

Application 4 (Python) : analyse de l'impact du programme "Head Start" sur la mortalité

Cette application consiste à évaluer l'impact du programme "**Head Start**" sur la mortalité. Les données proviennent de l'article de Jens Ludwig et Douglas Miller "Does Head Start improve children's life chances? Evidence from a regression discontinuity design" (The Quarterly journal of economics, 2007).

Le programme "Head Start" vise à améliorer les perspectives de vie des enfants issus de milieux défavorisés en leur fournissant des services éducatifs et sociaux. Le programme Head Start comprend des dépenses publiques dans plusieurs domaines. Environ 40 % du budget du programme est consacré à la scolarisation des jeunes enfants. Les autres services comprennent le programme nutritionnel, les services sociaux, les services de santé mentale et les services de santé. Cet ensemble de services du programme Head Start peut affecter la santé (notre objectif dans ce tutoriel) et la scolarisation (l'autre objectif de l'article de recherche) à travers une variété de canaux. En ce qui concerne la santé, la plupart de ces services ont augmenté les chances que les enfants voient un médecin et ont amélioré la détection précoce et le traitement de certaines maladies.

Depuis 1965, le programme américain "Head Start" offre des services préscolaires, sanitaires et autres services sociaux aux enfants pauvres âgés de trois à cinq ans et à leurs familles. Au moment de la rédaction de l'article de recherche, 900 000 enfants bénéficiaient du programme chaque année.

Il est difficile d'évaluer si le programme améliore les chances des enfants dans la vie, car la participation est probablement corrélée aux résultats.

Au moment de la mise en œuvre du programme en 1965, seuls les 300 comtés les plus pauvres des États-Unis recevaient **une assistance fédérale** dans les procédures de demande pour bénéficier du programme. Cette assistance augmente les chances de bénéficier des financements du programme "Head Start". Cette discontinuité dans la mise en œuvre a généré un groupe "de traitement" composé de comtés au-dessus du seuil de pauvreté et un groupe "de contrôle" composé de comtés en dessous.

Le but est d'exploiter cette discontinuité générée par l'assistance fédérale dans les procédures de demande pour estimer l'impact du programme "Head Start" sur la mortalité.

Informations générales :

Nous disposons d'observations au niveau des comtés concernant plusieurs variables. Le tableau suivant donne le descriptif des variables qui seront utilisées dans cette application.

Variables	Description
-----------	-------------

age5_9_sum2	Mortalité des enfants dans le groupe d'âge de cinq à neuf ans sur la période 1973-1983 (les enfants de cinq à neuf ans à cette période auraient été en âge de participer au programme Head Start après le lancement du programme)
rate_5964	Mortalité des enfants dans le groupe d'âge de cinq à neuf ans sur la période 1959 et 1964, avant la mise en œuvre du programme
Assistance	=1 si le comté fait partie des 300 plus pauvres
povrate60	Taux de pauvreté en 1960
povrate	Taux de pauvreté normalisé de sorte que les 300 comtés
PovAss	Interaction entre "povrate" et "Assistance"

I. Première partie : Statistiques descriptives et analyses graphiques

1. Produisez un histogramme pour la variable taux de pauvreté ("povrate") avant la mise en place du programme "Head Start" (en 1960). Que pouvez-vous conclure concernant cette distribution et notre stratégie d'estimation ?
2. Comparez les comtés proches du seuil. Pour cela, choisissez arbitrairement de n'utiliser que les comtés dont le taux de pauvreté est situé à au plus 8% en dessous et au-dessus du seuil, soit une largeur de bande dans le taux de pauvreté de 16% autour du seuil. Générez la variable appelée « bande » qui identifient ces comtés se trouvant autour du seuil.
3. Faites un tableau croisé entre « bande » et « Assistance ». Commentez
4. Observez-vous des différences significatives de moyenne pour la variable rate_5964 qui correspond à la variable de résultat qui nous intéresse (la mortalité) avant l'implémentation du "Head Start" autour du seuil ? Commentez en quoi le résultat de la comparaison est intéressant.
5. Faites une régression linéaire dont la variable expliquée est la variable de mortalité "rate_5964" et les variables explicatives sont "Assistance", "povrate" et "PovAss". Calculez les valeurs prédites que vous stockez dans la variable "LinearBefore". Représentez graphiquement "LinearBefore" en fonction "povrate".

II. Deuxième partie : mesure de l'impact

1. Faites une régression linéaire dont la variable expliquée est la variable de mortalité "age5_9_sum2" et les variables explicatives sont "Assistance", "povrate" et "PovAss". Calculez les valeurs prédites que vous stockez dans la variable "Linear". Représentez graphiquement "Linear" en fonction "povrate". Quel est l'effet du programme ?
2. Reprenez la question 1 en utilisant uniquement les comtés situés à au plus 8% en dessous et au-dessus du seuil, soit une bande de 16% autour du seuil. Commentez.
3. Reprenez la question 1 en utilisant uniquement les comtés situés à au plus 16% en dessous et au-dessus du seuil, soit une bande de 32% autour du seuil. Commentez.
4. Reprenez la question 1 en ajoutant le carré de "povrate" dans les variables explicatives. Commentez.