# An Introduction to Git, RStudio, and R Markdown

*Bayesian Psychometric Models, Lecture 2*

## In This Lecture:

1. Git (and how to access the course GitHub repo)
2. RStudio (and how to use Git within RStudio)
3. R Markdown (and how to make it run to play with syntax during class)

## Git Clients

- If you use a Mac/Linux machine, you already have a client as part of the OS.
- If you use Windows, you will have to download and install a client to use Git
  - A good client to use with RStudio is https://git-scm.com/download/win

## The Basics of Git

- Git is a version control system that helps to keep track of changes to files across the lifespan of a project
  - It is also great for using in collaboration with others
- An overview of Git is given in many places here are a few:
  - http://rogerdudler.github.io/git-guide/ (thank you, Dr. LeBeau)
  - https://git-scm.com/book/en/v2/Getting-Started-Git-Basics
- I will use these websites to help introduce you to Git concepts
- Also, I will highlight the GitHub site with our course repo: https://github.com/jonathantemplin/BayesianPsychometricModeling

## Using Git with R Studio

- Git makes our life easy when used with RStudio as we can use it to keep a current copy of our course notes available
  - We will use Git to download and update our course materials in RStudio
- Do do this:
  1. Open RStudio
  2. At the menu on the top, go to File…New Project
  3. Select "Version Control" on the "Create Project" window
  4. Select "Git" on the "Create Project from Version Control" window
  5. On the "Clone Git Repository" window input:
     - Repository URL: https://github.com/jonathantemplin/BayesianPsychometricModeling
     - Project directory name: Choose a directory name for the course materials (such as "BPM Git Repo")
       - NOTE: The master branch has an up-to-date R project file (.Rproj) that will have all files included.
     - Create project as a subdirectory of: Choose a location for your files to reside within on your local machine

## Using RStudio

- Next I will demonstrate RStudio for you using the contents of last week's R script file

- For more information about the RStudio Integrated Development Environment (IDE), see the following links:
  - https://www.rstudio.com/online-learning/
  - https://dss.princeton.edu/training/RStudio101.pdf
- Also, to really unlock RStudio's full potential, familiarize yourself with its keyboard shortcuts:
  - In the top menu, go to Tools... Keyboard Shortcuts Help

## Using R Markdown

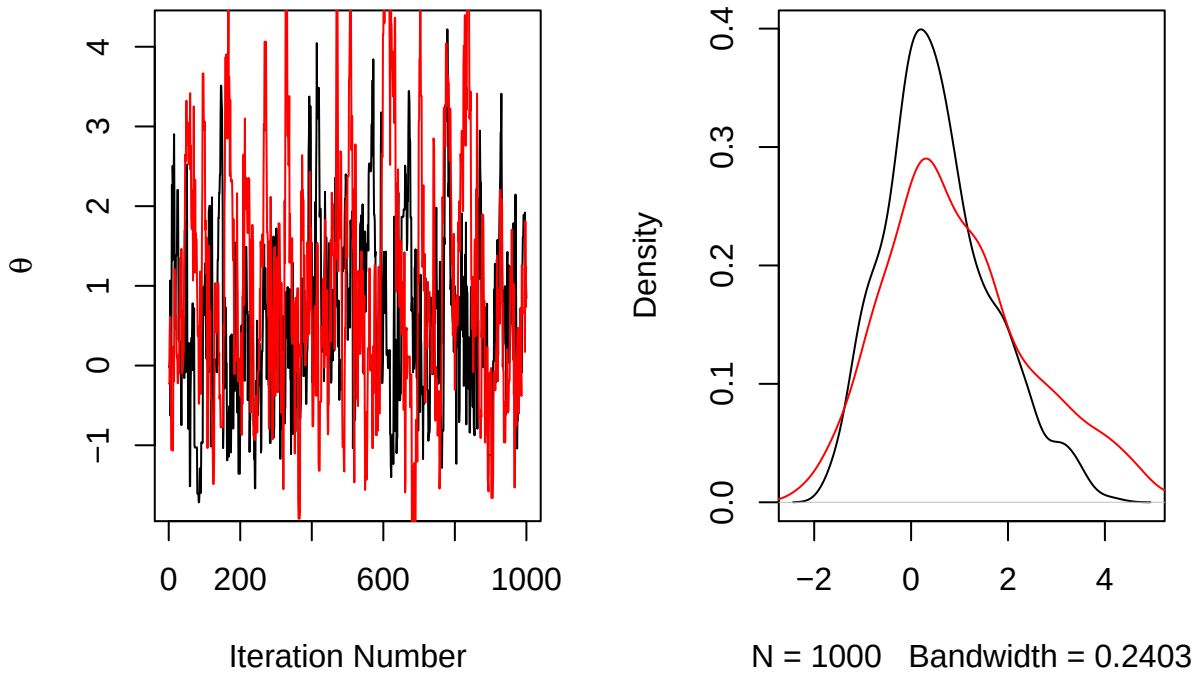
- Last week was the exception in that I did not provide course materials in R Markdown... today that changes
- R Markdown is a form of the Markdown language (https://en.wikipedia.org/wiki/Markdown)
  - Markdown was the counter to HTML (Hypertext Markup Language)... but is a markup language that is very easy to use
  - You can find Markdown nearly everywhere these days (see your Notes/OneNote application)
- Markdown makes writing very simple:
  - It works nearly everywhere (files are simple text)
  - It can incorporate more complicated markup languages, such as LaTex: $P(\theta|Y) \propto P(Y|\theta)\,P(\theta)$
- When you do need a type of document, you can then use any number of programs to make it look nice:
  - Pandoc (https://pandoc.org; converts to Word, PDF, LaTex, etc...)
  - The papaja R package (https://crsh.github.io/papaja_man/)
- R Markdown allows you to embed R script within the document, providing syntax snippets and output directly to your final document format
- You can find lots of helpful tips on R Markdown on some of these sites:
  - https://rmarkdown.rstudio.com
  - https://www.rstudio.com/wp-content/uploads/2015/02/rmarkdown-cheatsheet.pdf

## R Markdown Example

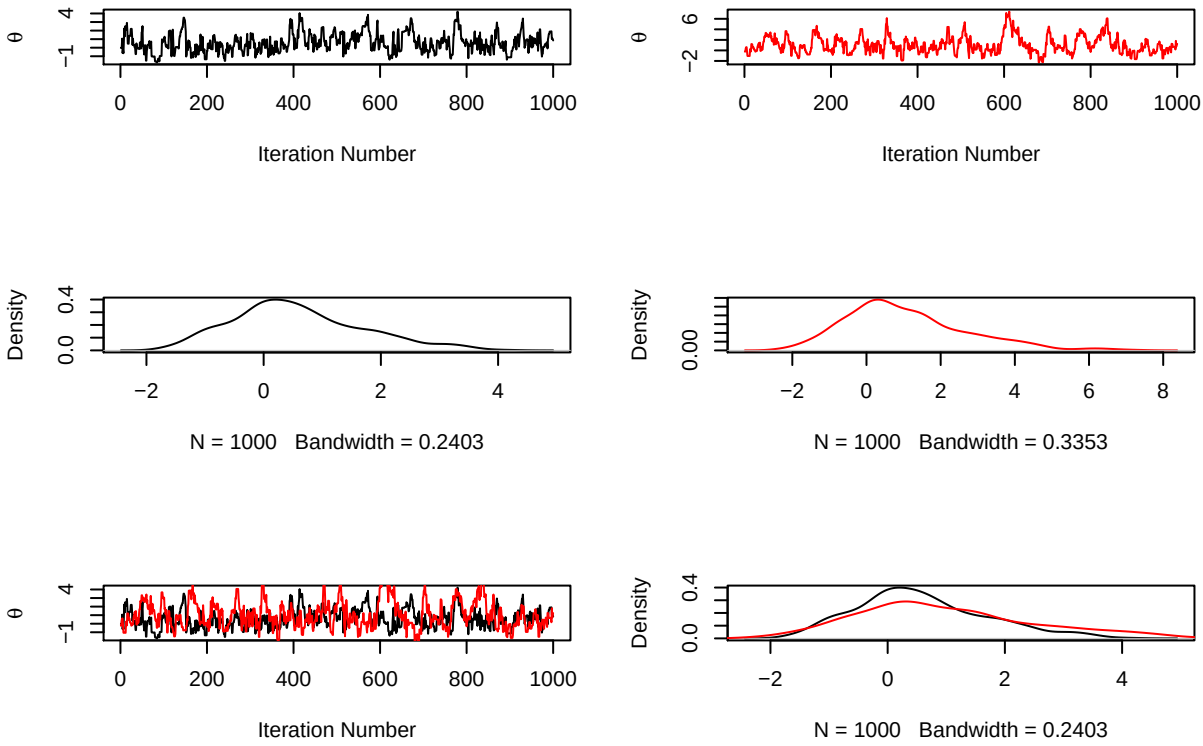
- Recalling last week's R script for theta, below is how to embed it in R Markdown
  - Note: I've changed the number of iterations to something very small to make it run fast

```
irtItemProb = function(a, b, c=0, theta){
  prob = c + (1-c) * exp(a*(theta-b))/(1+exp(a*(theta-b)))
  return(prob)
}

trueTheta = 0
nItems = 5
nItems = 5
bRange = c(-2,2)
aRange = c(1,2)
bSE = 1
aSE = 1
nSamples = 1000

# draw mean values of a, b
a = runif(n = nItems, min = aRange[1], max = aRange[2])
```

```
b = runif(n = nItems, min = bRange[1], max = bRange[2])

# draw items
itemResponses = rbinom(n = nItems, size = 1, prob = irtItemProb(a = a, b = b, theta = 1))
thetaChain = list(rep(NA, nSamples), rep(NA, nSamples))

# initialize theta values
curTheta = trueTheta
curThetaRand = trueTheta
for (iteration in 1:nSamples){

  # draw item parameters (if random)
  iterA = rnorm(n = nItems, mean = a, sd = aSE)
  iterB = rnorm(n = nItems, mean = b, sd = bSE)

  # calculate current likelihood of the data | theta
  curLogLike = sum(dbinom(x = itemResponses, size = 1, prob = irtItemProb(a = a, b = b, theta = curTheta
  curLogLikeRand = sum(dbinom(x = itemResponses, size = 1, prob = irtItemProb(a = iterA, b = iterB, the

  # draw new theta value
  propTheta = rnorm(n = 1, mean = curTheta, sd = 1)
  propThetaRand = rnorm(n = 1, mean = curThetaRand, sd = 1)

  # calculate proposed likelihood of the data | theta
  propLogLike = sum(dbinom(x = itemResponses, size = 1, prob = irtItemProb(a = a, b = b, theta = propTh
  propLogLikeRand = sum(dbinom(x = itemResponses, size = 1, prob = irtItemProb(a = iterA, b = iterB, th

  # do MH:
  if (log(runif(n = 1)) < (propLogLike-curLogLike)){
    # accept
    curTheta = propTheta
  }

  # do MH:
  if (log(runif(n = 1)) < (propLogLikeRand-curLogLikeRand)){
    # accept
    curThetaRand = propThetaRand
  }

  thetaChain[[1]][iteration] = curTheta
  thetaChain[[2]][iteration] = curThetaRand
}

par(mfrow = c(1,2))

plot(thetaChain[[1]], type="l", ylab = expression(theta), xlab = "Iteration Number")
lines(thetaChain[[2]], type="l", col = 2)
plot(density(thetaChain[[1]]), col = 1, main="")
lines(density(thetaChain[[2]]), col = 2)
```




```
par(mfrow = c(3,2))
plot(thetaChain[[1]], type="l", ylab = expression(theta), xlab = "Iteration Number")
plot(thetaChain[[2]], type="l", ylab = expression(theta), xlab = "Iteration Number", col =2)
plot(density(thetaChain[[1]]), col = 1, main="")
plot(density(thetaChain[[2]]), col = 2, main="")
plot(thetaChain[[1]], type="l", ylab = expression(theta), xlab = "Iteration Number")
lines(thetaChain[[2]], type="l", col = 2)
plot(density(thetaChain[[1]]), col = 1, main="")
lines(density(thetaChain[[2]]), col = 2)
```







## More R Markdown

- To compile the whole document (called "Knitting" as it uses a package named knitr), press the Knit button or use the keystroke command-shift-K
- To run a chunk of R code, find and press the button on the top right of the chunk
- Note: If chunks later in the document depend on chunks at the beginning, you will have to run the beginning ones first (chunks use the current Global R Environment for variables and functions)