

Dear \_\_\_\_\_,

Welcome to AP Statistics! In order to prepare for the upcoming school year, you are required to check out an AP Statistics book before you leave for summer break.

You are also required to complete the following assignment for the **FIRST** day of school. Bring it with you on the first day of class.

We will be starting the school year off with Chapter 5.

- Read and outline pages 327 – 371.
  - Make sure you include **all** information that has a box around it in the text. (Most of it is terminology.)
  - There are a lot of real world examples in the text. Make sure when you are outlining that you jot down a note **AND** try the example problems.
- There are exercises throughout the assignment. Complete the following problems as you come to them while reading:
  - Page 333: # 1 – 7
  - Page 341: # 9 – 11, 13 (Hint: You really have to read the section leading up to this to know how to use Table B!)
  - Page 347: # 15 – 20
  - Page 349: #21-28
  - Page 357: # 33 – 36
  - Page 365: #39 – 43
  - Page 371: #45-47
  - Page 374: #51-55
  - Chapter 5 Quizzes (Answers Attached)
  - Chapter 5 Review (Answers Attached)

\*All of these problems are not “heavy” on math, but there is content and terminology involved that I want you to be familiar with before we start the school year.

- There will be a **Chapter 5 test the first week of school**, so it is crucial that you complete these problems so that you are ready to review when you return and then take the test.

If you are stumped, I recommend using any of the following resources for any assistance you need:

- Khan Academy (<http://www.khanacademy.org/>)  
You will need to make an account, but it is completely free and will give you access to all of the videos.
- YouTube (<http://www.youtube.com/>)  
You would not believe all of the math help that is available on here. You may even find a specific “channel” to follow that will help you next school year!

You will need a graphing calculator every day in class in order to be successful. If you do not already own your own graphing calculator, we suggest TI-83+, TI-84+, or TI-Nspire models. You will be able to use any of those graphing calculators through college as well. Calculators are available at any office supply store such as Staples or Office Max. They are also available new or used through Amazon.com and Ebay.com.

If you have any questions before the school year is over, you can stop by my classroom 208). Enjoy your summer and I will see you in September!

Ms. Donelko [Donels01@wsdmi.org](mailto:Donels01@wsdmi.org)

## Quiz Chapter 5 Part 1

---

### OTHER

1. For each of the following situations involving sampling, identify—as precisely as possible—the population that the sample represents.
  - (a) A business school researcher wants to know what factors affect the survival and success of small businesses. She selects a sample of 150 eating-and-drinking establishments from those listed in the telephone directory for a large city.
  
  - (b) A member of Congress wants to know whether his constituents support proposed legislation on health care. His staff reports that 228 letters have been received on the subject, of which 193 oppose the legislation.
  
2. A local radio talk-show host asks viewers to call in and vote for or against a proposed plan to raise the prices charged by municipal parking meters in a downtown shopping district. 75% of the respondents are opposed to the increase. Describe one possible source of error or bias that might arise in this poll and indicate the direction in which the estimate might be biased. What is the name for this kind of bias?
  
3. Two different organizations conduct polls in a city whose mayor has been accused of taking bribes. One poll asks a SRS of city residents, “Do you think the mayor should resign because of accusations of his criminal activity?” The other asks, “Do you think the mayor should resign?” The first poll concluded that the majority of city residents think the mayor should resign. The second poll drew exactly the opposite conclusion. Explain why their results might be so different.

4. Your school will send a delegation of 35 seniors to a student life convention. 200 girls and 150 boys are eligible to be chosen. If a sample of 20 girls and separate sample 15 boys are each selected randomly, it gives each senior the same chance to be chosen to attend the convention.

(a) Is it an SRS? Explain.

(b) Explain clearly how you would use your calculator to choose a sample of 20 girls for this study.

(c) Beginning at line 108 in the random digits table, reproduced below, select the first three senior girls to be in the sample. Explain your procedures clearly.

108	60940	72024	17868	24943	61790	90656	87964
		18883					
109	36009	19365	15412	39638	85453	46816	83485
		41979					
110	38448	48789	18338	24697	39364	42006	76688
		08708					

## Quiz Chapter 5 Part 2

---

### OTHER

1. A family restaurant chain wants to test the market for a new menu item: a grilled chicken sandwich with chipotle salsa. They are interested in both how to market the item and the right price to charge for it. They decide to offer the sandwich at 60 different restaurants in the chain, using two different descriptions on the menu. Half the restaurants' menus will emphasize "healthy eating" and half will emphasize "value." These two groups of restaurants will be further divided in three groups, each charging either a High, Medium, or Low price for the sandwich. After a month, they will measure what proportion of customers order the new sandwich.

(a) Suppose the company plans to conduct a completely randomized design. List the experimental units, factors, and treatments in this experimental design.

(b) Suppose that 30 of the restaurants in the study are free-standing buildings and the other 30 are located inside malls. The company suspects that the different building types may have impact on how people respond to the advertising campaign and the price. How might they alter the design of this experiment to take this into account?

2. A college fitness center offers an exercise program for staff members who choose to participate. The program assesses each participant's fitness using a treadmill test, and also administers a personality questionnaire. There is a moderately strong positive correlation between fitness score and score for self-confidence. Explain why it would not be possible to conclude from this study that the exercise program increases one's self-confidence.
3. A medical study of heart surgery investigates the effect of a drug called a beta-blocker on the pulse rate of the patient during surgery. The pulse rate will be measured at a specific point during the operation. The investigators will use 20 patients facing heart surgery as subjects. Describe the design of a completely randomized, controlled experiment to test the effect of beta-blockers on pulse rate during surgery. Your answer should address all four basic principles of experimental design.

## Quiz Chapter 5 Part 3

---

### OTHER

1. For each study described below, comment on the extent to which inferences can be drawn about a larger population and whether cause and effect can be established.

(a) A football coach thinks lessons in yoga will improve the flexibility of his players and thereby reduce injuries. To test his theory, he randomly divides the players on the team into two groups. One group has 45 minutes of yoga training each day. The players in the other group do the standard stretching routine the team has used in the past. He compares flexibility in the two groups at the end of the experiment.

(b) Does lack of sleep affect your academic performance? A student explores this question by asking everyone in his statistics class to write down on a piece of paper his or her score on a recent test and total number of hours of sleep he or she got on the last three nights before taking the test.

(c) Does "Cold-Cut," a popular over-the-counter cold remedy that claims to reduce the length and severity of colds really work? A consumer advocacy group addresses this question by asking a random sample of 400 adults how many colds they'd had in the last six months, how long each cold lasted, and if they took "Cold-Cut" to treat the cold.

2. Preliminary observational studies have linked consumption of caffeine during pregnancy to a higher incidence of miscarriages. It would be unethical to run a controlled experiment to establish cause and effect in this situation. Describe two ways in which researchers can seek to establish cause and effect that do not involve experiments.

## Quiz Chapter 5 Part 1

1. ANS:

(a) Population is all small businesses. (b). Population is all the congressman's constituents.

2. ANS:

Only those listeners with strong opinions are likely to call in. The poll probably overestimates opposition to the increase. This is bias arising from voluntary response.

3. ANS:

The wording of the questions is different enough to produce different responses: mentioning bribery may cause a more negative reaction than not mentioning it, or some subjects might not even know about the accusations

4. ANS:

(a) No. Not every group of 35 seniors is equally likely to be selected. It's impossible, for example, to have a group that is all girls. This is a stratified random sample. (b) Number the girls from 1 to 200. Use the command `randInt(1,200)` to select girls until 20 different girls have been selected. (c) Answers may vary, depending on how the table is used. If we read across from left to right, the first three 3-digit numbers selected between 001 and 200 will be 179, 090, an 009.

## Quiz Chapter 5 Part 2

1. ANS:

(a) Experimental units: the 60 restaurants. Factors: menu description and price. Treatments: Healthy-High price, Healthy-Medium price, Healthy-Low price, Value-High price, Value-Medium price, Value-Low price. (b) Block for building type: randomly assign the six different treatments to the 30 free-standing buildings, the do the same to the 30 mall restaurants, so that there are exactly five of each building type assigned to each treatment

2. ANS:

Since this was an observational study, we cannot establish cause and effect. It's possible, for instance, that people with more self-confidence are more likely to choose to exercise. That is, people's personality might be a confounding variable.

3. ANS:

Randomly assign the 20 subjects to two groups. (Answers should describe a method of randomization, such as writing the subjects' names on slips of paper and drawing 10 slips from a hat.) The patients in one group will receive the beta-blocker during their operation while the other group will not receive the beta blocker. In all other ways, patients in the two groups should be treated identically, thus controlling for other variables that might affect pulse rate. Measure pulse rate of all patients at the specified point in the operation, and compare the difference in mean pulse rate for the two groups. Using 10 subjects in each group addresses the fourth principle, replication.

Quiz Chapter 5 Part 3

1. (a) Random assignment → cause and effect can be inferred. No random sampling → Cannot generalize beyond the subjects of the study. (b) No random assignment → cause and effect cannot be inferred. No random sampling → Cannot generalize beyond the subjects of the study. (c) No random assignment → cause and effect cannot be inferred. Random sampling → Can generalize to population from which the random sample was selected.
2. Answers will vary. Some good examples: establish a strong association between caffeine consumption and miscarriages in a wide variety of studies; establish a plausible mechanism for the impact of caffeine on miscarriages; show that the association exists in studies that stratify for other variables that might be confounded with caffeine consumption.

## Chapter 5 Review

---

### MULTIPLE CHOICE

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 simple random sample of 25 cars from a parking lot at Lee-Davis High School was taken. Which of the following statements is correct?
  - a. Since this is a simple random sample, it should be representative of all the cars in Mechanicsville.
  - b. If a simple random sample of 15 cars were taken, we would expect the same amount of variability in the proportion of red cars as we would with a sample of 25 cars.
  - c. An alternative method for getting a representative sample would be to select the 25 cars closest to a specified location, such as the entrance to the gymnasium.
  - d. A different team doing the sampling independently would probably obtain a slightly different answer for their sample proportion.
  - e. The results would 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. 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, six fruit packs, 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. Which of the following is the most plausible explanation for this observed association?
  - a. The 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. The 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. The association is due to confounding (Major League teams tend to be in large cities with more people, hence a greater number of divorces).
  - d. The association makes no sense, since many married couples go to the ballpark together.
  - e. The 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 accomplish which one of the following?
  - a. Limit the effects of variables other than the explanatory variable 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 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 following 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 a 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.

#### OTHER

1. 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.

- (b) Identify the explanatory and response variables.

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

(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?

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

2. 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?

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

(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.

3. 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).

(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.

1. ANS: C
2. ANS: B
3. ANS: D
4. ANS: C
5. ANS: B
6. ANS: D
7. ANS: D
8. ANS: C
9. ANS: A
10. ANS: D

Free Response

1. ANS:

(a) This is an experiment because treatments (aspirin and alcohol; alcohol only) are imposed on the subjects. (b) Explanatory variable: Aspirin consumption; Response: blood alcohol content. (c) Matched pairs experiment. Each subject was given both treatments (aspirin before drinking and no aspirin) and thus acted as his own "pair." (d) Significantly means that the difference found in the subject's blood alcohol level between the two treatments was large enough that it was unlikely to have arisen from chance variation. (e) Each subject acted as his or her own control, drinking alcohol alone and alcohol with aspirin. Comparing each subject's blood alcohol under both treatments allowed the researchers to isolate the impact of aspirin from any other variables.

2. ANS:

(a) A randomized block design—blocking by gender—will reduce the impact that differences between the responses of men and women to the treatment might have on variability arising from random assignment. (b) A larger number of subjects—greater replication—decreases the impact of random variation on experimental results, thereby increasing our ability to distinguish the effects of the treatment. (c) First create two blocks comprised of the 600 men and the 500 women. Then, within each block assign the men numbers from 001 to 600 and the women numbers from 001 to 500. Choose 3-digit numbers from the random number table, ignoring repeats and unassigned numbers, until you have selected 300 men. Then begin elsewhere in the table and follow the same procedure to randomly select 250 women. (Alternatively, use the command `randInt(1,600)` to select men until 300 different men have been selected, and `randInt(1,500)` to select 250 women.) These subjects will be treated with ACE, and the remaining subjects will receive a placebo. Compare changes in blood pressure between the ACE group and the control group after accounting for the difference in response for the men and women.

3. ANS:

(a) Wording of question bias: it's possible that using the term "constitutional right" generates a positive response more often than a question that does not mention the constitution, so that 91% is an overestimate of support. (b) Non-response bias: people without jobs outside the home are more likely to be home during the daytime, and probably have more time to bake. This would make 50% an overestimate.

