STATISTICS 8, MIDTERM EXAM 2

NAME:_____________________________________ Seat Number: ____________

Last six digits of Student ID#: ____________________ Circle your Discussion Section: 1 2 3 4

Make sure you have 6 pages, including Table A.1 on the last page.
You may use two pages of notes (both sides) and a calculator.
Multiple choice questions: There are 10 questions worth 4 points each (10 x 4 pts each = 40 pts). Instructions will be given when those begin on page 4.
Free response questions: Show all work. If you need extra space use the back of the page, but make sure to tell us it's there. Total of 60 points; points for each part of each question are shown.

1. (10 pts total) The sensitivity of a medical test for a certain disease is .95 and results are independent from one patient to the next. Three people who have the disease are tested.

   a. (4 pts) For each person tested, what is the probability that the test does *not* detect the disease, i.e. is a false negative?

   $$1 - .95 = .05$$

   b. (6 pts) What is the probability that at least one of the tests for three people with the disease comes out negative?

   $$1 - P(0 \text{ negative tests}) = 1 - (.95)^3 = 1 - .857375 = .142645$$

   (You could just leave it as a formula, and you could round off to 1 – .86 = .14.)

For use by graders: Free response points:

Question 1: ________ out of 10

Question 2: ________ out of 16

Question 3: ________ out of 26

Question 4: _________ out of 8

Free response score:__________

Multiple choice score:__________

Total score: ____________

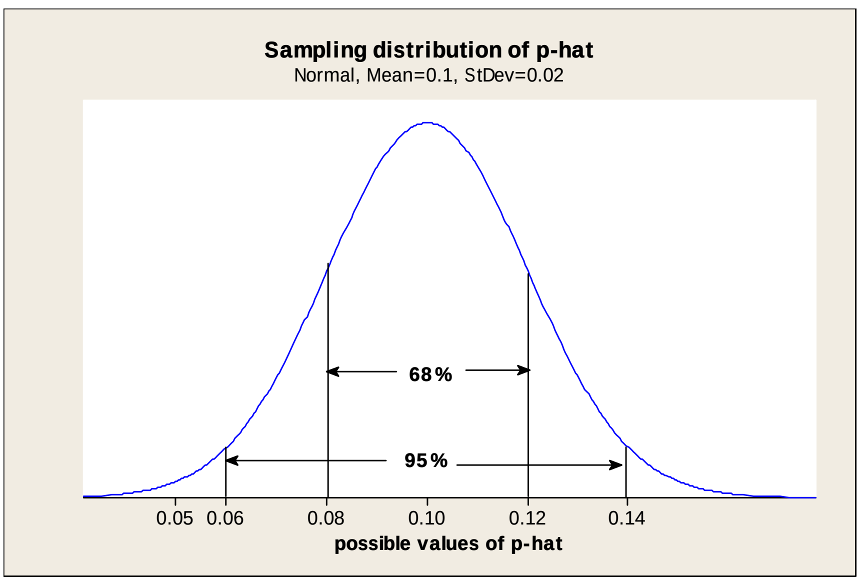
2. (16 pts total) A new drug is being proposed for the treatment of migraine headaches. Unfortunately some users in early tests of the drug have reported mild nausea as a side effect. The FDA will not approve the drug if it thinks that more than 5% (i.e. 0.05) of the population would suffer from this side effect. In an experiment to test this side effect, 225 people who suffer from migraine headaches receive the new drug and report any adverse effects. Suppose that in fact, 10% (i.e. 0.1) of the population would have mild nausea if they were to take the drug. Let $\hat{p}$ be the proportion of the sample of 225 drug users in this study who will suffer mild nausea.

a. (2 pts each) Give the mean and standard deviation of the sampling distribution of $\hat{p}$.

Mean = ______*.10*________ *(This is the population proportion who would have mild nausea.)*

Standard deviation = $\sqrt{\dfrac{(.1)(.9)}{225}} = .02$

b. (4 pts) Draw a picture of the sampling distribution of $\hat{p}$, marking the mean and the intervals that cover the middle 68% and 95% of the possible values.



c. (4 pts) On your figure in Part (b) show where the proportion of 0.05 would fall and compute the z-score for this value.

See graph for placement of 0.05. $z = \dfrac{.05 - .10}{.02} = -2.5$

d. (4 pts) The FDA will approve the drug if fewer than 5% of those in the sample experience mild nausea. What is the probability that the FDA will approve the drug based on this study?

$P(\hat{p} < .05) = P(z < -2.5) = .0062$ *(from Table A.1)*

3. (26 pts total) A medical school knows that students will be more successful in their program if they have studied a foreign language in college. They collect data on all of the students who graduated from their school in the past 15 years. Define the following events:

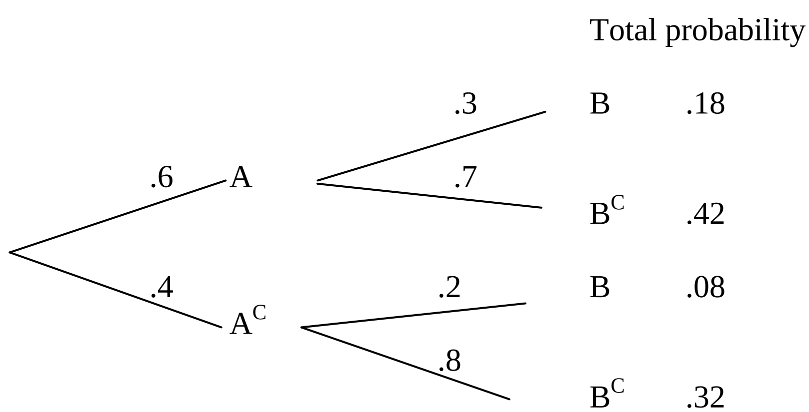
   A = Student studied a foreign language in college
   B = Student graduated from the medical school with honors

   Suppose 60% of the students studied a foreign language in college. Of those students, 30% graduated from medical school with honors. Of the 40% of students who did not study a foreign language in college, only 20% graduated from the medical school with honors.

   a. (2 pts each) Provide values for the following:

      $P(A) =$ ____.60______ $P(B \mid A) =$ ___.30_____ $P(B \mid A^C) =$ ___.20______

   b. (10 pts) Draw a tree diagram for this situation *or* construct a hypothetical hundred thousand table.



	Graduated with honors	Did not graduate w/honors	Total
Took language	18,000	42,000	60,000
No language	8,000	32,000	40,000
Total	26,000	74,000	100,000

   c. (6 pts) Using your diagram or table in Part (b), find the probability that a student who graduated with honors had studied a foreign language in college. In other words, find $P(A \mid B)$.

$$\textit{From the tree diagram: } \frac{.18}{.18+.08} = \frac{.18}{.26} = .6923\textit{; From the table: } \frac{18,000}{26,000} = .6923$$

   d. (4 pts) Based on this information, can the medical school conclude that studying a foreign language in college *causes* students to do better in medical school? Briefly explain.

      No. This was an observational study, so we cannot conclude cause and effect.

4. (8 pts total) The quiz for this class each week has 5 multiple choice questions with 4 choices each, worth 2 points each. Suppose someone is just guessing on every question. Define two random variables:
   X = total number of *questions* correct   Y = total *points* earned.

   a. (4 pts) Does X have a binomial distribution? If so, specify *n* and *p*. If not, explain which condition is not met.

      Binomial, $n = 5$ and $p = 1/4$

   b. (2 pts each) Fill in numerical values:

      E(X) = *$np = 5(1/4) = 1.25$*   E(Y) = *$2 \times E(X) = 2 \times (1.25) = 2.5$*

MULTIPLE CHOICE

- **You have Exam Version <u>B</u>. Write this on your Scantron *on the "SUBJECT" line*.**
- **Fill in and bubble your ID at the top of the Scantron. Put Discussion Section in "HOUR" line.**
- ***Circle* the best answer on this exam paper *and* bubble in the Scantron sheet.**

1. If A and B are independent events (both with probability greater than 0) then which of the following statements must be true?
   A. P(A) + P(B) = 1
   B. P(A and B) = 0
   C. P(A and B) = P(A)P(B)
   D. P(A and B) = P(A) + P(B)

2. Which of the following does *not* have a sampling distribution?
   A. Population proportion p
   B. The point estimate for the population proportion
   C. Sample mean $\bar{x}$
   D. Sample proportion $\hat{p}$

3. A sales person makes "cold calls" trying to sell a product by phone and is successful on each call with probability 1/50. Whether or not he is successful is independent from one call to the next. If he calls 50 people, the number of successful calls is:
   A. exactly 1, since he called 50 people and the probability of success is 1/50 each time.
   B. a binomial random variable.
   C. at most 1, because once he has been successful he can't be successful again in the 50 calls.
   D. equally likely to be 0, 1 or 2.

4. The expected value of a random variable is
   A. the mean value over an infinite number of observations of the variable.
   B. always one of the possible values for the random variable.
   C. the value that has the highest probability of occurring.
   D. always computed as *np*.

5. Suppose that a 95% confidence interval for the proportion of men over 60 who have high blood pressure is .30 to .40. Which of the following is the best interpretation of this information?
   A. We can be fairly confident (about 95%) that the proportion of men in the population who have high blood pressure is between .30 and .40.
   B. We can be fairly confident (about 95%) that the proportion of men in the sample who have high blood pressure is between .30 and .40.
   C. 95% of the men in the sample have blood pressure that is between 30% and 40% too high.
   D. 95% of the men in the population have blood pressure that is between 30% and 40% too high.

6. Suppose that the mean of the sampling distribution for the difference in two sample proportions is 0. This tells us that:
   A. The two sample proportions are equal to each other.
   B. The two sample proportions are both 0.
   C. The two population proportions are equal to each other.
   D. The two population proportions are both 0.

7. A swim club randomly chose one of its members to feature in its newsletter. The story provided the following information: *Allison is 28 years old, and has always loved water sports. She swam competitively when she was in high school, and she enjoys surfing. In college, she majored in exercise science.* Which one of the following statements about Allison is most likely?
   A. Allison teaches at a high school.
   B. Allison teaches at a high school and is the swim coach for the school.
   C. Allison teaches at a high school during the school year and spends the summers surfing.
   D. There is no way to know which of the above statements is most likely.

8. All of the following are ways to show that A and B are mutually exclusive events *except* one of them. Which one does not show that A and B are mutually exclusive? ***My mistake; there were 2 correct answers! You will get credit for either one of them.***
   A. P(A|B) = 0
   B. P(A) = 1 – P(A^C)
   C. P(A and B) = P(A)P(B)
   D. P(A or B) = P(A) + P(B)

9. When a random sample is to be taken from a population and a statistic is to be computed, the statistic can also be thought of as
   A. A point estimate
   B. A random variable
   C. Both of the above
   D. None of the above

10. You buy coffee at a kiosk at random times. The following table gives the probability distribution for X = number of customers in line when you show up (not including you):

k	0	1	2	3	4
***P*(*X* = *k*)**	0.15	0.20	0.40	0.20	0.05

   What is the probability that there will be at least two people in line when you show up?
   A. 0.25
   B. 0.40
   C. 0.65
   D. 0.75