Introduce Sweave into LyX Beamer Template

I Think Subtitles are Pretentious

Paul Johnson[1] [2]

[1]Department of Political Science
University of Kansas

[2]Center for Research Methods and Data Analysis
University of Kansas

Conference on LyX Presentations, 2014

Outline

What is Sweave?

- We will write R code into a LaTeX document
- When we compile the LaTeX document,
  - the system will trigger an R run
  - and the results are then "automatically" included in the document
- This was invented by Friedrich Leisch (Leisch, 2002a). See the homepage: `https://www.stat.uni-muenchen.de/~leisch/Sweave`.
- A newer variant of the same idea is the R package knitr
- See vignette "code_chunks" in the crmda stationery package.

LyX is not Necessary for this

- Can write in Raw LaTeX
- Instead of editing a *.tex file, edit an *.Rnw file.
- Look for *.Rnw in my Rcourse notes:
  `http://pj.freefaculty.org/guides/Rcourse/plot-1`

Here's the basic Sweave idea

Inside the document, there are "code chunks" in which we embed the R code.

```
<<eval=F>>=
x <- rnorm(100)
mean(x)
@
```

Sweave Details I

There are several arguments we can use inside the brackets. Trial and Error is required to master these. Here is a thumbnail sketch

```
<<chunk10, echo=TRUE, include=FALSE,
    fig=TRUE, height=4, width=5, eval=
    FALSE>>=
##whatever
@
```

chunk10 The chunk's name. It is optional. No spaces or funny symbols, otherwise, do what you like. Name will be handle for the chunk when referred to in the future. Helps in fixing errors because chunks have names.

Sweave Details II

echo If you want the R commands to be shown in the document, say T. To Conceal, say F.

include Set T if you want this chunk to show at the current location in the document. I often set to F so I can run something and manually insert results later

fig if you want this chunk to produce an R graphic, must be T. Otherwise omit

height,width In inches, size of graphic to be created

eval If set to F, then chunk is not "run" through R, but it is syntax-inspected for legal format. Often used for "teaching" purposes, where we don't want to launch a long calculation from a document, but we want to show the code.

Sweave Details III

At one time, it seemed necessary to use `T` and `F`, not "`TRUE`" and "`FALSE`" in the bracketed Sweave chunk starters. That is not true now. Can write full words.

Can reuse chunks later inside another chunk

```
< <chunk30>>=
    <<chunk10>>
@
```

What is the document "workflow", then?

- LyX file: A LyX document we edit.
  - In Ordinary LyX, we then convert to LaTeX and from there to PDF.
  - In Sweave, the LyX is converted to "noweb" literate programming document, Rnw
- Compile from Lyx:
  - export to "Rnw"
  - "R CMD Sweave file.Rnw" processes code chunks, inserts results, creating "file.tex"
  - "pdflatex file.tex" converts the tex to pdf

Is this better/different than knitr?

- There's a longish vignette about this in the crmda stationery package.
- knitr can also do the Rnw -> LaTeX conversion.
- knitr is distinguished by the fact that it can process documents prepared with both
  - LaTeX (Rnw files)
  - Markdown (Rmd files)
- In LaTeX documents, I've found Sweave more dependable (but perhaps that is just my age showing)

What's Good About Weaving (or knitting)?

- Results and graphs are automatically updated presentations!
- Never a mismatch between "pasted in" code and actual results
- "Literate Programming" = "Replicable Research Reports"
- This is easier to do in a LaTeX article than in a presentation, I suggest you might study the article examples before you get too busy on presentations.

This Sounds Easy, but

- Writing Beamer Slides–even without Sweave–is the most difficult kind of LaTeX work.
  - Typos result in compilation failures with arcane error messages
- Now we add
  - R code, which must be correct
  - Beamer markup must be customized to allow the Sweave
- LyX is a convenient LaTeX editor, but it is perhaps not so wonderfully helpful with Sweave and Beamer slides as it is with Articles.

Document Stages where Failures can Occur I

Rnw file: R code errors. No matter what, the Rnw file has to be passed to R for processing.

TeX file: If we run "R CMD Sweave yourFileHere.Rnw", it creates yourFileHere.tex file. R also creates figures and they are set aside in separate files, default PDF. The TeX file that R creates has elaborate markup that will collect the R code, results, plots, and so forth.

PDF: Run pdflatex (or similar) to convert the TeX file.

LyX has exporter for the Rnw file: File -> Export -> Sweave

Initiate Sweave

I have a template with bells and whistles, but if you start fresh this is not difficult.

- File -> New
- Menu: Document -> Settings -> Module. Choose Sweave and "Add"
- For basic Sweave use, this is the only required change in the LyX setup (since version 2.0.2 or so).
- Hit Ctrl-I and type in a code chunk, all is well. I've been using LyX with Sweave for a long time (wrote the original HOWTO on the LyX Wiki), and the programmers have done a beautiful job of making this work with articles.
- Repeat: I don't use the LyX code chunk environment, I use ERT

Initiate Sweave

I have a template with bells and whistles, but if you start fresh this is not difficult.

- File -> New
- Menu: Document -> Settings -> Module. Choose Sweave and "Add"
- For basic Sweave use, this is the only required change in the LyX setup (since version 2.0.2 or so).
- Hit Ctrl-I and type in a code chunk, all is well. I've been using LyX with Sweave for a long time (wrote the original HOWTO on the LyX Wiki), and the programmers have done a beautiful job of making this work with articles.
- Repeat: I don't use the LyX code chunk environment, I use ERT

Various other changes help

- Document -> Settings -> Text Layout -> Vertical Space
- Document -> Settings -> Modules -> Logical Markup

Presentations are Problematic

- Frames in LaTeX must have begin and end statements, LyX GUI has tested various ways to do that.
- LyX 2.1 introduced a new frame setup. All content must be "indented" in LyX, leading to giant margins in document while editing.
- Another problem is that the character styling on which R Sweave depends will not co-operate with a Beamer Frame.

The Error message you can expect I

- Output from R is rejected by Beamer slides

```
Paragraph ended before \FV@BeginScanning was
    complete.

LaTeX Error: \begin{Sinput} on input line 134
    ended by \end{beamer@framepause
Extra }, or forgotten \endgroup.
```

- Sweave uses the LaTeX package Verbatim to render input and output, and, unfortunately, Beamer frames are not compatible
- This is not a LyX problem, it is a LaTeX/Beamer problem
- Fix is to declare slides with an option, "containsverbatim".

Ordinary Slide Frame Bookends.

- Begin

```
\begin{frame}
\frametitle{My fabulous Slide}
```

- end:

```
\end{frame}
```

- Between those things, insert what you like

containsverbatim is a magic word

- A LyX Beamer document will fail if we don't adjust the begin command. This is the fix I incorporate:

```
\begin{frame}[containsverbatim]
```

- When I expect the output to go on for a few pages, I adjust the code like so

```
\begin{frame}[containsverbatim, allowframebreaks
```

- "allowframebreaks" is helpful if you are working quickly and want to generate a lot of slides with same title.

Self Defense Mechanisms

- I always start LyX in a terminal.
  - The R errors are viewable there.
- I configure the LyX temp directory to a place in my home folder, so I can see all the intermediate steps.

Insert some R now

```
> x <- rnorm(1000)
> mean(x)

[1] 0.00149445

> sd(x)

[1] 1.025294
```

Insert some more R now I

```
> x1 <- rnorm(1000)
> mean(x1)

[1] 0.006314803

> sd(x1)

[1] 0.9601718
```

Insert some more R now II

```
> se <- 1.5
> b0 <- 0.2
> b1 <- 0.04
> y <- b0 + b1 * x1 + se * rnorm(1000)
> dat <- data.frame(x1, y)
> m1 <- lm(y ~ x1, data = dat)
> summary(m1)

Call:
lm(formula = y ~ x1, data = dat)

Residuals:
    Min      1Q  Median      3Q     Max
-5.2742 -1.0173  0.0022  1.0306  4.1418
```

Insert some more R now III

```
Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  0.19296    0.04803   4.017 6.33e-05 ***
x1           0.02084    0.05005   0.416    0.677
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘

Residual standard error: 1.519 on 998 degrees of freedom
Multiple R-squared:  0.0001736,         Adjusted R-squared:
F-statistic: 0.1733 on 1 and 998 DF,  p-value: 0.6773
```

Insert some more R now IV

```
> dat$x2 <- rpois(1000, lambda = 2)
> m2 <- lm(y ~ x1 + x2, data = dat)
> summary(m2)

Call:
lm(formula = y ~ x1 + x2, data = dat)

Residuals:
   Min     1Q Median     3Q    Max
-5.391 -1.032  0.027  1.032  4.084

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  0.07301    0.08407   0.868   0.3854
```

Insert some more R now V

```
x1            0.01963    0.05000   0.393   0.6947
x2            0.05892    0.03391   1.738   0.0826 .
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘

Residual standard error: 1.517 on 997 degrees of freedom
Multiple R-squared:  0.003192,          Adjusted R-squared:
F-statistic: 1.596 on 2 and 997 DF,  p-value: 0.2031
```

```
> anova(m2, m1, test = "F")
```

Insert some more R now VI

```
Analysis of Variance Table

Model 1: y ~ x1 + x2
Model 2: y ~ x1
  Res.Df    RSS Df Sum of Sq      F Pr(>F)
1    997 2295.5
2    998 2302.5 -1   -6.9512 3.0191 0.0826 .
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘
```

Insert LaTeX output from R straight into the Document

- Many R functions exist that can write LaTeX tables and other LaTeX objects
- To let Sweave know this is happening, the chunk needs to include "results=tex".

Insert some more R now I

	The Small Model Estimate (S.E.)	The Big Model Estimate (S.E.)
(Intercept)	0.193***	0.073
	(0.048)	(0.084)
Hippotomi	0.021	0.020
	(0.050)	(0.050)
Elephants	.	0.059
		(0.034)
N	1000	1000
RMSE	1.519	1.517
R^2	0.000	0.003
adj R^2	-0.001	0.001

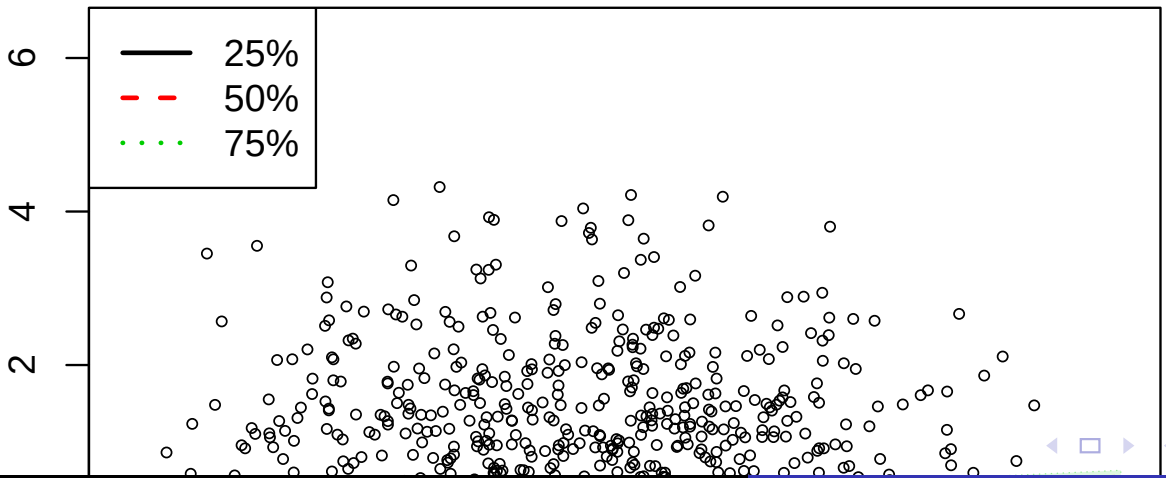
$*p \leq 0.05 ** p \leq 0.01 *** p \leq 0.001$

I want a figure now

I insert "fig=T" into the chunk declaration.

Note how the command runs off the screen. That's a basic Sweave problem.

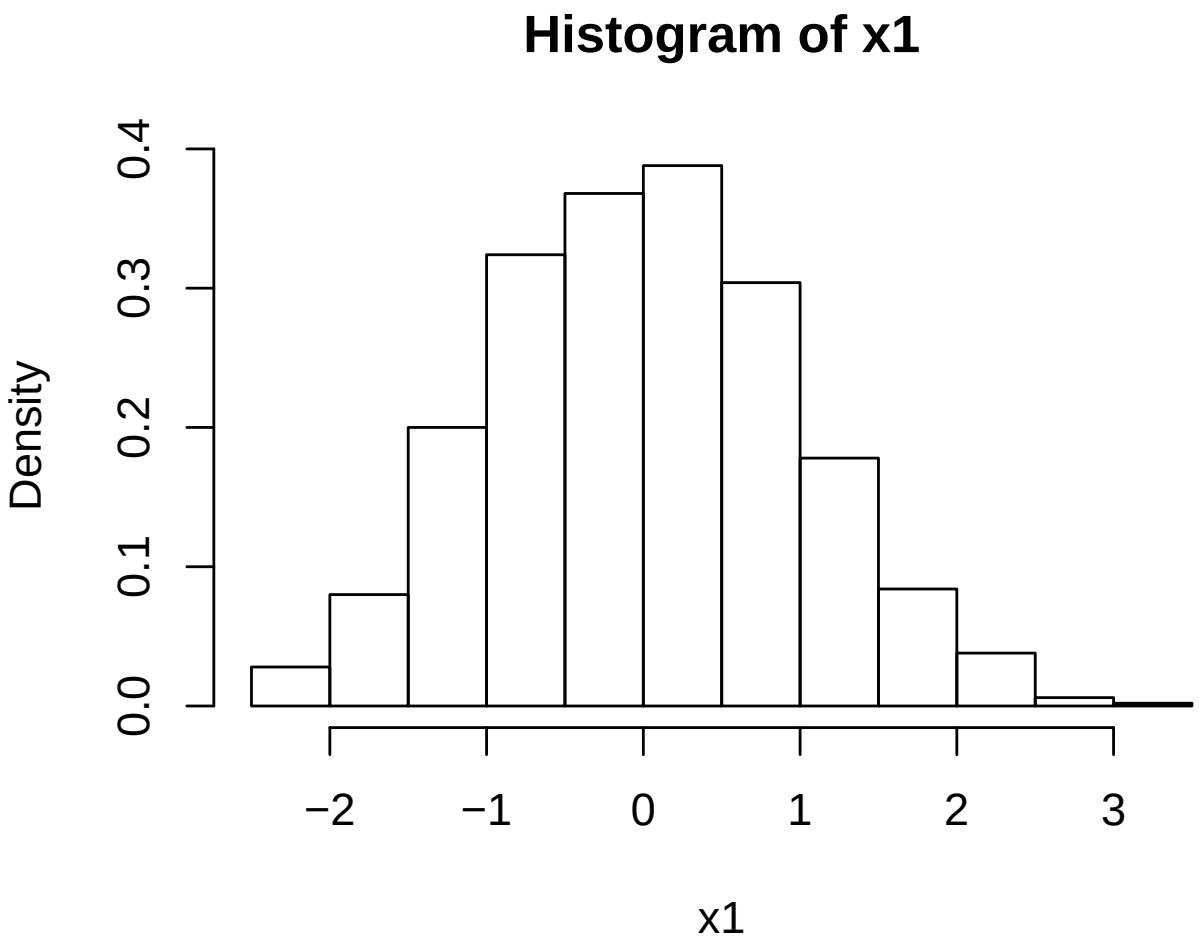
```
> library(rockchalk)
> plotSlopes(m2, plotx = x1, modx = x2, interval = "confide
```



A histogram I

```
> hist(x1, xlab="x1", main = "Histogram of x1", prob = TRUE
```

A histogram II

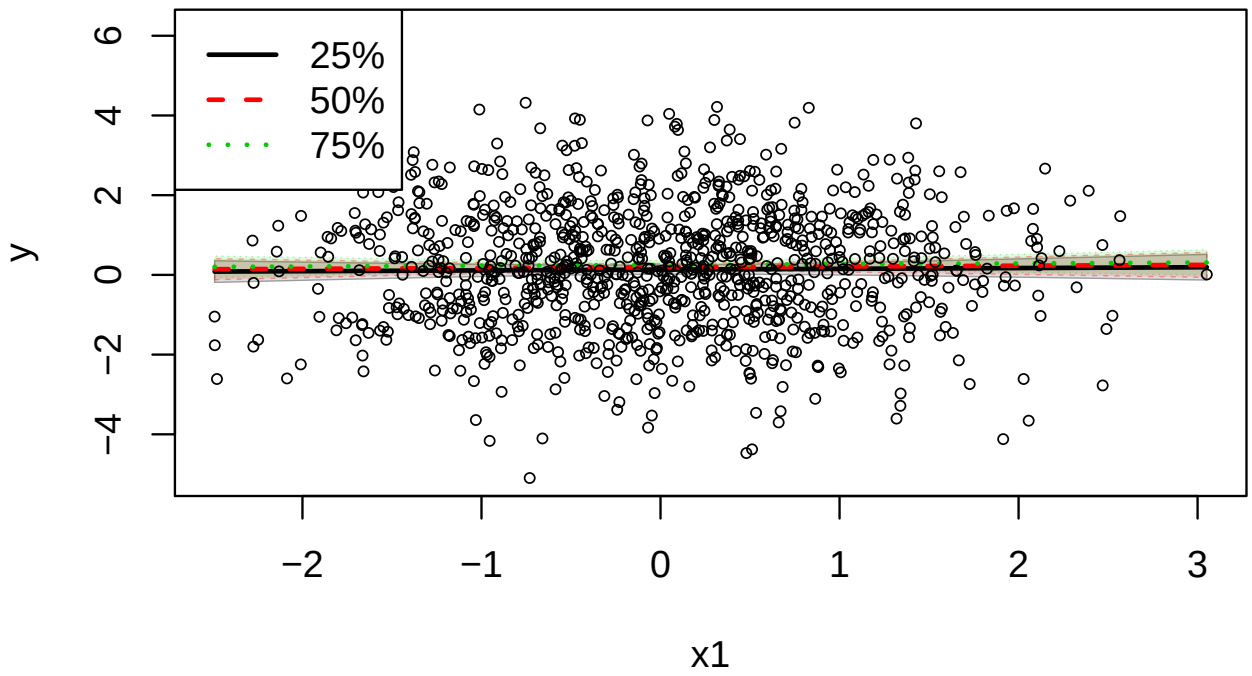


Hopefully, you see the problem

- Code chunks run out of the right margin. That's fatal. The Verbatim class has no "automatic line breaks".
- We can try to adjust the image shape and size when it is manufactured.
- We adjust the chunk
  <<fig=T,height=5,width=7>>=
- I generally find that does not come out "quite right". Resizing the output size often results in ridiculously large fonts. This time, however, the result is not so bad

I want a figure now that's inside the borders

```
> plotSlopes(m2, plotx = x1, modx = x2, interval = "confide
```



Better than Mediocre

- This is a good start for a new Beamer user.
- It is not perfect–not beautiful–but it is certainly better than PowerPoint.
- In beamer-3, I'll lay out some changes that can beautify the output and make the figures easier to wrangle.

References

Leisch, F. (2002a). Sweave: Dynamic generation of statistical reports using literate data analysis. In Härdle, W. and Rönz, B., editors, *Compstat 2002 — Proceedings in Computational Statistics*, pages 575–580. Physica Verlag, Heidelberg. ISBN 3-7908-1517-9.

Leisch, F. (2002b). Sweave, part I: Mixing R and LaTeX. *R News*, 2(3):28–31.