What we don't have but need. Some missing R functions in teaching econometrics

dr Rafał Woźniak

Faculty of Economic Sciences, University of Warsaw

Warsaw, 29-09-2019

why? and whyR?

- Despite buoyant number of R packages, some basic areas seem to be undeveloped.
- Some missing R functions over-complicate advanced econometrics teaching, and in consequence, can make R less appealing to students.
- Especially, students not majoring in econometrics might be tempted to use other statistical software.
- To justify why R is the language of choice in teaching, a new package is proposed. The package tries to bridge the gap of missing R functions.

what?

- This paper proposes a preliminary version of the new package (fesuw).
- It consists of a few functions that appear not to be found in available packages.
- Functions for marginal effects of
  - the binary choice models;
  - ordinal logits;
  - tobit models for a given set of values;
- R-squared statistics of static and wide panel models;
- the linktest for binary dependent models functions.
- https://github.com/rafalwozniaque/fesuw

linktest.R

```
Call:
glm(formula = y ~ yhat + yhat2,
    family = binomial(link = model$family$link))


Coefficients:
            Estimate Std. Error z value Pr(>|z|)
(Intercept) -0.03406    0.44867  -0.076 0.939490
yhat         1.03227    0.28856   3.577 0.000347 ***
yhat2        0.01598    0.08737   0.183 0.854863
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

R^2-statistics in R

A. Colin Cameron and Pravin K. Trivedi, *Microeconometrics Using Stata*, Stata Press, 2010, p. 264.

- Let $\hat{u}$ and $\hat{\beta}$ be estimates obtained from a panel model. Let $\rho^2(x, y)$ denote the squared correlation between x and y. Then
- Within R^2: $\rho^2 \left\{ (y_{it} - \bar{y}_i), (\boldsymbol{x}'_{it}\hat{\beta} - \bar{\boldsymbol{x}}'_i\hat{\beta}) \right\}$
- Between R^2: $\rho^2 \left(\bar{y}_i, \bar{\boldsymbol{x}}'_i\hat{\beta} \right)$
- Overall R^2: $\rho^2 \left(y_{it}, \boldsymbol{x}'_i\hat{\beta} \right)$

```
> source("static_wide_panels_R2.R")
> R2_stats = static_wide_panels_R2(fixed)

 R-squared:
          within  =  0.6566299
          between =  0.02762965
          overall =  0.04760428
```

Marginal effects – ome.R

```
> x = c(mean(randdata$income), mean(randdata$female),
        mean(randdata$num))
> ome(model, x)
        alternative1 alternative2 alternative3       at X=
income  -0.000003347 -0.000014684  0.000018031 7505.923082
female!  0.004040917  0.017776474 -0.021817392    0.000000
num     -0.000848907 -0.003724615  0.004573521    3.896849
(!) indicates marginal effect was calculated for discrete
change of dummy variable from 0 to 1

> x = c(7000, 0, 4)
> ome(model, x)
        alternative1 alternative2 alternative3 at X=
income  -0.000003274 -0.000014498  0.000017771  7000
female!  0.004257635  0.018310489 -0.022568124     0
num     -0.000830393 -0.003677344  0.004507737     4
(!) indicates marginal effect was calculated for discrete
change of dummy variable from 0 to 1
```

tobit_marginal_effects.R

```
>   tobit_marginal_effects(tobit1, x, dummies_indices=c())

 Marginal effects of the tobit model
 -----------------------------------
                         y*      E(y|x)  E(y|x,y>0)    Pr(y>0|x)      at X=
age             -0.1945103 -0.04938095 -0.04734533 -0.007553498 32.487521
yearsmarried     0.5764621  0.14634830  0.14031540  0.022385991  8.177696
religiousness   -1.6953493 -0.43040380 -0.41266132 -0.065836196  3.116473
female!         -1.0492278 -0.26637086 -0.25539028 -0.040745096  0.000000
occupation       0.1971279  0.05004549  0.04798247  0.007655148  4.194676
rating          -2.2874235 -0.58071560 -0.55677684 -0.088828459  3.931780
(!) indicates marginal effect was calculated for discrete change of
dummy variable from 0 to 1
```


AND THESE ARE THINGS YOU WANT?
NO, THESE ARE THINGS I NEED.
imgflip.com