# Answers to the Sample Final for Fall 2019

1. (a) B   (b) 20%   (c) 70%   (d) 70   (e) 70 (the same)   (f) 87   (g) 28 points

2. (a) The average is 4. So the deviations from average are $-3, -1, 0, 1, 3$. The SD is 2.

   (b) The average is $3 \times 4 + 7 = 19$, the SD is $3 \times 2 = 6$. (of course, you can work these numbers out directly.)

3. (a) 160. Work: 175 is 5 SDs above average at age 18, so the estimate at age 18 is above average by $r \times 5$ SDs. This is $0.8 \times 5 \times 15 = 60$ points.

   (b) 84. See section 10.3.

   (c) 100. The prediction for average at age 18 must be the average score at age 35: 100 points.

   (d) If the score at age 18 is unkown, predict the average score at age 35: 100 points.

4. (a) (i); multiplication rule.

   (b) (viii); chance is 0.

   (c) (iv); addition rule.

   (d) (viii); the events are not mutually exclusive.
   Comment: the chance is $2/52 - 1/52 \times 1/51$. (Needs a formula in footnote 2 to page 244 found on page A-16.)
   Our way around it is to find the chance of the opposite: not getting the king of spades on top and not getting the queen of spades on the bottom. It is restated as queen of spades first card, or any of the other 50 on top and anything but the queen of spades on the bottom. This chance is $1/52 + (50/52)(50/51) = 51/(51 \times 52) + 2500/(51 \times 52) = 2551/2652$. Subtract from 1 to get $101/2652$, which is the same as $2/52 - 1/52 \times 1/51 = 2(51)/(51 \times 52) - 1/(51 \times 52) = 101/2652$. (This was a very difficult problem; it was really about observing that one cannot add the chances of the two given things so (iv) is out; tricky logic will eliminate the other 6 answers, leaving (viii).)

   (e) (vii)

5. (a) False. These events aren't mutually exclusive, so you can't add the chances. (To find the chance, read section 14.4.)
   The actual chance is about 30.6%.

   (b) False: same reason.
   Actually, we know that the chance of getting a tail on the first toss is 50%; obviously the chance of getting at least one head is less than 100%, since you could also get a tail on the second toss after getting a tail on the first toss, and that would be a case of not getting at least one head among the two tosses. That proves that getting at least one head among the two tosses is not a sure thing and therefore does not have a chance of 100%. (To find the chance, read section 14.4.)
   The actual chance is 75%.

6. (a) Chance of no aces $= (5/6)^3 = 58\%$, so the chance of at least one ace $\approx 42\%$. Like de Méré, but with 3 rolls instead of 4.

   (b) $1 - (35/36)^{36} \approx 64\%$.

7. (a) (i) 3/4   (ii) 3/4   (iii) $3/4 \times 3/4 = 9/16$

   (b) 9/16

   (c) $1 - 9/16 = 7/16$

8. (a) The box model is drawing with replacement from the box $\boxed{1}\ \boxed{0}$.

   The average of the box is $\frac{1+0}{2} = 0.5$ and the SD is found by the short cut to be
   $$(1 - 0)\sqrt{(1/2)(1/2)} = (1)\sqrt{1/4} = 1/2 = 0.5.$$

   The sum of a large number of draws will follow the normal curve (Central Limit Theorem).

   For 10,000 tosses, the expected for the number of heads (a sum once we have set up the counting box) is $10,000 \times 0.5 = 5,000$; the SE for the number of heads is $\sqrt{10,000} \times 0.5 = 50$.

   We may optionally correct the numbers of heads under consideration by plus or minus one half since the only possible values are whole numbers. The chance is found to be about 82%.

$$\text{Percent} \approx \text{shaded area} \approx 82\%$$
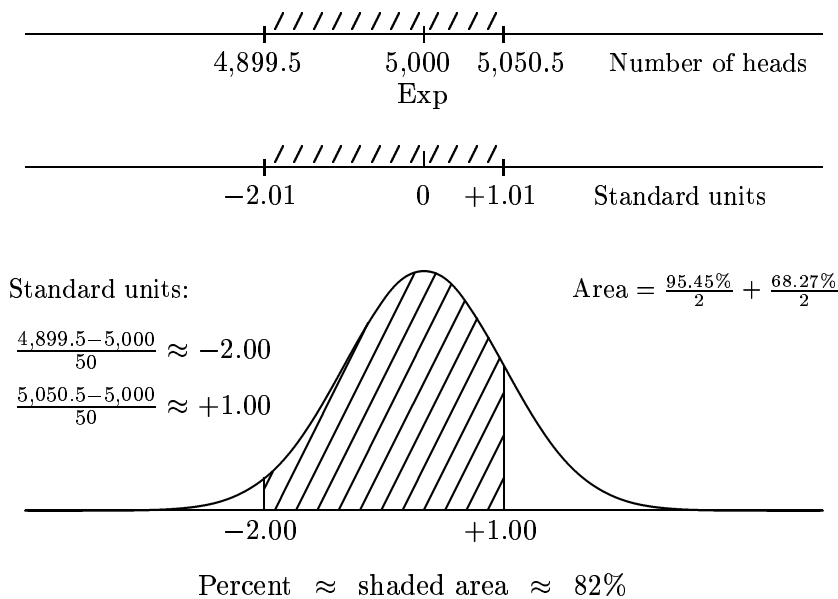
(b) The chance $\approx 2\%$, as the standard unit value was already found.

(c) The chance $\approx 16\%$, as the standard unit value was already found.

9. (a) False, because of chance error. The sample percentage is likely to be close to the population percentage, but not exactly equal. The SE for the percentage says how far off you can expect to be.

   (b) False. It is quite likely that the sample percentage is within 2 SDs of the population percentage, but the chance of that happening is only about 95%. That is not *extremely* likely. About 5% of the time, it will be off by more than 2 SDs. There is always some chance error in the sample percentage, no matter how well the sample is taken. Randomness cannot be eliminated, only tamed. Chance error will always be present and the estimate is not expected to be exactly correct even for large samples, because it is derived from a random process.

10. Sampling without replacement is to be desired. We do not want to sample an individual more than once.

   The calculations for "with replacement" will be simpler, because then independence applies and conditional probabilities are not needed. The normal approximation also assumes independence.

   Generally, if the sample is small relative to the size of the population, it does not make much of a difference whether the sample is taken with or without replacement. The associated probabilities are virtually the same.

11. The SD of the box can be estimated at 10, so the SE for the average of the 100 draws is estimated as 1. If the average of the box is 20, then the average of the draws is 2.7 SEs above its expected value. This isn't plausible.

12. (a) Null hypothesis: the number of correct guesses is like the sum of 1,000 draws from a box with one ticket marked 1 and nine 0's.

   (b) $\sqrt{0.1 \times 0.9}$. The null hypothesis tells you what's in the box. Use it.

   (c) $z \approx (173 - 100)/9.5 \approx 7.7$, and P is tiny.
   (SE sum $= \sqrt{1{,}000} \times$ SD of box $= 31.62 \times \sqrt{0.1 \times 0.9} = 31.62 \times 0.3 \approx 9.5$.)

   (d) Whatever it was, it wasn't chance variation.

   We might explain it like this: an average person would get one in 10 correct, so that is about 100 correct out of the 1,000. This individual comes along and gets considerably more correct and claims that he has ESP. The skeptic would counter by saying that he just got lucky. The statistician with his math can tell us that people do get lucky, but not that lucky. To get that lucky would be less than one chance in a trillion. Maintaining a position that this result was just due to chance error is untenable: it is without a doubt not a reasonable explanation.

13. No. The expected number of positives is 250, and the SE is $\sqrt{500} \times 0.5 \approx 11.18$. The observed number is 2.28 SEs above the expected, when the observed number is corrected to 275.5. (Here a one-sample test is appropriate.) We have assumed that the line is drawn at 5%; using 1% the answer is "Yes."