There is no recovery from poorly collected data.

So the first priority in a study is properly collecting and organizing the data to avoid the common pitfalls. On the advanced placement exam, using the standard vocabulary is paramount to earning a top score. Fully, yet concisely, explaining the methods as well as the reasons behind the methods is important. So what’s important and why?

**Randomization** – to reduce bias – def. the use of chance or probability during the selection process

Types of bias

1. voluntary response bias – when only those that choose to participate do participate. Those that choose to participate usually feel very strongly one way or the other.

2. response bias – when participants are put in position that makes them uncomfortable to respond truthfully. If a teacher asks for a show of hands of those who have ever cheated on a test many would not raise their hands even if they have cheated. Poorly worded questions would also lead to response bias. For instance, the question “Do you prefer essay questions or tricky worded multiple choice questions” would lead many to respond in favor of essay questions.

3. undercoverage bias – when certain groups are left out of a survey often due to the difficulty in including them. For instance, high school drop outs are rarely surveyed for issues on teenage opinions since most surveys are done at schools.

4. selection bias – when one group is more heavily studied than any other group. If only members of the Sierra Club are surveyed on their opinions of saving the rain forest, the results will be strongly skewed in an environmental direction.

To avoid bias we must randomly select subjects or experimental units from the population being studied. There are 4 basic systems of random selection we have studied.

**simple random samples** – the best method overall – number ALL possible subjects in the population. Then use a random number generator or table of random digits to select a specified number from the population. All possible combinations are possible. The chance of getting a biased group is small and taken into consideration with a statistic called the sampling error or standard deviation of the sample. Ask students to tell you what “the idiot factor” is.

**stratified random samples** – when we first group the subjects by some similar characteristic then take a random sample from each group. For instance, first group the subjects by gender and then randomly select 20 males and 20 females. This is done for comparison purposes.

**systematic random sample** – often done for convenience. Theoretically, line up the subjects and choose every, say, 10th one. Since you are alphabetically listed in my grade book, I could simply go down the list and choose every 5th student for a study.

**cluster sampling** – first splitting the population into similar groups, then completing a census of the groups selected. For instance, second block Westwood students are separated into clusters (classes). Randomly select 5 classes and survey everyone in each of the 5 classes.
**Blocking** – to reduce variation – def. creating groups that are similar with respect to a particular variable

**Blocking** is when groups that are already similar in some way are grouped together. This technique helps control certain lurking or confounding variables and limits the variation in the study statistics. (Note: blocking in an experiment is pretty much the same as stratifying when you choose a sample. It means you group subjects by something like gender, age, grade level, political party affiliation, since these differences often give different results due to the nature of the group.)

**Control group** – to reduce the effects of confounding variables
def. a group that receives no treatment or a placebo treatment

**Blindness** – to reduce bias – def. when the subject, the evaluator, or both (double blind) do not know which treatment is being administered. This is done so neither the subject nor the researcher can bias the study for or against the new drug. (Bias is often not intentional. We humans cannot help it)

Example 1: You are testing new pain reliever for use after wisdom teeth are pulled. How could blinding be used?

**Describing confounding variables:** When there is uncertainty with regard to which variable is causing an effect, we say the variables are confounded. IMPORTANT: In order to receive credit for a confounding variable, you must describe how it confounds the data AND relate the results to BOTH groups.

**Generalizability:** Results may only be generalized to the population randomly selected. If we study only Westwood students we may draw conclusions about only Westwood students, not all high school students.

**Experiments versus Observational Studies:** Experiments impose a treatment on the subject or experimental unit. Only a well designed, controlled experiment can show a causal relationship. One must randomly separate a control group from the experimental group for comparison. The control group may receive no treatment, a placebo, or an alternate treatment.

Example 2: Dr. Bicep is studying muscle growth. He randomly selects 30 patients to add instant protein to their daily diet and 30 patients to eat as they normally would. Both groups are required to hit the weight training room three times a week. The hypothesis is that the instant protein group will increase their muscle mass more than the group without the extra protein.

a) What is the treatment imposed in this experiment? *the instant protein vs. eating normally*

b) Describe a possible confounding variable. *starting physical ability, etc.*

c) Describe a possible observational study for the same problem. *see answers*
Design a study

a) Marine iguanas do not really pay attention to humans. Historically they have had no reason to fear them. Now with the influx of tourists, the iguanas are becoming more timid. Conservationists are interested in the distance at which an iguana begins to show alarm with and without exposure to tourists. Alarm is shown by a rapid head movement accompanied by a low clicking sound. Design an experiment to determine the distance at which iguanas become alarmed by human contact.

b) On the Galapagos islands, both marine iguanas and land iguanas are present. How could your design above be improved to include this knowledge? Why is this change necessary?

2004 #2

Researchers who are studying a new shampoo formula plan to compare the condition of hair for people who use the new formula with the condition of hair for people who use the current formula. Twelve volunteers are available to participate in this study. Information on these volunteers (numbered 1 through 12) is shown in the table below.

<table>
<thead>
<tr>
<th>Volunteer</th>
<th>Gender</th>
<th>Age</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Male</td>
<td>21</td>
</tr>
<tr>
<td>2</td>
<td>Female</td>
<td>20</td>
</tr>
<tr>
<td>3</td>
<td>Male</td>
<td>47</td>
</tr>
<tr>
<td>4</td>
<td>Female</td>
<td>60</td>
</tr>
<tr>
<td>5</td>
<td>Female</td>
<td>62</td>
</tr>
<tr>
<td>6</td>
<td>Male</td>
<td>61</td>
</tr>
<tr>
<td>7</td>
<td>Male</td>
<td>58</td>
</tr>
<tr>
<td>8</td>
<td>Female</td>
<td>44</td>
</tr>
<tr>
<td>9</td>
<td>Male</td>
<td>44</td>
</tr>
<tr>
<td>10</td>
<td>Female</td>
<td>24</td>
</tr>
<tr>
<td>11</td>
<td>Male</td>
<td>23</td>
</tr>
<tr>
<td>12</td>
<td>Female</td>
<td>46</td>
</tr>
</tbody>
</table>

a) These researchers want to conduct an experiment involving the two formulas (new and current) of shampoo. They believe that the condition of hair changes with age but not gender. Because researchers want the size of the blocks in an experiment to be equal to the number of treatments, they will use blocks of size 2 in their experiment. Identify the volunteers (by number) that would be included in each of the six blocks and give the criteria you used to form the blocks.

b) Other researchers believe that hair condition differs with both age and gender. These researchers will also use blocks of size 2 in their experiment. Identify the volunteers (by number) that would be included in each of the six blocks and give the criteria you used to form the blocks.

Female 1 - 2, 16 Ages 20, 24 Female 3 - 4, 5 Ages 60, 62 Male 2 - 3, 9 Ages 43, 47 Female 2 - 8, 12 Ages 48, 46 Male 1 - 1, 11 Ages 21, 23 Male 3 - 6, 7 Ages 61, 58

c) The researchers in part (b) decide to select three of the six blocks to receive the new formula and to give the other three blocks the current formula. Is this an appropriate way to assign treatments? If so, describe a method for selecting the three blocks to receive the new formula. If not, describe an appropriate method for assigning treatments.

No - defeat purpose of blocking

For each block (male/female) two random numbers are generated and assigned to the two volunteers. The volunteer with the smallest number is given the new formula and the other volunteer is given the current formula.
Researchers often mark wildlife in order to identify particular individuals across time or space. A study of butterfly migration is designed to determine which location on the butterflies’ wings is best for marking. The six possible locations are those shown as A through F in the figure below. The butterfly in the figure is a monarch (Danaus plexippus).

Because marks in certain locations may be more likely to attract predators or cause problems than marks in other locations, the goal is to determine whether the six marking locations result in equivalent chances of successful migration. To test this, researchers plan to mark 3,600 butterflies and release them, then count how many arrive displaying each marking location at the end of the migratory path.

a) Briefly describe a method you could use to assign the marking locations if you wanted to ensure that exactly 600 butterflies were marked in each location.

*See answers*

1. Way: Label each butterfly 1-3600, use random number table to select 600 for position A, 600 for position B, and so on...
2. Better: 3600 slips of paper in a bag, 600 for each letter. As each butterfly is captured, pull a slip of paper without replacement.

b) Briefly describe a method you could use to assign the marking locations if you wanted the location to be independent from one butterfly to the next, and wanted each location assigned with probability 1/6 each time.

Use a die and label 1=A, 2=B, 3=C, 4=D, 5=E, 6=F

As the butterfly is caught, roll the die and mark that location.

c) Using your method of assignment from part (b), explain how you would analyze the data collected from this study.

chi-squared test

<table>
<thead>
<tr>
<th></th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td>success</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>not succ</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

d) If butterflies are marked using your method of assignment from part (a), would you change your method of analysis? Explain your reasoning.

nope ⇒ Chi-squared test still applies
Sample Questions

1. In one study subjects were randomly given either 500 or 1000 milligrams of vitamin C daily, and the number of colds they came down with during a winter season was noted. In a second study people responded to a questionnaire asking about the average number of hours they sleep per night and the number of colds they came down with during a winter season.
   A) The first study was an experiment without a control group, while the second was an observational study.
   B) The first study was an observational study, while the second was a controlled experiment.
   C) Both studies were controlled experiments.
   D) Both studies were observational studies.
   E) None of the above is a correct statement.

2. Ann Landers, who wrote a daily advice column appearing in newspapers across the country, once asked her readers, “If you had it to do over again, would you have children?” Of the more than 10,000 readers who responded, 70% said no. (I’m certain your parents would say yes!) What does this show?
   A) The survey is meaningless because of voluntary response bias.
   B) No meaningful conclusion is possible without knowing something more about the characteristics of her readers.
   C) The survey would have been more meaningful if she had picked a random sample of the 10,000 readers who responded.
   D) The survey would have been meaningful if she had used a control group.
   E) This was a legitimate sample drawn from her readers and of sufficient size to allow the conclusion that most of her readers who are parents would have second thoughts about having children.

3. To survey the opinions of bleacher fans at Wrigley Field, a surveyor plans to select every one-hundredth fan entering the bleachers one afternoon. Will this result in a simple random sample of Cub fans who sit in the bleachers?
   A) Yes, because each bleacher fan has the same chance of being selected.
   B) Yes, but only if there is a single entrance to the bleachers.
   C) Yes, because the 99 out of 100 bleacher fans who are not selected will form a control group.
   D) Yes, because this is an example of systematic sampling, which is a special case of simple random sampling.
   E) No, because not every sample of the intended size has an equal chance of being selected.

4. A researcher planning a survey of heads of households in a particular state has census lists for each of the 23 counties in that state. The procedure will be to obtain a random sample of 10 heads of households from each of the 23 counties. Which of the following is a true statement about the resulting sample?
   I. This is not a proper study because children were not included.
   II. This stratified random sample is a type of simple random sample because subjects were randomly selected from each county.
   III. This is not a simple random sample because all possible groups of 230 subjects did not have the same probability of being selected.
   IV. This study may give important information about the similarities and differences of the 23 counties.
   A) II and IV  B) I and II  C) I and III  D) I, II, and III  E) None of these gives a complete set
5. A study is made to determine whether studying Latin helps students achieve higher scores on the verbal section of the SAT exam. In comparing records of 200 students, half of whom have taken at least 1 year of Latin, it is noted that the average SAT verbal score is higher for those 100 students who have taken Latin than for those who have not. Based on this study, guidance counselors begin to recommend Latin for students who want to do well on the SAT exam. Which of the following are true statements?

I. While this study indicates relation, it does not prove causation.
II. There could well be a confounding variable responsible for the seeming relationship.
III. Self-selection here makes drawing the counselors’ conclusion difficult.

A) I and II  B) I and III  C) II and III  D) I, II, and III  E) None of these gives a true complete set

6. A nutritionist believes that having each player take a vitamin pill before a game enhances the performance of the football team. During the course of one season, each player takes a vitamin pill before each game, and the team achieves a winning season for the first time in several years. Is this an experiment or an observational study?

A) An experiment, but with no reasonable conclusion possible about cause and effect.
B) An experiment, thus making cause and effect a reasonable conclusion.
C) An observational study, because there was no use of a control group.
D) An observational study, but a poorly designed one because randomization was not used.
E) An observational study, thus allowing a reasonable conclusion of association but not of cause and effect.

7. Researchers were interested to know whether internal vehicle temperatures vary by outside temperatures. To evaluate this, temperature rise was measured continuously over a 60-minute period in a dark sedan on 16 different clear, sunny days with outside temperatures ranging from 72°F to 96°F. The researchers’ method of analysis is best described as

A) a census
B) a survey
C) an observational study
D) a randomized comparative experiment
E) a single-blind randomized comparative experiment

(Pearson’s test prep, pg 67, #3)

8. Respondents to a randomly distributed questionnaire answered the question, “Do you agree that nuclear weapons should never be used because they are immoral?” The study that uses the results of this questionnaire will most likely suffer from which type(s) of bias?

A) understatement
B) voluntary response
C) response
D) nonresponse
E) all of the above

(Pearson’s test prep, pg 68, #4)

9. A newlywed couple is trying to choose one of two neighborhood supermarkets for their grocery shopping. They decide to randomly select 20 items, check their price at each store, then conduct a test to determine if one store is significantly less expensive than the other. What test should they conduct?

A) Two-sample z-test
B) Two-sample t-test
C) Matched-pairs t-test
D) $\chi^2$ goodness of fit test
E) Linear regression t-test

(Pearson’s test prep, pg 168, #6)
10. In a certain community, 20% of cable subscribers also subscribe to the company’s broadband service for their Internet connection. You would like to design a simulation to estimate the probability that one of six randomly selected subscribers has the broadband service. Using digits 0 through 9, which of the following assignments would be appropriate to model this situation? (Pearson’s test prep, pg 169, #7)

A) Assign even digits to broadband subscribers and odd digits to cable-only subscribers.  
B) Assign 0 and 1 to broadband subscribers and 2,3,4,5,6,7,8, and 9 to cable-only subscribers.  
C) Assign 0,1, and 2 to broadband subscribers and 3,4,5,6,7,8, and 9 to cable-only subscribers.  
D) Assign 1,2,3,4,5, and 6 to broadband subscribers and 7,8,9, and 0 to cable-only subscribers.  
E) Assign 0,1, and 2 to broadband subscribers; 3,4,5, and 6 to cable-only subscribers; and ignore digits 7,8, and 9.

11. The number of T-shirts a school store sells monthly has the following probability distribution:

<table>
<thead>
<tr>
<th>#of T-shirts, X</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>P(X)</td>
<td>0.02</td>
<td>0.15</td>
<td>0.18</td>
<td>0.21</td>
<td>0.14</td>
<td>0.08</td>
<td>0.08</td>
<td>0.04</td>
<td>0.03</td>
<td>0.02</td>
<td>0.05</td>
</tr>
</tbody>
</table>

If each T-shirt sells for $10 but costs the store $4 to purchase, what is the expected monthly T-shirt profit? (Pearson’s test prep, pg 169, #8)

\[ E(X) = \sum xP(x) = 3.78 \]

A) $3.78  
B) $15.12  
C) $22.68  
D) $30.00  
E) $37.80

12. A young woman works two jobs and receives tips for both jobs. As a hair-dresser, her distribution of weekly tips has mean $65 and standard deviation $5.75. As a waitress, her distribution of weekly tips has mean $154 and standard deviation $8.02. What are the mean and standard deviation of her combined weekly tips? (Assume independence for the two jobs.) (Pearson’s test prep, pg 170, #10)

\[ \text{Mean} = \frac{65 + 154}{2} = 109.5  
\[ \text{Standard Deviation} = \sqrt{5.75^2 + 8.02^2} \]

A) mean $167.16; standard deviation $9.87  
B) mean $167.16; standard deviation $13.77  
C) mean $219.00; standard deviation $2.27  
D) mean $219.00; standard deviation $9.87  
E) mean $219.00; standard deviation $13.77

13. A cause-and-effect relationship between two variables can best be determined from which of the following? (Pearson’s test prep, pg 170, #11)

A) A survey conducted using a simple random sample of individuals.  
B) a survey conducted using a stratified random sample of individuals.  
C) When the two variables have a correlation coefficient near 1 or −1.  
D) An observational study where the observational units are chosen randomly.  
E) A controlled experiment where the observational units are assigned randomly to treatments.

14. In a game of chance, three fair coins are tossed simultaneously. If all three coins show heads, then the player wins $15. If all three coins show tails, then the player wins $10. If it costs $5 to play the game, what is the player’s expected net gain or loss at the end of two games? (Pearson’s test prep, pg 175, #20)

\[ X\begin{array}{ccc} 10 & 5 & -5 \hline \frac{1}{8} & \frac{1}{8} & \frac{3}{4} \end{array} \]

\[ \mu_x = -1.88 \]

\[ \sigma_x = 5.555 \]

A) gain $15  
B) gain $1.88  
C) gain $3.75  
D) lose $1.88  
E) lose $3.75
15. Which of the following is a true statement about experimental design? (Pearson’s test prep, pg 210, #6)

A) Replication is a key component in experimental design. Thus, an experiment needs to be conducted on repeated samples before generalizing results. **not true** (large sample is good enough)

B) Control is a key component of experimental design. Thus, a control group that receives a placebo is a requirement for experimentation. **not true** (comparison can be control)

C) Randomization is a key component in experimental design. Randomization is used to reduce bias.

D) Blocking eliminates the effects of all lurking variables. **not true**

E) The placebo effect is a concern for all experiments. **not true**

16. An experimenter believes that two new exercise programs are more effective than any current exercise routines and wishes to compare the effectiveness of these two new exercise programs on physical fitness. The experimenter is trying to determine whether or not a control group, which follows neither of these new programs but continues with current exercise routines, would be beneficial. Which of the following can be said about the addition of a control group? (Pearson’s test prep, pg 210, #7)

A) A control group would eliminate the placebo effect.

B) A control group would eliminate the need for blinding in the study.

C) A control group would allow the experimenter to determine which of the two exercise programs improves physical fitness the most.

D) A control group would allow the experimenter to determine if either of the exercise programs is more effective than current programs for physical fitness.

E) There would be no added benefit to having a control group.

17. A drug company wishes to test a new drug. A researcher assembles a group of volunteers and randomly assigns them to one of two groups—one to take the drug and one to take a placebo. In addition, the company wants the experiment to be double-blind. What is the meaning of double-blind in this situation? (Pearson’s test prep, pg 215, #18)

A) The volunteers in both groups are blindfolded when they take the drug or placebo.

B) The volunteers in both groups do not know whether they are taking the drug or the placebo.

C) Neither the volunteers nor the drug company executives know which volunteers are taking the drug and which are taking the placebo. **X**

D) Neither the volunteers nor the evaluator know which volunteers are taking the drug and which are taking the placebo.

E) as long as the subjects are randomly assigned to the two groups, there is no need to make the experiment double-blind.
18. Since many individuals walk around their homes in their socks, a manufacturer has created a material for socks that is believed to be more resistant to wear than cotton. The manufacturer wishes to test this belief over a period of a month. Given a group of volunteers, which of the following designs will best test this new material's resistance to wear?

(Pearson's test prep, pg 220, #32)

\[ \times \] A) have the volunteers wear the socks made from the new material for a month, and check the wear on the socks at the end of the month.

\[ \times \] B) Allow half of the volunteers to wear cotton socks, while the other half wear socks made of the new material. Compare the wear on the socks at the end of the month.

\[ \times \] C) Randomly assign half of the volunteers to wear cotton socks, while the other half wear socks made of the new material. Compare the wear on the socks at the end of the month.

\[ \times \] D) Randomly assign half of the volunteers to wear cotton socks, while the other half wear socks made of the new material. At the end of two weeks, the volunteers should change sock types. Compare the wear on the socks at the end of the month.

\[ \times \] E) For each volunteer, randomly choose which foot wears a cotton sock, while the other foot wears a sock made of the new material. Compare the wear on the socks at the end of the month.

19. A psychologist from Austin, Texas interested in sleep's effect on the ability to learn randomly selects high school students from the area's local high schools. Half the students are randomly selected to sleep between 7 and 8 hours a night while the remaining half sleep as they normally would. At the end of the study, those students who got between 7 and 8 hours of sleep a night scored significantly higher on tests over the curriculum studied at their school. From this the psychologist can conclude

\[ \circ \] A) Austin high school students getting between 7 and 8 hours of sleep per night score higher on tests over their curriculum.

B) Students in the United States who get between 7 and 8 hours of sleep per night get better grades.

C) Austin students who sleep between 7 and 8 hours per night get better grades.

D) High school students getting between 7 and 8 hours of sleep per night score higher on tests over their curriculum.

? E) Since a placebo was not included in the study, no significant conclusions can be made.

20. Which of the following hypothesis tests automatically include blocking as part of the design.

\[ \circ \] A) two sample t-test

B) matched pairs t-test

C) 2 proportion z-test

D) linear regression t-test for slope

E) none of these