

<http://pedagogie.ac-limoges.fr/ses/spip.php?article398>



# Corrélation et causalité

- Méthodes - Déjouer les pièges ! -



Date de mise en ligne : mardi 9 mars 2010

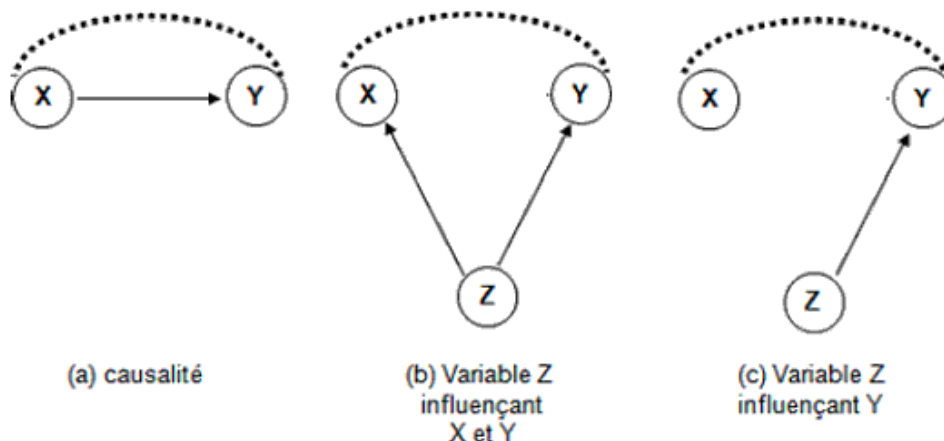
---

Copyright © SES Limoges - Tous droits réservés

---

Dans le glossaire de statistique descriptive figurant sur le site de [Fabrice Mazerolle](#), on peut lire cette présentation très claire.

« Le plus souvent, l'étude des relations entre deux variables a pour but plus ou moins avoué d'apprécier dans quelle mesure l'une des deux variables – dite variable explicative – exerce une influence causale sur l'autre – dite variable expliquée. Malheureusement, la corrélation n'implique pas la causalité, pour diverses raisons que nous allons maintenant évoquer. La figure ci-dessous illustre trois liens possibles entre les deux variables X et Y, liens qui sont tous compatibles avec un coefficient de corrélation identique, lequel ne permettra donc pas de discriminer entre les trois.



Source : d'après David S. MOORE et George P. McCABE, 2001, *Introduction to the Practice of Statistics*, W.H. Freeman & Company, New York, 3e édition, page 208.

Sur la figure, les lignes en pointillés indiquent l'existence d'une corrélation entre les variables X et Y. Les lignes en trait plein indiquent l'existence d'une causalité et la flèche indique le sens de la causalité. Dans le cas (a), nous voyons que la causalité sous-jacente va de X vers Y, c'est-à-dire que les variations de X expliquent celles de Y. La corrélation observée est donc bien le résultat d'une causalité directe.

Cependant, comme la causalité n'est pas observable, on ne peut pas conclure à l'existence d'une causalité de X vers Y à la simple mise en évidence d'une corrélation. En effet, comme l'illustrent les cas (b) et (c) de la figure, la corrélation peut aussi s'expliquer différemment. Dans le cas (b), c'est une variable Z, qui peut être inconnue ou connue mais non prise en compte, qui influence simultanément X et Y. Dans ce cas, on observera effectivement une corrélation entre X et Y, mais cette corrélation n'impliquera pas de causalité de X vers Y. Dans le cas (c), c'est une variable Z, qui peut être inconnue ou connue mais non prise en compte, qui influence uniquement Y. Dans ce cas, on observera effectivement une corrélation entre X et Y, mais cette corrélation n'impliquera pas de causalité de X vers Y, puisque la variation de X est autonome et celle de Y causée par la variable Z.

En conclusion, il faut retenir que corrélation n'est pas causalité. »

La confusion est pourtant très fréquente ce qui cause (je rigole) à la fois la joie et la fureur des commentateurs. Cela a même donné l'idée à [Etienne Wasmer](#) de créer un « Prix Causalis » dont le Règlement indique dans ces deux premiers articles :

Article 1. Le Prix Causalis récompense toute personne ou organe de presse confondant volontairement et en toute connaissance de cause corrélation et causalité afin de : promouvoir sa politique ; dénigrer celle du gouvernement ou du gouvernement précédent ; vendre son journal ; susciter une polémique.

Article 2. Le Prix Correlis récompense toute personne ou organe de presse confondant involontairement corrélation et causalité.

Pour lancer la compétition il fait lui même trois propositions et il ya eu de nombreux candidats. Pour voir l'étendu du désastre vous devez vous reporter à [Corrélation et causalité](#). Au même moment, avec un peu plus de précision puisque le questionnement portait sur une définition particulière (mais utilisée très souvent par les économistes) de la causalité il proposait aux lecteurs de donner le sens de la relation entre température et émission de CO2. L'une des conditions d'existence d'une causalité au sens de Granger est l'antériorité de la cause. Le graphique montre le contraire de ce qui est généralement admis : ce serait l'élévation de la température qui expliquerait l'augmentation des émissions de CO2. Les commentaires qui sont autant de réponses sont intéressants et vous pouvez les lire [avec le billet](#).

C'était en 2007 et comme la confusion a la peau dure, Etienne Wasmer en remet une couche dans un [billet du 3 janvier 2010](#) à propos d'un article publié dans Le Monde.

Si tout cela titille votre imagination et que vous voulez en savoir un peu plus sur les méthodes statistiques utilisées en particulier pour la causalité au sens de Granger vous pouvez lire ces deux articles trouvés sur le site d'Arthur Charpentier [1] :

- [Causalité et corrélation, Cum hoc ergo propter hoc](#)
- [Approche statistique de la causalité](#)

Vous pouvez aussi lire sans effort cet extrait de *Statistiques : méfiez vous !* de Nicolas Gauvrit [2] repris sur le site Â« Chralatans.info Â» et intitulé [Les Variables de confusion](#).

Enfin je vous recommande fortement la lecture de l'article que Claude Bordes a consacré en 2008 à la relation entre [Santé et pauvreté dans les pays en développement](#).

---

[1] Maître de conférence à l'université de Rennes, économiste, statisticien et actuair

[2] Ellipses Marketing (14 février 2007)