AP EXAM Review **Name _____KEY_____**
Free Response Ch 11-13 **Per ___**

Conducting a Study / Experiment

1. A manufacturer of boots plans to conduct an experiment to compare a new method of waterproofing to the current method. The appearance of the boots is not changed by either method. The company recruits 100 volunteers in Seattle, where it rains frequently, to wear the boots as they normally would for 6 months. At the end of the 6 months, the boots will be returned to the company to be evaluated for water damage.

a. Describe a design for this experiment that uses the 100 volunteers. Include a few sentences on how it would be implemented.

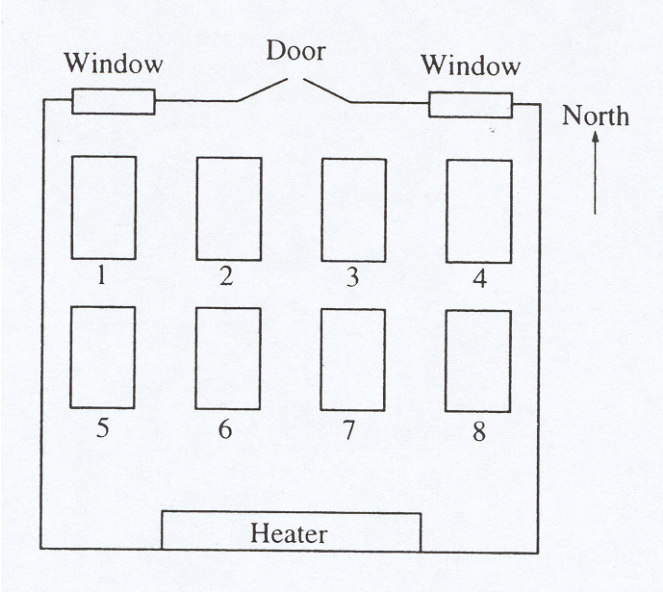
b. Could your design be double blind? Explain.

Because people were recruited for the study, they will be randomly assigned to treatments. Each person will be listed alphabetically on a list and assigned a unique #. I read random numbers off a table. The first 50 numbers get new waterproofing to left boot and last 50 numbers are assigned new waterproofing to right boot. I ignore duplicate random #s.

Because the water proofing doesn't change the appearance, the subjects will wear one of each boot for 6 months and then return them to manufacturer.

b) Yes, this can be double blind because the appearance is Not changed with new waterproofing. Also if the subjects receive the boots annonomously in the mail, and the person evaluating the water damage doesn't know which waterproofing is on the boots, then yes, this can be successfully double blind.

2. A new type of fish food has become available for salmon raised on fish farms. Your task is to design an experiment to compare the weight gain of salmon raised over a six-month period on the new and the old types of food. The salmon you will use for this experiment have already been randomly placed in eight large tanks in a room that has a considerable temperature gradient. Specifically, tanks on the north side of the room tend to be much colder than those on the south side. The arrangement of tanks is shown on the diagram below.



Describe a design for this experiment that takes account of the temperature gradient.

First I would block the fishtanks on homogenous conditions. 4 Blocks are (1,4) (2,3) (5,8) (6,7). Tanks (1,4) are cold and by the window, Tanks (2,3) are cold and infront of the door. Tanks (5,8) are by moderately close to heater, and tanks (6,7) are warmest in front of heater. For each block I will roll a die. If odds comes up, the odd # tank in block gets new food. If evens comes up, the even # tank in block gets new food. Second # in block gets the old food.

Before beginning I weigh all the salmon before randomly distributing them to the tanks. I then feed them new or old food per above. After six months I weigh all the salmon again. Lastly, I conduct a difference of two indep. means hypothesis test to see if new food had better growth.

3. In search of a mosquito repellent that is safer than the ones that are currently on the market, scientists have developed a new compound that is rated as less toxic than the current compound, thus making a repellent that contains this new compound safer for human use. Scientists also believe that a repellent containing the new compound will be more effective than the ones that contain the current compound. To test the effectiveness of the new compound versus that of the current compound, scientists have randomly selected 100 people from a state.

Up to 100 bins, with an equal number of mosquitoes in each bin, are available for use in the study. After a compound is applied to a participant's forearm, the participant will insert his or her forearm into a bin for 1 minute, and the number of mosquito bites on the arm at the end of that time will be determined.

a. Suppose this study is to be conducted using a completely randomized design. Describe a randomization process and identify an inference procedure for the study.

b. Suppose this study is to be conducted using a matched-pairs design. Describe a randomization process and identify an inference procedure for the study.

c. Which of the designs, the one in part **a** or the one in part **b,** is better for testing the effectiveness of the new compound versus that of the current compound? Justify your answer.

a. Begin with 100 volunteers and 100 bins already filled with mosquitoes.
I assign unique 2 digit ID numbers to each volunteer. 00 through 99
I read 2 digit numbers off a random number table. The numbers selected go into group #1, which means ONE arm receives the CURRENT mosquito repellant.
I ignore duplicate numbers.
I continue until fifty volunteers have current repellant.
The remaining fifty volunteers apply NEW repellant to one arm.
Each participant sticks arm into bin for one minute.
The number of bites is counted.
I conduct a DIFFERENCE OF TWO MEANS t-test to compare mean number of bites for each type of repellant.

b. Begin with 100 volunteers and 100 bins already filled with mosquitoes.
I assign unique 2 digit ID numbers to each volunteer. 00 through 99
I read 2 digit numbers off a random number table. The numbers selected go into group #1, which means LEFT arm receives the CURRENT mosquito repellant and RIGHT arm receives NEW repellant.
I ignore duplicate numbers.
I continue until fifty volunteers have current repellant on left arm.
The remaining fifty volunteers apply current repellant to right arm.
Each participant sticks both arms into bin for one minute.
The number of bites on each arm is counted and the DIFFERENCE BETWEEN ARMS is recorded.
I conduct a PAIRED MEANS t-test to compare difference in number of bites for each person.

c. B is better because a MATCHED PAIRS design (left and right arms) is better because blocking on a person REDUCES VARIABILITY. Person to person issues like scent & body order won't be confounding variables.

4. A biologist is interested in studying the effect of growth-enhancing nutrients and different salinity (salt) levels in water on the growth of shrimps. The biologist has ordered a large shipment of young tiger shrimps from a supply house for use in the study. The experiment is to be conducted in a laboratory where 10 tiger shrimps are placed randomly into each of 12 similar tanks in a controlled environment. The biologist is planning to use 3 different growth-enhancing nutrients (A, B, and C) and two different salinity levels (low and high).

a. List the treatments that the biologist plans to use in this experiment.

b. Using the treatments listed in part **a**, describe a completely randomized design that will allow the biologist to compare the shrimps' growth after 3 weeks.

c. Give one statistical advantage to having only tiger shrimps in the experiment. Explain why this is an advantage.

d. Give one statistical disadvantage to having only tiger shrimps in the experiment. Explain why this is a disadvantage.

a. There are 6 unique treatment combinations: A-Low, B-Low, C-Low, A-High, B-High, C-High

Remember: you can either:
Assign treatments to subjects (subjects volunteer first and know they may get new treatment, old treatment, or maybe a control) OR assign random subjects to a treatment (subjects are drafted, not volunteered)

b. It is easier to randomly assign the 6 treatments to 12 tanks than it is to randomly assign 120 shrimp to 12 tanks. Because there are 6 treatments, I will use a regular die to randomize. Starting at tank #1, I roll the die. If I get a 1, then the first treatment is added to the tank. If I roll a 2, then the second treatment is added to the tank, and so on. I continue rolling until each treatment is assigned twice and only twice. If I get a third roll of the die, I skip it and roll again. I weigh and measure the shrimp before treatment, and weigh and measure again after 3 weeks. I record the growth.

c. Advantage of using only tiger shrimp REDUCES VARIABILITY and increases the likelihood that the only growth DIFFERENCES ARE ONLY DUE TO TREATMENTS.

d. Disadvantage of using only tiger shrimp is a LIMITED GENERALIZATION TO OTHER SHRIMP populations. Difference species of shrimp may RESPOND DIFFERENTLY TO DIFFERENT TREAMENTS.

5. An apartment building has nine floors and each floor has 4 apartments. The building owner wants to install new carpeting in 8 apartments to see how well it wears before deciding whether to use that type of carpet in the whole building.

The figure below shows the floors of apartments in the building with their apartment numbers. Only the 9 apartments with **asterisk (*) have children in the apartment**.

11* 1st Floor 12 14 13	21 2nd Floor 22* 24 23*	31 3rd Floor 32 34 33
41 4th Floor 42 44 43	51* 5th Floor 52 54 53	61 6th Floor 62 64 63
71 7th Floor 72 74* 73*	81 8th Floor 82 84* 83	91 9th Floor 92* 94 93*

A) For convenience, the building owner wants to use a cluster sampling method, in which the floors are clusters, to select the 8 apartments. Describe a process for randomly selecting 8 apartments with a cluster method.

Each floor is assigned a unique 1-digit number 1-9. I will ignore 0.

From a random number table, read off the first digit, it is the first floor selected.

Read off the second digit. If the digits are duplicates, skip the second one and continue reading 1-digit at a time until both digits selected are unique.

Install carpet in each of the 8 apartments on the two floors.

B) An alternative sampling method would be to select a stratified random sample of 8 apartments, where strata are apartments with children and apartments with no children. A stratified random sample of size 8 might include two randomly selected apartments with children and six without children. In the context of this situation, give ONE statistical advantage of selecting a stratified sample as opposed to a cluster sample.

The advantage of sampling with strata (with children and child free) is that children playing, running, jumping, spilling food all is more wear and tear on carpet than adults would.

With the cluster sample it is possible to not get any apartments with children, so the building owner may mistakenly think the carpet is long wearing when it is not.

6. An administrator at a large university wants to conduct a survey to estimate the proportion of students who are satisfied with the appearance of the university buildings and grounds. The administrator is considering 3 methods of obtaining a sample of 500 students from the 70,000 at the university.

A) Because of financial constraints, the first method the administrator is considering consists of taking a convenience sample to keep the expenses low. A very large number of students will attend the first football game of the season. The first 500 students who enter the stadium could be used a convenience sample. Why might this sampling method be biased in producing an estimate of the proportion of students who are satisfied with the appearance of buildings and grounds?

Using a convenience sample of the first 500 people in the football stadium would be biased because those students are likely on the football team, the marching band, or cheer leaders. At the very least, the students are die-hard sports fans who believe everything about the university is the "greatest". Those students who think the university is fantastic will see everything (including buildings and grounds) as fantastic as well.

B) Because of the large number of students at the university, the second method the administrator is considering consists of using a computer with a random number generator to select a simple random sample of 500 students from the list of 70,000 names. Describe how to implement this method.

Assign each of the 70,000 unique five digit ID numbers (like a social security number).
Program the computer to generate a long list of 5 digit numbers.
Read the computers numbers and see whose name corresponds with the number.
Contact the student and ask them to take the survey.
If the computer generates a duplicate number, do NOT ask the student to take the survey twice.
Instead, skip the duplicate and continue down the list until 500 unique students were surveyed.

C) Because stratification can often provide a more precise estimate than a simple random sample, the third method the administrator is considering consists of selecting a stratified random sample of 500 students. The university has two campuses with males and females at both campuses. Under what circumstances would stratification by campus, rather than stratification by gender, provide a more precise estimate of the proportion of students who are satisfied with the appearance of buildings and grounds?

The two campuses may be located in different geographic areas, such as one in a downtown big city urban campus. While the sister campus is 100 miles away in a small university town. The urban campus likely does NOT have a large green leafy quadrangle or commons area. The urban campus probably has buildings close together, grime and garbage from the city, and parking garages as well. The small town campus probably has more trees and open spaces, benches. Stratifying by campuses takes into account dramatically different locations whose characteristics cannot be easily changed. Stratifying by campus also takes into account different expectations of cleanliness/green space in urban versus suburban.