Math 472: Homework 01

Due Wednesday, September 10 (at the beginning of class)

Problem 1. Install a bunch of software on your laptop:

(a) Install the statistical software `R`, available at the `https://cran.rstudio.com/`

(b) Install `RStudio Desktop`, available at `https://posit.co/download/rstudio-desktop/`

(c) Do the `R` tutorial

(d) DO `https://swcarpentry.github.io/shell-novice/` or `https://carpentries-incubator.github.io/git-novice-branch-pr/` or `https://cecileane.github.io/computingtools/pages/notes0922-markdown.html`

(e) Write an R function to generate and plot a 1 dimensional brownian motion

Remark 2. The next problem is intended to introduce/review some central concepts in probability. For the remainder of this course, any terms defined using *this format* are precise definitions and should be memorized (they are fair game for in-class quizzes) Any terms that **are bolded** are important and should be reviewed if you are not familiar with them.

Problem 3. An urn contains 5 balls, three red balls and two blue balls:

🔴🔴🔴🔵🔵

We consider the problem of sampling 3 balls from the urn, drawing the balls "without replacement". This means we draw one ball at random, then draw another ball at random, and then draw a third ball at random, without ever putting any of the balls back into the urn.

For $k = 1, 2$, and 3, we will use the notation R_k to denote the event that the k^{th} drawn ball is red, and B_k to denote the event that the k^{th} drawn ball is blue. Obviously, $\mathbb{P}[R_1] = 3/5$ and $\mathbb{P}[B_1] = 2/5$.

(a) Compute the conditional probabililities $\mathbb{P}[R_2 \mid B_1]$ and $\mathbb{P}[R_2 \mid R_1]$.

(b) Use the **Law of Total Probability** (Theorem 2.8 in the textbook, p70) and your answer to part (a) to compute $\mathbb{P}[R_2]$.

(c) If E and F are events, we use the notation EF to denote the event that both E and F occur (i.e., $EF = E \cap F$). Compute the probabilities of the four events R_1R_2, R_1B_2, B_1R_2 and B_1B_2.

(d) Use the Law of Total Probability and your answer to part (c) to compute $\mathbb{P}[R_3]$.

(e) In the remainder of this problem, we will compute the expected proportion of red balls among our 3 draws. To do this, we will introduce a standard technique: the use of indicator functions.

Given an event E, the *indicator function of E* is the function

$$\mathbf{1}_E = \begin{cases} 1 & : \text{ the event } E \text{ occurs} \\ 0 & : \text{ the event } E \text{ does not occur} \end{cases}$$

Indicator functions are random variables. Taking $E = R_k$, we have

$$\mathbf{1}_{R_k} = \begin{cases} 1 & : \text{ the } k^{\text{th}} \text{ ball drawn is red} \\ 0 & : \text{ the } k^{\text{th}} \text{ ball drawn is blue.} \end{cases}$$

A random variable is *discrete* if it can assume only a finite or countably infinite number of distinct values.

If X is a discrete random variable, and $S_X \subseteq \mathbb{R}$ is the set of possible values that X can take, then the *expectation* of X, denoted $\mathbb{E}[X]$, is defined as

$$\mathbb{E}[X] := \sum_{x \in S_X} x\mathbb{P}[X = x],$$

provided that this sum converges absolutely.

Using the above definition of expectation, prove that $\mathbb{E}[\mathbf{1}_{R_k}] = \mathbb{P}[R_k]$ for $k = 1, 2, 3$.

(f) Observe that

$$(\#\text{ of red balls in 3 draws}) = \mathbf{1}_{R_1} + \mathbf{1}_{R_2} + \mathbf{1}_{R_3},$$

and hence

$$(\text{the proportion of red balls in 3 draws}) = \frac{\mathbf{1}_{R_1} + \mathbf{1}_{R_2} + \mathbf{1}_{R_3}}{3} \tag{1}$$

The **linearity of expectation** says that if X, Y are random variables, and a, b are scalars, then $\mathbb{E}[X + Y] = \mathbb{E}[X] + \mathbb{E}[Y]$, and $\mathbb{E}[aX + b] = a\mathbb{E}[X] + b$. Use the linearity of expectations and Eq. (1) to compute the expected proportion of red balls in 3 draws. *(Note: if you've done all parts of this problem correct, you'll get 3/5.)*