Free response questions, 2014, second draft!

Note:
Some notes:

- Please make critiques, suggest improvements, and ask questions. This is a just one AP stats teacher's initial attempt at solving these. I, as you, want to learn from this process.
- I simply 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. But if they generate discussion and help others, than I've succeeded. They are simply one statistics teacher's attempt at the problems. I do this as a way to invite dialogue about the questions.
- You can go here to access the problems. This is a public site. http://media.collegeboard.com/digitalServices/pdf/ap/ap14_frq_statistics.pdf

1.
(a)

Proportion of those on campus who participated in at least one activity: $\frac{7+17}{33} \approx 0.727$

Proportion of those off campus who participated in at least one activity: $\frac{25+12}{67} \approx 0.552$

(b)
In this sample, students who reported living off campus were more likely to report participating in no extracurricular activities than those living on campus.

In this sample, students who reported living on campus were more likely to report participating in exactly one extracurricular activity than those who reported living off campus.

In this sample, students who reported living on campus were a bit more likely to report participating in two or more extracurricular activities than those reporting living off campus.

(c)
Because the p value of 0.23 is above any reasonable significance level (like 0.10), we do not have convincing evidence of an association between residential status and l the level of extracurricular participation for students at this university.

2.

(a)

P(select woman, then another woman, then another woman)

$=\left(\frac{3}{9}\right)\left(\frac{2}{8}\right)\left(\frac{1}{7}\right)\approx 0.0119$

(b)

Yes, there is reason to doubt the manager's claim. The chance that a truly random draw results in all three women being picked is rare: less than 2% of all draws of three students would produce a result as unusual as this.

(c)

This simulation does not work, because the probability of selecting a woman (or a man) needs to change after each draw. We are selecting people *without replacement* from a small group (only 9 people) in the real lottery, but by using a six-sided die to simulate the drawing, this simulation's plan assumes that we are drawing *with replacement.*

3.

(a)

Let X = # students absent at this school on a randomly selected day. *X* is approximately normal, with $\mu_X = 120, \quad \sigma_X = 10.5$.

$P(X > 140) = P(Z > \frac{140-120}{10.5}) \approx P(Z > 1.9048) \approx .0284$

(b)

This question requires that we consider the sampling distribution of $\bar{X}$, the sample mean, with *n = 3.* This distribution is approximately normal (because we took a random sample of 3 from an approximately normal population) In this case $\mu_{\bar{X}} = 140$ and $\sigma_{\bar{X}} = \frac{10.5}{\sqrt{3}}$.

Method I: Both sampling plans create estimates with the same mean: 120 students. But since the standard deviation of the sample mean is lower than the standard deviation of individuals from the population, we would expect that for part (b), we would be *less* likely to lose funding, because we are less likely to get a sample mean of 140 from three days than we are to get a result of 140 or higher from a single day.

Method II: For part (b), $P(\bar{X} > 140) = P(Z > \frac{140-120}{10.5/\sqrt{3}}) \approx P(Z > 3.299) \approx .000485$.

This is a lower probability from part (a), so we are *less* likely.

(c)

P(no T, W, or Th in three weeks)$=\left(\frac{2}{5}\right)^3 = .064$

4.
(a)
Samples of incomes tend to be skewed to the right. They also tend to have a few unusually high incomes.

Sample means tend to be pulled higher by right skew rightness or unusually high "outlier" incomes. Sample medians, however, are more resistant to outliers, and would not be pulled up by the skew rightness or outliers. This produces a statistical advantage because medians are more likely to produce values that are closer to an income near the middle of the distribution of incomes. Means might potentially produce incomes that are far higher than what most people at the university make.

(b)
Estimates are strong if they show *low bias and low variability.*

Assessing of bias: Method 1 is likely to result in bias, because those who respond to the sample do so voluntarily. It's reasonable to predict that those who choose to report their income are more satisfied with their income than those who refuse to share. I would predict that Method 1 would produce an estimate that is too high. Method 2, however, is done via random sampling, with a good plan for following up with each person selected. As long as responses are kept confidential, I would predict that Method 2 has much less bias than Method 1.

Assessing variability: A sample of 100 (as in method 2) would produce an estimate that has more variability than a sample of size 600 (as is predicted with method 1). However, I would be skeptical of the unsupported claim that the people proposing method 1 will get 600 responses for their online survey. Despite the *suspicion* of a larger sample size, their claim is not convincing. **I would choose method 2 over method 1 despite the possibility of lower variability because of the clear bias.**

5.
For this problem, we performed a study on a sample of *eight randomly selected car models*. We recorded:

W = amount paid by a randomly selected woman for this model
M = amount paid by a randomly selected man for this model

Since we record two variables on each model that are clearly associated, this is a *matched pairs design,* and thus requires a *one sample t test on paired differences.* The parameter of interest is

μ_{W-M} = mean amount that a randomly selected woman pays for a randomly selected car model at this dealership *above* what a randomly selected male pays for an identical car form the same model.

We are testing, at a 5% significance level:

$H_0 : \mu_{W-M} = 0$
$H_A : \mu_{W-M} > 0$

Conditions for a one-sample t test on paired differences:

Normality: The dot plot of differences appears to be roughly symmetric, with no outliers. It is reasonable to conclude that the underlying sampling distribution of differences is approximately normal).

Randomness: a random sample of ***eight car models*** was chosen from tax records of all car purchases in the county. Within that, we randomly selected one woman and one man who purchased a similarly equipped car of that model.

Independence: Since each difference was computed from a randomly selected woman and a randomly selected man for a randomly selected model, I feel confident that individual differences within the sample are independent. I also believe that 8 < (1/10) (number of car models in the county).

Mechanics: Using a *t-test* for paired differences with 7 degrees of freedom:
Sample mean: $585.00
Sample SD: $530.71
N = 8 ***t* statistic: 3.1178 p-value: 0.008**

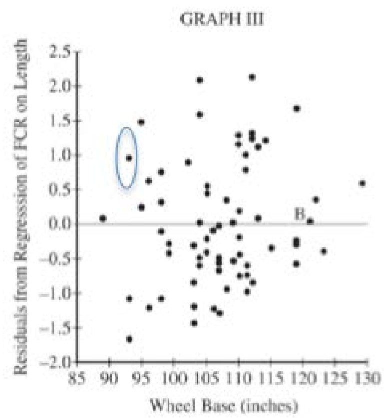
Conclusion: Our p-value 0.008 < 0.05, our significance (alpha) level. Therefore we have convincing evidence that women tend to pay more, on average, than men for car models in this county.

6.

(a)
$\widehat{FCR} = -1.595789 + 0.0372614(175) \approx 4.925.$

Residual = $FCR - \widehat{FCR} \approx 5.880 - 4.925 \approx 0.955$

This car consumes about 0.955 gallons per 100 miles more than what the linear model would predict for cars that are as long as this (175 inches).



(b)
i. The residual for car A, based on the regression equation predicting FCR from *length*, is +0.955. This car has a **wheel base** of about 93 inches. This car is circled to the right.

ii. For car B, the prediction of FCR based on car length is nearly perfect: The actual FCR for car B is just a bit higher than the predicted FCR.

(c)
In graph II, we are looking at the association between *engine size* and the *residuals from the regression model* to predict FCR from *car length.* We can see a moderate, positive linear association between these variables. This means that the residuals tend to be negative when engine sizes are shorter(meaning that the regression model with car length tends to *over-predict* FCR for these cars). When engine sizes are larger, the residuals are positive (meaning that the regression model with car length tends to *under-predict* FCR).

In graph III we look at the association between *wheel base* and the *residuals from the regression model to predict FCR from car length.* In this case, we see no clear pattern between these variables - this means that there is no systematic tendency for predictions from the regression equation to change when wheel base changes.

(d)

I recommend that Jamal choose engine size over wheel base size. Because the existing linear model's predictions systematically vary when engine size increases, we know that including engine size in a regression model will explain additional variability in FCR that is not currently explained by the existing regression model.

However, because there is no systematic pattern between wheel base and the residuals, we would not expect wheel base to explain much additional variability in FCR.