ECNS 561
Midterm
Fall 2016

105 possible points

______________________Name

1.) Consider the following function

$$f(x) = \begin{cases} cx^2 & 0 < x < 3 \\ 0 & \text{otherwise} \end{cases}$$

a.) (5 points) Find the constant *c* such that the function above is a density function.

b.) (5 points) Having solved for *c*, now compute $P(1 < X < 2)$.

2.) The triangular distribution is a continuous probability distribution with lower limit a, upper limit b, and mode c, where $a < b$ and $a \le c \le b$. We say that X has a triangular distribution if its pdf is

$$f(x) = \begin{cases} 0 & \text{for } x < a \\ \frac{2(x-a)}{(b-a)(c-a)} & \text{for } a \le x < c \\ \frac{2}{b-a} & \text{for } x = c \\ \frac{2(b-x)}{(b-a)(b-c)} & \text{for } c < x \le b \\ 0 & \text{for } b < x \end{cases}$$

a.) (8 points) For the values $a = 1$, $b = 4$, and $c = 2$ and using calculus, solve for E[X].

b.) (7 points) For the values $a = 1$, $b = 4$, and $c = 2$ and using calculus, solve for Var[X].

3.) (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.

4.) Let $\bar{Y}$ denote the sample average from a random sample with mean μ and variance σ^2. Consider two alternative estimators of μ:

$W_1 = (\frac{n-1}{n})\bar{Y}$

$W_2 = \frac{\bar{Y}}{2}$

a.) (6 points) Show that W_1 and W_2 are both biased. Solve for the biases and show what happens to them as $n \to \infty$.

b.) (4 points) Find the probability limits of W_1 and W_2. Which estimator is consistent?

c.) (4 points) Find Var(W_1) and Var(W_2).

d.) (6 points) Argue that W_1 is a better estimator than $\bar{Y}$ if μ is "close" to zero (hint: you need to consider both bias and variance).

5.) In matrix notation, suppose we have the following population model

$$\boldsymbol{y} = \boldsymbol{X\beta} + \boldsymbol{\varepsilon}$$

where $\boldsymbol{X}$ is an n by k matrix of k independent variables for n observations, $\boldsymbol{y}$ is an n by 1 vector of observations on the dependent variable, $\boldsymbol{\beta}$ is a k by 1 vector of unknown population parameters that we wish to estimate, and $\boldsymbol{\varepsilon}$ is an n by 1 vector of errors.

a.) (6 points) Using matrix notation, solve for the OLS estimator (i.e., the $\widehat{\boldsymbol{\beta}}$ that minimizes the sum of squared residuals).

b.) (7 points) Show that $\widehat{\boldsymbol{\beta}}$ is an unbiased estimator of $\boldsymbol{\beta}$.

c.) (7 points) Now, instead of the model above, suppose the true population model is given as

(1) $$\mathbf{y} = \mathbf{X}_1\boldsymbol{\beta}_1 + \mathbf{X}_2\boldsymbol{\beta}_2 + \boldsymbol{\varepsilon}.$$

However, suppose you only have data to estimate the following model

(2) $$\mathbf{y} = \mathbf{X}_1\boldsymbol{\beta}_1 + \mathbf{u} \qquad \text{where } \mathbf{u} = \mathbf{X}_2\boldsymbol{\beta}_2 + \boldsymbol{\varepsilon}.$$

Using matrix notation, show that the OLS estimator for equation (2) is biased.

6.) Consider the following county-level regression:

$$Crime_c = \beta_0 + \beta_1 Education_c + \varepsilon_c,$$

where *c* index counties. The dependent variable *Crime* is the crime rate per 1,000 population and the independent variable *Education* is the average number of years of education in the county.

The following table provides descriptive statistics for these variables for all 56 counties in MT for the year 2015:

Descriptive Statistics

	Mean	Definition
Dependent variable		
Crime	43.5	Number of crimes per 1,000 population
Independent variables		
Education	12.0	Ave. years of education
N = 56		

Suppose you use OLS to estimate the above equation and you obtain the following coefficient estimates (with standard errors in parentheses):

$$\widehat{\beta_0} = \ 60.0\ (15.0)$$
$$\widehat{\beta_1} = \ -2.25\ (0.25)$$

a.) (5 points) Interpret $\widehat{\beta_1}$. Given the descriptive statistics, would you say this effect is large in magnitude?

b.) (5 points) Would you say that $\widehat{\beta_1}$ is precisely estimated?

c.) (5 points) Given the coefficient estimates, plot the OLS regression line below. Make sure to label your graph and axis intercept points.

d.) (5 points) Suppose the raw data points for five MT counties are as follows:

County	*Crime*	*Education*
Fergus	40.0	12.0
Flathead	37.5	10.0
Gallatin	24.0	16.0
Missoula	20.5	14.0
Yellowstone	35.0	8.0

In the graph above, indicate the residuals for these five counties and provide the values of each residual.

7.) Recall the article by Mansour and Rees (2012) published in the *Journal of Development Economics* and entitled "Armed Conflict and Birth Weight: Evidence from the al-Aqsa Intifada." In this paper, the authors are interested in examining the relationship between intrauterine exposure to armed conflict and birth weight. Their results suggest that an additional conflict-related fatality 9-6 months before birth is associated with a modest increase in the probability of having a child who weighed less than 2500 g.

a.) (5 points) Briefly describe the natural experiment set up used to identify a causal relationship between exposure to armed conflict and birth weight.

b.) (5 points) Briefly describe the potential mechanisms through which armed conflict could influence birth weight. Which mechanism did the authors find support for? How did they determine this?