« ATT » et Contre-factuelles dans le cas discret

Présentation du Rapport de Stage

Leyth Akrout

Université d'Evry

6 juin 2022

Sommaire

1. Étude Clinique
   - Contexte Mathématique
   - Contre-factuelles
   - ATT
2. But du Stage
   - Article Initial
   - Calcul du biais
   - Généralisation
3. Travail réalisé
   - Modèle VINGARCH(p,q)
   - Estimation
   - Biais discret
4. Pistes d'Amélioration

On note $t \to D(t)$ le traitement à valeur dans $\{0, 1\}$

On note $t \to N(t)$ le processus qui compte les évènements d'intérêt :
Dépend de l'étude :

- Le nombres de morts
- Le nombres de rémission

On note $t \to \mu(t)$ l'intensité du processus N vérifiant que :

$$\mu(t)dt = \mathbb{E}(dN(t) = 1 \mid F_{t-})$$

Contexte Mathématique

Caractérisation des individus

Un individu est caractérisé par :

- $Z = (Z_1, \dots, Z_{d_Z})$ des covariables indépendantes du temps
- $X(t) = (X_1(t), \dots, X_{d_X}(t))$ des processus dépendant du temps
- S le temps d'administration du traitement

Contre-factuelles

On introduit :

- $X^{0|\widetilde{S}}$ le processus X si le traitement n'est pas été administré
- $X^{1|\widetilde{S}}$ le processus X si le traitement est administré

Avec ces notations :

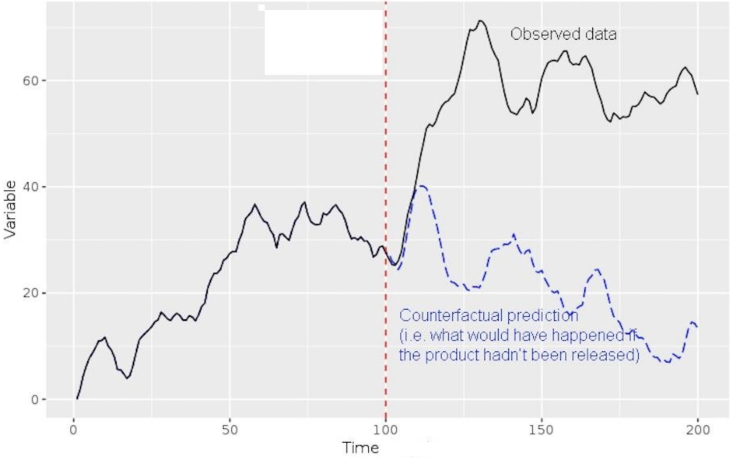
- $X(t) = X^{0|S}(t)$ si $t < S$
- $X(t) = X^{1|S}(t)$ si $t \geq S$

Les intensités sont notées :

- $\mu^{0|S}$ dans le cas sans traitement
- $\mu^{1|S}$ dans le cas avec traitement

Contre-factuelles

Le processus contre-factuelle de $(X(t))_{t \geq S}$ est donc $(X^{0|S}(t))_{t \geq S}$ qui est non observé.



ATT

L'ATT (Average treatment effect on treated) est un indicateur qui mesure l'efficacité d'un traitement parmi la population des individus traités

Avec les notations précédentes on peut définir l'ATT :

$$\mathrm{ATT}(t) = \mathbb{E}_S(\mu^{1|S}(t) - \mu^{0|S}(t))$$

Article Initial

Intensité du Processus

On se base sur l'article de Gran et al. [2], ils supposent que l'intensité s'écrit :

$$\mu(t) = \alpha_0(t) + \alpha_Z(t)Z + \alpha_X(t)X(t) + \alpha_D(t)D(t)$$

Donc si on introduit les contre-factuelles :

$$\mu^{1|S}(t) = \alpha_0(t) + \alpha_Z(t)Z + \alpha_X(t)X^{1|S}(t) + \alpha_D(t)D(t)$$

$$\mu^{0|S}(t) = \alpha_0(t) + \alpha_Z(t)Z + \alpha_X(t)X^{0|S}(t)$$

Article Initial

Calcul de l'ATT

Danc ce contexte pour $t \geq S$ l'ATT s'ecrit :

$$\text{ATT}(t) = \mathbb{E}_S(\mu^{1|S}(t) - \mu^{0|S}(t)) = \alpha_D(t) + \alpha_X(t)\Big[\mathbb{E}_S(X^{1|S}(t) - X^{0|S}(t))\Big]$$

Avec $X^{1|S}(t)$ qui est observé mais $X^{0|S}(t)$ qui ne l'est pas.

Il faut donc trouver un moyen de generer $X^{0|S}(t)$ pour $t \geq S$

Article Initial

Génération des contre-factuelles

L'idée de l'article de Gran et al. [2] est de supposer le processus $(X^{0|S}(t))_{t\geq 0}$ suit un modèle VAR

On peut à l'aide de cette supposition :

- Estimer les paramètres du modèle VAR
- Estimer les contre-factuelles $X^{0|S}(t)$ pour $t \geq S$ notée $\widetilde{X}^{0|S}(t)$
- Calculer un estimateur de l'ATT

Calcul du biais

Risque Quadratique

On peut réécrire :

$$\mu(t) = \alpha_0(t) + \alpha_Z(t)Z + \alpha_X(t)X(t) + \alpha_D(t)D(t)$$

De la maniere suivante :

$$\mu(t) = \alpha_0(t) + \alpha_Z(t)Z + \alpha_X(t)X^{0|S}(t) + d^S(t)D(t)$$

Avec :

$$d^S(t) = \alpha_D(t) + \alpha_X(t)\left[X^{1|S}(t) - X^{0|S}(t)\right]$$

Calcul du biais

Risque Quadratique

On note $A(t) = (\alpha_0(t), \alpha_Z(t), \alpha_X(t), d^S(t))^T$ les paramètres de la régression en fonction de la variable $W(t) = (1, Z(t), X^{0|S}(t), D(t))^T$.

Or ici $W(t)$ n'est pas observée pour $t \geq S$ on a uniquement accès à $\widetilde{W}(t) = (1, Z(t), \widetilde{X}^{0|S}(t), D(t))^T$

Calcul du biais

Mise en évidence du biais

Un biais est donc généré par le fait que :

$$W(t) = \widetilde{W}(t) + \xi(t)$$

Avec $\xi(t) = (0_1, 0_{d_Z}, \epsilon(t), 0_1)^T$ et $\epsilon(t)$ l'erreur $X^{0|S}(t) - \widetilde{X}^{0|S}(t)$

Calcul du biais

Expression du biais

Si on note $r_n(A, W)$ le risque quadratique recherché, Madame Agathe Guilloux et Madame Camille Nevoret ont montré que :

$$r_n(A, W) = r_n(A, \widetilde{W}) + \frac{1}{n}\sum_{i=1}^{n}\int_0^{\tau_i} \alpha_X(t)^T \mathbb{E}\left[\epsilon(t)\epsilon(t)^T\right]\alpha_X(t)dt$$

Avec n le nombre d'individus et τ_i le dernier temps ou l'on observe l'individu i.

Calcul du biais

Expression du biais

De plus, Madame Agathe Guilloux et Madame Camille Nevoret ont montré que :

$$\mathbb{E}\left[\epsilon(t)\epsilon(t)^T\right] = \sum_{j=0}^{t} \Pi^j \Sigma \Pi^{j\,T}$$

Avec Π la matrice des coefficients et Σ la matrice de variance-covariance de l'erreur du modèle VAR.

Jusqu'à présent le processus X était continu

But du Stage :

- Étendre l'article de Gran et al. [2] au cas discret
- Estimer et Corriger le biais dans le cas discret

Généralisation

Piste cas discret

Idée de Madame Agathe Guilloux : Remplacer le modèle VAR par un modèle INGARCH(p,q)

$$\lambda_t = \beta_0 + \sum_{i=1}^{p} b_i X_{t-i} + \sum_{i=1}^{q} c_i \lambda_{t-i}$$

Avec $\lambda_t = \mathbb{E}(X_t \mid F_{t-1})$, $F_t = \sigma(X_{t-j}, \lambda_0 \ / \ j = 1, .., t)$

et $X_t \mid F_{t-1}$ suive une loi de poisson de moyenne λ_t

D'après Ferland et al. [1] le modèle est :

$$\text{Stationnaire et ergodique} \iff \sum_{i=1}^{p} b_i + \sum_{i=1}^{q} c_i < 1$$

Généralisation

Problème subsistant

Il faut résoudre les problèmes suivants :

- Le modèle INGARCH(p,q) n'est défini qu'en dimension 1, peut on le généraliser ?
- Comment estimer les paramètres de cet éventuel modèle ?
- Comment exprimer le biais avec un nouveau modèle ?

Modèle VINGARCH(p,q)

Présentation du modèle

$(X_t)_{t\geq 0}$ suit un modèle VINGARCH(p,q) Poissonnien s'il existe un vecteur $\beta_0 \in \mathbb{R}^d$ et des matrices $B_1, ..B_p, C_1, .., C_q \in M_d(\mathbb{R}_+)$ tel que pour tout $t \geq max(p,q)$:

$$\lambda_t = \beta_0 + \sum_{i=1}^{p} B_i X_{t-i} + \sum_{i=1}^{q} C_i \lambda_{t-i}$$

avec $\lambda_t^i = \mathbb{E}(X_t^i \mid F_{t-1})$ et $F_t = \sigma(X_{t-j}^i, \lambda_0^i \ / \ j = 1, .., t, i = 1, .., d)$, de tel sorte que $X_t^i \mid F_{t-1}$ suive une loi de Poisson de paramètre λ_t^i et que $X_t^i \perp\!\!\!\perp X_t^j \mid F_{t-1}$, pour tout $t \geq 0$

Modèle VINGARCH(p,q)

Écriture Matricielle

$$\begin{pmatrix} X_t \\ X_{t-1} \\ \vdots \\ X_{t-p+1} \\ \lambda_t \\ \vdots \\ \lambda_{t-q+2} \\ \lambda_{t-q+1} \end{pmatrix} = \begin{pmatrix} \beta_0 \\ 0 \\ 0 \\ \vdots \\ \beta_0 \\ \vdots \\ 0 \\ 0 \end{pmatrix} + M \begin{pmatrix} X_{t-1} \\ X_{t-2} \\ \vdots \\ X_{t-p} \\ \lambda_{t-1} \\ \vdots \\ \lambda_{t-q+1} \\ \lambda_{t-q} \end{pmatrix} + \begin{pmatrix} \omega_t \\ 0 \\ \vdots \\ 0 \\ 0 \\ \vdots \\ 0 \\ 0 \end{pmatrix}$$

Avec ω_t l'erreur d'espérance nulle entre X_t et l'intensité λ_t de la loi de Poisson et M la matrice suivante :

Modèle VINGARCH(p,q)

Écriture Matricielle

$$M = \begin{pmatrix} B1 & B2 & \dots & B_p & C_1 & C_2 & \dots & C_q \\ Id_d & 0_d & \dots & 0_d & 0_d & 0_d & \dots & 0_d \\ 0_d & Id_d & \dots & 0_d & 0_d & 0_d & \dots & 0_d \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ B1 & B2 & \dots & B_p & C_1 & C_2 & \dots & C_q \\ 0_d & 0_d & \dots & 0_d & Id_d & 0_d & \dots & 0_d \\ 0_d & 0_d & \dots & 0_d & 0_d & Id_d & \dots & 0_d \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0_d & 0_d & \dots & 0_d & 0_d & \dots & Id_d & 0_d \end{pmatrix}$$

Proposition 1 : Les égalités suivantes sont vraies :

(i) $\mathbb{V}(\omega_t) = \text{diag}(\mathbb{E}(X_t))$

(ii) $Cov(\omega_{t_1}, \omega_{t_2}) = 0_d$ pour tout $\text{t}_1 \neq t_2, t_1, t_2 \geq 0$

(iii) $Cov(X_t, \omega_{t+h}) = 0_d$ pour tout h > 0 et $t \geq 0$

Proposition 2 : Soit $(X_t)_{t\geq 0}$ suivant un modèle VINGARCH(p,q) tel que :

(i) Les valeurs propres de la matrice M sont de module strictement plus petite que 1

(ii) Une seconde condition sur l'espérance et la variance du vecteur au temps initial

Alors la série $(X_t)_{t\geq 0}$ est stationnaire et ergodique pour la moyenne.

Estimation

Ensemble des Paramètres

On se place dans le cadre plus simple d'un modele VINGARCH(p,0).
L'ensemble des paramètres est :

$$\Theta = \left\{ \theta \in \mathbb{R}_+^{d(n+dp)} / \rho(M) < 1 \right\} =$$

$$= \left\{ \theta \in \mathbb{R}_+^{d(n+dp)} / \ \ \forall x \in \mathbb{C}, det\left(x^p Id_d - \sum_{i=1}^{p} x^{p-i} B_i \right) = 0 \implies \|x\| < 1 \right\}$$

On peut montrer que cet ensemble est ouvert.

Dans ce contexte on peut :

- Donner une expression explicite du logarithme de la vraisemblance
- Donner une expression explicite du gradient et de la matrice hessienne
- Montrer que la matrice hessienne est semi-définie négative

Ainsi le gradient s'annule $\Longleftrightarrow$ la vraisemblance admet un maximum global

La vraisemblance, le gradient et la matrice hessienne étant explicite on peut dont utiliser un algorithme de Newton pour estimer les paramètres du modèle.

Biais discret

Génération du processus discret

Si $(X_i^0(t))_{t\geq 0}$ suit un modèle VINGARCH(p,q) on peut montrer que le vecteur :

$$\begin{pmatrix} X_i^0(t_i+t) \\ X_i^0(t_i+t-1) \\ \vdots \\ X_i^0(t_i+t-p+1) \\ \lambda_i^0(t_i+t) \\ \vdots \\ \lambda_i^0(t_i+t-q+2) \\ \lambda_i^0(t_i+t-q+1) \end{pmatrix}$$

vaut :

Biais discret

Génération du processus discret

$$\sum_{j=0}^{t-1} M^j \begin{pmatrix} \beta_0^i \\ 0 \\ 0 \\ \vdots \\ \beta_0^i \\ \vdots \\ 0 \\ 0 \end{pmatrix} + M^t \begin{pmatrix} X_i^0(t_i) \\ X_i^0(t_i - 1) \\ \vdots \\ X_i^0(t_i - p + 1) \\ \lambda_i^0(t_i) \\ \vdots \\ \lambda_i^0(t_i - q + 2) \\ \lambda_i^0(t_i - q - 1) \end{pmatrix} + \sum_{j=0}^{t-1} M^j \begin{pmatrix} \omega_i^0(t_i + t - i) \\ 0 \\ \vdots \\ 0 \\ 0 \\ \vdots \\ 0 \\ 0 \end{pmatrix}$$

Biais discret

Génération du processus discret

Donc le processus contre-factuelle $(\widetilde{X}_i^0(t))_{t\geq 0}$ vérifie que :

$$\begin{pmatrix} \tilde{X}_i^0(t_i+t) \\ \tilde{X}_i^0(t_i+t-1) \\ \vdots \\ \tilde{X}_i^0(t_i+t-p+1) \\ \tilde{\lambda}_i^0(t_i+t) \\ \vdots \\ \tilde{\lambda}_i^0(t_i+t-q+2) \\ \tilde{\lambda}_i^0(t_i+t-q+1) \end{pmatrix} = \sum_{j=0}^{t-1} M^j \begin{pmatrix} \beta_0^i \\ 0 \\ 0 \\ \vdots \\ \beta_0^i \\ \vdots \\ 0 \\ 0 \end{pmatrix} + M^t \begin{pmatrix} X_i^0(t_i) \\ X_i^0(t_i-1) \\ \vdots \\ X_i^0(t_i-p+1) \\ \lambda_i^0(t_i) \\ \vdots \\ \lambda_i^0(t_i-q+2) \\ \lambda_i^0(t_i-q-1) \end{pmatrix}$$

Biais discret

Covariance de l'erreur

On obtient que :

$$\mathbb{V}(\epsilon_i(t)) = \mathbb{V}(X_i^0(t_i+t) - \tilde{X}_i^0(t_i+t)) = \mathbb{V}\Bigg(\sum_{j=0}^{t-1} UM^j \begin{pmatrix} \omega_i^0(t_i+t-i) \\ 0 \\ \vdots \\ 0 \\ 0 \\ \vdots \\ 0 \\ 0 \end{pmatrix}\Bigg)$$

Avec U la matrice $\begin{pmatrix} 1 & \dots & 0 & \dots & 0 \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & \dots & 1 & \dots & 0 \end{pmatrix}$, de taille $d \times d(p+q)$.

Biais discret

Expression du biais discret

A l'aide de la proposition 1, le biais vaut donc :

$$\frac{1}{n}\sum_{i=1}^{n}\int_0^{\tau_i}\alpha_X(t)^T\sum_{j=0}^{t-1}UM^j\begin{pmatrix}\mathrm{diag}(\mathbb{E}(X_i)) & 0 & \dots & 0\\ \vdots & \vdots & \ddots & \vdots\\ 0 & 0 & \dots & 0\end{pmatrix}{M^j}^TU^T\alpha_X(t)dt$$

Pistes d'Amélioration

- On peut améliorer les notions de stationnarité et d'ergodicité en utilisant des chaînes de Markov [3]
- En s'inspirant de la thèse "Some models for Time Series of Counts" de Liu [3] on peut simplifier la proposition 2
- On aimerait étendre le processus vectoriel X à un mélange de sous processus continus et discrets
- Un programme sous R est en cours de réalisation

Bibliographie

René Ferland, Alain Latour, and Driss Oraichi.
Integer-valued garch process.
Journal of Time Series Analysis, 27(6) :923–942, 2006.

J. M. Gran, R. Hoff, K. Røysland, B. Ledergerber, J. Young, and O. O. Aalen.
Estimating the treatment effect on the treated under time-dependent confounding in an application to the Swiss HIV Cohort Study.
arXiv :1604.01597 [stat], October 2016.
arXiv : 1604.01597.

Heng Liu.
Some Models for Time Series of Counts.
PhD thesis, Columbia University, 2012.