Econ 57
Gary Smith
Fall 2013

Final Examination Answers

1. Three of one sex is the most likely, with a probability of 8/16 = 1/2. There are 16 possible outcomes, and there is only 1 way to get all boys (BBBB), 1 way to get all girls (GGGG), 4 ways to get one girl and three boys (GBBB, BGBB, BBGB, and BBBG), four ways to get one boy and three girls (BGGG, GBGG, GGBG, and GGGB), and 6 ways to get two boys and two girls (BBGG, BGBG, BGGB, GBBG, GBGB, and GGBB).

2. This is like the in-class example of pedestrian accident victims wearing dark cloth, plus it has survivor bias. Their data are the chances of thinking ahead if survive; their recommendation concerns the chance of surviving if think ahead. In order to go from one to the other, we kneed to know the percentage of passengers who did not survive the crash who thought of an exit plan ahead of time (which we will never know since they did not survive). (We also don't know the accuracy of the survivor's recollection of such ill-defined thoughts.)

3. a. People who marry and divorce self-select.
   b. Women may be less likely to marry men who are in poor health and more likely to divorce men whose health suffers the most during marriage.

5. The probability of a particular Venus sequence is (0.39)(0.37)(0.12)(0.12). There are (4)(3)(2)(1) = 24 possible Venus sequences, since any of the four sides can appear on the first astragalus, any of the three remaining sides on the second astragalus, two on the third astragalus, and only the last side on the fourth astragalus. Thus, the probability of a Venus is 24(0.39)(0.37)(0.12)(0.12) = 0.0499, almost exactly 5%.

6. These are observational data. Perhaps, the state chose to put 75 mph speed limits on the safest highways (those that are straight, well lit, with light traffic).

7. Sum (d) minimizes the sum of squared prediction errors

8. The probability that a randomly selected tire will last at least 50,000 miles is

$$P = P\left[X > 50{,}000\right] = P\left[\frac{X-\mu}{\sigma} > \frac{50{,}000 - 40{,}000}{5{,}000}\right] = P\left[Z > 2\right] = 0.0228$$

   Therefore,
   a. P[at least one of the tires will last 50,000 miles] = 1 - P[none] = 1 - P^4 = 1 - 0.9772^4 = 0.088
   b. P[all four tires will last at least 50,000 miles] = P^4 = 0.0228^4 = 0.00000027
   c. Using the sampling distribution for the sample mean,

$$P\left[\bar{X} > 50{,}000\right] = P\left[\frac{\bar{X}-\mu}{\sigma/\sqrt{n}} > \frac{50{,}000 - 40{,}000}{5{,}000/\sqrt{4}}\right] = P\left[Z > 4\right] = 0.0000317$$

9. There should be separate 0-1 dummy variables for each state, instead of one dummy variable for all 50 states. There is no reason why D going from 1 to 2 (from Alabama to Alaska) has the same effect on

household income as does D going from 2 to 3 (from Alaska to Arizona).

10. a. We can't put a probability on the null hypothesis being true unless we do a Bayesian analysis. The p value relates to the probability of observing certain data if the null hypothesis is true, not the probability that the null hypothesis is true if we observe certain data.
    b. The t-value gauges statistical significance, not whether the magnitude of the estimated coefficient is substantial.

11. The coefficient for the roommate dummy measures the effect on GPA of having a roommate, holding constant happiness.

12. Regression to the mean does not assume that competition affects financial ratios. Regression is based on the fact that observed measurements that are far from the mean are probably farther from the mean than are the traits being measured.

13. This is the fallacious law of averages. [Waldemar Kaempffert, (1937, October 10). The Duke experiments in extra-sensory perception. *New York Times*, 102.]

14. There are 6 equally likely rolls for the defender: 1, 2, 3, 4, 5, and 6. Here are the probabilities that the attacker will lose an army in each of these cases:

defender die	Losing Rolls for Attacker	Probability
1	two 1s	(1/6)(1/6) =1/36
2	both dice are less than 3	(2/6)(2/6) = 4/36
3	both dice are less than 4	(3/6)(3/6) = 9/36
4	both dice are less than 5	(4/6)(4/6) = 16/36
5	both dice are less than 6	(5/6)(5/6) = 25/36
6	any roll	(6/6)(6/6) = 36/36

a. Because each of the defender rolls has a 1/6 probability the overall probability is
(1/6)(1/36) + (1/6)(4/36) + (1/6)(9/36) + (1/6)(16/36) + (1/6)(25/36) + (1/6)(36/36) = 91/216 = 0.4213
b. 1 - 91/216 = 125/216 = 0.5787
c. The expected value of the net gain from the attacker's viewpoint is (+1)(125/216) + (-1)(91/216) = 34/216 = 0.1574

15. We can use a chi-square test. The chi-square value is 0.7889. With 6 degrees of freedom, the p value is 0.9924, which is suspiciously high:

$$\chi^2 = \frac{(109-108.67)^2}{108.67} + \frac{(65-66.29)^2}{66.29} + \ldots + \frac{(0-0.01)^2}{0.01} = 0.7889$$

16. a. Omitting an important variable will reduce the model's ability to predict MPG accurately. This will be reflected in a relatively low R-squared.
    b. If the omitted variable is correlated with automobile weight, the estimated coefficient of weight will be biased.

17. Before we learn anything about Mr. Smith's children, there is a 0.25 probability that he has two boys (BB), a 0.50 probability that he has one boy and one girl (either BG or GB, depending on birth order), and a 0.25

probability that he has two girls (GG).

This is like the Monte Hall problem and can be answered by Bayes' Rule, where "WG" signifies that Smith is walking with a girl and we assume that if Smith a one boy and one girl, he is equally likely to walk with either.

$$\begin{aligned}
P[\text{GG if WG}] &= \frac{P[GG]P[\text{WG if GG}]}{P[GG]P[\text{WG if GG}]+P[BG]P[\text{WG if BG}]+P[GB]P[\text{WG if GB}]+P[BB]P[\text{WG if BB}]} \\
&= \frac{0.25(1.0)}{0.25(1.0)+0.25(0.5)+0.25(0.5)+0.25(0.0)} \\
&= 1/2
\end{aligned}$$

A contingency table gives the same answer:

	BB	BG	GB	GG	Total
Observe boy	100	50	50	0	200
Observe girl	0	50	50	100	200
Total	100	100	100	100	400

18. Data mining.

    McQueen, G., and S. Thorley, 1999, Mining fool's gold, Financial Analysts Journal, 55 (2), 61-72. found that the Foolish Four strategy was unimpressive during the years 1949-1972, the 24 years preceding the Gardner test period and also during the Gardner test period if the strategy was implemented on the first trading day of July instead of the first trading day of January. If the strategy really had any merit, it shouldn't be so sensitive to the choice of time period or starting month. (Motley Fool modified this theory one year after introducing it and abandoned it three years later.)

19. Most men are less intelligent than a highly intelligent woman. Unless the intelligence of spouses is perfectly correlated, so that nothing but intelligence matters in match-making, highly intelligent women will, on average, marry men who are of more average intelligence. [From Daniel Kahneman, Thinking Fast and Slow, p. 181.]

20. The figure compares Cornell's tuition and ranking over two different time periods! Tuition was over the 35-year period 1965-2000, while the college ranking was over the 12-year period, 1988-1999. There are two vertical axes—one for cost and one for ranking—and zero has been omitted from both axes to exaggerate the changes in costs and ranking. Not only is zero omitted, but the entire axis and the associated numbers have disappeared, so we have no way of gauging whether the changes were large or small. The *U.S. News* ranking went down, but in college rankings, number 1 is the best! Going from, say, 17th to 13th is actually an improvement.