

Part 1: Multiple Choice. *Circle the letter corresponding to the best answer.*

1. A new headache remedy was given to a group of 25 subjects who had headaches. Four hours after taking the new remedy, 20 of the subjects reported that their headaches had disappeared. From this information you conclude
 - (a) that the remedy is effective for the treatment of headaches.
 - (b) nothing, because the sample size is too small.
 - (c) nothing, because there is no control group for comparison.
 - (d) that the new treatment is better than aspirin.
 - (e) that the remedy is not effective for the treatment of headaches.
2. We wish to draw a sample of 5 without replacement from a population of 50 households. Suppose the households are numbered 01, 02, . . . , 50, and suppose that the relevant line of the random number table is 11362 35692 96237 90842 46843 62719 64049 17823.
Then the households selected are
 - (a) households 11 13 36 62 73
 - (b) households 11 36 23 08 42
 - (c) households 11 36 23 23 08
 - (d) households 11 36 23 56 92
 - (e) households 11 35 96 90 46
3. A maple sugar manufacturer wants to estimate the average trunk diameter of Sugar Maples trees in a large forest. There are too many trees to list them all and take a SRS, so he divides the forest into several hundred 10 meter by 10 meter plots, selects 25 plots at random, and measures the diameter of every Sugar Maple in each one. This is an example of a
 - (a) multistage sample.
 - (b) stratified sample.
 - (c) simple random sample.
 - (d) cluster sample.
 - (e) convenience sample.
4. A researcher for a consumer products company is field testing a new formula for laundry detergent. He has contracted with 60 families, each with two children, who have agreed to test the product. He randomly assigns 30 families to the group that will use the new formula and 30 to the group that will use the company's current detergent formula. The most important reason for this random assignment is that
 - (a) randomization makes the analysis easier since the data can be collected and entered into the computer in any order.
 - (b) randomization eliminates the impact of any confounding variables.
 - (c) randomization is a good way to create two groups of 30 families that are as similar as possible, so that comparisons can be made between the two groups.
 - (d) randomization ensures that the study is double-blind.
 - (e) randomization reduces the impact of outliers.

5. To test the effect of music on productivity, a group of assembly line workers are given portable mp3 players to play whatever music they choose while working for one month. For another month, they work without music. The order of the two treatments for each worker is determined randomly. This is
- (a) an observational study.
 - (b) a matched pairs experiment.
 - (c) a completely randomized experiment.
 - (d) a block design, but not a matched pairs experiment.
 - (e) impossible to classify unless more details of the study are provided.
6. A survey was done in the town of Mechanicsville to estimate the proportion of cars that are red and made by companies based in Japan. A random sample of 25 cars from a student parking lot at Lee-Davis High School was taken. Which of the following statements is not correct?
- (a) This sample may not be representative of the cars in Mechanicsville because mainly students park at Lee-Davis High School.
 - (b) If the particular parking space is vacant, we can simply select another parking space at random because it is unlikely that a space being vacant is related to the color or manufacturer of the car.
 - (c) It would be an error to simply select the first 25 parking spaces in the lot closest to the auditorium because there are a number of parking spaces there reserved for Drivers Ed vehicles, whose primary color is white.
 - (d) A different team doing the sampling independently would obtain different answers for their sample proportions.
 - (e) The results will be the same regardless of the time of day that the sample is taken.
7. A nutritionist wants to study the effect of storage time (6, 12, and 18 months) on the amount of vitamin C present in freeze dried fruit when stored for these lengths of time. Vitamin C is measured in milligrams per 100 milligrams of fruit. Six fruit packs were randomly assigned to each of the three storage times. The treatment, experimental unit, and response are respectively:
- (a) A specific storage time, amount of vitamin C, a fruit pack
 - (b) A fruit pack, amount of vitamin C, a specific storage time
 - (c) Random assignment, a fruit pack, amount of vitamin C
 - (d) A specific storage time, a fruit pack, amount of vitamin C
 - (e) A specific storage time, the nutritionist, amount of vitamin C
8. A researcher observes that, on average, the number of divorces in cities with Major League Baseball teams is larger than in cities without Major League Baseball teams. The most plausible explanation for this observed association is that the
- (a) presence of a Major League Baseball team causes the number of divorces to rise (perhaps husbands are spending too much time at the ballpark).
 - (b) high number of divorces is responsible for the presence of Major League Baseball teams (more single men means potentially more fans at the ballpark, making it attractive for an owner to relocate to such cities).
 - (c) association is due to the presence of a lurking variable (Major League teams tend to be in large cities with more people, hence a greater number of divorces).
 - (d) association makes no sense, since many married couples go to the ballpark together.

(e) observed association is purely coincidental. It is implausible to believe the observed association could be anything other than accidental.

9. Control groups are used in experiments in order to

- (a) control the effects of outside variables on the outcome.
- (b) control the subjects of a study to ensure that all participate equally.
- (c) guarantee that someone other than the investigators, who have a vested interest in the outcome, controls how the experiment is conducted.
- (d) achieve a proper and uniform level of randomization.
- (e) reduce the variability in results.

10. A survey is to be administered to recent graduates of a certain nursing school in order to compare the starting salaries of women and men. For a random sample of graduates, three variables are to be recorded: sex, starting salary, and area of specialization. Which of the follow best describes a conclusion that can be drawn from this study?

- (a) Whether being female causes graduates of this nursing school to have lower (or higher) starting salaries than males.
- (b) Whether being female causes graduates in this sample to have lower (or higher) starting salaries than males.
- (c) Whether choosing certain area of specialization causes females graduates of this nursing school to have lower (or higher) starting salaries than males.
- (d) Whether there is an association between sex and starting salary among graduates of this nursing school.
- (e) Whether there is an association between sex and starting salary at all nursing schools similar to this one.

Part 2: Free Response

Show all your work. Indicate clearly the methods you use, because you will be graded on the correctness of your methods as well as on the accuracy and completeness of your results and explanations.

11. Read the following brief article about aspirin and alcohol.

Aspirin may enhance impairment by alcohol

Aspirin, a long time antidote for the side effects of drinking, may actually enhance alcohol's effect, researchers at the Bronx Veterans' Affairs Medical Center say. In a report on a study published in the *Journal of the American Medical Association*, the researchers said they found that aspirin significantly lowered the body's ability to break down alcohol in the stomach. As a result, five volunteers who had a standard breakfast and two extra-strength aspirin tablets an hour before drinking had blood alcohol levels 30 percent higher than each had when they drank alcohol alone. Each volunteer consumed the equivalent of a glass and a half of wine.

That 30 percent could make the difference between sobriety and impairment, said Dr. Charles S. Lieber, medical director of the Alcohol Research and Treatment Center at the Bronx center, who was co-author of the report with Dr. Risto Roine.

(a) Explain why this is an experiment and not an observational study.

This is an experiment b/c treatments (aspirin & alcohol, alcohol only) are imposed.

(b) Identify the explanatory and response variables.

Exp - Aspirin consumption

Resp - blood alcohol content

(c) Identify the experimental design used in this study. Justify your answer.

Matched pairs. Each subject was given both treatments & thus acted as his own pair.

(d) In the second sentence above is the phrase, "...researchers said they found that aspirin significantly lowered the body's ability to break down alcohol..." What is the statistical meaning of the word "significantly" in the context of this study?

Significantly means that the difference found in the subjects blood alcohol level between the two treatments was large enough that it was unlikely to have arisen from chance.

(e) This was a controlled experiment. Describe how it was controlled and explain the purpose of doing so.

Each subject was his or her own control, drinking alcohol alone or alcohol & aspirin. ~~Temporarily~~ Comparing each subject BAC under both treatments allowed the researchers to isolate the impact of aspirin from other variables.

12. High blood pressure adds to the workload of the heart and arteries and may increase the risk of heart attacks. If not treated, this condition can also lead to heart failure, kidney failure, or stroke. We wish to test the effectiveness of Angiotensin-converting enzyme (ACE) inhibitors as a treatment for high blood pressure.

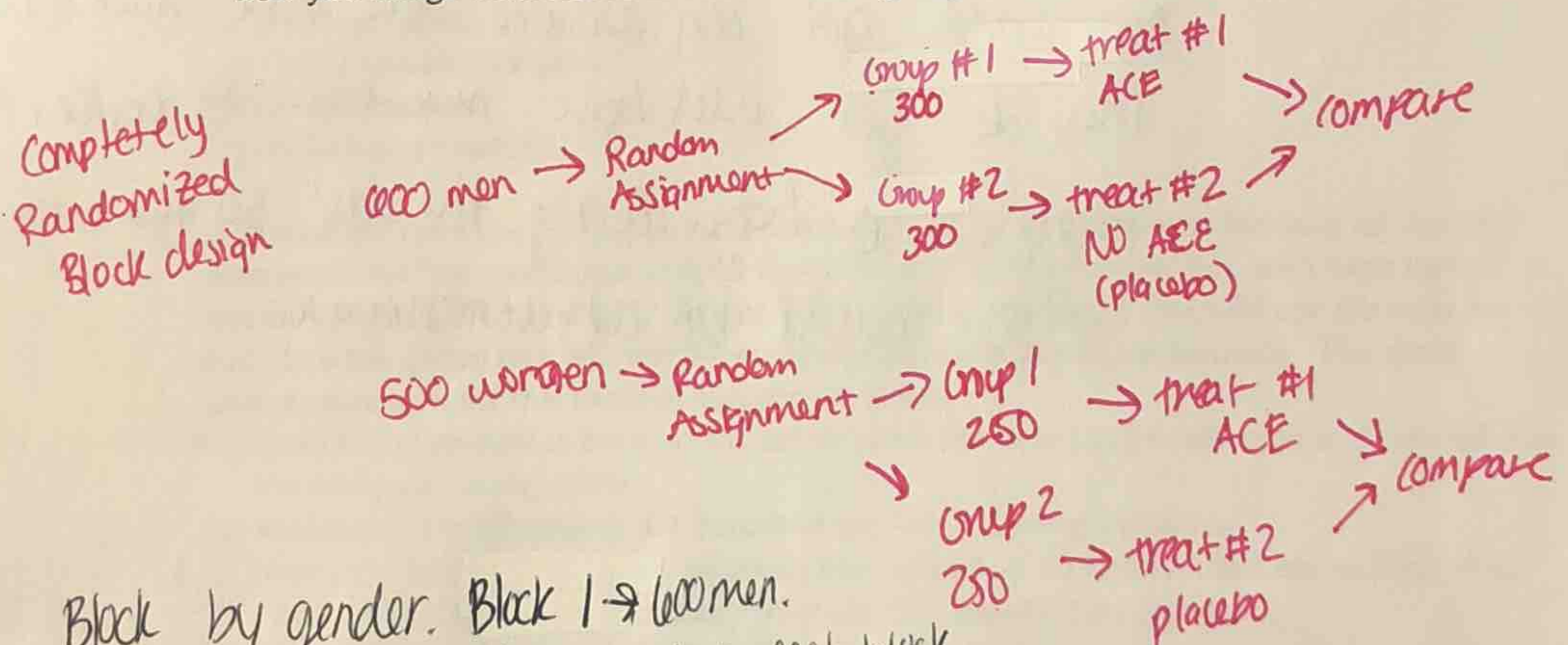
- (a) It is well known that men and women may react differently to common cardiovascular drug treatments. What sort of experimental design would you choose for this study, and why?

Randomized block design by gender will reduce the impact of differences between men and women to treatment that might have variability arising from random assignment.

- (b) Explain why an experiment involving 600 men and 500 women is preferable to one involving 60 men and 50 women.

A larger # of subjects, greater replication, decreases the impact of random variables/impact on experimental results, increasing our ability to distinguish the effects of treatment.

- (c) Assume that 600 men and 500 women suffering from high blood pressure are available for the study. Describe a design for this experiment. Be sure to include a description of how you assign individuals to the treatment groups.



Block by gender. Block 1 → 600 men.
Block 2 → 500 women. Assign within each block a # from 001 to 600 for men and 001 to 500 for women. Choose 3 digit numbers that are distinct from a random number generator until 300 men are in each

group and 250 women are in each group. Group 1 for both will be treated with ACE while Group 2 for both will get the placebo. Compare the changes in blood pressure between the two groups.

13. Bias is present in each of the following sampling designs. In each case, identify the type of bias involved and state whether you think the sample result obtained is lower or higher than the actual value for the population.

- (a) A political pollster seeks information about the proportion of American adults who oppose gun controls. He asks an SRS of 1000 American adults: "Do you agree or disagree with the following statement: Americans should preserve their constitutional right to keep and bear arms." A total of 910, or 91%, said, "Agree" (that is, 910 out of the 1000 oppose gun controls).

Wording of question bias - It is possible that the use of constitutional right can lead people to change their answer for the question, so the 91% would be an overestimate of support.

- (b) A flour company in Minneapolis wants to know what percent of local households bake at least twice a week. A company representative calls 500 randomly-selected households during the daytime and finds that 50% of those who responded bake at least twice a week.

Non-response bias - Not everyone who bakes is home during the week, so when they call those houses, they won't get an answer. Also, those houses who they do get will have more time to bake if they are @ home during the day, so the estimate of 50% would be an overestimate.