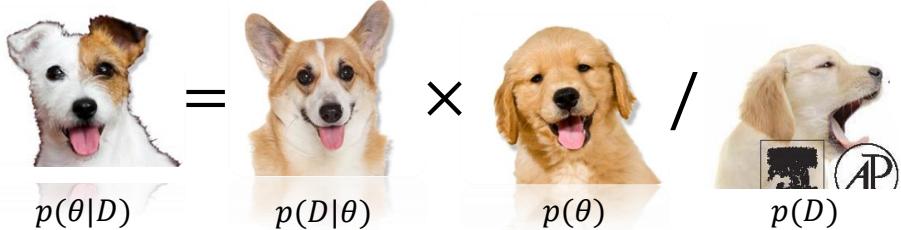


Doing Bayesian Data Analysis



John K. Kruschke

Outline of Talk:

- Bayesian reasoning generally.
- Bayesian estimation applied to two groups. Rich information.
 - Running the programs, and the flexibility of Bayesian software.
- The NHST t test: perfidious p values and the con game of confidence intervals.
- Conclusion: Bayesian estimation supersedes NHST.

Bayesian Reasoning

The role of data is to re-allocate credibility:

Prior Credibility with New Data
→ Posterior Credibility

via Bayes' rule

© John K. Kruschke, 2013

3

Bayesian Reasoning

The role of data is to re-allocate credibility:

Bayesian reasoning in everyday life is
intuitive:

© John K. Kruschke, 2013

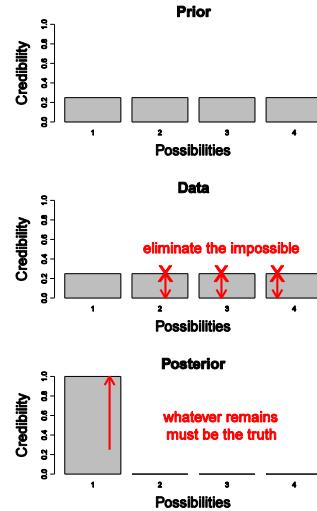
4

Bayesian Reasoning

The role of data is to re-allocate credibility:

Bayesian reasoning in everyday life is intuitive:

Sherlock Holmes: "How often have I said to you that when you have eliminated the impossible, whatever remains, however improbable, must be the truth?" (Doyle, 1890)



© John K. Kruschke, 2013

5

Bayesian Reasoning

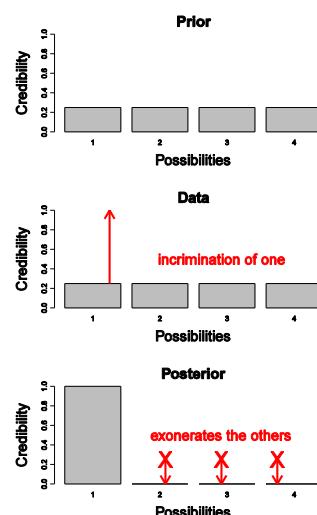
The role of data is to re-allocate credibility:

Bayesian reasoning in everyday life is intuitive:

Sherlock Holmes: "How often have I said to you that when you have eliminated the impossible, whatever remains, however improbable, must be the truth?" (Doyle, 1890)

Judicial exoneration: For unaffiliated suspects, the incrimination of one exonerates the others.

Credibility of the claim
that the suspect
committed the crime.



© John K. Kruschke, 2013

6

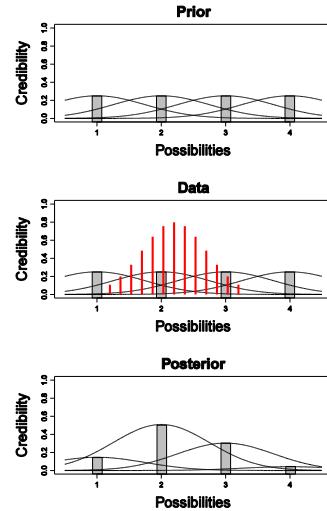
Bayesian Data Analysis

The role of data is to re-allocate credibility:

Bayesian reasoning in data analysis is intuitive:

Possibilities are *parameter values* in a model, such as the *mean* of a normal distribution.

We reallocate credibility to parameter values that are consistent with the data.



© John K. Kruschke, 2013

7

Bayesian Data Analysis

The role of data is to re-allocate credibility:

1. Define a meaningful descriptive model.
2. Establish prior credibility regarding parameter values in the model. The prior credibility must be acceptable to a skeptical scientific audience.
3. Collect data.
4. Use Bayes' rule to re-allocate credibility to parameter values that are most consistent with the data.

© John K. Kruschke, 2013

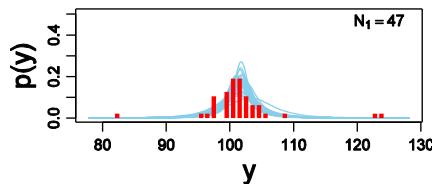
8

Robust Bayesian estimation for comparing two groups

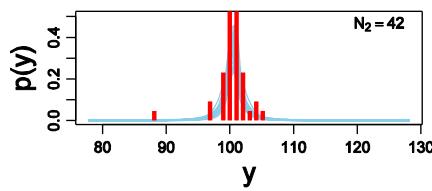
Consider two groups;
e.g.,
IQ of “smart drug” group
and of control group.

Step 1: Define a model
for describing the data.

Data Group 1 w. Post. Pred.



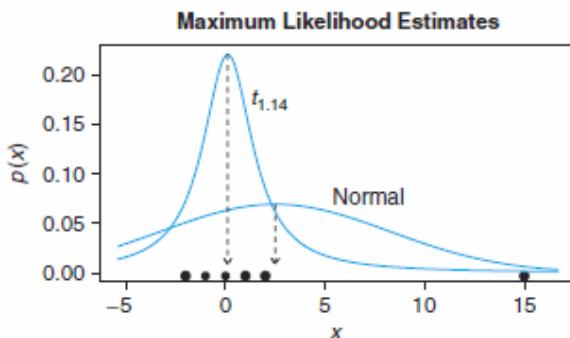
Data Group 2 w. Post. Pred.



© John K. Kruschke, 2013

0

Descriptive distribution for data with outliers



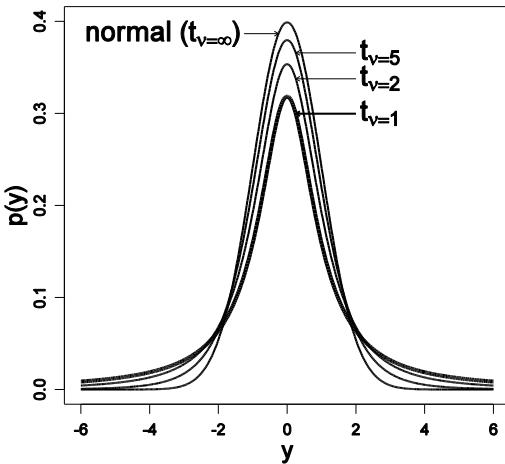
Normal is
pulled by
outliers, but t
distribution is
not.

t distribution is used here as a description of data,
NOT as a sampling distribution for p values!

© John K. Kruschke, 2013

10

Descriptive distribution for data with outliers

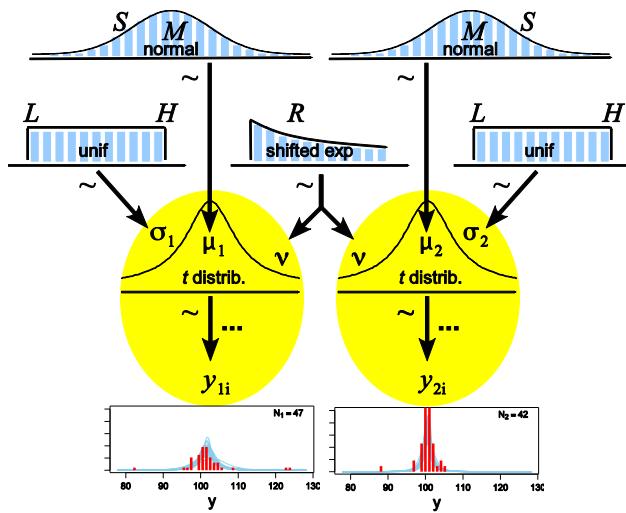


The t distribution has normality controlled by the parameter v .

© John K. Kruschke, 2013

12

Robust Bayesian estimation for comparing two groups

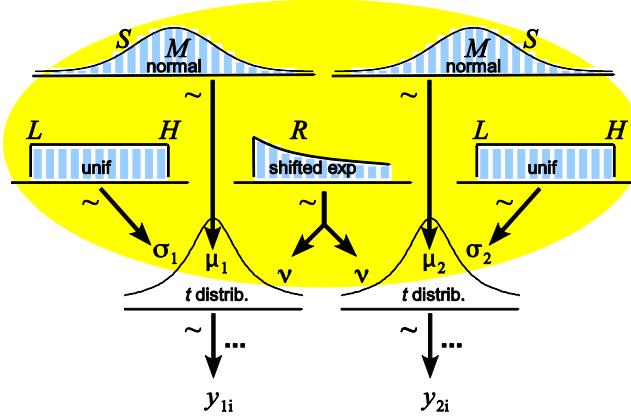


The data from each group are described by t distributions, using five parameters altogether.

© John K. Kruschke, 2013

13

Robust Bayesian estimation for comparing two groups

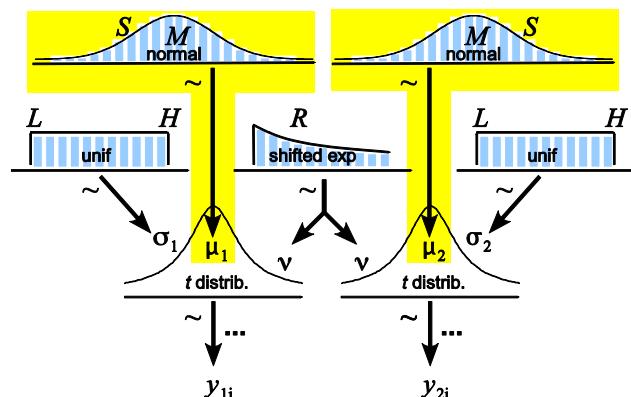


Step 2: Specify
the prior.

© John K. Kruschke, 2013

14

Robust Bayesian estimation for comparing two groups

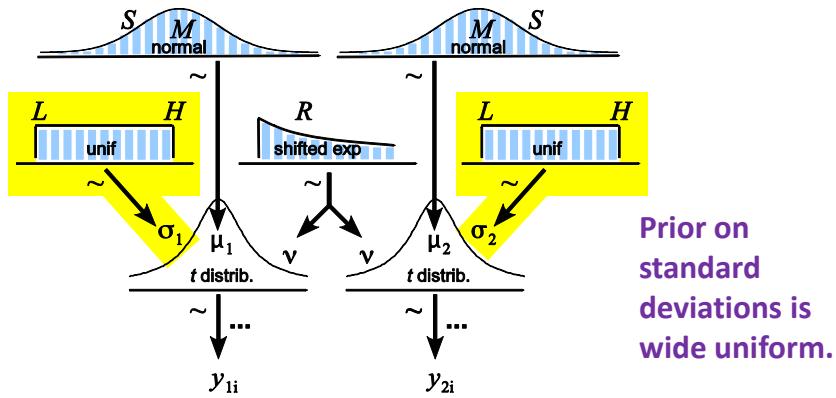


Prior on means
is wide normal.

© John K. Kruschke, 2013

15

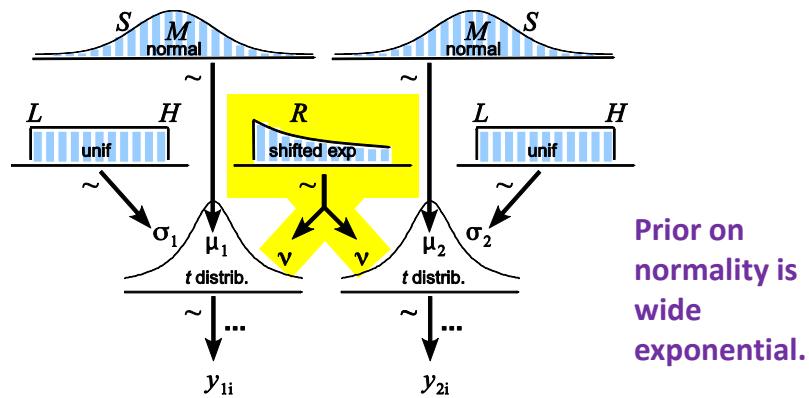
Robust Bayesian estimation for comparing two groups



© John K. Kruschke, 2013

16

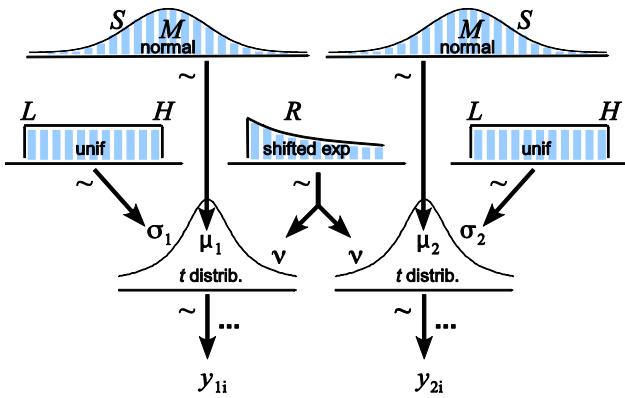
Robust Bayesian estimation for comparing two groups



© John K. Kruschke, 2013

17

Robust Bayesian estimation for comparing two groups

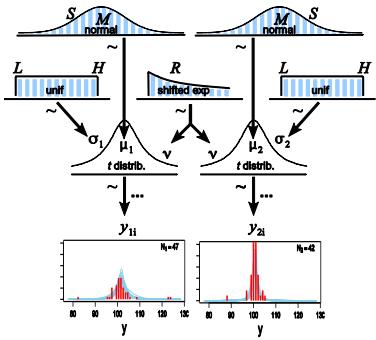


Parameter distributions will be represented by histograms: A huge number of representative parameter values.

© John K. Kruschke, 2013

18

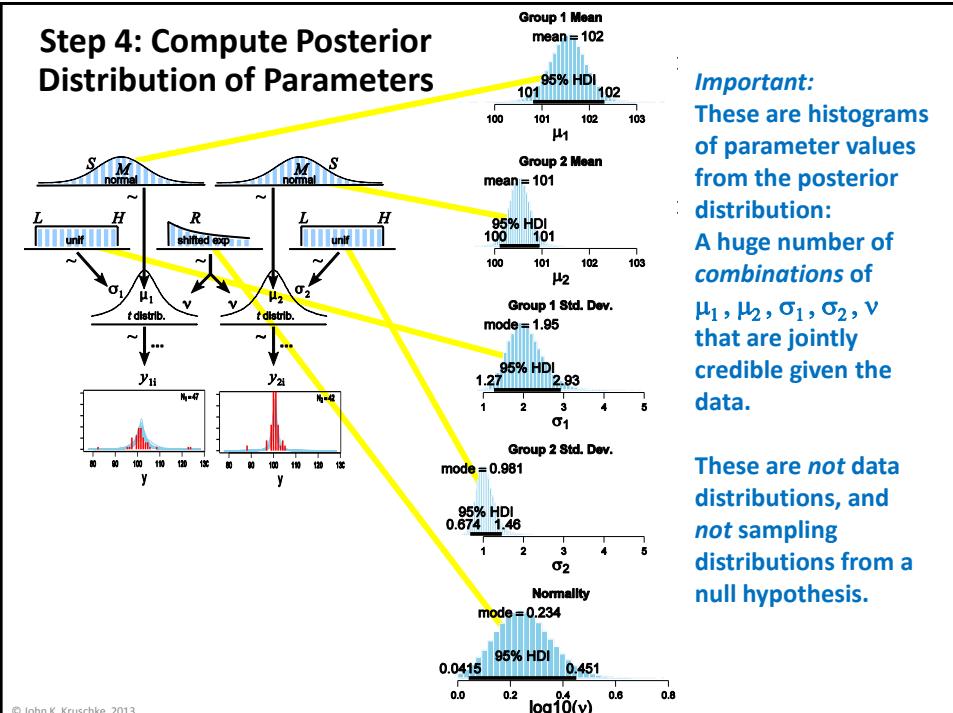
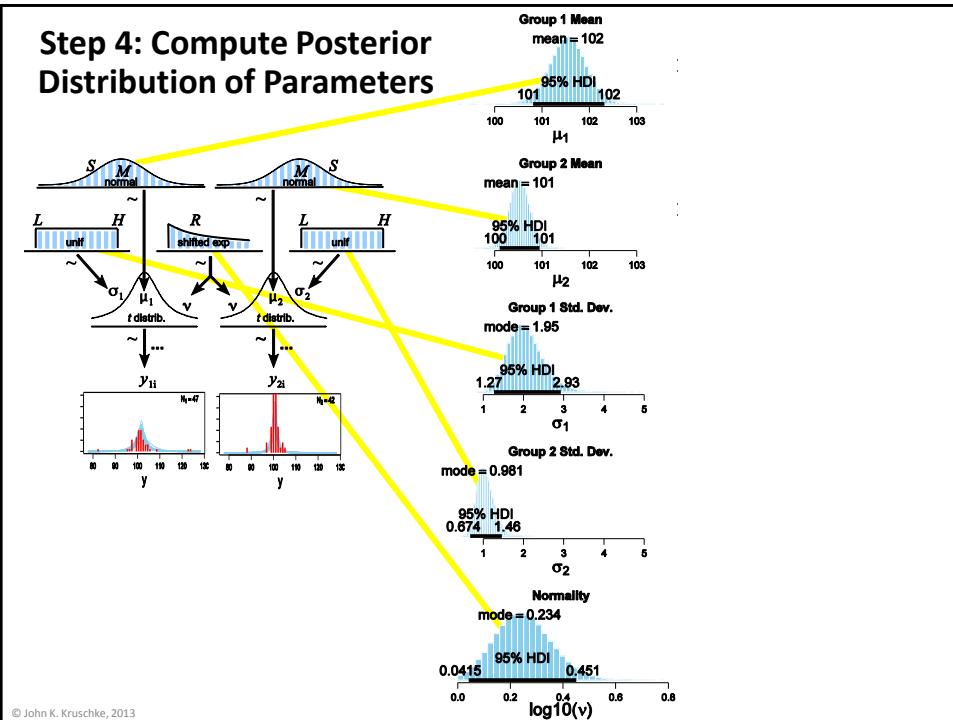
Step 3: Collect Data.



One fixed data set,
shown as red
histograms.

© John K. Kruschke, 2013

19

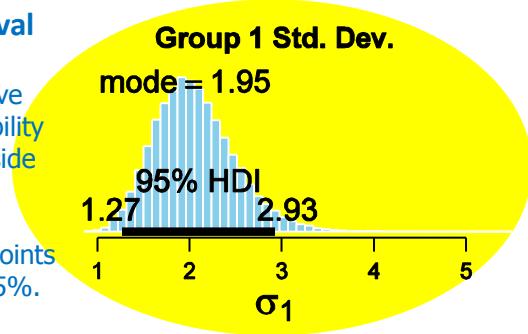


95% HDI: Highest density interval

Points within the HDI have higher credibility (probability density) than points outside the HDI.

The total probability of points within the 95% HDI is 95%.

Points outside the HDI may be deemed not credible.



© John K. Kruschke, 2013

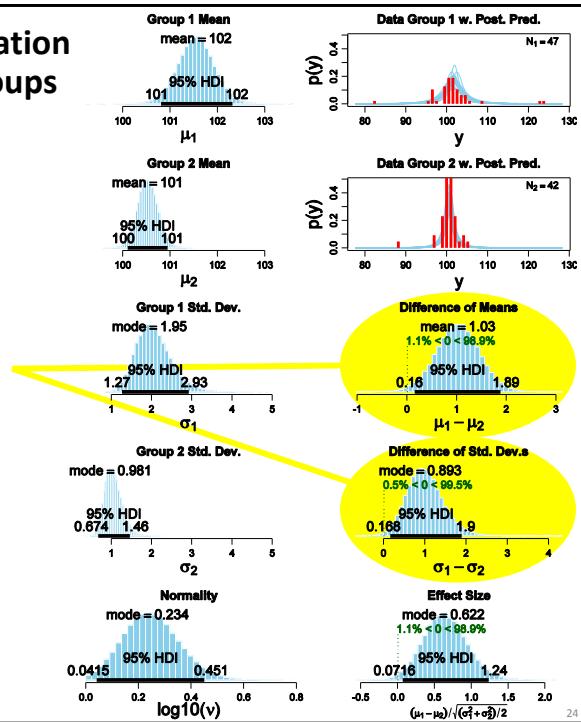
22

Robust Bayesian estimation for comparing two groups

Differences between groups?
Compute $\mu_1 - \mu_2$ and $\sigma_1 - \sigma_2$ at each of the many credible combinations.

Here, both differences are credibly non-zero.

(NHST would require two tests...)



© John K. Kruschke, 2013

24

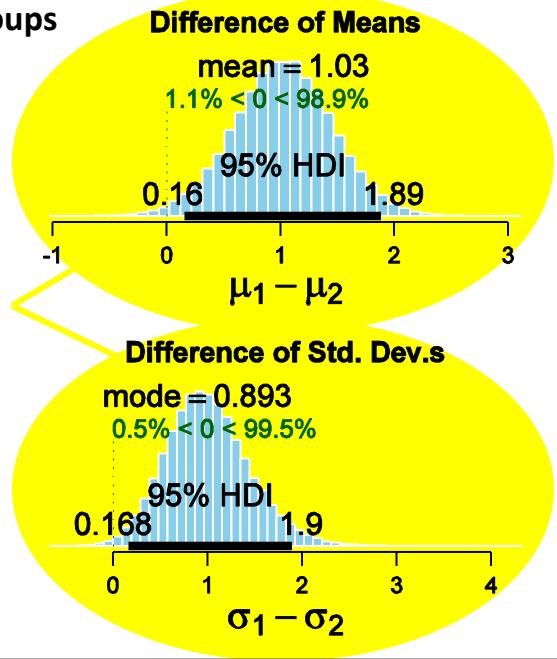
Robust Bayesian estimation for comparing two groups

Differences between
groups?

Compute $\mu_1 - \mu_2$
and $\sigma_1 - \sigma_2$
at each of the many
credible combinations.

Here, both differences
are credibly non-zero.

(NHST would require
two tests...)

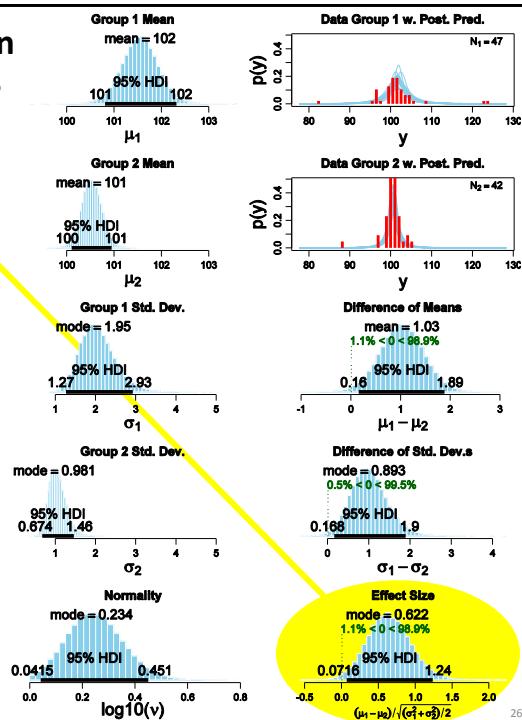


© John K. Kruschke, 2013

25

Robust Bayesian estimation for comparing two groups

Complete distribution
on effect size!

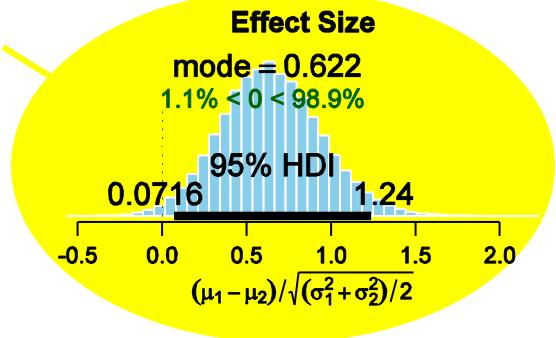


© John K. Kruschke, 2013

26

Robust Bayesian estimation for comparing two groups

Complete distribution
on effect size!



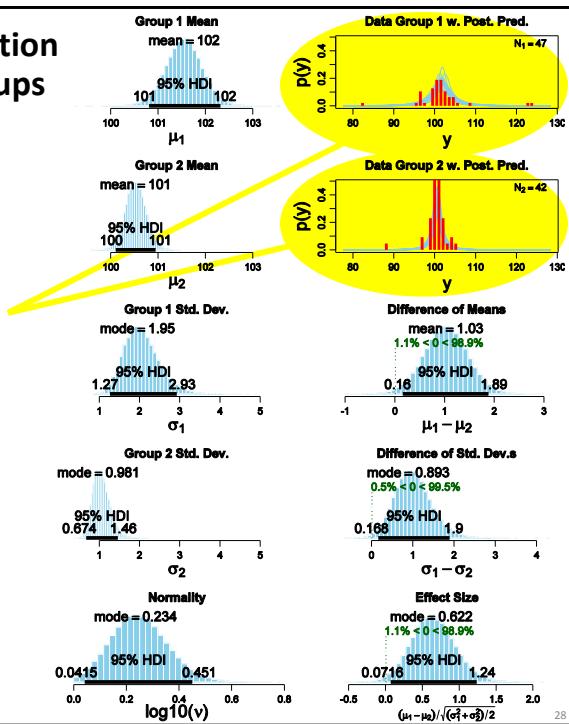
© John K. Kruschke, 2013

27

Robust Bayesian estimation for comparing two groups

Are the data described
well by the model?

Superimpose a
smattering of credible
descriptive distributions
on data.
= “posterior predictive
check”



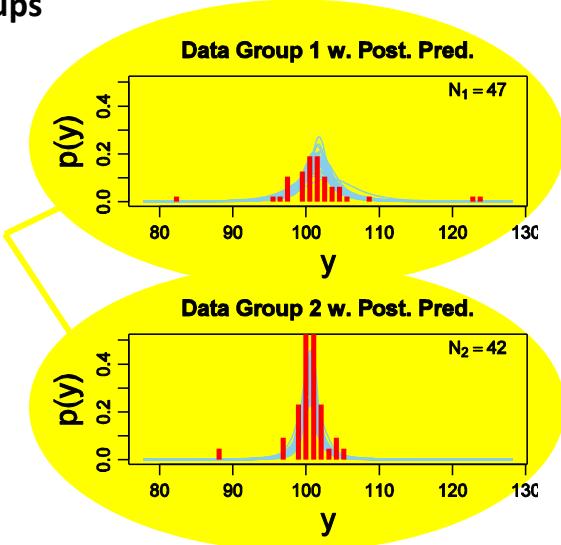
© John K. Kruschke, 2013

28

Robust Bayesian estimation for comparing two groups

**Are the data described
well by the model?**

Superimpose a
smattering of credible
descriptive distributions
on data.
= “posterior predictive
check”

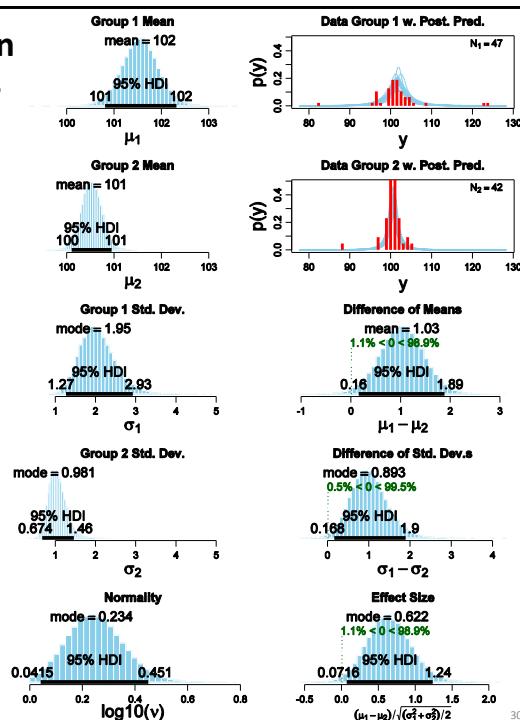


© John K. Kruschke, 2013

20

Robust Bayesian estimation for comparing two groups

Summary:
**→ Complete distribution
of credible parameter
values** (not merely point
estimate with ends of
confidence interval).
**→ Decisions about
multiple aspects of
parameters** (without
reference to p values).
**→ Flexible descriptive
model, robust to outliers**
(unlike NHST t test).



© John K. Kruschke, 2013

30

Computer Software:

Packaged for easy use!
Underlying program is never seen.

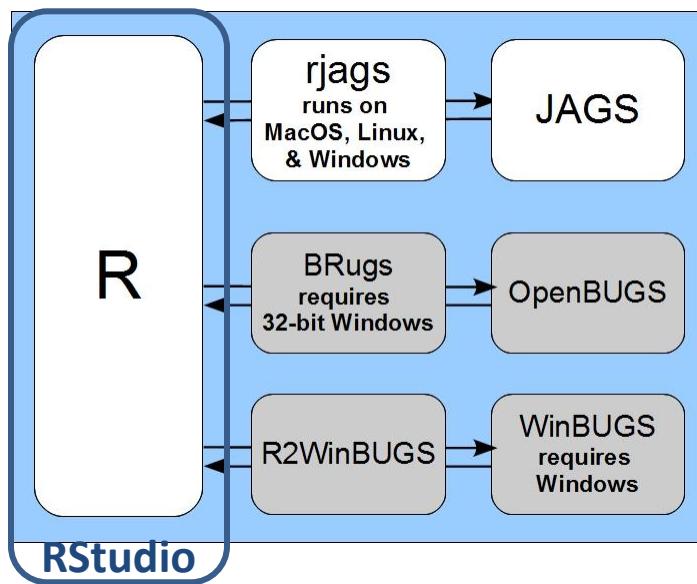
```
source("BEST.R") # load the program

# Specify data as vectors (replace with your own data):
y1 = c(101,100,102,104,102,97,105,105,98,...,101)
y2 = c(99,101,100,101,102,100,97,101,104,...,99)

# Run the Bayesian analysis:
mcmcChain = BESTmcmc( y1 , y2 )

# Plot the results of the Bayesian analysis:
BESTplot( y1 , y2 , mcmcChain )
```

Software: R and JAGS



<http://www.indiana.edu/~kruschke/BEST/>

Bayesian estimation supersedes the t test.

John K. Kruschke, Journal of Experimental Psychology: General

Abstract:
Bayesian estimation for two groups provides complete distributions of credible values for the effect size, group means and their difference, standard deviations and their difference, and the normality of the data. The method handles outliers. The decision rule can accept the null value (unlike traditional *t* tests) when certainty in the estimate is high (unlike Bayesian model comparison using Bayes factors). The method also yields precise estimates of statistical power for various research goals. The software and programs are free, and run on Macintosh, Linux, and Windows platforms.

The article:
Get the article [here](#), or at doi: [10.1037/a0029146](https://doi.org/10.1037/a0029146).
*Your click on this link constitutes your request to me for a personal copy of the linked article, and my delivery of a personal copy. Any other use is prohibited.

Videos:
Watch the [video](#), and this [additional video](#) that includes discussion of sequential testing. (Both presented at the 2012 Psychonomic Society meeting.)

Software (named "BEST" for "Bayesian estimation"):
There is (as of January 3, 2013) a [web app](#) created by Rasmus Bååth that provides the primary results of BEST without any need to install software. Just paste in your data and click a button. See the [blog post](#).

Get the [BEST programs from this zip file](#). Updated January 23, 2013. Be sure to unzip (extract) the zip file after it is saved on your computer. For information about running the programs, see the comments at the top of the file `BESTexample.R`.

The programs now (as of Sept. 3, 2012) include a version for estimating the parameters of a single group. The 1-group version is named `BEST1G.R`. See the example in `BEST1Gexample.R` and [this blog post](#).

For information about installing the software (R, JAGS, rjags, and RStudio), see the comments at the top of the file `BESTexample.R`, or at [this blog post](#).

Learn more:
For a complete tutorial about Bayesian methods, see [the book](#).

You can make comments at [the blog](#). Search the blog with the phrase "Bayesian estimation".

The posterior distributions showing complete distributions of the difference of means (right middle), the difference of standard deviations (right bottom), and posterior predictive check (right upper).

33

<http://doingbayesiandataanalysis.blogspot.com/2012/01/complete-steps-for-installing-software.html>

Doing Bayesian Data Analysis

Software Installation Exercise Solutions Workshops Book Reviews Humor

Saturday, January 28, 2012

Complete steps for installing software and programs

[Updated February 04, 2013]

To use the programs, there are six basic steps:

1. Install the general programming language R. Go to the [R web site](#) and install the latest version of R appropriate for your computer (Windows, Mac, Linux).
2. Install the R editor RStudio. R comes with its own built-in editor, but it is not very useful for dealing with long programs. Go to the [RStudio web site](#) and install the latest version of RStudio appropriate for your computer (Windows, Mac, Linux). This is optional, but highly recommended.
3. Install the Bayesian sampling program JAGS. Go to the [JAGS web site](#) and install the latest version of JAGS appropriate for your computer (Windows, Mac, Linux).
4. Install the package that lets R talk to JAGS. Invoke RStudio. Click the "Tools" menu and then "Install Packages." In the Packages slot type "jags" without the quotes, or, if a menu of packages appears, select jags. Then click install. (Alternatively, if working in R not RStudio, at the command line, type `install.packages("jags")`) You may be prompted to select an internet archive to get the package from; select a site geographically near you.)
5. Get the programs used in the book. For a zip file (named `ProgramsDoingBayesianDataAnalysis.zip`) that contains all the programs in a single file, click [HERE](#). Be sure to unzip (extract) the programs from the zip file.
6. Read the important tips for running the programs, below! Really, read them. Remember them. It'll save you frustration later.

Important tips for running the programs:

- I now recommend using JAGS instead of BUGS, and that's why the installation instructions (above) only mention JAGS, not BUGS. Any program that had "Bugs" or "Brugs" in its file name now has a JAGS equivalent with "Jags" in its file name. More info is available at [this blog post](#).
- Many of the programs call other programs when running, and write auxiliary files, so
 - put all the programs together in the same folder (they come together in one folder, so it's easiest to leave them there)
 - put that folder under a `writeable` directory (not in a protected system folder), such as where you would store ordinary research data, and
 - be sure that R has that folder as its working directory. Set the working directory in RStudio from the menu: Tools -> Set Working Dir. Or, if you're using R's built-in editor, set the working directory in R's

Search This Blog (forum search is separate)

Search

> DISCUSSION FORUM <

> Arrange a Workshop at Your University or Organization <-

The Book (click it for info):

Blog Archive

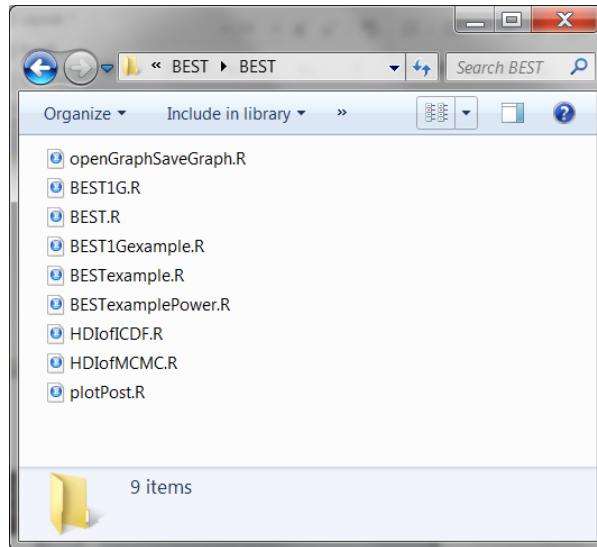
- 2013 (15)
- 2012 (57)
- December (1)
- November (3)
- October (2)
- September (4)
- August (4)
- July (2)

34

Example of running BEST

From your computer desktop, open the folder with the BEST files in it.

(Windows is shown; do equivalent in MacOS.)



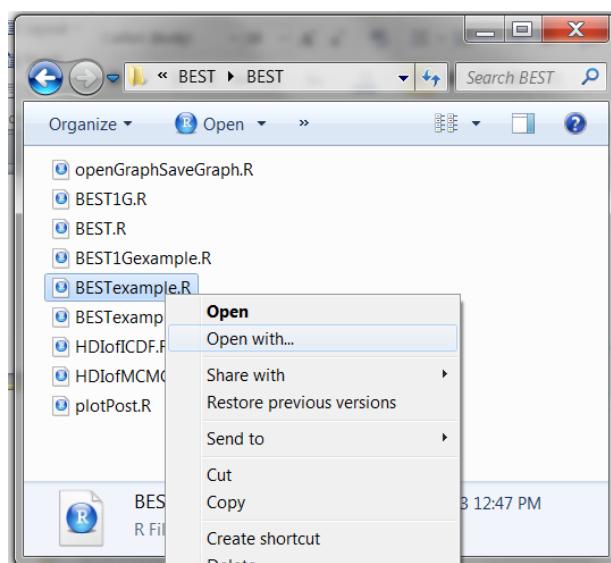
© John K. Kruschke, 2013

35

Example of running BEST

Select
BESTexample.R
and
Open with...

(or equivalent in
MacOS)



© John K. Kruschke, 2013

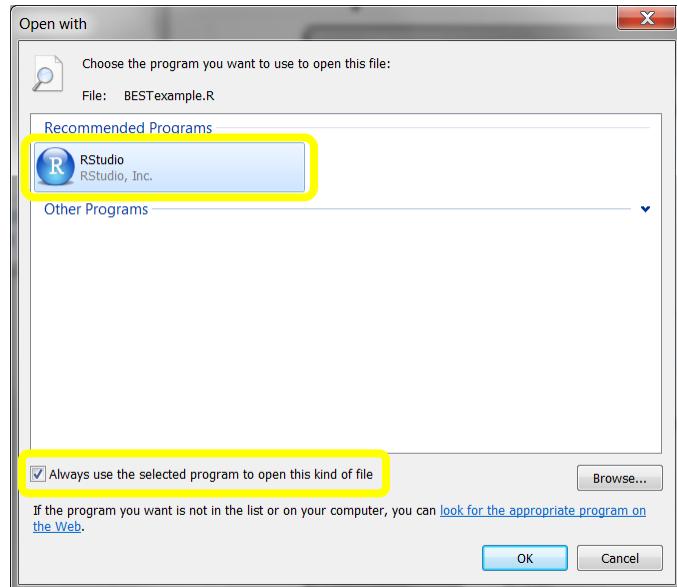
36

Example of running BEST

Browse to select RStudio
and be sure to check
“Always use...”

(or equivalent in
MacOS)

From now on, your
computer knows to
open .R files in
RStudio.



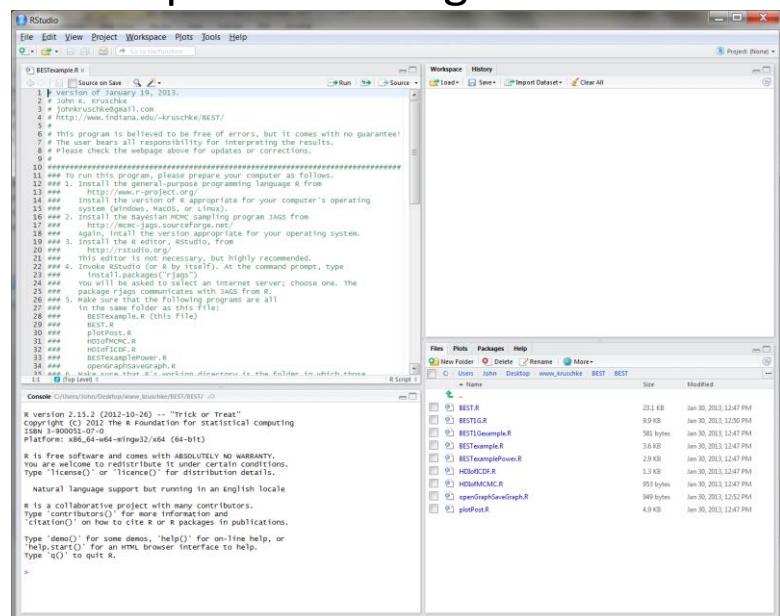
© John K. Kruschke, 2013

37

Example of running BEST

The program
then opens in
RStudio.

Notice the
four panes...



© John K. Kruschke, 2013

38

Example of running BEST

IMPORTANT:
Make sure that
the **working
directory** is
the folder with
BEST.R and
other
supporting
programs.

The screenshot shows the RStudio interface with the 'Tools' menu open. The 'Set Working Directory' option is highlighted, with a submenu showing 'To Source File Location', 'To Files Pane Location', and 'Choose Directory...'. The main menu bar includes 'File', 'Edit', 'View', 'Project', 'Workspace', 'Plots', 'Tools', and 'Help'. A code editor window titled 'BESTexample.R' is visible in the background.

```

1 # Version of January 19, 2013.
2 # John K. Kruschke
3 # johnkruschke@gmail.com
4 # http://www.indiana.edu/~kruschke
5 #
6 # This program is believed to be free of bugs.
7 # The user bears all responsibility for its use.
8 # Please check the webpage above if you have questions.
9 #
10 #####
11 ### To run this program, please prepare your computer as follows.
12 ## 1. Install R, a free, open-source programming language R from
13 ## http://www.r-project.org
14 ## Install the version of R appropriate for your computer's operating
15 ## system (Windows, Macos, or Linux).

```

The working directory is set automatically to the folder of the program if RStudio is invoked indirectly by clicking the program file (as we did).

© John K. Kruschke, 2013

39

Example of running BEST

In RStudio's
editing
window, scroll
down
BESTexample.R
to see these
lines:

The screenshot shows the RStudio interface with the code editor window titled 'BESTexample.R'. The window displays several lines of R code, starting with optional memory clearing and graphics closing, followed by loading functions from 'BEST.R', specifying data vectors, running Bayesian analysis using 'mcmcChain', plotting results with 'BESTplot', and saving data and results. The code editor has tabs for 'Run', 'Source', and 'Script'.

```

45 # OPTIONAL: Clear R's memory and graphics:
46 rm(list=ls()) # Careful! This clears all of R's memory!
47 graphics.off() # This closes all of R's graphics windows.
48
49 # Get the functions loaded into R's working memory:
50 source("BEST.R")
51
52 # Specify data as vectors:
53 # (Replace with your own data as needed. R can read many formats of data files,
54 # see the commands "scan" or "read.table" and its variants.)
55 y1 = c(101,100,102,104,102,97,101,100,123,105,103,100,95,102,106,
56 109,102,82,102,100,102,101,102,102,103,103,97,97,103,101,97,104,
57 96,103,124,101,101,100,101,101,104,100,101)
58 y2 = c(99,101,100,101,102,100,97,101,104,101,102,102,100,105,88,101,100,
59 104,100,100,101,102,103,97,101,101,100,101,99,101,100,100,
60 101,100,99,101,100,102,99,100,99)
61
62 # Run the Bayesian analysis:
63 mcmcChain = BESTmcmc( y1 , y2 )
64
65 # Plot the results of the Bayesian analysis:
66 postInfo = BESTplot( y1 , y2 , mcmcchain , pairsPlot=TRUE )
67 # Show detailed summary info on console:
68 show( postInfo )
69 # You can save the plot(s) using the pull-down menu in the R graphics window,
70 # or by using the following:
71 # saveGraph( file="BESTexample" , type="eps" )
72 # saveGraph( file="BESTexample" , type="jpeg" )
73
74 # Save the data and results for future use:
75 save( y1 , mcmcChain , postInfo , file="BESTexampleMCMC.Rdata" )
76 # To re-load the saved data and MCMC chain, type: load( "BESTexampleMCMC.Rdata"
77 )
78
79

```

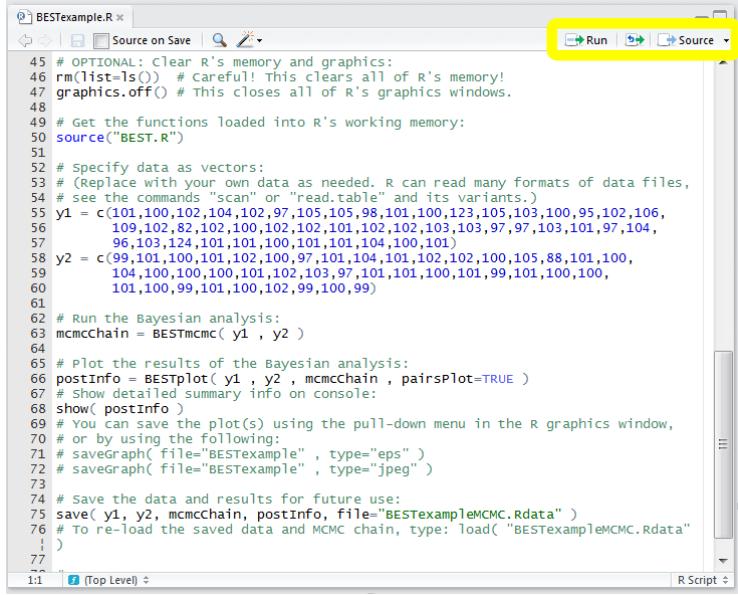
© John K. Kruschke, 2013

40

Example of running BEST

The Run button executes the line of code of the current cursor position, or whatever lines are selected (i.e., highlighted).

The Source button executes all lines.



```

@ BESTexample.R x | Source on Save | Run | Source
45 # OPTIONAL: Clear R's memory and graphics:
46 rm(list=ls()) # Careful! This clears all of R's memory!
47 graphics.off() # This closes all of R's graphics windows.
48
49 # Get the functions loaded into R's working memory:
50 source("BEST.R")
51
52 # Specify data as vectors:
53 # (Replace with your own data as needed. R can read many formats of data files,
54 # see the commands "scan" or "read.table" and its variants.)
55 y1 = c(101,100,102,104,102,97,105,105,98,101,100,123,105,103,100,95,102,106,
56 109,102,82,102,100,102,101,102,102,103,103,97,97,103,101,97,104,
57 96,103,124,101,101,100,101,104,100,101)
58 y2 = c(99,101,100,101,102,100,97,101,104,101,102,102,100,105,88,101,100,
59 104,100,100,100,101,102,103,97,101,101,100,101,99,101,100,100,
60 101,100,99,101,100,102,99,100,99)
61
62 # Run the Bayesian analysis:
63 mcmcChain = BESTmcmc( y1 , y2 )
64
65 # Plot the results of the Bayesian analysis:
66 postInfo = BESTplot( y1 , y2 , mcmcChain , pairsPlot=TRUE )
67 # Show detailed summary info on console:
68 show( postInfo )
69 # You can save the plot(s) using the pull-down menu in the R graphics window,
70 # or by using the following:
71 # saveGraph( file="BESTexample" , type="eps" )
72 # saveGraph( file="BESTexample" , type="jpeg" )
73
74 # Save the data and results for future use:
75 save( y1 , y2 , mcmcChain , postInfo , file="BESTexampleMCMC.Rdata" )
76 # To re-load the saved data and MCMC chain, type: load( "BESTexampleMCMC.Rdata"
| )
77
78
1:1 [ Top Level ] R Script

```

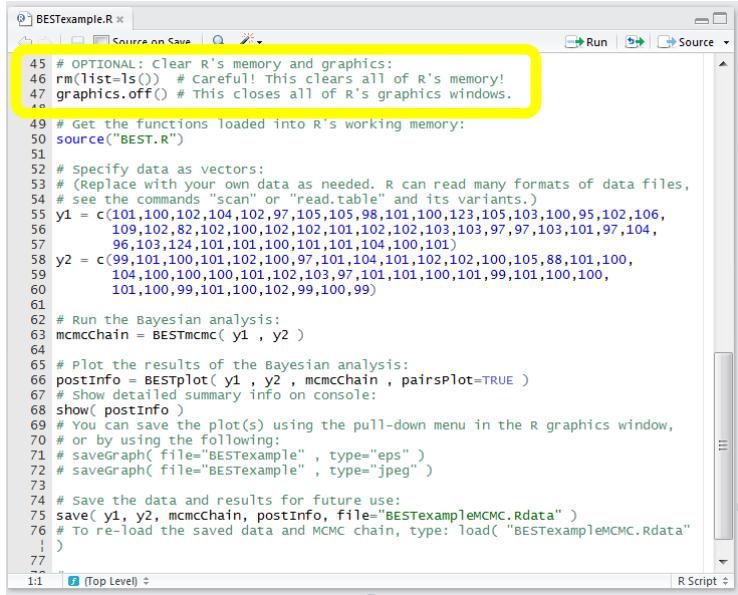
© John K. Kruschke, 2013

41

Example of running BEST

Optional clearing of R's memory and graphics.

indicates comment.



```

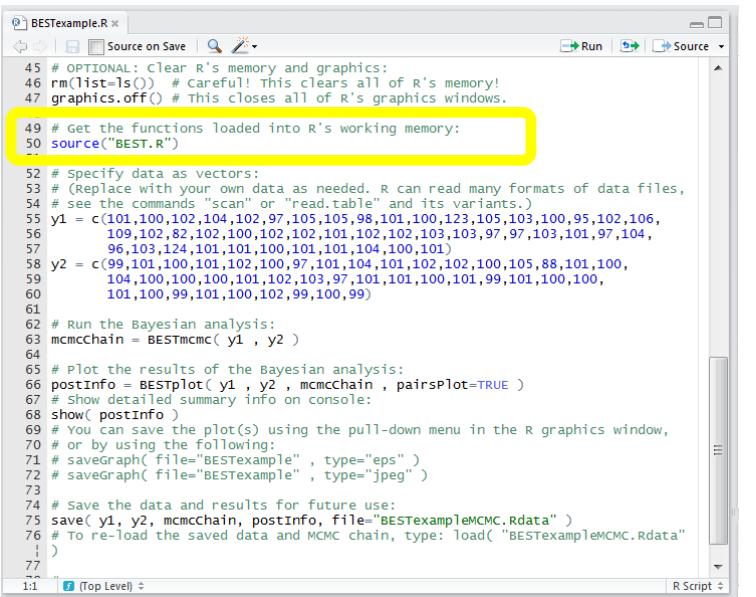
@ BESTexample.R x | Source on Save | Run | Source
45 # OPTIONAL: Clear R's memory and graphics:
46 rm(list=ls()) # Careful! This clears all of R's memory!
47 graphics.off() # This closes all of R's graphics windows.
48
49 # Get the functions loaded into R's working memory:
50 source("BEST.R")
51
52 # Specify data as vectors:
53 # (Replace with your own data as needed. R can read many formats of data files,
54 # see the commands "scan" or "read.table" and its variants.)
55 y1 = c(101,100,102,104,102,97,105,105,98,101,100,123,105,103,100,95,102,106,
56 109,102,82,102,100,102,101,102,102,103,103,97,97,103,101,97,104,
57 96,103,124,101,101,100,101,104,100,101)
58 y2 = c(99,101,100,101,102,100,97,101,104,101,102,102,100,105,88,101,100,
59 104,100,100,100,101,102,103,97,101,101,100,101,99,101,100,100,
60 101,100,99,101,100,102,99,100,99)
61
62 # Run the Bayesian analysis:
63 mcmcChain = BESTmcmc( y1 , y2 )
64
65 # Plot the results of the Bayesian analysis:
66 postInfo = BESTplot( y1 , y2 , mcmcChain , pairsPlot=TRUE )
67 # Show detailed summary info on console:
68 show( postInfo )
69 # You can save the plot(s) using the pull-down menu in the R graphics window,
70 # or by using the following:
71 # saveGraph( file="BESTexample" , type="eps" )
72 # saveGraph( file="BESTexample" , type="jpeg" )
73
74 # Save the data and results for future use:
75 save( y1 , y2 , mcmcChain , postInfo , file="BESTexampleMCMC.Rdata" )
76 # To re-load the saved data and MCMC chain, type: load( "BESTexampleMCMC.Rdata"
| )
77
78
1:1 [ Top Level ] R Script

```

© John K. Kruschke, 2013

42

Example of running BEST



```

@ BESTexample.R *
Source on Save | Run | Source
45 # OPTIONAL: Clear R's memory and graphics:
46 rm(list=ls()) # Careful! This clears all of R's memory!
47 graphics.off() # This closes all of R's graphics windows.
48
49 # Get the functions loaded into R's working memory:
50 source("BEST.R")
51
52 # Specify data as vectors:
53 # (Replace with your own data as needed. R can read many formats of data files,
54 # see the commands "scan" or "read.table" and its variants.)
55 y1 = c(101,100,102,104,102,97,105,105,98,101,100,123,105,103,100,95,102,106,
56 109,102,82,102,100,102,101,102,102,103,103,97,97,103,101,97,104,
57 96,103,124,101,101,100,101,101,104,100,101)
58 y2 = c(99,101,100,101,102,100,97,101,104,101,102,102,100,105,88,101,100,
59 104,100,100,100,101,102,103,97,101,101,100,101,99,101,100,100,
60 101,100,99,101,100,102,99,100,99)
61
62 # Run the Bayesian analysis:
63 mcmcChain = BESTmcmc( y1 , y2 )
64
65 # Plot the results of the Bayesian analysis:
66 postInfo = BESTplot( y1 , y2 , mcmcChain , pairsPlot=TRUE )
67 # Show detailed summary info on console:
68 show( postInfo )
69 # You can save the plot(s) using the pull-down menu in the R graphics window,
70 # or by using the following:
71 # saveGraph( file="BESTexample" , type="eps" )
72 # saveGraph( file="BESTexample" , type="jpeg" )
73
74 # Save the data and results for future use:
75 save( y1 , y2 , mcmcChain , postInfo , file="BESTexampleMCMC.Rdata" )
76 # To re-load the saved data and MCMC chain, type: load( "BESTexampleMCMC.Rdata"
77 )
78
79
1:1 [ Top Level ] R Script

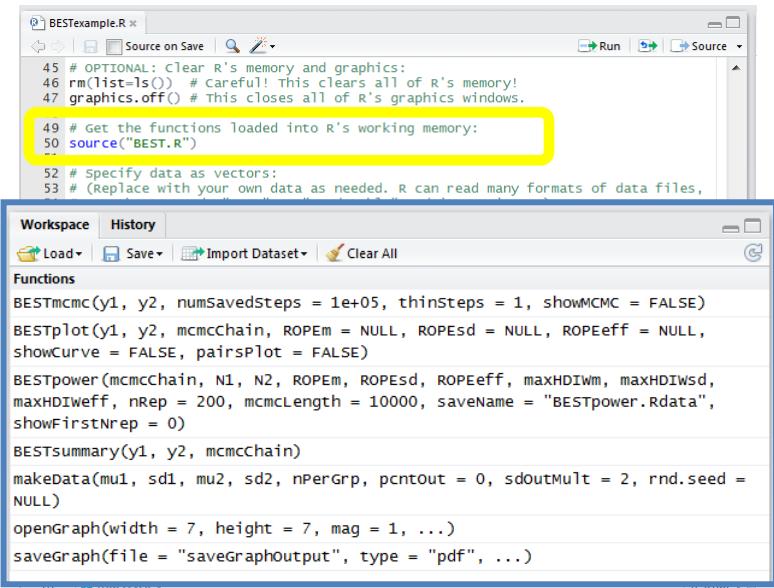
```

© John K. Kruschke, 2013

43

Example of running BEST

When executed, notice that functions appear in R's workspace, shown in another pane of RStudio.



```

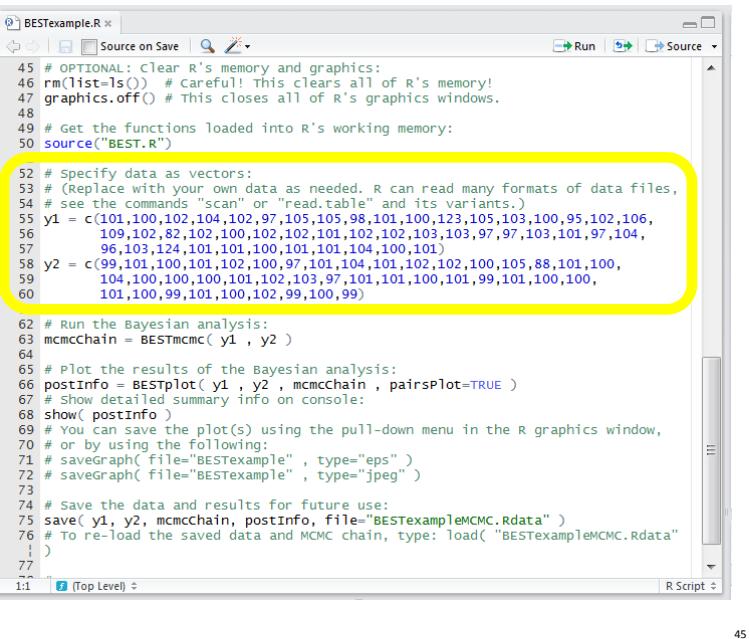
@ BESTexample.R *
Source on Save | Run | Source
45 # OPTIONAL: Clear R's memory and graphics:
46 rm(list=ls()) # Careful! This clears all of R's memory!
47 graphics.off() # This closes all of R's graphics windows.
48
49 # Get the functions loaded into R's working memory:
50 source("BEST.R")
51
52 # Specify data as vectors:
53 # (Replace with your own data as needed. R can read many formats of data files,
54 # see the commands "scan" or "read.table" and its variants.)
55 y1 = c(101,100,102,104,102,97,105,105,98,101,100,123,105,103,100,95,102,106,
56 109,102,82,102,100,102,101,102,102,103,103,97,97,103,101,97,104,
57 96,103,124,101,101,100,101,101,104,100,101)
58 y2 = c(99,101,100,101,102,100,97,101,104,101,102,102,100,105,88,101,100,
59 104,100,100,100,101,102,103,97,101,101,100,101,99,101,100,100,
60 101,100,99,101,100,102,99,100,99)
61
62 # Run the Bayesian analysis:
63 mcmcChain = BESTmcmc( y1 , y2 )
64
65 # Plot the results of the Bayesian analysis:
66 postInfo = BESTplot( y1 , y2 , mcmcChain , pairsPlot=TRUE )
67 # Show detailed summary info on console:
68 show( postInfo )
69 # You can save the plot(s) using the pull-down menu in the R graphics window,
70 # or by using the following:
71 # saveGraph( file="BESTexample" , type="eps" )
72 # saveGraph( file="BESTexample" , type="jpeg" )
73
74 # Save the data and results for future use:
75 save( y1 , y2 , mcmcChain , postInfo , file="BESTexampleMCMC.Rdata" )
76 # To re-load the saved data and MCMC chain, type: load( "BESTexampleMCMC.Rdata"
77 )
78
79
Workspace | History
Load | Save | Import Dataset | Clear All
Functions
BESTmcmc(y1, y2, numSavedSteps = 1e+05, thinSteps = 1, showMCMC = FALSE)
BESTplot(y1, y2, mcmcChain, ROPEm = NULL, ROPEsd = NULL, ROPEeff = NULL,
showCurve = FALSE, pairsPlot = FALSE)
BESTpower(mcmcChain, N1, N2, ROPEm, ROPEsd, ROPEeff, maxHDIwm, maxHDIwsd,
maxHDIwef, nRep = 200, mcmcLength = 10000, saveName = "BESTpower.Rdata",
showFirstNrep = 0)
BESTsummary(y1, y2, mcmcChain)
makeData(mu1, sd1, mu2, sd2, nPerGrp, pcntout = 0, sdoutMult = 2, rnd.seed =
NULL)
openGraph(width = 7, height = 7, mag = 1, ...)
saveGraph(file = "saveGraphoutput", type = "pdf", ...)

```

© John K. Kruschke, 2013

44

Example of running BEST



```

@ BESTexample.R *
Source on Save | Run | Source
45 # OPTIONAL: Clear R's memory and graphics:
46 rm(list=ls()) # Careful! This clears all of R's memory!
47 graphics.off() # This closes all of R's graphics windows.
48
49 # Get the functions loaded into R's working memory:
50 source("BEST.R")
51
52 # Specify data as vectors:
53 # (Replace with your own data as needed. R can read many formats of data files,
54 # see the commands "scan" or "read.table" and its variants.)
55 y1 = c(101,100,102,104,102,97,105,105,98,101,100,123,105,103,100,95,102,106,
56 109,102,82,102,100,102,101,102,102,103,103,97,97,103,101,97,104,
57 96,103,124,101,101,100,101,101,104,100,101)
58 y2 = c(99,101,100,101,102,100,97,101,104,101,102,102,100,105,88,101,100,
59 104,100,100,100,101,102,103,97,101,101,100,101,99,101,100,100,
60 101,100,99,101,100,102,99,100,99)

62 # Run the Bayesian analysis:
63 mcmcChain = BESTmcmc( y1 , y2 )
64
65 # Plot the results of the Bayesian analysis:
66 postInfo = BESTplot( y1 , y2 , mcmcChain , pairsPlot=TRUE )
67 # Show detailed summary info on console:
68 show( postInfo )
69 # You can save the plot(s) using the pull-down menu in the R graphics window,
70 # or by using the following:
71 # saveGraph( file="BESTexample" , type="eps" )
72 # saveGraph( file="BESTexample" , type="jpeg" )
73
74 # Save the data and results for future use:
75 save( y1 , y2 , mcmcChain , postInfo , file="BESTexampleMCMC.Rdata" )
76 # To re-load the saved data and MCMC chain, type: load( "BESTexampleMCMC.Rdata"
77 )
78

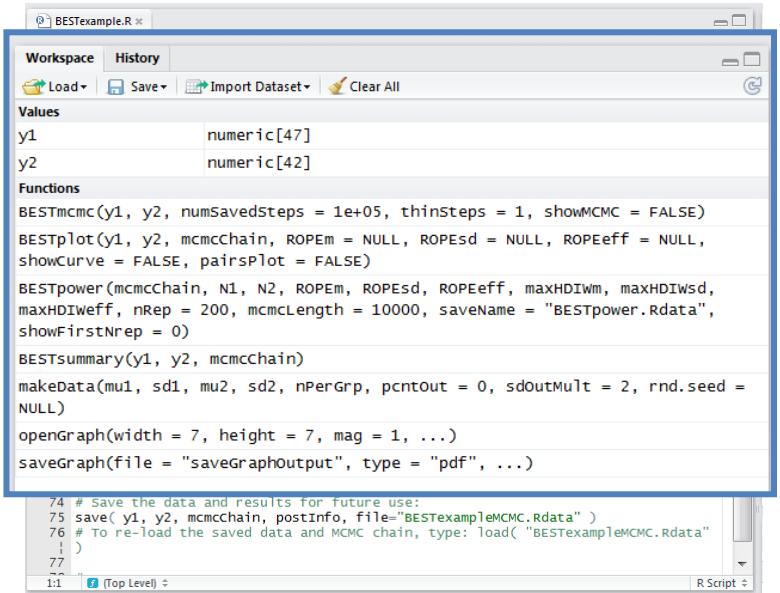
```

© John K. Kruschke, 2013

45

Example of running BEST

When executed, notice that variables y1 and y2 appear in R's workspace.



```

@ BESTexample.R *
Workspace | History
Load | Save | Import Dataset | Clear All
Values
y1      numeric[47]
y2      numeric[42]
Functions
BESTmcmc(y1, y2, numSavedSteps = 1e+05, thinSteps = 1, showMCMC = FALSE)
BESTplot(y1, y2, mcmcChain, ROPEm = NULL, ROPEsd = NULL, ROPEeff = NULL,
showCurve = FALSE, pairsPlot = FALSE)
BESTpower(mcmcChain, N1, N2, ROPEm, ROPEsd, ROPEeff, maxHDIwM, maxHDIwSd,
maxHDIwEff, nRep = 200, mcmcLength = 10000, saveName = "BESTpower.Rdata",
showFirstNrep = 0)
BESTsummary(y1, y2, mcmcChain)
makeData(mu1, sd1, mu2, sd2, nPerGrp, pcntout = 0, sdoutMult = 2, rnd.seed =
NULL)
openGraph(width = 7, height = 7, mag = 1, ...)
saveGraph(file = "saveGraphoutput", type = "pdf", ...)

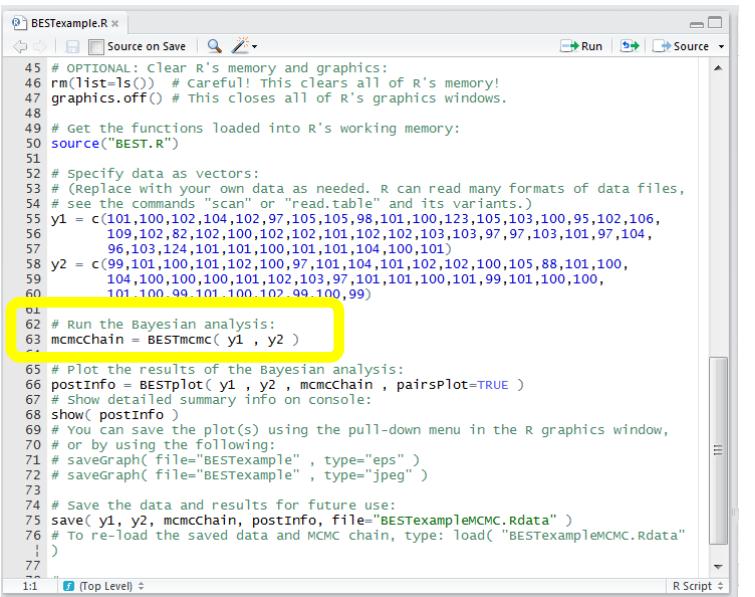
74 # Save the data and results for future use:
75 save( y1 , y2 , mcmcChain , postInfo , file="BESTexampleMCMC.Rdata" )
76 # To re-load the saved data and MCMC chain, type: load( "BESTexampleMCMC.Rdata"
77 )
78

```

© John K. Kruschke, 2013

46

Example of running BEST



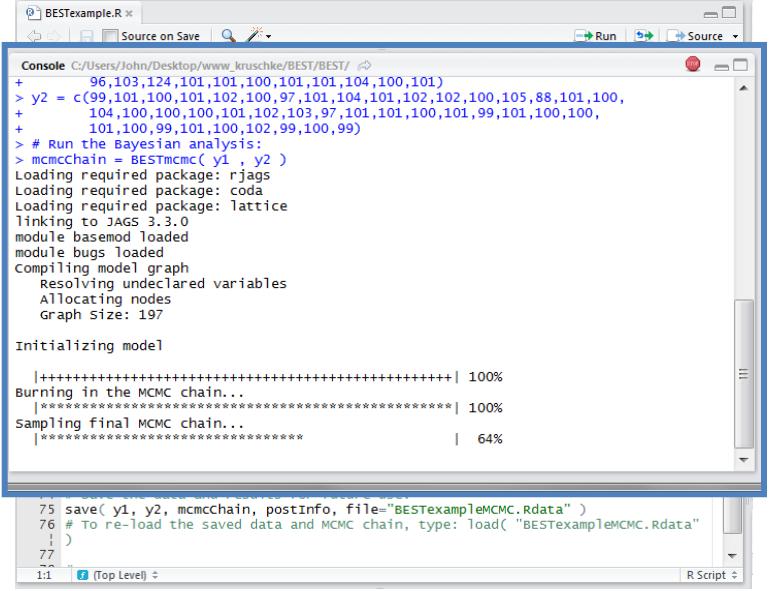
```

@ BESTexample.R x
Source on Save | Run | Source
45 # OPTIONAL: Clear R's memory and graphics:
46 rm(list=ls()) # Careful! This clears all of R's memory!
47 graphics.off() # This closes all of R's graphics windows.
48
49 # Get the functions loaded into R's working memory:
50 source("BEST.R")
51
52 # Specify data as vectors:
53 # (Replace with your own data as needed. R can read many formats of data files,
54 # see the commands "scan" or "read.table" and its variants.)
55 y1 = c(101,100,102,104,102,97,105,105,98,101,100,123,105,103,100,95,102,106,
56 109,102,82,102,100,102,101,102,102,103,103,97,97,103,101,97,104,
57 96,103,124,101,101,100,101,101,104,100,101)
58 y2 = c(99,101,100,101,102,100,97,101,104,101,102,102,100,105,88,101,100,
59 104,100,100,100,101,102,103,97,101,101,100,101,99,101,100,100,
60 101,100,99,101,100,102,99,100,99)
61
62 # Run the Bayesian analysis:
63 mcmcChain = BESTmcmc( y1 , y2 )
64
65 # Plot the results of the Bayesian analysis:
66 postInfo = BESTplot( y1 , y2 , mcmcChain , pairsPlot=TRUE )
67 # Show detailed summary info on console:
68 show( postInfo )
69 # You can save the plot(s) using the pull-down menu in the R graphics window,
70 # or by using the following:
71 # saveGraph( file="BESTexample" , type="eps" )
72 # saveGraph( file="BESTexample" , type="jpeg" )
73
74 # Save the data and results for future use:
75 save( y1 , y2 , mcmcChain , postInfo , file="BESTexampleMCMC.Rdata" )
76 # To re-load the saved data and MCMC chain, type: load("BESTexampleMCMC.Rdata"
77 )
78

```

© John K. Kruschke, 2013 47

Example of running BEST



When executed, the console shows dynamically the progress of the MCMC sampling.

Click the stop sign at the top of the console to interrupt the processing (but unfinished information is *not* saved)

```

Console C:/Users/John/Desktop/www.kruschke/BEST/BEST/
+ 96,103,124,101,101,100,101,101,104,100,101)
+ > y2 = c(99,101,100,101,102,100,97,101,104,101,102,102,100,105,88,101,100,
+ 104,100,100,100,101,102,103,97,101,101,100,101,99,101,100,100,
+ 101,100,99,101,100,102,99,100,99)
+ > # Run the Bayesian analysis:
+ > mcmcChain = BESTmcmc( y1 , y2 )
Loading required package: rjags
Loading required package: coda
Loading required package: lattice
Linking to JAGS 3.3.0
module basemod loaded
module bugs loaded
Compiling model graph
Resolving undeclared variables
Allocating nodes
Graph Size: 197

Initializing model
|+++++| 100%
Burning in the MCMC chain...
|*****| 100%
Sampling final MCMC chain...
|*****| 64%

```

```

75 save( y1 , y2 , mcmcChain , postInfo , file="BESTexampleMCMC.Rdata" )
76 # To re-load the saved data and MCMC chain, type: load("BESTexampleMCMC.Rdata"
77 )
78

```

© John K. Kruschke, 2013 48

Example of running BEST

When executed, two graphics windows appear, perhaps **behind** the RStudio window.

```

@ BESTexample.R x
Source on Save | Run | Source

45 # OPTIONAL: Clear R's memory and graphics:
46 rm(list=ls()) # Careful! This clears all of R's memory!
47 graphics.off() # This closes all of R's graphics windows.
48
49 # Get the functions loaded into R's working memory:
50 source("BEST.R")
51
52 # Specify data as vectors:
53 # (Replace with your own data as needed. R can read many formats of data files,
54 # see the commands "scan" or "read.table" and its variants.)
55 y1 = c(101,100,102,104,102,97,105,105,98,101,100,123,105,103,100,95,102,106,
56 109,102,82,102,100,102,102,101,102,102,103,103,97,97,103,101,97,104,
57 96,103,124,101,101,100,101,101,104,100,101)
58 y2 = c(99,101,100,101,102,100,97,101,104,101,102,102,100,105,88,101,100,
59 104,100,100,100,101,102,103,97,101,101,100,101,99,101,100,100,
60 101,100,99,101,100,102,99,100,99)
61
62 # Run the Bayesian analysis:
63 mcmcChain = BESTmcmc( y1 , y2 )

65 # Plot the results of the Bayesian analysis:
66 postInfo = BESTplot( y1 , y2 , mcmcChain , pairsPlot=TRUE )

68 show( postInfo )
69 # You can save the plot(s) using the pull-down menu in the R graphics window,
70 # or by using the following:
71 # saveGraph( file="BESTexample" , type="eps" )
72 # saveGraph( file="BESTexample" , type="jpeg" )
73
74 # Save the data and results for future use:
75 save( y1 , y2 , mcmcChain , postInfo , file="BESTexampleMCMC.Rdata" )
76 # To re-load the saved data and MCMC chain, type: load( "BESTexampleMCMC.Rdata"
| )
77

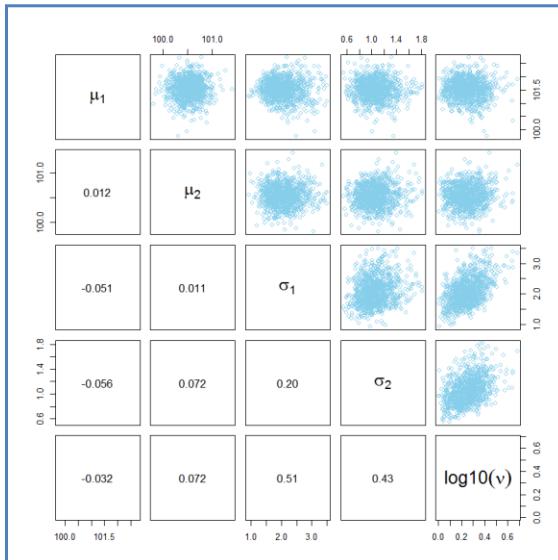
```

© John K. Kruschke, 2013

49

Example of running BEST

Here is a subset of points from the 50,000 step MCMC sample...



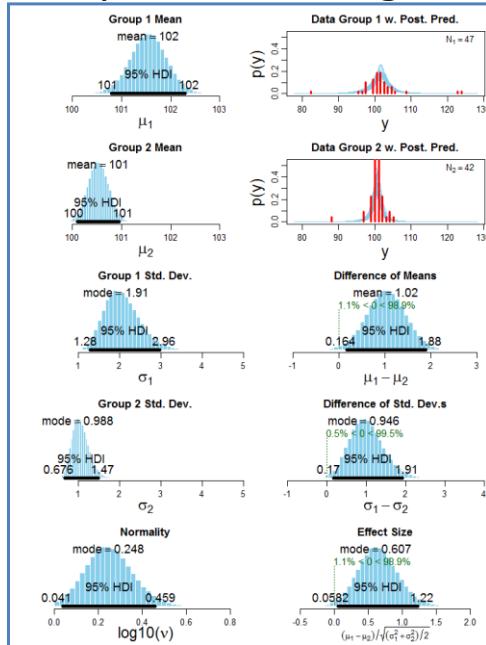
© John K. Kruschke, 2013

50

Example of running BEST

Here is a summary depiction of the posterior distribution, including difference of means, difference of standard deviations, effect size, and posterior predictive check.

© John K. Kruschke, 2013



51

Example of running BEST

```

BESTexample.R * | 
Source | Run | Source |
45 # OPTIONAL: Clear R's memory and graphics:
46 rm(list=ls()) # Careful! This clears all of R's memory!
47 graphics.off() # This closes all of R's graphics windows.
48
49 # Get the functions loaded into R's working memory:
50 source("BEST.R")
51
52 # Specify data as vectors:
53 # (Replace with your own data as needed. R can read many formats of data files,
54 # see the commands "scan" or "read.table" and its variants.)
55 y1 = c(101,100,102,104,102,97,105,105,98,101,100,123,105,103,100,95,102,106,
56 109,102,82,102,100,102,101,102,102,103,103,97,97,103,101,97,104,
57 96,103,124,101,101,100,101,101,104,100,101)
58 y2 = c(99,101,100,101,102,100,97,101,104,101,102,102,100,105,88,101,100,
59 104,100,100,100,101,102,103,97,101,101,100,101,99,101,100,100,
60 101,100,99,101,100,102,99,100,99)
61
62 # Run the Bayesian analysis:
63 mcmcChain = BESTmcmc( y1 , y2 )
64
65 # Plot the results of the Bayesian analysis:
66
67 # Show detailed summary info on console:
68 show( postInfo )

```

The code block shows an R script named 'BESTexample.R' with several lines highlighted by a yellow rectangle. Line 67 is specifically highlighted.

R Script

© John K. Kruschke, 2013

52

When executed, numerical details of the posterior appear in the console:

Example of running BEST

```
Console C:/Users/John/Desktop/www_kruschke/BEST/BEST/ ▷
Graph Size: 197

Initializing model
|+++++| 100%
Burning in the MCMC chain..
|*****| 100%
Sampling final MCMC chain..
|*****| 100%
> # Plot the results of the Bayesian analysis:
> postInfo = BESTplot( y1 , y2 , mcmcChain , pairsPlot=TRUE )
> # Show detailed summary info on console:
> show( postInfo )
      SUMMARY.INFO
PARAMETER   mean    median     mode   HDIlow   HDIhigh pcgtZero
mu1        101.5469112 101.5471794 101.5339003 100.79518100 102.2956758 NA
mu2        100.5253590 100.5230538 100.5123852 100.10553033 100.9423119 NA
mudiff     1.0215521  1.0231408  1.0339815  0.16445951  1.8831049  98.93302
sigma1     2.0642192  2.0242978  1.9119128  1.27706308  2.9570681 NA
sigma2     1.0521815  1.0324068  0.9882783  0.67612304  1.4730769 NA
sigmadiff   1.0120378  0.9820825  0.9461807  0.16984554  1.9114531  99.48701
nu         1.8448090  1.7676663  1.6252294  1.04256844  2.7786018 NA
nuLog10    0.2522517  0.2474003  0.2481552  0.04103825  0.4587576 NA
effsz      0.6424580  0.6320242  0.6068438  0.05819721  1.2216764  98.93302
>
67 # Show detailed summary info on console:
68 show( postInfo )
```

70 # or by using the following:
71 # saveGraph(file="BESTexample" , type="eps")
72 # saveGraph(file="BESTexample" , type="jpeg")
73
74 # Save the data and results for future use:
75 save(y1 , y2 , mcmcChain , postInfo , file="BESTexampleMCMC.Rdata")
76 # To re-load the saved data and MCMC chain, type: load("BESTexampleMCMC.Rdata")
77

1:1 [Top Level] ▷ R Script

© John K. Kruschke, 2013

53

Example of running BEST

Save graphs and MCMC results.

```
BESTexample.R * | Source on Save | Run | Source ▷
45 # OPTIONAL: Clear R's memory and graphics:
46 rm(list=ls()) # Careful! This clears all of R's memory!
47 graphics.off() # This closes all of R's graphics windows.
48
49 # Get the functions loaded into R's working memory:
50 source("BEST.R")
51
52 # Specify data as vectors:
53 # (Replace with your own data as needed. R can read many formats of data files,
54 # see the commands "scan" or "read.table" and its variants.)
55 y1 = c(101,100,102,104,102,97,105,98,101,100,123,105,103,100,95,102,106,
56       109,102,82,102,100,102,101,102,102,103,103,97,97,103,101,97,104,
57       96,103,124,101,101,100,101,101,104,100,101)
58 y2 = c(99,101,100,101,102,100,97,101,104,101,102,102,100,105,88,101,100,
59       104,100,100,101,102,103,97,101,101,100,101,99,101,100,100,100,
60       101,100,99,101,100,102,99,100,99)
61
62 # Run the Bayesian analysis:
63 mcmcChain = BESTmcmc( y1 , y2 )
64
65 # Plot the results of the Bayesian analysis:
66 postInfo = BESTplot( y1 , y2 , mcmcChain , pairsPlot=TRUE )
67 # Show detailed summary info on console:
```

69 # You can save the plot(s) using the pull-down menu in the R graphics window,
70 # or by using the following:
71 # saveGraph(file="BESTexample" , type="eps")
72 # saveGraph(file="BESTexample" , type="jpeg")
73
74 # Save the data and results for future use:
75 save(y1 , y2 , mcmcChain , postInfo , file="BESTexampleMCMC.Rdata")
76 # To re-load the saved data and MCMC chain, type: load("BESTexampleMCMC.Rdata")
77

1:1 [Top Level] ▷ R Script

© John K. Kruschke, 2013

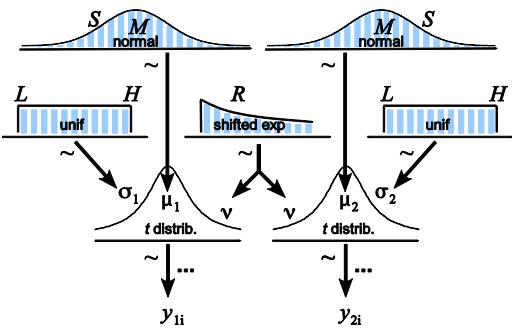
54

Robust Bayesian estimation for comparing two groups

Now for a look
under the hood



<http://www.autonationconnect.com/2010/07/backseat-mechanic-under-the-hood/>



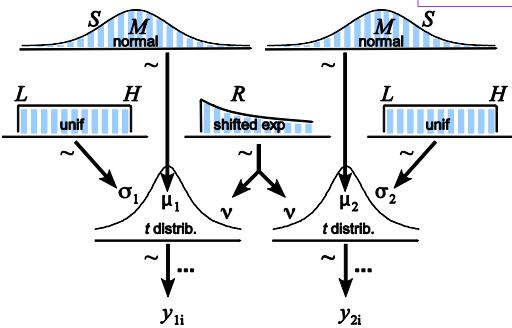
© John K. Kruschke, 2013

55

Robust Bayesian estimation for comparing two groups

Program BEST.R: JAGS model specification.

```
model {
  for ( i in 1:Ntotal ) {
    y[i] ~ dt( mu[x[i]] , tau[x[i]] , nu )
  }
  for ( j in 1:2 ) {
    mu[j] ~ dnorm( muM , muP )
    tau[j] <- 1/pow( sigma[j] , 2 )
    sigma[j] ~ dunif( sigmaLow , sigmaHigh )
  }
  nu <- nuMinusOne+1
  nuMinusOne ~ dexp(1/29)
}
```



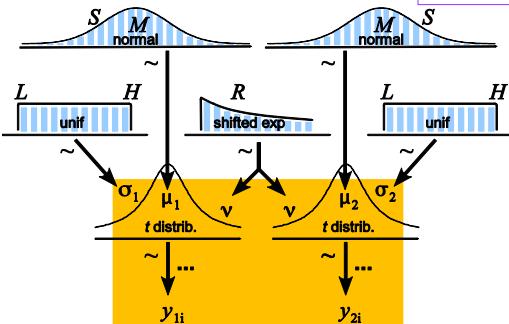
© John K. Kruschke, 2013

56

Robust Bayesian estimation for comparing two groups

Program BEST.R: JAGS model specification.

```
model {
  for ( i in 1:Ntotal ) {
    y[i] ~ dt( mu[x[i]] , tau[x[i]] , nu )
  }
  for ( j in 1:2 ) {
    mu[j] ~ dnorm( muM , muP )
    tau[j] <- 1/pow( sigma[j] , 2 )
    sigma[j] ~ dunif( sigmaLow , sigmaHigh )
  }
  nu <- nuMinusOne+1
  nuMinusOne ~ dexp(1/29)
}
```



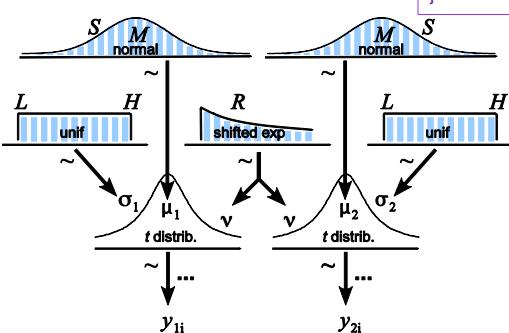
© John K. Kruschke, 2013

57

Robust Bayesian estimation for comparing two groups

Program BEST.R: JAGS model specification.

```
model {
  for ( i in 1:Ntotal ) {
    y[i] ~ dt( mu[x[i]] , tau[x[i]] , nu )
  }
  for ( j in 1:2 ) {
    mu[j] ~ dnorm( muM , muP )
    tau[j] <- 1/pow( sigma[j] , 2 )
    sigma[j] ~ dunif( sigmaLow , sigmaHigh )
  }
  nu <- nuMinusOne+1
  nuMinusOne ~ dexp(1/29)
}
```



Nested indexing:
 $x[i]$ is the group (1 or 2)
of the i^{th} score.

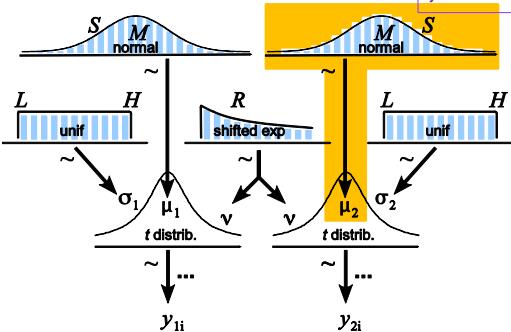
© John K. Kruschke, 2013

58

Robust Bayesian estimation for comparing two groups

Program BEST.R: JAGS model specification.

```
model {
  for ( i in 1:Ntotal ) {
    y[i] ~ dt( mu[x[i]] , tau[x[i]] , nu )
  }
  for ( j in 1:2 ) {
    mu[j] ~ dnorm( muM , muP )
    tau[j] <- 1/pow( sigma[j] , 2 )
    sigma[j] ~ dunif( sigmaLow , sigmaHigh )
  }
  nu <- nuMinusOne+1
  nuMinusOne ~ dexp(1/29)
}
```



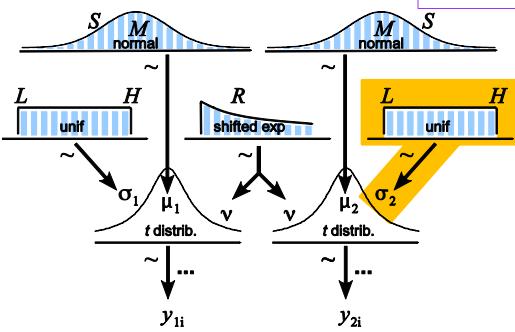
© John K. Kruschke, 2013

59

Robust Bayesian estimation for comparing two groups

Program BEST.R: JAGS model specification.

```
model {
  for ( i in 1:Ntotal ) {
    y[i] ~ dt( mu[x[i]] , tau[x[i]] , nu )
  }
  for ( j in 1:2 ) {
    mu[j] ~ dnorm( muM , muP )
    tau[j] <- 1/pow( sigma[j] , 2 )
    sigma[j] ~ dunif( sigmaLow , sigmaHigh )
  }
  nu <- nuMinusOne+1
  nuMinusOne ~ dexp(1/29)
}
```



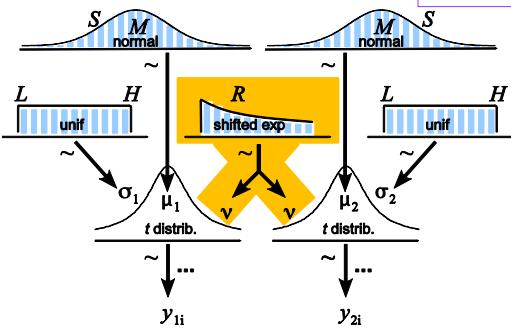
© John K. Kruschke, 2013

60

Robust Bayesian estimation for comparing two groups

Program BEST.R: JAGS model specification.

```
model {
  for ( i in 1:Ntotal ) {
    y[i] ~ dt( mu[x[i]] , tau[x[i]] , nu )
  }
  for ( j in 1:2 ) {
    mu[j] ~ dnorm( muM , muP )
    tau[j] <- 1/pow( sigma[j] , 2 )
    sigma[j] ~ dunif( sigmaLow , sigmaHigh )
  }
  nu <- nuMinusOne+1
  nuMinusOne ~ dexp(1/29)
}
```



© John K. Kruschke, 2013

61

Approximating the posterior by a random sample

Problem:

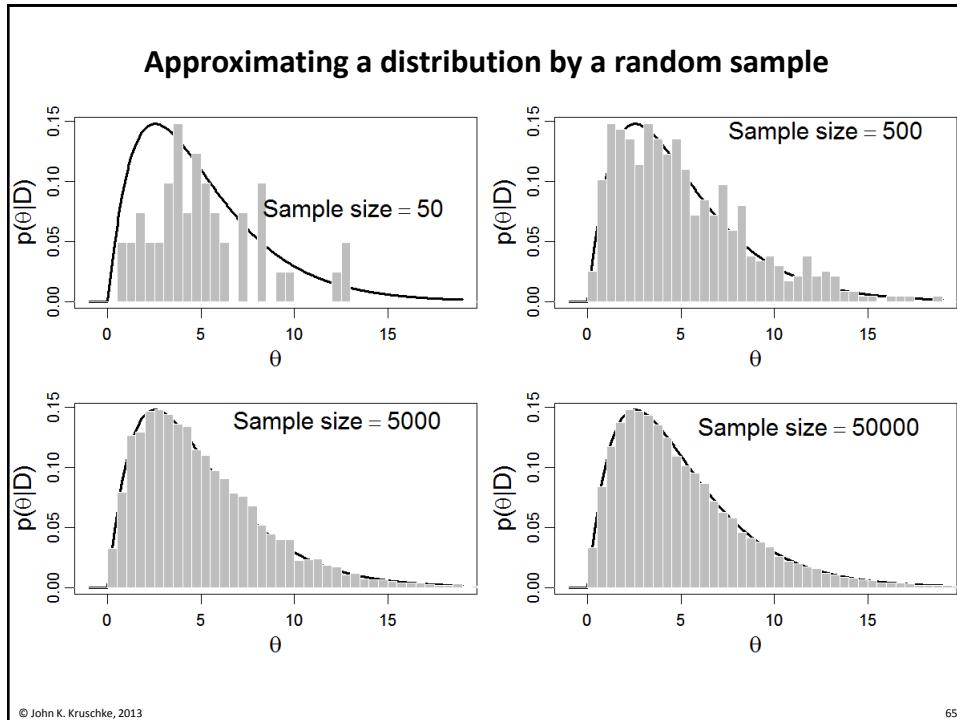
For realistically complex models, analytical mathematical derivation of the posterior distribution is unknown.

Solution:

- Approximate the posterior by a large random sample of representative values of the parameter.
- This approach works because we can generate random values from the posterior without solving the math.

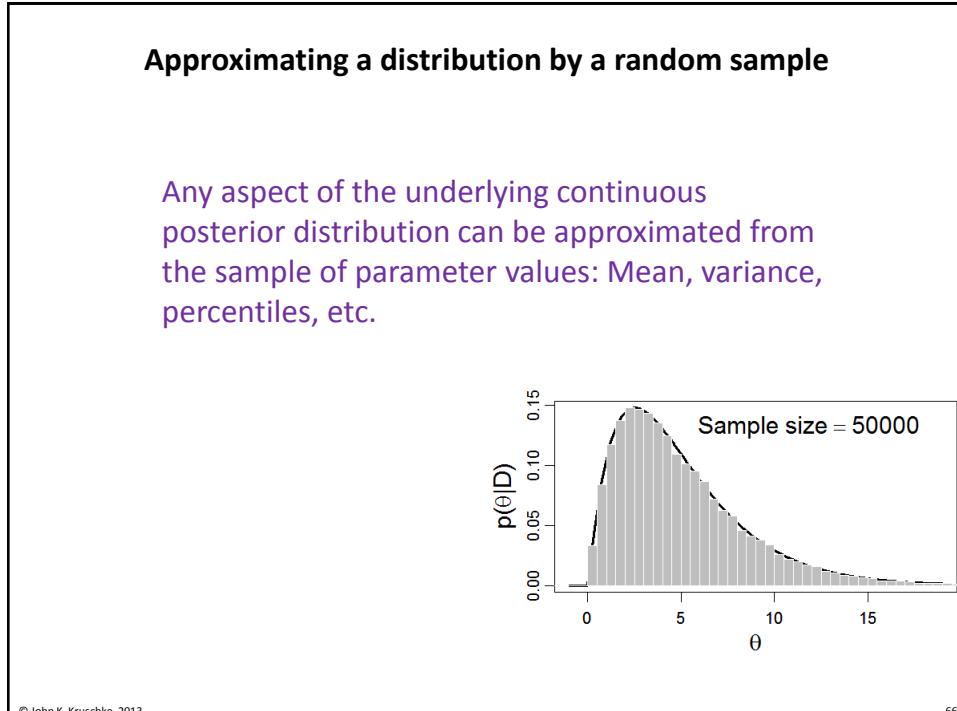
© John K. Kruschke, 2013

64



© John K. Kruschke, 2013

65



© John K. Kruschke, 2013

66

How to generate a random sample

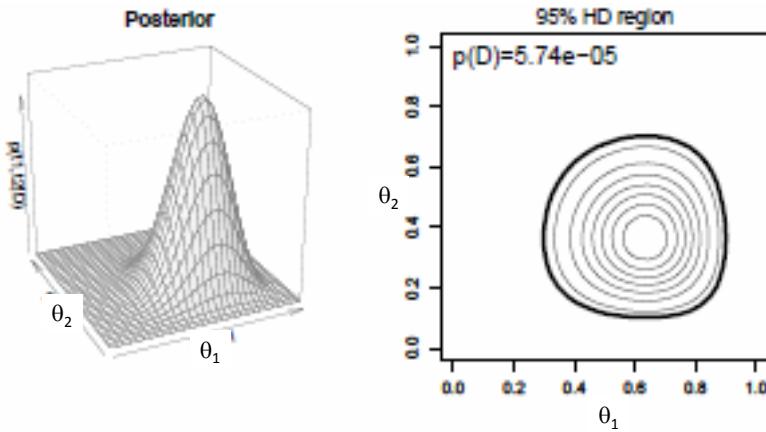
Some distributions have known, efficient random-sample generators:
 Uniform, normal, exponential, beta, gamma, etc.

Use the known random sample generators to indirectly sample from a complicated posterior distribution.
 There are several ways, of which we'll consider two:

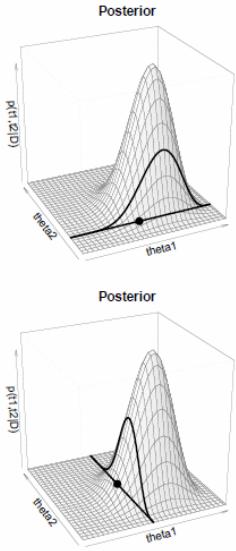
- *Gibbs Sampling*
- *Metropolis Algorithm*

Example

Consider this posterior distribution on two parameters, θ_1 and θ_2 . Goal is to generate a random sample without computing the integral $p(D)$.

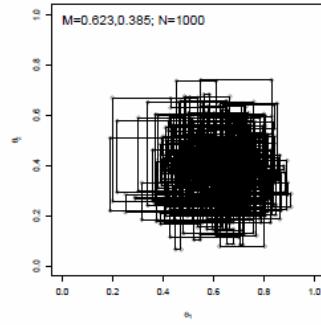


Gibbs Sampling



1. Derive conditional distribution for each parameter: $p(\theta_i | D, \{\theta_{j \neq i}\})$
2. Sample directly from each conditional distribution.

Figure 8.4: Two steps in a Gibbs sampling. The top panel shows a random generation of a value for θ_1 , conditional on a value for θ_2 . The heavy lines show a slice through the posterior at the conditional value of θ_2 , and the large dot shows a random value of θ_1 sampled from the conditional density. The bottom panel shows a random generation of a value for θ_2 , conditional on the value for θ_1 determined by the previous step. The heavy lines show a slice through the posterior at the conditional value of θ_1 , and the large dot shows a random value of θ_2 sampled from the conditional density.

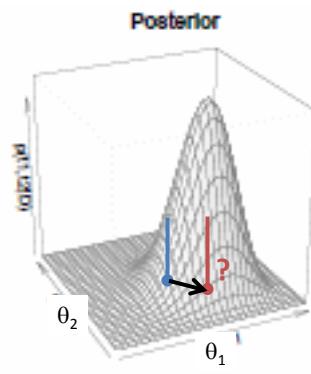
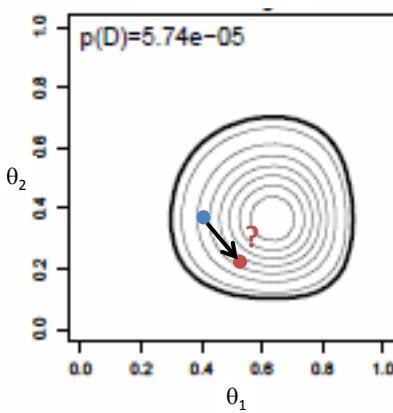


© John K. Kruschke, 2013

69

Metropolis Algorithm

1. Propose a new parameter position, θ_{pro} , directly sampled from a known distribution such as a normal.
2. Decide whether to accept the proposed position. Decision rule only needs to evaluate $p(D|\theta_{\text{pro}})$ and $p(\theta_{\text{pro}})$, not $p(D)$.



© John K. Kruschke, 2013

71

Metropolis Algorithm

1. Propose a new parameter position, θ_{pro} , directly sampled from a known distribution such as a normal.
2. Decide whether to accept the proposed position. Decision rule only needs to evaluate $p(D|\theta_{\text{pro}})$ and $p(\theta_{\text{pro}})$, not $p(D)$.

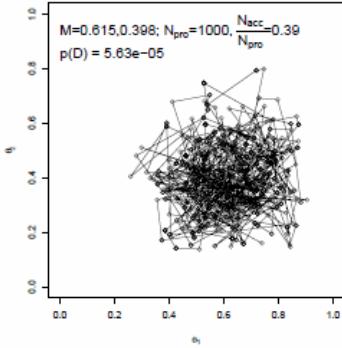


Figure 8.3: Metropolis algorithm applied to the prior and likelihood shown in Figure 8.1, p. 131. Compare this scatter plot with the lower right contour plot of Figure 8.1. N_{pro} is the number of proposed jumps (after burn in), and N_{acc} is the number of accepted proposals. R code is in Section 8.8.2 (BernTwoMetropolis.R).

Gibbs and Metropolis: Both “MCMC” Sampling

Both are “Markov chain Monte Carlo” (MCMC) methods:
 From current point (\leftarrow Markov chain),
 randomly generate a new point (\leftarrow Monte Carlo).

All MCMC methods generate a high-resolution picture of the single true posterior, but with different efficiencies.

JAGS makes it easy! We merely specify the model, and JAGS generates an MCMC representation of the posterior distribution.

Outline of Talk:

- Bayesian reasoning generally.
- Bayesian estimation applied to two groups. Rich information.
 - Running the programs, and the flexibility of Bayesian software.
- The NHST t test: perfidious p values and the con game of confidence intervals.
- Conclusion: Bayesian estimation supersedes NHST.

© John K. Kruschke, 2013

75

An example of a t test:

Data:

Group 1: 5.70 5.40 5.75 5.25 4.25 4.74; $M1 = 5.18$

Group 2: 4.55 4.98 4.70 4.78 3.26 3.67; $M2 = 4.32$

$t = 2.33$

Show of hands please:

Who bets that $p < .05$? Who bets that $p > .05$?

© John K. Kruschke, 2013

76

An example of a *t* test:

Data:

Group 1: 5.70 5.40 5.75 5.25 4.25 4.74; $M_1 = 5.18$

Group 2: 4.55 4.98 4.70 4.78 3.26 3.67; $M_2 = 4.32$

$t = 2.33$

Show of hands please:

Who bets that $p < .05$? Who bets that $p > .05$?

You're right!

You're right!

Null Hypothesis Significance Testing (NHST)

Consider how we draw conclusions from data:

- Collect data, *carefully insulated from our intentions*.
 - Double blind clinical designs.
 - No datum is influenced by any other datum before or after.
- Compute a summary statistic, e.g., for a difference between groups, the *t* statistic.
- Compute *p* value of *t*. If $p < .05$, declare the result to be “significant.”

Null Hypothesis Significance Testing (NHST)

Consider how we draw conclusions from data:

- Collect data, *carefully insulated from our intentions.*
 - Double blind clinical trials
 - No datum is influenced
- Compute a summary statistic between groups, the t statistic.
- Compute p value of t . If $p < .05$, declare the result to be "significant."

Value of p depends on the intention of the experimenter!

© John K. Kruschke, 2013

70

The road to NHST is paved with good intentions.

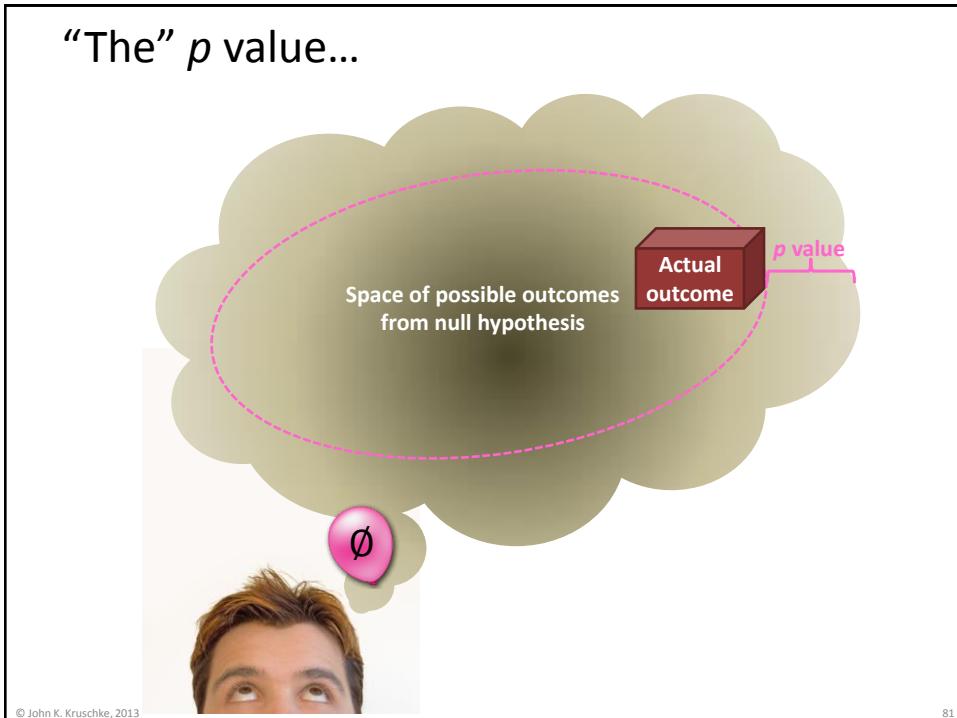
The p value is the probability that the actual sample statistic, or a result more extreme, would be obtained from the null hypothesis, *if the intended experiment were repeated ad infinitum.*

$$p \text{ value} = p(|t_{\text{null}}| > |t_{\text{actl}}|) \\ \text{for } t_{\text{null}} \text{ sampled according to} \\ \text{the intended experiment}$$

© John K. Kruschke, 2013

80

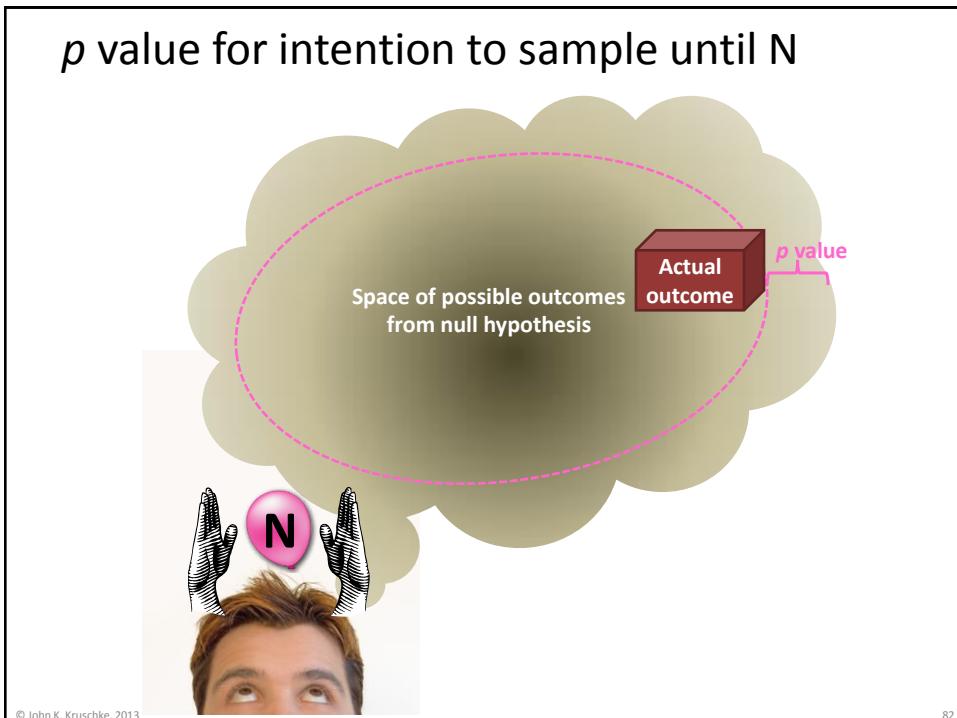
“The” p value...



© John K. Kruschke, 2013

81

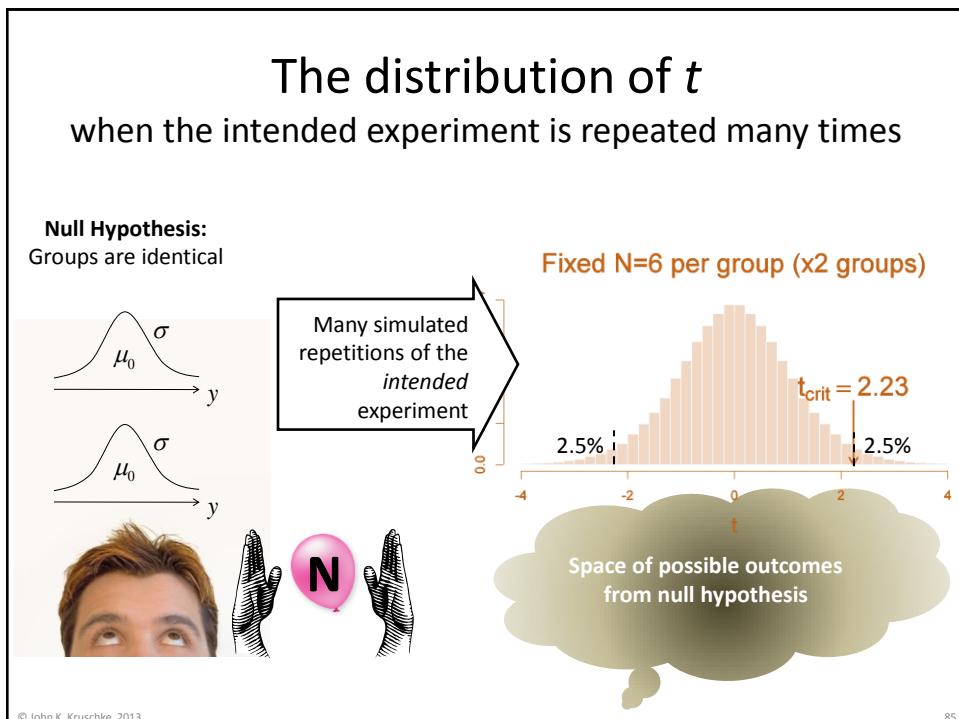
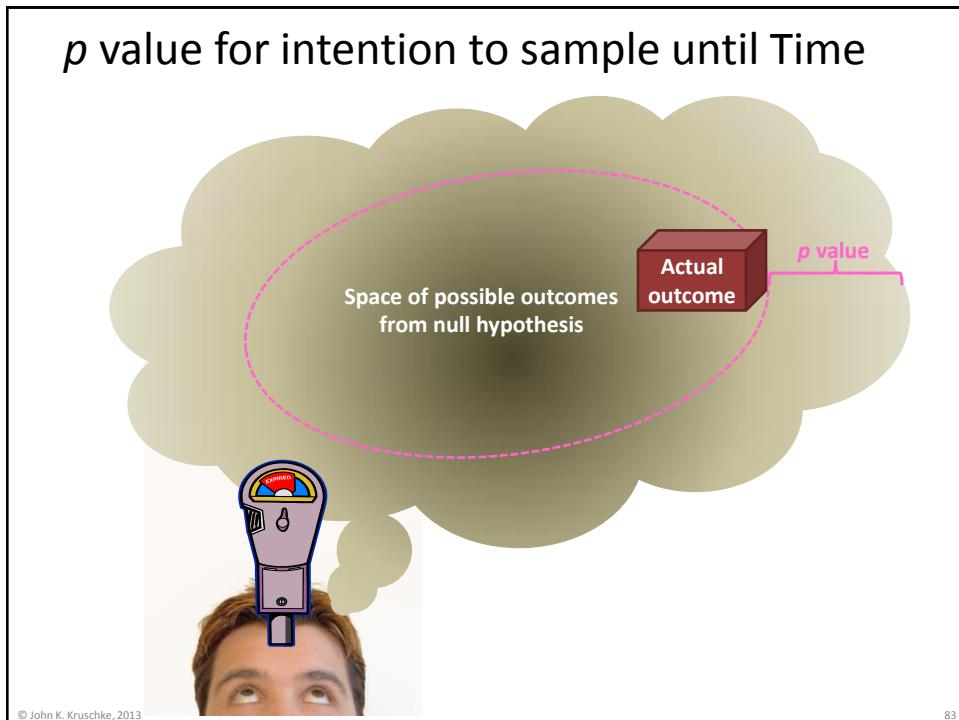
p value for intention to sample until N



© John K. Kruschke, 2013

82

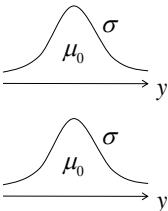
p value for intention to sample until Time



The distribution of t

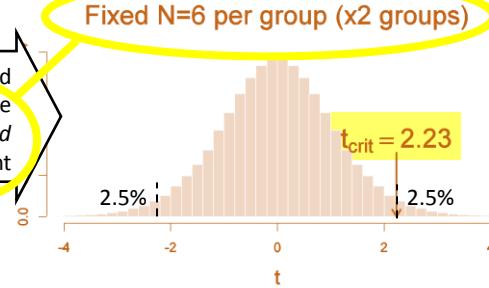
when the intended experiment is repeated many times

Null Hypothesis:
Groups are identical



Many simulated repetitions of the *intended experiment*

Fixed N=6 per group (x2 groups)



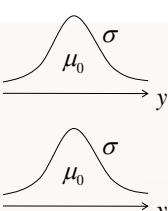
$t_{crit} = 2.23$

© John K. Kruschke, 2013

86

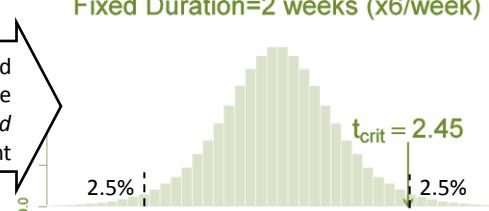
The intention to collect data until the end of the week

Null Hypothesis:
Groups are identical



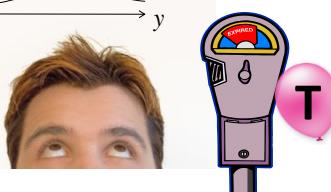
Many simulated repetitions of the *intended experiment*

Fixed Duration=2 weeks (x6/week)



$t_{crit} = 2.45$

Space of possible outcomes from null hypothesis

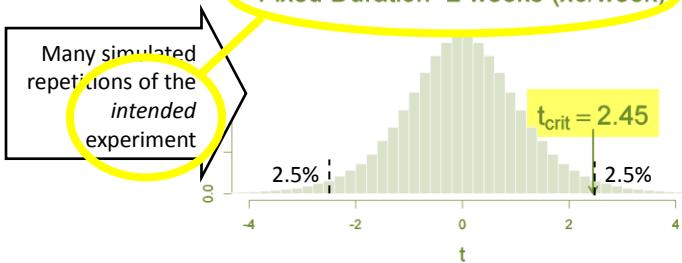
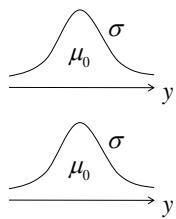


© John K. Kruschke, 2013

87

The intention to collect data until the end of the week

Null Hypothesis:
Groups are identical



© John K. Kruschke, 2013

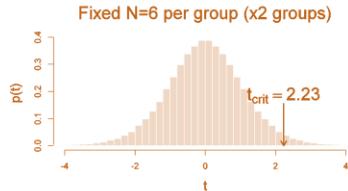
86

Two labs collect the same data:

Lab A: Collect data until N=6 per group.

Data:

Group 1: 5.70 5.40 5.75 5.25 4.25 4.74; M1 = 5.18
 Group 2: 4.55 4.98 4.70 4.78 3.26 3.67; M2 = 4.32
 $t = 2.33$



Lab A: Reject the null.

Lab B: Collect data for two weeks.

Data:

Group 1: 5.70 5.40 5.75 5.25 4.25 4.74; M1 = 5.18
 Group 2: 4.55 4.98 4.70 4.78 3.26 3.67; M2 = 4.32
 $t = 2.33$



Lab B: Do not reject the null.

© John K. Kruschke, 2013

90

The *real* use of the Neuralyzer:

You *meant* to collect data until N=12 !

Now *that's* significant!



© John K. Kruschke, 2013

91

Problem is not solved by “fixing” the intention

- All we need to do is decide in advance exactly what our intention is (or use a Neuralyzer after the fact), and have everybody chant a mantra to keep that intention fixed in their minds while the experiment is being conducted. Right?
- Wrong. The data don't know our intention, and the same data could have been collected under many other intentions.

© John K. Kruschke, 2013

92

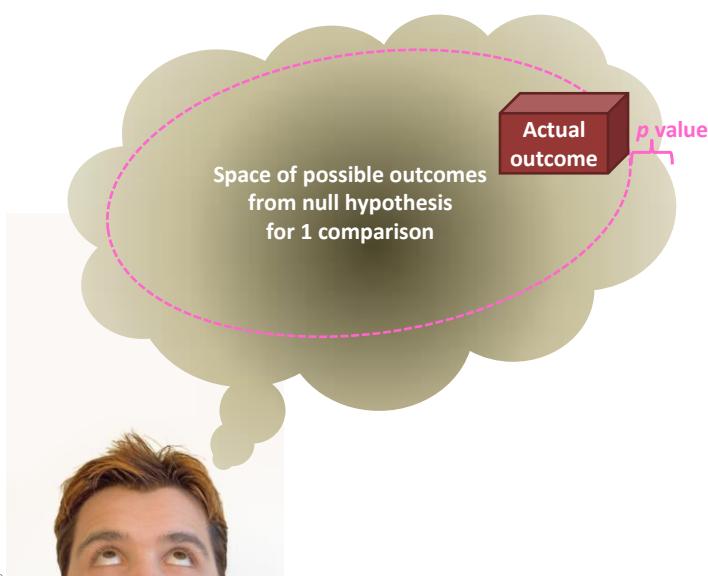
The intention to examine data thoroughly

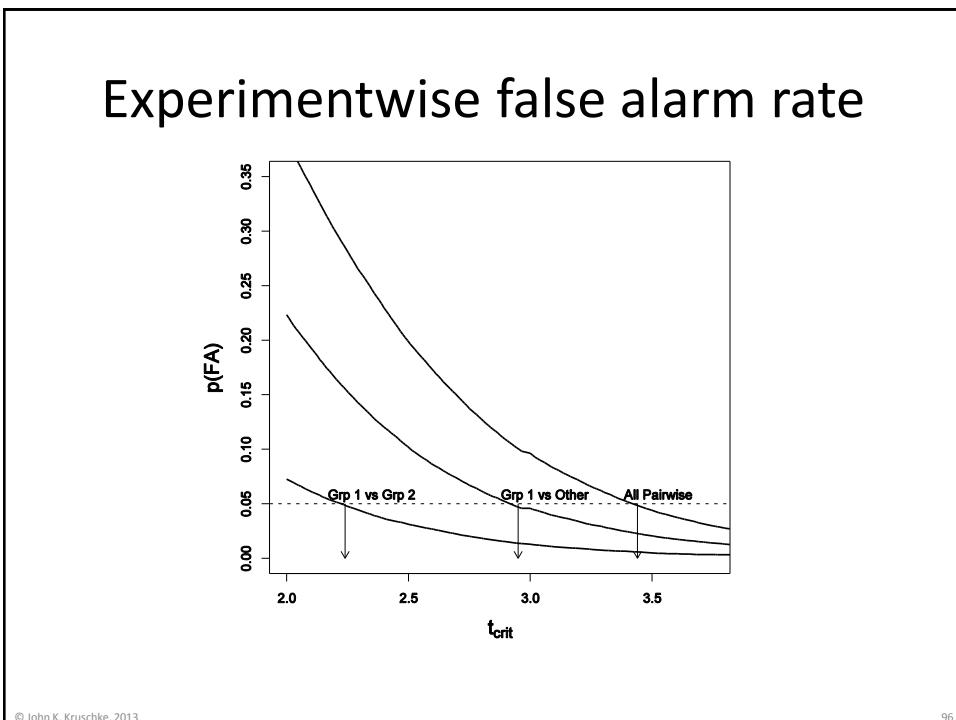
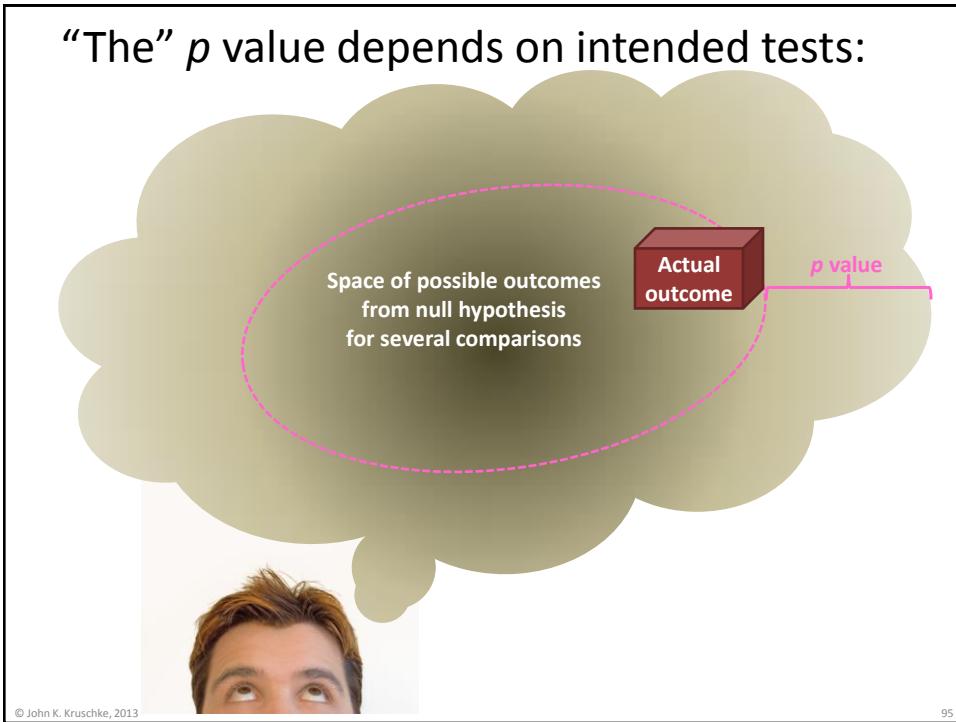
Many experiments involve multiple groups, and **multiple comparisons** of means.

Example: Consider 2 different drugs from chemical family A, 2 different drugs from chemical family B, and a placebo group. Lots of possible comparisons...

Problem: With every test, there is possibility of false alarm! False alarms are bad; therefore, keep the experimentwise false alarm rate down to 5%.

“The” p value depends on intended tests:





Multiple Corrections for Multiple Comparisons

Begin: Is goal to identify the best treatment?

Yes: Use Hsu's method.

No: Contrasts between control group and all other groups?

Yes: Use Dunnett's method.

No: Testing all pairwise and no complex comparisons (either planned or post hoc) and choosing to test only some pairwise comparisons post hoc?

Yes: Use Tukey's method.

No: Are all comparisons planned?

Yes: Use Scheffe's method.

No: Is Bonferroni critical value less than Scheffe critical value?

Yes: Use Bonferroni's method.

No: Use Scheffe's method (or, prior to collecting the data, reduce the number of contrasts to be tested).

Adapted from Maxwell & Delaney (2004). Designing experiments and analyzing data: A model comparison perspective. Erlbaum.

© John K. Kruschke, 2013

97

Multiple Corrections for Multiple Comparisons

Begin: Is goal to identify the best treatment?

Yes: Use Hsu's method.

No: Contrasts between control group and all other groups?

Yes: Use Dunnett's method.

No: Testing all pairwise and no complex comparisons (either planned or post hoc) and choosing to test only some pairwise comparisons post hoc?

Yes: Use Tukey's method.

No: Are all comparisons planned?

Yes: Use Scheffe's method.

No: Is Bonferroni critical value less than Scheffe critical value?

Yes: Use Bonferroni's method.

No: Use Scheffe's method (or, prior to collecting the data, reduce the number of contrasts to be tested).



Adapted from Maxwell & Delaney (2004). Designing experiments and analyzing data: A model comparison perspective. Erlbaum.

© John K. Kruschke, 2013

98

Good intentions make any result insignificant

- Consider an experiment with two groups.
- Collect data; compute t test on difference of means. Suppose it yields $p < .05$
- Now, think thoroughly about all the other comparison groups and other experiment groups you should and could meaningfully run.
- Earnestly intend to run them eventually, and to compare your current results with those results.
- **Poof! Your current data are no longer significantly different.**

© John K. Kruschke, 2013

99

Dang! I just wrecked the data!

Oh no! What happened?

I thought of another condition we could run!



© John K. Kruschke, 2013

100

Good intentions make many results *significant*

- Consider an experiment with two groups.
- Collect data; compute t test on difference of means, using df corresponding to actual N. Suppose $p > .05$, but not by much.
- *You had intended to collect a much larger sample size, but you were unexpectedly interrupted.*
- Use the larger intended N for df in the t test.
- ***Poof! Your current data are now significantly different!***

© John K. Kruschke, 2013

101

?

Confidence Intervals

provide no confidence

?

General definition of CI:

95% CI is the range of parameter values (e.g., $\mu_1 - \mu_2$) that would not be rejected by $p < .05$

Hence, *the 95% CI is as ill-defined as the p value.*

We see this dramatically in confidence intervals corrected for multiple comparisons.

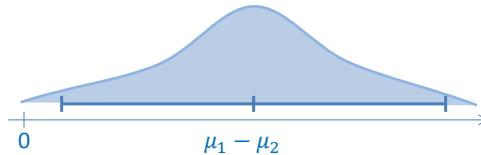
© John K. Kruschke, 2013

103

? **Confidence Intervals** **?**
provide no confidence

Confidence intervals provide no distributional information:

We have no idea whether a point at the limit of the confidence interval is any less credible than a point in the middle of the interval.



© John K. Kruschke, 2013

104

? **Confidence Intervals** **?**
provide no confidence

Confidence intervals provide no distributional information:

We have no idea whether a point at the limit of the confidence interval is any less credible than a point in the middle of the interval.

Implies

vast range for predictions of new data, and
 "virtually unknowable" power.

© John K. Kruschke, 2013

105

NHST autopsy

- p values are ill-defined: depend on sampling intentions of data collector. Any set of data has many different p values.
- Confidence intervals are as ill-defined as p values because they are defined in terms of p values.
- Confidence intervals carry no distributional information.

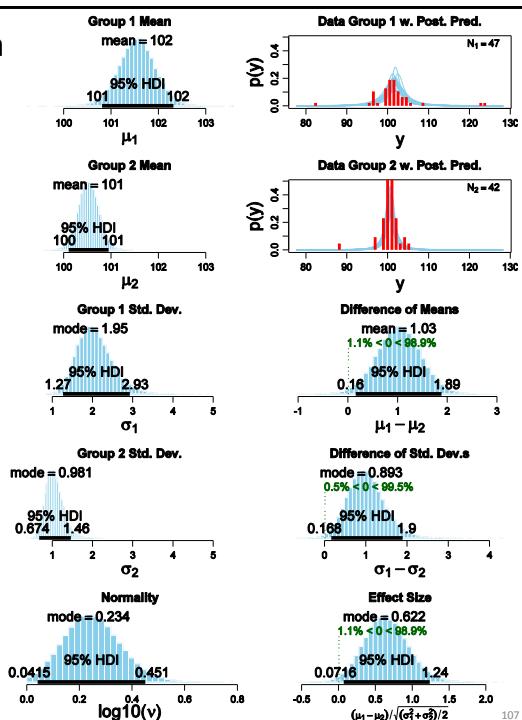
© John K. Kruschke, 2013

106

Recall Bayesian estimation for comparing two groups

Summary:

- Complete distribution of credible parameter values (not merely point estimate with ends of confidence interval).
- Decisions about multiple aspects of parameters (without reference to p values).
- Flexible descriptive model, robust to outliers (unlike NHST t test).



© John K. Kruschke, 2013

107

Recall Bayesian estimation:

What does NHST say?

t test of means:

$$t(87)=1.62, p=0.110 (>.05)$$

95%CI: -0.361 to 3.477.

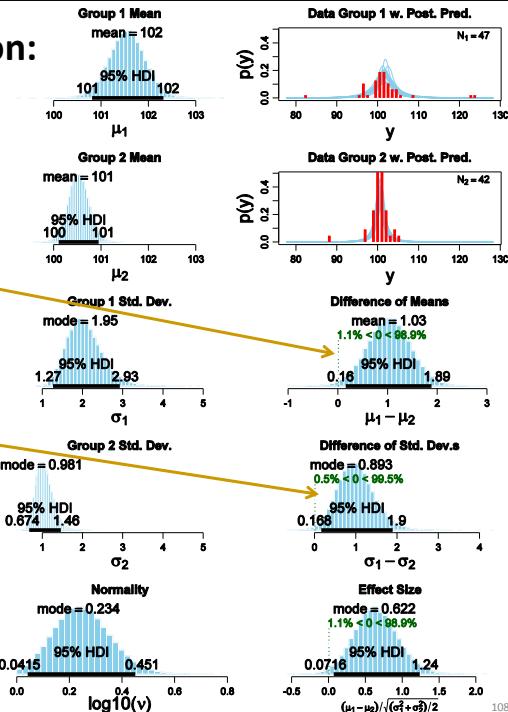
F test of variances:

$$F(46,41)=5.72, p < .001.$$

95%CI on *difference*: ?

But, must apply corrections
for multiple tests.

© John K. Kruschke, 2013



Recall Bayesian estimation:

What does NHST say?

Oops! Data are not normal,
so do resampling instead.

Resampling test of means:

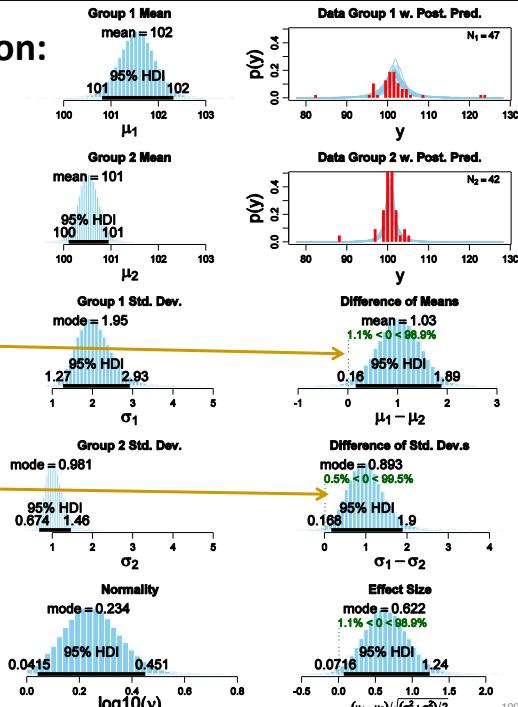
$$p=0.116 (>.05)$$

Resampling test of difference
of standard deviations:

$$p = 0.072 (>.05)$$

And, still must apply
corrections for multiple tests.
And there are no CI's.

© John K. Kruschke, 2013



Recall Bayesian estimation:

What does NHST say?

p values are fickle.

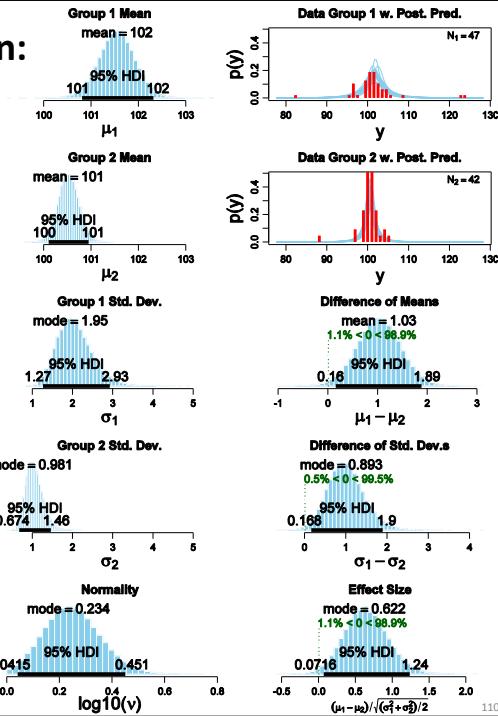
Confidence intervals change along with *p* values.

Confidence intervals carry no distributional info.

What does NHST say?

Not much!

© John K. Kruschke, 2013



Bayesian Estimation or NHST?

When Bayesian estimation and NHST *agree*, which should be used?

Bayesian estimation gives the most complete and informative answer. Answer from NHST is not informative and is fickle.

When Bayesian estimation and NHST *disagree*, which should be used?

Bayesian estimation gives the most complete and informative answer. Answer from NHST is not informative and is fickle.

© John K. Kruschke, 2013

111

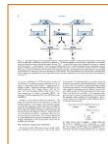
Conclusion

- p values are not well defined, nor are the limits of confidence intervals, and confidence intervals have no distributional info.
- Bayesian data analysis is the most complete and normatively correct way to estimate parameters in any model, for all your data.
- Bayesian data analysis is taking hold in 21st century science, from astronomy to zoology.

© John K. Kruschke, 2013

112

The blog: <http://doingbayesiandataanalysis.blogspot.com/>



Kruschke, J. K. (2013). **Bayesian estimation supersedes the t test.** *Journal of Experimental Psychology: General*, 142(2), 573-603.



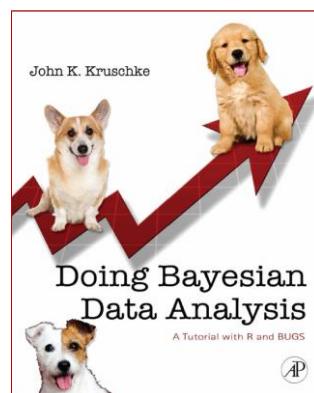
Kruschke, J. K. (2011). **Bayesian assessment of null values via parameter estimation and model comparison.** *Perspectives on Psychological Science*, 6(3), 299-312.



Kruschke, J. K. (2010). **What to believe: Bayesian methods for data analysis.** *Trends in Cognitive Sciences*, 14(7), 293-300.



Kruschke, J. K. (2010). **Bayesian data analysis.** *Wiley Interdisciplinary Reviews: Cognitive Science*, 1(5), 658-676.



Kruschke, J. K. (2011). ***Doing Bayesian Data Analysis: A Tutorial with R and BUGS.*** Academic Press / Elsevier.

© John K. Kruschke, 2013

113

Priors are not capricious

1. Priors are explicitly specified and must be acceptable to a skeptical scientific audience.
2. Typically, priors are set to be noncommittal and have very little influence on the posterior.
3. Priors can be informed by well-established data and theory, thereby giving inferential leverage to small samples.
4. When there is disagreement about the prior, then the influence of the prior on the posterior can be, and is, directly investigated. Different theoretically-informed priors can be checked.
5. Not using priors can be a serious blunder! E.g., drug/disease testing without incorporating prior knowledge of base rates.

© John K. Kruschke, 2013

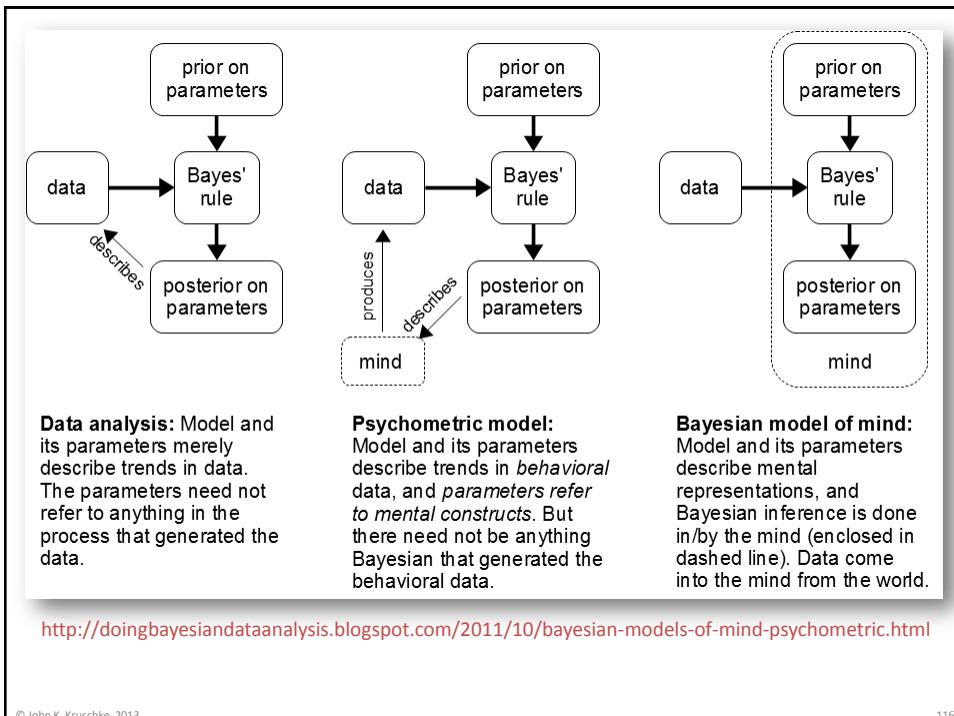
114

Prior credibility is *not* intentions

Bayesian Prior	NHST Intention (e.g., stopping rule, number of comparisons)
Explicit and supported by previous data.	Unknowable
Should influence interpretation of data.	Should <i>not</i> influence interpretation of data

© John K. Kruschke, 2013

115



© John K. Kruschke, 2013

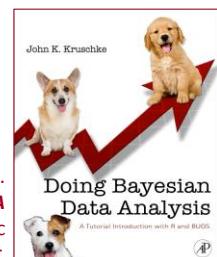
116

Bayesian estimation or Bayesian model comparison?

Bayesian estimation is also better than the “Bayesian *t* test,” which uses the “Bayes factor” from Bayesian model comparison...



Kruschke, J. K. (2011). Bayesian assessment of null values via parameter estimation and model comparison. *Perspectives on Psychological Science*, 6(3), 299-312.



Chapter 12 of Kruschke, J. K. (2011). *Doing Bayesian Data Analysis: A Tutorial with R and BUGS*. Academic Press / Elsevier.



Kruschke, J. K. (in press). Bayesian estimation supersedes the *t* test. *Journal of Experimental Psychology: General*. Appendix D.

© John K. Kruschke, 2013

117

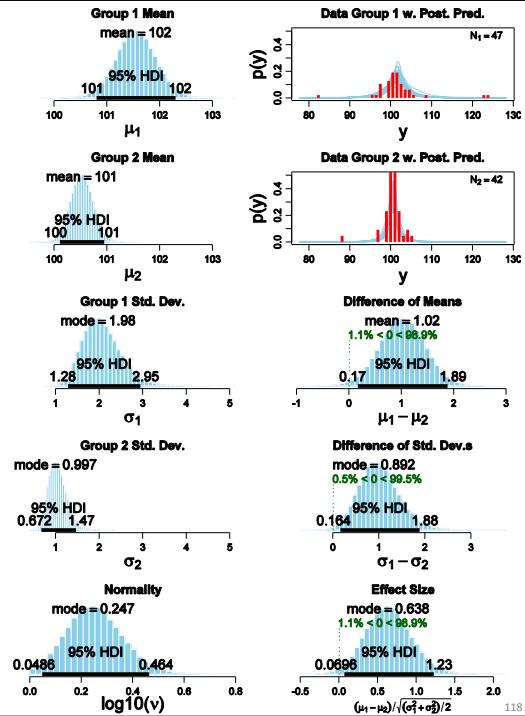
Example with outliers: BESTexample.R

Bayesian estimation:

- Credible differences between means and standard deviations.
- Complete distributional information on effect size and everything else.
- Non-normality indicated.

NHST t test:

- Outliers invalidate classic test.
- Resampling shows $p>.05$ for difference of means, $p>.05$ for difference of standard deviations.
- Need correction for multiple tests.
- No CI's. (And CI's would have no distributional info and fickle end points linked to fickle p values.)



© John K. Kruschke, 2013

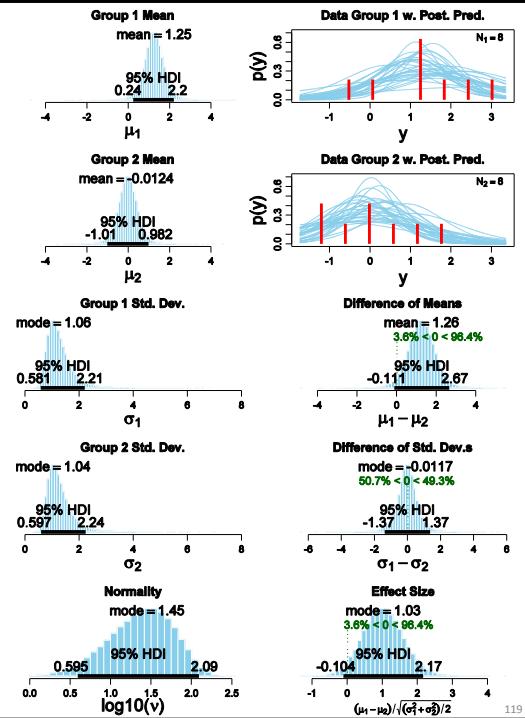
Example with small N

Bayesian estimation:

- Zero is among credible differences between means and standard deviations, and for effect size.**
- Complete distributional information on effect size and everything else.
- Normality is credible.

NHST t test:

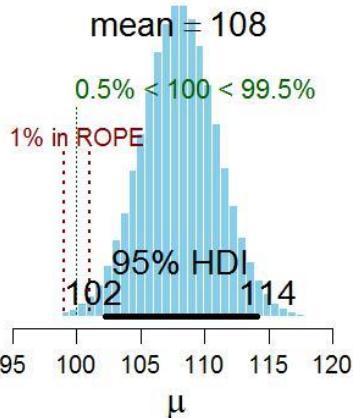
- $t(14)=2.33$, $p=0.035$, 95% CI: 0.099, 2.399. ($F(7,7)=1.00$, $p=.999$, CI on ratio: 0.20, 5.00.)
- Need correction for multiple tests, if intended.
- CI's have no distributional info and fickle end points linked to fickle p values.
- t test fails to reveal true uncertainty in parameter estimates when simultaneously estimating SD's.**



© John K. Kruschke, 2013

Region of Practical Equivalence (ROPE)

Marginal Posterior



Consider a landmark value. Values that are equivalent to that landmark for all practical purposes define the ROPE around that value.

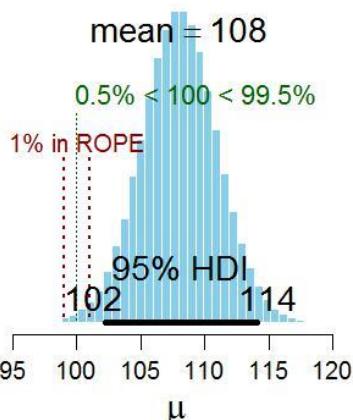
For example, the landmark value is 100, and the ROPE is 99 to 101.

© John K. Kruschke, 2013

120

Region of Practical Equivalence (ROPE)

Marginal Posterior



A parameter value is declared to be not credible, or rejected, if its entire ROPE lies outside the 95% HDI of the posterior distribution of that parameter.

A parameter value is declared to be accepted for practical purposes if that value's ROPE completely contains the 95% HDI of the posterior of that parameter.

© John K. Kruschke, 2013

121

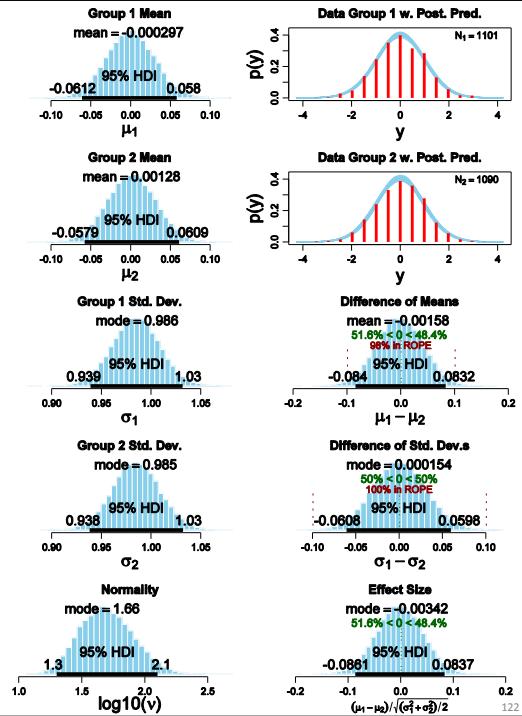
Example of accepting null value

Bayesian estimation:

- 95% HDI for difference on means falls within ROPE; same for SD's (enlarged next slide).
- Complete distributional information on effect size and everything else.
- Normality is credible.

NHST t test:

- p is large for both t and F tests, but NHST cannot accept null hypothesis.
- Need correction for multiple tests, if intended.
- CI's have no distributional info and fickle end points linked to fickle p values, and CI does not indicate probability of parameter value. Hence, **cannot use ROPE method in NHST**.



© John K. Kruschke, 2013

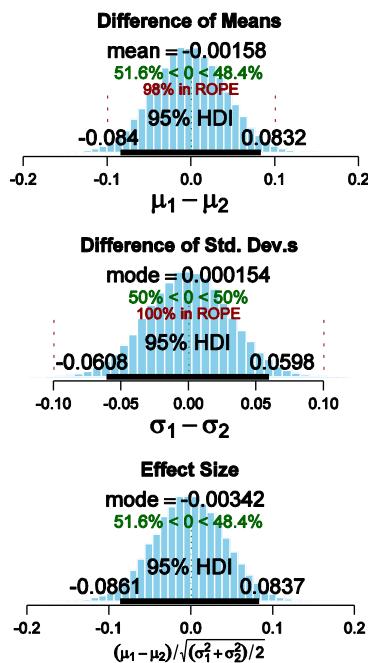
Example of accepting null value

Bayesian estimation:

- 95% HDI for difference on means falls within ROPE; same for SD's.
- Complete distributional information on effect size and everything else.
- Normality is credible.

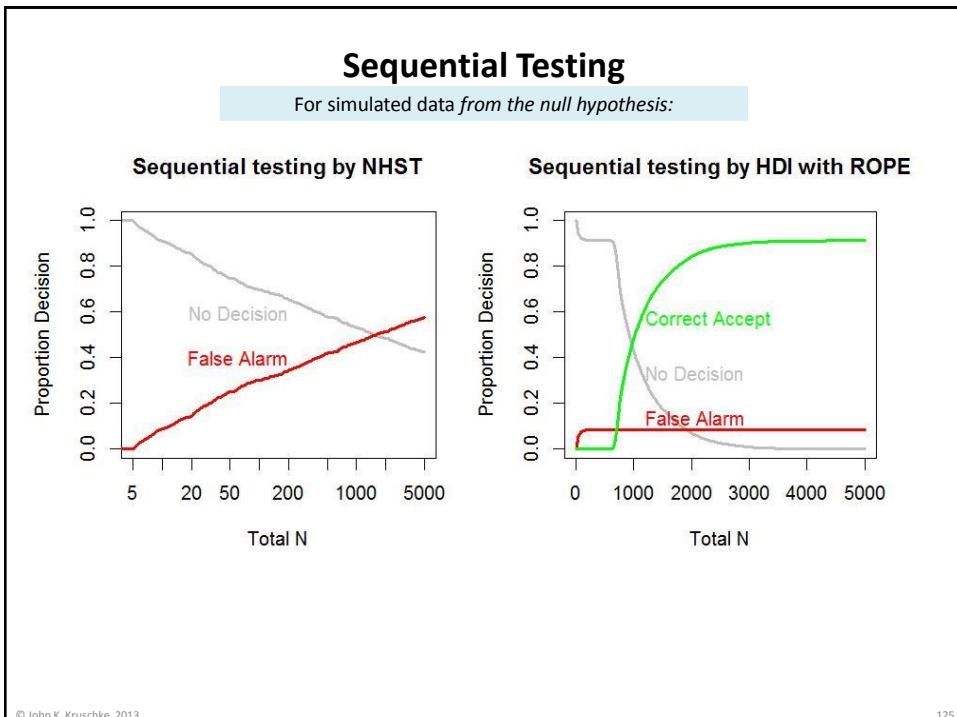
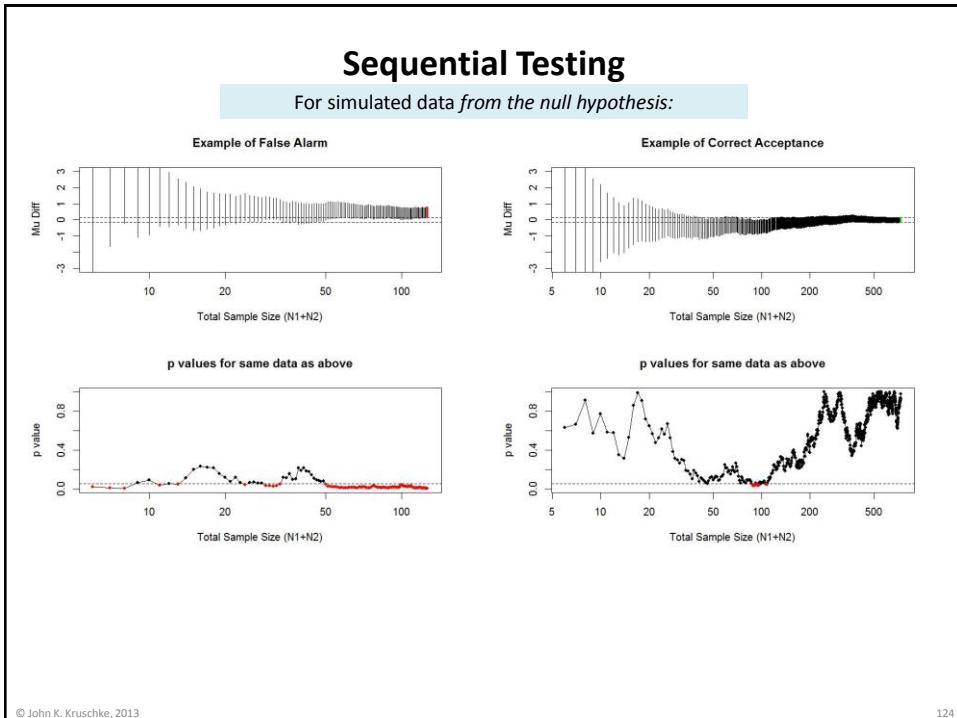
NHST t test:

- p is large for both t and F tests, but NHST cannot accept null hypothesis.
- Need correction for multiple tests, if intended.
- CI's have no distributional info and fickle end points linked to fickle p values, and CI does not indicate probability of parameter value. Hence, **cannot use ROPE method in NHST**.



© John K. Kruschke, 2013

123



Many other topics are in the book, e.g.

- ❖ Bayesian hierarchical **ANOVA, oneway and twoway with interaction contrasts.**
- ❖ The **generalized linear model.**
- ❖ Many types of **regression**, including multiple linear regression, logistic regression, ordinal regression.
- ❖ **Log-linear models vs chi-square test.**
- ❖ **Power:** Probability of achieving the goals of research.
- ❖ All preceded by **extensive introductory chapters** covering notions of probability, Bayes' rule, MCMC, model comparison, etc.

