Effets marginaux

Marc Thevenin

2022-10-07

Table of contents

# 1. **margins** et **marginaleffects**

**Package margins (2015? -)**

* Auteur: ***TJ.Leeper & al***
* Package maintenant assez ancien, peu ou pas de mise à jours récentes.
* Documentation: [Lien](https://cran.r-project.org/web/packages/margins/vignettes/Introduction.html)
* Assez embêtant à l’Ined: estimation des AME impossible pour le modèle multinomial logistique (non conditionnel) avec la fonction multinom du pkg nnet

**Package marginaleffects (2021-)**

* Auteur: ***VA.Bundock***
* Depuis septembre 2021. Mises à jour régulières
* Documentation très riche: [Lien](https://vincentarelbundock.github.io/marginaleffects/index.html)
* On l’attendait…estime les AME après un modèle multinomial logistique (non conditionnel) avec la fonction multinom du pkg nnet
  + Packages et fonctions prises en charge (>100): [Lien](https://vincentarelbundock.github.io/marginaleffects/articles/supported_models.html)
  + Notes de version (important): [Lien](https://vincentarelbundock.github.io/marginaleffects/news/index.html)
  + L’auteur propose également un package pour éditer les outputs **modelsummary**. Je n’y trouve pas encore mon compte, en particulier pour la sélection d’AME d’un sous ensemble d’outcomes pour un modèle de type ordinal ou multinomial. Les output sont pour l’instant édité de manière brute (il restent très lisibles).

**Installation des packages**

install.packages("margins")

install.packages("marginseffects")  
# ou   
remotes::install\_github("vincentarelbundock/marginaleffects")

library(margins)  
library(marginaleffects)  
library(ordinal)  
library(nnet)

Warning: le package 'marginaleffects' a été compilé avec la version R 4.2.3

Warning: le package 'ordinal' a été compilé avec la version R 4.2.3

Warning: le package 'nnet' a été compilé avec la version R 4.2.3

# 2. Exemples pour le calcul des AME

Base: petit extrait de la bases *nhanes2* (épidemio)

df <- read.csv("D:/D/Marc/SMS/STATA INED/Programmation/R/marginaleffect/nhanes.csv")  
df$hlthstat = as.factor(df$hlthstat)

## 2.1 Modèle logit (binaire)

**Estimation du modèle**

fit = glm(highbp ~ female + black + bmi, family=binomial, data=df)

**Calcul des AME**

## Stata - margins

qui use "D:\D\Marc\SMS\STATA INED\Programmation\R\marginaleffect\nhanes.dta" , clear  
qui logit highbp female black bmi   
margins, dydx(\*)

> , clear  
  
  
  
Average marginal effects Number of obs = 10,335  
Model VCE: OIM  
  
Expression: Pr(highbp), predict()  
dy/dx wrt: female black bmi  
  
------------------------------------------------------------------------------  
 | Delta-method  
 | dy/dx std. err. z P>|z| [95% conf. interval]  
-------------+----------------------------------------------------------------  
 female | -0.087 0.009 -9.65 0.000 -0.105 -0.069  
 black | 0.045 0.015 3.02 0.003 0.016 0.074  
 bmi | 0.032 0.001 36.31 0.000 0.030 0.034  
------------------------------------------------------------------------------

## R - margins

ame = margins(fit)  
summary(ame)

factor AME SE z p lower upper  
 black 0.0450 0.0149 3.0209 0.0025 0.0158 0.0742  
 bmi 0.0321 0.0009 36.3104 0.0000 0.0304 0.0338  
 female -0.0871 0.0090 -9.6499 0.0000 -0.1048 -0.0694

## R - marginaleffects

ame = marginaleffects(fit)  
summary(ame)

Term Contrast Estimate Std. Error z Pr(>|z|) 2.5 % 97.5 %  
 black mean(1) - mean(0) 0.0455 0.015219 2.99 0.00279 0.0157 0.0753  
 bmi mean(dY/dX) 0.0321 0.000884 36.32 < 0.001 0.0304 0.0339  
 female mean(1) - mean(0) -0.0876 0.009183 -9.54 < 0.001 -0.1056 -0.0697  
  
Columns: term, contrast, estimate, std.error, statistic, p.value, conf.low, conf.high

## 2.2 Modèle ordinal (odds proportionnels)

Seulement avec margineffects pour R

## Stata - margins

qui ologit hlthstat female black bmi   
margins, dydx(\*)

Average marginal effects Number of obs = 10,335  
Model VCE: OIM  
  
dy/dx wrt: female black bmi  
  
1.\_predict: Pr(hlthstat==1), predict(pr outcome(1))  
2.\_predict: Pr(hlthstat==2), predict(pr outcome(2))  
3.\_predict: Pr(hlthstat==3), predict(pr outcome(3))  
4.\_predict: Pr(hlthstat==4), predict(pr outcome(4))  
5.\_predict: Pr(hlthstat==5), predict(pr outcome(5))  
  
------------------------------------------------------------------------------  
 | Delta-method  
 | dy/dx std. err. z P>|z| [95% conf. interval]  
-------------+----------------------------------------------------------------  
female |  
 \_predict |  
 1 | -.0235471 .0062116 -3.79 0.000 -.0357216 -.0113726  
 2 | -.0089765 .0023674 -3.79 0.000 -.0136165 -.0043365  
 3 | .0093555 .0024826 3.77 0.000 .0044896 .0142214  
 4 | .0145253 .0038254 3.80 0.000 .0070276 .0220229  
 5 | .0086428 .0022926 3.77 0.000 .0041493 .0131363  
-------------+----------------------------------------------------------------  
black |  
 \_predict |  
 1 | -.1206233 .0102194 -11.80 0.000 -.1406529 -.1005936  
 2 | -.0459833 .0039391 -11.67 0.000 -.0537038 -.0382628  
 3 | .0479248 .0042586 11.25 0.000 .0395781 .0562716  
 4 | .0744078 .0062401 11.92 0.000 .0621774 .0866383  
 5 | .0442739 .003994 11.08 0.000 .0364457 .0521021  
-------------+----------------------------------------------------------------  
bmi |  
 \_predict |  
 1 | -.0088687 .0006413 -13.83 0.000 -.0101257 -.0076117  
 2 | -.0033809 .00025 -13.52 0.000 -.0038709 -.0028908  
 3 | .0035236 .0002694 13.08 0.000 .0029956 .0040517  
 4 | .0054707 .0003941 13.88 0.000 .0046983 .0062432  
 5 | .0032552 .0002564 12.69 0.000 .0027526 .0037578  
------------------------------------------------------------------------------

## R - marginaleffects

J’utilise la fonction **clm** du package ordinal pour estimer le modèle.

fit = clm(hlthstat ~ female + black + bmi, family=binomial, data=df)  
ame = marginaleffects(fit)  
summary(ame)

Group Term Contrast Estimate Std. Error z Pr(>|z|) 2.5 %  
 1 black mean(1) - mean(0) -0.10334 0.007366 -14.03 <0.001 -0.11778  
 2 black mean(1) - mean(0) -0.05951 0.005847 -10.18 <0.001 -0.07098  
 3 black mean(1) - mean(0) 0.02743 0.001697 16.16 <0.001 0.02410  
 4 black mean(1) - mean(0) 0.08029 0.007114 11.29 <0.001 0.06635  
 5 black mean(1) - mean(0) 0.05514 0.005875 9.39 <0.001 0.04362  
 1 bmi mean(dY/dX) -0.00887 0.000641 -13.83 <0.001 -0.01012  
 2 bmi mean(dY/dX) -0.00338 0.000250 -13.52 <0.001 -0.00387  
 3 bmi mean(dY/dX) 0.00352 0.000269 13.08 <0.001 0.00300  
 4 bmi mean(dY/dX) 0.00547 0.000394 13.88 <0.001 0.00470  
 5 bmi mean(dY/dX) 0.00326 0.000256 12.69 <0.001 0.00275  
 1 female mean(1) - mean(0) -0.02358 0.006230 -3.78 <0.001 -0.03579  
 2 female mean(1) - mean(0) -0.00896 0.002366 -3.79 <0.001 -0.01360  
 3 female mean(1) - mean(0) 0.00940 0.002511 3.74 <0.001 0.00448  
 4 female mean(1) - mean(0) 0.01453 0.003833 3.79 <0.001 0.00702  
 5 female mean(1) - mean(0) 0.00861 0.002274 3.79 <0.001 0.00415  
 97.5 %  
 -0.08891  
 -0.04805  
 0.03075  
 0.09423  
 0.06665  
 -0.00761  
 -0.00289  
 0.00405  
 0.00624  
 0.00376  
 -0.01137  
 -0.00433  
 0.01432  
 0.02205  
 0.01307  
  
Columns: group, term, contrast, estimate, std.error, statistic, p.value, conf.low, conf.high

Niveau output cela se complique pour sélectionner un ou un sous ensemble d’outcome…On va y arriver.

## 2.3 Modèle multinomial (logistique)

***Remarque***: l’estimation des AME avec un modèle multinomial de type *stereotype logit model*, adapté aux variables à structure ordinale n’est pas possible avec marginaleffects (modèle estimé avec le pck VGAM). Avec Stata, il est estimé avec la commande **slogit**.

## Stata - margins

qui mlogit hlthstat female black bmi   
margins, dydx(\*)

Average marginal effects Number of obs = 10,335  
Model VCE: OIM  
  
dy/dx wrt: female black bmi  
  
1.\_predict: Pr(hlthstat==Excellent), predict(pr outcome(1))  
2.\_predict: Pr(hlthstat==Very\_good), predict(pr outcome(2))  
3.\_predict: Pr(hlthstat==Good), predict(pr outcome(3))  
4.\_predict: Pr(hlthstat==Fair), predict(pr outcome(4))  
5.\_predict: Pr(hlthstat==Poor), predict(pr outcome(5))  
  
------------------------------------------------------------------------------  
 | Delta-method  
 | dy/dx std. err. z P>|z| [95% conf. interval]  
-------------+----------------------------------------------------------------  
female |  
 \_predict |  
 1 | -.0435641 .0082163 -5.30 0.000 -.0596679 -.0274604  
 2 | .0088463 .008501 1.04 0.298 -.0078154 .0255081  
 3 | .0233105 .0088631 2.63 0.009 .0059391 .0406819  
 4 | .0261863 .0072332 3.62 0.000 .0120095 .040363  
 5 | -.014779 .0050284 -2.94 0.003 -.0246345 -.0049235  
-------------+----------------------------------------------------------------  
black |  
 \_predict |  
 1 | -.1107437 .0162243 -6.83 0.000 -.1425427 -.0789447  
 2 | -.0668401 .0155164 -4.31 0.000 -.0972515 -.0364286  
 3 | .0635409 .0141258 4.50 0.000 .0358548 .091227  
 4 | .060461 .0105785 5.72 0.000 .0397274 .0811946  
 5 | .0535818 .0065457 8.19 0.000 .0407524 .0664112  
-------------+----------------------------------------------------------------  
bmi |  
 \_predict |  
 1 | -.0100225 .0009394 -10.67 0.000 -.0118636 -.0081814  
 2 | -.0026028 .0008992 -2.89 0.004 -.0043652 -.0008403  
 3 | .0037662 .0008894 4.23 0.000 .0020231 .0055094  
 4 | .006369 .0006729 9.46 0.000 .0050501 .0076879  
 5 | .0024901 .0004731 5.26 0.000 .0015627 .0034174  
------------------------------------------------------------------------------

## R - marginaleffects

fit = multinom(hlthstat ~ female + black + bmi, data=df)

# weights: 25 (16 variable)  
initial value 16633.540825   
iter 10 value 15780.594333  
iter 20 value 15554.684483  
final value 15554.562787   
converged

ame = marginaleffects(fit)  
summary(ame)

Group Term Contrast Estimate Std. Error z Pr(>|z|) 2.5 %  
 1 black mean(1) - mean(0) -0.10209 0.011639 -8.77 < 0.001 -0.12490  
 2 black mean(1) - mean(0) -0.07310 0.012772 -5.72 < 0.001 -0.09813  
 3 black mean(1) - mean(0) 0.05028 0.015066 3.34 < 0.001 0.02075  
 4 black mean(1) - mean(0) 0.05872 0.012889 4.56 < 0.001 0.03345  
 5 black mean(1) - mean(0) 0.06620 0.010456 6.33 < 0.001 0.04570  
 1 bmi mean(dY/dX) -0.01002 0.000939 -10.67 < 0.001 -0.01186  
 2 bmi mean(dY/dX) -0.00260 0.000899 -2.89 0.00379 -0.00437  
 3 bmi mean(dY/dX) 0.00377 0.000889 4.23 < 0.001 0.00202  
 4 bmi mean(dY/dX) 0.00637 0.000673 9.46 < 0.001 0.00505  
 5 bmi mean(dY/dX) 0.00249 0.000473 5.26 < 0.001 0.00156  
 1 female mean(1) - mean(0) -0.04372 0.008273 -5.28 < 0.001 -0.05993  
 2 female mean(1) - mean(0) 0.00894 0.008515 1.05 0.29355 -0.00775  
 3 female mean(1) - mean(0) 0.02344 0.008865 2.64 0.00821 0.00606  
 4 female mean(1) - mean(0) 0.02620 0.007197 3.64 < 0.001 0.01209  
 5 female mean(1) - mean(0) -0.01486 0.005061 -2.94 0.00331 -0.02478  
 97.5 %  
 -0.079277  
 -0.048070  
 0.079808  
 0.083978  
 0.086688  
 -0.008181  
 -0.000841  
 0.005509  
 0.007688  
 0.003418  
 -0.027502  
 0.025634  
 0.040811  
 0.040306  
 -0.004944  
  
Columns: group, term, contrast, estimate, std.error, statistic, p.value, conf.low, conf.high