22S:30/105, Statistical Methods and Computing
Spring 2015, Instructor: Cowles
Midterm 2

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

Show your work on any problems that involve calculations. If your answer to a multiple choice or true-false question would vary under different conditions, write an explanation. I will grade on a curve and will give partial credit wherever possible.

1. A *sampling distribution* is: (Circle one.)
   (a) the probability of getting a simple random sample that is representative of the population.
   (b) the distribution of values of a variable among all the individuals in the population.
   (c) the distribution of values of a statistic in all possible samples of the same size from the same population.
   (d) the distribution of values of a variable in a sample
   (e) none of the above
2. The distribution of blood cholesterol level in the population of young men aged 20 to 34 years is close to Normal with standard deviation $\sigma = 41$ milligrams per deciliter (mg/dl). You measure the blood cholesterol of 14 cross-country runners. The mean level is $\bar{x} = 172$ mg/dl. Assume that σ in cross-country runners is the same as in the general population.
   (a) Compute a 90% confidence interval for the mean level μ among cross-country runners. (Numeric answer; show your work.)
   (b) How large a sample would be needed to reduce the margin of error to 5 mg/dl? (Numeric answer; show your work.)

3. Hallux abducto valgus (HAV) is the name for a deformity of the joint at the base of the big toe that causes the toe to angle toward the outside of the foot. This dataset contains the angle of deformity (in degrees) in 38 patients with HAV. The larger the angle, the more severe the condition. The 38 patients can be treated as a simple random sample of patients with HAV. Refer to the SAS output below in answering the following questions.

   (a) Podiatrists wish to use these data to draw inference about the mean angle of deformity in the population of people with HAV. From the SAS output, give the numeric values of:

      i. the point estimate of the population mean

      ii. a 95% confidence interval for the population mean

   (b) The correct interpretation of the interval that you specified in the previous question is: (Circle one.)

      i. The probability that $\bar{x}$ lies in the interval is .95.
      ii. The probability that μ lies in the interval is .95.
      iii. 95% of patients with HAV will have an angle of deformity in the interval.
      iv. None of the above

   (c) The type of confidence interval that `proc means` produces is: (Circle one.)

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

```
                The UNIVARIATE Procedure
                    Variable:  angle

 Stem Leaf                     #             Boxplot
    5 0                        1                0
    4
    4
    3 88                       2                |
    3 00012224                 8             +-----+
    2 55556666888             11             *--+--*
    2 000111123                9             +-----+
    1 66788                    5                |
    1 34                       2                |
      ----+----+----+----+
  Multiply Stem.Leaf by 10**+1


            Analysis Variable : angle

                           Lower 95%       Upper 95%
 N            Mean     CL for Mean     CL for Mean
-------------------------------------------------------
38      25.4210526      22.9641615      27.8779438
-------------------------------------------------------
```

4. The common earthworms that are seen on sidewalks after a rainstorm are of the species *Lumbricus terrestris*.

Figure 1: Earthworm

Researchers studying this species believe that the population mean length of worms of this type is less than 200 millimeters (mm). They plan to gather a simple random sample of 100 earthworms and measure each one.

The researchers will use their data to test the following hypotheses:

$$H_0: \quad \mu \geq 200$$
$$H_A: \quad \mu < 200$$

(a) The researchers will use the sample mean $\bar{x}$ to estimate the population mean μ. Here $\bar{x}$ is: (Circle one.)

i. a parameter
ii. a population
iii. a sample
iv. a statistic
v. a test statistic
vi. none of the above

(b) Which range of values of $\bar{x}$ would provide evidence against H_0? (Circle one.)

i. large values
ii. small values
iii. impossible to tell from the information given

(c) The researchers will conduct their hypothesis test at significance level $\alpha = .10$. They believe that the population standard deviation of length in this type of earthworms is 50 mm. They will have 100 worms in their sample. What is the critical value of $\bar{x}$ at which they should reject H_0? (Numeric answer; show your work.)

(d) What is the power of their test against the alternative hypothesis $H_A : \mu = 190$? (Numeric answer; show your work.)