Homework 10: 4/17 - 4/19

STA 336 Due 11:59pm Thursday, April 27 Name:

Instructions: Write-up complete solutions to the following problems and submit answers on Gradescope. Your solutions should be neatly-written, show all work and computations, include figures or graphs where appropriate, and include some written explanation of your method or process (enough that I can understand your reasoning without having to guess or make assumptions). A rubric for homework problems appears on the final page of this assignment.

- Unless otherwise noted, problem numbers are taken from the 4th edition of DeGroot and Schervish's *Probability and Statistics*.

Monday 4/17

Additional Problems

AP1. Suppose $X \sim \text{Bin}(20, p)$ with p unknown. In this problem, we will construct a one-sided 95% confidence interval for p, by inverting an equivalent hypothesis test. For each $0 < p_0 < 1$, let δ_{p_0} denote the test of the hypotheses

$$H_0 : p \leq p_0 \qquad H_1 : p > p_0$$

using the criteria "Reject H_0 if $X \geq c_{p_0}$", where c_{p_0} is chosen to be the smallest integer c so that $P(X > c|p = p_0) \leq 0.05$; that is, c_{p_0} is the 0.95 quantile of $\text{Bin}(20, p_0)$.

(a) Use the `pbinom` function in R to compute $P(X > c|p = p_0)$ for each $c \in \{0, 1, \ldots, 20\}$, when $p_0 = 0.5$. Then identify the smallest value of c for which $P(X > c|p = 0.5) \leq 0.05$. Confirm that this value of c is the 0.95 quantile of $\text{Bin}(20, 0.5)$ using `qbinom`.

(b) Run the following code in R to create a vector of equally spaced values of p_0 between 0 and 1:

```
p0 <- seq(0, 1, by = 0.001)
```

Then combine the `qbinom` function with your vector `p0` to create a vector called `cp0` of the values c_{p_0} for your collect of p_0 values. Create a plot in R showing the graph of c_{p_0} as a function of p_0.

(c) Suppose $X = 13$ is observed. Does the procedure $\delta_{0.5}$ reject H_0? Does the procedure $\delta_{0.4}$ reject H_0?

(d) Suppose $X = x$ is observed. Let $\omega(x)$ denote the set

$$\omega(x) = \{p_0 \in (0, 1) : x \leq c_{p_0}\}$$

Explain why $\omega(x)$ consists of all values of p_0 for which the procedure δ_{p_0} does not reject H_0 when $X = x$ is observed.

(e) Use your graph from part (b) to explain why $\omega(x)$ is an **interval** with right endpoint of 1; That is, explain why if $p_0 \in \omega(x)$ and if $p_0' > p_0$, then $p_0' \in \omega(x)$.

(f) Use Theorem 9.1.13 to explain why $\omega(X)$ is a 95% confidence interval for p.

(g) Suppose $X = 13$ is observed. Create the observed 95% confidence interval for p using the vector `cp0` you created in part (b). Your answer should take the form $(p_0, 1)$ for some p_0 in the list your created in part (b). *Hint: for a vector `v` and number `t`, the function `min(which(v >= t))` returns the position of the first element in `v` which is greater than or equal to `t`.*

AP2. **(Optional; will not be collected)** Complete Exercise 9.1.17.

Wednesday 4/19

Section 9.5: 2, 6, 17

Friday 4/21

Section 9.6: 4, 11 (For part (b), use Welch's approximation on page 593 - 594)
Section 9.7: 1, 15

Homework 10: 4/17 - 4/19

STA 336 Due 11:59pm Thursday, April 27 Name:

General Rubric

Points	Criteria
5	The solution is correct **and** well-written. The author leaves no doubt as to why the solution is valid.
4.5	The solution is well-written, and is correct except for some minor arithmetic or calculation mistake.
4	The solution is technically correct, but author has omitted some key justification for why the solution is valid. Alternatively, the solution is well-written, but is missing a small, but essential component.
3	The solution is well-written, but either overlooks a significant component of the problem or makes a significant mistake. Alternatively, in a multi-part problem, a majority of the solutions are correct and well-written, but one part is missing or is significantly incorrect
2	The solution is either correct but not adequately written, or it is adequately written but overlooks a significant component of the problem or makes a significant mistake.
1	The solution is rudimentary, but contains some relevant ideas. Alternatively, the solution briefly indicates the correct answer, but provides no further justification
0	Either the solution is missing entirely, or the author makes no non-trivial progress toward a solution (i.e. just writes the statement of the problem and/or restates given information)
Notes:	For problems with multiple parts, the score represents a holistic review of the entire problem. Additionally, half-points may be used if the solution falls between two point values above.