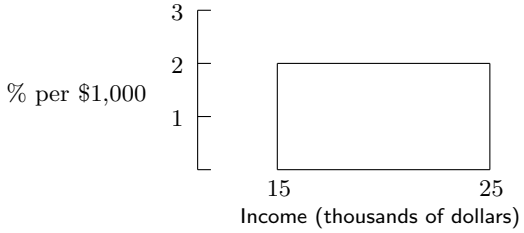
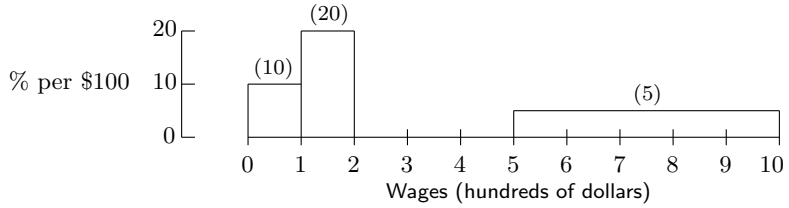
Review Problems for the Math 125 Final

On the final to get full credit you **must show your work**. No work, no credit.
The final has ten or eleven questions, each worth about 9 or 10 points, and possibly a bonus essay question.

1. The sketch below shows one block of the family-income histogram for a certain city. About what percent of the families in the city had incomes between \$18,000 and \$25,000?



2. A histogram of monthly wages for part-time employees is shown below (densities are marked in parentheses). Nobody earned more than \$1,000 a month. The block over the class interval from \$200 to \$500 is missing. How tall must it be? Find the median wage and the 75th percentile wage.



3.

Number of children				3	4–5	6					
Percent of women				12	7	3.5					

Consider the histogram for the above group of women according to the number of children they had. Only part of the data set is given here.

True or false:

(a) The height of the rectangle for 4 or 5 children will be 7%.

(b) The area of the rectangle for 3 children will be 36%.

(c) The area of the rectangle for 4 or 5 children will be 14%.

(d) Since the percentage for 4 or 5 children is twice the percentage for 6 children, its rectangle will be twice as high.

(e) To get endpoints for the rectangles of this histogram, in every case a correction of $\pm 1/2$ is needed.

4. The observed ages of a group of interns at a local hospital were:

$$\{34, 37, 30, 27, 29, 36, 38\}.$$

(a) Find the median of this list.

(b) Find the mode of this list.

(c) Find the average and the standard deviation of this list of ages. (Show your work.)

(d) How many of the ages were within $\frac{1}{2}$ SD of the average age?

(e) How many of the ages were within 1.75 SDs of the average age?

(f) Using the results of part (c), convert the observed value of 30 to standard units.

5. A list of numbers consists of: 13, 9, 11, 7, and 10.
   (a) Find the average and the SD of this list.
      i. Which entries are within 1 SD of the average of the list?
      ii. Which entries are within 2 SDs of the average of the list?
   (b) Convert each entry on the list to standard units (that is, using the average and SD of the list), making a new list of five numbers.
      i. Find the average and SD of the converted list.

6. A survey of a large group of men found that their average height was 67 inches with an SD of 4 inches.
   (a) A 72.4-inch man has what percentile rank in height? (Assume normal curve.)
   (b) How tall is a man whose percentile rank in height is 40%? (Assume normal.)
   (c) Assuming the normal curve, about what percentage of the men in this group are between 62 and 66 inches tall?

7. Among freshmen at a certain university, scores on the Math SAT followed the normal curve, with an average of 500 and an SD of 100. Fill in the blanks; explain briefly.
   (a) A student who scored 505 on the Math SAT was at the _____ percentile of the score distribution.
   (b) To be at the 54th percentile of the distribution, a student needed a score of about _____ on the Math SAT.

8. Among freshmen at a certain university, scores on the Math SAT followed the normal curve, with an average of 500 and an SD of 100.
   (a) About what percentage of the class scored below 666?
   (b) One freshman was 0.7 SDs below average on the Math SAT. About what percentage of the freshmen had lower scores than he did?

9. For the data set below:

x	y
3	4
3	2
1	1
3	5
4	5
4	7

   (a) find r, the correlation coefficient.
   (b) based on these calculations, predict y when $x = 1.5$.
   (c) find the root-mean-square error of the regression line that is used to predict y from x.
   (d) find the slope of that regression line.

10. For the data set below:

x	y
4	4
3	2
4	1
3	5
1	5
3	7

   (a) find the correlation coefficient and the equation of the regression line that predicts y given x.
   (b) find the root-mean-square error of that line.
   (c) predict y when $x = 1.5$.
   (d) determine if the point $(1, 5)$ is above or below that line.

11.

x	y
10	15
8	25
7	19
5	13
6	17
9	23
4	21

Find the correlation coefficient for this data set.

Then use that number to predict y when $x = 8.75$.

Also use that number to find the root-mean-square error of the regression line that is used to predict y from x.

12. Suppose that the correlation between weight (in pounds) and years of schooling completed is about -0.10. True or false and explain:
    (a) Heavier persons tend to be more educated.
    (b) Persons with more education tend to weigh less.
    (c) The correlation between years of schooling completed and weight (in pounds) is about -0.10.
    (d) Lighter persons tend to be more educated.
    (e) If you eat and put on 25 pounds, you will become less educated.

13. For the first-year students at a certain university, the correlation between SAT scores and first-year GPA was 0.60. The scatter diagram is football-shaped. Predict the percentile rank on the first-year GPA for a student whose percentile rank on the SAT was
    (a) 70% (b) 50% (c) unknown

14. In a certain class, midterm scores average out to 60 with an SD of 15, as do scores on the final. The correlation between midterm scores and final scores is about 0.60.
    (a) Estimate the average final score for the students whose midterm scores were 90.
    (b) Predict the final score for a student whose midterm score was 90.
    (c) True or false, and explain: the students who averaged 78 on the final must have averaged 90 on the midterm.

15. A survey of heights and weights of a group of men led to the following results:

$$\text{average height} \approx 67 \text{ inches}, \quad \text{SD} \approx 4 \text{ inches}$$
$$\text{average weight} \approx 152 \text{ pounds}, \quad \text{SD} \approx 30 \text{ pounds}, \quad r \approx 0.57$$

    (a) Find the regression equation for predicting height from weight. Use it to predict the height of a 200-lb. man.
    (b) What is the slope of that equation? Explain how it could be intepreted in the context of predicting the height of a man who is 30 pounds heavier than his friend.
    (c) A prediction made by this equation is likely to be off by about how many inches?

16. Two draws will be made at random from a deck of cards. The unconditional probability of getting the king of hearts on the first draw is 1/52. The unconditional probability of getting the king of hearts on the second draw is 1/52.
    (a) The chance that the first draw turns out to be the king of hearts and the second draw turns out to be the king of hearts is $1/52 \times 1/52 = 1/2704$ when the draws are made _____ replacement, because then the events are _____.
    (b) The chance that at least one of the two draws turns out to be the king of hearts must be calculated as $1 - (51/52 \times 51/52) = 103/2704$ when the draws are made _____ replacement, because then the events are not _____.
    (c) The chance that neither of the two draws turns out to be the king of hearts is $51/52 \times 50/51 = 50/52 = 25/26$ when the draws are made _____ replacement, because then the draws are not _____.

    Fill in the blanks using one option from each pair below for each sentence.

    (with, without) (independent, mutually exclusive)

17. (a) A die is rolled twice. Getting a five on the first roll and getting a five on the second roll are independent.

    A student reasons that the chance of getting fives on both rolls equals $1/6 \times 1/6$ because the rolls are independent.

    Is he or she reasoning correctly?

    (b) A die is rolled twice. Getting a five on the first roll and getting a five on the second roll are not mutually exclusive.

    A student reasons that the chance of getting at least one five on the two rolls does not equal $1/6 + 1/6$ because the rolls are not mutually exclusive.

    Is he or she reasoning correctly?

    (c) A die is rolled twice. Getting a five on the first roll and getting a five on the second roll are independent.

    A student reasons that the chance of getting a five on the second roll given that the first roll came up five is the same as the unconditional chance of getting a five on the second roll because the two draws are independent.

    Is he or she reasoning correctly?

18. Each of two boxes contains ten tickets, marked 1 through 10. One ticket is drawn at random from each box.

    Let A be the result: the ticket drawn from the first box is marked 5.

    Let B be the result: the ticket drawn from the second box is marked 9.

    - Find P(A) and P(B). Are A and B independent? mutually exclusive?
    - Find P(both A and B). Would multiplying the unconditional probabilities, P(A) and P(B), give the correct answer, or are conditional probabilities required? Why or why not?
    - Find P(at least one of A or B)=P(A or B). Would adding the individual probabilities, P(A) and P(B), give the correct answer? Why or why not? Should the actual calculation be done with conditional or unconditional probabilities?

19. A poker hand of five cards is selected at random without replacement.

    Find the chance that the hand contains (defining a picture card as a Jack, Queen, or King):

    (a) All picture cards.

    (b) No picture cards.

    (c) At least one picture card.

    (d) Not all picture cards.

    (e) At least one ace.

20. A die is rolled ten times. What is the chance of never getting a six?

21. A certain panhandler calculates that about 1% of the people he approaches will give him one dollar or more. Find the chance that at least one of 150 individuals, selected randomly and independently, will give him one dollar or more.

22. A box contains one red marble and nine green ones. Five draws are made at random with replacement.

    True or false: the chance that exactly two draws will be red is

$$10 \times \left(\frac{1}{10}\right)^2 \left(\frac{9}{10}\right)^3.$$

23. A die will be rolled 7 times. What is the chance of obtaining exactly two 6's?

24. A coin is tossed 37 times. Estimate the chance of getting 15 heads and 22 tails.

25. A coin is tossed 400 times. Estimate, to the nearest percent, the chance of getting exactly 200 heads. Show all calculations.

26. Bartholemew will draw one hundred and sixty tickets at random with replacement from the box $\boxed{\boxed{0}\ \boxed{0}\ \boxed{0}\ \boxed{1}\ \boxed{2}}$.

    (a) Estimate the chance that $\boxed{1}$ turns up on exactly 27 draws. Show your work. You may round your answer to the nearest whole percent.

    (b) Suppose that the total of the numbers on the one hundred and sixty tickets drawn from the original box came out to 111. The expected sum was _____, the observed sum is _____, the chance error is _____, SE is _____, and the observed sum in standard units is _____.

        i. Which number was subject to chance error and has the chance error built into it: the expected value of the sum, or the observed sum?

27. A die will be rolled 720 times.

    (a) The number of aces (⚀) should end up around _____, give or take _____ or so.

    (b) The percentage of aces (⚀) should end up around _____%, give or take _____% or so.

    (c) Half the time, the percentage of aces (⚀) in 720 rolls will be bigger than _____.

    (d) Find the chance that the percentage of aces (⚀) in the 720 rolls is 18% or more.

28. Five hundred draws will be made at random with replacement from a box that contains ten tickets: three with the number 1 on them, and seven with the number 0 on them.

    Which of these are known exactly before the draws are made, and which are subject to the luck of the draw?

    (i) The expected value for the percentage of 1's among the draws.

    (ii) The observed percentage of 1's among the draws.

    (iii) The percentage of 1's in the box.

29. A simple random sample of 1,000 persons is taken to estimate the percentage of Democrats in a large city. It turns out that 543 of the people in the sample are Democrats.

    (a) True or false: since only one sample was taken, it is not possible to find confidence intervals.

    (b) If you answered false, find a 95%-confidence interval for the percentage of Democrats in the city.

30. A survey organization takes a simple random sample of 625 households from a city of 80,000 households. On the average, there are 2.30 persons per sample household, and the SD is 1.75. Say whether each of the following statements is true or false, and explain.

    (a) The SE for the sample average is 0.07.

    (b) Since only one sample was taken, it would be improper to use a confidence interval to estimate the average household size in the city.

    (c) A 95%-confidence interval for the average household size in the city is 2.16 to 2.44.

    (d) The 95%-confidence level is about right because household size follows the normal curve.

    (e) The 95%-confidence level is about right because, with 625 draws from the box, the probability histogram for the average of the draws follows the normal curve.

31. A box of tickets has an average of 80, and an SD of 6. One hundred draws will be made at random with replacement from this box. The statistician wants to estimate the chance that the average of the draws will be in the range 79.1 to 80.9.

    (a) True or false: to get probabilities for the sum of the draws, the normal approximation will work reasonably well.

    (b) Somebody did the following work: the standard units of the endpoints are $\frac{79.1-80}{6} = -0.15$ and $\frac{80.0-80}{6} = 0.15$ and the chance is about 11.92%, the area under the curve between -0.15 and 0.15.

    Was that correct?

    (c) If incorrect, find the correct estimation for the chance that the average of the draws will be between 79.1 and 80.9. Show your work.

32. Sixty-four measurements are made on the speed of light. These average out to 300,007 and the SD is 10, the units being kilometers per second. (You may assume the Gauss model, with no bias.)

    (a) The speed of light is estimated as _____; this estimate is likely to be off by _____ or so.

    (b) Find an approximate 95%-confidence interval for the speed of light.

33. Four hundred draws are made at random with replacement from a box. The average of the draws is 152.4, and their SD is 18. Someone claims that the average of the box equals 150. Is this plausible?

34. An ESP experiment asks the subject to guess unseen which target was chosen randomly from 10 targets by the researcher.

    (A subject without ESP guesses as if at random; a subject with ESP has a greater than random chance of guessing correctly.)

    Suppose that in 1,000 trials, a subject scores 173 correct guesses.

    (a) Set up the null hypothesis as a box model.

    (b) The SD of the box is _____. Fill in the blank, using one of the options below, and explain briefly.

    $$\sqrt{0.1 \times 0.9} \qquad \sqrt{0.173 \times 0.827}$$

    (c) Make the z-test.

    (d) What do you conclude?

35. A coin is tossed 90 times, and it lands heads 53 times. Is the chance of heads equal to 50%? Or are there too many heads for that?

    (a) Formulate the null and alternative hypotheses in terms of a box model.

    (b) Compute z and P.

    (c) What do you conclude?

36. One day, upon tossing a single die 600 times, I got:

    108 ones, 93 twos, 114 threes, 120 fours, 93 fives, and 72 sixes.

    Compute χ^2 and find P for this experiment. Is the die biased, based on those 600 tosses?

    Justify your conclusion by comparing the P-value you obtained to the benchmark (5% or 1%) that you are using.