

# Pol Sci 630: Problem Set 2 - Properties of Random Variables

Prepared by: Anh Le (anh.le@duke.edu)

Due Date: Tuesday, September 8, 2015, 10 AM (Beginning of Class)

Note 1: It is absolutely essential that you show all your work, including intermediary steps, and comment on your R code to earn full credit. Showing all steps and commenting on code them will also be required in future problem sets.

Note 2: Please use a \*single\* PDF file created through knitr to submit your answers. knitr allows you to combine R code and  $\text{\LaTeX}$  code in one document, meaning that you can include both the answers to R programming and math problems. Also submit the source code that generates the PDF file (i.e. either .Rnw or .Rmd files)

Note 3: Make sure that the PDF files you submit do not include any references to your identity. The grading will happen anonymously. You can submit your answer at the following website: <http://ps630-f15.herokuapp.com/>

## 1. Expected Value and Its Properties

**a.**

(1 point) (DeGroot, p. 216) Suppose that one word is to be selected at random from the sentence “the girl put on her beautiful red hat”. If  $X$  denotes the number of letters in the word that is selected, what is the value of  $E(X)$ ?

**b.**

(2 point) (Degroot p. 216) Suppose that one letter is to be selected at random from the 30 letters in the sentence given above. If  $Y$  denotes the number of letters in the word in which the selected letter appears, what is the value of  $E(Y)$ ?

Hint: 1a) and 1b) force you to think carefully about the definition of expectation value. For each problem, think about what is your random variable ( $X$ ), which values it takes on ( $x \in \{?, ?, \dots\}$ ) and with what probability ( $P(X = x) = ?$ )

**c.**

(1 point) (Degroot, p. 224) Suppose that three random variables  $X_1, X_2, X_3$  are uniformly distributed on the interval  $[0, 1]$ . They are also independent. Determine the value of  $E[(X_1 - 2X_2 + X_3)^2]$ .

## 2. Variance and its properties

For this problem, you can use the properties of expected value.

**a.**

(1 point) Prove that  $Var(aX + b) = a^2 Var(X)$ .

**b.**

(2 points) Implement in R two functions that calculates the variance of the sum of two variables in two ways. The first calculates `Var(X + Y)`. The second calculates `Var(X) + Var(Y) + 2Cov(X, Y)`.

You should use vectorized operation and check that two functions return the same result. You may not use R's built-in `var()` and `cov()` functions.

**c.**

(1 point) (Degroot, p. 232) Suppose that one word is selected at random from the sentence 'the girl put on her beautiful red hat'. If  $X$  denotes the number of letters in the word that is selected, what is the value of  $Var(X)$ ?

## 3. Binomial distribution (4 points)

(Credit to Jan) This problem is taken from Pitman (1993) Probability

Suppose a fair coin is tossed  $n$  times. Find a simple formula in terms of  $n$  and  $k$  for the following probability:  $Pr(k \text{ heads} | k-1 \text{ heads or } k \text{ heads})$ . Please pay close attention to the formula, particularly what event is conditioned on what events. (Ch. 2.1, Problem 10 b) (p. 91)

Hint 1: Use the binomial distribution to model this.

Hint 2: Use  $Pr(A|B) = \frac{Pr(A \cap B)}{Pr(B)}$  with  $A = k \text{ heads}$  and  $B = k-1 \text{ heads or } k \text{ heads}$

## 4. Plotting distribution (4 points)

For this problem, you'll need to Google some R techniques (e.g. side-by-side / overlapping plot). Also, label the axes and the plots accordingly.

**a.**

(1 point) Download a variable you are interested in, using WDI. Plot the histogram, density plot, boxplot, and normal quantile plot.

**b.**

(1 point) Plot the density of that variable for Europe and Asia, 1) side by side (Hint: `par(mfrow=c(?, ?))`), and 2) overlapping in the same plot.

**c.**

(1 point) Draw the scatterplot of that variable against another variable of interest. Is the bi-variate relationship as you expected?

**d.**

(1 point) In the scatterplot, label the point that represents your country (Hint: Tutorial) and color it red (some Googling may be necessary)