

Question 1: Multiple choice: circle the best answer to each question:

(1)[2 points] A census attempts to include data from

- (a) a simple random sample.
- (b) a stratified random sample.
- (c) all adults of voting age in a population.
- (d) all individuals in a population.
- (e) all special interest groups in a population.

(2)[2 points] 65% of 400 first year Malaspina students surveyed said that the experience of first year students is positive. This value of 65% is

- (a) a parameter.
- (b) a voluntary response.
- (c) a convenience sample.
- (d) a completely randomized sample.
- (e) a statistic.

(3)[2 points] In what type of data gathering are treatments imposed on individuals?

- (a) a census.
- (b) a sample survey.
- (c) an experiment.
- (d) a an observational study.
- (e) a telephone poll.

(4)[2 points] In a recent study of 83 heavy trucks, 31% were found to have tire pressure less than 80% of the recommended level. The heavy trucks in this example represent

- (a) explanatory variables.
- (b) response variables.
- (c) individuals.
- (d) strata.
- (e) a census.

(5)[2 points] A surveyor is attempting to assess the popularity of Britney Spears by standing outside a music store and asking customers if they would or would not buy her latest release. The sample in this case is

- (a) a simple random sample.
- (b) a voluntary response sample.
- (c) a convenience sample.
- (d) a matched sample.

(e) an unbiased sample.

(6)[2 points] A researcher wants to use a computer program based random number generator to select an SRS of size 5 from a population of 100 labelled  $0, 1, \dots, 99$ . The random number generator returns the five values  $0, 1, 2, 3, 4$ . Assuming the computer program is working correctly, the researcher should

(a) reject the random numbers since they're obviously not random.

(b) use the given random numbers since that combination is just as likely to come up as any other combination of five numbers.

(c) start again using a bigger sample size.

(d) keep drawing random numbers, keeping only the even ones.

(e) keep drawing random numbers until five come up that look truly random.

(7)[2 points] Using line 117 of the table of random digits (use Table A from your book) we select an SRS of size 4 from a class of 30 students labelled  $00, 01, \dots, 29$ . What individuals make up the SRS?

(a) 38, 16, 79, 85

(b) 16, 18, 06, 23

(c) 16, 18, 26, 21

(d) the table cannot be used since all the numbers are five digits long.

(e) the table cannot be used because the individuals are labelled improperly.

(8)[2 points] A random survey 387 workers at a national polling firm found that 39% of them disliked participating in surveys. Assuming a 95% level of confidence, the margin of error in this study is

(a) approximately 95%.

(b) approximately 39%.

(c) exactly 5.1%.

(d) approximately 5.1%.

(e) cannot be determined because the sample is biased.

} NOTE: EARLIER  
VERSION OF SOLUTIONS  
INCORRECTLY GAVE  
"C" AS THE ANSWER

(9)[2 points] Suppose two surveys are done to determine the percentage of the population in favour of new transit fees. The second survey uses a sample size four times as large as the first. The margin of error in the second survey is approximately

(a) four times the size of the first.

(b) twice the size of the first.

(c) half the size of the first.

(d) one quarter the size of the first.

(e) cannot be determined because we need at least one sample size to be given.

(10)[2 points] Using a completely randomized experimental design, we wish determine the effect of washing clothes using different combinations of cleaning agents. The combinations are: soap alone, soap plus bleach, and finally soap plus bleach plus baking soda. Water temperature may have an effect on the outcome, so the clothes will be washed at different temperatures: cold, warm, and hot water. There are 9 treatments in total. The explanatory variables in this case are

- (a) water temperature and cleaning agents used.
- (b) nine in number.
- (c) lurking.
- (d) the same as the response variables.
- (e) cannot be determined because of the randomized design.

**Question 2** A cell phone company wishes to determine the market share of its competitor, so a random sample of Nanaimo residents is drawn from the telephone book (which does not list cell phone numbers). The individuals are contacted and asked whether they have a cell phone, and if so, with what service provider. 35.8% of the people surveyed state that they use the competitor's service.

(a)[4 points] The margin of error in the study was 2.5% (assuming a 95% confidence level). Approximately how large was the sample? Round your answer to the nearest person.

$$\frac{1}{\sqrt{n}} = 0.025$$

$$\frac{1}{0.025} = \sqrt{n}$$

$$\therefore n = \left(\frac{1}{0.025}\right)^2 = 1600$$

(b)[3 points] How many people sampled said that they use the competitor's service?

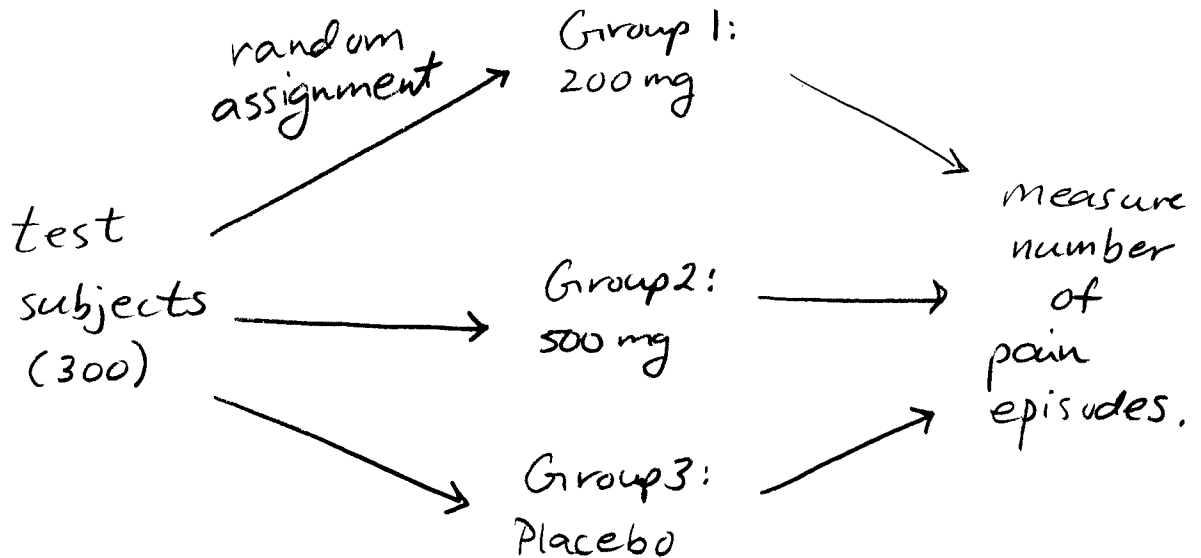
$$\begin{aligned} \text{Using (a), number of people} &= (0.358)(1600) \\ &= 572.8 \\ &\doteq 573 \end{aligned}$$

(b)[3 points] What might cause bias in this result?

The sampling frame is the phone book which does not list cell phone numbers, and so excludes those individuals who use cell phones exclusively - perhaps a substantial part of the competitor's market.

**Question 3** A medical study is designed to test the effectiveness of a new drug which is supposed to reduce the number of pain episodes in arthritis sufferers. A test group of 300 subjects (all arthritis sufferers) will be randomly assigned to three treatments: 200 mg of drug, 500 mg of drug, or a placebo.

(a)[4 points] Draw a diagram to outline the design of the study. Make sure your diagram includes all of the information above.



(b)[3 points] What are the explanatory variables in this case? What possible values can they take?

The explanatory variable is the dose of drug received.

Possible values are 200 mg, 500 mg, 0 mg.

(c)[3 points] What are the response variables? What would be an appropriate way to assign values to them?

The response variable is the number of pain episodes, which could be counted in pain episodes per day following treatment.

**Question 4** In a recent telephone poll, 491 of 1002 eligible Americans voters surveyed indicated that they wished to see the Democrats in the White House next year.

(a)[2 points] State the population, in this survey.

Eligible American voters.

(b)[3 points] Compute  $\hat{p}$ , the estimated proportion of Americans voters who wish to see the Democrats in the White House next year.

$$\hat{p} = \frac{491}{1002} \doteq 0.49 = 49\%$$

(c)[3 points] Calculate the margin of error in this survey assuming a 95% confidence level.

$$\text{margin of error} = \frac{1}{\sqrt{1002}} \doteq 3.2\%$$

(d)[2 points] State your findings, including the estimated proportion of Americans voters who wish to see the Democrats in the White House next year, the sample size, and a confidence statement.

In a recent survey, 49% of eligible American voters indicated that they wished to see the Democrats in the White House next year. This survey used a random sample of 1002 eligible American voters, and has a margin of error of 3.2% nineteen times out of twenty.