

Name: \_\_\_\_\_

## Practice Final 1

Read all instructions carefully. Calculators *are* allowed, but you may also leave answers unevaluated provided you do not need the decimal for additional work; e.g.,  $6(\frac{1}{2})^5$ . This is a closed book exam and no notes are allowed; the tables of areas under the normal curve and the  $\chi^2$  curve from the book are provided. You are not to provide or receive help from any outside source during the exam except that you may ask the instructor for clarification of a problem. You have three hours and you should attempt all 14 problems (total 120 points). There will be partial credit, so show all work.

(1) (16 pts) A survey of 400 consumers selected at random from a community of 150,000 found that 42% used coupons regularly.

(a) Find an 80% confidence interval for the percentage of people in the community as a whole who use coupons regularly.

(b) If a sample of 300 consumers from a different community of 100,000 found 39% used coupons regularly, should you conclude there is a difference in the underlying population percentages between the two communities?

(2) (10 pts) Much of mathematics publishing (as well as this exam) is prepared in a markup language called LaTeX that is compiled into a text file. For a certain collection of books, the average LaTeX pages per chapter was 57.8 with a standard deviation of 16.4 pages. The average text pages per chapter was 64 with a standard deviation of 19.6 pages, and the correlation was given by  $r = 0.95$ .

(a) A new chapter has 65 pages of LaTeX writing. Using the information given, predict how many pages of text it will compile into.

(b) Give a 95% confidence interval for the length of the text file that is created from a 65-page LaTeX file. What assumption must you make about the data to do so?

(c) Looking at the raw data you note every chapter's text file is at least as long as its LaTeX file. Does this change your answers to parts (a) and (b)? Why or why not?

- (3) (5 pts) A newspaper advertisement says that sales for retailers who contract with a certain advertising firm increase by \$2500 per month. A sample of 50 businesses that contracted with that firm shows an average increase by \$1600, with a standard deviation of \$300. Can the claim in the newspaper be accurate?

(4) (12 pts) In a certain game of chance, a deck of 52 cards is shuffled and the top cards turned over one by one until the total reaches or passes 3, with an ace counting as 1 and face cards all counting as 10. If you get a 3 on the first card, you win \$10. If your total is 3 as of the second card, you win \$20, and if your total is 3 as of the third card, you win \$40. Otherwise (if your total exceeds 3 at any point) you lose \$1.

(a) Make a probability distribution or box model for your winnings in this game.

(b) What is the expected value after 10 rounds?

(c) For each of the following alterations to the deck, say whether the expected value for your winnings will increase, decrease, or stay the same (you do not have to say how much it will change if it changes).

(i) An ace is removed.

(ii) A king is removed.

(iii) Only the hearts and spades are used.

(5) (4 pts) Suppose you have a data set with five-number summary as follows.

minimum		-5
lower quartile		2
median		10
upper quartile		15
maximum		22

What happens to these values if you do the following to each entry in the data set?

(a) Add 3.

(b) Multiply by  $-2$ .

- (6) (6 pts) A survey of people participating in community service work found the following results on their reasons. Is the reason for participating independent from sex?

	strictly voluntary	court ordered	other
male	79	5	16
female	85	2	13

- (7) (19 pts) At Probability Camp, the day's activities are assigned by a small computer that simulates drawing from a box model with replacement and prints out a ticket with the results, individual for each camper. The box model is as follows.

activity	# tickets
campfire building	2
nature hike	5
swim lessons	5
boating trip	1
lanyard making	3

- (a) What is the probability a camper will make lanyards on exactly two days of a 5-day camping session?

- (b) What is the expected number of times a camper will take a nature hike in a 5-day camp session?

(c) Annie is nervous about fire and Bill is edgy about water. They agree to trade tickets if Annie gets “campfire building” and Bill does not or Bill gets “swim lessons” or “boating trip” and Annie does not. What is the probability that they trade tickets on Monday?

(d) In the same situation as part (c), suppose Annie and Bill do not trade on Monday. Given that, what is the probability that Annie is building campfires or Bill is participating in a water-based activity?

- (8) (12 pts) A computer program is designed to simulate four draws with replacement from a bin containing 15 ping-pong balls painted with the numbers 1 through 5. Each number is painted on the corresponding number of balls (i.e., 1, 2, 2, 3, 3, 3, etc.), and we consider the number of 3s drawn. 500 rounds of machine drawing gave the following results; does the program appear to be working correctly?

# 3s	frequency
0	196
1	211
2	80
3	15
4	0

(9) (4 pts) You are allowed to choose any four whole numbers between 0 and 20 (repeats allowed), with the goal of controlling the standard deviation of that four-element data set.

(a) What do you choose to minimize the standard deviation? Is there more than one option?

(b) What do you choose to maximize the standard deviation? Is there more than one option?

(10) (7 pts) A population of roughly 200,000 people is 22% retirees.

(a) In a sample of 400 people, what percentage of retirees do you expect to find?

(b) What is the standard error for the percentage of retirees in a sample of 400 people from this population?

(c) Is the number in (b) exact or an approximation? If the latter, what in the calculation makes it inexact and is it possible to find the exact value?

(11) (5 pts) Suppose you have a set of imperfectly correlated paired data points. Let  $z_x$  be the result of converting  $x$  into standard units with respect to  $\bar{x}$  and  $SD_x$ , and likewise define  $z_y$ . Assume for  $z_x = 0.85$  that the average  $y$  associated with  $x$  has  $z_y = -0.4$ .

(a) If  $z_x = -0.85$ , what is  $z_y$  for the average  $y$  associated to  $x$ ?

(b) If  $z_x = 0.3$ , what is  $z_y$  for the average  $y$  associated to  $x$ ?

(12) (9 pts) The heights of a collection of brother/sister pairs are correlated with  $r = 0.6$ . The brothers' average height was 69 inches with a standard deviation of 2.6, and the sisters' average height was 64 inches with a standard deviation of 2.4.

(a) What percentage of sisters have height no more than 62 inches?

(b) What percentage of brothers of height 70 inches have sisters with height no more than 62 inches?

(13) (8 pts) A large set of exam scores are summarized by an average of 100 and a standard deviation of 15.5 points. The distribution is roughly normal.

(a) What percentage of the scores are less than 80 points?

(b) What percentage of scores are between 90 and 120 points?

- (14) (3 pts) For a certain data set you are testing, the sample statistic is the 90th percentile of the probability distribution for the null hypothesis. Based on that amount of information can you draw conclusions about accepting or rejecting the null hypothesis?