Counting the Seconds

Working with Reaction Time Data

Matthew Sigal

YORK UNIVERSITÉ UNIVERSITY

Department of Psychology

Canadian Psychological Association Convention
Halifax, Nova Scotia
June 14-16, 2012

Slides available online at `www.matthewsigal.com`

Got time?

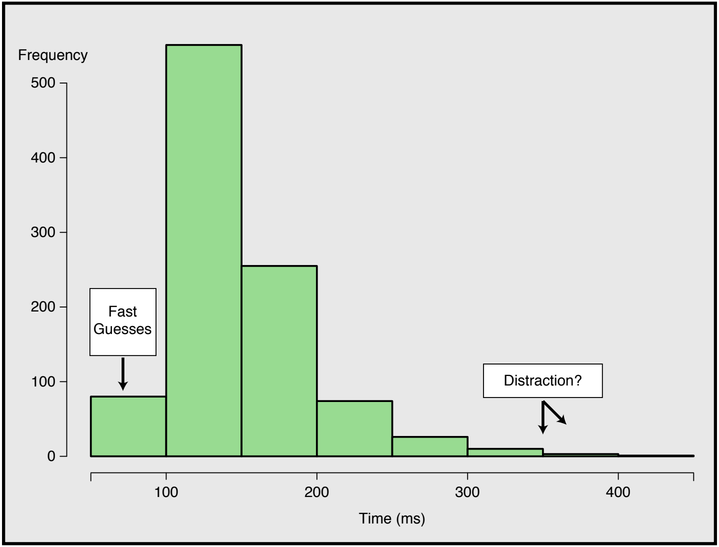
- Directly observable, with high levels of precision.
- Easy to collect.
- Time is a fairly stable construct.
- Allows researchers to make inferences about some fairly substantial psychological matters.
- Ratio level data!

The Bad News

Unfortunately, RT data have some issues...

- Independence between Observations
- Non-Normal Distribution
- Uninformative Outliers

A typical reaction time distribution:



Typical Solutions

Approaches that deal with the non-independence:

- Aggregate or average over trial data.

Approaches that deal with the non-normality of the response:

- Simple transformations (e.g. model $\log RT$).
- Fit data to non-normal distribution (e.g. the "ex-Gaussian").

Approaches that deal with outliers:

- Manual deletion of obvious extreme values.
- Robust or non-parametric tests of mean differences (e.g. *t*- and *F*-tests based upon trimmed means and Windsorized variances).

Each of these approaches treat a particular symptom of RT data.

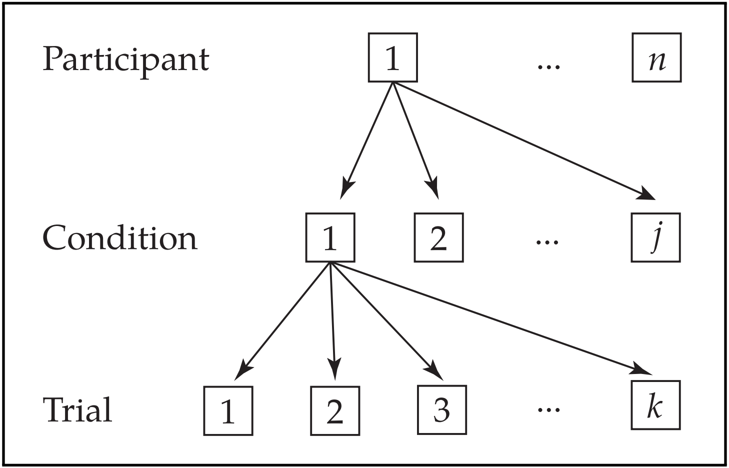
Why Multilevel Modeling?

- Multilevel modeling is a statistical procedure that is a generalization of typical regression methods.
- To account for "nesting", which is common in social science research, regression coefficients are given a model.
- This allows for prediction at even the lowest level of analysis.
- Predictive effects of a variable can be decomposed into direct and contextual (or group) effects.

Model Specification in R

- Using `library(lme4)`, a basic MLM object is specified as:
  - `model <- lmer(Y ~ X + (X | W), data)`
    where, the `X` portion is the "fixed effect" part of the model, and `(X | W)` specifies the "random effect" or nesting structure of variable X being nested within grouping variable W.
- Overall tests of significance and summary statistics on each component of the model can easily be obtained:
  - `anova(model)`
  - `summary(model)`
- Comparisons between nested models can also be conducted:
  - `anova(model1, model2)`

Typical RT Nesting Structure:



MLM Benefits for RT Data:

- We can account for the non-independence of observations.
  - This is the traditional benefit of MLM, as we partition the regressions into within- and between-participant components.
  - Within an RT paradigm, this would allow for between trial comparisons, within- and between-condition comparisons, as well as between-participant comparisons.
- We can account for missing data.
  - For MLM, the data must be entered in long (as opposed to wide) format. This has benefits for analysis:
    - It uses all available information (no listwise/casewise deletion).
    - If missingness is at random, parameter estimates stay robust, even with large amounts of missingness.
    - Otherwise, traditional approaches to missing data can still be utilized (e.g. multiple imputation).

MLM Benefits for RT Data:

- We can account for non-normal distributions.
  - Using `library(lme4)` or `library(MCMCglmm)`, for instance, we can fit generalized models in which the response is fitted as being distributed from a Gamma distribution.
  - This requires a `family` command, e.g.:
    `model <- glmer(Y ~ X + (X | W), data, family = Gamma)`
- We have multiple approaches for dealing with outliers.
  - If we are using a generalized model, then outliers become less of a nuisance since they will not have a large impact on fit.
  - Standard approaches: Condition level Windsorizing.
  - More advanced: *M*-estimators used to apply a weighting algorithm (`weights` command is available in both `lme4` and `nlme`).

MLM Benefits for RT Data:

- We can test hypotheses about group differences.
  - Such tests are highly flexible, and allow for the control of Type I errors (e.g. through the use of a Bonferroni correction).
  - The user needs to supply an `L` matrix with the desired contrasts.
  - This process can be somewhat automated using the `Ldiff()` function in the `R` library `spidadev`.

- We can include and test the usefulness of interesting covariates, as we would in a typical regression model (e.g., gender, education level, ethnicity, et cetera).
  - This is simply done by including the variable (and possibly its interaction with other variables) in the fixed effects portion of the model command.
  - e.g.: `model <- lmer(Y ~ X * V + (X | W), data)`

MLM Benefits for RT Data:

- MLM approach has higher power than traditional ANOVA.
    - For more information see Lachaud & Renaud (2011).
- And, finally, all of this can be fairly easily (and freely) implemented using `R`.

  Some Useful `R` Packages:
    - `foreign` for loading datasets from a variety of sources.
    - `nlme` or `lme4` for basic MLM.
    - `lme4` and `MCMCglmm` for modeling generalized models.
    - `MICE` for multiple imputation of missing values.

  More help: The `R-sig-mixed-models` listserv at

  `https://stat.ethz.ch/mailman/listinfo/r-sig-mixed-models`

Conclusion

Multilevel modeling seems like a perfect fit for analyzing reaction time data.

Utilizing a MLM framework allows researchers to address and actually model many of the typical idiosyncrasies that arise in the study of reaction time.

Future work will yield step-by-step instructions on conducting such an analysis, with emphasis on interpreting the various parameters, and highlighting the gains in interpretation in contrast to more traditional methodologies.

Additional References and Resources:

- Croon, M. A., & van Veldhoven, M. J. (2007). Predicting group-level outcome variables from variables measured at the individual level. *Psychological Methods*, *12*(1): pp. 45-57.
- Baayen, R. H., & Milin, P. (2010). Analyzing reaction times. *Int'l Journal of Psychological Research*, *3*(2): pp. 12–28.
- Snijders, T. A., & Bosker, R. (2011). *Multilevel Analysis*, 2^{nd} Ed. Sage Publications.
- Lachaud, C. M., & Renaud, O. (2011). A tutorial for analyzing human reaction times. *Applied Psycholinguistics*, *32*(2), pp. 389–416.

Slides are available at `www.matthewsigal.com`.

Please send questions or comments to `msigal@yorku.ca`.