



STOR 320 Programming II

Lecture 21

Yao Li

Department of Statistics and Operations Research

UNC Chapel Hill

Introduction

- Reading
 - Chapters 19-21 in R4DS
 - Chapters 14-18 in RP4DS
 - Chapter 7 in AoRP
 - Chapter 4 in FCSPR
- Programming Steps
 - Understand the Problem
 - Inputs and Outputs
 - Create Code
 - Test the Code (Simple Case)
 - Generalize the Code
 - Test Problematic Cases
 - Edit Code to Handle Issues
 - Consider Efficiency

Setup for Lecture

- Open Tutorial 10
- Packages Required:
 - Tidyverse
 - Ecdat
- Knit Document As You Go
- Read Introduction
- Prepare Your Minds for the Matrix



Part 1: Loops

- Correlation Matrix
 - Definition: Matrix Which Shows the Correlation Between Every Pair of Numeric Variables
 - Used to Understand Strength of Linear Relationships Between Numeric Variables
 - Helpful in Measuring Collinearity
- Run Chunk 1
 - Inspect the Variables in Cigar
 - Inspect the Correlation Matrix
 - Which Variable(s) is Inappropriate for a Correlation Analysis? Why?

Part 1: Loops

- Run Chunk 2
 - Run First Half – Loops through Every Combination of Columns and Computes Correlation
 - Examine Second Half – Loops Through Every Combination of Columns Excluding the First Column
 - Fill in Blanks with Appropriate Indices so Second Loop Works
 - Run Second Half
- Run Chunk 3
 - Inspect the Variables in H1
 - Uncomment to Print Correlation Matrix
 - What is the Problem?

Part 1: Loops

- Run Chunk 4
 - Observe the Difference Between the Printed Tibbles
 - What is the Difference?
 - How Would You Explain the First Loop to a Toddler?
 - What is `cat()` doing?
 - How Would You Explain the Second Loop to an Infant?
 - Remember: There Are an Infinite Number of Ways to Do the Same Thing.

Part 2: SRS

- Important For Simulation Studies
- Known Distributions

Distribution	Density/pmf	cdf	Quantiles	Random Numbers
Normal	<code>dnorm()</code>	<code>pnorm()</code>	<code>qnorm()</code>	<code>rnorm()</code>
Chi square	<code>dchisq()</code>	<code>pchisq()</code>	<code>qchisq()</code>	<code>rchisq()</code>
Binomial	<code>dbinom()</code>	<code>pbinom()</code>	<code>qbinom()</code>	<code>rbinom()</code>

- “d” -> Useful for Plotting Density Curve for Continuous Variables or Probability Mass Function for Discrete Variables
- “p” -> Finds the Probability Less Than Or Equal to a Given Number
- “q” -> Finds Cutoff Points
- “r” -> Generates a Random Sample from the Distribution

Part 2: SRS

- For SRS, Use “r”
- Run Chunk 1
 - Scenario for x1: You Ask BLANK Number of Students Their Grades where Grades Follow a Normal Distribution with Mean=82 and SD=2
 - Scenario for x2: You Ask BLANK Number of Students to Roll a Fair Die 10 Times and Tell You the Number of 6's that Appeared.

Part 2: SRS

- Sampling From Finite Set of Possible Outcomes
- Run Chunk 2
 - Scenario: Flip k Coins
 - $P(\text{Heads}) = \text{BLANK}$
 - $P(\text{Tails}) = 1 - \text{BLANK}$
 - How would You Explain What the Figure is Showing to a Politician?