Name: ____________________________

Directions: **Work only on this sheet** (on both sides, if needed); do not turn in any supplementary sheets of paper. There is actually plenty of room for your answers, as long as you organize yourself BEFORE starting writing. In order to get full credit, SHOW YOUR WORK.

1. (10) Look at the rogue network router example in Sec. 2.11.1 (p.18) of the current PLN. The sample mean for the wireless data was 11.520. From this find an approximate 95% confidence interval for the corresponding population mean. Write your answer as a numerical expression, e.g. 8 ± 1.68.

2. (15) On p.8 of our current PLN, the mean wait μ for a bus is estimated via simulation. An approximate 95% confidence interval for μ is printed out. Write R code that prints out the standard error of $\overline{W}$.

3. (15) Write an R function with prototype

```
propci <- function(x,alpha)
```

that will compute and print out a confidence interval of approximate level $100(1-\alpha)\%$ for a population proportion p for a certain trait. Here **x** is a vector of 1s and 0s, 1 meaning the trait is present and 0 meaning not.

4. Suppose $fX(t) = ct^{-2}$, for $c < t < \infty$. We wish to test $H_0 : c = 1$ vs. $H_A : c > 1$, *using just a single observation X*. We will use a significance level α of 0.10.

(a) (10) In order to achieve $\alpha = 0.10$, our rule should be to reject H_0 if X ________________________.

(b) (10) If in actuality c = 5, what is the probability that we make the right decision, i.e. that we reject H_0?

5. (10) Consider the lottery example. Suppose someone asks whether our estimator $\widehat{c}$ is biased. Which answer is correct? (i) Yes, it's biased upward; (ii) yes, it's biased downward; (iii) no, it is not biased; (iv) it is sometimes biased upward, sometimes downward and sometimes unbiased; (v) the term "bias" does not make sense in this context; (vi) none of the above. (Note: Our s^2 is biased downward as an estimator of σ^2.)

6. (15) Candidates X, Y and Z are vying for election. Let p_1, p_2 and p_3 denote the fractions (proportions!) of people planning to vote for them. We poll n people at random, yielding estimates $\widehat{p_1}$, $\widehat{p_2}$ and $\widehat{p_3}$. Y claims that she has more supporters than the other two candidates combined. Give a formula for an approximate 95% confidence interval for $p_2 - (p_1 + p_3)$.

7. (15) Think of sales of books A and B at an online bookseller. The vendor is considering recommending B to anyone who buys A. Define

$$\begin{aligned} p_1 &= \text{P(buy A and buy B)} \\ p_2 &= \text{P(buy A and not buy B)} \\ p_3 &= \text{P(not buy A and buy B)} \\ p_4 &= \text{P(not buy A and not buy B)} \end{aligned}$$

The seller has decided to recommend B if P(buy B | buy A) > 0.4. Devise an approximate $\alpha = 0.05$ test of

$$H_0 : \text{P(buy B | buy A)} = 0.4 \tag{1}$$

vs. H_A : P(buy B | buy A) > 0.4, based on our sample estimators $\widehat{p_i}$, i = 1,2,3,4.

Solutions:

1. $11.520 \pm 1.96 \cdot 9.939/\sqrt{344}$

2.

```
cat("the standard error is",s/sqrt(nreps))
```

3.

```
propci <- function(x,alpha) {
   phat <- mean(x)
   se <- sqrt(phat*(1-phat)/length(x))
   cat(phat,"+-",qnorm(1-alpha/2)*se)
}
```

4.a Set

$$\alpha = P(X > r) = \int_r^\infty ct^{-2}\ dt = \frac{c}{r} \tag{2}$$

Under H_0 we have c = 1. So, $r = 1/\alpha = 10$.

4.b

$$P(\text{reject}H_0) = P(X > 10) = \int_{10}^\infty 5t^{-2}\ dt = 0.5 \tag{3}$$

5. The answer is (iii).

$$E(\widehat{c}) = 2E(\overline{X}) - 1 = 2EX - 1 = c \tag{4}$$

6. Our point estimate is

$$\widehat{p_2} - (\widehat{p_1} + \widehat{p_3}) \tag{5}$$

In order to get a confidence interval, we need the standard error:

$$Var[\widehat{p_2} - (\widehat{p_1} + \widehat{p_3})] = Var(\widehat{p_2}) + Var(\widehat{p_1}) + Var(\widehat{p_3}) - 2Cov(\widehat{p_2}, \widehat{p_1}) - 2Cov(\widehat{p_2}, \widehat{p_3}) - 2Cov(\widehat{p_1}, \widehat{p_3}) \tag{6}$$

Now note that $Var(\widehat{p_i}) = p_i/n$ and for $i \neq j$ $Cov(\widehat{p_i}, \widehat{p_j}) = -p_i p_j/n$. Plug these into (6), substitute the sample estimates $\widehat{p_i}$, and take the square root. You now have the standard error. Add and subtract 1.96 times the standard error from the point estimate.

7. First,

$$P(\text{buy B}|\text{buy A}) = \frac{p_1}{p_1 + p_2} \tag{7}$$

So, we are testing

$$H_0 : \frac{p_1}{p_1 + p_2} = 0.4 \tag{8}$$

which after some algebra is seen to be the same as testing

$$H_0 : 1.5p_1 - p_2 = 0 \tag{9}$$

So, we set

$$Z = \frac{1.5\widehat{p_1} - \widehat{p_2}}{s.e.} \tag{10}$$

where s.e. is the standard error of the numerator, calculated in a manner similar to what is done in Problem 6. We then reject H_0 if Z > 1.65.