required to stop a vehicle is an important characteristic of a tire. A study was done to compare the stopping distances of four brands of tire (from a fixed speed on a controlled wet pavement). Part of the analysis is shown below.

```
One-way ANOVA: BrandA, BrandB, BrandC, BrandD

Source  DF      SS     MS     F      P
Factor   3   95.36  31.79  4.78  0.012
Error   19  126.47   6.66
Total   22  221.83

S = 2.580


                           Individual 95% CIs For Mean Based on
                           Pooled StDev
Level   N    Mean  StDev  -----+---------+---------+---------+----
BrandA  6  36.167  2.858   (--------*-------)
BrandB  5  38.800  2.168            (--------*---------)
BrandC  6  36.000  3.406  (--------*--------)
BrandD  6  40.833  1.329                    (-------*--------)
                          -----+---------+---------+---------+----
                             35.0      37.5      40.0      42.5

Pooled StDev = 2.580
```

What can you conclude from this analysis? Use $\alpha = 0.05$.

(a) We have evidence that Brand C has the shortest mean stopping distance.

(b) * The four brands differ in mean stopping distance, but we can't say which is the most or least effective brand of tire.

(c) There is no evidence of any differences in mean stopping distance.

(d) We have evidence that Brand A has shorter mean stopping distance than Brand D.