Mega HW #1
ECNS 561
Fall 2018

Due Date: September 25th

NOTE: You may hand your homework in by groups of two or three. You are not required to work on your assignment with another person, but it is recommended.

Your handwriting MUST be legible. If it isn't, I will refuse to grade your homework. You will have to rewrite it and will be docked points for it being late.

1.) Work the following problems from Wooldridge Appendix B (pgs. 752-754):

> B2 (note: you can use Table G.1 to calculate the final probabilities), B4, B5, B6, B10 (parts i and ii only).

2.) Find the expectation and variance of a continuous random variable with a pdf of $f(x) = 2(1-x)$ for all $0 < x < 1$, and zero elsewhere.

3.) Suppose that 90% of all batteries from a certain supplier have acceptable voltages. A certain type of flashlight requires two type-D batteries, and the flashlight will work only if both its batteries have acceptable voltages. Among ten randomly selected flashlights, what is the probability that at least nine will work?

4.) Consider the following pdf for a discrete random variable (this is a distribution statisticians would call "heavy-tailed)

$$p(x) = \begin{cases} c/x^3 & x = 1, 2, 3, \ldots \\ 0 & \text{otherwise} \end{cases}$$

Is E(X) finite? Justify your answer. (Hint: You might need to brush up on your sums of infinite series).

5.) A result called **Chebyshev's inequality** states that for any probability distribution of a random variable X and any number k that is at least 1, $P(|X - \mu| \geq k\sigma) \leq 1/k^2$. In words, the probability that the value of X lies at least k standard deviations from its mean is at most $1/k^2$.

i.) Suppose the pdf of X is as follows

x	0	1	2	3	4	5	6
P(x)	.1	.15	.2	.25	.2	.06	.04

Compute μ and σ for this distribution. Then evaluate $P(|X - \mu| \geq k\sigma)$ for $k = 2$. What does this suggest about the upper bound in Chebyshev's inequality relative to the corresponding probability?

ii.) Now, let X have three possible values -1, 0, and 1, with probabilities 1/18, 8/9, and 1/18, respectively. What is $P(|X - \mu| \geq 3\sigma)$, and how does it compare to the corresponding bound?

iii.) Give a distribution for which $P(|X - \mu| \geq 5\sigma) = .04$.

iv.) Chebyshev's inequality is valid for continuous as well as discrete distributions. Obtain $P(|X - \mu| \geq k\sigma)$ in the case of a normal distribution for k = 1, and compare to the Chebyshev upper bound.

6.) Suppose a continuous random variable X has the following pdf

$$f(x) = \begin{cases} .15e^{-.15(x-.5)} & x \geq 0.5 \\ 0 & \text{otherwise} \end{cases}$$

Using calculus, show that the area under f(x) between $-\infty$ and ∞ is equal to 1.

7.) The event $\{X^2 \leq y\}$ is equivalent to $\{-\sqrt{y} \leq X \leq \sqrt{y}\}$. If X has a standard normal distribution, use the previous statement to write the integral that equals $P(X^2 \leq y)$. Then differentiate this with respect to y to obtain the pdf of X^2 [the square of a N(0, 1) variable]. Finally, show that X^2 has a chi-squared distribution with v = 1 degrees of freedom.
(Hint #1: Use the following identity:

$$d/dy\left\{\int_{a(y)}^{b(y)} f(x)dx\right\} = f[b(y)] \cdot b'(y) - f[a(y)] \cdot a'(y).$$

Hint #2: One important property of the gamma function $\Gamma(\alpha)$ is that $\Gamma(1/2) = \sqrt{\pi}$.)

8.) Recall our definition of a continuous random variable, X, that follows a Chi-squared distribution

$$f(x) = \begin{cases} cx^{n/2-1}\exp(-x/2) & x \geq 0 \\ 0 & \text{otherwise} \end{cases}$$

where c is a constant:

$$c = 1/(2^{n/2}\Gamma(n/2))$$

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

i.) Show that the expected value of a Chi-square random variable X is

$$E[X] = n$$

Make sure to show your steps for this proof.

(Hint #1. Don't forget about integration by parts…I know you brushed up on this in math camp! Hint #2. Remember, c is a constant…don't make your life more complicated than it already is!)

ii.) Show that the variance of a Chi-square random variable X is

$$Var[X] = 2n$$

Again, make sure to show your steps.

9.) This problem requires the use of the "KS Per Capita Income 1969 to 2012" data file that can be access from the course webpage. In case you are running an older version of STATA, I have attached an excel file that you can copy and paste into your STATA data editor. The variables in the data set are defined as follows:

> fips: unique county identifier
> county: county name
> yr_i: nominal income per capita in year i

For the following problem, you are required to turn into me your STATA do file (either through email or share it with me via dropbox). Your entire answer should be self-contained within one do file. That is, I should be able to open the original data set in STATA, open your do file, click "run", and have the final cleaned data set right in front of me.

i.) First, you will need to convert the data to a "long" format. The data in the excel file are currently in a "wide" format. Check out the following link that shows how to use the "reshape" command in STATA to convert data from wide to long: http://www.ats.ucla.edu/stat/stata/modules/reshapel.htm.

ii.) Next, you will need to convert the nominal income per capita data to real dollars (use dollars in 2000 as your reference year). Here is link to the historical consumer price index (http://www.inflationdata.com/inflation/consumer_price_index/historicalcpi.aspx?reloaded=true). When converting the nominal income figures into real dollars, make sure to use the CPIs in the "Annual" column of the table found at the link above.

iii.) Now that you have a data set in long format where one of your variables is income in real dollars, let's look at some real income distributions.

> a.) Create a histogram for real income across all Kansas counties for the year 1970. The command you will be using in STATA is the "histogram" command. Perhaps the easiest way to learn how to create histograms is to go the drop down tab "Graphics" and then click on "Histogram." From there you can play around until you get the command you want for your final do file that you will turn into me. Make your graph look good! I want to see a graph title, the y-axis and x-axis labelled, etc. If you want to see some

examples of good looking STATA graphs check out Figures 1-3 in this awesome paper: http://www.dmarkanderson.com/Medical_Marijuana__Accidents_and_Alcohol_7-8-13_v1.pdf. Or, you can check out Figures 2a and 2b in this equally awesome paper: http://www.dmarkanderson.com/Deployments_Combat_Exposure_and_Crime_07_30_14_v2.pdf. I might even throw in a few extra credit points for the homework with the best looking graphs. Print this graph and turn it in with your homework.

b.) Does real income in 1970 appear to be normally distributed? Why or why not?

c.) Now, create a histogram for the natural log of real income per capita in 1970. Print this graph and turn it in with your homework. Does the natural log of real income per capita appear to be normally distributed? Which histogram appears to be a closer approximation of a normal distribution?

d.) Has the income distribution changed much over time in KS? You do not need to turn in any graphs here…just a short description will be fine.