Answers to the Sample of Final Exam Questions for Spring 2020

1. 12%.

2. (a) average: 12.
   SD: 4.
   (b) 11,13,13,14,17 are within 1.5 SDs of average.
   All six members of the list are within 2.5 SDs of average.

3. 1.58%. With interpolation, the answer comes out to about 1.515%.

4. 0.3.

5. (a) 160.
   (b) 148.
   (c) The regression effect.

   In each case there are just a few very-high-scoring people. Since the correlation is not perfect, the score at the other age regresses toward 100.

   The actual number of 175's at age 18 is extremely small; so they barely make a dent in the total of people scoring 160 at age 35. The majority of 160's at age 35 can, in general, have lower scores at age 18; and—still—these few people who scored 175 at age 18 will be opposite the overall trend and have higher scores at age 18.

   Simply stated, the answers to (a) and (b) are for the overall or average of each group. There still could be some small portion in each case with a higher score at the age being predicted.

   There is no contradiction. Don't get fooled by the regression fallacy.

6. (a) $\left(\frac{12}{13}\right)^{40} \approx \frac{1.46977 \times 10^{43}}{3.61189 \times 10^{44}} \approx 0.04069 \approx 4.07\%$.
   (b) $100\% - 4.07\% = 95.93\%$.
   (c) 95.93%.
   (d) $\left(\frac{25}{26}\right)^{40} \approx 0.208289 \approx 20.83\%$.
   (e) $13.97\% + 14.44\% = 28.41\%$. Use the binomial formala twice and, since they are mutually exclusive, add the results.

7. (a) True. The draws are independent, so just cube 1/4.
   (b) False; it's 57.8%.

8. (a) 480, give or take 20 or so.
   (b) $16\frac{2}{3}\%$, give or take 0.694% or so.

9. $94.1\% \pm 0.6\%$ or 93.5% to 94.7%.

10. (a) False: the SE is indeed estimated from the sample data, as in section 23.1; but the SE for the sample average is 0.005. The value 0.152 was the SD, not the SE for the sample averge.

    (b) False. There is no such thing as a 95%-confidence interval for the *sample average*; you *know* the sample average. It's the population average that you have to worry about (refer to pages 385–386 in chapter 21, section 3).

    (c) True. The 95%-confidence interval is the sample average plus or minus 2 SEs.

    (d) False, This confuses the SD with the SE. And it's ridiculous in the first place, because a student must have a whole number of cats (0, or 1, or 2, and so forth). The range 0.216 to 0.824 is impossible for any particular student, let alone 95% of them; although this range is fine for the average of all the students.

    (e) False. For instance, if number of cats followed the normal curve, there'd be some students with a negative number of cats since from symmetry we'd have roughly the same area in the histogram for the blocks representing 2 cats and -1 cat; we're not ready for that.

    (f) True. See section 23.1, page 411; and section 23.2, pages 418–419. Even though number of cats does not follow the normal curve, you can still use the normal curve to approximate the probability histogram for the sample average.

11. No. The expected number of heads is 500,000, and the SE is $\sqrt{1{,}000{,}000} \times 0.5 = 500$. The observed number is 4.03 SEs above the expected.

    P is about 0.00255 of 1%, very small and highly significant.