

# Study Guide for the Final Examination

Math 125, May 24, 2013 from 3 to 6 p.m. on Chaps. 3–5,10,12,13,17,18,20,21,26,28 emphasizing: 5,14,18

Also see final review (with detailed solutions): problems 6c,8,14a,15a,16a,16b,20,21,25,26bi,27b,28,29,33,36

To get full credit on the final examination you **must show your work**. No work, no credit.

Each question is worth about 9 points. The actual final has 11 questions and a 5-pt. essay.

1. (a) For each list below, work out the average, the deviations from average, and the SD.
  - i. 1, 3, 4, 5, 7
  - ii. 6, 8, 9, 10, 12(b) How is list (ii) related to list (i)? How does this relationship carry over to the average? the deviations from average? the SD?
2. (a) For each list below, work out the average, the deviations from average, and the SD.
  - i. 1, 3, 4, 5, 7
  - ii. 3, 9, 12, 15, 21(b) How is list (ii) related to list (i)? How does this relationship carry over to the average? the deviations from average? the SD?
3. (a) For each list below, work out the average, the deviations from average, and the SD.
  - i. 5, -4, 3, -1, 7
  - ii. -5, 4, -3, 1, -7(b) How is list (ii) related to list (i)? How does this relationship carry over to the average? the deviations from average? the SD?
4. Among applicants to one law school, the average LSAT score was about 169, the SD was about 9. The histogram of LSAT scores followed the normal curve reasonably well.

About what percentage of the applicants scored below:

- (a) 182?
  - (b) a student who scored at the 72nd percentile?
  - (c) a student whose whose score was 1.15 SDs below the average of all the applicants?
  - (d) the median score?
  - (e) the two-thirds of all the applicants who scored the highest?
5. The heights of the men age 18 and over in HANESS averaged 69 inches; the SD was 3 inches. Use the normal curve to estimate the percentage of these men with heights:
- (a) below 76.5 inches (expressed as a percent with accuracy to 2 decimal places).
  - (b) within 0.8 SDs of the average.
  - (c) below +1.55, when expressed in standard units.
  - (d) below the 11th percentile height.
  - (e) that round off to the average, when expressed as the nearest whole number of inches.
  - (f) below the top one-eighth of all the men.
6. An instructor standardizes her midterm and final each semester so the class average is 50 and the SD is 10 for both tests. The correlation between the tests is about 0.50. Estimate the average final score for the students whose midterm scores were

- (a) 30    (b) 86    (c) 60    (d) 50

Plot your regression estimates.

7. In a study of the stability of IQ scores, a large group of individuals is tested once at age 18 and again at age 35. The following results are obtained.

age 18: average score  $\approx 100$ , SD  $\approx 15$

age 35: average score  $\approx 100$ , SD  $\approx 15$ ,  $r \approx 0.80$

- (a) Estimate the average score at age 35 for all the individuals who scored 135 at age 18.
  - (b) Estimate the average score at age 35 for all the individuals who scored 85 at age 18.
  - (c) Estimate the average score at age 35 for all the individuals whose score at age 18 was average.
  - (d) For each of the three estimates above, explain why it supports the concept underpinning the regression method: that is, show that the results are in conformance with the regression effect.
8. In one study of identical twins, the average height was found to be about 68 inches, with an SD of about 3 inches. The correlation between the heights of the twins was about 0.95, and the scatter diagram was football-shaped.
- (a) One member of a pair of twins was measured and found to be 74 inches tall. Predict the height of his twin.
  - (b) Estimate the average height of the twins of individuals in that study who are 74 inches tall.
9. A university has made a statistical analysis of the relationship between Math SAT scores (ranging from 200 to 800) and the first-year GPA (ranging from 0 to 4.0) for students who complete the first year. The results

average Math SAT score = 550    SD = 80

average first-year GPA = 2.6,    SD = 0.6,     $r = 0.4$

Find the regression equation for predicting first-year GPA from Math SAT score.

10. Find the regression equation for predicting weight from height for men age 18–24 in HANESS, based on the following information:

average height  $\approx 70$  inches,    SD  $\approx 3$  inches

average weight  $\approx 180$  pounds,    SD  $\approx 45$  pounds,     $r \approx 0.40$

11. Find the regression equation for predicting the height of a husband from the height of his wife, based on the following information:

average height of husband  $\approx 68$  inches,    SD  $\approx 2.7$  inches

average height of wife  $\approx 63$  inches,    SD  $\approx 2.5$  inches,     $r \approx 0.25$

12. In one study, 100 people had heights measured. It turned out that 25 of them were 67 inches tall, 50 were 68 inches tall, and 25 were 69 inches tall.

Draw a histogram for the results.

13. A box has 10 tickets: 7 of them with the number 1, 2 of them with the number 2, and 1 of them with the number 9.

Draw the probability histogram for the numbers in the box.

14. Toss a hundred pennies in the air and record the number of heads that come up when they fall. Do this several thousand times and plot a histogram for the numbers that you get. You will have a histogram that closely resembles the normal curve, and the more times you toss the hundred pennies the closer your histogram will get to the curve.

A. R. Jensen, *Harvard Educational Review* (1969)

If you keep on tossing this group of a hundred pennies, will your histogram get closer and closer to the normal curve? Or will it converge to the probability histogram for the number of heads in 100 tosses of a coin? Choose one option, and explain briefly.

15. Draws are being made at random with replacement from a box. The number of draws is getting larger and larger. Say whether each of the following statements is true or false, and explain. (“Converges” means “gets closer and closer.”)
- (a) The probability histogram for the sum of the draws (when put in standard units) converges to the normal curve.
  - (b) The histogram for the numbers in the box (when put in standard units) converges to the normal curve.
  - (c) The histogram for the numbers drawn (when put in standard units) converges to the normal curve.
  - (d) The histogram for the numbers drawn converges to the histogram for the numbers in the box.

16. A die is rolled 30 times. What is the chance of never getting a 5?

Find a numeric answer rounded to the nearest hundredth of one percent (nearest 0.01%). This is the preferred solution here.

Also show the other two forms of the answer:

the decimal and the fraction  $1/n$  (rounded off to the nearest integer).

For a small chance the  $1/n$  form is usually the most meaningful representation, expressed verbally like “one chance out of four million,” for example.

17. Two tickets are going to be drawn at random without replacement from the box

1	2	3	4
---	---	---	---

.

- (a) What is the chance that the second ticket turns out to be a 4?
- (b) True or false: since the chance that the second ticket turns out to be a 4 is  $1/4$  and obviously the chance that the first ticket turns out to be a 3 is  $1/4$ , the chance of drawing the 3 first and then the 4 second is  $1/4 \times 1/4$ .

If true, name the rule that was properly used to get this result.

If false, explain what was wrong with the calculation shown, and supply the correct chance of drawing the 3 followed by the 4.

18. A box contains five tickets, with letters A, B, C, D, E. Two draws will be made at random without replacement.

- (a) The things “the first draw results in the A” and “the second draw results in the B” are \_\_\_\_\_ and \_\_\_\_\_.

Fill in using one option from each pair below for each blank.

(independent, dependent) (mutually exclusive, not mutually exclusive)

- (b) To get the probability that the ticket with the A is drawn first and the ticket with the B is drawn second (both of them happen), you would need to \_\_\_\_\_.

Choose one: (add, multiply).

Based on your first answer in part (a), conditional probabilities will be \_\_\_\_\_.

Choose one: (required, unnecessary).

- (c) Calculate the probability that the A is drawn first and the B is drawn second.

19. A coin will be tossed 1200 times. Estimate the chance of getting—

- (a) exactly 590 heads
- (b) 576 to 630 heads, inclusive
- (c) 575 heads or fewer
- (d) 631 heads or more

20. Five hundred draws are made at random with replacement from the box 

0	0	0	1
---	---	---	---

.

True or false and explain:

- (a) The percentage of 1's among the draws is exactly equal to the sum of the draws times 0.2%.
- (b) The expected value for the percentage of 1's among the draws exactly equals 25%.
- (c) The probability histogram for the percentage of 1's among the draws closely resembles the normal curve.

21. You are drawing at random from a large box of red and blue marbles.

- (a) Fill in the blanks.

The expected value for the percentage of reds in the \_\_\_\_\_ equals the percentage of reds in the \_\_\_\_\_. *Options:* sample, population.

- (b) True or false: this expected value and this percentage are only approximately equal.

22. In a certain town there are 30,000 registered voters, of whom 12,000 are Democrats. A survey organization is about to take a simple random sample of 1,000 registered voters.

- (a) The expected value for the percentage of Democrats in the sample is \_\_\_\_\_.  
The SE for the percentage of Democrats in the sample is \_\_\_\_\_.

- (b) The percentage of Democrats in the sample is likely to be around \_\_\_\_\_, give or take \_\_\_\_\_ or so.

23. A university has 25,000 students, of whom 17,000 are undergraduates. The housing office takes a simple random sample of 500 students; 357 of the 500 are undergraduates. Fill in the blanks.

For the percentage of undergraduates in the sample, the observed value is \_\_\_\_\_ but the expected value is \_\_\_\_\_.

24. According to one investigator's model, the data are like 400 draws made at random from a large box. The null hypothesis says that the average of the box equals 50; the alternative says that the average of the box is more than 50. In fact, the data averaged out to 52.7, and the SD was 25. Compute  $z$  and  $P$ . What do you conclude?

25. A gambler is accused of using a loaded die, but he pleads innocent. A record has been kept of the last 60 throws. There is disagreement about how to interpret the data and a statistician is called in.

The observed frequencies for the six numbers on the die are summarized in this table.

<i>Value</i>	<i>Observed frequency</i>
1	9
2	11
3	10
4	8
5	12
6	10

Make a  $\chi^2$ -test of the null hypothesis that the die is fair.

26. **(Essay, worth 5 points)** A simple random sample of size 625 was taken to determine the percentage in a large population who supported a particular candidate.

Describe the steps needed to come up with a confidence interval, justifying each step. Does this single sample—properly taken—yield useful information?