Second Midterm Examination

Econ 103, Statistics for Economists

March 21st, 2017

You will have 70 minutes to complete this exam. Graphing calculators, notes, and textbooks are not permitted.

I pledge that, in taking and preparing for this exam, I have abided by the University of Pennsylvania's Code of Academic Integrity. I am aware that any violations of the code will result in a failing grade for this course.

Name: ______________________________

Signature: ______________________________

Student ID #: ______________ Recitation #: ______________

Question:	1	2	3	4	5	6	Total
Points:	20	25	20	20	30	25	140
Score:							

Instructions: Answer all questions in the space provided, continuing on the back of the page if you run out of space. Show your work for full credit but be aware that writing down irrelevant information will not gain you points. Be sure to sign the academic integrity statement above and to write your name and student ID number on *each page* in the space provided. Make sure that you have all pages of the exam before starting.

Warning: If you continue writing after we call time, even if this is only to fill in your name, twenty-five points will be deducted from your final score. In addition, a point will be deducted for each page on which you do not write your name and student ID.

1. Let Y and Z be discrete RVs with the following joint pmf:

		Z		
		0	1	2
Y	0	1/4	3/8	1/8
	1	0	1/8	1/8

$\boxed{5}$ (a) Write down the support set and marginal pmf of Y.

$\boxed{5}$ (b) Write down the support set and marginal pmf of Z.

$\boxed{5}$ (c) Calculate $E[YZ]$.

$\boxed{5}$ (d) Write down the conditional pmf of Y given that $Z = 1$.

Name: ____________________ Student ID #: ____________________

2. Answer each part. No explanation is required.

$\boxed{4}$ (a) If X and Y are RVs with variance one and correlation $1/2$ what is $Var(X-Y)$?

$\boxed{4}$ (b) Suppose $X_1, \ldots, X_n \sim$ iid with mean μ and variance σ^2. Write down the formula for the standard error of $\bar{X}_n$.

$\boxed{4}$ (c) If $X \sim N(\mu = 3, \sigma^2 = 9)$ approximately what is the value of $P(0 \leq X \leq 6)$?

$\boxed{4}$ (d) Write R code to calculate the probability that a standard normal RV takes on a value between -1 and 3.

$\boxed{4}$ (e) Let $Z \sim N(0,1)$. Write R code to calculate c such that $P(-c \leq Z \leq c) = 0.8$.

$\boxed{5}$ (f) Suppose we observe a sample of 16 iid observations from a $N(\mu, \sigma^2)$ population and we know that $\sigma^2 = 1$. Our sample mean $\bar{x}$ turns out to equal 3. Write down an approximate 95% confidence interval for μ.

Name: ____________________ Student ID #: ____________________

3. Let X be a continuous RV with support set $[0,1]$ and pdf $f(x) = a + bx^2$ where $a, b > 0$.

$\boxed{6}$ (a) Calculate the CDF of X, $F(x_0)$, in terms of a and b.

$\boxed{6}$ (b) Calculate $E[X]$ in terms of a and b.

$\boxed{8}$ (c) Since a and b are positive, $f(x) \geq 0$. This is one of the two conditions required to ensure that $f(x)$ is a valid pdf. What is the other condition? What restriction must we place on a and b to ensure that the condition is satisfied?

$\boxed{20}$ 4. Let $X_1, \ldots, X_n \sim$ iid $N(\mu, \sigma^2 = 100)$ and define $\bar{X}_n = \frac{1}{n}\sum_{i=1}^n X_i$ as usual. Find the value of n for which $P(\mu - 5 \leq \bar{X}_n \leq \mu + 5) \approx 0.95$. Your answer should give a specific numeric value for n and should *not* involve any R commands.

Name: ______________________ Student ID #: ______________________

This problem was part of your homework for Lectures 13–14. I have reworded the problem slightly to make it clearer, but the details and solution are unchanged.

5. Let S denote the total number of successes in n iid Bernoulli trials, each with probability of success π. Consider two estimators of π: the sample proportion $P = S/n$ and an alternative estimator $P^* = \frac{S+1}{n+1} = \left(\frac{n}{n+2}\right) P + \left(\frac{1}{n+2}\right)$.

$\boxed{5}$ (a) Calculate the bias of P.

$\boxed{5}$ (b) Calculate $Var(P)$.

$\boxed{5}$ (c) Is P consistent for π? Explain briefly.

$\boxed{5}$ (d) Calculate the bias of P^*.

$\boxed{5}$ (e) Calculate $Var(P^*)$.

$\boxed{5}$ (f) Is P^* consistent for π? Explain briefly.

6. At the time of this writing approximately 50% of Americans disapprove of President Trump according to a weighted average of polls constructed by FiveThirtyEight. Assume for the purposes of this question that this estimate is *exactly correct* so that precisely half of the US population disapproves of Trump. This means that we can model a poll based on a random sample of n Americans as an iid sequence of n Bernoulli(1/2) RVs where a 1 indicates that a given individual in the poll *disapproves* of Trump.

$\boxed{10}$ (a) Write an R function called `disapprove_prop` that simulates a poll asking a random sample of n Americans if they disapprove of Trump. Your function should take a single input argument `n` the sample size of the poll and then make n iid simulation draws from the population described in the problem statement. The output of `disapprove_prop` should be the sample proportion of individuals who disapprove of Trump calculated from your `n` simulation draws.

$\boxed{5}$ (b) Suppose that you wanted to learn about the accuracy of the sample proportion in the poll described in the problem statement when $n = 10$. Write R code to approximate its sampling distribution based 10000 Monte Carlo simulations using your function `disapprove_prop` from part (b). Store your simulations in a vector called `poll_sims`.

$\boxed{10}$ (c) Continuing from (b) suppose I entered `mean(poll_sims)` and `var(poll_sims)` at the R console. Approximately what result would I get for each? Explain briefly.

Name: ______________________ Student ID #: ______________________