Solutions to Sample of Final Exam Problems

(Final on December 15)
Math 125, Fall 2023

Multiple Choice. $4\frac{1}{2}$ points for each correct response, 1 point deducted for each wrong answer

1. Answer: (D) 1.23% per person

   The percent is found to be (1,476/40,000) $\times$ 100% = 3.69%.

   The blocks for 5 to 7 start at 4.5, because the family size is a discrete variable. Each block is one family size wide and is centered on the number. For 5, it must be centered on 5 and start at 5.5. The number 5 represents one possible size of family.

   The height of the block for families of size 5 to 7 is 1.23% per family size count.

   Here the endpoints are 4.5 people and 7.5 people, as the counting categories are discrete. (Refer to the last paragraph on page 43 and observe the histogram pictured on page 44.) The width is 7.5 people $-$ 4.5 people = 3 people.

   The height is found by dividing the percentage by the width of the base. The result will be densities: percent per unit category. Here it is in % per family size count.

   (Distributing the % over the three separated counts (5, 6, and 7) produces the density: % per person. Each of those bases is 1 family size wide.)

   So the answer is 3.69% divided by 3 family counts, which gives 1.23% per family count.

2. The area of the block for family size from 5 to 7 (inclusive) represents:

   (D) percent of families with that size.

   This is the basic definition. See the box on page 32 of the text.

3. False.

   The vertical scale is a density scale and the height of the block is a density.
   See section 3.3 in the text.

   The area of the block represeents the percentage. See page 32 in the text.

4. (a) The average is $\frac{2+3+4+5+6+7+8+9+10}{6} = \frac{54}{9} = 6$.
   The SD is $\sqrt{\frac{16+9++4+1++0+1+4+9+16}{9}} = \sqrt{\frac{60}{9}} = \sqrt{\frac{20}{3}} = \sqrt{60}/3 = 2.582$.
   The standard units equals $\frac{\text{obs-avg}}{\text{SD}} = \frac{10-6}{2.582} = \frac{4}{2.582} = 1.549$.
   The answer is: (C) 1.549.

   Alt. Method to get SD (top of pg. 74): avg of squares is (4+9+16+25+36+49+64+81+100)/9=384/9=128/3.
   Then the square of the average is $6^2 = 36$. SD = $\sqrt{36-(128/3)} = \sqrt{20/3} = 2.582$.

   (b) Within 1.25 SDs of the average means within 1.25 $\times$ 2.882 = 3.23 or less away from 6. The limits are 2.77 and 9.23. In that range are 7 of the entries, so the percentage is 7/9 $\times$ 100% = 78%.
   The answer is: (D) 78%.

   (c) For the normal curve, the percent within 1.25 SDs of the average means the area under the curve between $z = -1.25$ and $z = +1.25$ (it starts at 1.25 SDs below 0, which is standard units of -1.25, and continues to 1.25 SDs above 0, which is standard units of $+1.25$).
   The required area from the Normal Table is 78.87%, so the answer is: (D) 79%.

5. (A) 0.5 It is the basic definition of standard units. See the box on page 79.

6. (C) 43%

   The endpoints in standard units are:

   $\frac{64-65}{3.32} = \frac{-1}{3.32} = -0.30$ and

   $\frac{68-65}{3.32} = \frac{3}{3.32} = 0.90$.

   The area from -0.30 to 0.90 is broken up into two parts:

   The area from -0.30 to 0 (one half of the area from -0.30 to $+0.30$), which comes out to 11.79%, and

   the area from 0 to 0.90 (one half of the area from -0.90 to $+0.90$), which comes out to about 31.595%.

   Add the two areas to get 11.79% + 31.595%, or about 43%.

7. (C) 0.375

   First find the average x and the average y:

   avg $x = \frac{6+8+2+10+14}{5} = \frac{40}{5} = 8$ and avg $y = \frac{3+11+12+9+15}{5} = \frac{50}{5} = 10$.

   The find SD of x and SD of y:

   SD of $x = \sqrt{\frac{(-2)^2+0^2+(-6)^2+2^2+6^2}{5}} = \sqrt{\frac{80}{5}} = 4$. Same for SD of y; it equals 4.

   Now convert each of the ten numbers in the original table to standard units, and put the ten results in a new table.

   Proceed like this: $\frac{6-8}{4} = \frac{-2}{4} =$ -0.5.

   The ten results are put in the new table and each row is multiplied across to get the products.

x std. units	y std. units	product
-0.5	-1.5	0.875
0	0.25	0
-1.5	0.5	-0.75
0.5	-0.25	-0.125
1.5	1.25	1.875

   Then get the average of the 5 products. They add to 1.875; divide by 5: 0.375.

   That is r. The correlation coefficient, r, for the above data set is: (C) 0.375

   Alt. Method to get SD for x (top of pg. 74): avg of squares is (36+64+4+100+196)/5=400/5=80.

   Then the square of the average is $8^2 = 64$. SD $= \sqrt{80-64} = \sqrt{16} = 4$.

   Alt. Method to get SD for y (top of pg. 74): avg of squares is (9+121+144+81+225)/5=580/5=116.

   Then the square of the average is $10^2 = 100$. SD $= \sqrt{116-100} = \sqrt{16} = 4$.

   Alt. Method to get r (pg. 134, Techincal Note): avg of products is (18+88+24+90+210)/5=430/5=86.

   Then the product of the averages is $8 \times 10 = 80$. $r = (86-80)/(\text{SD of } x \times \text{SD of } y) = 6/(16) = 0.375$.

8. Use the simple formula:

   predicted y = average y + [(x in standard units) $\times$ r $\times$ (the SD of y)].

   The weight of 225 pounds is $\frac{225-175}{45} = \frac{50}{45} = 10/9$ in standard units.

   The formula yields: $69 + [10/9 \times 0.42 \times 3] = 69 + 1.4 = 70.4$ inches.

   Alternative Solution without the Formula: (See the top half of page 160 in the text.)

   The height of 225 pounds is 50 pounds above 175 pounds, the average. That is $50/45 = 10/9$ SDs above average.

   These men who weigh 225 pounds should be r times 10/9 SDs above average in height. So they are $0.42 \times 10/9 = 0.46667$ SDs above average in height.

   That's 0.46667×3 inches = 1.4 inches. So their average is around $69 + 1.4 = 70.4$ inches. Use that estimated average for the prediction. (See page 165 to top of page 166 in the text.)

   Answer: (D) 70.4 inches

9. Answer: (A) 72 inches.

   His height stayed the same.

   Correlation measures association. But association is not the same as causation. (See page 150.)

10. Answer: (C) 2.72257 inches

    The formula is $\sqrt{1-r^2}$ $\times$ the SD of the variable being predicted. (Top box on page 186.)

    Here: $\sqrt{1-r^2} = \sqrt{1-0.42^2} = \sqrt{1-0.1764} = \sqrt{0.8236} = 0.907524$.

    Multiply 0.907524 by 3 inches to get 2.72257 inches, answer (C).

11. Answer: (A) $y = 0.028x + 64.1$

    The slope of the regression line is: $\frac{r \times \text{SD of } y}{\text{SD of } x}$.

    Here, the slope is $\frac{0.42 \times 3}{45} = 0.028$.

    Next write the equation as $y = 0.028x + b$, plug in a known point (x, y) on the line, and solve for b. The point of averages is always on the regression line. That would be a good choice here. It is $x = 175$, $y = 69$.

    $69 = 0.028(175) + b$, $69 = 4.9 + b$, $b = 69 - 4.9 = 64.1$. The line is $y = 0.028x + 64.1$, answer (A).

12. Answer: (A) $0.028(225) + 64.1 = 70.4$.

    This is the same as the answer to problem 8, as it should be.

13. Answer: (C) 37.8 pounds

    Here we are predicting weight from height so we need the slope of the regression line that predicts weight from height. That will be when x is the height variuable and y is the weight variable.

    The slope of the regression line is then: $\frac{r \times \text{SD of } y}{\text{SD of } x}$.

    Here, the slope is $\frac{0.42 \times 45}{3} = 6.3$.

    Next, use the equation: slope = (change in y)/(change in x).

    $6.3 =$ (change in weight)/6

    $6.3 \times 6 =$ change in weight.

    The change in weight will be 37.8 pounds. That is from the regresssion line.

14. (D) 65%

    Do not add up four chances of hitting the bullseye. The outcomes are not mutually exclusive; he could hit the bullseye more than once. (See the box on page 242 in the text.)

    Instead look to the opposite outcome. The opposite of hitting the bullseye at least once is not hitting it at all in the four throws. That outcome is "he missed the bullseye on all four throws."

    Since the results are independent, just multiply together four factors, each of them the unconditional chance of missing (0.77). (See the box on page 232 in the text.)

    We have just found the chance of the opposite of "at least one of the four throws results in a bullseys" to be equal to $(0.77)^4 \approx 0.35 = 35\%$. (Now refer to the second box on page 223 of the text.)

    The chance originally requested (of hitting the bullseye at least once) is

    $100\% - 35\% = 65\%$, answer (D).

15. (F) 42%

    This is a binomial probability. The clue is the word "exactly" in the question. Find n, k, and p.

    $n = 4$, $k = 1$, $n - k = 3$, $p = 0.23$, and $1 - p = 0.77$.

    This is hitting the bullseye once, so $k = 1$ and $p = 0.23$, the chance of hitting the bullseye on each throw.

    The binomial formula is: The probability $= \frac{n!}{k! \times (n-k)!} p^k (1-p)^{n-k}$. (page 259)

    Here $\frac{4!}{1! \times 3!} (0.23)^1 (0.77)^3 \approx 0.42 = 42\%$, answer (F).

16. Answer: (F) 31/36.

    - One idea is to apply the de Méré method, detailed in the box on page 250.
    - Another method is possible. It is presented in footnote 2 to Chapter 14 on page A-16.
    - A third, more common method, is not in this text. It is only hinted at in Example 5 on p. 242 and in the *Technical Notes* on pg. 245. P(A or B) = P(A) + P(B) − P(A and B). Those that have already learned this formula elsewhere are permitted to use it here:

      $1/6 + 5/6 - (1/6 \times 5/6) = 31/36$. (A and B are independent.)

      (Be aware that this formula only holds for combining two outcomes, while the de Méré method may be used for a chance when there are more than two outcomes being combined (P(A or B or C), for example).

    The solution shown below is for the de Méré method.

    Do not add the chances of events A and B: they are not mutually exclusive. A four could happen on the first roll with a diffrent result on the second roll.

    Instead find the chance of the opposite event to "A or B", and subtract that probability from 1 (or 100%). See the second box on page 223 and the box on page 250.

    State the opposite result: a non-four on the first roll and the same result on the second roll.

    Events A and B are indepndent: it does not matter whether or not a four happened on the first roll, the chances of not getting a different number on the second roll will be 1/6.

    The chance of the opposite is P(not A) × P(not B).

    That is $5/6 \times 1/6$ using the multiplication rule for independent events (on page 232).

    The chance of the opposite is 5/36; so the chance of A or B is $1 - 5/36 = 31/36$, answer (F).

17. (B) 3%

    This is a binomial probability, with the clue the word "exactly."

    However, most calculators will not return the value of 350 factorial (350!).

    The normal approximation to the probability histogram for the binomial should be used, as the number of tosses is large enough.

    Set up the box model: one 0 and one 1. The average and the SD of this box are both 1/2.

    The number of heads is modeled by the sum of the draws from this 0–1 box.

    Find the EV for the sum of the draws and the SE for the sum of draws.

    They are: $350 \times 0.5 = 175$, and $\sqrt{350} \times 0.5 = 9.354$.

    The box of the probability histogram for 182 heads runs from 181.5 to 182.5 heads.

    (See page 317, example 1(a).)

    The standard units for the endpoints are:

    $(181.5 - 175)/9.354 \approx 0.70 \quad \text{and} \quad (182.5 - 175)/9.354 \approx 0.80.$

    The area under the normal curve between 0.70 and 0.80 $= 57.63\%/2 - 51.61\%/2 = 3.81\% \approx 3\%$.

18. Answer: (D).

    For the difference between the number of heads observed and the expected number of heads, see figure 2 on page 276 and the last comment by *Assistant* on pages 276 to 277.

    That difference tends to get larger as the number of tosses increases.

    For the difference between the percentage of heads observed and 50%, see figure 1 on page 275 and the last comment by *Assistant* on pages 276 to 277.

    That difference tends to get smaller as the number of tosses increases.

    This agrees with (D).

19. (D) 3.44%

    To get the proper counting box needed to model the percentage, put 1 on the even-numbered cards and put 0 on all the other cards.

    The give or take for the sample percentage is just the SE for the sample percentage, which has the formula: $\dfrac{\text{SD of the box}}{\sqrt{\text{no. of draws}}} \times 100\%$.

    (See the technical note on page 362 of the text.)

    Use the shortcut to find the SD of the box:

    The fraction of tickets with 1 is 20/52, and the fraction of tickets with 0 is 32/32.

    The formula gives: $(1-0)\sqrt{\frac{32}{52} \times \frac{20}{52}} = \sqrt{0.38462 \times 0.61538} = \sqrt{0.23669} = 0.4865.$

    (See the formula on page 298.)

    Then the answer is: $\frac{0.4865}{\sqrt{200}} \times 100\% = 3.44\%$, answer (D).

20. (B) 66.4 to 67.6%

   (a) Set up a 0–1 box where 1 is on each reader of newspapers and 0 on the others.
   (b) Find the sample percent: $(2310/3500) \times 100\% = 66\%$.
   (c) Use the bootstrap to estimate the fraction of 1's in the box: 2310/3500 or 0.66. The fraction of 0's in the box will be 0.34.
   (d) Estimate the SD of the box by the shortcut: $\sqrt{0.66 \times 0.34} = \sqrt{0.2244} = 0.4737$.
   (e) Estimate the SE for the percentage: $\frac{\text{SD of box}}{\sqrt{\text{no. of draws}}} \times 100\% = \frac{0.4737}{\sqrt{3500}} \times 100\% = \frac{47.37\%}{59.161} = 0.8\%$.
   (f) Write 95%-confidence interval as sample percent plus or minus two SEs.
   (g) Answer is: $66\% \pm 1.6\%$. That's 66%-1.6% to 66%+1.6%, or 64.4% to 67.6%, answer (B).

21. (C) $574

   The standard error for the average is (using technical note on page 415) the SD of the box, divided by the square root of the number of draws. With a large random sample, the SD of the sample may be used to estimate the SD of the box.

   Here it gives $\$36{,}300/\sqrt{4000}$=$\$36{,}300/63.246\approx\$574$, answer (C).

   (The second bullet on page 417 explains what the SE says about how far off the average income in the entire city (the population average) is from the sample average.)

22. (D) 82.30%

   The standard error for the average is (using technical note on page 415) the SD of the box, divided by the square root of the number of draws.

   Here it gives $20/\sqrt{81} = 20/9 \approx 2.2222$.

   A sample of size 81 should be large enough for the normal approximation to apply.

   Find the standard units for 117 and 123: $(117 - 120)/2.2222 = -1.35$ and $(123 - 120)/2.2222 = 1.35$.

   The area under the normal curve from -1.35 to $+1.35$ is just the table entry: 82.30%, answer (D).

23. (B) 0.119 pounds

   The SD of the weighings estimates the SD of the error box.

   (See the box on page 451 in the text.)

   A weighing is modeled by drawing a ticket from the error box and adding its value to the exact weight. A typical ticket will be off from 0 (the average of the error box) by about the SD: 0.65 pounds. (See the box on page 67 in the text.)

   The actual weight of the patient is the population average for weighings. The sample average is 187.63 pounds. Now refer to the second bullet on page 417 of the text.

   The SE for the sample average (using the Technical Note on the top of page 415) is found to be $\frac{\text{SD of box}}{\sqrt{\text{no. of draws}}} = \frac{0.65}{\sqrt{30}} = 0.119$ pounds. (Refer to page 416 to estimate the SD of the box.)

   So, the sample average (187.63 pounds) is off from the population average (the actual weight of the patient) by about 0.119 pounds.

   Answer: (B) 0.119 pounds

24. (D) $p = 5.48\%$, fair

The null hypothesis is that the coin is fair, with a box having one 0 and one 1.

The alternative hypothesis is that the coin gets too many heads.

Model the number of heads for the coin by the sum of the draws from the box.

The average and SD of the box are both 1/2.

The EV for the sum of the draws is number of draws $\times$ average of the box.

The SE for the sum of the draws is $\sqrt{\text{number of draws}} \times$ SD of the box.

Here: EV $= 55 \times (1/2) = 27.5$, and SE $= \sqrt{55} \times (1/2) = 3.708$.

Now, to approximate the chance of getting 34 or more heads, draw the block of the probability histogram for 34 heads. It runs from 33.5 to 34.5 heads. Then the endpoint for 34 or more heads (the tail) is 33.5. The chance for this result or any result more extreme is needed.

Using 34 in the standard units calculation is wrong. The continuity correction ($\pm 1/2$) is needed here because the number of tosses is small (large enough for the normal curve, but at the same time small enough to need the $\pm 1/2$ correction). Read the first paragraph on page 317 of the text and review example 1 below.

The normal curve may be used to approximate the chances for the sum of the draws because the number of draws is large enough (not too large but more than 25). See the Central Limit Theorem in the box on page 325 of the text.

The standard units are: $\frac{33.5-27.5}{3.708} \approx 1.60$.

The central area for $z = \pm 1.60$ is 89.04%. Two tails are $100\% - 89.04\% = 10.96\%$.
The right tail—which is P—is half of that: 5.48%.

Since $P > 5\%$, conclude it was just chance error from a fair coin.

Answer: (D) $P = 5.48\%$, fair

25. (A) 1%, unfair

There are 5 degees of freedom: one less than the number of categories.

The expected values are all $600 \times (1/6)$, or 100.

For each category, subtract the expected value from the observed, square the result, and then divide by 100, the expected.

Here is the first calculation: $\frac{(87-100)^2}{100} = 169/100 = 1.69$.

Add the five results. This is χ^2: $1.69 + 2.25 + 0.36 + 6.25 + 4.84 + 0.01 = 15.4$.

Or, since they are all over 100, add the six squares of the (observed $-$ expected) and then divide by 100: $\frac{169+225+36+625+484+1}{100} = \frac{1540}{100} = 15.4$.

From the χ^2-table, the value of P is estimated at 1%. It's a bit less than 1%.

Since P is less than 1%, we do not attribute the difference to chance error. Instead, we reject the null hypothesis and conclude that the die is unfair.

Answer: (A) 1%, unfair