Program Evaluation

Exercise 1: Answer the following questions. When using math formula, make sure you define the variables used. (10 pts)

a. (1.5 pts) Explain what is the difference between one-sided and two-sided compliance problems. Give an example of each case.

b. (1.5 pts) What is a randomized controlled experiment. Give an example.

c. (2 pts) Define the propensity score? What is it used for?

d. (2 pts) Specify which treatment effect we can identify in a fuzzy regression discontinuity design and give its expression.

e. (1.5 pts) Explain what is the difference between matching with replacement and matching without replacement. Which one delivers the most efficient estimate?

f. (1.5 pts) Explain what is the difference between parametric and non-parametric estimation.

Exercise 2: (10 pts)

We want to investigate whether higher education causes higher wages. We use a dataset in which for n individuals we observe their wage rate (in logarithms, y), whether they have been to college ($c = 1$ if they have been, 0 if not), age and gender ($s = 1$ are males).

We regress the logarithm of the wage rate, y_i, on the college variable c_i and we get for the coefficient of education, an estimate of 0.048 with a standard error (s.e.) of 0.0030. This estimate is denoted $\hat{\beta}_1$. We call x the other variables age and gender. We also linearly regress the logarithm of the wage rate, y_i, on the college variable c_i AND variables x_i and we get for the coefficient of education, an estimate of 0.068 with a standard error (s.e.) of 0.0034. This estimate is denoted $\hat{\beta}_2$.

Question 1: (2 pts) Define precisely what the treatment is, what the potential outcomes are and what the average treatment effect on the treated (ATT) is. Explain with words or with formula.

Question 2: (2 pts) Describe the linear regression that would lead to estimates of the ATT in the population when the potential outcomes are given by:

$$y(d) = \alpha(d) + \beta(d)c + x\gamma(d) + \varepsilon(d)$$

when x consists of age and gender.

Question 3: (1 pt) Under what condition, $\hat{\beta}_1$ is a consistent estimate of ATT.

Question 4: (1 pt) Under what conditions, $\hat{\beta}_2$ is a consistent estimate of ATT.

Question 5: (1 pt) Explain what the common support assumption means in this context.

Question 6: (3 pts) Compare the estimated ATT obtained using Question 2 to $\hat{\beta}_1$ and $\hat{\beta}_2$ in terms of their statistical properties. You might need to formulate some additional conditions.

Program Evaluation

Exercice 1: Répoondez aux questions suivantes. Si vous utilisez des formules mathématiques, veillez à bien définir les variables utilisées. (10 pts)

a. (1.5 pts) Expliquer la différence entre les problèmes de one-sided et two-sided compliance. Donner un exemple pour chacun.

b. (1.5 pts) Qu'est-ce qu'une expérience aléatoire contrôlée? Donner un exemple.

c. (2 pts) Définissez le score de propensité (propensity score)? À quoi sert-il?

d. (2 pts) Spécifiez quel effet de traitement on peut identifier dans un fuzzy regression discontinuity design et donner son expression mathématique.

e. (1.5 pts) Expliquer la différence entre matching avec remise et matching sans remise (with/without replacement). Lequel donne l'estimateur le plus efficace?

f. (1.5 pts) Expliquer la différence entre estimation paramétrique et estimation non-paramétrique..

Exercice 2: (10 pts)

On cherche à savoir si un niveau d'étude élevé implique un salaire élevé. On utilise des données sur n individus pour qui on observe le salaire (en logarithme, y), s'ils sont allés à l'université ($c = 1$ s'ils y sont allés, 0 sinon), l'âge et le sexe ($s = 1$ pour les hommes).

On régresse le logarithme du salaire, y_i sur la variable université (college) c_i et on obtient un estimateur de 0.048 pour cette variable avec un écart-type (s.e.) de 0.0030. On appelle cet estimateur $\hat{\beta}_1$. Soient x les autres variables âge et sexe. On régresse le logarithme du salaire, y_i sur la variable université c_i ET les autres variables x_i et on obtient comme estimateur du coefficient de la variable d'éducation 0.068 avec un écart-type (s.e.) de 0.0034. On appelle cet estimateur $\hat{\beta}_2$.

Question 1: (2 pts) Définir precisément quel est le traitement, quels sont les outcomes potentiels et qu'est-ce-que que l'effet de traitement sur les traités (ATT). Expliquer avec des mots ou des formules mathématiques.

Question 2: (2 pts) Décrire la régression linéaire qui donnerait un estimateur de l'ATT quand les outcomes potentiels sont donnés par:

$$y(d) = \alpha(d) + \beta(d)c + x\gamma(d) + \varepsilon(d)$$

où x représente l'âge et le sexe.

Question 3: (1 pt) Sous quelle condition $\hat{\beta}_1$ est un estimateur convergent de l'ATT.

Question 4: (1 pt) Sous quelles conditions, $\hat{\beta}_2$ est un estimateur convergent de l'ATT.

Question 5: (1 pt) Expliquer ce que l'hypothèse de support commun (common support) signifie dans ce contexte.

Question 6: (3 pts) Comparer l'estimateur d'ATT obtenu à la Question 2 avec $\hat{\beta}_1$ et $\hat{\beta}_2$ et leurs propriétés statistiques. Vous pourriez avoir besoin de formuler des hypothèses supplémentaires.