Answers to the Sample Final for Spring 2018

1. (a) around 85% (b) 40–50 cigarettes (c) 0–10 cigarettes (d) 30%

2. (a) 97th percentile (b) 620 points

3. (a) -0.80 (b) 0.3 (c) 1.00

   Comment: In (c), all the points lie on the line $y = 2x$, so there is no need to do any arithmetic.

4. (a) 71 inches—regression method.

   (b) 2.3 inches. The scatter diagram is homoscedastic, so the sons' heights are off the regression line by similar amounts, for any father's height. The amount off is the r.m.s. error of the line, which is $\sqrt{1-0.5^2} \times 2.7 \approx 2.3$ inches.

   (c) About 68%, using the normal curve for a value within one r.m.s. error or less (meaning that the standard units will be between -1 and $+1$).

5. (a) Provided they are independent, the chance of both is 5%, obtained by the multiplication rule.

   (b) Provided they are mutually exclusive, the chance of at least one is 60%, obtained by the addition rule.

6. (a) $36/52 \times 35/51 \times 34/50 \times 33/49 \times 32/48 = 66/455 \approx .145$, about $14\frac{1}{2}\%$.

   (b) $16/52 \times 15/51 \times 14/50 \times 13/49 \times 12/48 = 1/595 \approx 0.00168$, about one-sixth of one percent.

   (c) $389/455 \approx .855$, about $85\frac{1}{2}\%$; it's the opposite of the result in part (a).

   Comment. "all not numbered" and "not all numbered" are two different propositions, and part (b) asked for "all not numbered" in different words.

7. The expected value is 400, the observed value is 431, the chance error is 31, the standard error is 20.

   First find the average and the SD of the box: 4 and 2. Then multiply the average of the box by the number of draws to get the expected value, and multiply the SD of the box by the square root of the number of draws to get the standard error.

   The observed value is just what sum of the draws turned out to be, and the chance error is just the observed value minus the expected value.

8. (a) 60, give or take 6.794 or so.

   (b) 23.077%, give or take 2.613% or so.

   (c) 23.077%, because the normal approximation applies and the median of the normal curve is at the average. From page 326, we know that the expected value is at center of the normal curve: 0 in standard units.

9. No, the normal curve applies because it is a simple random sample of large enough size.

   The 95%-confidence interval is 1.86 ± 0.06.

10. The SE for the average is 1. The answer to (a) is almost 100%. The answer to (b) is 68%. Don't confuse the SE for the average of the draws with the SD of the box.

11. (a) Tossing the coin is like drawing at random with replacement 10,000 times from a 0–1 box, with 0 = tails and 1 = heads. The fraction of 1's in the box is unknown.
    Null hypothesis: this fraction equals 1/2. Alternative: the fraction is bigger than 1/2.
    The number of heads is like the sum of the draws.

    (b) $z = 3.34$, $P \approx 4/10{,}000 = 0.04$ of 1%.

    (c) There are too many heads to explain as chance variation.