ECNS 561
Midterm
Fall 2017

______________________Name

1.) (15 points) Show that the sample average, $\bar{Y}$, is an unbiased and consistent estimator of μ.

2.) (15 points) Gamma random variables are characterized as follows:

Let X be a continuous random variable and its support be the set of positive real numbers:

$$R_X = [0, \infty).$$

We say that X has a Gamma distribution with parameters n and h if its probability density function is:

$$f_X(x) = \begin{cases} c x^{n/2 - 1} \exp\left(-\frac{n}{h}\frac{1}{2}x\right) & \text{if } x \in R_X \\ 0 & \text{if } x \notin R_X \end{cases}$$

where c is a constant and is equal to $\frac{(n/h)^{n/2}}{2^{n/2}\Gamma(n/2)}$

and $\Gamma(\ )$ is the Gamma function.

Show that the expected value of a Gamma random variable X is $\mathrm{E}[X] = h$.

3.) The uniform distribution is characterized as follows:

Let X be a continuous random variable and its support be a closed interval of real numbers:

$$R_X = [l, u].$$

We say that X has a uniform distribution on the interval $[l, u]$ if its probability density function is:

$$f_X(x) = \begin{cases} \frac{1}{u-l} & \text{if } x \in R_X \\ 0 & \text{if } x \notin R_X \end{cases}$$

a.) (7 points) Show that the expected value of a uniform random variable X is $E[X] = \frac{u+l}{2}$.

b.) (8 points) Show that the variance of a uniform random variable X is $\text{Var}[X] = \frac{(u-l)^2}{12}$.

c.) (5 points) Let X be a uniform random variable with support:

$$R_X = [4, 12].$$

Calculate the following probabilities:

$$P(5 \leq X \leq 10)$$

and

$$P(X > 6)$$

4.) (10 points) Let $\{Y_1, Y_2, ..., Y_n\}$ be a random sample of size n on annual earnings from the population of male workers with a high school education and denote the population mean as μ_Y. Let $\{Z_1, Z_2, ..., Z_n\}$ be a random sample of size n on annual earnings from the population of male workers with a college education and denote the population mean as μ_Z.

Let $\{W_1, W_2, ..., W_n\}$ be a random sample of size n on annual earnings from the population of female workers with a high school education and denote the population mean as μ_W. Let $\{X_1, X_2, ..., X_n\}$ be a random sample of size n on annual earnings from the population of female workers with a college education and denote the population mean as μ_X.

Suppose we want to estimate the difference in the percentage difference in annual earnings between the two groups across gender, $\delta = [100*(\mu_Z - \mu_Y)/\mu_Y] - [100*(\mu_X - \mu_W)/\mu_W]$. Show that

$$G_n = [100*(\overline{Z_n} - \overline{Y_n})/\overline{Y_n}] - [100*(\overline{X_n} - \overline{W_n})/\overline{W_n}]$$

is a consistent estimator of δ.

5.) (10 points) To investigate possible gender discrimination in a firm, a sample of 100 men and 64 women with similar job descriptions are selected at random. A summary of the resulting monthly salaries follows:

	Average Salary ($\bar{Y}$)	**Standard Deviation (S_Y)**	**N**
Men	$3100	$200	100
Women	$2900	$320	64

What do these data suggest about wage differences in the firm? Do they represent statistically significant evidence that wages of men and women are different? To answer this question, first state the null and alternative hypothesis; second, compute the relevant test statistic; and finally use the test statistic to answer the question.

Notes: • se($\bar{Y}_1 - \bar{Y}_2$) = $\sqrt{\frac{S_1^2}{N_1} + \frac{S_2^2}{N_2}}$

•

Critical values		
p-value	one-tailed test	two-tailed test
.050	1.65	1.96
.025	1.96	2.24
.010	2.33	2.58
.005	2.58	2.81
.001	3.08	3.3

6.) (10 points) The following question is based on the article by Prickett et al. (2014) in the *American Journal of Public Health*. In this paper, the authors found that, when interacted with stronger firearm legislation, Child Access Prevention (CAP) laws were associated with safer household gun storage. Should their estimates be given a causal interpretation? If so, explain why you find their identification strategy convincing. If not, explain why you do not find their identification strategy convincing and provide at least two specific reasons as to why their estimates may be biased.

7.) In 1986, Barton County, KS voted to allow the sale of alcohol consumption by the drink in bars and restaurants. In other words, Barton County went from being classified as a "dry" county to being classified as a "wet" county. Before 1986, alcohol by the drink could only be purchased in special clubs and lounges that required paying a membership fee; this greatly limited the opportunities to consume alcohol in public settings. A Barton County public official has asked you to evaluate whether going "wet" increased violent crime.

To evaluate going "wet", you have decided that you must first select an appropriate control group (i.e. a county that remained "dry" during this same period). Consider the following characteristics of Barton County and the bordering counties of Pawnee, Stafford, and Rice.

	Barton County	Pawnee County	Stafford County	Rice County
% of population that is white	85%	90%	75%	82%
Whether Sunday liquor sales are legal	Yes	No	Yes	No
% of population that voted democrat	40%	35%	35%	85%
County-level unemployment rate	8%	3%	8%	12%
% of population with a HS diploma	81%	90%	75%	60%

a.) (5 points) Assuming you can only choose one, which of the three bordering counties (Pawnee, Stafford, or Rice) would you choose as your control county? Why did you choose this county?

b.) (7 points) To evaluate the impact on violent crime, you have data on average monthly violent crimes committed per 1,000 people for both Barton County and your control county (which we will just refer to as County X). You have these data for 1985 and 1987 as follows:

	Barton County		County X	
	1985	1987	1985	1987
Violent crime rate	1.5	3.0	2.3	1.8

If you were to **only** use within-Barton County variation in crime, would you likely be over- or under-estimating the true effect of going "wet" on violent crime rates? Why?

c.) (8 points) Now let's assume you have more "pre-treatment" data on average monthly violent crime rates for Barton County and County X:

	Barton County					County X				
	1979	1981	1983	1985	1987	1979	1981	1983	1985	1987
Violent crime rate	1.0	1.2	1.35	1.5	3.0	2.3	2.3	2.3	2.3	1.8

Given this new information, how does inference from part b.) change? That is, without the consideration of pre-existing trends in violent crime, does a "difference-in-difference" estimate based solely on the information provided in part b.) overstate or understate the true impact of going "wet" on violent crime? You do not necessarily have to do any calculations, I am mainly looking for the correct intuition.