Practice Exam 2

Name: ______________________________

Section: ______________________________

You may use a calculator and two 8.5” by 11” sheets of notes (front and back), which you will turn in with your exam.

Please show all your work, including all calculations, and explain your answers. Whenever needed, please round numbers (including intermediate calculations) to the nearest 0.001.

Cell phones and any other electronic devices (aside from your calculator) are not permitted. No interaction of any sort is allowed with your classmates.

I Conceptual Questions

Please answer the following in no more than 1-2 sentences each.

1. (4 points) Describe stratified sampling and give an example of when it might be used.

2. (4 points) Describe the difference between a population and a sample.

3. (4 points) Describe the difference between disjoint events and independent events.

II Word Problems

1. Police often set up sobriety checkpoints (roadblocks where drivers are asked a few brief questions to allow the officer to judge whether or not the person may have been drinking). If the officer does not suspect a problem, drivers are released to go on their way. Otherwise, drivers are detained for a Breathalyzer test that will determine whether or not they will be arrested. The police say that based on the brief initial stop, trained officers can make the right decision 80% of the time. Suppose the police operate a sobriety checkpoint after 9:00 p.m. on a Saturday night, a time when national traffic safety experts suspect that about 12% of drivers have been drinking.

   Define A to be the event that a driver who is stopped has been drinking.

   Define B to be the event that the police decide a driver they stop has been drinking.

   a) Write a sentence describing what the event A^c means, in a similar style to the definition of the event A above.

   b) Write a sentence describing what the event B^c means, in a similar style to the definition of the event B above.

   c) For this part of the problem, you're just matching up numbers in the problem statement to the probabilities of events. No calculations are necessary.

   Find $P(A)$:

   Find $P(B|A)$:

   Find $P(B^c|A^c)$:

   d) For this part of the problem, you may find it helpful to draw a probability tree diagram.

   Find $P(B^c|A)$:

   Find $P(B|A^c)$:

   e) What's the probability that a driver who is stopped has been drinking, and the police decide (correctly) that he has been drinking?

f) If I know that the police decided a particular driver has been drinking, what is the probability that the driver actually had been drinking? That is, find $P(A|B)$

g) Are the events A and B independent? Justify your answer with a comparison of numeric probabilities that you have already calculated in previous parts of the question.

2. After menopause, some women take supplemental estrogen. There is some concern that if these women also drink alcohol, their estrogen levels will rise too high. Twelve volunteers who were receiving supplemental estrogen were randomly divided into two groups, as were 12 other volunteers not on estrogen. In each case, one group drank an alcoholic beverage, the other a nonalcoholic beverage. An hour later, everyone's estrogen level was checked. Only those on supplemental estrogen who drank alcohol showed a marked increase.

   a) Was this study an experiment or an observational study?

   b) Was blocking used? If so, describe how and why.

   c) Why was it important to randomize the assignment of the study participants to the groups that either drank or did not drink an alcoholic beverage? What problem is randomization trying to prevent?

   d) What is blinding? From the description above, is there any evidence that blinding was used in this study?

   e) Is it possible to make a conclusion about a causal connection between estrogen supplements, alcoholic beverages, and estrogen levels based on this study? Why or why not?

3. In the 1980s, it was generally believed that congenital disorders (that is, a health condition present at or before birth) affected about 5% of the nation's children. Some people believe that the increase in the number of chemicals in the environment has led to an increase in the incidence of congenital disorders. A recent study examined randomly selected sample of 392 children and counted the number who showed signs of a congenital disorder.

   a) In this study, what is the population parameter of interest? (We don't have a number for this; just describe the parameter in a sentence.)

   b) Let X be a random variable equal to the number of children in the sample who show signs of a congenital disorder. What would the sampling distribution of X be, assuming the proportion of children born with congenital disorders has not changed since the 1980's? Check all assumptions.

   c) Let $\widehat{p}$ be the proportion of children in the sample who show signs of a congenital disorder. What would the sampling distribution of $\widehat{p}$ be, assuming the proportion of children born with congenital disorders has not changed since the 1980's? Check all assumptions (if you already checked any in part b, you do not need to check them again).

d) In the sample, 24 of the children showed signs of a congenital disorder. Using your choice of the distribution from part b) or part c) and the R output below, find the probability of getting at least 24 children with congenital disorders in this sample, assuming that the proportion of children born with congenital disorders has not changed since the 1980's. You will only need one of the numbers below to calculate your answer.

```
> pbinom(q = 23, size = 392, prob = 0.05)

[1] 0.8189341

> pbinom(q = 24, size = 392, prob = 0.05)

[1] 0.8704152

> pnorm(q = 0.0609, mean = 0.05, sd = 0.218)

[1] 0.5199388

> pnorm(q = 0.0609, mean = 0.05, sd = 0.044)

[1] 0.5978273

> pnorm(q = 0.0609, mean = 0.05, sd = 0.011)

[1] 0.839135
```

4. Twix Minis have, on average, 2.00 grams of saturated fat per bar, with a standard deviation of 0.17 grams. The amount of saturated fat per bar follows a normal distribution. The day after Halloween I ate 3 Twix Minis.

   a) What is the mean and standard deviation of the **total amount** of saturated fat I consumed?

   b) What distribution would you use to model the **sample mean** of the number of grams of saturated fat per bar? Be as specific as possible.

   c) Using the 68/95/99.7 Rule, find the probability that the average number of grams of saturated fat per bar in my four candy bars was greater than 1.902 grams.