22S:30/105, Statistical Methods and Computing
Spring 2013, Instructor: Cowles
Midterm 2

Name: _____ Course no. (30 or 105) _____

Show your work on any problems that involve calculations.

I will grade on a curve and will give partial credit wherever possible.

1. The number of flaws per square yard in a type of carpet material varies with mean 1.6 flaws per square yard and standard deviation 1.2 flaws per square yard. The population distribution cannot be Normal, because a count takes only whole-number values. An inspector plans to sample 200 square yards of the material, to record the number of flaws found in each square yard, and to calculate $\bar{x}$, the sample mean number of flaws per square yard inspected. Use the central limit theorem to find the approximate probability that inspector's $\bar{x}$ value will exceed 2.

2. A chemical engineer is designing the production process for a new product. The chemical reaction that produces the product may have higher or lower yield, depending on the temperature and the stirring rate in the vessel in which the reaction takes place. The engineer decides to investigate the effcts of combinations of two temperatures (50 degrees C and 60 degrees C) and three stirring rates (60 rpm, 90 rpm, and 120 rpm) on the yield of the process. She will process two batches of the product at each combination of temperature and stirring rate.

   (a) Is the engineer undertaking an experiment or an observational study? Explain briefly.

   (b) What is the response variable?

   (c) What are the factors?

(d) Identify all of the treatments.

(e) How many experimental units will be needed?

3. t-tests and t confidence intervals are used when (circle all that apply):

   (a) the variable of interest is quantitative
   (b) the variable of interest is binary
   (c) the parameter of interest is a population mean
   (d) the parameter of interest is a population standard deviation
   (e) the population mean is assumed known
   (f) the population standard deviation is assumed known
   (g) none of the above
   (h) all of the above

4. I was taught in junior high health classes that the population mean body temperature in healthy adult humans is 98.6 degrees F. I want to use temperature data from a sample of healthy adults to test whether the population mean is actually different from 98.6. Write the appropriate null and alternative hypotheses that I should test, using standard statistical symbols.

5. A die is a small 6-sided cube with a different number of dots (between 1 and 6) on each of its faces. Two Boy Scouts plan to do the following experiment: they will flip a coin once and roll the die once, and will record whether the coin came up heads or tails and the number of dots on the upward face of the die.

   What is the sample space of this experiment?

6. ICON, UI's course management software, displays the mean, median, mode, and standard deviation of the scores for each graded item in a course. Would it be useful for ICON to also display a confidence interval for the mean score? Why or why not?

7. Researchers studied 11 people who had been diagnosed as dependent on caffeine. One measurement that was taken on each person was the beats per minute the subject achieved when asked to press a button 200 times as quickly as possible. Refer to the attached SAS output when answering the following questions.

   (a) Which confidence interval procedure does `proc means` use? (Circle one).

      i. b interval
      ii. t interval
      iii. z interval
      iv. none of the above

   (b) Is there any evidence in the SAS output that the confidence interval procedure that you chose should not be used for these data? Explain briefly.

   (c) What quantity are the researchers 99% confident lies in the interval that SAS produced? (Circle one.)

      i. n
      ii. $\bar{x}$
      iii. μ
      iv. s
      v. σ
      vi. none of the above

   (d) Would a 95% confidence interval calculated from the same data be wider or narrower than the 99% interval given here? Briefly explain.

(e) The number 326.64 given in the output below is a: (Circle one).

i. population
ii. sample
iii. parameter
iv. statistic
v. random variable
vi. none of the above

```
                The UNIVARIATE Procedure
                   Variable:  beatcaf

  Stem Leaf                     #             Boxplot
     4 12                       2                |
     3 8                        1             +-----+
     3 0044                     4             *--+--*
     2 889                      3             +-----+
     2 4                        1                |
       ----+----+----+----+
   Multiply Stem.Leaf by 10**+2

                The MEANS Procedure

            Analysis Variable : beatcaf

                           Lower 99%       Upper 99%
 N            Mean     CL for Mean     CL for Mean
-------------------------------------------------------
11     326.6363636     272.3143665     380.9583608
-------------------------------------------------------
```