Econ 57
Gary Smith
Fall 2011

Final Examination Answers

1. Hypothesis tests, confidence intervals, and other statistical inferences depend on the *number* of people in the sample, not the percentage of the population.

2. a. binomial
   b. simple regression
   c. chi-square
   d. multiple regression

3. a. ANOVA
   b. chi-square
   c. binomial
   d. multiple regression

4. a. chi-square
   b. binomial
   c. matched-pair t test
   d. binomial

5. P-values are probabilities and cannot be larger than one. We want the probability that the t-value would be so far from zero, in either direction. Here we want the probability that t would be *less* than or equal to –0.636, which is 0.265. The 2-sided p-value is 2(0.265) = 0.530.

6. X and Y are reversed; a t test should have been used.

7. Both tests show statistical significance and, in each case, the fact that the p value is lower, the farther the test statistic is from zero. In fact, the two tests are equivalent.
8. Using the subtraction rule,
   a. $1-\left(\frac{5}{6}\right)^4 = 0.518$
   b. $1-\left(\frac{35}{36}\right)^{24} = 0.491$

9. When A increases, T tends to increase too and both have a negative effect on D. Therefore, the omission of D will make the coefficient of A more negative (increase its absolute value). It will seem that A has a very large effect on D when part of the effect is due to T.

10. The probability of six blackened or six not blackened is $0.5^6 + 0.5^6 = 2/64$. The probability of five blackened or five not blackened is $6(0.5^6) + 6(0.5^6) = 12/64$. Thus the expected value is 5(2/64) + 1(12/64) + 0(40/64) = 22/64.

11. First, $p < 0.001$ for the intercept strongly *rejects* the null hypothesis that the intercept is zero. Second, does it make perfect sense that a box of sugar-less cereal would be free?

12. This is identical to the Monte Hall problem. A's probability stays at one-third because he did not learn anything useful from the warden. This can be confirmed using Bayes' Rule.

13. The coefficients are ceteris paribus. The coefficient of the dummy variable measures whether, for people with the same income, rent is affected by having a full-time job.

14. Regression to the mean. Differences in observed performance tend to exaggerate differences in abilities, so that subsequent performance tends to be closer to the mean.

15. The alternative hypothesis is that not all means are equal, not that every mean is different,

16. a. The test assumes that the samples are independent, but home and away wins are not independent since if a team wins at home the away team necessarily loses.
    b. No, the test should take into account the number of games played. Data for 20 teams each playing 19 games is different from data for 20 teams each playing 38 games.
    c. The standard deviations should be squared

17. The probability of winning is

$$\frac{6}{44}\frac{5}{43}\frac{4}{42}\frac{3}{41}\frac{2}{40}\frac{1}{39}=\frac{1}{7{,}059{,}052}$$

The expected value of a ticket is

$$\$27{,}000{,}000\left(\frac{1}{7{,}059{,}052}\right)+\$0\left(\frac{7{,}059{,}051}{7{,}059{,}052}\right)=\$3.82$$

Net of the \$1 cost, the expected value is \$2.82. (The prize was paid in 20 annual installment of \$1,350,000 beginning immediately, so the present value of the prize was less than \$27 million.)

18. a. Of the 145 players, 80 batted better at night and 65 batted better during the day. Using the binomial distribution to test the null hypothesis that a player is equally likely to bat better at night, the probability of 80 or more successes is 0.1224:

$$P[X \geq 80]=\sum_{x=80}^{145}\binom{145}{x}0.5^{x}0.5^{n-x}=0.1224$$

b. We can use a chi-square test or a difference-in-proportions test:

$$Z=\frac{\frac{14}{31}-\frac{66}{114}}{\sqrt{\frac{\hat{p}(1-\hat{p})}{31}+\frac{\hat{p}(1-\hat{p})}{114}}} \text{ where } \hat{p}=\frac{14+66}{31+114}=0.5517$$

$$=1.2640$$

The two-sided p value is 0.2062.

19. He divided by the observed percent, rather than the expected percent. More importantly, he should have used the number of observations in each category, not the percentage. Statistical significance depends on the size of the sample!

20. There is survivor bias in that those workers still with the company are not a random sample of those hired in 2000. Some employees may have been fired, others may have taken better jobs elsewhere.