Solutions to Supplementary Practice for the Final Exam

(Final on Dec. 15)
Math 125, Fall 2023

Multiple Choice. 5 points for each correct response, 1 point deducted for each wrong answer

1. Answer: (A) 5.5, 1.15

   The percent is found to be (414/12,000) $\times$ 100% = 3.45%.

   The blocks for 6 to 8 start at 5.5, because the family size is a discrete variable. Each block is one family size wide and is centered on the number. For 6, it must be centered on 6 and start at 5.5. The number 6 represents one possible size of family.

   The height of the block for families of size 6 to 8 is 1.15% per family size count.

   Here the endpoints are 5.5 people and 8.5 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 8.5 people $-$ 5.5 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 (6, 7, and 8) produces the density: % per person. Each of those bases is 1 family size wide.)

   So the answer is 3.45% divided by 3 family counts, which gives 1.15% per family count.

2. The average is $\frac{(-1)+(-1)+2}{3} = \frac{0}{3} = 0.$

   The SD is $\sqrt{\frac{1+1+4}{3}} = \sqrt{\frac{6}{3}} = \sqrt{2}.$

   Alt. Method to get SD (top of pg. 74): avg of squares is (1+1+4)/3=6/3=2.
   Then the square of the average is $0^2 = 0$. SD $= \sqrt{2-0} = \sqrt{2}$.

   The answer is: (C) $\sqrt{2}$.

3. The average is $\frac{(-1)+(-1)+2}{3} = \frac{0}{3} = 0.$

   The SD is $\sqrt{\frac{1+1+4}{3}} = \sqrt{\frac{6}{3}} = \sqrt{2}.$

   The standard units equals $\frac{\text{obs-avg}}{\text{SD}} = \frac{2-0}{\sqrt{2}} = \frac{2}{\sqrt{2}} = \frac{\sqrt{2}\times\sqrt{2}}{\sqrt{2}} = \sqrt{2}.$

   Alt. Method to get SD (top of pg. 74): avg of squares is (1+1+4)/3=6/3=2.
   Then the square of the average is $0^2 = 0$. SD $= \sqrt{2-0} = \sqrt{2}$.

   The answer is: (B) $\sqrt{2}$.

4. The average is $\frac{(-1/2)+(-1/2)+(-1/2)+(-1/2)+2}{5} = \frac{0}{5} = 0.$

   The SD is $\sqrt{\frac{(1/4)+(1/4)+(1/4)+(1/4)+4}{5}} = \sqrt{\frac{5}{5}} = \sqrt{1} = 1.$

   Alt. Method to get SD (top of pg. 74): avg of squares is ((1/4)+(1/4)+(1/4)+(1/4)+4)/5=5/5=1.
   Then the square of the average is $0^2 = 0$. SD $= \sqrt{1-0} = \sqrt{1} = 1$.

   The answer is: (B) 1.

5. Answer: (F) 91%.

   A height that is 1.35 SDs above average will have standard units of +1.35.
   (See the box on page 79.)

   The area below 1.35 on the normal curve will approximate the percent of the group with heights less than that woman. (See examples 8 and 9 on pages 85 to 87.)

   The area below 1.35 is broken down into the area below 0 and the area between 0 and 1.35. The area below 0 is 50%. The area from 0 to 1.35 is half the area between −1.35 and 1.35. The area from −1.35 to 1.35 is about 82%.

   So the area from 0 to 1.35 is, by symmetry of the normal curve, half of 82%: 41%.

   Add 50% and 41%. The answer is: (F) 91%.

6. Answer: (B) equal to

   Observe that 97 percent are below the 97th percentile and 95 percent are below the 95th percentile. That leaves 2 percent of the women between them.

   Similarly for the 75th and 77th percentiles: 2 percent between them.

   Both are exactly equal to 2 percent, so equal.

7. Answer: (E) 4.75 inches

   For insight into this problem, see Example 10 on pages 90 and 91.

   First find both percentiles in standard units.

   The 5th percentile is $-z$ with 5% of the area below it. So above z is also 5% of the area. That leaves $100\% - 5\% - 5\% = 90\%$ between $-z$ and Z. The area of 90% is between −1.65 and 1.65. Conclude 5th percentile in standard units is about −1.65.

   The 60th percentile is a positive value of z that has 60% of the area below it. Break the area into below 0 (50%) and from 0 to z (10%).

   When 0 to z is 10%, $-z$ to 0 is also 10%. That makes $-z$ to z about 20%. The area of 20% is between −0.25 and 0.25. Conclude that the 60th percentile in standard units is about +0.25.

   The difference between the percentiles in standard units is $0.25 - (-1.65) = 1.9$.

   Standard units tell how many SDs the height is above or below average. (See the box on page 79.)

   The difference in standard units, 1.9, tells how many SDs apart the height are.

   In inches, it is 1.9(2.5 inches)=4.75 inches, answer (E).

   Alternate Solution: replace last 4 paragraphs, starting from "The difference between ⋯ " with below.

   Given standard units = −1, 65, find the height. $-1.65 = (x - 65)/2.5$. $x = 60.875$ inches.

   Given standard units = 0.25, find the height. $0.25 = (x - 65)/2.5$. $x = 65.625$ inches.

   Then subtract the heights: $65.625 - 60.875 = 4.75$ inches, answer (E).

8. Answer: (A) −1.

   Before doing any calculations plot the three points on a graph.

   They clearly lie on a straight line that has a negative slope. That means that the correlation coefficient is −1. (See the last sentence of the second paragraph on page 128.)

   Also, without graphing, note that the points wiil be on the graph of the linear equation $y = 4 - x$, which has negative slope of −1. A graph of a linear equation is known to be a straight line.

   Not noting that the points are on a straight line leads to calculations that are extremely messy. The standard units come out to be −1.224744871, etc.

9. Use the simple formula:

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

   A midterm score of 80 points $\frac{80-50}{25} = 1.2$ in standard units.

   The formula yields: $55 + [1.2 \times 0.6 \times 15] = 55 + 10.8 = 65.8$ points.

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

   A midterm score of 80 points is 30 points above 50 points, the average. That is $30/25 = 1.2$ SDs above average.

   Students who scored 80 points on the midtem should score r times 1.2 SDs above average on the final exam. So they are $0.6 \times 1.2 = 0.72$ SDs above average on the final exam.

   That's 0.72×15 points $= 10.8$ points. So their final score is around $55 + 10.8 = 65.8$ points. Use that estimated average for the prediction. (See page 165 to top of page 166 in the text.)

   Answer: (B) 65.8 points.

10. Answer: (C) 12 points

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

    Here: $\sqrt{1-r^2} = \sqrt{1-0.6^2} = \sqrt{1-0.36} = \sqrt{0.64} = 0.8$.

    Multiply 0.8 by 15 points to get 12 points, answer (C).

11. Answer: (D) $y = 0.36x + 37$

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

    Here, the slope is $\frac{0.6 \times 15}{25} = 0.36$.

    Next write the equation as $y = 0.36x + b$, 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 $x = 50$, $y = 55$.

    $55 = 0.36(50) + b$, $55 = 18 + b$, $b = 55 - 18 = 37$. The line is $y = 0.36x + 37$, answer (D).

12. Answer: (B) $0.36(80) + 37 = 65.8$.

    This is the same as the answer to problem 9, as it should be.

13. Use the simple formula:

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

    The height of 71 inches is $\frac{71-65}{3} = 2$ in standard units.

    The formula yields: $130 + [2 \times 0.4 \times 12] = 130 + 9.6 = 139.6$ pounds.

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

    The height of 71 inches is 6 inches above 65 inches, the average. That is 2 SDs above average.

    These women who are 71-inches tall should be r times 2 SDs above average in weight. So they are $0.4 \times 2 = 0.8$ SDs above average in weight.

    That's 0.8×12 pounds $= 9.6$ pounds. So their average is around $130 + 9.6 = 139.6$ lbs. Use that estimated average for the prediction. (See page 165 to top of page 166 in the text.)

    Answer: (D) 139.6 pounds

14. Answer: (E) 11 pounds

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

    Here: $\sqrt{1-r^2} = \sqrt{1-0.4^2} = \sqrt{1-0.16} = \sqrt{0.84} = 0.9165$.

    Multiply 0.9165 by 12 pounds to get 11 pounds, answer (E).

15. Answer: (D) weight = 1.6 × height +26

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

    Here, the slope is $\frac{0.4 \times 12}{3} = 1.6$.

    Next write the equation as $y = 1.6x + b$, 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 $x = 65$, $y = 130$.

    $130 = 1.6(65) + b$, $130 = 104 + b$, $b = 130 - 104 = 26$. The line is $y = 1.6x + 26$, answer (D).

16. Answer: (D) $1.6(71) + 26 = 139.6$.

    This is the same as the answer to problem 13, as it should be.

17. Use the simple formula:

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

    The weight of 150 pounds is $\frac{150-130}{12} = \frac{20}{12} = 5/3$ in standard units.

    The formula yields: $65 + [5/3 \times 0.4 \times 3] = 65 + 2 = 67$ inches.

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

    The height of 150 pounds is 20 pounds above 130 pounds, the average. That is $20/12 = 5/3$ SDs above average.

    These women who weigh 150 pounds should be r times 5/3 SDs above average in height. So they are $0.4 \times 5/3 = 2/3$ SDs above average in height.

    That's $2/3 \times 3$ inches = 2 inches. So their average is around $65 + 2 = 67$ inches. Use that estimated average for the prediction. (See page 165 to top of page 166 in the text.)

    Answer: (B) 67 inches

18. Answer: (B) 2.75 inches

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

    Here: $\sqrt{1-r^2} = \sqrt{1-0.4^2} = \sqrt{1-0.16} = \sqrt{0.84} = 0.9165$.

    Multiply 0.9165 by 3 inches to get 2.75 inches, answer (B).

19. Answer: (A) height = 0.1 × weight +52

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

    Here, the slope is $\frac{0.4 \times 3}{12} = 0.1$.

    Next write the equation as $y = 0.1x + b$, 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 $x = 130$, $y = 65$.

    $65 = 0.1(130) + b$, $65 = 13 + b$, $b = 65 - 13 = 52$. The line is $y = 0.1x + 52$, answer (A).

20. Answer: (A) $0.1(150) + 52 = 67$.

    This is the same as the answer to problem 17, as it should be.

21. Answer: (B) or (C).

    Just because 73-inch-tall women averaged 142.8 pounds in weight, do not conclude that 142.8-pound women averaged 73 inches in height. That is the regression fallacy. The connection between 73 inches and 142.8 pounds is not important. It is just the result from regression toward the mean. Continuing in the other direction, further regression toward the mean will happen. The two regressions will get even closer to the mean, not back to the start. (This is discussed in sections 10.4 and 10.5.)

22. Answer: (C) 67 inches.

    Her height stayed the same.

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

23. Answer: (B). (See section 10.5 and summary point 6 on page 179.)

24. Answer: (B) 38%

    First find the root-mean-square error of the regression line.

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

    Here: $\sqrt{1-r^2} = \sqrt{1-0.62^2} = \sqrt{1-0.3884} = \sqrt{0.6116} = 0.7846$.

    Multiply 0.7946 by 27.78 pounds to get 21.796 pounds.

    Now, 10.9 pounds is within about $10.9/21.796 = 0.5$ r.m.s. errors of the regression line. Using the ideas for 1 r.m.s. error found on pages 182 (middle) to 183 (middle), the estimate is about 38%. This rule of thumb holds for many data sets, but not all.

25. Answer: (D) 74th

    First find the 91st percentile in standard units.

    The 91st percentile is a positive value of z that has 91% of the area below it. Break the area into below 0 (50%) and from 0 to z (41%).

    When 0 to z is 41%, $-z$ to 0 is also 41%. That makes $-z$ to z about 82%. The area of 82% is between -1.35 and 1.35. Conclude that the 91st percentile in standard units is about $+1.35$.

    This man is 1.35 SDs above average in height. The regression method predicts he will be $0.49 \times 1.35 \approx 0.65$ SDs above average in weight. Finally this can be translated back into a percentile rank. (See Example 2 on page 166.)

    Break the area below $+0.65$ into two parts: below 0 (50%) and from 0 to $+0.65$ (half of -0.65 to 0.65: half of 48% = 24%). Together there is 74% of the area below $+0.65$.

    That is the answer: The percentile rank on weight is predicted as 74%, answer (D).

26. Answer: (A) less than

    The events are not mutually exclusive; it is possible to roll five spots multiple times.

    See the box on page 242. The same idea applies to more than two events. If they are all mutually exclusive, the individual chances may be added to get the chance that at least one event occurs. If they are not—as is the case here—the sum will be too big.

    The sum, 2/3, will be too big. The chance of at least one five will be less.

27. Answer: (B) equal to

    The events are mutually exclusive. The tickets are drawn without replacement; it is impossible to draw the ticket numbered 5 more than once.

    See the boxes on page 241. The same idea applies to more than two events. If they are all mutually exclusive, the individual chances may be added to get the chance that at least one event occurs.

    The chance that at least one five will be drawn is equal to the sum of the draws, 2/3.

28. Answer: (B), (D), or (F).

(A) is incorrect because, while the chance of at least one will be the sum of the draws (the results being mutually exclusive), the chance of drawing the red ticket on the second draw must be unconditional. Just as the chance of drawing the red on the first draw is without conditions, the chance of drawing the red on the second draw must be without conditions.

To understand what that chance will be, read carefully Example 2(a) on pages 226 and 227, especially the paragraph on page 227 staring with "The 1/51 $\cdots$."

Once it is established that the unconditional chance of drawing the red card on the second draw is 20%, (B) is shown to be correct.

Answers (D) and (F) are absolutely correct. They are complicated methods of arriving at the same chance as in part (B), showing that often there are several valid methods to solve a problem. Consider these fancy alternate solutions as challenge or enrichment problems, beyond the standard material of the course.

However, in some sense, the method of (D) is to be preferred because of the confusion between conditional and unconditional chances in the statements of (A) and (B).

29. Answer: (C) or (E).

The method in (C) is the de Méré method, detailed in the box on page 250.

The method in (E) is another method, 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 pg. 242 and in the *Technical Notes* on pg. 245. It is: $\text{P(A or B)} = \text{P(A)} + \text{P(B)} - \text{P(A and B)}$.

Those that have already learned this formula elsewhere are permitted to use it here:

$1/5 + 1/5 - 1/25 = 9/25 = 36\%$.

(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).

30. Answer: (B) $1 - (\frac{5}{6} \times \frac{2}{3} \times \frac{3}{4}) = 1 - \frac{5}{12} = \frac{7}{12}$

The three events are not mutually exclusive, so to find the chance that at least one event occurs, do not add the chances (box on page 242).

Instead use the de Méré system detailed on page 250.

The opposite of at least one championship team is no championship team. That means all three teams did not win championships. They are independent, so use the method of page 232 for three events.

The chance that the Red Sox do not win is $1 - 1/6 = 5/6$ (page 223, second box). With fractions instead of percents, subtract from 1 instead of 100%.

The chance tht the Patriots do not win is similarly $1 - 1/3 = 2/3$.

The chance that the Bruins do not win is similarly $1 - 1/4 = 3/4$.

The product of the three probabilities will give the chance that all three did not win championships: $(5/6)(2/3)(3/4) = 5/12$.

This is the chance of the opposite of at least one championship.

The chance of at least one championship is $1 - 5/12 = 7/12$, answer (B).

31. (E) 73.8%

Do not add up six chances of answering correctly. The outcomes are not mutually exclusive; the student could get more than one question correct. (See the box on page 242 in the text.)

Instead look to the opposite outcome. The opposite of answering at least one question correctly is not answering any of the six correctly. That outcome is "the student answered all six questions incorrectly."

Since the results are independent, just multiply together six factors, each of them the unconditional chance of answering incorrectly (0.80). (See the box on page 232 in the text.)

We have just found the chance of the opposite of "getting at least one question correct" to be equal to $(0.80)^6 \approx 0.262 = 26.2\%$. (Now refer to the 2nd box on page 223 of the text.)

The chance originally requested (of being late at least once) is $100\% - 26.2\% = 73.8\%$, answer (E).

32. (D) 39.3%

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

$n = 6$, $k = 1$, $n - k = 5$, $p = 0.2$, and $1 - p = 0.8$.

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

Here $\frac{6!}{1! \times 5!}(0.2)^1(0.8)^5 \approx 0.393 = 39.3\%$, answer (D).

33. Answer: (D) 34.5%

The chance of at least one may be broken down into the chance of exactly one or at least two.

These two results are mutually exclusive, so P(at least one) = P(exactly one) + P(at least two).

This gives 73.8% = 39.3% + P(at least two).

Solving for P(at least two), gives

P(at least two) = 73.8% − 39.3% = 34.5%, answer (D).

34. (C) 21.76%.

Without replacement, the draws are dependent. Use the Multiplication Rule found in the box on page 229 of the text.

Calculating the chance using conditional probabilities:

$$\frac{36}{52} \times \frac{35}{51} \times \frac{34}{50} \times \frac{33}{49} = \frac{99}{455} \approx 21.76\%, \text{ answer (C).}$$

35. (E) 78.24%.

The chance of a face card on any draw is 30.8%. But do not add the chances of a face card on each draw to find the chance of at least one, since the results are not mutually exclusive. (See the box on page 242 in the text.) It is possible to get a face card on the first draw and a face card on the second draw.

Instead, look to the opposite event. The opposite of getting at least one face card is getting no face cards: getting all numbered cards.

Without replacement, the draws are dependent. Use the Multiplication Rule found in the box on page 229 of the text.

Calculating that chance using conditional probabilities:

$$\frac{36}{52} \times \frac{35}{51} \times \frac{34}{50} \times \frac{33}{49} = \frac{99}{455} \approx 21.76\%.$$

The chance of the original outcome is 100% minus the chance of the opposite.

So, the probability originally sought is 78.24%, answer (E).

36. (G) 99.33%.

The chance of a numbered card on any draw is 69.2%. But do not add the chances of a numbered card on each draw to find the chance of at least one, since the results are not mutually exclusive. (See the box on page 242 in the text.) It is possible to get a numbered card on the first draw and a numbered card on the second draw.

Instead, look to the opposite event. The opposite of getting at least one numbered card is getting no numbered cards: getting all face cards.

Without replacement, the draws are dependent. Use the Multiplication Rule found in the box on page 229 of the text.

Calculating that chance using conditional probabilities:

$$\frac{16}{52} \times \frac{15}{51} \times \frac{14}{50} \times \frac{13}{49} = \frac{4}{595} \approx 0.0067 \approx 0.67\%.$$

The chance of the original outcome is 100% minus the chance of the opposite: 100% − 0.67% = 99.33%.

It is less confusing if the fraction is subtracted: $1 - 4/595 = 591/595 = 99.33\%$.

So, the probability originally sought is 99.33%, answer (E).

37. (D) 27.23%

With replacement, the draws are independent.

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

$n = 4$, $k = 2$, $n - k = 2$, $p = 36/52 = 9/13$, and $1 - p = 4/13$.

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

Here $\frac{4!}{2! \times 2!}(9/13)^2(4/13)^2 \approx 0.2723 = 27.23\%$, answer (D).

(The calculation ended up as $6 \times (81/169)(16/169) = 7776/28{,}561 \approx 0.2723$.)

38. Answer: (E) 69.23%

The expected value for the sample percentage equals the population percentage (page 359).

The box model has 36 ones for the numbered cards and a 16 zeros for all others.

The percentage of ones in the box is $36/52 \times 100\% = 69.23\%$, answer (E).

39. (C) 4.615%

To get the proper counting box needed to model the percentage, put 1 on the tickets representing numbered cards and put 0 on the tickets representing face cards.

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

(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 36/52, and the fraction of tickets with 0 is 16/52.

The formula gives: $(1 - 0)\sqrt{0.6923 \times 0.3077} = \sqrt{0.213} = 0.4615$.

(See the formula on page 298.)

Then the answer is: $\frac{0.4615}{\sqrt{100}} \times 100\% = 4.615\%$, answer (C).

40. Answer: (C) 96.8%

Use the normal curve for the sample percentage. From problems 36 and 37, we know that the expected value for the sample percentage is 69.23% and the standard error for the sample percentage is 4.615%.

The standard units are found from the endpoints of the regions by the formula

$$(\text{endpoint} - \text{EV\%})/(\text{SE\%}) = (60 - 69.23)/4.615 = -2 \quad \text{and}$$
$$(80 - 69.23)4.615 \approx 2.35.$$

The area from -2 to 0 is 47.725%, and the area from 0 to 2.35 is 49.06%.

Add the two half-areas to get about 96.8%, answer (C).

41. (C) 1/32

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

$n = 8$, $k = 7$, $n - k = 1$, $p = 1/2$, and $1 - p = 1/2$.

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

Here $\frac{8!}{7! \times 1!}(1/2)^7(1/2)^1 = 8 \times (1/2)^8 = \frac{8}{256} = 1/32$, answer (C).

42. (C) 1/16

At most means either one head or no heads. The two outcomes are mutually exclusive. Find the chance for each of them and add.

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

$n = 7$, $k = 1$, $n - k = 6$, $p = 1/2$, and $1 - p = 1/2$.

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

Here $\frac{7!}{1! \times 6!}(1/2)^1(1/2)^6 = 7 \times (1/2)^7 = 7/128$.

No heads means tails on all tosses. Since the outcomes are independent, just multiply together seven factors—each the chance of heads on any one toss—to get $(1/2)^7 = 1/128$. (The box on page 232 may be extended to more than two things.)

The chance for a most one head is now $7/128 + 1/128 = 8/128 = 1/16$, answer (C).

43. (B) 7.355%

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

However, most calculators will not return the value of 103 factorial (103!).

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: $102 \times 0.5 = 51$, and $\sqrt{102} \times 0.5 = 5.05$.

The box of the probability histogram for 53 heads runs from 52.5 to 53.5 heads. (See page 317, example 1(a).)

The standard units for the endpoints are:

$(52.5 - 51)/5.05 \approx 0.30$ and $(53.5 - 51)/5.05 \approx 0.50$.

The area under the normal curve between 0.30 and 0.50 $= 38.29\%/2 - 23.58\%/2 = 7.355\%$, answer (B).

44. (a) (B) 1.13%

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

$n = 67$, $k = 42$, $n - k = 25$, $p = 0.5$, and $1 - p = 0.5$.

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

Here $\frac{67!}{42! \times 25!}(0.5)^{42}(0.5)^{25} \approx 0.0113 = 1.13\%$, answer (B).

(b) (A) 1.17%

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: $67 \times 0.5 = 33.5$, and $\sqrt{67} \times 0.5 = 4.0927$.

The box of the probability histogram for 42 heads runs from 41.5 to 42.5 heads.
(See page 317, example 1(a).)

The standard units for the endpoints are:
$(41.5 - 33.5)/4.0927 \approx 1.95$ and $(42.5 - 33.5)/4.0927 \approx 2.20$.

The area under the normal curve between 1.95 and 2.20 = 97.22%/2 − 94.88%/2 = 1.17%, answer (A).

45. (D) 300, 18.26

First find the average and SD of the box: avg=(2+3+4+5+6+7+8+9+10)/9=54/9=6, and $\text{SD} = \sqrt{(16+9+4+1+0+1+4+9+16)/9} = \sqrt{60/9} = \sqrt{20/3} = 2.582$.

From the boxes on pages 289 and 291 and the top box on page 292, the answers are the EV for the sum, which is $50 \times 6 = 300$, and the SE for the sum, which is $\sqrt{50} \times 2.5812 = 18.26$. The answer is (D) 300, 18.26.

46. Answer: (C) 90%

Fifty draws is large enough to use the normal curve. Use the EV for the sum that was found in problem 45. It is 18.26.

Within 30 of the expected number means within $30/18.26 \approx 1.65$ SEs of the expected value. That translates into standard units of ± 1.65. The chance is about 90%, answer (C).

47. Answer: (D) 2, 1

Set up a box model to count the number of heads. Put a 1 for heads and a 0 for tails. (See the box on page 301.) The sum of the draws from this box will model the number of heads.

The average and the SD of this box are both 1/2. The SD is clearly 1/2, because every ticket in the box is exactly 1/2 from the average (1/2). It is not necessary to calculate $\sqrt{\frac{(1/2)^2+(1/2)^2}{2}} = \sqrt{\frac{(1/4)+(1/4)}{2}} = \sqrt{\frac{1/2}{2}} = \sqrt{1/4} = 1/2$.

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

They are: $4 \times 0.5 = 2$, and $\sqrt{4} \times 0.5 = 2 \times 0.5 = 1$.
(See the box on page 289 and the lower box on page 291.)

The number of heads (the sum of the draws) is likely to be around 2, give or take 1.
(See the upper boxes on pages 291 and 292.)

Answer: (D) 2, 1

48. Answer: (C) 30, 5

Set up a box model to count the number of fours. Put a 1 for four and a 0s for each of the other five possible results. (See the box on page 301.) The sum of the draws from this box will model the number of fours.

The average of this box is 1/6.

The SD of the box is found by the shortcut on page 298.

$$\left(\begin{matrix}\text{bigger}\\ \text{number}\end{matrix} - \begin{matrix}\text{smaller}\\ \text{number}\end{matrix}\right) \times \sqrt{\begin{matrix}\text{fraction with}\\ \text{bigger number}\end{matrix} \times \begin{matrix}\text{fraction with}\\ \text{smaller number}\end{matrix}}$$

The SD is $(1-0)\sqrt{(1/6)\times(5/6)} = \sqrt{5/36} = \sqrt{5}/\sqrt{36} = \sqrt{5}/6 \approx 0.37268$.

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

They are: $180 \times 1/6 = 30$, and $\sqrt{180} \times (\sqrt{5}/6) = 5$.
(See the box on page 289 and the lower box on page 291.)

The number of fours (the sum of the draws) is likely to be around 30, give or take 5.
(See the upper boxes on pages 291 and 292.)

Answer: (C) 30, 5

49. (C) $16\frac{2}{3}$, 1

To get the proper counting box needed to model the percentage, put 1 on the ticket for the 4 and put 0 on the tickets representing the other five possible results..

The percent of fours should come out to around the expected value of the sample percentage for the above counting box. (See the boxes on pages 259 and 264.)

The EV for the sample percentage is $(1/6) \times 100\% = 16\frac{2}{3}\%$.

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 1/6, and the fraction of tickets with 0 is 5/6.

The formula gives the SD: $(1-0)\sqrt{(1/6)\times(5/6)} = \sqrt{0.16667 \times 0.83333} = \sqrt{0.13889} \approx 0.37268$.

(See the formula on page 298.)

Then the answer is: $\frac{0.37268}{\sqrt{1386}} \times 100\% \approx 1.00\%$, answer (C).

50. (D) 20.275%

The sample size is sufficiently large that the normal curve could be used to compute chance for the sample percentage.

To get the proper counting box needed to model the percentage, put 1 on the ticket for the 4 and put 0 on the tickets representing the other five possible results..

The percent of fours should come out to around the expected value of the sample percentage for the above counting box. (See the boxes on pages 259 and 264.)

The EV for the sample percentage is $(1/6) \times 100\% = 16\frac{2}{3}\%$.

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 1/6, and the fraction of tickets with 0 is 5/6.

The formula gives the SD: $(1-0)\sqrt{(1/6)\times(5/6)} = \sqrt{0.16667\times 0.83333} = \sqrt{0.13889} \approx 0.37268$. (See the formula on page 298.)

Then the SE for the sample percentage (the give or take) is: $\frac{0.37268}{\sqrt{240}} \times 100\% \approx 2.4056\%$.

Now, 6.68% is the chance of getting a sample percent larger than a certain percent. If 6.68% is the upper tail, 6.68% is also the lower tail, leaving 86.64% of the area in the middle. That sample percent is 1.5 standard units on the normal curve. So it is at 1.5 SDs above the average percent: $16.667\% + 1.5 \times 2.4056\% = 16.667\% + 3.608\% = 20.275\%$, answer (D).

51. (C) 64.03 to 69.47%

    (a) Set up a 0–1 box where 1s are on tickets that model the chance it lands heads and 0s are on tickets that model the chance it lands tails.
    (b) Find the sample percent: $(801/1200)\times 100\% = 66.75\%$.
    (c) Use the bootstrap to estimate the fraction of 1's in the box: 801/1200 or 0.6675.
    (d) Estimate the SD of the box by the shortcut: $\sqrt{0.6675\times 0.3325} = \sqrt{0.22194} = 0.4711$.
    (e) Estimate the SE for the percentage: $\frac{\text{SD of box}}{\sqrt{\text{no. of draws}}} \times 100\% = \frac{0.4711}{\sqrt{1200}} \times 100\% = \frac{47.11\%}{34.64} = 1.36\%$.
    (f) Write 95%-confidence interval as sample percent plus or minus two SEs.
    (g) Answer is: $66.75\% \pm 2.72\%$.

        That's 66.75%−2.72% to 66.75%+2.72%, or 64.03% to 69.47%, answer (C).

52. (C) 2.4475

    The 95%-confidence interval will be: sample average ± 2 SEs for the average.
    (See pg. 381 2nd bullet, and the paragraph on the bottom of pg. 416, continued to the top of pg. 417.)

    The formula for the standard error for the sample average is: (SD of the box)/$\sqrt{\text{no. of draws}}$. (This is the simplified form of the formula from the *Technical note* on page 416.)

    The SD of the box is not known. Because the sample size is large, apply the bootstrap method and use the SD of sample to estimate the SD of the box.

    SE for sample average = 27.364 pounds/$\sqrt{500}$ = 1.223755 pounds.

    The 95%-confidence interval is the sample average plus or minus two SEs.

    Two SEs come out to be 2×1.223755 pounds ≈ 2.4475 pounds, answer (C).

53. Answer: (C) a and d only

    The expected value and the average of the box are exact numbers, so (a) and (d) are true, while (b) and (e) are false.

    The average of the draws will be around the expected value, but it is subject to chance error. Generally, it will not be exactly 2. So (c) is false.

    The average of the box is $(1+2+2+3)/4 = 2$, and the expected value for the average of the draws is equal to the average of the box. (See box on page 410.)

    Answer: (C) a and d only

54. Answer: (C)

    See page 451.

    (A) is contradicted on page 451; (B) the standard error of the sample average estimates how far a typical sample average is from the average of the box; that is not the SD of the box; (D) the average of the error box is assumed to be 0 (page 450); (E) the first sentence is true (page 450), but the second sentence is false (page 451).

    (C) is true, as stated in the box on page 451.

    Answer: (C)

55. (D) $P = 12.5\%$, fair

The null hypothesis is that the coin is fair, modeled by 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 $= 7500 \times (1/2) = 3750$, and SE $= \sqrt{7500} \times (1/2) = 43.3$.

To approximate the chance of getting 3800 or more heads, draw the block for 3800 heads for the probability histogram. It runs from 3799.5 to 3800.5 heads. Then the endpoint for 3800 or more heads (the tail) is 3799.5. The chance for this result or any result more extreme is needed.

Just using 3800 heads for the endpoint is probably close enough. The continuity correction ($\pm 1/2$) is not absolutely needed here because the number of tosses is large. Read the first paragraph on page 317 of the text.

The normal curve may be used to approximate the chances for the sum of the draws because the number of draws is large. See the Central Limit Theorem in the box on page 325.

The standard units are: $\frac{3799.5-3750}{43.3} \approx 1.15$. (Using 3800 gives the same answer.)

The central area for $z = \pm 1.15$ is 74.99%. Two tails are $100\% - 74.99\% = 25.01\%$.
The right tail—which is P—is half of that: 12.5%.

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

Answer: (D) $P = 12.5\%$, fair

56. (B) $P = 3.5\%$, not fair

The null hypothesis is that the coin is fair.

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

Find the chance that a fair coin gets 7 heads or more. That is 7 heads or 8 heads. Add the chances as the results are mutually exclusive.

That chance will be the obesrved signifcance level, P. (See the box on page 480.)

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

$n = 8$, $k = 7$, $n - k = 1$, $p = 1/2$, and $1 - p = 1/2$.

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

Here $\frac{8!}{7! \times 1!} (1/2)^7 (1/2)^1 = 8 \times (1/2)^8 = 8/256$.

Eight heads means heads on all tosses. Since the outcomes are independent, just multiply together eight factors—each the chance of heads on any one toss—to get $(1/2)^8 = 1/256$. (The box on page 232 may be extended to more than two things.)

The chance for 7 or 8 heads is now $8/256 + 1/256 = 9/256 \approx 3.5\%$.

With P less than 5%, the result is statistically significant. Conclude that the 7 heads observed is too many for a fair coin. (Section 26.4.)

Generally, when $P < 5\%$, the null hypothesis is not supported. Answer: (B), $P = 3.5\%$, not fair.

57. (D) 25%, fair

There are 9 degees of freedom: one less than the number of categories.

The expected values are all $400 \times (1/10)$, or 40.

For each category, subtract the expected value from the observed, square the result, and then divide by 40, the expected.

Here is the first calculation: $\dfrac{(34-40)^2}{40} = 3/40 = 0.9$.

Add the ten results. This is χ^2: $0.9+0.025+1.6+0+4.9+0.225+0.4+0.9+0.625+2.025 = 11.6$.

Or, since they are all over 40, add the ten squares of the (observed − expected) and then divide by 40: $\frac{36+1+64+0+196+9+16+36+25+81}{40} = \frac{464}{40} = 11.6$.

From the χ^2-table, the value of P is estimated at 25%. It's between 10% and 30%, closer to 30%.

Answer: (D) 25%, fair

58. (a) True.
 (b) True.
 (c) True.
 (d) False.
 (e) False.
 (f) False.

A: boxes on pages 40 and 41.

B: endpoints for 20 years, for example: 20.000 to 20.999. Use 20 as the left endpoint.

C: box on page 67.

D: False. It will be off by the SE for the average, not the SD of the box.
See the last sentence on page 410 and summary note 2 on page 437.

E: See figure 1 on page 275 and the last comment by *Assistant* on pages 275 to 276.

F: See to top box on page 412.