

Séance 2: Introduction à l'Économétrie

Sandra Nevoux

Sciences Po

Judi 10 Septembre 2015

L'essentiel à retenir

- 1 Régression
- 2 Moindres carrés ordinaires
- 3 Corrélation vs. Causalité
- 4 Causalité au sens de Granger
- 5 Différence-en-Différences
 - Expériences naturelles
 - Expériences aléatoires

Mesurer un impact

- Procédure:
 - 1e étape: Théorie économique.
 - 2e étape: Collecte des données relatives à la théorie économique qu'on veut tester.
 - 3e étape: Méthode(s) empirique(s) appliquée(s) à ces données en vue de tester la théorie.
- Régression de Y sur X : formalisation d'une relation mathématique et graphique entre X et Y :

$$Y = f(X)$$

- On considère dans le cours exclusivement la régression linéaire, c'est-à-dire une approximation de la relation entre X et Y par une fonction affine:

$$Y = a + bX$$

- b mesure la corrélation entre X et Y .
- Causalité:
 - Régression "satisfaisante" $\Rightarrow b$ mesure l'effet causal.
 - Régression "insatisfaisante" \Rightarrow Méthodes empiriques pour mesurer le "véritable" effet causal.

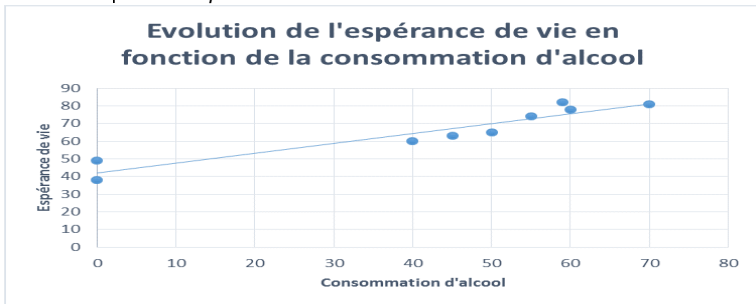
Régression - Définition

- **Définition:** Représentation graphique et mathématique simplifiée de la relation entre, d'une part, une variable Y , et d'autre part, une ou plusieurs variable(s) X . Il s'agit de la courbe et de la fonction approchant au mieux les observations de X et Y .
- Dimension de la régression:
 - Régression simple: régression de la variable Y sur une seule variable X .
 - Régression multiple: régression de la variable Y sur plusieurs variables X .
- Type de régression:
 - Régression linéaire: relation entre la variable Y et la (ou les) variable(s) X représentée graphiquement par une droite et mathématiquement par une fonction affine.
 - Régression exponentielle, logarithmique,...

⇒ Régression linéaire simple: "meilleures" droite et fonction affine qui relient les observations de X aux observations ajustées de Y .
- Terminologie des variables:
 - La variable Y est **endogène** = déterminée par une (ou plusieurs) autre(s) variable(s).
 - La (ou les) variable(s) X est (sont) **exogène(s)** = déterminée(s) indépendamment d'autres variables.

Régression - Application

- Données \Rightarrow Nuage de points dans le plan (*Conso*, *Espérance*) \Rightarrow Régression linéaire simple de l'*Espérance* sur la *Conso*:



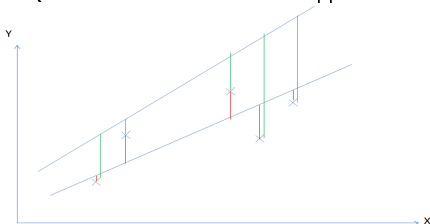
- Droite passant au plus près des observations relatives à la *Conso* et à l'*Espérance*.
- Equation linéaire: fonction affine correspondant à cette droite:
$$\text{Espérance} = a + b * \text{Conso}$$

Moindres carrés ordinaires

Définition

Comment déterminer la droite qui passe au plus près des observations de X et Y ?
⇒ Méthode des moindres carrés ordinaires (MMCO).

- **Définition:** La MMCO consiste à rechercher la droite qui minimise la somme des carrés des écarts verticaux entre les observations de la variable Y et les observations ajustées de la variable Y .
 - Notation:
 - Les observations réelles de la variable Y sont notées Y .
 - Les observations ajustées de la variable Y sont notées \hat{Y} .
- ⇒ Quelle droite est la meilleure approximation des observations? Pourquoi?

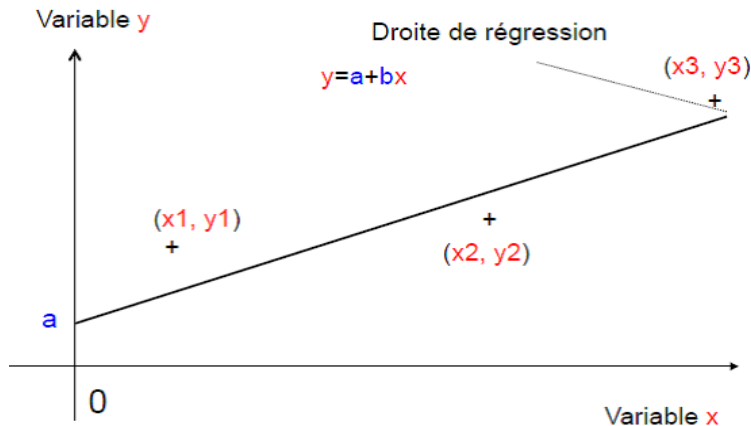


Équation linéaire

$$Y = a + bX$$

- L'équation linéaire est composée:
 - Des variables X et Y , qui sont connues.
 - Des paramètres a et b qui sont inconnus et que l'on cherche à calculer afin d'estimer une corrélation, voire une causalité, entre X et Y .
- Interprétation des paramètres a et b de la droite:
 - b :
 - **Définition:** coefficient directeur (ou pente) de la droite.
 - **Interprétation:** une augmentation d'une unité de X est associée à une variation de b unités de Y .
 - a :
 - **Définition:** ordonnée à l'origine de la droite.
 - **Interprétation:** Quand la variable X est nulle, l'agent dispose de a unités de Y .

Représentation graphique et mathématique



Calcul des paramètres a et b

$$b = \frac{(x_1 - \bar{x})(y_1 - \bar{y}) + (x_2 - \bar{x})(y_2 - \bar{y}) + (x_3 - \bar{x})(y_3 - \bar{y}) + \dots}{(x_1 - \bar{x})^2 + (x_2 - \bar{x})^2 + (x_3 - \bar{x})^2 + \dots}$$

$$a = \bar{y} - b\bar{x} \Leftrightarrow \bar{y} = a + b\bar{x}$$

- \bar{x} est la moyenne de X et \bar{y} est la moyenne de Y .
- La droite de régression passe par (\bar{x}, \bar{y}) .

Calcul des paramètres a et b - Application

⇒ Quelle est la valeur de b ? Comment l'interpréter?

⇒ Quelle est la valeur de a ? Comment l'interpréter?

Elève	1	2	3	4	5	6	7	8	9
<i>pointure</i>	38	35	37	32	30	33	38	31	35
<i>Nombre de fautes lors d'une dictée</i>	5	6	7	10	12	8	5	12	8

Modèle vs. Réalité

Modèle	Réalité
$\hat{Y} = a + bX$	$Y = a + bX + \epsilon$
\hat{Y} est la valeur de Y prédite par le modèle.	Y est la valeur réelle de Y . ϵ est le résidu, c'est-à-dire l'écart (vertical) entre la valeur prédite \hat{Y} et la valeur réelle Y .

Qualité de l'ajustement linéaire - Coefficient de détermination R^2

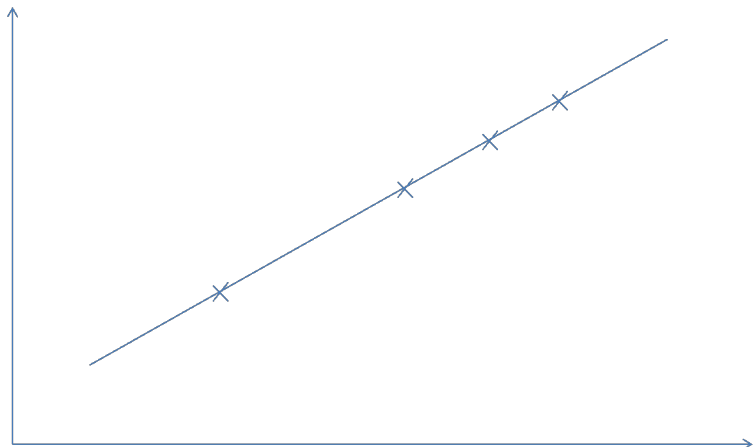
$\epsilon \Rightarrow$ Coefficient de détermination R^2 .

- **Définition:** R^2 est le coefficient de détermination, c'est-à-dire la part de la variabilité de Y expliquée par les variables du modèle (ici, X) et se définit comme:

$$R^2 = \frac{(\hat{Y}_1 - \bar{Y})^2 + (\hat{Y}_2 - \bar{Y})^2 + (\hat{Y}_3 - \bar{Y})^2 + \dots}{(Y_1 - \bar{Y})^2 + (Y_2 - \bar{Y})^2 + (Y_3 - \bar{Y})^2 + \dots}$$

- R^2 est compris entre 0 et 1.
- plus R^2 est proche de 1, plus la part de la variabilité de Y expliquée par les variables du modèle est importante (plus Y est proche de \hat{Y}), et plus R^2 révèle une bonne qualité d'ajustement linéaire et un pouvoir explicatif important.
- plus R^2 est proche de 0, plus la part de la variabilité de Y expliquée par les variables du modèle est faible (plus Y est éloignée de \hat{Y}), et plus R^2 révèle une mauvaise qualité d'ajustement linéaire et un faible pouvoir explicatif.

R^2 - Représentation graphique



Précision des paramètres a et b - Écart-type et intervalle de confiance

Économétrie \Rightarrow Les paramètres a et b sont eux aussi des variables aléatoires et la MMCO va seulement permettre d'approcher ces coefficients avec une erreur \Rightarrow Écart-type.

- Lorsque l'on reporte les paramètres \hat{a} et \hat{b} estimés d'une régression, on met en général en dessous du paramètre l'écart-type entre parenthèses.
- **Définition:** l'écart-type d'un paramètre est la racine carrée de la variance de ce paramètre.
- **Interprétation:** l'écart-type, dénoté e_a (respectivement e_b), constitue un indicateur de la précision du paramètre et permet de définir l'intervalle de confiance de ce paramètre.

\Rightarrow Le paramètre réel a a 95% de probabilité d'appartenir à l'intervalle de confiance $[\hat{a} - 2e_a; \hat{a} + 2e_a]$.

\Rightarrow Le paramètre réel b a 95% de probabilité d'appartenir à l'intervalle de confiance $[\hat{b} - 2e_b; \hat{b} + 2e_b]$.

Calcul de l'intervalle de confiance - Application

⇒ Quel est l'intervalle de confiance à 95% de a ?

⇒ Quel est l'intervalle de confiance à 95% de b ?

$$\text{Stress} = 0,28 + 0,02 \text{ Chômage}$$

(0,03) (0,005)

Interprétation des paramètres a et b

Paramètres a et b pas toujours exprimés en unités \Rightarrow Interprétation différente de ces paramètres.

Équation linéaire	Interprétation
$\hat{Y} = a + bX$	Une hausse d'une unité de X est associée à une variation de b unités de \hat{Y} .
$\log(\hat{Y}) = a + b\log(X)$	Une hausse de 1% de X est associée à une variation de b % de \hat{Y} .
$\log(\hat{Y}) = a + bX$	Une hausse d'une unité de X est associée à une variation de $100 * b$ % de \hat{Y} .
$\hat{Y} = a + b\log(X)$	Une hausse de 1% de X est associée à une variation de $\frac{b}{100}$ unités de \hat{Y} .

Résumé

- Régression linéaire (simple et multiple): définition \Rightarrow représentation graphique et mathématique:

$$Y = a + bX + e$$

$$\hat{Y} = \underset{(\alpha)}{a} + \underset{(\beta)}{b} X$$

$$R^2 = \gamma$$

- Méthode des moindres carrés ordinaires (MMCO): définition \Rightarrow représentation graphique.
- Paramètres a et b : calcul, définition et interprétation (en fonction des unités).
- Coefficient de détermination R^2 : définition et interprétation.
- Écarts-types α et β : interprétation et définition de l'intervalle de confiance.

Résumé - Application

Regressions of Log Wine Price on Climate Variables

Independent variables

Age of vintage	0.2400	(0.0075)
Average temperature over growing season (April–September)	0.6080	(0.1160)
Rain in August	-0.00380	(0.00095)
Rain in the months preceding the vintage (October–March)	0.00115	(0.00051)
Average temperature in September	0.0077	(0.0565)
R-squared	0.828	
Root mean squared error	0.293	

Causalité vs. Corrélation

Causalité

- **Définition:** mécanisme théorique particulier, identifié par la théorie économique et dont on souhaite tester s'il se vérifie empiriquement. Lorsqu'on tente d'identifier un effet causal précis, on teste dans quelle mesure X cause Y lorsqu'on sait que, d'après la théorie économique, X doit causer Y .
- Dynamique:
 - Causalité positive: une hausse (respectivement baisse) de X est identifiée comme entraînant une hausse (respectivement baisse) de Y .
 - Causalité négative: une hausse (respectivement baisse) de X est identifiée comme entraînant une baisse (respectivement hausse) de Y .
 - Causalité nulle: une variation de X est identifiée comme n'ayant aucun effet sur la variation de Y .
- Exemples:
 - Travail et effort \Rightarrow Admission à Sciences Po.
 - Bonne hygiène de vie \Rightarrow Bonne santé.

Corrélation

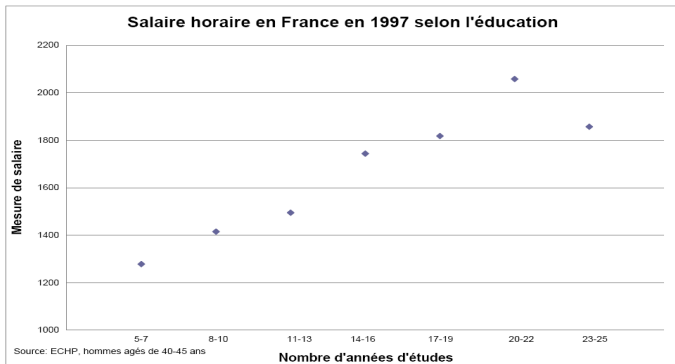
- **Définition:** accumulation des causalités existant potentiellement entre X et Y et qui se traduit par un lien statistique. La corrélation est donc la part de la variation de Y expliquée par la variation de X .
⇒ La corrélation ne témoigne en rien d'un effet causal de X sur Y .
- Dynamique:
 - Corrélation positive: une hausse (respectivement baisse) de X est associée à une hausse (respectivement baisse) de Y .
 - Corrélation négative: une hausse (respectivement baisse) de X est associée à une baisse (respectivement hausse) de Y .
 - Corrélation nulle: une variation de X n'est associée à aucune variation de Y .
- Exemples:
 - Parapluie \Leftrightarrow Pluie.
 - Prix du tabac \Leftrightarrow Consommation de tabac.
 - Etudes \Leftrightarrow Salaire.

Corrélation vs. Causalité

- Causalité implique nécessairement corrélation.
- Corrélation n'implique pas nécessairement causalité:
 - Erreur(s) de mesure de X et/ou Y .
 - Facteurs manquants: variables Z , P , D qui ne sont pas incluses dans la régression et qui pourtant ont un effet causal sur Y . L'omission de ces facteurs explicatifs supplémentaires peut renforcer la corrélation entre X et Y , sans toutefois qu'il y ait un lien de causalité entre X et Y .
 - Simultanéité:
 - Causalité double: X cause Y et Y cause X (causalité simultanée (X cause Y et Y cause X en même temps); causalité réciproque (X cause Y qui à son tour cause X qui à son tour cause Y et ainsi de suite)).
 - Causalités multiples: X cause Y , Z cause Y , Z cause X , Y cause X .
 - Causalité inverse: Y cause X .

Corrélation vs. Causalité - Application (1)

- ⇒ Quelle corrélation observe-t-on sur ce graphique ?
- ⇒ Comment pourrait-on interpréter cette relation ?
- ⇒ l'accès facilité aux études supérieures permet-il d'améliorer la situation économique des travailleurs ?



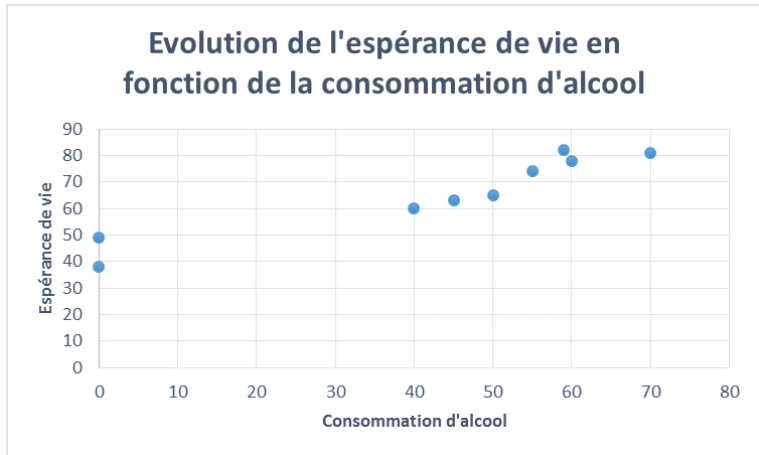
Corrélation vs. Causalité - Application (2)

⇒ Existe-t-il une corrélation entre l'espérance de vie et la consommation d'alcool ? Si oui, est-elle positive ou négative? (on pourra s'aider d'un graphique)

⇒ Existe-t-il une causalité? Si oui, dans quel sens?

Pays	France	USA	Nigéria	Ethiopie	A	B	C	D	E
<i>Espérance de vie</i>	82	78	49	38	60	65	81	63	74
<i>Consommation d'alcool par habitant et par an (unité arbitraire)</i>	59	60	0	0	40	50	70	45	55

Corrélation vs. Causalité - Application (2)



Corrélation vs. Causalité - Application (3)

⇒ Existe-t-il une corrélation entre la pointure d'un élève et le nombre de fautes réalisées lors d'une dictée? Si oui, est-elle positive ou négative? (On pourra s'aider d'un graphique)

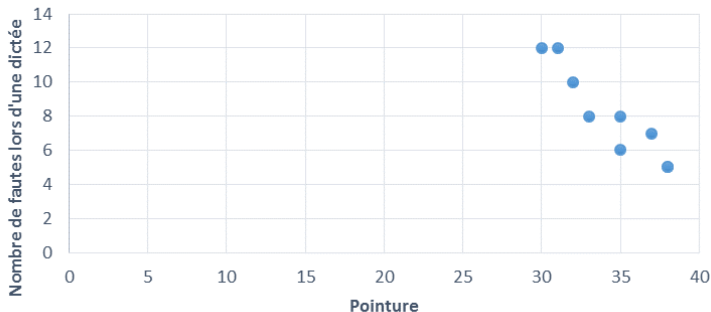
⇒ Qu'est ce qu'un point aberrant?

⇒ Existe-t-il une causalité? Si oui, dans quel sens?

Elève	1	2	3	4	5	6	7	8	9
<i>pointure</i>	38	35	37	32	30	33	38	31	35
<i>Nombre de fautes lors d'une dictée</i>	5	6	7	10	12	8	5	12	8

Corrélation vs. Causalité - Application (3)

Evolution du nombre de fautes lors d'une dictée en fonction de la pointure



Corrélation vs. Causalité - Application (4)

- Dans le guide de l'étudiant de Freakonomics, les auteurs observent une très forte corrélation entre la vente de glaces et le nombre de noyades.

⇒ Peut-on en déduire une causalité? Pourquoi?

- Dans *Le Monde* du 23 septembre 2007, "*l'usage régulier du cannabis augmenterait de 41% les risques de développer ultérieurement une affection psychiatrique de nature psychotique*". En outre, ce risque augmenterait avec l'intensité et la durée de la consommation.

⇒ Les auteurs de ce travail se sont refusés à établir une relation directe de causalité entre la consommation de cannabis et les affections psychotiques. Pourquoi?

Identification

Identification de l'effet causal de X sur $Y \Rightarrow$ Plusieurs méthodes:

- Causalité au sens de Granger.
- Différence-en-Différences:
 - Expériences naturelles.
 - Expériences aléatoires.

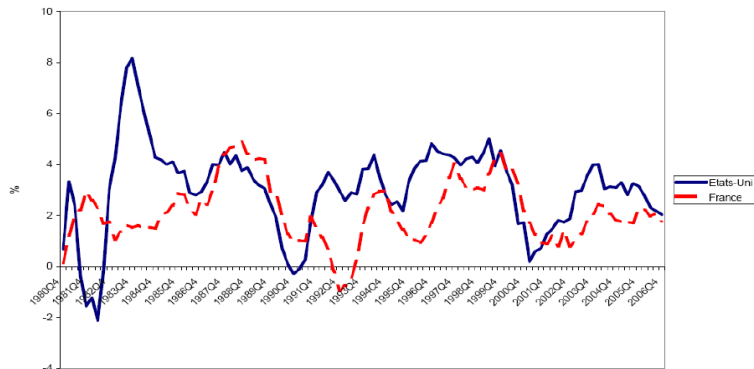
Causalité au sens de Granger - Définition

- **Définition:** X cause Y au sens de Granger si les valeurs passées de X ont un impact statistique sur la valeur actuelle ou future de Y .
- Dynamique:
 - Causalité positive au sens de Granger: les valeurs passées de X ont un impact statistique positif sur la valeur actuelle ou future de Y .
 - Causalité négative au sens de Granger: les valeurs passées de X ont un impact statistique négatif sur la valeur actuelle ou future de Y .
- Timing et séquentialité des événements.
- Compréhension de la causalité double (simultanée et/ou réciproque).
- Limites:
 - Corrélation vs. Causalité.
 - Anticipations des agents \Rightarrow Fausse causalité.
- Exemples:
 - Parapluie \Leftrightarrow Pluie.
 - Sport: Salaire \Leftrightarrow Victoires.
 - Macroéconomie: Taux de croissance économique \Leftrightarrow Consommation.

Causalité au sens de Granger - Application (1)

- ⇒ Observe-t-on un lien entre le taux de croissance aux États-Unis et en France?
- ⇒ La relation est-elle synchronisée?
- ⇒ Comment pourrait-on interpréter cette relation?

Taux de croissance du PIB en France et aux Etats-Unis



Source: OECD main economic indicators; PIB à prix constants (2000) ajusté des variations

Causalité au sens de Granger - Application (2)

- Football: $\%Victoires_t = 0,327 + 0,191 * Salaire_{t-1}$

- Baseball: $\%Victoires_t = 0,404 + 0,097 * Salaire_{t-1}$

⇒ Observe-t-on un lien entre le salaire et le pourcentage de victoires au football et au baseball?

⇒ La relation est-elle synchronisée?

⇒ Comment pourrait-on interpréter cette relation?

Différence-en-Différences

Contrefactuel

Identification de l'effet causal de X sur $Y \Rightarrow$ Contrefactuel.

- **Définition:** le contrefactuel est la situation généralement non observée qui se serait produite en l'absence d'un phénomène particulier dont on cherche à mesurer l'impact causal. L'écart entre la situation observée, affectée par le phénomène en question, et le contrefactuel, est précisément l'impact causal.
- En comparant la valeur de Y quand X a agi avec la valeur du contrefactuel de Y (c'est-à-dire quand X n'a pas agi), on obtient l'impact causal.
- Théorie vs. Réalité:
 - En théorie, le contrefactuel de Y est le clone de Y : il lui est identique en tout point, sauf au regard de la variable X dont on veut estimer l'effet sur Y .
 - Dans la réalité, il est impossible d'avoir un contrefactuel parfait. On essaye donc de s'en approcher le plus possible.
- Exemples:
 - Jumeaux \Leftrightarrow Rendements de l'éducation.
 - Un individu bénéficiant ou non d'une politique d'aide au retour à l'emploi \Leftrightarrow Retour à l'emploi.

Groupes de traitement et de contrôle

- Le **groupe de traitement** est le groupe qui bénéficie d'une mesure de politique économique dont on souhaite tester l'effet.
- Le **groupe de contrôle** est le groupe aussi proche que possible du groupe de traitement mais qui s'en distingue par le fait qu'il n'a pas bénéficié de la mesure de politique économique dont on souhaite tester l'effet.
- Le groupe de contrôle illustre ce qui se serait passé pour le groupe de traitement en l'absence de la mesure de politique économique. Il sert de contrefactuel.

Différence-en-Différences - Définition

- La méthode de **différence-en-différences** consiste à comparer l'évolution d'un certain comportement du groupe de traitement à l'évolution du même comportement au cours de la même période du groupe de contrôle.

⇒ On conclut sur l'impact de la mesure de politique économique en comparant la différence dans le comportement du groupe de traitement sur la période de référence à la différence du comportement du groupe de contrôle sur la même période: c'est la différence de la différence. L'effet causal est identifié par la DIFFÉRENCE entre le groupe de traitement et le groupe de contrôle.

Différence-en-Différences - Estimation

	Traitement	Contrôle
Avant	$Y_{T,avant}$	$Y_{C,avant}$
Après	$Y_{T,apr}$	$Y_{C,apr}$

$$\Rightarrow \text{Différence de traitement} = Y_{T,apr} - Y_{T,avant}$$

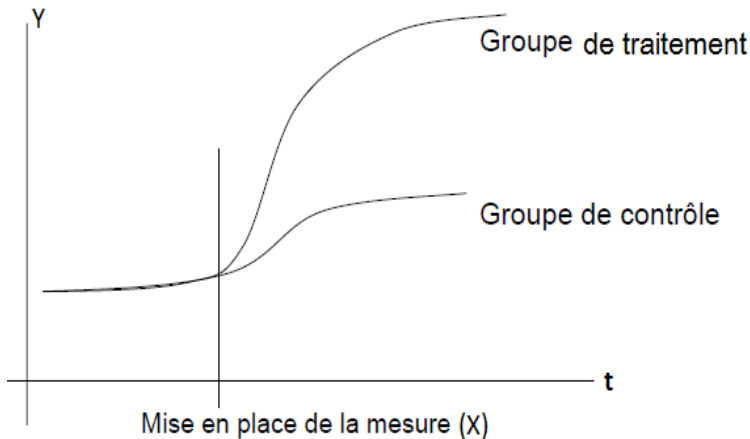
$$\Rightarrow \text{Différence de contrôle} = Y_{C,apr} - Y_{C,avant}$$

$$\begin{aligned} \Rightarrow \text{Effet causal} &= \text{Différence de traitement} - \text{Différence de contrôle} \\ &= (Y_{T,apr} - Y_{T,avant}) - (Y_{C,apr} - Y_{C,avant}) \end{aligned}$$

Différence-en-Différences - Explications

- Effet causal vs. Différence de traitement ($Y_{T,apr} - Y_{T,avant}$):
 - La différence de traitement ($Y_{T,apr} - Y_{T,avant}$) ne peut être interprétée comme un effet causal, car le contexte peut aussi jouer sur l'évolution du groupe de traitement.
 - Le changement de contexte sera pris en compte par la différence de contrôle, qui représente ce qui serait arrivé au groupe de traitement s'il n'avait pas bénéficié de la mesure de politique économique.
- Effet causal vs. Différence entre les groupes de traitement et de contrôle après le traitement ($Y_{T,apr} - Y_{C,apr}$):
 - La différence entre les groupes de traitement et de contrôle après le traitement ($Y_{T,apr} - Y_{C,apr}$) ne peut être interprétée comme un effet causal, car ces deux groupes peuvent avoir ex ante des caractéristiques différentes qui influent sur l'évolution de Y .

Différence-en-Différences - Représentation graphique



Différence-en-Différences - Application

On cherche à estimer l'impact du fait d'avoir le manuel d'Etienne Wasmer sur les résultats de microéconomie des étudiants de 1e année. Je choisis aléatoirement dans la classe un groupe de 10 étudiants auquel j'offre le livre d'Etienne Wasmer.

⇒ Afin d'identifier l'impact causal de cette mesure, quelle(s) hypothèse(s) faut-il faire sur les caractéristiques des étudiants de la classe? Cette hypothèse est-elle vérifiée?

On mesure la moyenne des deux groupes avant et après la mesure:

- Groupe de traitement:
 - Avant la mesure: moyenne de 12/20.
 - Après la mesure: moyenne de 15/20.
- Groupe de contrôle:
 - Avant la mesure: moyenne de 13/20.
 - Après la mesure: moyenne de 15/20.

⇒ Est-ce que la différence entre avant et après pour le groupe de traitement permet d'estimer l'effet causal de la mesure?

⇒ Est-ce que la différence entre les groupes de traitement et de contrôle après permet d'estimer l'effet causal de la mesure?

⇒ Quel est l'effet causal de la mesure?

Différence-en-Différences - Expériences

La méthode de la différence-en-différences est applicable à deux types d'expériences:

Expériences naturelles	Les groupes de traitement et de contrôle sont "naturels".
Expériences aléatoires	Les groupes de traitement et de contrôle sont "construits".

Expériences naturelles - Définition

- **Définition:** Une expérience naturelle (ou méthode quasi-expérimentale) consiste en l'identification d'un groupe de contrôle et d'un groupe de traitement suite à un choc exogène qui produit une variation substantielle dans les données. Le groupe de traitement est celui qui a subi l'événement alors que le groupe de contrôle est un groupe qui a été épargné par l'événement. Le comportement des deux groupes en amont de l'événement est comparé afin de s'assurer de la similarité des deux groupes.
- Dans le cadre d'une expérience naturelle, les groupes de traitement et de contrôle sont naturels.
- Théorie vs. Réalité:
 - En théorie, le choc est exogène en ce sens qu'il n'est pas anticipé et est potentiellement subi par les deux groupes de la même manière.
 - Dans la réalité, il est très rare que le choc soit parfaitement exogène, dans la mesure où il est en général corrélé à d'autres caractéristiques du groupe de traitement qui peuvent directement affecter Y .
- Exemples:
 - Évènement naturel.
 - Mesure réglementaire ou loi: instauration du salaire minimum.

Expériences naturelles - Application

IMPACT DU SALAIRE MINIMUM SUR L'EMPLOI

- Y représente le taux d'emploi (la population active occupée divisée par la population en âge de travailler).
- X représente la hausse du salaire minimum:
 - Avant: $X = 0$ pour les deux groupes.
 - Après: $X = 1$ pour le groupe de traitement et $X = 0$ pour le groupe de contrôle.

⇒ Quel est l'impact causal de la hausse du salaire minimum X sur le taux d'emploi Y ?

	Traitement	Contrôle
Avant	63, 1	61, 7
Après	64, 2	62, 2

Expériences aléatoires - Définition

- **Définition:** Une expérience aléatoire (ou expérience contrôlée) consiste en la détermination du groupe qui bénéficiera de la mesure et du groupe qui n'en bénéficiera pas au sein d'une population d'individus très similaires. Les groupes de traitement et de contrôle sont choisis par le chercheur de manière aléatoire et selon un protocole.
- Dans le cadre d'une expérience aléatoire, les groupes de traitement et de contrôle sont fabriqués.
- les deux groupes sont choisis de sorte que les individus des deux groupes sont identiques en tout point, hormis au regard de la mesure dont on veut tester l'effet causal.
- Lorsque l'expérience aléatoire est impossible (pour des raisons éthiques ou de faisabilité), le choix des groupes de traitement et de contrôle doit être « imaginé » par le chercheur de manière à ce que les groupes aient des caractéristiques aussi proches l'un de l'autre que possible.
- Exemples:
 - Expérimentations de Pôle Emploi.
 - Self-sufficiency project.

Expériences aléatoires - Application

En 1992, le Canada a lancé un programme nommé le Self-Sufficiency Project (SSP). Le programme était destiné à 9 000 familles monoparentales vivant dans deux provinces du pays, le but étant de réduire la pauvreté, d'augmenter l'emploi et d'aider à la transition vers une situation meilleure. Son principe était d'inciter au retour à l'emploi les personnes vivant d'allocations sociales, en leur versant un complément de salaire si elles retrouvaient seules un emploi à temps plein. Ce complément de salaire pouvait atteindre 25% du salaire obtenu. Une particularité de ce programme est que la sélection des 9 000 familles a été purement aléatoire et qu'un suivi de familles répondant aux critères mais ne bénéficiant pas de l'aide a également été mis en place. Ce programme contenait donc dans son principe celui de sa propre évaluation.

De nombreuses études ont tenté d'évaluer si cette politique publique a eu des effets significatifs sur la situation de ces familles. Nous allons ci-dessous en étudier la pertinence.

Expériences aléatoires - Application

⇒ Une façon d'évaluer la politique serait d'analyser la dimension temporelle, c'est-à-dire qu'on va comparer la situation des familles qui ont bénéficié de l'aide au moment où elles ont reçu cette aide et leur situation un certain nombre de mois après. Supposons qu'on remarque que le taux d'emploi a augmenté parmi ces individus. Peut-on conclure à l'efficacité du programme?

⇒ Une autre façon de procéder serait de considérer le taux d'emploi des adultes de familles monoparentales dans la province ayant bénéficié de la mesure et celui d'une province n'en ayant pas bénéficié. Est-ce une façon satisfaisante de procéder?

⇒ Qu'entend-on par « groupe de traitement » et « groupe de contrôle »? Expliquez.

Expériences aléatoires - Application

⇒ En quoi le fait que ces familles ont été choisies de manière aléatoire rend l'étude de cas pertinente? Par exemple, pourquoi douterions-nous des résultats présentés si l'aide financière avait été attribuée en fonction de la richesse du ménage ou si les familles avaient été choisies de façon discrétionnaire par des organismes sociaux locaux?

⇒ Qu'entend-on par « variables de contrôle »? Pourquoi un grand nombre d'études doivent s'appuyer sur de telles variables? L'étude décrite ci-dessus doit-elle considérer de telles variables?

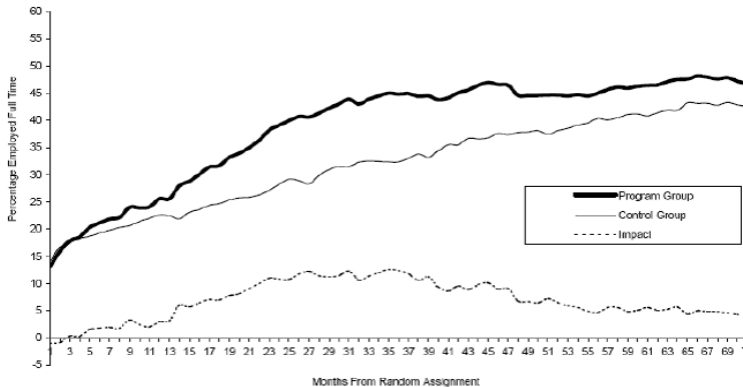
⇒ Le résultat principal de l'étude est le graphique qui se trouve ci-dessous. Le programme a démarré au mois numéroté 12 et s'est achevé au mois 48. Les deux groupes ont été suivis 24 mois après la fin du programme (jusqu'au mois 71). Commentez ce graphique.

Régression
Moindres carrés ordinaires
Corrélation vs. Causalité
Causalité au sens de Granger
Différence-en- Différences

Contrefactuel
Groupes de traitement et de contrôle
Différence-en-Différences
Expériences naturelles
Expériences aléatoires

Expériences aléatoires - Application

Figure 3.2: Percentage Employed Full Time, by Months From Random Assignment, in the SSP Applicant Study



Sources: Baseline survey and 12-month, 30-month, 48-month, and 72-month follow-up survey data.