003-Model Output

Randomized Control Trial on Epileptic Seizures

June 1, 2015

Abstract

Outputting regression estimates to LaTeXtables is one of the most tedious tasks an analyst has to perform. If an error is made during the data-cleaning or analysis steps, one is forced to manually re-enter the parameter estimates and their standard errors. Furthermore, manual entry of so many numbers is prone to error. In this example, we illustrate how to automate the task of creating tables from model output using the `xtable::xtable` (Dahl, 2014) and `texreg::texreg` (Leifeld, 2013) functions. We illustrate this task using the epileptic seizures data in 59 individuals randomized to the anti-epileptic drug progabide or placebo (Thall and Vail, 1990). For each patient, the number of epileptic seizures was recorded during a baseline period of 8 weeks, and then for 4 consecutive 2-week periods post-randomization.

1 The Models

Let Y_{ij} be the number of seizures on patient i during period j for $i = 1, \ldots, 59$ and $j = 0, 1, 2, 3, 4$, and T_j is the observation time during period j, for $j = 0, 1, 2, 3, 4$ with $T_0 = 8$ weeks and $T_j = 2$ weeks for $j = 1, 2, 3, 4$.

1.1 GLM

The response model is $Y_{ij}|\boldsymbol{\beta}, \mu_{ij} \sim_{ind} \text{Poisson}(\mu_{ij})$ with

$$\begin{aligned}\log(\mu_{ij}) = &\log(T_j) + \beta_0 + \beta_1 \cdot \mathbb{1}_{\{\text{progabide}\}}(\text{treatment}_i) + \\ &\beta_2 \cdot \mathbb{1}_{\{\text{post}\}}(\text{randomization}_i) + \\ &\beta_3 \cdot \mathbb{1}_{\{\text{progabide}\}}(\text{treatment}_i) \times \mathbb{1}_{\{\text{post}\}}(\text{randomization}_i)\end{aligned} \tag{1}$$

1.2 GLMM

We consider the following two-stage model (Wakefield, 2013):
Stage One: The response model is $Y_{ij}|\boldsymbol{\beta}, \mu_{ij}, b_i \sim_{ind} \text{Poisson}(\mu_{ij})$ with

$$\begin{aligned}\log(\mu_{ij}) = &\log(T_j) + \beta_0 + \beta_1 \cdot \mathbb{1}_{\{\text{progabide}\}}(\text{treatment}_i) + \\ &\beta_2 \cdot \mathbb{1}_{\{\text{post}\}}(\text{randomization}_i) + \\ &\beta_3 \cdot \mathbb{1}_{\{\text{progabide}\}}(\text{treatment}_i) \times \mathbb{1}_{\{\text{post}\}}(\text{randomization}_i) + b_i \end{aligned} \tag{2}$$

Stage Two: The random effects model is $b_i \,|\, \sigma_0^2 \sim_{iid} \mathcal{N}(0, \sigma_0^2)$.

1.3 GEE

Marginal mean is given by (1), and variance model $var(Y_{ij}) = \alpha\mu_{ij}$. We consider various forms for the working correlation.

2 Results

	GLM	GLMM	GEE Ind	GEE Ex.	GEE Unst.	GEE AR1
β_0	**1.35** (0.03)***	**1.03** (0.15)***	**1.35** (0.16)***	**1.35** (0.16)***	1.15 (1.64)	**1.31** (0.16)***
β_1	0.03 (0.05)	−0.02 (0.21)	0.03 (0.22)	0.03 (0.22)	−0.14 (2.06)	0.02 (0.21)
β_2	**0.11** (0.05)*	**0.11** (0.05)*	0.11 (0.12)	0.11 (0.12)	−0.06 (0.50)	0.16 (0.11)
β_3	−0.10 (0.07)	−0.10 (0.06)	−0.10 (0.21)	−0.10 (0.21)	0.13 (0.83)	−0.13 (0.27)
AIC	4643.05	2030.67				
BIC	4657.80	2049.11				
Log Likelihood	-2317.53	-1010.34				
Deviance	3574.07					
Num. obs.	295	295	295	295	295	295
Num. groups: subject		59				
Variance: subject.(Intercept)		0.61				
Variance: Residual		1.00				
Num. clust.			59	59	59	59

$^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$

Table 1: Comparing model estimates

References

David B. Dahl. *xtable: Export tables to LaTeX or HTML*, 2014. URL http://CRAN.R-project.org/package=xtable. R package version 1.7-4. 1

Philip Leifeld. texreg: Conversion of statistical model output in R to LaTeX and HTML tables. *Journal of Statistical Software*, 55(8):1–24, 2013. URL http://www.jstatsoft.org/v55/i08/. 1

Peter F Thall and Stephen C Vail. Some covariance models for longitudinal count data with overdispersion. *Biometrics*, pages 657–671, 1990. 1

Jon Wakefield. *Bayesian and frequentist regression methods*. Springer Science & Business Media, 2013. 2

Yihui Xie. *Dynamic Documents with R and knitr*. Chapman and Hall/CRC, Boca Raton, Florida, 2013. URL http://yihui.name/knitr/. ISBN 978-1482203530.

Yihui Xie. knitr: A comprehensive tool for reproducible research in R. In Victoria Stodden, Friedrich Leisch, and Roger D. Peng, editors, *Implementing Reproducible Computational Research*. Chapman and Hall/CRC, 2014. URL http://www.crcpress.com/product/isbn/9781466561595. ISBN 978-1466561595.

Yihui Xie. *knitr: A General-Purpose Package for Dynamic Report Generation in R*, 2015. URL http://yihui.name/knitr/. R package version 1.10.5.

A R Code

```
sessionInfo()
getPckg <- function(pckg) install.packages(pckg, repos = "http://cran.r-project.org")

pckg = try(require(knitr))
if (!pckg) {
    cat("Installing 'knitr' from CRAN\n")
    getPckg("knitr")
    require(knitr)
}

pckg = try(require(texreg))
if (!pckg) {
    cat("Installing 'texreg' from CRAN\n")
    getPckg("texreg")
    require(texreg)
}

pckg = try(require(MASS))
if (!pckg) {
    cat("Installing 'MASS' from CRAN\n")
    getPckg("MASS")
    require(MASS)
}

pckg = try(require(xtable))
if (!pckg) {
    cat("Installing 'xtable' from CRAN\n")
    getPckg("xtable")
    require(xtable)
}

pckg = try(require(geepack))
if (!pckg) {
    cat("Installing 'geepack' from CRAN\n")
    getPckg("geepack")
    require(geepack)
}
```

```
pckg = try(require(lme4))
if (!pckg) {
    cat("Installing 'lme4' from CRAN\n")
    getPckg("lme4")
    require(lme4)
}

pckg = try(require(data.table))
if (!pckg) {
    cat("Installing 'data.table' from CRAN\n")
    getPckg("data.table")
    require(data.table)
}

pckg = try(require(dplyr))
if (!pckg) {
    cat("Installing 'dplyr' from CRAN\n")
    getPckg("dplyr")
    require(dplyr)
}
#################################### R Source code for seizure data Created: May 27, 2015
#################################### Updated: NOTES:
data("epil")
DT <- epil %>% as.data.table
DT.base <- DT %>% distinct(subject)
DT.base[, `:=`(period = 0, y = base)]
DT.epil <- rbind(DT, DT.base)
setkey(DT.epil, subject, period)
DT.epil[, `:=`(post = as.numeric(period > 0), tj = ifelse(period ==
    0, 8, 2))]
fit.glm <- glm(y ~ trt * post, data = DT.epil, family = "poisson",
    offset = log(tj))

fit.glmm <- lme4::glmer(y ~ trt * post + (1 | subject), data = DT.epil,
    family = "poisson", offset = log(tj))

fit.gee.ind <- geepack::geeglm(y ~ trt * post, id = subject,
    offset = log(tj), data = DT.epil, family = "poisson", corstr = "independence")
```

```r
fit.gee.ex <- geepack::geeglm(y ~ trt * post, id = subject, offset = log(tj),
    data = DT.epil, family = "poisson", corstr = "exchangeable")

fit.gee.unst <- geepack::geeglm(y ~ trt * post, id = subject,
    offset = log(tj), data = DT.epil, family = "poisson", corstr = "unstructured")

fit.gee.ar1 <- geepack::geeglm(y ~ trt * post, id = subject,
    offset = log(tj), data = DT.epil, family = "poisson", corstr = "ar1")

texreg::texreg(list(fit.glm, fit.glmm, fit.gee.ind, fit.gee.ex,
    fit.gee.unst, fit.gee.ar1), custom.model.names = c("GLM",
    "GLMM", "GEE Ind", "GEE Ex.", "GEE Unst.", "GEE AR1"), custom.coef.names = c("$\\beta_0$",
    "$\\beta_1$", "$\\beta_2$", "$\\beta_3$"), fontsize = "scriptsize",
    caption = "Comparing model estimates", single.row = T, bold = 0.05,
    float.pos = "H")
extract.geepack <- function(model) {
    s <- summary(model)
    names <- rownames(s$coef)
    co <- s$coef[, 1]
    se <- s$coef[, 2]
    pval <- s$coef[, 4]

    n <- nrow(model.frame(model))
    nclust <- length(s$geese$clusz)

    gof = c(n, nclust)
    gof.names = c("Num. obs.", "Num. clust.")

    tr <- createTexreg(coef.names = names, coef = co, se = se,
        pvalues = pval, gof.names = gof.names, gof = gof, gof.decimal = rep(FALSE,
            length(gof)))
    return(tr)
}

setMethod("extract", signature = className("geeglm", "geepack"),
    definition = extract.geepack)

## example from the geeglm documentation data(dietox)
## dietox£Cu <- as.factor(dietox£Cu) gee1 <- geeglm(Weight ~
```

```
## Cu * (Time + I(Time^2) + I(Time^3)), data = dietox, id =
## Pig, family = poisson('identity'), corstr='ar1') # What the
## heck are we fitting??  summary(gee1) screenreg(gee1)
## texreg(gee1)
```

B Session Information

```
sessionInfo()
```

```
## R version 3.2.0 (2015-04-16)
## Platform: x86_64-pc-linux-gnu (64-bit)
## Running under: Ubuntu 14.04 LTS
##
## locale:
##  [1] LC_CTYPE=en_CA.UTF-8       LC_NUMERIC=C
##  [3] LC_TIME=en_CA.UTF-8        LC_COLLATE=en_CA.UTF-8
##  [5] LC_MONETARY=en_CA.UTF-8    LC_MESSAGES=en_CA.UTF-8
##  [7] LC_PAPER=en_CA.UTF-8       LC_NAME=C
##  [9] LC_ADDRESS=C               LC_TELEPHONE=C
## [11] LC_MEASUREMENT=en_CA.UTF-8 LC_IDENTIFICATION=C
##
## attached base packages:
## [1] stats     graphics  grDevices utils     datasets
## [6] methods   base
##
## other attached packages:
##  [1] dplyr_0.4.1      data.table_1.9.4 lme4_1.1-7
##  [4] Rcpp_0.11.6      Matrix_1.2-0     geepack_1.2-0
##  [7] xtable_1.7-4     MASS_7.3-39      texreg_1.35
## [10] knitr_1.10
##
## loaded via a namespace (and not attached):
##  [1] magrittr_1.5    splines_3.2.0   lattice_0.20-31
##  [4] minqa_1.2.4     highr_0.5       stringr_1.0.0
##  [7] plyr_1.8.2      tools_3.2.0     parallel_3.2.0
## [10] grid_3.2.0      nlme_3.1-120    DBI_0.3.1
## [13] lazyeval_0.1.10 assertthat_0.1  nloptr_1.0.4
## [16] reshape2_1.4.1  formatR_1.2     evaluate_0.7
## [19] stringi_0.4-1   chron_2.3-45
```