格兰杰因果性检验评述

曹永福*

摘要： 格兰杰因果性检验是计量经济学中最常用的因果性检验方法，但是因为种种原因，目前存在着对该检验方法的模糊认识和不正确运用。本文回顾了格兰杰因果性检验的发展过程，并从信息集、非平稳变量、即期因果性等方面对运用格兰杰因果性时存在的一些问题进行了讨论。

关键词： 因果性 格兰杰因果性 信息集

JEL 分类： C0、C3、C4

1.引言

探讨因果性是科学研究的重要使命，我们的日常生活也几乎离不开“因果”，但是到底什么叫“因果性”，从哲学上并没有一个普遍接受的概念。伽利略认为原因是结果的充分必要条件，休谟主张因果律是一种“恒常的联系（constant conjunction）”，而后来 Hart 和 Honore 则主张“原因是导致结果发生变动的一种变动”①。因果性的概念可谓众说纷纭，哲学家们直到今天仍然在为建立更合理的因果性概念而不断努力。

与哲学家的纯粹逻辑思辨不同，经济学中的因果性必须是可以通过实际的数据进行检验的。一般来讲经济学中的因果性概念与其具体的检验方法是统一的，其中最有影响力的是格兰杰因果性。但是因为缺乏对原始文献的准确把握，目前对格兰杰因果性存在一些模糊认识，这个问题已经引起部分学者的注意，如庞皓、陈述云（1999），周建、李子奈（2004）等等。但是，从实际情况看，格兰杰因果性仍然存在很多不恰当的应用。

为了更清楚的认识格兰杰因果性的全貌，本文首先简要介绍了因果性定义的概率论方法，然后全面介绍了格兰杰因果性的由来以及操作方法，并且从信息集、变量非平稳性、即期因果性等方面对国内目前实证研究中存在的问题进行了探讨。

2.因果性的概率论方法

因为经济学研究无法象自然科学那样可以进行重复的、可控制的试验，而且研究的现象往往具有不确定性，因此基于概率论的方法成为因果性研究的重要方法。其中格兰杰因果性就属于概率论方法的范畴。

因果性的概率论方法中最简单的就是 Suppes（1970）的定义。Suppes 认为，如果事件 A 的发生增加了事件 B 发生的概率，即 $P(B|A)>P(B)$，那么事件 A 就构成事件 B 的原因。比如吸烟为事件 A，患癌症为事件 B，如果整体患癌症的概率 $P(B)=2\%$，而吸烟患癌症的概率 $P(B|A)=5\%$，吸烟者患癌症的概率更大，则认为吸烟就构成患癌症的原因。

---

*作者单位：中国社科院世界经济与政治研究所统计分析研究室。通讯地址：北京建国门内大街 5 号，邮编 100732；Email：caoyongfu@eyou.com。

① 对这几种因果性概念的详细探讨参见倪培民（2004）。

Suppes（1970）还认为 P(B|A)<P(B)，则事件 A 构成事件 B 的反向因果性，此外，还有评论者主张只要 P(B|A)≠P(B)，即如果 A 的发生改变了 B 的概率分布，那么就可以认为事件 A 构成事件 B 的因果性。

Suppes（1970）定义的一个缺点是他没有考虑事件发生的先后，有时从统计上无法区分原因和结果。因为从公式推导上看，如果 P(B|A)>P(B)则 P(A|B)>P(A)一定成立[①]，也就是说如果 A 是 B 的原因，则必然可以推导出 B 是 A 的原因。沿用前面的例子，如果 P(患癌症|吸烟)>P（患癌症），那么按照公式一定可以推出 P(吸烟|患癌症)>P（吸烟），患癌症的人吸烟的概率更大，我们能否说患癌症是导致吸烟的原因呢？显然这个结论是无法接受的。这个问题在统计学上被称为观测性等价（Observational equivalence）。解决这个问题的办法是考虑原因和结果之间的先后时序，或者从理论上默认因果性的方向。

需要指出的是，Suppes（1970）称其定义的这种因果性为有初步证据的因果性（prima facie causality）。为什么要加这个限定词呢？因为有的时候尽管 A 的发生增加了 B 发生的概率，但是有时候可能 A 并不是 B 的原因。可能实际上有一个“共同因素”C 影响 A 和 B 造成了 A 与 B 之间的联系，也有可能 A 除了直接影响 B 之外，还通过影响 C 来间接影响 B 等等，情况可能非常复杂，因为世界是错综复杂的。

这里仅以“共同因素”的情形为例加以说明。比如观察到气压计变动(A)，下雨(B)的概率就会增加，即 P(B|A)>P(B)。那么气压计变动是不是下雨的真正原因呢？显然不是，因为气压计变动和下雨背后有一个共同的原因，那就是大气压的变动。如果把大气压变动（C）作为一个事件加以考虑，A 就不再增大 B 发生的概率[②]，这时候 A 实际上就不是 B 的原因。可见，因为 Suppes（1970）这个定义仅仅考察两个事件的简单情形，难免会造成虚假的结果。因此只能称之为有初步证据的因果性。

3.格兰杰因果性的定义与操作

与 Suppes（1970）的基于随机事件的定义相比，格兰杰的因果性定义（Granger，1980）运用了信息集的概念，而且强调了事件发生的时序。下面给出格兰杰因果性的定义。

要检验 X 和 Y 之间的因果性，令 Ω_n 为到 n 期为止宇宙中的所有信息，Y_n 为到 n 期为止所有的 Y_t (t=1…n)，X_{n+1} 为第 n+1 期 X 的取值，Ω_n-Y_n 为除 Y 之外的所有信息。如果我们承认“现在和过去可以影响未来，而未来不能影响过去”，并且假设 Ω_n 中不包含任何冗余的信息。则如果：

$$F(X_{n+1} \mid \Omega_n) \neq F(X_{n+1} \mid (\Omega_n - Y_n)) \qquad (1)$$

即可以认为变量 Y 对变量 X 有格兰杰因果性。可以看出格兰杰因果性的定义与 Suppes（1970） 关于 P(B|A)≠P(B)的定义非常相似，但是他的概念是基于信息集的，这样就把需要考

[①]因为 P(AB)=P(B|A)P(A)=P(A|B)P(B)，P(B|A)/P(B)=P(A|B)/P(A)。因此如果 P(B|A)>P(B)，则必然可以推出 P(A|B)>P(A)。

②实际上此时 P(B|AC)=P(B|C)。

虑的因素大大拓展而不是单纯考察两个事件；而且他强调事件发生的时序性，如果 Y 构成 X 的原因，那么本期的 Y 会影响下一期 X 的概率分布。换句话说，如果 Y_n 含有 X_{n+1} 特有的预测信息，那么 Y 对 X 就构成格兰杰因果性。

上文中的信息集 Ω_n 不仅包括所有的相关变量并且包括变量的无限滞后值，但现实当中，我们无法得到所有信息 Ω_n，我们必须大大缩小信息集的范围，将 Ω_n 改为目前可获得的信息集 J_n。如果：

$$F(X_{n+1} \mid J_n) \neq F(X_{n+1} \mid (J_n - Y_n)) \qquad (2)$$

也就是说在信息集 J_n 内，Y 的信息可以改善对 X 的预测，则相对于信息集 J_n，Y 构成 X 的有初步证据的因果性（prima facie causality）。将宇宙信息集 Ω_n 改为可获得信息集 J_n 以后，格兰杰跟 Suppes 一样也加上了有初步证据的这个限制词。原因跟上文中的论述比较相似。因为这时我们的信息集是不完整的，我们无法确定是否在现有信息集 J_n 以外还有其他变量会对 X 产生影响，也就无法确定 Y 是否是真正的“原因”所在（还可以以上文中的共同因素的情况作为例子）。

格兰杰（Granger，1980）指出，信息集中遗漏重要解释变量很可能导致虚假的因果性推断，如果适当的拓展信息集合，把重要的变量引入信息集，原来的虚假因果性可能消失。因此，在现有可获得的信息集中做出因果性的推断必须要谨慎对待，这是格兰杰在定义时加上限定词的缘由。

在实际操作的时候，检验变量的分布函数是否相等是非常困难的，更简便的方法是按照公式（3）从期望值的角度来处理[①]。如果：

$$E(X_{n+1} \mid J_n) \neq E(X_{n+1} \mid (J_n - Y_n)) \qquad (3)$$

也就是说，如果预测误差 $\delta_{n+1} = E(X_{n+1} \mid J_n) - E(X_{n+1} \mid J_n - Y_n)$ 显著的不为 0，那么就认为 Y 含有对 X 的预测信息，Y 就相对于信息集 J_n 构成对 X 有初步证据的因果性。

后来的发展则逐渐过渡到从预测精度的角度来检验，即令 $\sigma^2(X_{n+1} \mid J_n)$ 为在给定信息集下对 X 的预测误差的方差，$\sigma^2(X_{n+1} \mid (J_n - Y_n))$ 为没有 Y 的情况下对 X 的预测误差，如果 $\sigma^2(X_{n+1} \mid J_n) < \sigma^2(X_{n+1} \mid (J_n - Y_n))$，也就是说如果变量 Y 的存在能够显著地减小对 X 的预测误差的方差（或者说能够显著地改善对 X 的预测精度），那么就可以认为 Y 对 X 有因果性。到此为止，就可以得到通常计量经济学教科书上关于格兰杰因果性检验的方法，多数教科书中都以两个变量为例（见张晓峒，2000）：

[①] 公式（2）基于分布函数的定义被称做全面因果性，公式（3）的定义被称做均值因果性。Covey（1992）曾经就全面因果性的检验方法进行过探讨。

建立一个回归方程：

$$x_t = \sum_{i=1}^{k} \alpha_i x_{t-i} + \sum_{i=1}^{k} \beta_i y_{t-i} + u_{1t} \qquad (4)$$

原假设为 Y 不构成对 X 的因果性即 H_0: $\beta_1 = \beta_2 = ...\beta_k = 0$，则在原假设成立的情况下：

$$F = \frac{(SSE_r - SSE_u)/k}{SSE_u/(T-2k)} \sim F_{(k,T-2k)} \qquad (5)$$

其中 SSE_r 为施加约束时的残差平方和（也就是没有 Y 的情况下 X 自身做回归的残差平方和），SSE_u 为没有施加约束时的残差平方和，T 为样本容量，k 为最大滞后阶数。这个检验的思路是，如果考虑 Y 的情况下的残差平方和显著的小于没有 Y 的情况下的残差平方和，就认为 Y 的存在显著的提高了对 X 的预测精度，Y 对 X 就有因果性。

综上所述，为了操作的方便，格兰杰的因果性经过了一系列的演化，尽管基本保留了原来的思想，但与最初的定义已经有较大差别。

4.格兰杰因果性应用时必须注意的问题

因果性在经济分析中占有至关重要的地位，格兰杰因果性因为其简单实用而得到了普遍青睐。然而，很多应用都存在一些问题，连格兰杰（1980）也承认：“最近的一些作者因为缺少对原始文献的阅读，很多应用都是不恰当的，就象一个聚会中的传言一样，以讹传讹，最后与初始的东西面目全非”。下面讨论目前运用中出现的常见问题。

（1）关于信息集的问题

上文中已经提到，格兰杰在把信息集从宇宙信息集 Ω_n 改为可获得的信息集 J_n 以后，加了“相对与某某信息集的有初步证据的”等限制词，因为如果在信息集中遗漏重要的变量，推导出来的因果性可能是虚假因果性。但是随着人们应用的越来越多，这些修饰词逐渐被人们所遗忘，直接称之为因果性，但是有些结论显然是欠妥当的。

目前教科书和实际应用中多数是两个变量之间检验，很容易出现遗漏重要相关变量的情形。比如，有计量结果显示上证指数和深证成指之间存在因果性，也就是说两者之间相互包含对方的预测信息，但是这种因果性可能是虚假的。如果我们承认股票市场的走势主要取决于另外一个变量即资金供求状况，那么资金供求状况就可能同时影响这两个指数。如果把资金供求状况这个变量纳入信息集，上证指数和深证成指之间的因果性或许就消失了。

因此，在进行因果性推断时，合理的定义信息集是非常重要的。因为我们无法采用宇宙中所有的信息，而在有限的信息集中如何保证不遗漏重要的相关变量，是非常关键的。在选择变量的时候，重要的是从经济理论出发进行取舍。

（2）关于非平稳变量的问题

这个问题其实已经困绕实证研究者很长时间。原始的格兰杰因果性定义并没有规定变量必须是平稳的。当格兰杰（1969）首先提出他的因果性定义的时候，非平稳变量的问题在计量经

济学领域还刚刚起步。在 20 世纪 70 年代，对变量平稳性的探讨逐渐在计量经济学占据了重要地位。而在格兰杰（1980）重新回顾他的因果性定义的时候，对变量平稳性也没有再深入分析，他只是指出："非平稳变量带来的问题过于复杂，不便展开讨论"。

但是目前有一点学术界是有定论的，就是如果变量是非平稳的，那么应用公式（5）中的 F 统计量来做推断会产生问题。He（2001）运用维纳过程推导出，当变量为非平稳时间序列时，该统计量的渐进分布不再是 F 分布；周建、李子奈（2004）运用蒙特卡洛模拟也得出当变量为非平稳时间序列时，任何无关的两个的变量间都很容易得出有因果性的结论。因此，在实证研究时，一般认为只有平稳变量才能应用公式（5）中的 F 统计量进行推断，否则结论可能是不可靠的。

最典型的例子就是货币供应量与 *GDP* 之间的因果性检验，对于"货币是否中性的"（也就是货币是否影响实际产出）的讨论很多都是采用因果性检验来讨论。因为这两个变量都是非平稳的，如果不对变量进行平稳性处理，那么一般都可以得出两者之间存在因果关系的结论，也就不存在"货币是否中性"的争论了，这显然是不妥当的。

然而国内的研究中还是经常可以见到直接对非平稳变量运用公式（5）的 F 统计量进行因果性检验做法。笔者认为，对这种方法得出的结论须谨慎对待。

（3）关于即期因果性（Instantaneous Causality）的概念

有人认为，一个变量的变化在当期就可以影响另一个变量的变化，因此，应该建立即期因果性的概念。即把公式（1）改为 $F(X_n \mid \Omega_n) \neq F(X_n \mid (\Omega_n - Y_n))$，只要 Y 对当期的 X 能产生预测作用，就可以认为 Y 对 X 有即期因果性。

这个定义的设想是好的，因为它可以考察变量之间的当期联系。但问题是如果接受这个定义，Y 对 X 有即期因果性那么必然意味着 X 对 Y 也有即期因果性，除非有另外的信息来区分因果性的方向，否则这个定义将是不能令人满意的。

格兰杰（Granger，1988）指出只要我们承认原因在前，结果在后，那么很多表面上的即期因果性实际上是由于观测频率引起的。比如一个变量引起另外一个变量的变化只需要一周的时间，但是我们只能每月月底观测一次数据，那么就会观察到两个变量同时变化，即表现为即期因果性。但是如果我们改成一周观测一次数据，就可以观察到两者的时滞了。经济时间序列的因果传递有的时滞可能非常短，甚至可能在瞬间完成，但是只要承认有时滞，那么从道理上讲这种因果性定义的必要性就值得商榷。

与格兰杰因果性概念相联系的问题还有外生性、因果性与理性预期等等诸多问题，在此不再详细论述。

5.结论

因果性是一个历史悠久而又不断争论的问题。直到今天，包括经济学在内的各个学科对因果性并没有一个普遍接受的定义。格兰杰的因果性思想是原因对结果应该有预测性。这种思想其实尚可讨论，比如 Hoover（2001）提出了基于控制变量的定义，以及 Glymour 等基于图论方

法的定义[①]。可见在计量经济学领域因果性检验方法仍然在不断发展。

翻遍国内外普通的计量经济学教材，涉及到因果性概念的章节一般只有格兰杰因果性这一部分，这就容易给人造成一种错觉，即只有格兰杰因果性检验才可以发现变量间的因果性，其他的分析都没有考虑因果性。其实除了前面提到的种种定义以外，包括相关性分析、普通的回归分析都与因果性的概念有联系[②]，因果性分析应该是计量经济学分析的基础考虑。

总之，格兰杰因果性检验并非发现因果性的灵丹妙药，也不是唯一药方。连格兰杰本人(Granger, 1988)也承认，当从理论上相信变量之间有因果性的时候，这种检验可以增强对因果性的信心。但是如果你不相信这种因果性的定义，也不必强求。这应该是对格兰杰因果性的准确定位。

参考文献

庞皓、陈述云："格兰杰因果性检验的有效性及其应用"，《统计研究》1999 年 11 期。

倪培民："再铸因果概念"，《世界哲学》2004 年第 1 期。

张晓峒：《计量经济分析》，南开大学出版社，2000 年。

周建、李子奈："Granger 因果关系检验的适用性"，《清华大学学报（自然科学版）》2004 年期。

Covey, Ted, David Bessler, *Testing for Granger's full causality*, The review of Economics and Statistics, Vol.74, issue1,1992.

He,Zonglu, Maekawa Koichi, *On spurious Granger causality*, Economic Letter, 73, 2001

Hoover, *Causality in Macroeconomics*, Cambridge University Press, 2001.

Granger, *Investigating causal relations by econometric models and cross-spectral methods,* Econometrica, Vol.37,NO.3, 1969.

Granger, *Testing for causality: a personal viewpoint*, Journal of Economic Dynamics and control, 2,1980.

Granger, *Some recent developments in a concept of causality*, Journal of Econometrics, 39,1988.

Suppes, *A probabilistic theory of causality*, North-Holland, Amsterdam,1979.

[①]见 Hoover(2001)第 155 页。

[②]历史上曾经主张高度的相关性就意味着因果性，后来多数的计量经济学家认为相关性和因果性是有区别的。