Module 1 - MTH 461

Contents

Instructions:

- Please provide complete solutions for each problem. If it involves mathematical computations, explanations, or analysis, please provide your reasoning or detailed solutions.

- Note that some problems have multiple solutions or ways to solve it. Make sure that your solutions are clear enough to showcase your work and understanding of the material.

- Creativity and collaborations are encouraged. Use all of the resources you have and what you need to complete the module. Each student must take personal responsibility and submit their work individually. Please abide by the University of Portland Academic Honor Principle.

- There are two ways you can write your answers, a: by handwriting (either physically or digitally), or b: by typing on a template document with file type options, which can be downloaded from the course website.

- If you had handwritten your answers/solutions on a physical paper, make sure to label it properly and please scan your document using a scanner app for convenience. Suggestions: (1) "Tiny Scanner" for Android or (2) "Scanner App" for iOS.

- **Please save your work as one pdf file, don't put your name in any part of the document, and submit it to the Teams Assignments for this course. Your document upload will correspond to your name automatically in Teams.**

- If you have questions or concerns, please feel free to ask the instructor.

R Packages:

- Below are pre-loaded general packages required for this module assignment. You can load more packages here or throughout the module if necessary.

- Note that you need to install R packages before you can use them. You can use the `install.packages()` in the R console, or go to the "Tools" tab and click "Install Packages..." in R Studio.

- Be careful on loading R packages because sometimes any two packages can have conflicting functions when calling them.

```r
# pre-load packages here
library(tidyverse)
```

I. Calculus Review

Objective: To review what you have learned and to rejuvenate your skills in Calculus.

Exercises

1. The function $p(a)$ is defined for all real numbers a as

$$p(a) = \begin{cases} 1/36, & a = 1 \\ 3/36, & a = 2 \\ 5/36, & a = 3 \\ 7/36 & a = 4 \\ 9/36, & a = 5 \\ 11/36, & a = 6 \\ 0, & \text{otherwise.} \end{cases}$$

   a. Sketch the graph of the function $p(a)$.
   b. What are the values of $p(-0.5)$, $p(0)$, $p(2)$, $p(3.5)$, and $p(7)$?

2. Let $p(a)$ be the same function as in the previous problem. Define another function $F(a)$ in term of $p(a)$ for all real numbers a as

$$F(a) = \sum_{a_i \leq a} p(a_i), \quad \text{if } p(a_i) > 0.$$

   a. What are the values of $F(-0.5)$, $F(0)$, $F(1)$, $F(1.6)$, $F(2)$, $F(3.5)$, and $F(7)$?
   b. Write down the formula for $F(a)$ as a piecewise function (similar to the one above for $p(a)$).
   c. Sketch the graph of $F(a)$.
   d. Find the limits $\lim_{a \to -\infty} F(a)$ and $\lim_{a \to \infty} F(a)$.

3. Given the function $f(x)$ below.

$$f(x) = \begin{cases} cx + 3, & -3 \leq x \leq -2 \\ 3 - cx, & 2 \leq x \leq 3 \\ 0, & \text{otherwise} \end{cases}$$

   If $\int_{-\infty}^{\infty} f(x)\, dx = 1$, what is the value of c?

4. Let $f(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$. The function $F(a)$ is defined, for any $a \geq 0$, in terms of f as

$$F(a) = \int_{-\sqrt{a}}^{\sqrt{a}} f(x)\, dx.$$

   Find the expression for the derivative $F'(a)$, $a \geq 0$.

II. Simulating Probabilities

Objectives:

- Perform sampling with or without replacement.
- Perform randomization and probability simulations.

Materials

- `sampling-from-urns.R` - This R script provides R codes that simulates random sampling of balls from an urn and visualizes the result. It also provides example R codes on how to code in in the theoretical probabilities of the balls and an urn scenario.

Exercises

Consider the scenario where you have an urn with a total of 20 balls: 9 black balls and 11 red balls. You will draw 7 balls from this urn without replacement. Estimate the probability of drawing 0, 1, 2, 3, 4, 5, 6, or 7 black balls and compare this estimate to the true probabilities.

1. Shown below is an example R code block that runs the R script named `sampling-from-urns.R`. Run it and provide your initial output below. Describe what the output shows you. Note that the initial simulation does not match the given scenario.

```
source("sampling-from-urns.R")
```
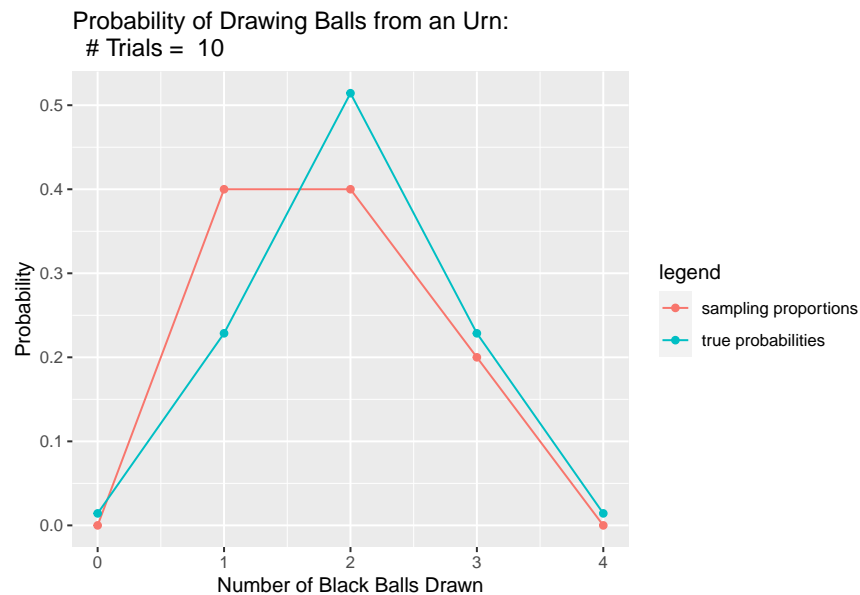

2. Calculate by hand the true probabilities in this scenario. (The code you received provides a lot of reasoning in the comments.)

3. Modify the R script you were given to simulate the given scenario. When you modify the script, make sure you create a new file with a new file name and run it similarly from problem 1. Provide an output below and describe your observations.

4. Provide outputs from your code (the figures comparing both probabilities) for at least these two different scenarios:

    a. `numTrials = 10^2`
    b. `numTrials = 10^4`

5. A short explanation of what is shown in your graphs and what you have learned from this exercise. (Example: Do you feel you agree well with the theoretical results? Which one is better? How many trials do you think you need to get a "good" agreement to the true probability?)

6. If you did not know the true probabilities for this experiment how would you decide your simulations had considered enough trials.

III. BibTeX References and Citations in R Markdown

Objective: Part of any scientific work in Data Science and Applied Mathematics, requires you to write your work with citations. The purpose of this part of the Module teaches you how to use bibTeX references and citations.

Materials

This part of the module uses the following files.

- The file named `references.bib` contains an entry of a BibTeX formatted reference shown below. The bib key to this entry is `R-base`.

```
@Manual{R-base,
  title = {R: A Language and Environment for Statistical Computing},
  author = {{R Core Team}},
  organization = {R Foundation for Statistical Computing},
  address = {Vienna, Austria},
  year = {2016},
  url = {https://www.R-project.org/},
}
```

- The file named `apa.csl` is a CSL (Citation Style Language) file which was downloaded from Zotero Style Repository. It contains code to generate citations and references within your document using the American Psychological Association (APA) 7th edition style.

Exercises

1. Visit Google Scholar and search for "data science". Using the first item in the results, click the "Cite" button. Look for the BibTeX link, where it leads to another page showing the BibTeX formatted citation of the reference. Copy and paste the citation into a new line in the `references.bib` file. Show what you copied and pasted by filling in the verbatim code block below.

```
put your BibTeX formatted citation here and in the .bib file
```

2. We can cite references listed in our BibTeX file by using the bib keys of the references. We can start using those bib keys, using the following syntax:

   - `[@key]` for single citation or `[@key1; @key2]` for multiple citations
   - `@key` is also a valid syntax when using it in a sentence

   For citing the `R-base` bib key, we use `[@R-base]`, which is rendered as (R Core Team, 2016). Notice the `@` key character before the bib key to indicate that we are citing an entry from the BibTeX file which will show in the References section.

   Identify the bib key of the BibTeX reference you just added and cite it in two different ways: (a) as a single citation, and (b) use it in a sentence. You can use a made-up sentence for the purpose of this exercise.

References

R Core Team. (2016). *R: A language and environment for statistical computing.* R Foundation for Statistical Computing. https://www.R-project.org/