validateHOT - an R package for validating validation tasks and choice modelling tools

08 January

Summary

validateHOT is an R package that provides functions for preference measurement techniques like (adaptive) choice-based conjoint analyses (hereafter (A)CBC) and maximum difference scaling (hereafter MaxDiff). More specifically with the package, users can validate validation tasks, perform market simulations, and rescaleg raw utility scores. The package works with data obtained using, for example, the ChoiceModelR package (Sermas 2022) or Sawtooth's Lighthouse Studio.[1]

Statement of need

Researchers and practitioners use preference measurement techniques for many reasons, for example, simulating markets or to determine the importance of attributes to name a few (Steiner and Meißner 2018). Their ultimate goal is to predict future behavior (Green and Srinivasan 1990). In order to predict accurately and make the right decision, it is essential to ensure that the collected data is valid. One way to test the model's validity is by including validation tasks (e.g., Orme 2015), which are usually fixed tasks (i.e., same across participants) and excluded for utility estimation in hierarchical Bayes (HB) estimation.

The validateHOT provides the relevant tools for these steps: (1) it assesses the model's validity, (2) runs relevant market simulations, (3) converts raw utilities scores into scores that are easy to interpret. Finally, it is an open source tool helping researchers reporting accompanied scripts for their research papers.

State of the field in R

Other packages provide functions to calculate validation metrics, however, these are not always specified for individual part-worth utilities. The Metrics package (Hamner and Frasco 2018), for example, provide functions to run validation metrics such as *mean absolute error* or the five metrics of the confusion matrix. However, converting the output of, for example, estimations using Sawtooth Software or the ChoiceModelR package (Sermas 2022) into the right format, requires some data wrangling. The conjoint package (Bak and Bartlomowicz 2012) provides functions most similar to validateHOT's ones, but no validation functions are included and the package focuses on classical conjoint analysis. Thus, it is limited when applying more common conjoint methods. The logitr package (Helveston 2023) provides market simulations tools, however, no validation metrics such as mean hit probability or hit rate. Figure 1 shows a comparison of validateHOT's functions with current R packages. To the best of our knowledge, a package that converts raw utility scores into validation metrics or running a variety of marketing simulations (especially TURF and TURF ladder) is missing.

validateHOT is introduced with data estimated with Lighthouse Studio. It, however, can be used with data estimated with the packages ChoiceModelR (Sermas 2022), bayesm (Rossi 2023), or STAN (2023), if used with similar settings.

[1]We refer to both validation and holdout tasks interchangeably.

Key functionalities	validateHOT (v 1.0.4)	Metrics (v 0.1.4)	caret (v 7.0-1)	conjoint (v 1.41)	philentropy (v 0.9.0)	logitr (v 1.1.2)	mlogit (v 1.1-1)	support.bws (v 0.4-6)
Confusion matrix	✓	✓	✓					
Creating design matrix				✓		✓	✓	✓
Creating holdout / market scenario	✓			✓		✓		
Estimate utilities				✓		✓	✓	✓
Estimate WTP						✓		
Hit rate	✓							
Kullblack-Leibler-Divergence	✓				✓			
MAE, MedAe, RMSE	✓	✓	✓					
Market Shares	✓		✓	✓		✓		✓
Mean hit probability	✓							
TURF	✓							

Figure 1: Comparison of validateHOT's functions to existing R packages

Key functions

validateHOT's functions can be categorized into four main components, see Table 1. To bring the data into the right format for most functions, we created the `create_hot()` function, which creates each alternatives' total utility by applying the additive utility model.

Table 1: Overview of validateHOT's main components and their corresponding functions

Validation metrics	Confusion matrix	Market simulations	Rescaling scores
hitrate() kl() mae() medae() mhp() rmse()	accuracy() f1() precision() recall() specificity()	freqassort() marksim() reach() turf() turf_ladder()	att_imp() prob_scores() zc_diffs() zero_anchored()

Typical workflow

We provide the workflow for a MaxDiff study (Schramm and Lichters 2024) and a CBC study with a linear-coded price attribute (the vignette provides further examples; Sablotny-Wackershauser et al. (2024)). To run the following code chunks, please install and load the magrittr package (Bache and Wickham 2022).

MaxDiff

Creating validation task / market scenario

After running the HB estimation, the raw utilities must be read into R. For the first example, we assume a validation task with seven alternatives plus the no-buy alternative.

```
hot_mxd <- create_hot(
  data = maxdiff,
  id = "id",
  none = "none",
  prod.levels = list(2, 9, 10, 14, 15, 16, 17),
  method = "maxdiff",
  varskeep = "group",
  choice = "hot"
)
```

Validation metrics

To get, for example, the hit rate (`hitrate()`), we provide the data, the alternatives in the validation task (`opts`), and the actual choice (`choice`).

```
hitrate(
  data = hot_mxd,
  opts = c(option_1:none),
  choice = choice
)
```

Market simulations

We also introduce two functions for market simulations, namely `marksim()` and `turf()`. In the following example, we simulated market shares according to the multinomial logit model (McFadden 1974).

```
marksim(
  data = hot_mxd,
  opts = c(option_1:none),
  method = "sop",
  res = "agg"
)
```

Next, `turf()`, a "product line extension model" (Miaoulis, Parsons, and Free 1990, 29), is a tool to find the perfect assortment that creates the highest reach. This method is useful for MaxDiff studies (Chrzan and Orme 2019, 108). Users can specify the arguments `fixed` (i.e., alternatives that must be part of the assortment) and `prohib` (i.e., forbid specific combinations).

Assuming the user conducted an anchored MaxDiff analysis with ten items (`opts`) and now wants to find the best assortment with a size of three items (`size`). As a threshold that needs to be exceeded (`none`), the user uses the anchor (no-buy alternative).

```
turf(
  data = maxdiff,
  opts = c(option_01:option_10),
  none = none,
  size = 3L,
  approach = "thres"
) %>%
  head(n = 5)
```

CBC

Creating validation task / market scenario

For a CBC, the setup of `create_hot()` is almost the same, only the arguments `prod.levels`, `lin.p`, `coding`, and `method` are new.

```
hot_cbc_linear <- create_hot(
  data = cbc_linear,
  id = "id",
  none = "none",
  prod.levels = list(
    c(3, 6, 10, 13, 16, 20, 24, 32, 248.55),
    c(3, 5, 10, 14, 16, 18, 22, 27, 237.39),
    c(4, 6, 9, 14, 15, 20, 25, 30, 273.15),
    c(4, 5, 10, 11, 16, 19, 26, 32, 213.55),
    c(2, 6, 8, 14, 16, 17, 26, 31, 266.10),
    c(2, 5, 7, 12, 16, 20, 26, 29, 184.50)
  ),
  coding = c(rep(0, times = 8), 1),
  lin.p = "price",
  interpolate.levels = list(c(seq(from = 175.99, to = 350.99, by = 35))),
  method = "cbc",
  choice = "hot"
)
```

Rescaling scores

We can also display the attributes importance scores (`att_imp()`). Therefore, we need to define the attribute levels (`attrib`) and again the coding of the attributes (`coding`).

```
att_imp(
  data = cbc_linear,
  attrib = list(
    paste0("att1_lev", c(1:3)),
    paste0("att2_lev", c(1:2)),
    paste0("att3_lev", c(1:4)),
    paste0("att4_lev", c(1:4)),
    paste0("att5_lev", c(1:2)),
    paste0("att6_lev", c(1:4)),
    paste0("att7_lev", c(1:6)),
    paste0("att8_lev", c(1:6)),
    "price"
  ),
  coding = c(rep(0, times = 8), 1),
  interpolate.levels = list(c(seq(from = 175.99, to = 350.99, by = 35))),
  res = "agg"
)
```

Availability

The package validateHOT is available on GitHub.

Acknowledgments

We would like to thank Sawtooth Software for their great transparent documentation.

References

Bache, Stefan Milton, and Hadley Wickham. 2022. "Magrittr: A Forward-Pipe Operator for r." https://CRAN.R-project.org/package=magrittr.

Bak, A., and T. Bartlomowicz. 2012. "Conjoint Analysis Method and Its Implementation in Conjoint r Package," 239–48. http://keii.ue.wroc.pl/pracownicy/tb/Bak_A_and_Bartlomowicz_T_Conjoint_analysis_method_and_its_implementation_in_conjoint_R_package.pdf.

Chrzan, Keith, and Bryan K. Orme. 2019. *Applied MaxDiff: A Practitioner's Guide to Best-Worst Scaling*. Provo, UT: Sawtooth Software.

Green, Paul E., and V. Srinivasan. 1990. "Conjoint Analysis in Marketing: New Developments with Implications for Research and Practice." *Journal of Marketing* 54 (4): 3–19. https://doi.org/10.1177/002224299005400402.

Hamner, Ben, and Michael Frasco. 2018. "Metrics: Evaluation Metrics for Machine Learning." https://CRAN.R-project.org/package=Metrics.

Helveston, John Paul. 2023. "{Logitr}: Fast Estimation of Multinomial and Mixed Logit Models with Preference Space and Willingness-to-Pay Space Utility Parameterizations" 105. https://doi.org/10.18637/jss.v105.i10.

McFadden, Daniel. 1974. "Conditional Logit Analysis of Qualitative Choice Behavior." In *Frontiers in Econometrics*, edited by Paul Zarembka, 105–42. Economic Theory and Mathematical Economics. New York: Academic Press.

Miaoulis, George, Henry Parsons, and Valerie Free. 1990. "Turf: A New Planning Approach for Product Line Extensions." *Marketing Research* 2 (1): 28–40.

Orme, Bryan K. 2015. "Including Holdout Choice Tasks in Conjoint Studies."

Rossi, Peter. 2023. "Bayesm: Bayesian Inference for Marketing/Micro-Econometrics." https://CRAN.R-project.org/package=bayesm.

Sablotny-Wackershauser, Verena, Marcel Lichters, Daniel Guhl, Paul Bengart, and Bodo Vogt. 2024. "Crossing Incentive Alignment and Adaptive Designs in Choice-Based Conjoint: A Fruitful Endeavor." *Journal of the Academy of Marketing Science*, January. https://doi.org/10.1007/s11747-023-00997-5.

Schramm, Joshua Benjamin, and Marcel Lichters. 2024. "Incentive Alignment in Anchored MaxDiff Yields Superior Predictive Validity." *Marketing Letters*, January. https://doi.org/10.1007/s11002-023-09714-2.

Sermas, Ryan. 2022. "ChoiceModelR: Choice Modeling in r." https://CRAN.R-project.org/package=ChoiceModelR.

Stan Development Team. 2023. "{RStan}: The {r} Interface to {Stan}." https://mc-stan.org/.

Steiner, Michael, and Martin Meißner. 2018. "A User's Guide to the Galaxy of Conjoint Analysis and Compositional Preference Measurement." *Marketing ZFP* 40 (2): 3–25. https://doi.org/10.15358/0344-1369-2018-2-3.