

Méthodes quantitatives d'analyse (POL 2809)

Séance 9, 6 novembre 2019

Enseignante: Florence Vallée-Dubois

Bureau: C-3114

Dispos: mercredis, 10h-11h30

florence.vallee-dubois@umontreal.ca

Aujourd'hui

GOA

On entame le Devoir 4.

Inférence causale

Une relation entre X et Y.

Direction (temporalité): la cause arrive avant l'effet.

La relation n'est pas le résultat d'un facteur commun.

D'abord et avant tout: des principes théoriques.

Graphiques orientés acycliques (GOA)

Graphiques orientés acycliques (GOA)

Manière d'illustrer des relations
causales (des théories).

Graphiques orientés acycliques (GOA)

Manière d'illustrer des relations causales (des théories).

Ensuite, on peut analyser un GOA pour répondre à la question: Est-il possible d'estimer l'effet causal?

Graphiques orientés acycliques (GOA)

Permet:

Graphiques orientés acycliques (GOA)

Permet:

D'illustrer une relation causale.

Graphiques orientés acycliques (GOA)

Permet:

D'illustrer une relation causale.

De vérifier s'il est possible d'identifier la relation causale.

Graphiques orientés acycliques (GOA)

Permet:

D'illustrer une relation causale.

De vérifier s'il est possible d'identifier la relation causale.

De savoir quelles variables inclure ou exclure d'une analyse statistique.

Un GOA est: **orienté**

Un GOA est: **orienté**

Noeuds (descendants ou ancêtres):
variables

Un GOA est: **orienté**

Noeuds (descendants ou ancêtres):
variables

Arcs: flèches

Un GOA est: **orienté**

Noeuds (descendants ou ancêtres):
variables

Arcs: flèches

Chemins: séquence d'arcs (suite de
flèches)

Un GOA est: **orienté**

Noeuds (descendants ou ancêtres):
variables

Arcs: flèches

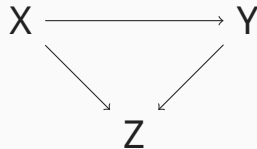
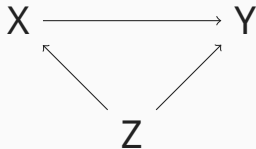
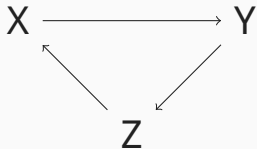
Chemins: séquence d'arcs (suite de
flèches)

La cause est toujours l'ancêtre d'un
effet.

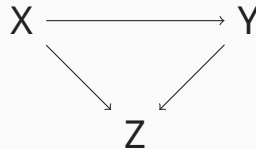
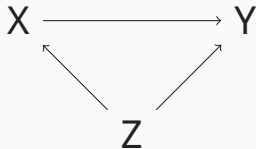
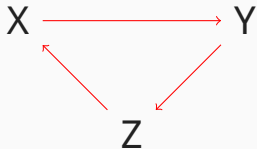
Un GOA est: **acyclique**

Il ne peut y avoir aucun circuit: les arcs ne doivent pas pouvoir nous faire revenir sur nos pas.

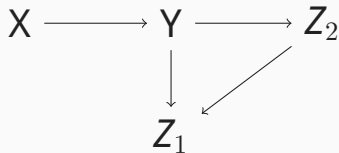
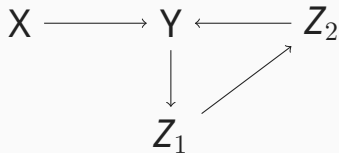
Un GOA est: `acyclic`



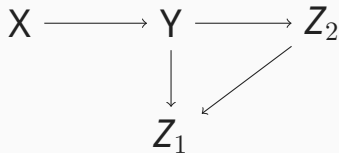
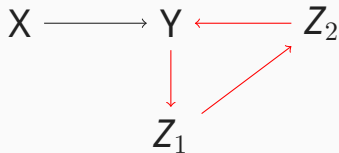
Un GOA est: `acyclic`



Un GOA est: acyclique



Un GOA est: acyclique



Information statistique

La relation causale est unidirectionnelle, mais l'information statistique peut circuler dans les deux sens.

Information statistique

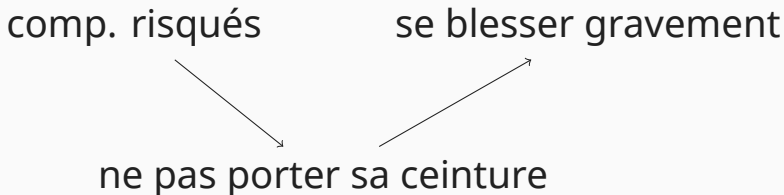
La relation causale est unidirectionnelle, mais l'information statistique peut circuler dans les deux sens.

Exemple:

Information statistique

La relation causale est unidirectionnelle, mais l'information statistique peut circuler dans les deux sens.

Exemple:



Chemins ouverts et chemins fermés

Quand le chemin laisse circuler l'info
stat.: il est ouvert

Chemins ouverts et chemins fermés

Quand le chemin laisse circuler l'info stat.: il est ouvert

Quand le chemin ne laisse pas circuler l'info stat.: il est fermé

Chemins ouverts et chemins fermés

Quand le chemin laisse circuler l'info stat.: il est ouvert

Quand le chemin ne laisse pas circuler l'info stat.: il est fermé

Pensez aux arcs comme à des tuyaux par lesquels circulent l'information.

Types de chemins

Chaînes:

$X \longrightarrow Z \longrightarrow Y$

Types de chemins

Chaînes:

$X \longrightarrow Z \longrightarrow Y$

Les chaînes sont des chemins ouverts.

Connaître l'information centrale (Z)
ferme le chemin.

Exemple de chaîne

Feu —————> Fumée —————> Alarme

Exemple de chaîne

Feu —————> Fumée —————> Alarme

Contrôler pour "fumée" ferme le chemin.

Exemple de chaîne

Feu —→ **Fumée** —→ **Alarme**

Contrôler pour "fumée" ferme le chemin.

Si on prend 2 cas où il y a de la fumée,
l'information sur le feu ne nous aidera pas à savoir
si l'alarme s'est enclenchée. Et vice-versa.

Types de chemins

Fourchettes:



Types de chemins

Fourchettes:



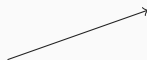
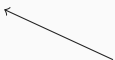
Les fourchettes sont des chemins ouverts.

Connaître l'information centrale (Z) ferme le chemin.

Exemple de fourchette

Pointure

Habiletés en lecture

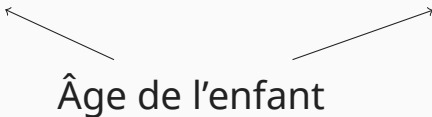


Âge de l'enfant

Exemple de fourchette

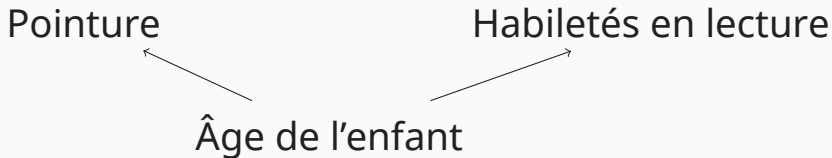
Pointure

Habiletés en lecture



Contrôler pour "l'âge de l'enfant" ferme le chemin.

Exemple de fourchette



Contrôler pour "l'âge de l'enfant" ferme le chemin.

Si on prend 2 enfants du même âge, l'information sur leur pointure de souliers ne nous aidera pas à prédire leurs habiletés en lectures.

Types de chemins

Collision:



Types de chemins

Collision:



Les collisions sont des chemins fermés.

Connaître l'information centrale (Z)
ouvre le chemin.

Exemple de collision

Talent $\longrightarrow$ Célébrité $\longleftarrow$ Beauté

Exemple de collision

Talent $\longrightarrow$ Célébrité $\longleftarrow$ Beauté

Contrôler pour "la célébrité" ouvre le chemin.

Exemple de collision

Talent $\longrightarrow$ **Célébrité** $\longleftarrow$ **Beauté**

Contrôler pour "la célébrité" ouvre le chemin.

Si on prend 2 acteurs tout aussi célèbres, avoir de l'information sur leur talent nous permet de prédire leur beauté.



À quoi ça sert?

C'est la pause!

Utilité des GOA

Vérifier s'il est possible d'identifier la relation causale.

Savoir quelles variables inclure ou exclure d'une analyse statistique.

Chemins ouverts et chemins fermés

Les chaînes et les fourchettes sont ouvertes.

Chemins ouverts et chemins fermés

Les chaînes et les fourchettes sont ouvertes.

Les collisions sont fermées.

Chemins ouverts et chemins fermés

Les chaînes et les fourchettes sont ouvertes.

Les collisions sont fermées.

Un chemin composé uniquement de chaînes et de fourchettes est **ouvert**.

Chemins ouverts et chemins fermés

Les chaînes et les fourchettes sont ouvertes.

Les collisions sont fermées.

Un chemin composé uniquement de chaînes et de fourchettes est **ouvert**.

Dès qu'il y a au moins une collision, le chemin est **fermé**.

Chemins ouverts et chemins fermés: exemples

$$X \longrightarrow Z_1 \longrightarrow Z_2 \longrightarrow Z_3 \longrightarrow Y$$

Chemins ouverts et chemins fermés: exemples

Ouvert.

$$X \longrightarrow Z_1 \longrightarrow Z_2 \longrightarrow Z_3 \longrightarrow Y$$

Chemins ouverts et chemins fermés: exemples

$$X \longrightarrow Z_1 \longleftarrow Z_2 \longrightarrow Z_3 \longrightarrow Y$$

Chemins ouverts et chemins fermés: exemples

Fermé.

$$X \xrightarrow{\text{red}} Z_1 \xleftarrow{\text{red}} Z_2 \longrightarrow Z_3 \longrightarrow Y$$

Chemins ouverts et chemins fermés: exemples

$$X \longleftarrow Z_1 \longleftarrow Z_2 \longrightarrow Z_3 \longrightarrow Y$$

Chemins ouverts et chemins fermés: exemples

Ouvert.

$$X \xleftarrow{\quad} Z_1 \xleftarrow{\quad} Z_2 \xrightarrow{\quad} Z_3 \xrightarrow{\quad} Y$$

Chemins ouverts et chemins fermés: exemples

$$X \longrightarrow Z_1 \longleftarrow Z_2 \longrightarrow Z_3 \longleftarrow Y$$

Chemins ouverts et chemins fermés: exemples

Fermé.

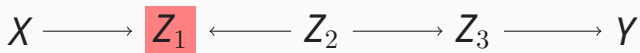
$$X \longrightarrow Z_1 \longleftarrow Z_2 \longrightarrow Z_3 \longleftarrow Y$$

Chemins ouverts et chemins fermés

Contrôler pour le maillon central d'une collision ouvrira le chemin.

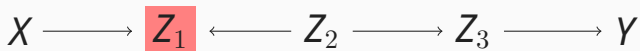
Chemins ouverts et chemins fermés

Contrôler pour le maillon central d'une collision ouvrira le chemin.



Chemins ouverts et chemins fermés

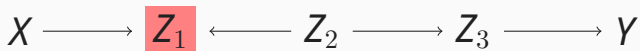
Contrôler pour le maillon central d'une collision ouvrira le chemin.



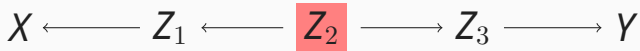
Contrôler pour au moins un maillon d'un chemin ouvert fermera le chemin.

Chemins ouverts et chemins fermés

Contrôler pour le maillon central d'une collision ouvrira le chemin.



Contrôler pour au moins un maillon d'un chemin ouvert fermera le chemin.



Pour identifier une relation causale, il y a deux étapes

Quand il y a de nombreuses variables qui sont liées à X et/ou Y, identifier l'effet causal peut être difficile.

Pour identifier une relation causale, il y a deux étapes

Quand il y a de nombreuses variables qui sont liées à X et/ou Y, identifier l'effet causal peut être difficile.

Les GOA nous offrent une marche à suivre pour savoir pour quelles variables contrôler, ou pas.

Pour identifier une relation causale,
il y a deux étapes

1. NE PAS contrôler pour un descendant de X.

Pour identifier une relation causale, il y a deux étapes

1. NE PAS contrôler pour un descendant de X.
2. Fermer les "chemins par la porte arrière".

1. Ne pas contrôler pour un descendant de X

Contrôler pour un descendant de X ne permet pas d'identifier l'effet causal total de X sur Y .

1. Ne pas contrôler pour un descendant de X

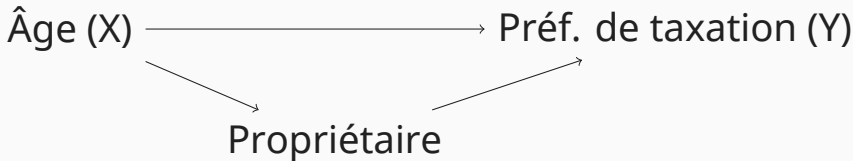
Contrôler pour un descendant de X ne permet pas d'identifier l'effet causal total de X sur Y .

Exemple.

1. Ne pas contrôler pour un descendant de X

Contrôler pour un descendant de X ne permet pas d'identifier l'effet causal total de X sur Y.

Exemple.



2. Fermer les "chemins par la porte arrière"

Un chemin par la porte arrière est un chemin qui lie X et Y , avec une flèche qui pointe vers X .

2. Fermer les "chemins par la porte arrière"

Un chemin par la porte arrière est un chemin qui lie X et Y , avec une flèche qui pointe vers X .

S'il est ouvert, la relation entre X et Y est biaisée. Il faut s'assurer que ces chemins soient fermés.

2. Fermer les "chemins par la porte arrière"

Un chemin par la porte arrière est un chemin qui lie X et Y , avec une flèche qui pointe vers X .

S'il est ouvert, la relation entre X et Y est biaisée. Il faut s'assurer que ces chemins soient fermés.

Exemple.

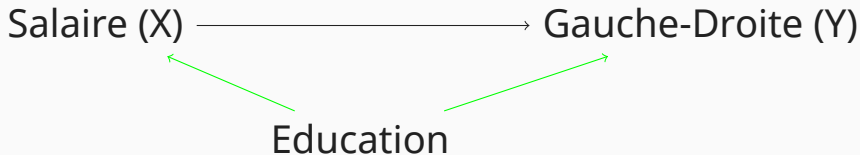
2. Fermer les "chemins par la porte arrière"

Exemple.

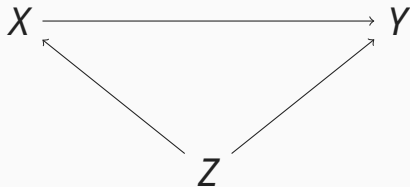


2. Fermer les "chemins par la porte arrière"

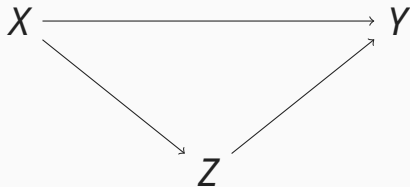
Exemple.



Exercise 1



Exercise 2



Exercices complexes en classe.

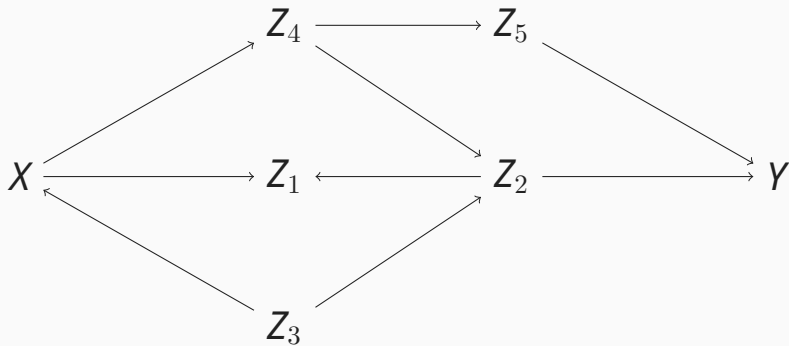
Résumé: Marche à suivre pour identifier l'effet causal

1. Faites la liste des variables pour lesquelles il faut éviter de contrôler.
 - Tous les descendants de X (enfants et petits-enfants).
2. Faites la liste de tous les chemins par la porte arrière.
 - Tous les chemins qui vont de X à Y et qui ont une flèche qui pointe vers X.
3. Est-ce qu'il y a des chemins par la porte arrière qui sont ouverts?
 - Dès qu'il y a une collision, le chemin est fermé.

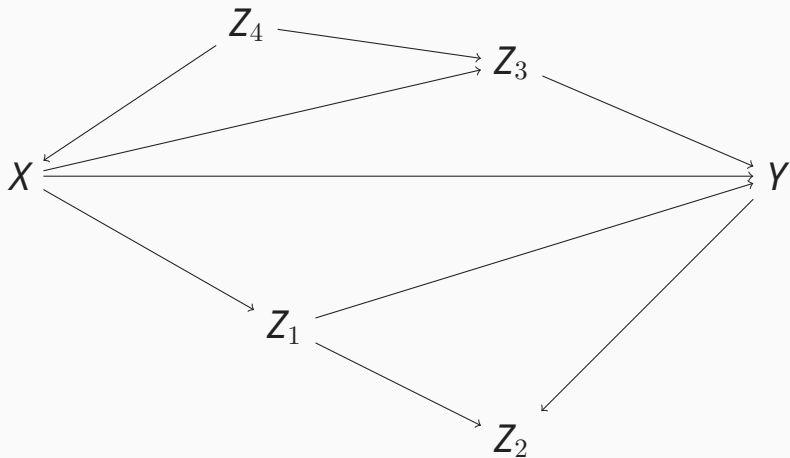
Résumé: Marche à suivre (suite)

4. Si oui, est-il possible de les bloquer en contrôlant pour des variables qui n'ont pas été identifiées à la question 1?
 - On peut fermer un chemin en contrôlant pour au moins une variable par chemin.
 - Si on contrôle pour une variable au centre d'une collision, cela ouvrira le chemin.
 - Il ne faut pas contrôler pour un descendant.
 - Dites pour quelles variables on peut contrôler.
5. Est-ce que l'effet causal de X sur Y est identifiable?
 - Si on a réussi à fermer les chemins par la porte arrière ouverts = oui.
 - Sinon = non.

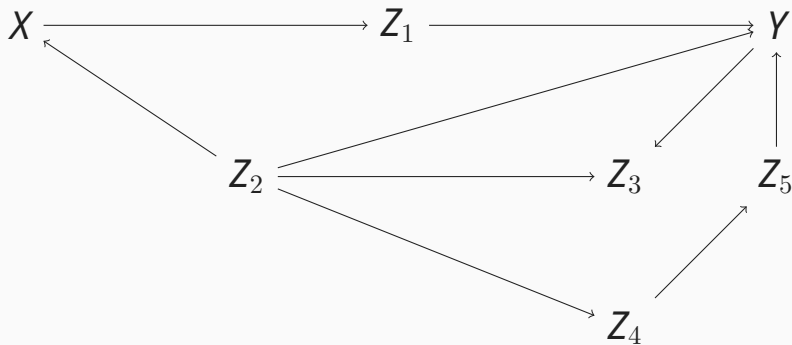
Exercices en classe 3 (les réponses seront données en classe)



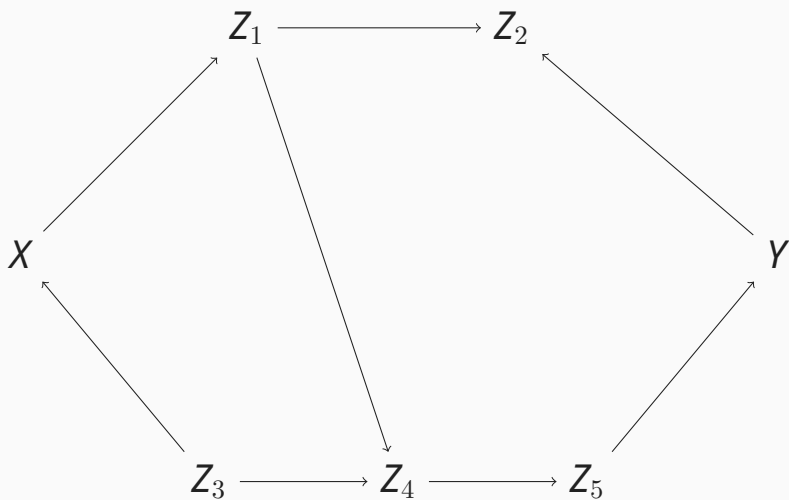
Exercices en classe 4 (les réponses seront données en classe)



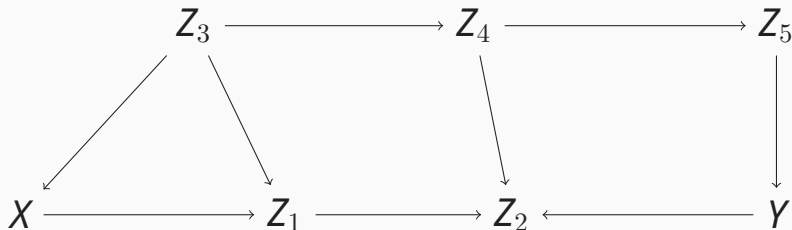
Exercices en classe 5 (les réponses seront données en classe)



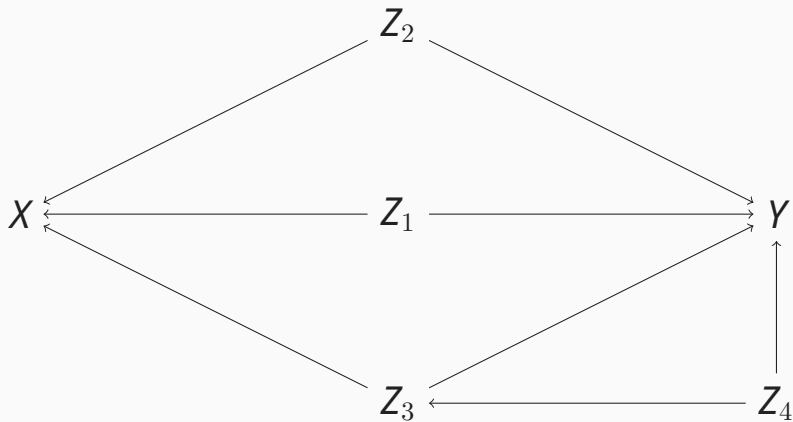
Exercices en classe 6 (les réponses seront données en classe)



Exercices en classe 7 (les réponses seront données en classe)



Exercices en classe 8 (les réponses seront données en classe)





Rappelez-vous

Rappelez-vous

Les expériences aléatoires sont le "gold standard" de la recherche de la causalité.

Rappelez-vous

Les expériences aléatoires sont le "gold standard" de la recherche de la causalité.

Les GOA ne remplacent pas les expériences aléatoires.

Rappelez-vous

Les expériences aléatoires sont le "gold standard" de la recherche de la causalité.

Les GOA ne remplacent pas les expériences aléatoires.

Ils nous aident à déterminer si l'effet causal est identifiable.

Rappelez-vous

Les expériences aléatoires sont le "gold standard" de la recherche de la causalité.

Les GOA ne remplacent pas les expériences aléatoires.

Ils nous aident à déterminer si l'effet causal est identifiable.

Mais même si on pense avoir contrôlé pour les bonnes variables dans une régression, on peut toujours en avoir oublié d'autres.

Prochain cours

Les biais.

**À la semaine
prochaine!**

