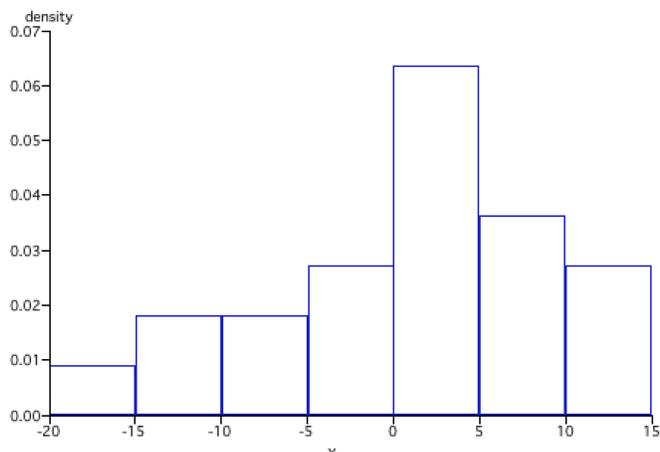


Final Examination (150 minutes)

No calculators allowed. Just set up your answers, for example, $P = 49/52$. BE SURE TO EXPLAIN YOUR REASONING. If you want extra time, you can buy time at a price of 1 point a minute; for example, if your test is handed in 10 minutes after the scheduled finish time, 10 points will be subtracted from the test score.

1. Here is a histogram of 1,000 observations. If the smallest value, -18 , was changed to -14 , indicate whether each of the following would increase, decrease, or stay the same:



- mean
- median
- standard deviation
- interquartile range
(the length of the box in a box plot)

2. A 2016 survey of adults between the ages of 20 and 100 concluded that the elderly have the greatest satisfaction with life. What statistical problem might make this conclusion misleading?
3. Here are several situations where there is an incorrect application of statistical reasoning. Explain briefly what is wrong and why it is wrong.
- 100 students are asked if dining-hall food is better or worse this year than last year. A difference-in-proportions test compared the 57% that chose “improved” with the 43% that chose “worsened.”
 - A one-sample t-test rejected the null hypothesis that the sample mean is 25.
 - A matched-pair t-test reported that the results are statistically significant at the 5% level because the p value is 0.95.
 - A researcher calculated the price-earnings (P/E) ratios of the 30 stocks in the Dow Jones Industrial Average, as of January 1, 2017. A difference-in-means test was then used to compare the average price increase in 2017 for the 15 stocks with the lowest P/Es with the average price increase of all 30 stocks.

4. Two homework problems are assigned for each class period, and a computer program randomly chooses students to answer the problems in class. Every student is equally likely to be chosen for each problem, and the same student can be chosen for both problems. In the seventh class of the semester, the same student was chosen for the 2 homework assignments, leading some students to suspect that there was a bug in the computer program. If there are 24 students in the class and 15 class periods and the computer program is indeed random, what is the probability that, in at least one of these 15 class periods, the same student will be chosen for the 2 homework assignments?

5. The equation $Y = \alpha + \beta X + \varepsilon$ was estimated by ordinary least squares. Explain why you either agree or disagree with these statements :

a. Least squares regression minimizes the sum $\sum_{i=1}^n (X_i - \bar{X})^2$

b. R^2 is equal to the correlation coefficient squared.

c. A doubling of the number of observations, with the new data exactly replicating the original data, would increase R^2 .

d. $R^2 = 1$ means $Y = X$.

6. In 1992, there were 504 large stock mutual funds. Lipper Analytical reported that the average annual return for these 504 funds over the preceding six years was 18.2%. What statistical problem do you see?

7. The Super Bowl Indicator says that the stock market goes up if a team in the National Football Conference or a team in the American Football Conference that used to be in the National Football League wins the Super Bowl; the market goes down otherwise. This indicator worked for 20 of the first 22 Super Bowls (1967 – 1988). In 1990 two finance professor published an article in the *Journal of Finance*, one of the premier peer-reviewed finance journals, arguing that, “although the theoretical relationship connecting the Super Bowl and subsequent stock market movements is not obvious,” the 20-out-of-22 record was highly statistically significant and would have very profitable if followed by investors. To disprove the possibility that it might be a coincidence, they calculated the binomial probability of 20 out of 22 successes as

$$P[X = 20] = \binom{22}{20} 0.5^{20} 0.5^2 = 0.000055$$

a. Why is the correct p value not 0.000055?

b. What is wrong with their logic?

8. A study asked 60 students to predict their scores on a high school math test. It turned out that:
- The average actual score (59.1) was lower than the average predicted score (74.9).
 - There was a 0.65 correlation between the predicted and actual scores.

The author concluded that

- Students over-estimate their ability.
- Test scores can be increased by raising students' self-esteem:

The educational implication of this observation is that there is correlation between self-esteem and self-efficacy, suggesting that teachers strive to raise students' self-esteem in mathematics by way of motivation.... This observation was confirmed by the fact that two students who predicted that they would fail actually did so. This implies that there is also a correlation between negative self-perception or inferiority complex and inadequate motivation to learn or failure.

Provide completely different statistical interpretations of results (a) and (b).

a.

b.

9. The Unified Parkinson Disease Rating Scale (UPDRS) gauges the degree of Parkinson's disease in patients on a scale of 0 to 199, with 0 representing no disability and 199 total disability. Eleven subjects with Parkinson's disease were put on a regiment of riding stationary bikes 3 days a week for 12 weeks:

| Subject ID | UPDRS at week 0 | UPDRS at week 12 |
|------------|-----------------|------------------|
| TA | 16 | 3 |
| JC | 50 | 19 |
| LD | 23 | 5 |
| CE | 14 | 9 |
| FE | 18 | 5 |
| SF | 48 | 24 |
| KO | 20 | 3 |
| HP | 7 | 6 |
| DR | 2 | 3 |
| RS | 13 | 0 |
| JT | 17 | 10 |

Show exactly how you would test whether the cycling was effective. Do not do the calculations, but show the calculations that must be done.

10. Two researchers estimated the following equation for 763 U. S. business industries:

$$P = 2.76 + 0.03789A + 5.94C + 0.318L, \quad R^2 = 0.983$$

[4.35] [40.04] [1.95] [3.98]

where P = industry profits, A = industry assets, C = industry concentration (percent of revenue held by top 50 firms in industry), L = industry profits 5 years earlier, and $[]$ = t -values. Explain why you either agree or disagree with their conclusion: “The reason that the coefficient of industry assets is so statistically significant may be that if industry profits were large 5 years earlier, assets will be large today.”

11. Continuing the previous exercise, explain why you either agree or disagree with their conclusion: “We fail to reject the null hypothesis, at the 5% significance level, that concentration has no effect on profitability. As such, we conclude that market concentration is not related to profitability.”

12. You are going to toss m standard two-sided coins in the air n times. Each time you toss the m coins, you will record x , the number of coins that landed heads side up. When you finish the n trials, you will calculate the mean and standard deviation of the n values of x and draw a histogram.

- a. Which do you expect to have a larger mean value of x , 100 tosses of 4 coins or 1,000 tosses of 4 coins?
- b. Which do you expect to have a larger standard deviation of x , 100 tosses of 4 coins or 1,000 tosses of 4 coins?
- c. Which histogram for x do you expect to have a larger area, 100 tosses of 4 coins or 100 tosses of 40 coins?
- d. Which histogram for x do you expect to look more like a normal distribution, 10,000 tosses of 4 coins or 1,000 tosses of 40 coins?

13. Identify the most appropriate statistical test for each of these studies (for example, two-sample t -test, ANOVA, or multiple regression):

- a. A study of 463 stock splits between January 1, 2007, and December 31, 2016, compared the return on these stocks on the day the split was announced to the return on the S&P 500 index that day.
- b. A study predicted the number of hours a person spent on volunteer work each month based on the person's gender, race, age, and income.
- c. A study compared average starting salaries for recent college graduates hired by Facebook, Amazon, Netflix, Google, and Apple.
- d. A study looked at whether home runs in Major League Baseball games in Fenway Park in Boston tended to be hit to left field, center field or right field.
- e. A study used data for all 58 California counties to compare the difference between the 2009 and 2017 divorce rate with the difference between the 2009 and 2017 the unemployment rate.

14. In a mock courtroom experiment, 144 volunteer jurors were told that a liquor store had been robbed by a man wearing a ski mask. The police arrested a suspect near the store whose height, weight, and clothing matched the clerk's description. The ski mask and money were found in a nearby trash can. After hearing this evidence, the jurors were asked to write down their estimate of the probability that the arrested man "really did it." The average probability was 0.25.

Then a forensic expert testified that samples of the suspect's hair matched a hair found inside the ski mask and that only two percent of the population has hair that microscopically matches the hair found in the ski mask. After hearing this evidence, the jurors' average revised probability of the defendant's guilt was 0.63. What is the Bayesian posterior probability?

15. The 5 positions on a basketball team are shown below. As of 2012, 167 of the 299 head coaches in the NBA played in the NBA. A study of their playing position obtained these data:

| | number | % |
|----------------|-----------|-------------|
| point guard | 61 | 36.5 |
| shooting guard | 42 | 25.1 |
| small forward | 25 | 14.9 |
| power forward | 21 | 12.6 |
| center | <u>18</u> | <u>10.8</u> |
| | 167 | 100.0 |

Test the null hypothesis that NBA coaches who played in the NBA are equally likely to have played each of these five positions. (Just set up.)

16. Since 1937, the Gallup polling organization has been asking Americans, "Do you approve or disapprove of the way XYZ is handling his job as president?" Of the 10 people elected president since 1937, all 10 had lower favorability ratings at the end of their first term than at the beginning. Provide a statistical explanation.

17. In 1996 The Motley Fool advised readers that their Foolish-Four system, based on data for 1973 to 1995, “should grant its fans the same 25 percent annualized returns going forward that it has served up in the past.” The Foolish Four system is:

On January 1, calculate the dividend yield for each Dow stock

Choose the 10 stocks with the highest dividend yields

Of these 10, choose the 5 with the lowest price per share

Of these 5, cross out the one with the lowest price

Invest 40% in the stock with the next lowest price

Invest 20% of your wealth in each of the other three stocks

Why, as a statistician, are you skeptical?

18. A student estimated a regression model using annual data for 1990 through 2015,

$$C = \beta_0 + \beta_1 Y + \beta_2 W + \varepsilon$$

where C = consumption spending; Y = disposable income, and W = household wealth, all in billions of dollars. The initial estimate of β_2 was not statistically significant; so the student omitted wealth from the equation, arguing that, “There could have been multicollinearity between the two variables of income and wealth. Therefore, the least-squares regression is more accurate without wealth included.” Explain why you either agree or disagree with this reasoning.

19. A study found that people who practice the violin more than 2 hours a day have fewer friends than do people who play in weekly bowling leagues. As a statistician, what problem do you see with the conclusion that playing the violin makes a person more introverted?

20. What is visually misleading about this figure?

