Bayesian Analysis of Response Time Data

Mario Peruggia, Trisha Van Zandt, & Meng Chen

The Ohio State University
Outline

- Problems with RT data
- Hierarchical Bayesian Analysis
- Some Data
- Standard Analysis (ANOVA)
- Bayesian Analysis
- Summary and Conclusions
Characteristics of RT

- Positively skewed distributions
  - gamma, Weibull, Wald, ex-Gaussian

- Serial dependencies
  - learning, fatigue (long-term)
  - sequential effects (short-term)
Skewed Distributions

Bayesian Analysis of Response Time Data – p.4
Learning
Sequential Dependencies

Bayesian Analysis of Response Time Data – p.6
Problems with RT Analysis

- No common practice for isolating or acknowledging “nuisance” effects
- No inferential techniques incorporating RT models
Outline

- Problems with RT data
- Hierarchical Bayesian Analysis
- Some Data
- Standard Analysis (ANOVA)
- Bayesian Analysis
- Summary and Conclusions
Hierarchical Bayesian Analysis

\[ f(t|\lambda) = \lambda \exp\{-t\lambda\} \]
\[ \pi(\lambda) = \beta \exp\{\lambda/\beta\} \]

where \( \beta \) is a known constant.

This is a standard conjugate model and the calculations of the posterior can be done in closed form.
Why a Hierarchical Analysis?

Suppose for a particular experimental context

\[ f(t|\lambda) = \lambda \exp\{-t\lambda\}. \]

- Covariates \( X \) and \( Y \)

- \( \ln \lambda = \beta_0 + \beta_1 X + \beta_2 Y + \beta_3 XY \)
Why a Hierarchical Analysis?

Suppose for a particular experimental context
\[ f(t|\lambda) = \lambda \exp\{-t\lambda\}. \]

- Covariates \( X \) and \( Y \)
- \[ \ln \lambda = \beta_0 + \beta_1 X + \beta_2 Y + \beta_3 XY \]

What probabilities can we assign to the \( \beta_i \)s?
Suppose for a particular experimental context
\[ f(t|\lambda) = \lambda \exp\{ -t\lambda \} . \]

- Covariates \( X \) and \( Y \)
- \( \ln \lambda = \beta_0 + \beta_1 X + \beta_2 Y + \beta_3 XY \)

What probabilities can we assign to the \( \beta_i \)s?

Provides a way to estimate effects of fixed factors while maintaining modeling accuracy. (Computations of the posterior now proceed numerically, e.g., using MCMC methods.)
Outline

- Problems with RT data
- Hierarchical Bayesian Analysis
- Some Data
  - An experiment
  - Characteristics of the data
- Standard Analysis (ANOVA)
- Bayesian Analysis
- Summary and Conclusions
Experiment

- Recognition memory: participants studied lists of words and then were tested
- Two tests per day for each of 10 days
- Dependent variable: RT
- Covariates:
  - Response accuracy (correct or incorrect)
  - Word type (old or new)
Learning

Bayesian Analysis of Response Time Data – p.13
Sequential Effects

Bayesian Analysis of Response Time Data – p.14
Covariates: Word type and Response Accuracy

- Average across observations within each condition to obtain subject means
- Use means in repeated measures ANOVA
## ANOVA Table

Type 3 Analysis of Variance for Avg.RT

<table>
<thead>
<tr>
<th>Source</th>
<th>DF</th>
<th>SS</th>
<th>MS</th>
<th>F</th>
<th>Pr &gt; F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Word</td>
<td>1</td>
<td>97335</td>
<td>97335</td>
<td>4.27</td>
<td>0.1308</td>
</tr>
<tr>
<td>Acc</td>
<td>1</td>
<td>135369</td>
<td>135369</td>
<td>2.36</td>
<td>0.2224</td>
</tr>
<tr>
<td>Word*Acc</td>
<td>1</td>
<td>30972</td>
<td>30972</td>
<td>0.61</td>
<td>0.4931</td>
</tr>
<tr>
<td>Sub</td>
<td>3</td>
<td>3780997</td>
<td>1260332</td>
<td>**</td>
<td></td>
</tr>
<tr>
<td>Word*Sub</td>
<td>3</td>
<td>68452</td>
<td>22817</td>
<td>**</td>
<td></td>
</tr>
<tr>
<td>Acc*Sub</td>
<td>3</td>
<td>172398</td>
<td>57466</td>
<td>**</td>
<td></td>
</tr>
<tr>
<td>Word<em>Acc</em>Sub</td>
<td>3</td>
<td>153385</td>
<td>51128</td>
<td>**</td>
<td></td>
</tr>
<tr>
<td>Error</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>15</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Outline

- Problems with RT data
- Hierarchical Bayesian Analysis
- Some Data
- Standard Analysis (ANOVA)
- Bayesian Analysis
  - Hierarchical model
  - Model evaluation
  - Covariate effects
- Summary and Conclusions
The Weibull Distribution

A theoretically motivated likelihood:

\[ f(t \mid r, \lambda) = r \lambda t^{r-1} \exp\{-\lambda t^r\} \]

- Logan’s (1988, 1992) “race” model of automaticity
- One of the limiting distributions for a minimum statistic
- Empirical RT distributions decrease (stochastically) as a power function of time
A Hierarchical Bayesian Model

\[ RT_{i,d,l,e} \sim \text{Weibull}(r_i, \lambda_{i,d,l,e}) \]
A Hierarchical Bayesian Model

\[ RT_{i,d,l,e} \sim \text{Weibull}(r_i, \lambda_{i,d,l,e}) \]

\[ r_i \sim \text{Exponential}(\lambda_r), \quad \lambda_r \sim \Gamma(10^{-1}, 10^{-1}) \]
A Hierarchical Bayesian Model

\[ RT_{i,d,l,e} \sim \text{Weibull}(r_i, \lambda_{i,d,l,e}) \]

\( r_i \sim \text{Exponential} (\lambda_r) \), \( \lambda_r \sim \Gamma (10^{-1}, 10^{-1}) \)

Regression model:

\[ \ln \lambda_{i,d,l,e} = \alpha_{i,d,l} + \beta_1 W_{i,d,l,e} + \beta_2 A_{i,d,l,e} + \beta_3 (W A)_{i,d,l,e} + \eta_{i,d,l,e} \]
A Hierarchical Bayesian Model

\[ RT_{i,d,l,e} \sim \text{Weibull}(r_i, \lambda_{i,d,l,e}) \]

\[ r_i \sim \text{Exponential}(\lambda_r), \quad \lambda_r \sim \Gamma(10^{-1}, 10^{-1}) \]

Regression model:

\[ \ln \lambda_{i,d,l,e} = \alpha_{i,d,l} + \beta_1 W_{i,d,l,e} + \beta_2 A_{i,d,l,e} + \beta_3 (WA)_{i,d,l,e} + \eta_{i,d,l,e} \]

Fixed effects: \[ \beta_1, \beta_2, \beta_3 \sim N(0, 10^3) \]
A Hierarchical Bayesian Model

\[ RT_{i,d,l,e} \sim \text{Weibull}(r_i, \lambda_{i,d,l,e}) \]

\( r_i \sim \text{Exponential}(\lambda_r), \; \lambda_r \sim \Gamma(10^{-1}, 10^{-1}) \)

Regression model:

\[
\ln \lambda_{i,d,l,e} = \alpha_{i,d,l} + \beta_1 W_{i,d,l,e} + \beta_2 A_{i,d,l,e} + \beta_3 (W A)_{i,d,l,e} + \eta_{i,d,l,e}
\]

Fixed effects: \( \beta_1, \beta_2, \beta_3 \sim N(0, 10^3) \)

Autoregressive error:

\( \eta_{i,d,l,e} \sim N(\phi_i \eta_{i,d,l,e-1}, \tau) \)
A Hierarchical Bayesian Model

\[ RT_{i,d,l,e} \sim \text{Weibull}(r_i, \lambda_{i,d,l,e}) \]

\[ r_i \sim \text{Exponential}(\lambda_r), \quad \lambda_r \sim \Gamma(10^{-1}, 10^{-1}) \]

Regression model:

\[
\ln \lambda_{i,d,l,e} = \alpha_{i,d,l} + \beta_1 W_{i,d,l,e} + \beta_2 A_{i,d,l,e} + \beta_3 (WA)_{i,d,l,e} + \eta_{i,d,l,e}
\]

Fixed effects: \( \beta_1, \beta_2, \beta_3 \sim N(0, 10^3) \)

Autoregressive error:

\[ \eta_{i,d,l,e} \sim N(\phi_i \eta_{i,d,l,e-1}, \tau) \]

Learning: \( \alpha_{i,d,l} \sim N(\alpha_{0,d}, \tau_\alpha) \)
Shape and Scale Priors

$r$

$\ln \lambda$

RT

Bayesian Analysis of Response Time Data – p.2
Regression Parameters

\[ \lambda_r, \beta_1, \beta_2, \beta_3, \alpha, \eta \]

\[ r, \ln \lambda, \text{RT} \]
Learning Hyperpriors

\[ \beta_1, \beta_2, \beta_3, \lambda_r, r, \alpha, \alpha_0, \tau_\alpha, \ln \lambda, \eta, \text{RT} \]
Autoregressive errors
Interesting Parameters

- $\lambda_r$
- $r$
- $\beta_1, \beta_2, \beta_3$
- $\alpha_0$
- $\tau_{\alpha}$
- $\phi_0$
- $\tau_{\phi}$
- $\phi$
- $\eta$
- $\ln \lambda$
- $\ln$
Model Evaluation

Cross-validation

- One (randomly chosen) list per day used to fit the model
Cross-validation

- One (randomly chosen) list per day used to fit the model
- Individual differences: one shape parameter or four?
Cross-validation

- One (randomly chosen) list per day used to fit the model
- Individual differences: one shape parameter or four?
- Generate posterior predictive mean RTs for each participant for each day and each list
No individual differences
Fixed Effects: The Rest

- (New & Right) - (New & Wrong)

- (Old & Wrong) - (New & Wrong)

- (New & Right) - (Old & Wrong)
<table>
<thead>
<tr>
<th></th>
<th>mean</th>
<th>2.5% lower limit</th>
<th>97.5% upper limit</th>
<th>P(&gt;0)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\beta_1$</td>
<td>0.4029</td>
<td>-0.0522</td>
<td>0.8510</td>
<td>.96</td>
</tr>
<tr>
<td>$\beta_2$</td>
<td>0.4317</td>
<td>0.0237</td>
<td>0.8408</td>
<td>.98</td>
</tr>
<tr>
<td>$\beta_3$</td>
<td>0.5604</td>
<td>0.0135</td>
<td>1.1129</td>
<td>.98</td>
</tr>
<tr>
<td>$r_1$</td>
<td>12.2014</td>
<td>10.7471</td>
<td>13.3318</td>
<td></td>
</tr>
<tr>
<td>$r_2$</td>
<td>12.5136</td>
<td>11.2307</td>
<td>13.5632</td>
<td></td>
</tr>
<tr>
<td>$r_3$</td>
<td>12.5063</td>
<td>11.0980</td>
<td>13.5992</td>
<td></td>
</tr>
<tr>
<td>$r_4$</td>
<td>6.8793</td>
<td>6.0611</td>
<td>7.6068</td>
<td></td>
</tr>
<tr>
<td>$\phi_1$</td>
<td>0.0519</td>
<td>-0.0289</td>
<td>0.1331</td>
<td>.89</td>
</tr>
<tr>
<td>$\phi_2$</td>
<td>0.0860</td>
<td>0.0015</td>
<td>0.1707</td>
<td>.98</td>
</tr>
<tr>
<td>$\phi_3$</td>
<td>0.0332</td>
<td>-0.0511</td>
<td>0.1184</td>
<td>.79</td>
</tr>
<tr>
<td>$\phi_4$</td>
<td>0.0409</td>
<td>-0.0371</td>
<td>0.1189</td>
<td>.85</td>
</tr>
</tbody>
</table>
Learning Effects

Subject 1

Subject 2

Subject 3

Subject 4
Outline

- Problems with RT data
- Hierarchical Bayesian Analysis
- Some Data
- Standard Analysis (ANOVA)
- Bayesian Analysis
- Summary and Conclusions
RT data have special problems that aren’t well addressed by ANOVA.

ANOVA of mean RTs shows no significant fixed effects.

A Bayesian model directly addresses effects of learning and sequentially dependent errors within a theoretically motivated framework.

Within this framework it is easy to see strong effects of Word and Accuracy on RT.
Conclusions

- Trends and sequentially dependent errors cannot be ignored in analysis.
- Effects of experimental covariates can be masked in the traditional procedure.
- Explicitly modeling nuisance effects within a theoretically motivated framework can be very powerful.
Under certain regularity conditions

- the posterior is asymptotically normal,
- the posterior mean approaches the MLE of the parameter.