Econ 57
Gary Smith
Spring 2017

Final Examination Answers

1. The bank data are for a single year, the divorce data are not.

2. The calculation assumes that the tires rotate independently, which is debatable.

3. [Ben Baskin, "Best Year Ever," Sports Illustrated, August 24, 2015] It is not clear how they came up with 522,675, since the product of these 5 probabilities is 1/1,346,464,516. Nonetheless, the more fundamental problem is that this is a perfect example of the Feinman trap. If *Sports Illustrated* (or Ed Feng) had predicted these 5 events before they happened, that would have been incredible. Not so, predicting them after they already occurred. There were thousands of sporting events in 2015 and the probability of being able to identify some unlikely events *after* they occurred is one.

4. a. This has no effect, as can be visualized with a simple regression scatter plot.
   b. If we divide both sides of the equation by 1,000 in order to convert the data from billions to trillions, we can see that there is no effect on the coefficients of income or wealth, but that the intercept is reduced by a factor of 1,000:
   $$\left(\frac{C}{1{,}000}\right)=\left(\frac{\beta_0}{1{,}000}\right)+\beta_1\left(\frac{Y}{1{,}000}\right)+\beta_2\left(\frac{W}{1{,}000}\right)+\left(\frac{\varepsilon}{1{,}000}\right)$$
   c. This will affect the estimate of the coefficient of income unless income and wealth are uncorrelated.

5. Regression to the mean plus selective recall. No team has a 0% probability of fumbling, so a 0% fumble rate exaggerates their ability not to fumble. We remember when announcers say something that is immediately contradicted and forget the times when they are not contradicted. [Kevin Kaduk, "Broadcasters jinx Browns' turnover-free streak," Yahoo Sports, November 16, 2014.]

6. If this person had a 5% chance of getting into each college and their decisions were independent, the probability of getting into at least one would be 1 minus the probability of being rejected by all three colleges: $1 - 0.95^3$. However, acceptances are not randomly drawn from applications. Some students have more than a 5% chance of being accepted and some less. Nor are the decisions independent. We can imagine a very weak candidate who is almost certain to be rejected by all three colleges.

7. You either win \$1 or lose \$2,047. The probability of losing \$2,047 is the probability of losing 11 bets in a row: $(1/2)^{11} = 1/2{,}048$. The probability of winning \$1 is consequently 2,047/2,048, and the expected value of the Martingale system is
   $$\mu=(\$1)\left(\frac{2{,}047}{2{,}048}\right)+(-\$2{,}047)\left(\frac{1}{2{,}048}\right)=\$0$$
   This illustrates the general principle that if the expected value of a wager is \$0, then the expected value of any sequence of such wagers (no matters how one adjusts the size of the bets as the wagers proceed) is \$0.

8. The large number of students who ended up with a longest streak of 3 or fewer suggests an unfounded belief in the fallacious law of averages. We can multiply the given probabilities by 24 to determine the expected value of the number of students whose longest streak would be in each of these three categories. The chi-square distribution can then be used for a statistical test:

$$\chi^2 = \frac{(4-(.174)24)^2}{(.174)24} + \frac{(17-(.361)24)^2}{(.361)24} + \frac{(3-(.465)24)^2}{(.465)24} = 13.99$$

The p-value is 0.0008.

9. The first student can be anyone. The second student must be one of the remaining 23 students, and the probability of this happening is 23/24. So the probability of electing 17 different students is

$$P = 1\frac{23}{24}\frac{22}{24}\cdots\frac{8}{24} = 0.000423$$

10. Not rejecting a null hypothesis does not prove that the null hypothesis is true. Here, the *p* values may have been large due to the small sample size. It is puzzling that they kept their sample small by looking at 36 days a single year, and that they chose 1977 in a paper published in 1982.

11. a. A 95% confidence interval is:

$$0.589 \pm 1.96\sqrt{\frac{(0.589)(0.411)}{353}} = 0.589 \pm 0.051$$

The binomial model might be inappropriate because the success probability varies from golfer to golfer.

b. A difference-in-proportions test is

$$Z = \frac{.589-.168}{\sqrt{\frac{\hat{p}(1-\hat{p})}{353}+\frac{\hat{p}(1-\hat{p})}{167}}}, \qquad \hat{p} = \frac{0.589(353)+0.168(167)}{353+167}$$

12. a. No, the null hypothesis that the slope is zero is rejected by estimates that are several standard deviations from zero, in either direction.

b. Yes, the formula for the correlation coefficient is

$$r = \frac{\sum_{i=1}^{n}(x_i-\bar{x})(y_i-\bar{y})/(n-1)}{s_x s_y}$$

The calculated value doesn't depend on which variable is X and which is Y.

c. No, the least squares procedure does not assume the explanatory variable is normally distributed.

13. a. two-sample t-test
b. simple regression
c. difference-in-proportions or chi-square
d. chi-square
e. multiple regression

14. [Gus Sauter, Vanguard, quoted in Carol Vincent, "Funds that Make Sense," *Fortune*, March 15, 1999, p. 133.] For each person, the probability of 10 heads is $(1/2)^{10} = 1/1{,}024$

a. If we now consider the 1,000 flippers to be $n = 1{,}000$ Bernoulli trials, each with a probability $\pi = 1/1{,}024$ of successes, then the expected value of the number of people who will get 10 heads is $\pi n = 1{,}000/1{,}024$.

b. The probability that at least one person will get 10 heads is equal to 1 minus the probability that no one gets 10 heads: $1 - (1 - 1/1.024)^{1{,}000} = 0.376$.

15. U is the standardized Z-value. U = –0.8 means that this manager is 0.8 standard deviations below average.

16. The correct answer is (e). The margin for error depends on the size of the sample, not the size of the population (unless the sample is a substantial part of the population, which is not the case here).

17. Using Bayes' Rule, your revised probability is

$$\begin{aligned} P[a \text{ if } 3\text{H}] &= \frac{P[a]P[3\text{H if } a]}{P[a]P[3\text{H if } a] + P[b]P[3\text{H if } b] + P[c]P[3\text{H if } c]} \\ &= \frac{(0.2)0.75^3}{(0.2)0.75^3 + (0.3)0.5^3 + (0.5)1.0} \\ &= 0.1357 \end{aligned}$$

Notice that your probability drops from 0.2 to 0.1357, which is reminiscent of Thomas Paine's argument that, "Is it more probable that nature should go out of her course, or that a man should tell a lie?"

18. The average value of the residuals is not equal to zero.

19. The first student's answer can be anything. If their excuse was untrue, the probability that the second student would select the same tire is 1/4.

20. Data grubbing. This pattern is surely a coincidence, and there are always coincidental patterns, even in random data.