Practice Quiz 11

Statistics 1000
Fall 2008 (take and self-check by Dec. 1)
Dr. Nancy Pfenning

1. (5 pts.) Some shoppers were observed in supermarket bakery departments that provided tongs and others were observed in departments that provided tissues. A researcher recorded how many people used their hands to withdraw baked goods instead of the tongs or tissues provided:

	Hands	No Hands	Total
Tongs	97	11	108
Tissues	83	49	132
Total	180	60	240

   (a) Which *two* of these is a correct formulation of the null hypothesis?
      i. Use of hands, and whether tongs or tissues are provided, are not related.
      ii. Use of hands, and whether tongs or tissues are provided, are related.
      iii. Proportions who use their hands are the same for all shoppers in stores that provide tongs and stores that provide tissues.
      iv. Proportions who use their hands are different for all shoppers in stores that provide tongs and stores that provide tissues.

   (b) Explain how the study's results may be biased if observations were made in the morning for stores with tongs and in the evening for stores with tissues.

   (c) Explain how the study's results may be biased if stores with tongs tended to be located in areas with a large student population.

   (d) If proportions using their hands were actually equal for shoppers in stores providing tongs and tissues, then the proportions would both be ______.

   (e) Complete this table of counts expected under the null hypothesis.

	Hands	No Hands	Total
Tongs			108
Tissues			132
Total	180	60	240

   (f) Calculate the chi-square statistic; its size is
   (i) large (ii) not large (iii) borderline

   (g) The p-value is (i) small (ii) not small (iii) borderline

   (h) Draw your conclusions, first in terms of a relationship, then in terms of population proportions using their hands.

2. (5 pts.) Salary (in millions) was regressed on batting average for a sample of 6 baseball players in 2004.

```
The regression equation is
Salary = - 28.9 + 122 BattingAverage
Predictor        Coef     SE Coef          T        P
Constant      -28.947       8.700      -3.33    0.029
BattingA       121.55       30.32       4.01    0.016
S = 1.574       R-Sq = 80.1%     R-Sq(adj) = 75.1%
```

(a) Explain why it makes sense for the relationship to be positive.

(b) The p-value and the value of R-Sq together tell us that there is

  i. weak evidence of a weak relationship between batting average and salary
  ii. weak evidence of a strong relationship between batting average and salary
  iii. strong evidence of a weak relationship between batting average and salary
  iv. strong evidence of a strong relationship between batting average and salary

(c) We seek evidence regarding the slope of the regression line for the
(i) sample of 6 players (ii) population of all players

(d) Inference for regression leads us to conclude that the slope
(i) may equal zero (ii) equals zero (iii) does not equal zero

(e) Would a confidence interval for the slope contain zero? (Answer yes or no.)

(f) Output is shown when interval estimates are requested for a batting average of .3. Which interval estimates the mean salary of all players whose batting average is .3?

```
New Obs     Fit     SE Fit         95.0% CI              95.0% PI
1         7.518      0.767    (   5.387,   9.648)  (   2.656,  12.379)
Values of Predictors for New Observations
New Obs  BattingA
1           0.300
```

(g) One particular player with a batting average of .3 earned a salary of 4.917 million. Based on the appropriate interval, is this surprisingly low, or is it "in the right ballpark"?