

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. Answer this student's question: "Is a random sample of 5% of the students at my school large enough, or should I use 10%?"

2. Identify the most appropriate statistical test for each of these studies (for example, two-sample t test, ANOVA, or multiple regression):
 - a. A researcher investigated a strategy of betting against professional football teams that won the previous week by counting the number of bets won and lost using this strategy during the years 2000-2010.
 - b. A researcher investigated the relationship between price/earnings (P/E) ratios and stock returns by looking at the P/E ratios of 50 popular stocks on December 31, 1972, and the percentage returns on each of these 50 stocks from 1973 through 2010.
 - c. A researcher investigated whether Jews are able to postpone their deaths until after Passover by counting the number of Jews who died during each of the 8 weeks surrounding Passover: 4 weeks before, 3 weeks before, 2 weeks before, 1 week before, 1 week after, 2 weeks after, 3 weeks after, and 4 weeks after.
 - d. A researcher investigated the socioeconomic mobility of California women by using data on the each woman's income, her parent's income, whether or not her mother was born in the United States, and whether or not her father was born in the United States.

3. Identify the most appropriate statistical test for each of these studies (for example, two-sample t test, ANOVA, or multiple regression):
 - a. A researcher looked at daily returns in the stock market over the past 20 years to see if returns tend to be relatively good or bad on certain days of the week.
 - b. A researcher investigated the hot hands phenomenon by tabulating how often professional bowlers followed a strike, spare, or open frame with a strike, spare or open frame.
 - c. A researcher looked at 50 stocks that had been removed from the Dow Jones Industrial Average and, in each case, identified whether, after the substitution, the stock that was removed did better or worse than the stock it replaced it.
 - d. A researcher investigated discrimination against women by a certain employer by comparing male and female salaries, taking into account education, years of experience, and other factors the employer claimed affected salaries.

4. Identify the most appropriate statistical test for each of these studies (for example, two-sample t test, ANOVA, or multiple regression):
- A researcher wanted to see if natural births are more likely at certain times of the day. She divided the day into 24 one-hour periods and recorded the number of births in each one-hour period.
 - A researcher compared professional poker player behavior after big wins and losses by tabulating the number of players who played more aggressively after big losses than after big wins.
 - A researcher looked at the performance of stocks with clever ticker symbols (such as MOO for United Stockyards) by comparing the stock daily returns on a portfolio of clever ticker stocks with the S&P500.
 - A comparison of two methods of forecasting interest rates tabulated the number of times each method was more accurate.
5. What is impossible in the results reported below? How did this error occur?
- $H_0: \mu = 98.6$
t-value: -0.636
degrees of freedom: 36
 $P[t > -0.636]$: 0.735
2-sided p-value: 1.470
6. A study looked at the weight and miles per gallon of 25 cars randomly selected from the *Consumer Reports* 2010 car guide. The authors reported the estimated OLS equation to be $Y = 6418.018 - 122.303X$, with a chi-square test giving a chi-square value of 9.90 with 24 degrees of freedom. What is wrong with the estimated equation? What is wrong with the statistical test?
7. Explain why you either agree or disagree with this reasoning: “I used the F test because the two-sample t test only shows the significance whereas the further the F test is from zero, the stronger is the evidence against the null hypothesis.”
8. De Mere asked the great French mathematician, Blaise Pascal, to compare these probabilities:
- rolling at least one 6 in four throws of a single six-sided die; or
 - rolling at least one double-6 in twenty-four throws of a pair of six-sided dice
- Calculate these two probabilities (just set up).

9. A regression equation was used to estimate the extent to which age at death D is negatively affected by an individual's annual consumption of alcohol A and tobacco T : $D = \alpha + \beta_1 A + \beta_2 T$. Worried about a possible multicollinearity problem due to the positive correlation between A and T , the researcher dropped T . How will this omission affect the estimated coefficient of A ? Explain your reasoning.
10. A traditional game uses six peach pits that have been blackened on one side with fire. The six pits are placed in a cup, shaken thoroughly, and then inspected. The player receives five points if all six pits have a similar side face up (either all blackened or all not blackened) and gets one point if five of the six pits have a similar side face up, with one pit different. Otherwise, the player receives no points. Assuming that each pit is equally likely to land blackened or not-blackened, what is the expected value for one roll?
11. A researcher estimated the equation

$$Y = 0.3834 - 0.0074X$$

[16.81] [3.55]

for 74 cereals, where Y = price per ounce, X = grams of sugar per ounce, and the t values are in brackets. Explain why you either agree or disagree with this reasoning: "The Y -intercept of 0.3834 is statistically significant at the 1% level (with a p -value < 0.001). This result makes perfect sense because one would expect a box of cereal with zero grams of sugar to cost zero cents."

12. Three prisoners, A , B and C , are in separate cells and sentenced to death. The governor has randomly selected one of them to be pardoned. The warden knows which one is pardoned, but is not allowed to tell. Prisoner A begs the warden to let him know the identity of one of the others who is going to be executed. "If B is to be pardoned, give me C 's name. If C is to be pardoned, give me B 's name. And if I'm to be pardoned, flip a coin to decide whether to name B or C ."
- The warden tells A that B is to be executed. Prisoner A is pleased because he believes that his probability of surviving has gone up from $1/3$ to $1/2$, as it is now between him and C . Prisoner A tells C the news, who is also pleased, because he reasons that A still has a chance of $1/3$ to be the pardoned one, but his chance has gone up to $2/3$. What is the correct answer?

13. A multiple regression equation was used to explain rents paid by people who live and work in Phoenix:

$$Y = \alpha + \beta_1 X + \beta_2 D + \varepsilon$$

where Y = monthly rent, X = monthly income, $D = 1$ if full-time job, 0 if part-time. Explain why you either agree or disagree with this interpretation of β_2 : “The coefficient of the dummy variable measures the extent to which renters can afford to pay a higher rent because their full-time job gives them more income.”

14. A researcher investigated the question of whether a Major League Baseball (MLB) player’s performance deteriorates the year after signing a long-term contract. He measured the performance of nonpitchers the year before signing the contract and the year after by the OBPS (on base plus slugging percentage)—the higher, the better. A least regression yielded these results $Y = 0.3770 + 0.5033X$, with $R^2 = 0.268$, where Y = OBPS the year after the contract is signed and X = OBPS the year before the contract is signed. The researcher concluded, “The low R^2 and the fact that slope is less than 1 confirm my theory that OBPS declines after a long-term contract is signed.”

Provide an alternative statistical explanation for why the R^2 might be low and the slope less than 1 even if performance is not affected by signing a contract.

15. A researcher looked at home sales in each quarter: “The null hypothesis for my F test is $H_0: \mu_1 = \mu_2 = \mu_3 = \mu_4$, meaning that average sales in each quarter should equal all other quarters. My alternative hypothesis is that $\mu_1 \neq \mu_2 \neq \mu_3 \neq \mu_4$, meaning that the average difference in sales in each quarter is significant with respect to each quarter.” What is wrong with this interpretation?

16. A researcher looked at the 20 English soccer clubs that played in the Premier League during the 2009-2010 season. These teams won an average of 9.65 out of 19 home games, with a standard deviation of 3.79, and won 4.55 out of 19 away games, with a standard deviation of 3.15. (Tie games were ignored.) He then calculated the t-value for two-sample test of the null hypothesis that the average number of wins in home games is equal to the average number of wins in away games:

$$t = \frac{9.65 - 4.55}{\sqrt{\frac{3.79^2}{20} + \frac{3.15^2}{20}}} = 8.82$$

- What assumption underlying the two-sample t-test is violated?
- Do you agree that 20 is the correct sample size?
- What is wrong with the formula used to calculate this t-value?

17. In many state lotteries, the grand prize carries over to the next drawing if there is no winner. In 1992 the Virginia prize grew to \$27 million. A ticket cost \$1 and would win the grand prize if the buyer correctly picked 6 of 44 numbers, 1 to 44, not necessarily in order. What is the expected value of buying a ticket?
18. A researcher looked at whether Major League Baseball (MLB) players had higher batting averages during games played during the day or games played at night. He found that 14 out of 31 blue-eyed players batted better at night and that 66 out of 114 of the non-blue-eyed players batted better at night. Use these data to test these null hypotheses:
- Overall, there is no difference between day and night games.
 - Eye color does not affect day/night performance.

19. A researcher believed that students whose last names begin earlier in the alphabet are given higher grades than students whose names come later in the alphabet. To test this theory, last names were divided into three groups: early (A - H), middle (I - P), and late (Q - Z). The following data were collected for students receiving A grades in a large introductory college course:

Last Name	Students	A Grades
A - H	40.5%	43.2%
I - P	35.7%	29.2%
Q - Z	23.8%	27.6%
Total	100.0%	100.0%

He calculated the chi-square value as

$$\chi^2 = \frac{(40.5 - 43.2)^2}{43.2} + \frac{(35.7 - 29.2)^2}{29.2} + \frac{(23.8 - 27.6)^2}{27.6} = 2.14$$

What is wrong with his chi-square calculation?

20. A 2010 study of a large California company identified white male employees who had been hired to entry-level positions in 2000. This study found that men with noticeable southern accents had received smaller salary increases over this 10-year period than did men without southern accents. What is the biggest problem with this study?