Homework due Oct. 13

Assigned exercises: Ch.5, OpenIntro Stats, ex. 5.2, 5.4, 5.5, 5.8, 5.10, 5.11, 5.12, 5.13. (8 exercises)
Graded exercises: 5.2 (a-c), 5.4 (a-c), 5.5 (a-c), 5.8, 5.13.
Total (maximum) possible points = 20.
3 pt for each of 5 graded problems, plus 5 for completion of the rest.

Exercises from Ch.5, OpenIntro Stats

(5.2) The question asks to determine whether the parameter of interest is a mean or proportion.

- (a) Proportion. Because the response to the survey question (whether or not a respondent worries about "whatever") will be a categorical variable.
- (b) Mean. Because the survey is interested in the percentage increase in revenue for each respondent. This will lead to a quantitative variables in their dataset.
- (c) Proportion. Reason is very similar to (a), since survey respondents are asked whether or not they do something, which leads to a categorical variable.

  (d) and (e) are not graded, but here are the answers:
- (d) Proportion. Like (a) and (c).
- (e) Mean. Respondents are asked for quantitative information.

(5.4)
- (a) The population is all adults in the US.
- (b) The parameter of interest is the proportion of all US adults who would not be able to cover a $400 unexpected expense without borrowing or going into debt.
- (c) The point estimate is the sampled value, which is $\hat{p} = \frac{322}{765} = 0.421$.

  (d)-(g) are not graded, but here are the answers:
- (d) The standard error.
- (e) $SE = \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} = \sqrt{\frac{0.421(1-0.421)}{765}} = 0.0179$
- (f) Yes, she should be surprised, because a value of 0.5 is more than 4 standard errors away from the sampled estimate.
- (g) Recomputing the standard error with 0.4 in place of 0.421 gives: $SE = 0.0177$, which is very close to the previous value.

(5.5)
- (a) This is the sampling distribution (of sample proportions).
- (b) Provided the conditions are met, the CLT guarantees the sampling distribution will be symmetric (in fact, normal). In this exercise, the sample is random and if its size is $< 10\%$ of the population (which has not been clarified), we may assume it is independent. Whether its size is large enough depends on the parameter value. Here we are told this value is at least 5%. That means: $n \cdot p = 800 \cdot 0.05 > 10$,

and also $n \cdot (1-p) = 800 \cdot 0.95 > 10$. On the other end, the largest value of the parameter is 30%, and we have: $n \cdot p = 800 \cdot 0.3 = 240 > 10$ and $n \cdot (1-p) = 800 \cdot 0.7 = 560 > 10$. Thus, the sampling distribution should be symmetric.

(c) From the CLT, the formula for the standard deviation is

$$\sqrt{\frac{p(1-p)}{n}} = \sqrt{\frac{0.08(0.92)}{800}} = 0.0096$$

(d)-(e) are not graded, but here are the answers:

(d) The standard error.

(e) The standard error will be higher, which means there will be higher variability in sampled estimates.

(5.8)

- To construct a confidence interval, we must start by checking the conditions:

  This exercise doesn't give enough information, but it does say a normal distribution may be used. So we proceed with the assumption the conditions are met.

- Computations:

  Confidence interval $= \hat{p} \pm z^* \cdot ME$
  Sampled proportion is $\hat{p} = 0.52$. Margin of Error $= z^* \cdot SE$

  For 99% confidence, we do an inverse lookup for 0.995 and get: $z* = 2.58$.

  Margin of Error $= 2.58 \cdot 0.024 = 0.0619$

  Therefore, the confidence interval is: $0.52 \pm 0.0619 = [0.458, 0.582]$

- Conclusion: We are 99% confident that between 45.8% and 58.2% of all US adult Twitter users get at least some of their news on Twitter.

(5.13) (a) The sample is random and $< 10\%$ of the visitors to the website, according to the given information. So it is reasonable to assume it satisfies the independence condition. The sample size is large enough, since it has 64 successes, and $752 - 64$ failures, both being more than 10. Thus the conditions are met.

(b) We have $\hat{p} = 64/752 = 0.085$, and $SE = \sqrt{\frac{0.085(1-0.085)}{752}} = 0.0102$

(c) For 90% confidence, we do an inverse lookup for 0.95 and get: $z* = 1.645$.

Margin of Error $= 1.645 \cdot 0.0102 = 0.0168$

Therefore, the confidence interval is: $0.085 \pm 0.0168 = [0.0682, 0.1018]$

Conclusion: We are 90% confident that between 6.8% and 10.1% of new visitors to the website will register.