Kyle Beitz

STAT 4155 Homework 1

Part A

1. SMOG 4.5 (a,c,e)

a)

Depending on how the researcher threw the grains, the distribution of grains may not fully encapsulate the whole population, which could result in the possible unequal representation of the total population and skewed results.

c)

With this method, teachers may refuse to answer if their students performed poorly. This makes results of the data biased towards those who did well, while leaving out the poor-performing population. Also, this method does not encapsulate the whole population, biasing results to only the teachers (and their students) on that specific messaging platform.

e)

This method lends itself to responses from only those women whose opinions on their marriage lie on the extremes. This would result in an unequal representation of the population because those who do not write in are not represented.

2. SMOG 4.9

a.

We do not trust the results of the survey. One reason surveys are difficult to trust is because responses can be heavily biased. A lot of the time, people are most motivated to fill out surveys when they are unhappy with the topic in question. So, in this example, many respondents of this survey may be partaking only because they are frustrated with the offensiveness of television. Those who are happy with television may not feel the need to respond, as they don't feel it needs to change. Collecting responses in this voluntary manner creates bias that may not encapsulate the total population.

b.

We think quite a bit less than 92% of the U.S. television watching public would say that "today's shows are more offensive than ever" because that would mean that around 304,000,000 out of 331,000,000 would have to say that TV is more offensive than ever. That is a very large proportion and we believe the distribution would be more equally distributed at around 50%.

3. SMOG 4.14

Estimate p:

$\hat{p} = \frac{Y}{n}$ where Y = number of successes

$$\hat{p} = \frac{25}{30} = 0.83$$

Place a bound on the error of estimation:

$$\hat{p} \pm 2\sqrt{(1-\frac{n}{N})\frac{\hat{p}(1-\hat{p})}{n-1}}$$

$$\hat{p} \pm 2\sqrt{(1-\frac{30}{300})\frac{.83(1-.83)}{29}}$$

$$\hat{p} \pm .132$$

4. SMOG 4.15

If we have a population size of 300 and the number of successful events is 25 with a 95% confidence level, we would need a sample size of 132 to get a margin of error of 0.05.

5. SMOG 4.20

Estimate p:

$\hat{p} = \frac{Y}{n}$where Y = number of successes

$$\hat{p} = \frac{430}{1000} = 0.43$$

Place a bound on the error of estimation:

$$\hat{p} \pm 2\sqrt{(1-\frac{n}{N})\frac{\hat{p}(1-\hat{p})}{n-1}}$$

$$\hat{p} \pm 2\sqrt{(1-\frac{1000}{99000})\frac{.43(1-.43)}{999}}$$

$$\hat{p} \pm .031$$

6. SMOG 4.21

If we have a population size of 99000 and the number of successful events is 430 with a 95% confidence level, we would need a sample size of 1800 to get a margin of error of 0.02.

7. SMOG 2.2

Within the definition of sampling units, we are unsure of what nonoverlapping means exactly.

Part B

(1)

With a mean of 250, SD of 150, and population size of 1005, in order to achieve a 95% confidence interval with a relative margin of error no larger than 15% you will need a sample size of 60 (using fpc to be safe).

(2)

Using RANDBETWEEN function we drew a random number of 157.

(3)
We would recommend using the fpc because the margin of error is lower.

(4)
Yes. We got a relative margin of error of 8.7% using the fpc and our actual data sample size of 157, which is less than the desired 15%.

(5)
With a 90% confidence interval, we would need a smaller sample size to achieve a relative margin of error of 15%, because as confidence level decreases, sample size decreases to keep the relative margin of error at 15%.

(6)
Given a 95% confidence interval and a targeted relative margin of error of 5%, this would require a larger sample size. Since margin of error and sample size have an inverse effect, it only makes sense to increase sample size to achieve a smaller margin of error. Intuitively, as your sample size gets larger, you get more information about the population (your sample is closer to the the true population); consequently, your relative margin of error will decrease.