Statistical Analysis of Performance Data – Linear Group

Tuesday 21 January, 2020

Methods and Software

Statistical analysis was based on the local odds ratios Generalized Estimating Equations method (Touloumis, Agresti, and Kateri 2013) as this is implemented in the R (R Core Team 2019) package multgee (Touloumis 2015). In addition, the function `ComparisonStats` was developed to evaluate the statistical significance of the desired comparisons for the accuracy and time data.

```
> ComparisonStats <- function(FittedModel, Lmatrix, alpha = 0.05) {
+     Lmatrix <- matrix(Lmatrix, nrow = 1)
+     ModelBetas <- FittedModel$coefficients
+     ModelVCov <- FittedModel$robust.variance
+     Estimate <- drop(Lmatrix %*% ModelBetas)
+     SdError <- sqrt(drop(Lmatrix %*% ModelVCov %*% t(Lmatrix)))
+     CBs <- Estimate + qnorm(c(alpha/2, 1 - alpha/2)) * SdError
+     pvalue <- 2 * pnorm(-abs(Estimate/SdError))
+     ans <- c(exp(c(Estimate, CBs)), round(pvalue, 4))
+     names(ans) <- c("Estimate", paste0((1 - alpha) * 100, "% LB"), paste0((1 - alpha) *
+         100, "% UB"), "p-value")
+     ans
+ }
```

Import Data

The preference data for the Linear group were obtained by executing the following commands:

```
> library("readxl")
> preference_data <- read_excel("data/mainAllPreference_GS.xlsx")
> names(preference_data) <- gsub("[.]", "_", make.names(names(preference_data), unique = TRUE))
> linear_preference_data <- preference_data[preference_data$Notation == "Linear", ]
> linear_preference_data <- linear_preference_data[linear_preference_data$Preference_Value ==
+     1, ]
```

Analysis of Preference Data

First, the following regression model was fitted to the preference data

$$\log\left[\frac{\Pr(Y_{ij}=k)}{\Pr(Y_{ij}=4)}\right] = \beta_{0k} + \beta_{1k}x_{ij1}$$

for $k = 1, 2, 3$ and where

- $\Pr(Y_{ij} = k)$ is the probability for participant i to find preferable diagram k (1=NWM, 2=NWMA, 3=WM and 4=WMA) for task j (1=subset, 2=disjoint),
- x_{ij1} is the indicator for the *subset* task type,

for $i = 1, \ldots, 102$, corresponding to the individual participants.

```
> library("multgee")
> fitted_model <- nomLORgee(formula = Treatment ~ Question_Type, id = Participant,
+                           data = linear_preference_data, LORstr = "independence")
```

By comparing this model to the null model

```
> reduced_model <- nomLORgee(formula = Treatment ~ 1, id = Participant,
+                            data = linear_preference_data, LORstr = "independence")
> Waldts <- waldts(fitted_model, reduced_model)
> Waldts
Goodness of Fit based on the Wald test

Model under H_0: Treatment ~ 1
Model under H_1: Treatment ~ Question_Type

Wald Statistic=7.2943, df=3, p-value=0.0631
> fitted_model <- reduced_model
```

we can conclude that the preference of a treatment does not depend on the task type (p - value $= 0.0631$). The regression model is further simplified to

$$\log\left[\frac{\Pr(Y_{ij} = k)}{\Pr(Y_{ij} = 4)}\right] = \beta_{0k}$$

for $k = 1, 2, 3$.

Comparison of treatments

Ranking: WM>WMA≈NWM≈NWMA

```
> ## NWMA vs. NWM
> ComparisonStats(fitted_model, c(1, 0, -1))
 Estimate    95% LB    95% UB   p-value
1.0000000 0.4529629 2.2076863 1.0000000
> ## WM vs. NWM
> ComparisonStats(fitted_model, c(0, 0, -1))
Estimate   95% LB   95% UB  p-value
26.71429 11.46271 62.25869  0.00000
> ## WMA vs. NWM
> ComparisonStats(fitted_model, c(1, 0, -1))
 Estimate    95% LB    95% UB   p-value
1.0000000 0.4529629 2.2076863 1.0000000
> ## NWMA vs. WM
> ComparisonStats(fitted_model, c(1, 0, 0))
  Estimate     95% LB     95% UB    p-value
0.03743316 0.01779235 0.07875526 0.00000000
```

```r
> ## NWMA vs. WMA
> ComparisonStats(fitted_model, c(1, -1, 0))
 Estimate    95% LB    95% UB   p-value
0.6363636 0.2820362 1.4358393 0.2763000
> ## WM vs. WMA
> ComparisonStats(fitted_model, c(0, -1, 0))
 Estimate    95% LB    95% UB   p-value
17.000000  9.084984 31.810734  0.000000
```

References

R Core Team. 2019. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. http://www.R-project.org/.

Touloumis, A. 2015. "R Package Multgee: A Generalized Estimating Equations Solver for Multinomial Responses." *Journal of Statistical Software* 64 (8): 1–14.

Touloumis, A., A. Agresti, and M. Kateri. 2013. "Generalized Estimating Equations for Multinomial Responses Using a Local Odds Ratios Parameterization." *Biometrics* 69 (3): 633–40.