Possible responses to the 2015 AP Statistics Free Resposne questions, Draft #1.

You can access the questions here at AP Central.

Note: I construct these as a service for both students and teachers to start discussions. There is nothing "official" about these solutions. I certainly can't even guarantee that they are correct. They probably have typos and errors. If you catch some, let me know! But if they generate discussion and help others, then I've succeeded.

1. Accounting Salaries, two companies.

Centers: There's not much difference in median yearly salaries between the two companies (each about $52,000). Variability: Salaries at company A are more variable than a company B; the IQR at A is about $55,000 - $47,000 = $8,000. At B, the IQR about $53,000-47,000 = $6000. Both distributions are roughly symmetric. There are no outlier salaries at company B, but at company A, There are two outlying high salaries: two entry-level accountants made over $70,000 that year.

b)

i) Because these box plots display salaries of accountant five years after being hired in 2009 at $36,000 a year, they show what might happen to me at each company five years from now. Suppose I am a great accountant, and expect to be one of the very best. Then I might prefer corporation A over B because I might want to aim for a very high salary (over $60,000, say). The chances of that happening are better at corporation A than at corporation B, because salaries are more variable in A.

ii) I might prefer a job at corporation B over A if I am a poor accountant, and I want greater reassurance that I attain a baseline salary of at least $40,000 a year. Notice that because of the lower spread for corporation B, all salaries are over $40,000, and this is not the case at corporation A: I might end up with a lower salary if I am a weak performer.

2. Confidence interval for customer discounts.

a) No it does not; the interval is (0.09, 0.21), and 0.20 is inside the interval. O.20 It's a plausible value for p.

b) No, it does not. According to the confidence interval, 0.20 is one of many plausible values for p. A lack of evidence that p is not equal to 0.2 does not necessarily imply that the interval provides convincing evidence that p = 0.2

b) When we replace n with 4n in the formula for the standard deviation of the sample proportion $z^*\sqrt{\frac{p(1-p)}{4n}} = \frac{1}{2}z^*\sqrt{\frac{p(1-p)}{n}}$ So the margin of error = (1/2)(0.06) = 0.03.

c) This new margin or error produced a confidence interval of (0.12, 0.18), and 0.2 is not in this interval. This provides convincing evidence that p is not equal to 0.02, and the program is not working as intended.

3. ATMS.

a. $0.21 + 0.40 + 0.24 = 0.85$.

b. $0(0.15)+(1)(0.21)+2(0.40)+3(0.24) = 1.73$.

c. $0.24/ 0.85 = .2824$

d. The expected value will be greater. The new expected value is $1.73/0.85 = 2.0353$

4. Aspirin and colon cancer.

This requires a two-proportion z test for the difference in two proportions. Let

p_1 = the proportion of all adults like our volunteers who develop colon cancer when taking a low dose aspirin

p_2 = the proportion of all adults like our volunteers who develop colon cancer when taking a placebo.

We test: $H_0 : p_1 = p_2 \quad H_A : p_1 < p_2$

Conditions:

The problem states that the 1000 volunteers were assigned at random to one of the two treatment groups. So the random assignment condition is met.

Group	Number successes	Number failures	sample proportion
aspirin	15	485	.03
placebo	26	474	.052

Large Sample Size: Is the sampling distribution of $\hat{p}_1 - \hat{p}_2$ is approximately normal? We see above that the number of successes and failures in each group is greater than10. Or, using the pooled sample proportion, we see that $(1000)\frac{15+26}{500+500} = 41$ and $(1000)\frac{485+474}{500+500} = 959$ are both greater than 10. Using either check, we feel confident that the sampling distribution of $\hat{p}_1 - \hat{p}_2$ is approximately normal.

Mechanics: Using the test statistic, $z = \dfrac{\frac{15}{500} - \frac{24}{500}}{\sqrt{\frac{41}{1000}\left(\frac{1}{500} + \frac{1}{500}\right)}} = -1.75 \; p-value \approx 0.03969$.

P value is < 0.05, so we reject H_0. We have convincing evidence at the $\alpha = .05$ level that taking a low-dose aspirin reduces the risk of colon cancer among individuals like our volunteers in this study.

5. Heights and arm-spans

a) There's a moderate, positive, linear association between heights and arm spans. Taller kids tend to have longer arm spans.

b)

i) Graph 2 is better. The line y = x separates the students into the three groups mentioned. Those above the line are short rectangles. Those below the line are short rectangles. Students on the line have equal arm spans and heights. The regression line from Graph 1 makes no reference to whether heights and arm spans are similar, so this line won't work.

ii) 3 square, 4 tall rectangles, 5 short rectangles

c) predicted arm span = 11.74 + 0.8427(61) = about 62.0467 inches.

6. Totrillas.

a. For all tortillas *on that day,* the second method will not work – we will get a sample of 200 tortillas coming from a population with a mean of 5.9 or a sample of 200 tortillas coming from a population with a mean of 6.1. But the mean of the population appears to be at/near 6.0.

(PS: It's important to know that estimates of the mean diameter, in repeated sampling, will produce an unbiased sampling distribution of the population mean. But this is not the question posed. No statistic is mentioned in this problem, and we're not being asked about whether the sample mean is unbiased)

b. This sample is bimodal. In order to get two clumps of data in our sample, we need to be sampling from both groups. This will happen from Method 1, not Method 2.

c. Method 2 has less variability in tortilla diameters than Method 1. In method 1, we will typically get a bimodal sample. The overall distribution will be centered at 6.0 inches, but one mode will be centered at 5.9 inches, and another will be centered at 6.1 inches. We'll get lots of diameters far away from the sample mean. With method 2, we will get either a unimodal sample centered at 5.9 inches, or a unimodal sample centered at 6.1 inches… but not *both* in the same sample. For this reason, diameters in samples from Method 2 will fall closer to their sample mean than will the diameters from Method 1.

d. $\mu_{\bar{X}} = \mu_X = 6.0$, because random samples were taken. Because sample sizes are less than 10% of the population (200 < (0.1)200,000), we can say that the standard deviation is approximately $\sigma_{\bar{X}} = \frac{0.11}{\sqrt{200}} \approx 0.00778$.
Because we are taking a ***simple random sample*** of size of 200 > 30, the central limit theorem allows us to assume that the sampling distribution of $\bar{X}$ is approximately normal.

e. Method 2 will have a sampling distribution with more variability. With Method 1, we will be getting 365 sample means, each with an average near 6.0. This sampling distribution is roughly normal with $\sigma_{\bar{X}} = \frac{\sigma_X}{\sqrt{200}}$. With method 2, things are more complicated; we're *not* taking simple random samples, so we can't use $\sigma_{\bar{X}} = \frac{\sigma_X}{\sqrt{200}}$ to compute the standard deviation. In method 2, we will also be getting 365 sample means. But because half of the sample means will come from a population centered at 5.9 inches, and half will be coming from a population centered at 6.1 inches, we will see more variation in the sample means - even though in the long run the sampling distribution will balance out with a mean of 6.0. This sampling distribution from Method 2 will be bimodal, and the sampling distribution from Method 1 will be unimodal and symmetric with a mean at 6.0. This makes the SD of the sampling distribution larger for Method 2.

f. Method 1 will. Both sampling distributions have a mean of 6.0 inches. But as we said in part e., the standard deviation of the sampling distribution of $\bar{X}$ will be lower than with Method 2. So an individual sample mean has a higher chance of being near 6.0 with method 1 than with method 2.