



THE WORLD BANK



Session technique

Session I:

inférence causale

Motivation

- ❑ Les questions de recherche qui génèrent le plus d'études dans les sciences de la santé sont de nature causale
- ❑ Par exemple:
 - Quelle est l'efficacité d'un médicament donné à une population donnée?
 - Quelle est la proportion de décès causés par une maladie donnée qui auraient pu être évités par un traitement donné ou une mesure politique?

Motivation

- Les questions empiriques les plus difficiles en économie se rapportent aux relations de cause-à-effet:
 - La décentralisation scolaire améliore-t-elle la qualité des écoles?
 - Une année d'enseignement supplémentaire est –elle *la cause* d'un revenu plus élevé?
 - Les transferts conditionnels(ex. allocations familiales) sont-ils *la cause* d'une meilleure santé infantile?

Une définition intuitive de la cause

- Ian a pris le médicament le 1er Septembre 2003
 - Cinq jours plus tard, il meurt

- Si Ian n'avait pas pris le médicament le 1er Septembre 2003 (toutes choses étant égales par ailleurs)
 - Cinq jours plus tard, il serait encore vivant

- La prise de médicament est-elle la cause de la mort d'Ian?

Une définition intuitive de la cause

- Jim n'a pas pris le médicament le 1er Septembre 2002
 - Cinq jours après il était vivant

- Si Jim avait pris le médicament le 1er Septembre 2002(toutes choses étant égales par ailleurs)
 - Cinq jours après il aurait été vivant

- Le médicament est-il cause de la survie de Jim?

Raisonnement humain relatif à l'inférence causale

- Nous comparons (souvent , seulement mentalement)
 - Le résultat quand l'action *A est présente*
 - Le résultat quand l'action *A est absente*
 - Toutes choses étant égales par ailleurs

- Si les deux résultats diffèrent, nous disons que l'action *A a un effet causal*
 - causatif ou préventif

- En épidémiologie, *A s'appelle communément* l'exposition ou le traitement

Motivation

L'intérêt de ces questions est motivé par des:

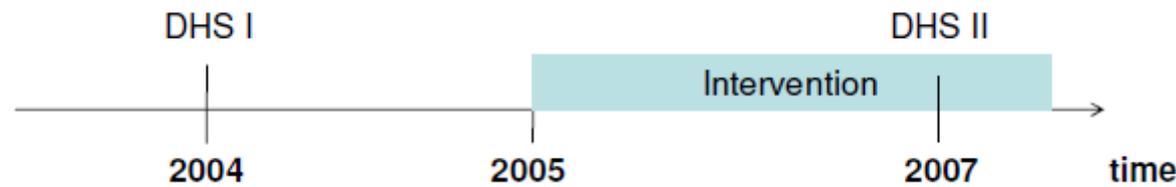
- Préoccupations des politiques
 - ▣ Les programmes publiques réduisent-ils la pauvreté?

- Considérations théoriques

- Problèmes rencontrés par les décideurs individuels

Intuition du problème: un exemple hypothétique

- Supposons que nous voulons estimer les effets de programme de vaccinations. Après des recherches, nous découvrons qu'une ONG fortement motivée a donné des vaccins à tous les nouveaux-nés depuis 2005 dans un district A d'un pays en développement. Nous découvrons qu'il y a une enquête démographique et sanitaire (enquête de référence DHS) avant l'intervention et après. La ligne temporelle se présente comme suit:



Vaccins et mortalité infantile

- ❑ Après quelques recherches sur le Web, vous trouvez une brochure de l'ONG qui déclare *”grâce à notre intervention, la mortalité infantile a chuté de 122 en 2004 à 97 en 2007”*.
- ❑ Est-ce convaincant? Pourquoi oui ou pourquoi non?

Vaccins et mortalité infantile

- ❑ “les différences simples” dans le temps sont difficiles à interpréter; il se pourrait que:
 - a. le programme ait effectivement été causal pour l’effet total
 - b. d’autres facteurs survenus entretemps dans le pays et la région pourraient avoir eu une influence;
 - c. le programme a eu en fait un effet négatif, mais d’autres facteurs ont occulté ce dernier
- ❑ Les changements absolus sont de faibles preuves d’effets causaux -> il serait beaucoup plus convaincant de montrer que la zone vaccinée s’est mieux comportée que des zones similaires (avoisinentes): le déclin de la mortalité est-il oui ou non plus important dans le district A que dans les autres districts sur la même période de temps?

Analyse statistique normalisée

- **Outils**: probabilité et autres techniques d'estimation
- **But**: inférer les paramètres d'une distribution à partir d'échantillons tirés de cette distribution.
- **Utilisations**: à l'aide de tels paramètres on peut:
 - Inférer une association parmi les variables,
 - Estimer la probabilité d'évènements passés et futurs,
 - Mettre à jour la probabilité des évènements à la lumière de mesures et preuves nouvelles.
- **Condition** pour que ceci marche bien: les conditions expérimentales doivent rester les mêmes.

Analyse causale

- ❑ L'analyse causale va plus loin que l'analyse statistique normalisée:
- ❑ Son but est d'inférer les aspects du processus de production de données.
- ❑ à l'aide de tels aspects , on peut déduire
 - la probabilité d'évènements sous *conditions stationnaires* (comme dans l'analyse statistique normalisée) ,
 - et aussi la dynamique des évènements soumise à des *conditions changeantes*.

Analyse causale

- “*dynamique des évènements soumise à conditions changeantes*” comprend:
 1. Prédire les effets des interventions
 2. Prédire les effets de changements spontanés
 3. Identifier les causes d'évènements rapportés

Causalité par opposition à corrélation

- Analyse statistique normalisée/théorie des probabilités:
 - Le mot “cause “ n’est pas dans son vocabulaire
 - Nous permet de dire que deux évènements sont mutuellement corrélés ou dépendants
 - Si nous trouvons l’un, nous pouvons nous attendre à trouver l’autre
- Ceci n’est pas suffisant pour les responsables politiques
 - Ils recherchent une logique à leurs décisions politiques: si nous faisons XXX, alors obtiendrons-nous YYY?
 - Ainsi devons-nous ajouter un vocabulaire de causalité au langage des probabilités.

Généralisations de la théorie du contrefactuel

- (a) Les effets causaux dans une sous-population
- (b) Exposition et résultat non-dichotomiques
- (c) Résultats contrefactuels non-déterministes
- (d) Expositions variant dans le temps

Théories du contrefactuel

- Neyman (1923)
 - Effets des expositions ponctuelles dans des expériences aléatoires
- Rubin (1974)
 - Effets des expositions ponctuelles dans des études observationnelles et aléatoires
- Robins (1986)
 - Effets des expositions de temps variable dans des études observationnelles et aléatoires



Le modèle causal de Rubin

- Définissons la population par U .
Chaque unité de U est notée u .
- Pour chaque $u \in U$, il existe une valeur associée $Y(u)$ de la variable d'intérêt Y , que nous appelons la variable-réaction.
- Rubin démontre que les causes ne sont que les choses qui pourraient être des traitements dans des expériences hypothétiques.

- Pour simplifier , nous supposons qu'il y a seulement deux causes ou niveaux de traitement.
- Soit D une variable qui indique la cause à laquelle chaque unité dans U est exposée:

$$D = \begin{cases} 1 & \text{if unit } u \text{ is exposed to treatment} \\ 0 & \text{if unit } u \text{ is exposed to control} \end{cases}$$

- Dans une étude contrôlée, D est définie par l'expérimentateur.
- Dans une étude non-contrôlée, D est déterminée par des facteurs échappant au contrôle de l'expérimentateur.

- La réponse Y est potentiellement affectée selon que u ait reçu le traitement ou non (témoin).
- Ainsi, nous avons besoin de deux variables-réactions:

$Y_1(u)$ is the outcome if unit u is exposed to treatment

$Y_0(u)$ is the outcome if unit u is exposed to control

$$D = \begin{cases} 1 & \text{si l'unité } u \text{ est exposée au traitement} \\ 0 & \text{si l'unité } u \text{ est exposée au témoin} \end{cases}$$

$Y_1(u)$ est le résultat si l'unité u est exposée au traitement

$Y_0(u)$ est le résultat si l'unité est exposée au témoin



Alors, le résultat de chaque unité u peut s'écrire

$$Y(u) = DY_1(u) + (1 - D)Y_0(u)$$

Remarque: cette définition suppose que le statut de traitement d'une unité n'affecte pas les résultats potentiels des autres unités.

Définition: pour chaque unité u , le traitement cause l'effet

$$\delta_u = Y_1(u) - Y_0(u)$$

Problème fondamental d'inférence causale:

Pour un u donné u , nous observons soit $Y_1(u)$ OU $Y_0(u)$

Nous ne pouvons pas observer la valeur de $Y_1(u)$ et $Y_0(u)$ sur la même unité u

\Rightarrow *Il est impossible d'observer l'effet de traitement sur u par lui-même*

Problème: nous n'avons pas de preuve contrefactuelle pour u

P.ex. ce qui se serait arrivé à u en l'absence de traitement

- Étant donné que l'effet de traitement pour une seule unité u ne peut pas être observé, nous tentons d'identifier *l'effet moyen de traitement* pour la population U (ou pour des sous-populations).
- *L'effet moyen de traitement (average treatment effect) ATE* sur U (ou sous-populations de U):

$$TE_u = \delta_u = Y_1(u) - Y_0(u)$$



$$\begin{aligned}ATE_U &= E_U [Y_1(u) - Y_0(u)] \\ &= E_U [Y_1(u)] - E_U [Y_0(u)] \\ &= \bar{Y}_1 - \bar{Y}_0 \\ &= \bar{\delta} \quad (1)\end{aligned}$$

- La solution statistique remplace les effets de traitement non-observables de t sur une unité spécifique u avec l'effet *moyen* de traitement de t sur une population U composée de tels u , qui lui peut être observé.
- Bien que $E_U(Y_1)$ et $E_U(Y_0)$ ne peuvent toutes deux être calculées, elles peuvent être estimées.
- La plupart des méthodes économétriques tentent d'établir à partir de données d'observation des estimations cohérentes de

$$E_U(Y_1) = \bar{Y}_1 \quad \text{and} \quad E_U(Y_0) = \bar{Y}_0$$

Donc nous essayons d'estimer:

$$\begin{aligned}ATE_U &= E_U [Y_1(u)] - E_U [Y_0(u)] \\ &= \bar{Y}_1 - \bar{Y}_0\end{aligned}\tag{1}$$

Considérer l'estimateur simple suivant d' ATE_U :

$$\hat{\delta} = [\hat{Y}_1 | D = 1] - [\hat{Y}_0 | D = 0]\tag{2}$$

- équation (1) est définie pour la population totale,
- équation (2) est un estimateur devant être évalué sur un échantillon tiré de cette population.

Lemme: si nous supposons que:

$$[\bar{Y}_1 | D = 1] = [\bar{Y}_1 | D = 0]$$

$$\text{est } [\bar{Y}_0 | D = 1] = [\bar{Y}_0 | D = 0]$$

Alors

$$\hat{\delta} = [\hat{Y}_1 | D = 1] - [\hat{Y}_0 | D = 0]$$

est un estimateur cohérent de

$$\bar{\delta} = \bar{Y}_1 - \bar{Y}_0$$

- Ainsi, une condition suffisante de l'estimateur simple pour estimer de manière cohérente le véritable effet moyen de traitement ATE est que:

$$[\bar{Y}_1 | D = 1] = [\bar{Y}_1 | D = 0]$$

et

$$[\bar{Y}_0 | D = 1] = [\bar{Y}_0 | D = 0]$$

- Le résultat moyen sous traitement \bar{Y}_1 est le même pour les groupes de traitement ($D=1$) et témoin ($D=0$)
- Le résultat moyen sous contrôle \bar{Y}_0 est le même pour les groupes de traitement ($D=1$) et témoin ($D=0$)

Quand ces conditions seront-elles satisfaites?

- Il est suffisant que l'assignation au traitement D soit non-corrélée avec les distributions du résultat potentiel de Y_0 et de Y_1 .
 - Intuitivement: il ne peut y avoir aucune corrélation entre
 - Si oui ou non quelqu'un bénéficie du traitement
 - Dans quelle proportion cette personne bénéficie potentiellement de ce traitement

- La manière la plus facile d'obtenir cette absence de corrélation est de pratiquer l'assignation aléatoire de traitement.

Une autre manière de voir

- Après un peu d'algèbre, on peut montrer que:

$$\underbrace{\hat{\delta}}_{\text{Estimateur simple}} = \underbrace{\bar{\delta}}_{\text{véritable impact}} + \underbrace{\left([\bar{Y}_0 \mid D = 1] - [\bar{Y}_0 \mid D = 0] \right)}_{\text{différence au départ}} + (1 - \pi) \underbrace{\left(\bar{\delta}_{\{D=1\}} - \bar{\delta}_{\{D=0\}} \right)}_{\text{hétérogénéité de traitement}}$$

Une autre manière de voir(avec des mots)

- Il existe deux sources de biais qu'il est nécessaire d'éliminer des estimations des effets causaux tirés d'études observationnelles.
 - Différence au départ. (biais de sélection)
 - Hétérogénéité de traitement.

- La plupart des méthodes disponibles traite seulement du biais de sélection

Traitement sur les unités traitées

- ATE n'est pas toujours le paramètre d'intérêt.
- Souvent, c'est un effet moyen de traitement pour les *unités traitées* qui est d'un intérêt substantiel:

$$\begin{aligned}TOT &= E [Y_1(u) - Y_0(u) \mid D = 1] \\ &= E [Y_1(u) \mid D = 1] - E [Y_0(u) \mid D = 1]\end{aligned}$$

Traitement sur les unités traitées

Si nous devons estimer TOT

$$TOT = E [Y_1(u) | D = 1] - E [Y_0(u) | D = 1]$$

Alors l'estimateur simple (2)

$$\hat{\delta} = [\hat{Y}_1 | D = 1] - [\hat{Y}_0 | D = 0]$$

Estime TOT de manière cohérente si:

$$[\bar{Y}_0 | D = 1] = [\bar{Y}_0 | D = 0]$$

“aucune différence de départ entre les groupes d'unités traitées et témoins”

Références

- ❑ Judea Pearl (2000): Causality: Models, Reasoning and Inference, Cambridge University press. (Book) Chapters 1, 5 and 7.
- ❑ Trygve Haavelmo (1944): “The probability approach in econometrics,” *Econometrica* 12, pp. iii-vi+1-115.
- ❑ Arthur Goldberger (1972): “Structural Equations Methods in the Social Sciences,” *Econometrica* 40, pp. 979-1002.
- ❑ Robins JM, Greenland S. The role of model selection in causal inference from nonexperimental data. *American Journal of Epidemiology* 1986; 123: 392-402.
- ❑ Donald B. Rubin (1974): “Estimating causal effects of treatments in randomized and nonrandomized experiments,” *Journal of Educational Psychology* 66, pp. 688-701.
- ❑ Paul W. Holland (1986): “Statistics and Causal Inference,” *Journal of the American Statistical Association* 81, pp. 945-70, with discussion.