July 8, 2013

Class goals

- This course provides an introduction to Generalized Linear Model (GLM) and Generalized Linear Mixed Model (GLMM)
  - Mathematical backgroung
  - Intuition and conceptualization
  - Geometical interpretation
  - Common issues and solutions for GLM and GLMM analyses
  - Relation to ANOVA
- We will learn
  - how to conduct, interpret and report GLM and GLMM analyses in R
  - how to visualize data in R
  - how to prepare data for visualization and analysis (transformation)
  - The course will be part lecture, part learning by doing.

Course overview

Lecture 1:

- (re-)introducing Generalized Linear Models (GLM)
- relation between GLM and ANOVA
- example (linear) models and geometric interpretation
- Generalized Linear Mixed Models (GLMM)

Course overview

Lecture 1:

- (re-)introducing Generalized Linear Models (GLM)
- relation between GLM and ANOVA
- example (linear) models and geometric interpretation
- Generalized Linear Mixed Models (GLMM)

Lecture 2:

- relation between GLMM and ANOVA
- logistic regression
- common issues and solutions in GLMs and GLMMs
  - identifying and removing outliers
  - coding your predictors
- example models

Course overview

Lecture 3: Beyond linear models

- Binomial models (logistic regression and mixed logit models)
- Empirical logit weighted linear regression
- Poisson models

Course overview

Lecture 3: Beyond linear models

- Binomial models (logistic regression and mixed logit models)
- Empirical logit weighted linear regression
- Poisson models

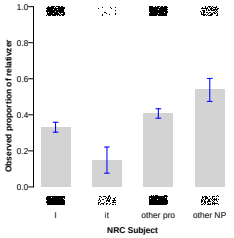
Lecture 4: Tools for data analysis, exploration, and transformation

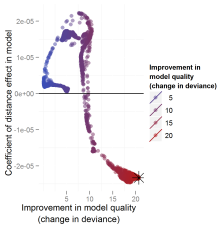
- `plyr`
- `reshape2`

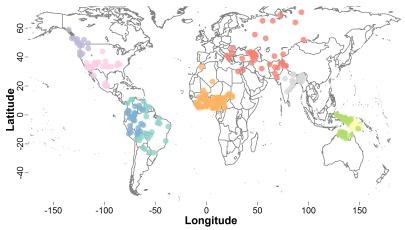
Course overview

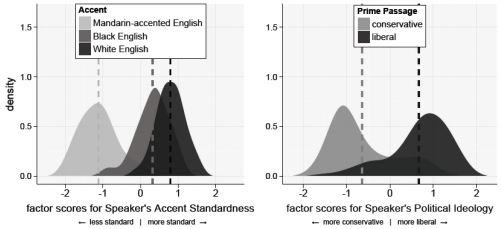
Lecture 5: Visualizing and summarizing your data

- library `ggplot2`





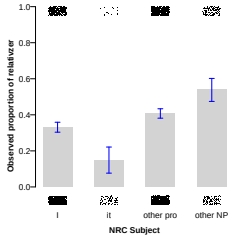


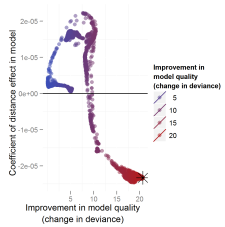


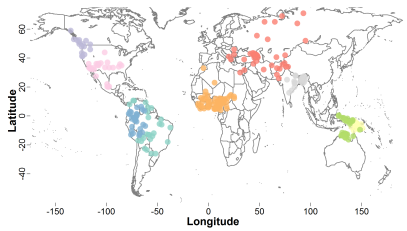
Course overview

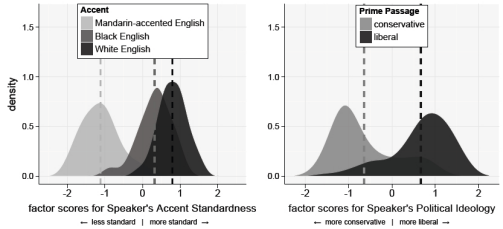
Lecture 5: Visualizing and summarizing your data

- library `ggplot2`









- library `stargazer`
- library `knitr`

Table: Example Stargazer table generated from R

	Dependent variable:		
	(logged) RT		Correct response?
	OLS		*logistic*
	(1)	(2)	(3)
Intercept	6.497***	6.466***	1.664**
	(0.030)	(0.028)	(0.666)
Word frequency (logged)	-0.031***	-0.031***	0.412***
	(0.006)	(0.006)	(0.154)
Native language	0.285***	0.286***	-1.642*
	(0.042)	(0.042)	(0.886)
Trial position	-0.0003**		
	(0.0001)		
Word frequency (logged):Native language	-0.027***	-0.027***	0.261
	(0.009)	(0.009)	(0.212)
Observations	1,659	1,659	1,659
R^2	0.161	0.158	
Adjusted R^2	0.159	0.157	
Akaike Inf. Crit.			520.100
Note:			*p<0.1; **p<0.05; ***p<0.01

Course overview

Lecture 6 and 7: Common issues and solutions in GLMs and GLMMs

- collinearity
- model evaluation
- random effect structure

Course overview

Lecture 6 and 7: Common issues and solutions in GLMs and GLMMs

- collinearity
- model evaluation
- random effect structure

Lecture 8: Remaining issues and continued discussion

- Reporting GLMMs in your article

Oh, we are all so different . . .

Folks in this class represent varied linguistic interest and varied degrees of expertise in statistics and R.

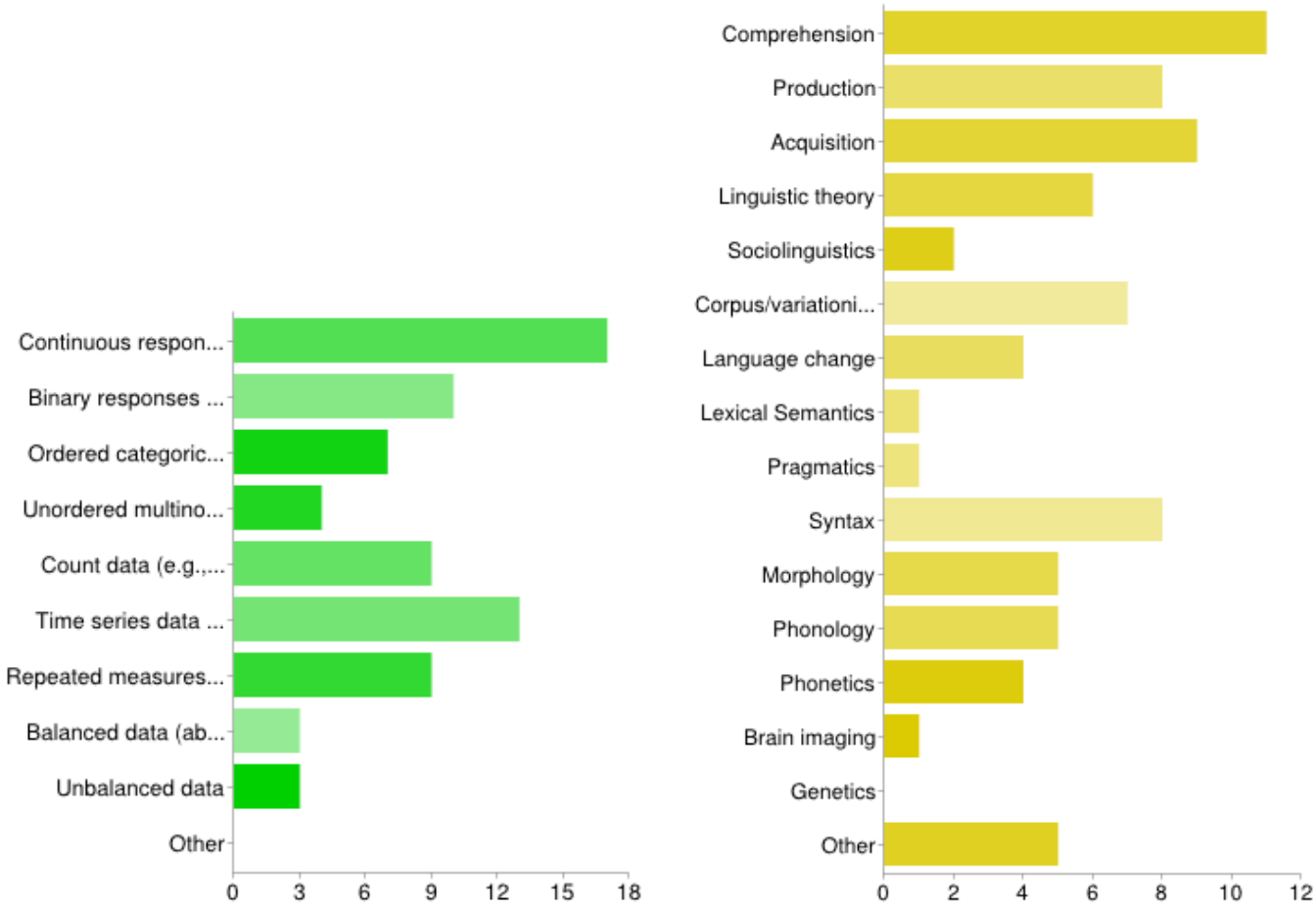


Figure: Background 7 and areas of interest fig-survey:subfig6

Oh, we are all so different . . .

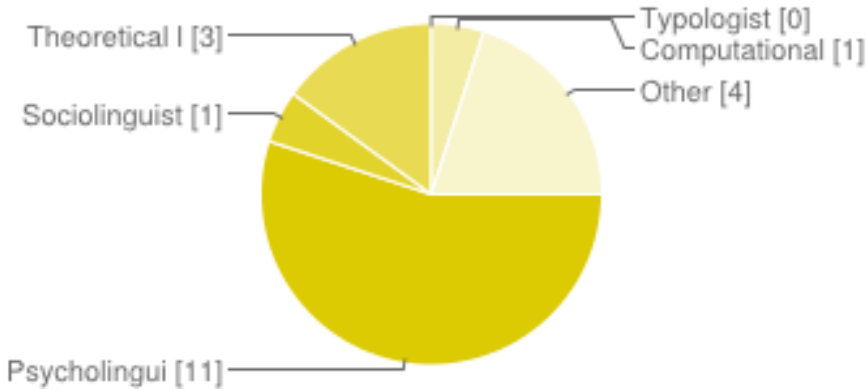
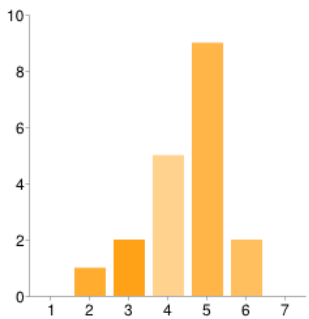
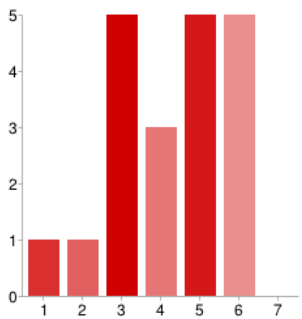


Figure: Background fig-survey:subfig4





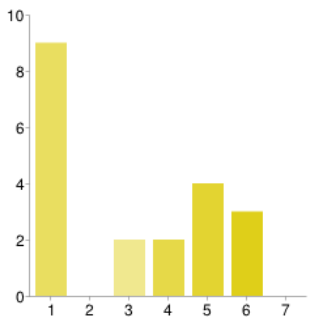


Figure: Expertise in regression fig-survey:subfig1, R fig-survey:subfig2, and lme4 fig-survey:subfig3

→ Please be patient and help each other out. (Change seating arrangement?)

Acknowledgments

- The slides for this class include (usually modified) materials prepared by:
  - Judith Degen (Rochester)
  - Maureen Gillespie (New Hampshire)
  - Dave Kleinschmidt (Rochester)
  - Victor Kuperman (Stanford)
  - Roger Levy (UCSD)

  ... with their permission
- I am also grateful for feedback from:
  - Austin Frank (Rochester)
  - Previous audiences to similar workshops at CUNY, Haskins, Rochester, Buffalo, UCSD, MIT, Iowa, and Groningen.