CBC in R with Individual Utilities & Survey Mockups

Chris Chapman, Steven Ellis
ccchapman@google.com, elliss@google.com

Contribution: Improved tools to work with CBC in R

Rcbc 0.2 lets analysts: (1) mock up CBC surveys easily; (2) simulate CBC designs and responses; (3) estimate aggregate and individual-level Hierarchical Bayes utilities easily; (4) import designs and responses from commercial CBC software (e.g., Sawtooth Software [3]) and do additional and parallel analyses in R. Rcbc [1] is Open-Source Software, available from the authors under the GNU General Public License.

A Working CBC Mockup in 5 Lines of R

Rcbc 0.2 adds the ability to tag attributes with friendly names, and to write them to a CSV file as specified by this matrix. This allows easy visualization of the CBC format, and for testing of the CBC using a typical spreadsheet program.

### Example: HB for CBC in 7 Lines of R

We assume that you have fielded a CBC study using Sawtooth Software SSI/Web, and saved the resulting "TAB" file with the Sawtooth-generated design and responses to "MyCBCtabFileFromSawtooth.tab".

```r
# Step 1: Import the data
tmp.rav ← read.csv("~/somedir/MyCBCtabFileFromSawtooth.tab") # load the data
tmp.tab ← tmp.rav[,1:24] # get the design matrix from the relevant columns
tmp.attr ← findSSIattrs(tmp.tab) # get the relevant columns
tmp.win ← tmp.rav[,25] # get the winners from the relevant column

# Step 2: Estimate the HB model
tmp.logitHB ← estimateMNLfromDesignHB(tmp.tab, tmp.win, kCards =3, kTrials =8, kResp =200)

# Step 3: Get the aggregate mean beta utilities and individual-level mean betas
tmp.RHBaavnbeta ← apply(tmp.logitHB$betadraw, 2, mean) # means across draws/respondents
tmp.RHBindubetas ← extractHBbetas(tmp.logitHB, tmp.attr) # mean of draws per respondent
```

These functions use the R package ChoiceModelR [4] (which builds on and updates bayesm [2]). Even when Rcbc’s assumptions don’t fit a project, our code may be a starting point to work with ChoiceModelR.

Limitations and Future Work

Rcbc is not a substitute for best-of-breed commercial software for CBC; it is a supplement.

Primary limitations:
- Only rectangular CBC designs
- Slower estimation than commercial software
- No checks on data quality; good data assumed

Future plans: (1) Do attribute impact estimation from HB models [1]. (2) Handle data from mockup surveys more robustly. (3) Refactor design structure to handle respondent IDs and meta-data in a smart way instead of assuming rectangular blocks.

References