M 140 **Test 2** **B** Name_____

SHOW YOUR WORK FOR FULL CREDIT!

Problem	Max. Points	Your Points
1-10	10	
11	4	
12	4	
13	15	
14	23	
15	8	
16	6	
17 Extra Credit	4	
Total	70	

Multiple Choice Questions (1 point each)

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 there is no control group for comparison.
   c. nothing, because the sample size is too small.
   d. that the new treatment is better than aspirin.
   e. that the remedy is not effective for the treatment of headaches.

2. There are ten departments at SMART University. To conduct a survey, a researcher selects ten faculty members randomly from each department. The sample selected in this way is called a:
   a. Simple random sample.
   b. Stratified random sample.
   c. Systematic random sample.
   d. Voluntary response sample.
   e. Multistage sample.

3. Refer to Question 2. If the researcher selects every 10^{th} person from the alphabetical list of faculty members, then the sample selected this way would be:
   a. Simple random sample.
   b. Stratified random sample.
   c. Systematic random sample.
   d. Voluntary response sample.
   e. Multistage sample.

4. Glamour magazine asked readers to fill out a questionnaire printed in the magazine and mail it in. The questionnaire asked about childhood sexual abuse. 42.6 % percent of the 176 respondents said they had been sexually abused as children. This figure is probably biased because,
   a. the 42.6% of 176 does not make sense here.
   b. the sample size was too small.
   c. the sample was voluntary.
   d. None of the above, there is no reason to doubt the accuracy of the results.

5. There are two statistics classes. The first has 350 students and the second has 250 students. In the first class the students are instructed to each toss a coin 20 times and record the value of $\hat{p}$, the proportion of heads. The instructor them makes a histogram of the 350 values of $\hat{p}$ obtained. The second class did the same, except that each student tossed a coin 40 times. The histogram of $\hat{p}$ values for the first class should be
   a. more biased since it is based on a smaller number of tosses.
   b. less biased since it is based on a larger number of students.
   c. more variable since it is based on a smaller number of tosses.
   d. less variable since it is based on a larger number of students.

6. A polling agency took a random sample of 1000 likely voters in Florida (population: about 18 million), and another random sample of 1000 likely voters in New Mexico (population: about 2 million). Other things being equal, the sample from Florida would tend to be _____ as the sample from New Mexico.
   a. 3 times as accurate
   b. 9 times as accurate
   c. as accurate
   d. one third as accurate
   e. one ninth as accurate

7. Right before a presidential election, the Gallup Poll increases its sample size to 4000 instead of the usual 1200. The reason Gallup does this is,
   a. to increase the margin of error.
   b. to reduce the margin of error.
   c. to get better coverage of special interest groups.
   d. to use better trained interviewers.

The **next three questions** refer to the Harvard Nurses Health Study, which began in 1976. The study included more than 120,000 women who in successive years were asked to answer questions about overall health, smoking, diet, use of birth control pills, and estrogen supplements. Epidemiologists reviewed the medical records of these women every year. Today after more than a decade the research team found that women on an extremely high-fat diet (nearly 50% of energy intake as fat) are at no more risk of breast cancer than those who adhere to an extremely lean diet (less than 29% fat). They did find a clear association between high fat diets and colon cancer, however.

8. What type of study is this?
   a. opinion polls
   b. observational study
   c. randomized experiment

9. What two response variables were mentioned in the description?
   a. nurses and medical records
   b. breast cancer and colon cancer
   c. high fat diet and low fat diet
   d. smoking, diet, use of birth control pills, and estrogen supplements
   e. epidemiologists and Harvard University Nurses

10. Amount of fat in a person's diet is,
   a. a lurking variable
   b. a response variable
   c. an explanatory variable
   d. a placebo

11. (4 points) Decide if each of the following is an observational study or a designed experiment.

a. A study compared the heights of adults who regularly smoked as adolescents to the height of adults who never smoked as adolescents.

i. Observational Study
ii. Designed Experiment

b. A biologist measured the increasing amounts of phosphorus in Percy Priest Lake and recoreded a decreasing number of lake trout over a 7-year period.

i. Observational Study
ii. Designed Experiment

c. Thirty tuna were randomly assigned to two tanks of water, 15 tuna in each tank. One tank was polluted with methyl mercury, while the other tank was not polluted. The survival times of the fish in the two tanks were compared.

i. Observational Study
ii. Designed Experiment

d. Fifty children were randomly assigned to listen to heavy metal music (25 children) or a relaxation tape (25 children). The IQs of the students were measured and compared afterward.

i. Observational Study
ii. Designed Experiment

12. (4 points) Identify each of the following as a *parameter* or a *statistic*.

a. the mean number of states visited by all students at CSUN
parameter

b. the proportion of a random sample of CSUN students who prefer to hear "good news" before bad
statistic

c. the mean number of cats in a random sample of 300 American households
statistic

d. the proportion of American adults who told a Gallup pollster that they prefer to abolish rather than retain the penny
statistic

13. (15 points) A recent research project at Lenox Hill Hospital in New York City provided some information about the laser procedure surgeons have been using in heart patients for more than a decade. Using a laser beam, they drill a little hole in the patient's heart to relieve chest pain. A group of 298 heart patients volunteered for this experiment. They were sedated during the procedure but awake. They could hear the doctors discuss the laser process, and each patient thought he or she was receiving the treatment, although the surgeons actually drilled the hole only in half of the patients' hearts.

a. What are the explanatory and response variables?

Explanatory variable: whether or not the surgeons drilled a hole in the patient's heart

Response variable: chest pain

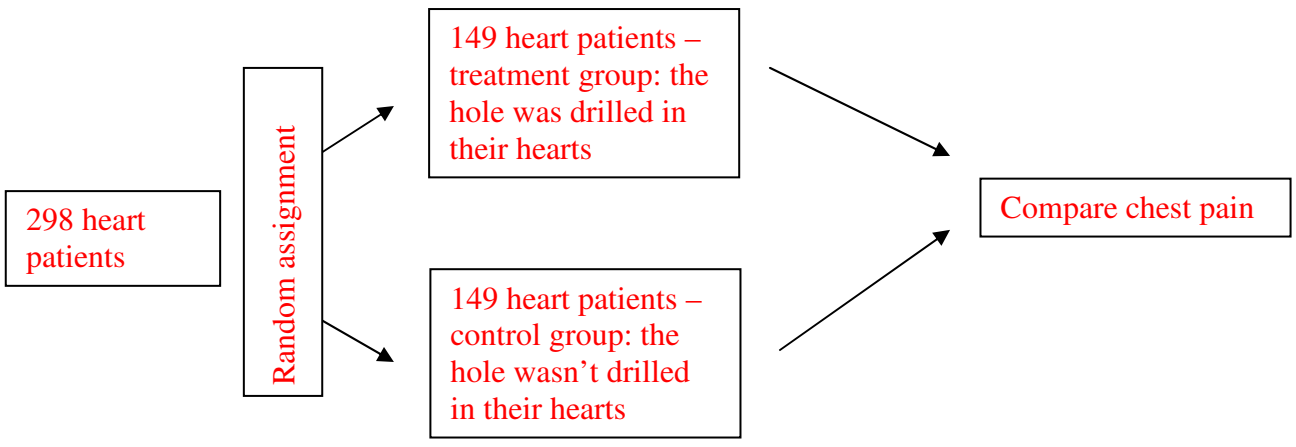
b. Explain what the placebo was in this study, and why they included it.

The placebo here was that they pretended that drill a hole into the patient's heart in the control group.

c. Was this also a double-blind study? Explain your answer.

No, obviously, the surgeons knew when they actually drilled the hole.

d. Draw the outline of the study. Make sure you indicate the response variable, and group sizes.



e. Explain where at what point in the study the experimenters should use ***random assignment***:

When they divided the 298 patients into the two groups. They should have randomly assign each patient into one of the groups.

f. The experimenters suspected that gender could be one of the lurking variables. Explain briefly what they could do to control for this variable.

They could use a block design. First they can separate the patients by gender (blocking), then do the experiment with the females and the males separately.

14. (23 points) Suppose that in order to study/estimate the proportion of all pennies in the U.S. that are older than 10 years, each student in a statistics class was asked to bring in 100 randomly collected pennies and check whether each was older than 10 years or not.

a. Clearly identify **in words** the population, the sample, the parameter of interest, and the statistic in this situation.

Population: all pennies in the U.S.

Sample: the 100 pennies the students randomly picked the brought to class

Parameter: the proportion of all pennies in the U.S. that are older than 10 years

Statistic: the proportion of the 100 pennies in each student's sample that are older than 10 years

b. One of the students, Leslie, counted 39 pennies in her sample that were older than 10 years. Let's assume that in reality 53% of all pennies in the U.S. are older than 10 years.

Fill in the blanks: $\hat{p}$ = _0.39__ p = __0.53__

c. Another student had 57 pennies in his sample that were older than 10 years. This fact that different random samples from a population will give different results is called

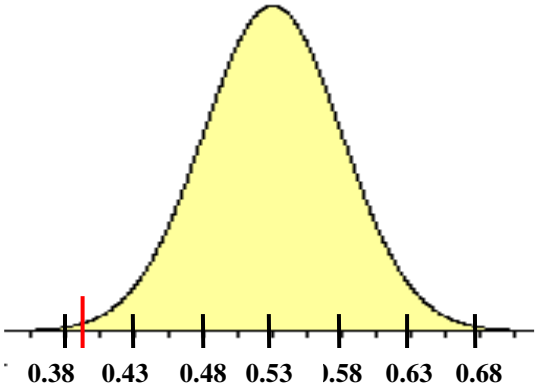
_sampling variability_____.

d. If we could create the sampling distribution of the sample proportions for all possible samples of size 100, then according to the Central Limit Theorem what would be the shape, the mean and the standard deviation of this sampling distribution?
Shape: bell shaped because np = 100(0.53) >10, and n(1-p) = 100(1 – 0.53) >10

Mean: p = 0.53

Standard deviation: $\sqrt{\frac{p(1-p)}{n}} = \sqrt{\frac{0.53(1-0.53)}{100}} = 0.0499 \approx 0.05$

e. Draw the sampling distribution of the sample proportions. Mark the mean, and three standard deviations below and above the mean, and label them with values.



d. Based on the sampling distribution, was Leslie's sample unusual? Explain.

Yes, it's unusual. Her sample proportion, 0.39, is more than two standard deviations below the mean (see read mark).

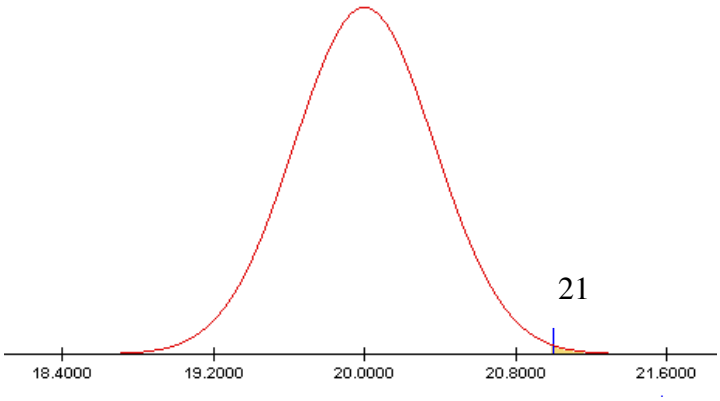
e. If we would create the sampling distribution of the sample proportions for all possible samples of size of 1000 instead of 100, how would the mean and standard deviation of this sampling distribution compare to the sampling distribution of $n = 100$ you described in part (d). Fill in the blanks with one of the following: ***smaller than, greater than, the same as.***

The mean of the sampling distribution for $n = 1000$ would be _the same as_____ the mean of the sampling distribution for $n = 100$.

The standard deviation of the sampling distribution for $n = 1000$ would be _smaller than__ the standard deviation of the sampling distribution for $n = 100$.

15. (8 points) The weights of the contents of a cereal box are normally distributed with a mean weight of 20 ounces and a standard deviation of 0.37 ounce. (Don't forget to draw and shade in both parts!)

a. What percent of cereal boxes weigh more than 21 ounces?

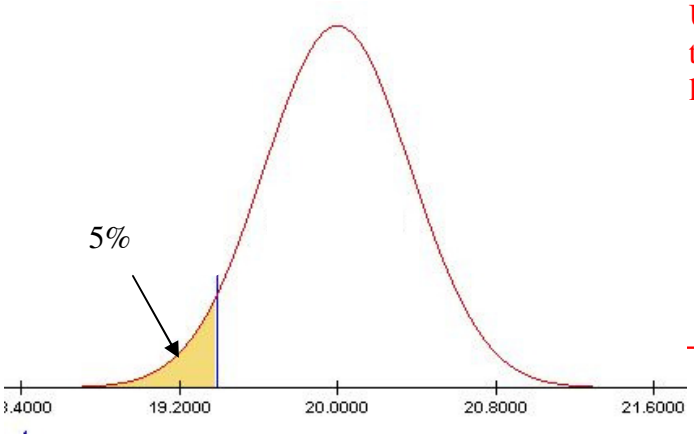


$$z = \frac{21 - 20}{0.37} = 2.70$$

Probability: 0.0035

About 0.35% of cereal boxes weigh more than 21 ounces.

b. Boxes in the lower 5% do not meet the minimum weight requirements and must be repackaged. What is the minimum weight requirement for a cereal box?



Using the table, or the calculator, the z-score corresponding the lower 5% is -1.64 (or -1.65)

$$-1.64 = \frac{x - 20}{0.37} \quad \Rightarrow \quad x = 19.39$$

The minimum weight requirement for a cereal box is 19.39 ounces.

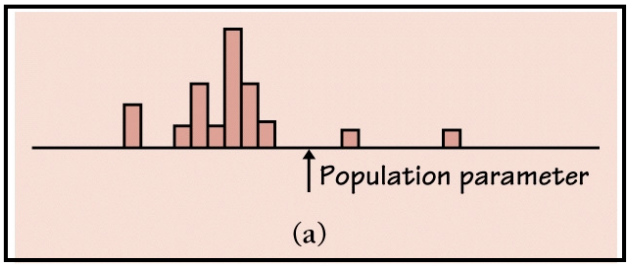
16. (6 points) Adam and his girlfriend, Betty, often teased each other about their heights: he called Betty "Shortie" because she was short , and she called him "Beanstalk" because he was tall. The heights of men are approximately normally distributed with mean 70.1 in and standard deviation of 2.7in. The heights of women are approximately normally distributed too with mean of 64.8in and 2.5 in. If Adam is 75.9 in tall, and Betty's height is 59.7 in, who has "more right" to call the other one as "Shortie" or "Beanstalk" (i.e. whose height is more extreme)?

The standardized value for Adam's height is $z = \frac{75.9 - 70.1}{2.7} = 2.15$.

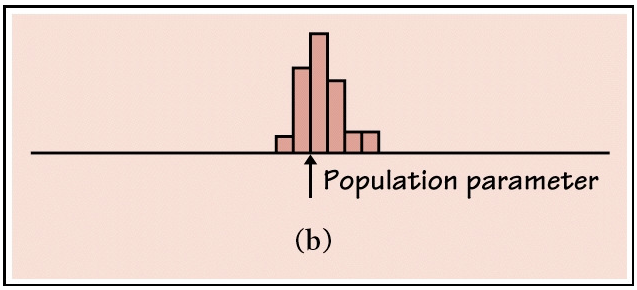
The standardized value for Betty's height is $z = \frac{59.7 - 64.8}{2.5} = -2.04$.

Since Adam's height is farther from the mean than Betty's height, she has more right to call him Beanstalk. Adam's height is more extreme.

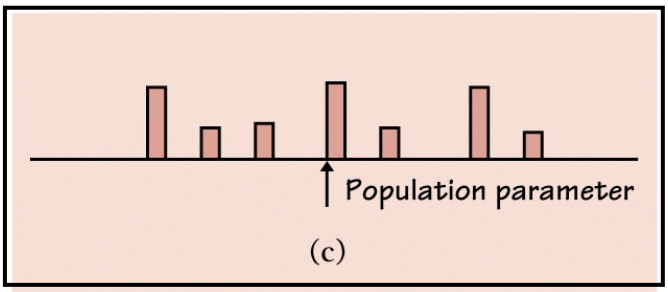
17. EXTRA CREDIT: The histograms below show four sampling distributions of statistics intended to estimate the same parameter. Label each distribution relative to the others as having large or small bias and as having large or small variability. (Circle the correct choices.)



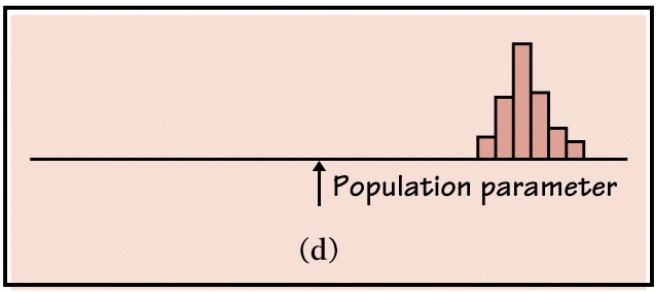
Bias:	~~large~~	small
Variability:	large	small



Bias:	large	small
Variability:	large	small



Bias:	large	small
Variability:	large	small



Bias:	large	small
Variability:	large	small