

COMMENT SAVOIR SI UNE RELATION DE CORRELATION CACHE UNE RELATION DE CAUSALITE ?

L'analyse des corrélations place l'économiste devant deux problèmes. D'une part, il lui faut déterminer si une relation de corrélation correspond ou non à une relation de causalité, et d'autre part, si relation de causalité il y a, quel est le sens de causalité existant entre les deux variables. L'objectif de ce TD est de vous familiariser avec les deux principales techniques utilisées par les économistes pour répondre à ces questions. Quel est leur intérêt ? Il est absolument essentiel, car faire ce travail, c'est être capable de débusquer l'erreur derrière nombre de discours publics...

Ainsi, à la fin de ce TD, vous serez capable de mettre en place les conditions d'analyse pour déterminer si on doit croire les affirmations suivantes :

- « Si tu révises bien tes cours, tes notes seront meilleures »
- « Grâce à la politique de baisse des charges sociales mise en place par le gouvernement, le nombre d'emplois a augmenté de 5% »
- « Grâce à la mise en place des caméras de vidéo-surveillance dans la ville, la délinquance a diminué de 10% »
- « Quand X joue en équipe de France de football, l'équipe gagne, donc il faut systématiquement faire jouer X »

Tentant, non ? Alors laissons-nous guider...

I- LES DEUX GRANDES TECHNIQUES DE MESURES DE LA CAUSALITE

1) Technique 1 : la causalité au sens de Granger

• Le principe

Clive Granger a reçu le prix Nobel d'économie en 2003. La relation de causalité au sens de Granger part d'une intuition simple. S'il y a causalité entre deux variables, on doit généralement trouver une corrélation entre l'évolution de ces deux variables. Par exemple, s'il y a causalité entre le nombre d'objets produits par un salarié et le salaire de ce salarié, alors nécessairement si le salarié produit plus, son salaire doit alors augmenter. Ainsi :

Document 1

	Nombre de produits fabriqués par un travailleur dans le mois	Salaire mensuel
Janvier	20	1 000
Février	30	1 500
Mars	40	2 000
Avril	50	2 500

Question 1 : Y a-t-il corrélation entre le nombre de produits fabriqués de travail et le salaire obtenu ?

Vous avez du répondre qu'il y a une corrélation, qui montre qu'à chaque produit fabriqué est associé 50 euros de salaire. Donc à chaque produit réalisé en plus, le salarié pourra espérer obtenir 50 euros supplémentaires. Il semble donc y avoir une relation de causalité claire, qui montre que c'est le nombre de produits fabriqués qui détermine le salaire obtenu

Cependant, les choses sont peut-être plus compliquées qu'il n'y paraît... ainsi, imaginons qu'on nous apprenne qu'en fait le salaire mensuel est déterminé *avant* le début du mois. Ainsi, c'est en décembre qu'il a été décidé de donner 1 000 euros au salarié pour le mois de janvier, puis en janvier qu'il a été décidé de lui donner 1 500 euros pour le mois de février... dans ce cas là c'est en fait la causalité inverse qui a joué : c'est le niveau de salaire qui détermine la motivation et le niveau d'effort du salarié, et donc sa production !

On en déduit que la causalité qui nous paraissait au départ « évidente » au regard de la corrélation était en fait fautive... Comment l'avons-nous découvert ? En regardant quelle variable variait *avant* l'autre. En effet, si une variable varie systématiquement avant l'autre, on peut logiquement en déduire que le sens de causalité va dans ce sens et non dans l'autre. Par exemple, il y a corrélation entre la pratique du sport et le fait d'être essoufflé. En suivant notre méthode, il n'y a pas de doute possible sur le sens de causalité : c'est bien parce qu'on a fait du sport *avant* qu'on est essoufflé *après*, et non l'inverse !

Cette manière de savoir s'il y a causalité et quelle est son sens au sens de Granger peut être définie ainsi : **une variable X cause la variable Y si les valeurs passées de X ont un impact statistique sur la valeur actuelle ou future de Y.**

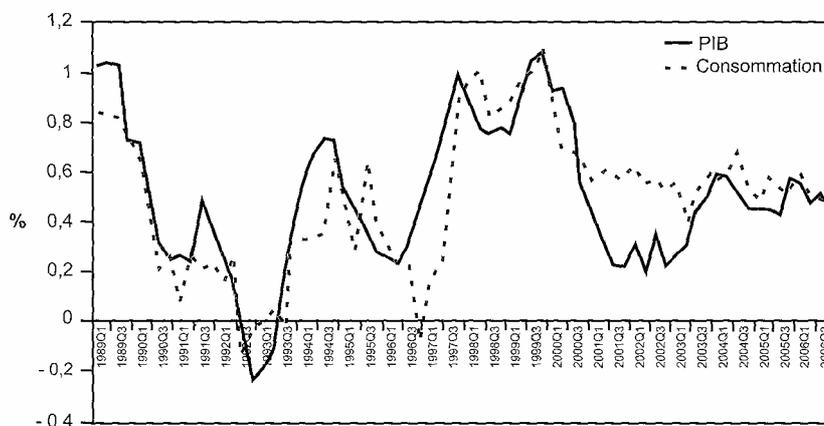
Pour ceux qui ne sont pas rebutés par les mathématiques, on peut écrire : $X(t) = a + bY(t-1)$
Si b est significativement différent de 0, alors on peut dire que Y cause X au sens de Granger avec un retard d'une période.

L'idée est de regarder l'évolution de certaines variables dans les séries chronologiques. Si une variable semble précéder en termes d'évolutions une autre variable, alors on peut supposer un lien de causalité.

Exemple : le lien consommation / production.

⇒ Question : est-ce la hausse de la consommation qui génère la hausse de la production (approche keynésienne), ou l'inverse (approche classique) ?

Document 2 : Evolution du PIB réel et de la consommation en France



Sources : Main Economic Indicators, OCDE; variables ajustées des variations saisonnières; prix constants 2000.

Source : Main Economic Indicators, OCDE.

Il semble ici que les variations de la production précèdent de deux ou trois trimestres ceux de la consommation : la causalité serait donc de la production à la consommation.

- Les limites

Cette technique connaît cependant deux principales limites :

- 1) Il se peut que Y influe sur X en $t-2$... auquel cas on aurait une causalité circulaire ! D'où la question : quel horizon temporel retenir ? Ainsi, dans notre document 1, imaginons que le salaire proposé en décembre pour janvier ait lui-même été décidé à la suite... de la production réalisée en novembre, alors dans ce cas la causalité irait de la production au salaire, et non l'inverse...
- 2) Les anticipations viennent perturber le test... exemple : la météo un soir me dit qu'il va pleuvoir demain après midi. En partant de chez moi le matin je prends mon parapluie, et l'après midi il pleut. Au sens de Granger, on en déduira que c'est le fait de prendre le parapluie qui génère l'apparition de la pluie ! La relation de Granger peut donc prédire des causalités inverses à celles qui existent réellement.

2) Technique 2 : les expériences contrôlées

En économie, la méthode qui permet de se rapprocher le plus des sciences exactes est l'expérience aléatoire contrôlée. L'idée s'inspire directement des études médicales où il y a un groupe de contrôle qui reçoit un placebo, et un groupe de traitement (qui reçoit le vrai médicament). A priori, si on observe une différence dans le niveau de santé ou le développement de la maladie entre les deux groupes, celle-ci devrait être une bonne approximation de l'impact du traitement, puisque les deux groupes sont théoriquement identiques.

La condition (et la limite) de cette stratégie est que les deux groupes soient au préalable les plus ressemblant possibles pour qu'il n'y ait pas de biais sur les résultats...

Ces expériences contrôlées reposent sur la méthode du « contrefactuel ».

Contrefactuel

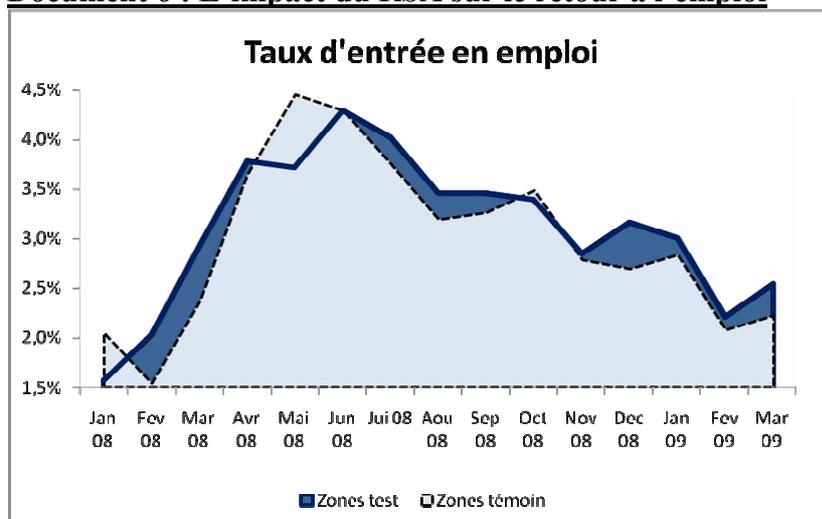
On mesure la causalité en utilisant la notion de contrefactuel. Le contrefactuel est une situation virtuelle qui se serait produite si le phénomène dont on cherche à mesurer l'impact causal n'avait pas eu lieu. L'écart entre la réalité, affectée par le phénomène en question, et le contrefactuel, correspond alors à l'impact causal.

Exemple : le diplôme a-t-il une influence sur le revenu ? Le contrefactuel est ici le revenu d'un individu s'il n'avait pas de diplôme. La différence entre son revenu réel avec diplôme et ce revenu contrefactuel mesurera alors précisément l'impact causal du diplôme sur les revenus. Le problème est qu'on ne saura jamais quel aurait été le revenu de CET individu s'il n'avait pas eu de diplôme... il va donc falloir développer des méthodes statistiques pour tenter de mesurer ce contrefactuel.

Un bon exemple d'oubli du contrefactuel est quand un homme politique dit : « grâce à ma mesure l'emploi a augmenté de 5% ». Certes l'emploi a augmenté de 5%... mais que ce serait-il passé si la mesure n'avait pas été prise ? Peut-être l'emploi aurait-il « spontanément » augmenté de 4%, auquel cas en fait la mesure n'a eu un impact qu'à hauteur de 1% ! Et peut-être même l'emploi aurait-il augmenté de 8%... auquel cas la mesure a en fait eu un impact négatif sur l'emploi !

Exemple : le test de Marc Gurgand sur le RSA

Document 6 : L'impact du RSA sur le retour à l'emploi



Source : Rapport d'évaluation du RSA