

Econometrie Appliquée (M1-S2)

TD4 - LPM, Modèle logit et probit

Jean-Baptiste Guiffard

2023-04-03

Description de la base de données IndiaSchool

- Nombre d'observation : 902
- Nombre de variables : 20

Variable	Description
chage	Child's age
cedasp	Child's educational aspirations
trtime	Travel time to school in minutes
hstudy	Weekly hours of home study
numsib	Number of siblings
vacwrk	Child works out of school hours
sraven	Student ability (Raven test)
wealth	Parents wealth
male	Childs' gender
snum	School identifier number
parhelp	Parents help with studies at home
bookhom3	More than 100 books in the house
lowcaste	Belongs to low caste
muslim	Religion is muslim
medyrs	Mothers education in years
medyrsq	Mothers education in years squared
lachieve	Log of student achivement
lsraven	Log of student ability (Raven test)
puaind	Dummy Variable =1 if Private School
lachpua	Log of student achievement in private schools

```
setwd("C:/Users/Giffarrd/OneDrive - Université Paris 1 Panthéon-Sorbonne/COURS_DISPENSES/IEDES_2022_2023")  
  
#install.packages('MASS')  
library("MASS")  
# Charger les données  
data_indiaschool <- read.csv2("IndiaSchool.csv")
```

Question 1 : Estimer avec un modèle logit, un modèle probit et Modèle de Probabilité Linéaire (LPM), le lien entre la probabilité de fréquenter une école privée et le nombre de frères et soeurs, les “capacités” de l’élève, la mesure de la richesse du ménage, l’âge de l’élève, le sex de l’élève (=1 si c’est un garçon), la caste de l’élève, la religion de l’élève (=1 si musulman) et le nombre d’années d’éducation de la mère.

Question 2 : Comparez le coefficient estimé de l’éducation de la mère dans les trois modèles. Pour cela vous pouvez faire apparaître tous les coefficients côte à côte en utilisant la commande `stargazer` qui porte le même nom.

```
#install.packages("stargazer")
library(stargazer)
```

Question 3 : Calculer l’effet estimé d’une année supplémentaire d’éducation de la mère sur la probabilité de s’inscrire dans une école privée (à la moyenne des autres caractéristiques).

```
#install.packages('margins')
library("margins")
```

1. Première étape : Récupérer l’objet qui compile les résultats de régression du modèle probit.
2. Créer un objet list qui pour chaque variable donne la moyenne, de cette manière :

```
at_mean <- list(numsib = mean(data_indiaschool$numsib),
                sraven = mean(data_indiaschool$sraven))
```

3. Avec la fonction `margins` du package qui porte le même nom, calculer les marges (dérivées partielles) au point moyen.
4. Interpréter

Et à la main ?

- Récupérer les coefficients estimés avec le modèle probit avec la commande `coefficients` appliqués à l’objet résultat d’estimation.
- Créer un vecteur `means` qui contiendra toutes les moyennes des variables explicatives que vous utilisez dans votre modèle probit
- Calculer la densité de probabilité pour chaque observation du modèle Probit, en utilisant les coefficients du modèle et les moyennes des variables explicatives, en utilisant la fonction `dnorm` et en multipliant chaque coefficient par la moyenne de la variable, en ajoutant aussi le coefficient de la constante.
- Afficher le résultat en multipliant la densité de probabilité pour chaque observation du modèle Probit par le coefficient associé à `medyrs`

Question 4 : En calculant à la main les valeurs, tracer l’évolution de la probabilité au fur et à mesure que le niveau d’éducation de la mère évolue sur l’ensemble de l’échelle.

- Tracer le graphique avec `ggplot2` et la fonction `geom_line`.

Question 5 : Evaluer la qualité de l'estimation

- Récupérer les probabilités estimés pour chaque individu, puis construire une variable binaire $\hat{y} = 1$ si la proba estimée est au-dessus de 0.5, 0 sinon.
- Faire construire un tableau de correspondance, pour mettre en lien la probabilité estimée avec l'outcome binaire observé, avec la fonction table.

Question 6 : Calculer l'effet estimé d'une année supplémentaire d'éducation de la mère sur la probabilité de s'inscrire dans une école privée, à des valeurs moyennes pour toutes les autres variables indépendantes:

- lorsque l'enfant est un garçon
- et lorsque l'enfant est une fille

Question 7 : Calculer l'effet estimé d'une année supplémentaire d'éducation de la mère si l'enfant a un frère ou une sœur et s'il a deux frères ou sœurs, sur la probabilité de s'inscrire dans une école privée, pour des valeurs moyennes de toutes les autres variables indépendantes.