'Extended' Bradley-Terry models

David Firth and Heather Turner

Department of Statistics
University of Warwick

Munich, 2010–02–25

Introduction: Bradley-Terry model and extensions

Pair-comparison studies

Sport: player \( i \) beats player \( j \)

Psychometrics: object \( i \) is preferred to object \( j \)

Sport (etc.): interest in players and their attributes
Psychometrics (etc.): interest in judges (subjects) and their attributes

Bradley-Terry model

The basic model:

\[
\text{pr}(i \text{ beats } j) = \frac{\alpha_i}{\alpha_i + \alpha_j},
\]

with \( \alpha_i \) the relative ‘ability’ of object \( i \).

Work with log abilities:

\[
\text{logit}\{\text{pr}(i \text{ beats } j)\} = \log(\alpha_i) - \log(\alpha_j) = \lambda_i - \lambda_j.
\]

Extensions?

We will focus here on three possible directions from the basic model:

1. (Log-)abilities \( \lambda_i \) determined/predicted by object covariate vector \( x_i \).
2. \( \lambda_i \rightarrow \lambda_{ik} \): the ability of object \( i \) varies between different comparisons \( k \).
3. \( i \) versus \( j \), no preference? (‘tied’ comparisons)

'Structured' Bradley-Terry model

\[
\lambda_i = f_i(\beta) + U_i = \sum_r \beta_r x_{ir} + U_i \quad \text{(for example)}
\]

- attributes of objects/players predict ability
- \( U_i \) is random error, with variance \( \sigma^2 \), say — needed in order to allow for imperfect prediction
- \( \Rightarrow \) complex random effects model, with linear predictor

\[
\sum_r (x_{ir} - x_{jr}) \beta_r + (U_i - U_j)
\]

Ability varying between comparisons

\[
\lambda_i \rightarrow \lambda_{ik}
\]

e.g., time-varying covariates,

\[
\lambda_{ik} = \sum_r \beta_r x_{ikr} + U_i
\]

e.g., subject-specific abilities,

\[
\lambda_{ik} = \lambda_{is}
\]

where \( s = s(k) \) identifies the subject who makes comparison \( k \).

e.g., abilities predicted by subject covariates,

\[
\lambda_{is} = \sum_t \gamma_{ist} z_{st} + E_{is}
\]
Extended Bradley-Terry models
Introduction: Bradley-Terry model and extensions

Ability varying between comparisons (continued)

e.g., still with abilities $\lambda_{i\alpha}$ varying between subjects, a particular form likely to be useful is multiplicative interaction,

$$\lambda_{i\alpha} = \lambda_i \exp \left( \sum_t \gamma_{zt} \right) + E_{i\alpha}$$

This last form is not yet implemented in the BradleyTerry2 package; it will require features from the gnm (generalized nonlinear models) package.

Implementation in R: The BradleyTerry2 package

Main new features

- flexible formula interface to modelling fitting function BTm(): allows object-specific, subject-specific, contest-specific variables and random effects [limited implementation]
- efficient data management of multiple data frames

Best of original BradleyTerry package

- translation of formula to appropriate design matrix
- methods for fitted model object, e.g. anova, BTabilities
- missing data handling

Data Structure

> library(BradleyTerry2); data(CEMS); str(CEMS)

List of 3

$ preferences: 'data.frame': 4545 obs. of 8 variables: ...
 ..$ student : num [1:4545] 1 1 1 1 1 1 1 1 1 1 ...
 ..$ school1 : Factor w/ 6 levels "Barcelona","London",...
 ..$ school2 : Factor w/ 6 levels "Barcelona","London",...
 ..$ win1 : num [1:4545] 1 1 NA 0 0 0 1 1 0 1 ... ...

$ students: 'data.frame': 303 obs. of 8 variables:
 ..$ STUD: Factor w/ 2 levels "other","commerce": 1 2 1 2 1 1 1 2 ...
 ..$ ENG: Factor w/ 2 levels "good","poor": 1 1 1 2 1 1 2 1 ...

$ schools: 'data.frame': 6 obs. of 7 variables:
 ..$ Barcelona: num [1:6] 1 0 0 0 0 ...
 ..$ London: num [1:6] 0 1 0 0 0 ...

Model Specification

Model specification is controlled by four arguments to BTm()

outcome a binomial response as accepted by glm().

player1, player2 specify the players in each contest and any other player-specific contest variables in data frames with the same attributes.

id the name of the factor in player1/player2 that gives the identity of the player.

formula a one-sided formula for player ability.

Ties

What to do when neither $i$ nor $j$ is preferred?

Elaborate the Bradley-Terry model? (Rao and Kupper, 1967; Davidson, 1970)

A crude alternative approach/approximation:

$$\text{tie} = \frac{1}{2} \text{‘win’ for each of } i \text{ and } j$$

Suggests a generalization: half $\rightarrow$ some other fraction?

CEMS Data

The CEMS data (Dittrich et al, 1998) concern the preferences of students in selecting a school from the Community of European Management Schools for their international visit.

- 6 CEMS schools are covered in the survey
- students were to choose between each pair of schools (ties allowed)
- further data collected on students e.g. type of degree, language skills
### Standard Bradley Terry Model

A Bradley-Terry model with a separate ability for each player can be specified as follows:

```r
> standardBT <- BTM(outcome = cbind(win1.adj, win2.adj),
+ player1 = data.frame(school = school1),
+ player2 = data.frame(school = school2),
+ id = "school", formula = ~ ., refcat = "Stockholm",
+ data = CEMS$preferences)
```

Or we can use the default id, ". . ."

```r
> standardBT <- BTM(outcome = cbind(win1.adj, win2.adj),
+ player1 = school1, player2 = school2,
+ formula = ~ ., refcat = "Stockholm",
+ data = CEMS$preferences)
```

### Model Summaries

For models with no random effects, `BTM` returns an object which is essentially a "glm" object, hence the usual model summaries can be obtained, e.g. `print()`:

```r
Bradley Terry model fit by glm.fit
Call: BTM(outcome = cbind(win1.adj, win2.adj), player1 = school1,
+ player2 = school2, formula = ~ ., refcat = "Stockholm",
+ data = CEMS$preferences)

Coefficients:
... Barcelona .London .Milano ..Paris ..St.Gallen
0.5379 1.5975 0.3878 0.9064 0.5261

Degrees of Freedom: 446 Total (i.e. Null); 4449 Residual
Null Deviance: 8749 AIC: 9654
Warning message:
In eval(expr, envir, enclos) : non-integer counts in a binomial glm!
```

### Object and Subject Variables

The final model in Dittrich et al, incorporating interactions with subject-variables, can be estimated as follows:

```r
> interactionBT <- BTM(outcome = cbind(win1.adj, win2.adj),
+ player1 = school1, player2 = school2,
+ formula = ~ .. +
+ DEG[student] * St.Gallen[.] +
+ STU[student] * (Paris[.] + St.Gallen[.]) +
+ ENG[student] * St.Gallen[.] +
+ ITA[student] * (London[.] + Milano[.] +
+ SEX[student] * Milano[.] +
+ refcat = "Stockholm", data = CEMS)
```

### Interaction Model

```r
> summary(interactionBT)$coef[, 1:2] / 1.75

Coefficients:
... Barcelona .London .Milano ..Paris ..St.Gallen
1.2476 1.2945

Degrees of Freedom: 42 Total (i.e. Null); 36 Residual
Null Deviance: 78.02
Residual Deviance: 44.05 AIC: 140.5
```

### Baseball Data

The baseball data (Agresti, 2002) gives the results for 7 teams of the Eastern Division of the American League during the 1987 season:

```r
> str(baseball)
'data.frame': 42 obs. of 4 variables:
$ home.team: Factor w/ 7 levels "Baltimore","Boston",..: 5 5 5 5 5
$ away.team: Factor w/ 7 levels "Baltimore","Boston",..: 4 7 6 2 3
$ home.wins: int 4 4 6 4 6 3 4 6 4 6 ...
$ away.wins: int 3 2 3 2 3 0 3 2 3 0 ...
```
Player-specific Contest Variables

```r
> baseball$home.team <- data.frame(team = baseball$home.team, + at.home = 1)
> baseball$away.team <- data.frame(team = baseball$away.team, + at.home = 0)
> baseballModel2 <- update(baseballModel1, + formula = ~ team + at.home)
```

Coefficients:

<table>
<thead>
<tr>
<th>teamBoston</th>
<th>teamCleveland</th>
<th>teamDetroit</th>
<th>teamMilwaukee</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1438</td>
<td>0.7047</td>
<td>1.4754</td>
<td>1.6196</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>teamNew York</th>
<th>teamToronto</th>
<th>at.home</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.2813</td>
<td>1.3271</td>
<td>0.3023</td>
</tr>
</tbody>
</table>

Degrees of Freedom: 42 Total (i.e. Null); 35 Residual
Null Deviance: 78.02
Residual Deviance: 38.64 AIC: 137.1

Comparing Models

```r
> anova(baseballModel1, baseballModel2)
```

Analysis of Deviance Table

<table>
<thead>
<tr>
<th>Model 1: ~team</th>
<th>Resid. Df</th>
<th>Resid. Dev</th>
<th>Df</th>
<th>Deviance</th>
</tr>
</thead>
<tbody>
<tr>
<td>cbind(home.wins, away.wins)</td>
<td>36</td>
<td>44.053</td>
<td>1</td>
<td>5.4106</td>
</tr>
</tbody>
</table>

Springall Data

The `springall` data (Springall, 1973) gives the results of an experiment in which assessors were asked to determine which of two samples had the lesser flavour strength.

Samples were determined by a 3 x 3 factorial design, with factors flavour contentration and gel concentration.

The aim of the experiment was to describe the response surface over the two factors.

Random Effects

The flavour strength over the design region can be modelled by a second order response surface model, with random effects to allow for variation between samples with the same covariates:

```r
> springall.model <- BTm(cbind(win.adj, loss.adj), col, row, + ~ flav[..] + gel[..] + flav.2[..] + gel.2[..] + flav.gel[..] + (1 | ..), + data = springall)
```

Response Surface Model

Bradley Terry model fit by glmxPQL.fit
PQL algorithm converged to fixed effects model

Call: BTm(outcome = cbind(win.adj, loss.adj), player1 = col, + player2 = row, formula = ~flav[..] + gel[..] + flav.2[..] + gel.2[..] + flav.gel[..] + (1 | ..), data = springall)

Coefficients:

<table>
<thead>
<tr>
<th>flav[..]</th>
<th>gel[..]</th>
<th>flav.2[..]</th>
<th>gel.2[..]</th>
<th>flav.gel[..]</th>
</tr>
</thead>
<tbody>
<tr>
<td>-0.4194</td>
<td>-0.32578</td>
<td>0.01665</td>
<td>0.10506</td>
<td>0.02376</td>
</tr>
</tbody>
</table>

Degrees of Freedom: 36 Total (i.e. Null); 31 Residual
Null Deviance: 327.9
Residual Deviance: 15.47 AIC: 113
**Extended Bradley-Terry models**

Implementation in R: The BradleyTerry2 package

**Simplified Model**

```r
> springall.model2 <- update(springall.model, ~ . - flav.2[..])
```

Bradley Terry model fit by glmmPQL.fit

Call:

```r
BTm(outcome = cbind(win.adj, loss.adj), player1 = col, player2 = row, 
formula = ~flav[..] + gel[..] + gel.2[..] + flav.gel[..] + 
         (1 | ..), data = springall)
```

Fixed effects:

```
flav[..] gel[..] gel.2[..] flav.gel[..]  
-0.26366 -0.32690 0.10416 0.02476
```

Random Effects Std. Dev.: 0.1406561

**Extended Bradley-Terry models**

Implementation in R: The BradleyTerry2 package

**Fitted Response Surface**

![Response Surface Plot]

**New work on ties (not yet in BradleyTerry2)**

Davidson (1970) formulation:

\[
pr(tie) = \frac{\nu \sqrt{\alpha_i \alpha_j}}{\alpha_i + \alpha_j + \nu \sqrt{\alpha_i \alpha_j}}
\]

\[
pr(i beats j | not tied) = \frac{\alpha_i}{\alpha_i + \alpha_j}
\]

For inference: either

- discard ties, use the conditional likelihood (robust?)
- ML for all parameters including \( \nu \) (efficient?)

A log-linear model. But too restrictive?

\[
\nu \to \infty: pr(tie) \to 1
\]

\[
\nu \to 0: pr(tie) \propto \nu \sqrt{\alpha_i \alpha_j/(\alpha_i + \alpha_j)} \quad \text{(approx.)}
\]

The single extra parameter \( \nu \) conflates

- overall (max) probability of a tie
- strength of dependence of \( pr(tie) \) on \( \alpha_i, \alpha_j \).

And the strongest dependence allowed (i.e., as \( \nu \to 0 \)) is actually rather weak.

(Same comments apply to the Rao-Kupper model for ties.)

**A ‘2-parameter’ model for ties**

Details omitted here — paper in preparation, preprint to appear soon at http://go.warwick.ac.uk/dfirth