Modèle de sélection en présence de non-réponse non ignorable et population hétérogène

7e colloque francophone sur les sondages - ENSAI

Eric Gautier (CREST (ENSAE))
gautier@ensae.fr

le 7 novembre 2012

Non-réponse et sondage

- Considérons un modèle de superpopulation où chaque unité i dans la population totale est une réalisation iid dans la loi jointe de caractéristiques (Y, X^T, Z^T, U^T).
  - Y = variable sujette à non-réponse partielle dans l'enquête étudiée.
  - X, Z = vecteurs de caractéristiques observés pour tous dans l'échantillon final. X = *variable de contrôle*, Z = *variable instrumentale*. (X^T, Z^T) rend compte de *l'hétérogénéité observée*.
  - U est un vecteur de caractéristiques inobservées, il rend compte de *l'hétérogénéité inobservée*.
- Définissons 2 autres variables aléatoires : D et B.
  - $D = 1$ si l'unité se trouve dans l'échantillon et 0 sinon. L'échantillon est celui obtenu après tirage de l'enquête puis non-réponse totale. Pour simplifier, nous supposons que les poids obtenus après redressement (y compris le calage sur marge) sont les vrais inverses des probabilités de sélection.
  - $B = 1$ si l'unité dans l'échantillon fourni la valeur de son Y et 0 sinon.

Non-réponse et sondage

- Considérons un modèle de superpopulation où chaque unité i dans la population totale est une réalisation iid dans la loi jointe de caractéristiques (Y, X^T, Z^T, U^T).
  - Y = variable sujette à non-réponse partielle dans l'enquête étudiée.
  - X, Z = vecteurs de caractéristiques observés pour tous dans l'échantillon final. X = *variable de contrôle*, Z = *variable instrumentale*. (X^T, Z^T) rend compte de *l'hétérogénéité observée*.
  - U est un vecteur de caractéristiques inobservées, il rend compte de *l'hétérogénéité inobservée*.
- Définissons 2 autres variables aléatoires : D et B.
  - $D = 1$ si l'unité se trouve dans l'échantillon et 0 sinon. L'échantillon est celui obtenu après tirage de l'enquête puis non-réponse totale. Pour simplifier, nous supposons que les poids obtenus après redressement (y compris le calage sur marge) sont les vrais inverses des probabilités de sélection.
  - $B = 1$ si l'unité dans l'échantillon fourni la valeur de son Y et 0 sinon.

Modélisation hierarchique pour l'inférence

Gautier, E. (2011) : "Hierarchical Bayesian estimation of inequality measures with non-rectangular censored survey data with an application to wealth distribution of the French households". *Annals of Applied Statistics* **5** 1632–1656.

1. $G = \widehat{G}\left(Y_1, \ldots, Y_n\right) + \sqrt{\widehat{V\left(\widehat{G}\right)}\left(Y_1, \ldots, Y_n\right)}\epsilon$ où $\epsilon \sim \mathcal{N}(0,1)$, $\epsilon \perp (Y_1, \ldots, Y_n)|(X_1, \ldots, X_n)$ (sélection sur observables, la non-réponse totale est supposée MAR), $\widehat{G}\left(y_1, \ldots, y_n\right) = \frac{\sum_{k=1}^n (2\hat{r}(k)-1) w_k y_k}{\sum_{k=1}^n w_k \sum_{k=1}^n w_k y_k} - 1$, $\hat{r}(k) = \sum_{j=1}^n w_j \mathbb{1}\left\{y_j \leq y_k\right\}$ et $\widehat{V\left(\widehat{G}\right)}\left(y_1, \ldots, y_n\right)$ est obtenu sur le linéarisé en tenant compte du calage et du plan (exemple : utiliser POULPE) (ref. Deville (1999), Shao (1994))
2. $F_{Y|X,D=1}(\cdot|X)$ (que l'on estime, c.f. plus loin).

Nous fournissons un intervalle $[b, h]$, le plus petit possible, tel que $\widehat{\mathbb{P}}(G \in [b,h] | Y_1 = y_1, \ldots, Y_r = y_r; X_1 = x_1, \ldots, X_n = x_n) \geq 1 - \alpha$.
Cette probabilité est calculée par Monte-Carlo à partir de la distribution empirique en simulant m complétions dans $\widehat{F_{Y|X,D=1}}(\cdot|x_k)$ puis G pour ces m complétions.

Modélisation hierarchique pour l'inférence

Gautier, E. (2011) : "Hierarchical Bayesian estimation of inequality measures with non-rectangular censored survey data with an application to wealth distribution of the French households". *Annals of Applied Statistics* **5** 1632–1656.

1. $G=\widehat{G}\left(Y_1,\dots,Y_n\right)+\sqrt{\widehat{V\left(\widehat{G}\right)}\left(Y_1,\dots,Y_n\right)}\epsilon$ où $\epsilon\sim\mathcal{N}(0,1)$, $\epsilon\perp(Y_1,\dots,Y_n)|(X_1,\dots,X_n)$ (sélection sur observables, la non-réponse totale est supposée MAR), $\widehat{G}\left(y_1,\dots,y_n\right)=\frac{\sum_{k=1}^n(2\hat{r}(k)-1)w_ky_k}{\sum_{k=1}^n w_k\sum_{k=1}^n w_ky_k}-1$, $\hat{r}(k)=\sum_{j=1}^n w_j\mathbb{1}\left\{y_j\le y_k\right\}$ et $\widehat{V\left(\widehat{G}\right)}\left(y_1,\dots,y_n\right)$ est obtenu sur le linéarisé en tenant compte du calage et du plan (exemple : utiliser POULPE) (ref. Deville (1999), Shao (1994))
2. $F_{Y|X,D=1}(\cdot|X)$ (que l'on estime, c.f. plus loin).

Nous fournissons un intervalle $[b,h]$, le plus petit possible, tel que $\widehat{\mathbb{P}}(G\in[b,h]|Y_1=y_1,\dots,Y_r=y_r;X_1=x_1,\dots,X_n=x_n)\ge 1-\alpha$.
Cette probabilité est calculée par Monte-Carlo à partir de la distribution empirique en simulant m complétions dans $\widehat{F_{Y|X,D=1}}(\cdot|x_k)$ puis G pour ces m complétions.

Modélisation hierarchique pour l'inférence

Gautier, E. (2011) : "Hierarchical Bayesian estimation of inequality measures with non-rectangular censored survey data with an application to wealth distribution of the French households". *Annals of Applied Statistics* **5** 1632–1656.

1. $G = \widehat{G}\left(Y_1, \ldots, Y_n\right) + \sqrt{\widehat{V\left(\widehat{G}\right)}\left(Y_1, \ldots, Y_n\right)}\epsilon$ où $\epsilon \sim \mathcal{N}(0,1)$, $\epsilon \perp (Y_1, \ldots, Y_n) | (X_1, \ldots, X_n)$ (sélection sur observables, la non-réponse totale est supposée MAR), $\widehat{G}\left(y_1, \ldots, y_n\right) = \frac{\sum_{k=1}^{n}(2\hat{r}(k)-1)w_k y_k}{\sum_{k=1}^{n} w_k \sum_{k=1}^{n} w_k y_k} - 1$, $\hat{r}(k) = \sum_{j=1}^{n} w_j \mathbb{1}\left\{y_j \leq y_k\right\}$ et $\widehat{V\left(\widehat{G}\right)}\left(y_1, \ldots, y_n\right)$ est obtenu sur le linéarisé en tenant compte du calage et du plan (exemple : utiliser POULPE) (ref. Deville (1999), Shao (1994))
2. $F_{Y|X,D=1}(\cdot|X)$ (que l'on estime, c.f. plus loin).

Nous fournissons un intervalle $[b, h]$, le plus petit possible, tel que $\widehat{\mathbb{P}}(G \in [b, h] | Y_1 = y_1, \ldots, Y_r = y_r; X_1 = x_1, \ldots, X_n = x_n) \geq 1 - \alpha$.
Cette probabilité est calculée par Monte-Carlo à partir de la distribution empirique en simulant m complétions dans $\widehat{F_{Y|X,D=1}}(\cdot|x_k)$ puis G pour ces m complétions.

Modélisation hierarchique pour l'inférence

Gautier, E. (2011) : "Hierarchical Bayesian estimation of inequality measures with non-rectangular censored survey data with an application to wealth distribution of the French households". *Annals of Applied Statistics* **5** 1632–1656.

1. $G = \widehat{G}(Y_1, \ldots, Y_n) + \sqrt{\widehat{V\left(\widehat{G}\right)}(Y_1, \ldots, Y_n)}\epsilon$ où $\epsilon \sim \mathcal{N}(0,1)$, $\epsilon \perp (Y_1, \ldots, Y_n) | (X_1, \ldots, X_n)$ (sélection sur observables, la non-réponse totale est supposée MAR), $\widehat{G}(y_1, \ldots, y_n) = \frac{\sum_{k=1}^n (2\hat{r}(k)-1) w_k y_k}{\sum_{k=1}^n w_k \sum_{k=1}^n w_k y_k} - 1$, $\hat{r}(k) = \sum_{j=1}^n w_j \mathbb{1}\{y_j \leq y_k\}$ et $\widehat{V\left(\widehat{G}\right)}(y_1, \ldots, y_n)$ est obtenu sur le linéarisé en tenant compte du calage et du plan (exemple : utiliser POULPE) (ref. Deville (1999), Shao (1994))
2. $F_{Y|X,D=1}(\cdot|X)$ (que l'on estime, c.f. plus loin).

Nous fournissons un intervalle $[b, h]$, le plus petit possible, tel que $\widehat{\mathbb{P}}(G \in [b, h] | Y_1 = y_1, \ldots, Y_r = y_r; X_1 = x_1, \ldots, X_n = x_n) \geq 1 - \alpha$.
Cette probabilité est calculée par Monte-Carlo à partir de la distribution empirique en simulant m complétions dans $\widehat{F_{Y|X,D=1}}(\cdot|x_k)$ puis G pour ces m complétions.

Non-réponse partielle

- La non-réponse est MAR (Rubin) si $\exists X$, toujours observé lorsque $D = 1$, tel que $\forall \phi \in C_b(\mathbb{R})$,

$$\mathbb{E}[\phi(Y)B|X, D = 1] = \mathbb{E}[\phi(Y)|X, D = 1]\mathbb{E}[B|X, D = 1]$$

ou de manière equivalente si

$$\mathbb{E}[\phi(Y)|X, B = 1, D = 1] = \mathbb{E}[\phi(Y)|X, D = 1].$$

- Nous nous intéressons à des situations non-MAR. Dans ce cas, nous ne pouvons pas faire d'inférence sans modéliser explicitement le mécanisme de non-réponse : on dit que la non-réponse est "non-ignorable".
- Un modèle qui semble flexible : $B = \mathbb{1}\{g(Z) > \epsilon\}$ et $Z \perp (\epsilon, Y)|X, D = 1$ où g et la loi de ϵ sont nonparamétriques. Ce modèle est équivalent à $B = \mathbb{1}\{F(Z) > U\}$ où $U|X = x, D = 1 \sim \mathcal{U}(0, 1)$ et $Z \perp (U, Y)|X, D = 1$.
- Les variables constituant le vecteur d'instruments Z doivent être liées à la non-réponse mais pas à Y, les variables X peuvent être introduites pour justifier l'utilisation des instruments. Par exemple l'heure de passage de l'enquêteur, l'identité de l'enquêteur (par exemple s'il n'y en avait que 2), etc. La valeur de l'instrument doit varier pour les unités de l'échantillon.

Non-réponse partielle

- La non-réponse est MAR (Rubin) si $\exists X$, toujours observé lorsque $D = 1$, tel que $\forall \phi \in C_b(\mathbb{R})$,

$$\mathbb{E}[\phi(Y)B|X, D = 1] = \mathbb{E}[\phi(Y)|X, D = 1]\mathbb{E}[B|X, D = 1]$$

ou de manière equivalente si

$$\mathbb{E}[\phi(Y)|X, B = 1, D = 1] = \mathbb{E}[\phi(Y)|X, D = 1].$$

- Nous nous intéressons à des situations non-MAR. Dans ce cas, nous ne pouvons pas faire d'inférence sans modéliser explicitement le mécanisme de non-réponse : on dit que la non-réponse est "non-ignorable".
- Un modèle qui semble flexible : $B = \mathbf{1}\{g(Z) > \epsilon\}$ et $Z \perp (\epsilon, Y)|X, D = 1$ où g et la loi de ϵ sont nonparamétriques. Ce modèle est équivalent à $B = \mathbf{1}\{F(Z) > U\}$ où $U|X = x, D = 1 \sim \mathcal{U}(0, 1)$ et $Z \perp (U, Y)|X, D = 1$.
- Les variables constituant le vecteur d'instruments Z doivent être liées à la non-réponse mais pas à Y, les variables X peuvent être introduites pour justifier l'utilisation des instruments. Par exemple l'heure de passage de l'enquêteur, l'identité de l'enquêteur (par exemple s'il n'y en avait que 2), etc. La valeur de l'instrument doit varier pour les unités de l'échantillon.

Non-réponse partielle

- La non-réponse est MAR (Rubin) si $\exists X$, toujours observé lorsque $D=1$, tel que $\forall \phi \in C_b(\mathbb{R})$,
$$\mathbb{E}[\phi(Y)B|X,D=1]=\mathbb{E}[\phi(Y)|X,D=1]\mathbb{E}[B|X,D=1]$$
ou de manière equivalente si
$$\mathbb{E}[\phi(Y)|X,B=1,D=1]=\mathbb{E}[\phi(Y)|X,D=1].$$
- Nous nous intéressons à des situations non-MAR. Dans ce cas, nous ne pouvons pas faire d'inférence sans modéliser explicitement le mécanisme de non-réponse : on dit que la non-réponse est "non-ignorable".
- Un modèle qui semble flexible : $B=\mathbf{1}\{g(Z)>\epsilon\}$ et $Z\perp(\epsilon,Y)|X,D=1$ où g et la loi de ϵ sont nonparamétriques. Ce modèle est équivalent à $B=\mathbf{1}\{F(Z)>U\}$ où $U|X=x,D=1\sim\mathcal{U}(0,1)$ et $Z\perp(U,Y)|X,D=1$.
- Les variables constituant le vecteur d'instruments Z doivent être liées à la non-réponse mais pas à Y, les variables X peuvent être introduites pour justifier l'utilisation des instruments. Par exemple l'heure de passage de l'enquêteur, l'identité de l'enquêteur (par exemple s'il n'y en avait que 2), etc. La valeur de l'instrument doit varier pour les unités de l'échantillon.

Non-réponse partielle

- La non-réponse est MAR (Rubin) si $\exists X$, toujours observé lorsque $D = 1$, tel que $\forall \phi \in C_b(\mathbb{R})$,
$$\mathbb{E}[\phi(Y)B|X, D = 1] = \mathbb{E}[\phi(Y)|X, D = 1]\mathbb{E}[B|X, D = 1]$$
ou de manière equivalente si
$$\mathbb{E}[\phi(Y)|X, B = 1, D = 1] = \mathbb{E}[\phi(Y)|X, D = 1].$$
- Nous nous intéressons à des situations non-MAR. Dans ce cas, nous ne pouvons pas faire d'inférence sans modéliser explicitement le mécanisme de non-réponse : on dit que la non-réponse est "non-ignorable".
- Un modèle qui semble flexible : $B = \mathbf{1}\{g(Z) > \epsilon\}$ et $Z \perp (\epsilon, Y)|X, D = 1$ où g et la loi de ϵ sont nonparamétriques. Ce modèle est équivalent à $B = \mathbf{1}\{F(Z) > U\}$ où $U|X = x, D = 1 \sim \mathcal{U}(0, 1)$ et $Z \perp (U, Y)|X, D = 1$.
- Les variables constituant le vecteur d'instruments Z doivent être liées à la non-réponse mais pas à Y, les variables X peuvent être introduites pour justifier l'utilisation des instruments. Par exemple l'heure de passage de l'enquêteur, l'identité de l'enquêteur (par exemple s'il n'y en avait que 2), etc. La valeur de l'instrument doit varier pour les unités de l'échantillon.

Monotonie

- Cas non-MAR si $U \not\perp Y|X$, i.e. le paramètre d'hétérogénéité inobservée U (des variables manquantes dans le modèle de non-réponse, etc.) est responsable de la sélection endogène.
- Ce modèle est en fait extrêmement restrictif !
- Il est équivalent (Vytlacil 02) à l'hypothèse de monotonie introduite par Imbens & Angrist 94 : $\forall z, z' \in supp(Z)$, si on change la valeur des instruments pour tous de z à z'
  ♠ $\forall u \in [0,1]$, $\mathbf{1}\{F(z) > u\} \geq \mathbf{1}\{F(z') > u\}$ (si $F(z) \geq F(z')$)
  ♠ ou $\forall u \in [0,1]$, $\mathbf{1}\{F(z) > u\} < \mathbf{1}\{F(z') > U\}$ (si $F(z) < F(z')$).
- Dans le cas où il n'y a que deux enquêteurs A et B et que l'on utilise l'identitié de l'enquêteur comme instrument, la monotonie signifie que tout individu dévoilant son Y à A le dévoilerait aussi à B (ou le contraire). Exemple de raison d'échec : un individu particulier trouve l'enquêteur A plus sympathique que B (ex. car il lui ressemble) et accepte de répondre. Dans ce cas une caractéristique supplémentaire intervient dans le choix de répondre ou non. Celle-ci n'est pas observée par le statisticien $\Rightarrow$ important d'introduire plusieurs sources d'hétérogénéité dans l'équation de sélection.

Monotonie

- Cas non-MAR si $U \not\perp Y|X$, i.e. le paramètre d'hétérogénéité inobservée U (des variables manquantes dans le modèle de non-réponse, etc.) est responsable de la sélection endogène.
- Ce modèle est en fait extrêmement restrictif !
- Il est équivalent (Vytlacil 02) à l'hypothèse de monotonie introduite par Imbens & Angrist 94 : $\forall z, z' \in supp(Z)$, si on change la valeur des instruments pour tous de z à z'

  ♠ $\forall u \in [0,1]$, $\mathbf{1}\{F(z) > u\} \geq \mathbf{1}\{F(z') > u\}$ (si $F(z) \geq F(z')$)

  ♠ ou $\forall u \in [0,1]$, $\mathbf{1}\{F(z) > u\} < \mathbf{1}\{F(z') > U\}$ (si $F(z) < F(z')$).
- Dans le cas où il n'y a que deux enquêteurs A et B et que l'on utilise l'identitié de l'enquêteur comme instrument, la monotonie signifie que tout individu dévoilant son Y à A le dévoilerait aussi à B (ou le contraire). Exemple de raison d'échec : un individu particulier trouve l'enquêteur A plus sympathique que B (ex. car il lui ressemble) et accepte de répondre. Dans ce cas une caractéristique supplémentaire intervient dans le choix de répondre ou non. Celle-ci n'est pas observée par le statisticien $\Rightarrow$ important d'introduire plusieurs sources d'hétérogénéité dans l'équation de sélection.

Monotonie

- Cas non-MAR si $U \not\perp Y|X$, i.e. le paramètre d'hétérogénéité inobservée U (des variables manquantes dans le modèle de non-réponse, etc.) est responsable de la sélection endogène.
- Ce modèle est en fait extrêmement restrictif !
- Il est équivalent (Vytlacil 02) à l'hypothèse de monotonie introduite par Imbens & Angrist 94 : $\forall z, z' \in supp(Z)$, si on change la valeur des instruments pour tous de z à z'
  ♠ $\forall u \in [0,1]$, $\mathbf{1}\{F(z) > u\} \geq \mathbf{1}\{F(z') > u\}$ (si $F(z) \geq F(z')$)
  ♠ ou $\forall u \in [0,1]$, $\mathbf{1}\{F(z) > u\} < \mathbf{1}\{F(z') > U\}$ (si $F(z) < F(z')$).
- Dans le cas où il n'y a que deux enquêteurs A et B et que l'on utilise l'identitié de l'enquêteur comme instrument, la monotonie signifie que tout individu dévoilant son Y à A le dévoilerait aussi à B (ou le contraire). Exemple de raison d'échec : un individu particulier trouve l'enquêteur A plus sympathique que B (ex. car il lui ressemble) et accepte de répondre. Dans ce cas une caractéristique supplémentaire intervient dans le choix de répondre ou non. Celle-ci n'est pas observée par le statisticien ⇒ important d'introduire plusieurs sources d'hétérogénéité dans l'équation de sélection.

Monotonie

- Cas non-MAR si $U \not\perp Y|X$, i.e. le paramètre d'hétérogénéité inobservée U (des variables manquantes dans le modèle de non-réponse, etc.) est responsable de la sélection endogène.
- Ce modèle est en fait extrêmement restrictif !
- Il est équivalent (Vytlacil 02) à l'hypothèse de monotonie introduite par Imbens & Angrist 94 : $\forall z, z' \in supp(Z)$, si on change la valeur des instruments pour tous de z à z'
  ♠ $\forall u \in [0,1]$, $\mathbf{1}\{F(z) > u\} \geq \mathbf{1}\{F(z') > u\}$ (si $F(z) \geq F(z')$)
  ♠ ou $\forall u \in [0,1]$, $\mathbf{1}\{F(z) > u\} < \mathbf{1}\{F(z') > U\}$ (si $F(z) < F(z')$).
- Dans le cas où il n'y a que deux enquêteurs A et B et que l'on utilise l'identitié de l'enquêteur comme instrument, la monotonie signifie que tout individu dévoilant son Y à A le dévoilerait aussi à B (ou le contraire). Exemple de raison d'échec : un individu particulier trouve l'enquêteur A plus sympathique que B (ex. car il lui ressemble) et accepte de répondre. Dans ce cas une caractéristique supplémentaire intervient dans le choix de répondre ou non. Celle-ci n'est pas observée par le statisticien $\Rightarrow$ important d'introduire plusieurs sources d'hétérogénéité dans l'équation de sélection.

Le modèle à choix binaire et coefficients aléatoires

Gautier, E., et Y. Kitamura (2008) : "Nonparametric estimation in random coefficients binary choice models". A paraître dans *Econometrica*.

Gautier, E., et S. Hooderlein (2011) : "Estimating treatment effects with a nonparametric random coefficients selection equation". Preprint (v1) arXiv :1109.0362.

Gautier, E., et E. Le Pennec (2011) : "Adaptive estimation in random coefficients binary choice models using needlet thresholding". Preprint arXiv :1106.3503.

- Heckman & Vytlacil 05, le modèle "benchmark" non additivement séparable et avec plusieurs sources d'hétérogénéité inobservé est un modèle à choix binaire et coefficients aléatoires :

$$B = \mathbf{1}\{Z_1\overline{\Gamma}_1 + Z_2{}^T\overline{\Gamma}_2 + \overline{\Gamma}_0 > 0\}$$

avec Z_1 scalaire et Z_2 de taille $L-1$ où $L \geq 2$. Supposons un modèle nonparamétrique pour la loi de $(\Gamma_1, \Gamma_2^T, \Gamma_0)$ et $\overline{\Gamma}_1 > 0$ p.s.
Dans ce modèle, chaque individu a son propre vecteur de coefficients (préférences).

- Ré-écrivons : $B = \mathbf{1}\{Z_1 - Z_2{}^T\Gamma - \Theta > 0\}$ puis, en posant $\widetilde{S} = (Z_2^T, 1)^T/\|(Z_2^T, 1)\|$, $\widetilde{V} = Z_1/\|(Z_2^T, 1)\|$, et $\widetilde{\Gamma} = (\Gamma^T, \Theta)^T$, $B = \mathbf{1}\{\widetilde{S}^T\widetilde{\Gamma} < \widetilde{V}\}$.

Le modèle à choix binaire et coefficients aléatoires

Gautier, E., et Y. Kitamura (2008) : "Nonparametric estimation in random coefficients binary choice models". A paraître dans *Econometrica*.

Gautier, E., et S. Hooderlein (2011) : "Estimating treatment effects with a nonparametric random coefficients selection equation". Preprint (v1) arXiv :1109.0362.

Gautier, E., et E. Le Pennec (2011) : "Adaptive estimation in random coefficients binary choice models using needlet thresholding". Preprint arXiv :1106.3503.

- Heckman & Vytlacil 05, le modèle "benchmark" non additivement séparable et avec plusieurs sources d'hétérogénéité inobservé est un modèle à choix binaire et coefficients aléatoires :

$$B = \mathbf{1}\{Z_1\overline{\Gamma}_1 + {Z_2}^T\overline{\Gamma}_2 + \overline{\Gamma}_0 > 0\}$$

  avec Z_1 scalaire et Z_2 de taille $L-1$ où $L \geq 2$. Supposons un modèle nonparamétrique pour la loi de $(\overline{\Gamma}_1, \overline{\Gamma}_2^T, \overline{\Gamma}_0)$ et $\overline{\Gamma}_1 > 0$ p.s.
  Dans ce modèle, chaque individu a son propre vecteur de coefficients (préférences).

- Ré-écrivons : $B = \mathbf{1}\{Z_1 - {Z_2}^T\Gamma - \Theta > 0\}$ puis, en posant $\widetilde{S} = (Z_2^T, 1)^T/\|(Z_2^T, 1)\|$, $\widetilde{V} = Z_1/\|(Z_2^T, 1)\|$, et $\widetilde{\Gamma} = (\Gamma^T, \Theta)^T$, $B = \mathbf{1}\{\widetilde{S}^T\widetilde{\Gamma} < \widetilde{V}\}$.

Le modèle à choix binaire et coefficients aléatoires

Gautier, E., et Y. Kitamura (2008) : "Nonparametric estimation in random coefficients binary choice models". A paraître dans *Econometrica*.

Gautier, E., et S. Hooderlein (2011) : "Estimating treatment effects with a nonparametric random coefficients selection equation". Preprint (v1) arXiv :1109.0362.

Gautier, E., et E. Le Pennec (2011) : "Adaptive estimation in random coefficients binary choice models using needlet thresholding". Preprint arXiv :1106.3503.

- Heckman & Vytlacil 05, le modèle "benchmark" non additivement séparable et avec plusieurs sources d'hétérogénéité inobservé est un modèle à choix binaire et coefficients aléatoires :

$$B = \mathbf{1}\{Z_1\overline{\Gamma}_1 + {Z_2}^T\overline{\Gamma}_2 + \overline{\Gamma}_0 > 0\}$$

  avec Z_1 scalaire et Z_2 de taille $L-1$ où $L \geq 2$. Supposons un modèle nonparamétrique pour la loi de $(\Gamma_1, \Gamma_2^T, \Gamma_0)$ et $\overline{\Gamma}_1 > 0$ p.s.
  Dans ce modèle, chaque individu a son propre vecteur de coefficients (préférences).
- Ré-écrivons : $B = \mathbf{1}\{Z_1 - {Z_2}^T\Gamma - \Theta > 0\}$ puis, en posant $\widetilde{S} = (Z_2^T, 1)^T/\|(Z_2^T, 1)\|$, $\widetilde{V} = Z_1/\|(Z_2^T, 1)\|$, et $\widetilde{\Gamma} = (\Gamma^T, \Theta)^T$, $B = \mathbf{1}\{\widetilde{S}^T\widetilde{\Gamma} < \widetilde{V}\}$.

Le modèle à choix binaire et coefficients aléatoires

Gautier, E., et Y. Kitamura (2008) : "Nonparametric estimation in random coefficients binary choice models". A paraître dans *Econometrica*.

Gautier, E., et S. Hooderlein (2011) : "Estimating treatment effects with a nonparametric random coefficients selection equation". Preprint (v1) arXiv :1109.0362.

Gautier, E., et E. Le Pennec (2011) : "Adaptive estimation in random coefficients binary choice models using needlet thresholding". Preprint arXiv :1106.3503.

- Heckman & Vytlacil 05, le modèle "benchmark" non additivement séparable et avec plusieurs sources d'hétérogénéité inobservé est un modèle à choix binaire et coefficients aléatoires :

$$B = \mathbf{1}\{Z_1\overline{\Gamma}_1 + {Z_2}^T\overline{\Gamma}_2 + \overline{\Gamma}_0 > 0\}$$

  avec Z_1 scalaire et Z_2 de taille $L-1$ où $L \geq 2$. Supposons un modèle nonparamétrique pour la loi de $(\Gamma_1, \Gamma_2^T, \Gamma_0)$ et $\overline{\Gamma}_1 > 0$ p.s.
  Dans ce modèle, chaque individu a son propre vecteur de coefficients (préférences).

- Ré-écrivons : $B = \mathbf{1}\{Z_1 - {Z_2}^T\Gamma - \Theta > 0\}$ puis, en posant $\widetilde{S} = (Z_2^T, 1)^T/\|(Z_2^T, 1)\|$, $\widetilde{V} = Z_1/\|(Z_2^T, 1)\|$, et $\widetilde{\Gamma} = (\Gamma^T, \Theta)^T$,
  $B = \mathbf{1}\{\widetilde{S}^T\widetilde{\Gamma} < \widetilde{V}\}$.

Le modèle à choix binaire et coefficients aléatoires

Gautier, E., et Y. Kitamura (2008) : "Nonparametric estimation in random coefficients binary choice models". A paraître dans *Econometrica*.

Gautier, E., et S. Hooderlein (2011) : "Estimating treatment effects with a nonparametric random coefficients selection equation". Preprint (v1) arXiv :1109.0362.

Gautier, E., et E. Le Pennec (2011) : "Adaptive estimation in random coefficients binary choice models using needlet thresholding". Preprint arXiv :1106.3503.

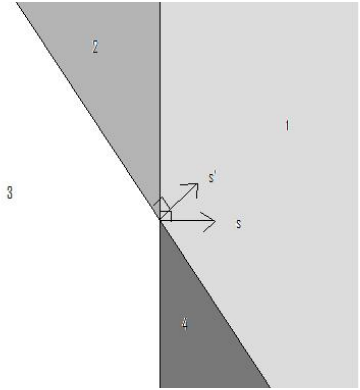
- Heckman & Vytlacil 05, le modèle "benchmark" non additivement séparable et avec plusieurs sources d'hétérogénéité inobservé est un modèle à choix binaire et coefficients aléatoires :

$$B = \mathbf{1}\{Z_1\overline{\Gamma}_1 + {Z_2}^T\overline{\Gamma}_2 + \overline{\Gamma}_0 > 0\}$$

  avec Z_1 scalaire et Z_2 de taille $L-1$ où $L \geq 2$. Supposons un modèle nonparamétrique pour la loi de $(\Gamma_1, \Gamma_2^T, \Gamma_0)$ et $\overline{\Gamma}_1 > 0$ p.s.
  Dans ce modèle, chaque individu a son propre vecteur de coefficients (préférences).
- Ré-écrivons : $B = \mathbf{1}\{Z_1 - {Z_2}^T\Gamma - \Theta > 0\}$ puis, en posant $\widetilde{S} = (Z_2^T, 1)^T/\|(Z_2^T, 1)\|$, $\widetilde{V} = Z_1/\|(Z_2^T, 1)\|$, et $\widetilde{\Gamma} = (\Gamma^T, \Theta)^T$,
  $B = \mathbf{1}\{\widetilde{S}^T\widetilde{\Gamma} < \widetilde{V}\}$.

Non-monotonie

Le modèle permet des situations non-monotones dans le mécanisme de non-réponse. Fixons $v \in supp(\widetilde{V})$, s et s' dans $supp\left(f_{\widetilde{S}|\widetilde{V}}(\cdot|v)\right)$, et posons $D_s(\gamma) = \mathbf{1}\{s^T\gamma < v\}$
Dans la zone 2 $D_s = 0$ et $D_{s'} = 1$, dans la 4 $D_s = 1$ et $D_{s'} = 0$



Hypothèses

(H-1) La loi conditionnelle de $(Z_1, Z_2, \Gamma^T, \Theta)$ sachant $X = x, D = 1$ est absolument continue par rapport à la mesure de Lebesgue $\forall x \in supp\left(f_{X|D=1}\right)$. Z_1 et Y peuvent en fait être discrets.
$\Rightarrow$ *restrictions d'exclusion*

(H-2) $(Z_1, Z_2) \perp (Y, \Gamma^T, \Theta) | X, D = 1$.

(H-3) $0 < \mathbb{P}(B = 1 | X, D = 1) < 1 \quad p.s.$

(H-4) $\forall x \in supp\left(f_{X|D=1}\right)$, $supp\left(f_{\widetilde{S}|X,D=1}(\cdot|x)\right) = H^+$ et
$\forall s \in \mathbb{S}^L : \; s^T(0, \dots, 0, 1) \geq 0$, $supp\left(f_{\widetilde{V}|\widetilde{S},X,D=1}(\cdot|s,x)\right) \supset$
$\left[\inf_{\gamma \in supp\left(f_{\widetilde{\Gamma}|X,D=1}(\cdot|x)\right)} s^T\gamma, \sup_{\gamma \in supp\left(f_{\widetilde{\Gamma}|X,D=1}(\cdot|x)\right)} s^T\gamma\right]$.

HYPOTHÈSES

(H-1) La loi conditionnelle de $(Z_1, Z_2, \Gamma^T, \Theta)$ sachant $X = x, D = 1$ est absolument continue par rapport à la mesure de Lebesgue $\forall x \in supp\left(f_{X|D=1}\right)$. Z_1 et Y peuvent en fait être discrets.
$\Rightarrow$ *restrictions d'exclusion*

(H-2) $(Z_1, Z_2) \perp (Y, \Gamma^T, \Theta) | X, D = 1$.

(H-3) $0 < \mathbb{P}(B = 1 | X, D = 1) < 1 \quad p.s.$

(H-4) $\forall x \in supp\left(f_{X|D=1}\right)$, $supp\left(f_{\widetilde{S}|X,D=1}(\cdot|x)\right) = H^+$ et
$\forall s \in \mathbb{S}^L : \ s^T(0,\ldots,0,1) \geq 0$, $supp\left(f_{\widetilde{V}|\widetilde{S},X,D=1}(\cdot|s,x)\right) \supset$
$\left[\inf_{\gamma \in supp\left(f_{\widetilde{\Gamma}|X,D=1}(\cdot|x)\right)} s^T\gamma, \sup_{\gamma \in supp\left(f_{\widetilde{\Gamma}|X,D=1}(\cdot|x)\right)} s^T\gamma\right]$.

Hypothèses

(H-1) La loi conditionnelle de $(Z_1, Z_2, \Gamma^T, \Theta)$ sachant $X = x, D = 1$ est absolument continue par rapport à la mesure de Lebesgue $\forall x \in supp\left(f_{X|D=1}\right)$. Z_1 et Y peuvent en fait être discrets.
$\Rightarrow$ *restrictions d'exclusion*

(H-2) $(Z_1, Z_2) \perp (Y, \Gamma^T, \Theta) | X, D = 1$.

(H-3) $0 < \mathbb{P}(B = 1 | X, D = 1) < 1 \quad p.s.$

(H-4) $\forall x \in supp\left(f_{X|D=1}\right)$, $supp\left(f_{\widetilde{S}|X,D=1}(\cdot|x)\right) = H^+$ et
$\forall s \in \mathbb{S}^L: \ s^T(0, \ldots, 0, 1) \geq 0$, $supp\left(f_{\widetilde{V}|\widetilde{S},X,D=1}(\cdot|s, x)\right) \supset$
$\left[\inf_{\gamma \in supp\left(f_{\widetilde{\Gamma}|X,D=1}(\cdot|x)\right)} s^T\gamma, \sup_{\gamma \in supp\left(f_{\widetilde{\Gamma}|X,D=1}(\cdot|x)\right)} s^T\gamma\right]$.

Hypothèses

(H-1) La loi conditionnelle de $(Z_1, Z_2, \Gamma^T, \Theta)$ sachant $X = x, D = 1$ est absolument continue par rapport à la mesure de Lebesgue $\forall x \in supp\left(f_{X|D=1}\right)$. Z_1 et Y peuvent en fait être discrets.
$\Rightarrow$ *restrictions d'exclusion*

(H-2) $(Z_1, Z_2) \perp (Y, \Gamma^T, \Theta) | X, D = 1$.

(H-3) $0 < \mathbb{P}(B = 1 | X, D = 1) < 1 \quad p.s.$

(H-4) $\forall x \in supp\left(f_{X|D=1}\right)$, $supp\left(f_{\widetilde{S}|X,D=1}(\cdot|x)\right) = H^+$ et
$\forall s \in \mathbb{S}^L : \; s^T(0, \dots, 0, 1) \geq 0$, $supp\left(f_{\widetilde{V}|\widetilde{S},X,D=1}(\cdot|s,x)\right) \supset \left[\inf_{\gamma \in supp\left(f_{\widetilde{\Gamma}|X,D=1}(\cdot|x)\right)} s^T\gamma, \sup_{\gamma \in supp\left(f_{\widetilde{\Gamma}|X,D=1}(\cdot|x)\right)} s^T\gamma\right]$.

Identification

Notons $Y_0 = YB$, dans ce cas la non-réponse correspond à un 0 (on suppose que $\mathbb{P}(Y = 0|D = 1, X) = 0$ p.s.) et $\mathbb{E}_D$ l'espérance conditionnelle sachant $D = 1$.

Theorem

$\forall \phi : \ \mathbb{E}_D[|\phi(Y)|] < \infty$, *p.p.* $x \in supp(X)$

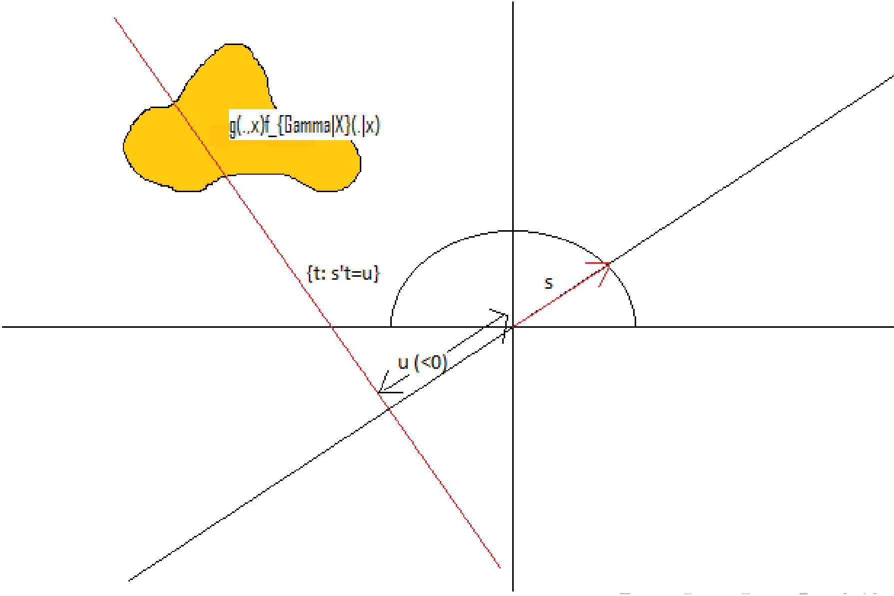
$$f_{\widetilde{\Gamma}|X,D=1}(\cdot|x) = R^{-1}\left[\overline{\partial_v \mathbb{E}_D\left[B\left|\left(\widetilde{S},\widetilde{V}\right) = \cdot, X = x\right.\right]}\right]$$

$$\overline{\mathbb{E}_D\left[\phi(Y)\left|\widetilde{\Gamma} = \cdot, X = x\right.\right] f_{\widetilde{\Gamma}|X,D=1}(\cdot|x)} = R^{-1}\left[\overline{\partial_v \mathbb{E}_D\left[\phi(Y_0)B\left|\left(\widetilde{S},\widetilde{V}\right) = \cdot, X = x\right.\right]}\right].$$

$\overline{g}$ = extension de g qui vaut 0 en dehors du domaine de définition
$\Rightarrow$ (ex)

$$F_{Y|X,D=1}(y|x) = \int_{\mathbb{R}^L} R^{-1}\left[\overline{\partial_v \mathbb{E}_D\left[\mathbf{1}\left\{Y_0 \leq y\right\} B\left|\left(\widetilde{S},\widetilde{V}\right) = \cdot, X = x\right.\right]}\right](\gamma)d\gamma$$

$\Rightarrow$ (ex) mesure d'inégalité, quantiles, etc.


g(.,x)f_{Gamma|X}(.|x)
{t: s't=u}
s
u (<0)

Estimateur

- $A_T[f](\gamma) := \int_{s\in\mathbb{S}^L:\ s^T(0,\dots,0,1)\geq 0}\int_{-\infty}^{\infty} K_T(s^T\gamma - u)f(s,u)dud\sigma(s) =$ inverse régularisée de R, $K_T(u) := 2(2\pi)^{-L}\int_0^T \cos(tu)t^{L-1}\Psi(t/T)dt$ où $\Psi \in \mathcal{S}(\mathbb{R})$ est symmetrique et $\Psi(0) = 1$ (*ex.* $\psi = \psi_0$ où $\psi_0 :\ x \mapsto \exp\left(1 - \max\left\{\frac{1}{1-x^2}, 0\right\}\right)$).
- Estimateur : $A_{T_n}\left[\overline{\widehat{\partial_v\mathbb{E}_D[\phi(Y_0)B|(\widetilde{S},\widetilde{V}) = \cdot, X = x]}}\right](\gamma)$, l'estimateur de la dérivée de la fonction de régression est obtenu par polynômes locaux, etc.
- $F_{Y|X,D=1}(y|x) =$ $\int_{\mathbb{R}^L} A_{T_n}\left[\overline{\widehat{\partial_v\mathbb{E}_D[\mathbf{1}\{Y_0 \leq y\}B|(\widetilde{S},\widetilde{V}) = \cdot, X = x]}}\right](\gamma)\mathbf{1}\{\gamma \in \mathcal{B}_n\}d\gamma$ où $\mathcal{B}_n$ est un fermé borné de $\mathbb{R}^L$.

Estimateur

- $A_T[f](\gamma) := \int_{s\in\mathbb{S}^L:\ s^T(0,\dots,0,1)\geq 0} \int_{-\infty}^{\infty} K_T(s^T\gamma - u)f(s,u)dud\sigma(s) =$ inverse régularisée de R, $K_T(u) := 2(2\pi)^{-L}\int_0^T \cos(tu)t^{L-1}\Psi(t/T)dt$ où $\Psi\in\mathcal{S}(\mathbb{R})$ est symmetrique et $\Psi(0)=1$ (*ex.* $\psi=\psi_0$ où $\psi_0:\ x\mapsto \exp\left(1-\max\left\{\frac{1}{1-x^2},0\right\}\right)$).
- Estimateur : $A_{T_n}\left[\overline{\partial_v\mathbb{E}_D[\phi(Y_0)B|\widehat{(\widetilde{S},\widetilde{V})}=\cdot,X=x]}\right](\gamma)$, l'estimateur de la dérivée de la fonction de régression est obtenu par polynômes locaux, etc.
- $F_{Y|X,D=1}(y|x) = \int_{\mathbb{R}^L} A_{T_n}\left[\overline{\partial_v\mathbb{E}_D[\mathbf{1}\{Y_0\leq y\}B|\widehat{(\widetilde{S},\widetilde{V})}=\cdot,X=x]}\right](\gamma)\mathbf{1}\{\gamma\in\mathcal{B}_n\}d\gamma$ où $\mathcal{B}_n$ est un fermé borné de $\mathbb{R}^L$.

Estimateur

- $A_T[f](\gamma) := \int_{s\in\mathbb{S}^L:\ s^T(0,\dots,0,1)\geq 0}\int_{-\infty}^{\infty} K_T(s^T\gamma-u)f(s,u)dud\sigma(s) =$ inverse régularisée de R, $K_T(u) := 2(2\pi)^{-L}\int_0^T \cos(tu)t^{L-1}\Psi(t/T)dt$ où $\Psi\in\mathcal{S}(\mathbb{R})$ est symmetrique et $\Psi(0)=1$ (*ex.* $\psi=\psi_0$ où $\psi_0:\ x\mapsto \exp\left(1-\max\left\{\frac{1}{1-x^2},0\right\}\right)$).
- Estimateur : $A_{T_n}\left[\overline{\partial_v \widehat{\mathbb{E}_D[\phi(Y_0)B|(\widetilde{S},\widetilde{V})}=\cdot,X=x]}\right](\gamma)$, l'estimateur de la dérivée de la fonction de régression est obtenu par polynômes locaux, etc.
- $F_{Y|X,D=1}(y|x) =$ $\int_{\mathbb{R}^L} A_{T_n}\left[\overline{\partial_v \widehat{\mathbb{E}_D[\mathbf{1}\left\{Y_0\leq y\right\}B|(\widetilde{S},\widetilde{V})}=\cdot,X=x]}\right](\gamma)\mathbf{1}\{\gamma\in\mathcal{B}_n\}d\gamma$ où $\mathcal{B}_n$ est un fermé borné de $\mathbb{R}^L$.

Estimateur simple

- (H-5) p.p. pour $s \in \mathbb{S}^L$: $s^T(0,\dots,0,1) \geq 0$ et p.s. pour $x \in supp(X)$, $supp\left(f_{\widetilde{V}|\widetilde{S},X,D=1}(\cdot|s,x)\right) = \mathbb{R}$; Pour la fonction ϕ considérée, p.p. en $s \in \mathbb{S}^L$: $s^T(0,\dots,0,1) \geq 0$ et p.s. en $x \in supp(f_{X|D=1})$, $v \mapsto \mathbb{E}_D[\phi(Y_0)B|(\widetilde{S},\widetilde{V}) = (s,v), X = x]$ est continue et $v \mapsto \mathbb{E}_D[\phi(Y_0)B|(\widetilde{S},\widetilde{V}) = (s,v), X = x]$ et $v \mapsto \partial_v \mathbb{E}_D[\phi(Y_0)B|(\widetilde{S},\widetilde{V}) = (s,v), X = x]$ sont bornées par des polynômes en v.
- Sous (H-5) nous pouvons utiliser l'estimateur

$$\frac{1}{n}\sum_{i=1}^{n} \frac{\widetilde{K}_{T_n}(\widetilde{s}_i^T\gamma - \widetilde{v}_i)T_{\tau_n}(\phi(y_{0i}))b_i}{\widehat{f_{\widetilde{S},\widetilde{V},X|D=1}}(\widetilde{s}_i,\widetilde{v}_i,x)} \mathbf{1}\left\{\left|\widehat{f_{\widetilde{S},\widetilde{V},X|D=1}}(\widetilde{s}_i,\widetilde{v}_i,x)\right| > t_n\right\}\mathcal{K}_{\eta_n}(x_i - x)$$

où $\widetilde{K}_T(u) := 2(2\pi)^{-L}\int_0^T \sin(tu)t^L\Psi(t/T)dt$, $\widehat{f_{\widetilde{S},\widetilde{V},X|D=1}}$ est un estimateur plug-in de $f_{\widetilde{S},\widetilde{V},X|D=1}$, $T_\tau(x) = -\tau\mathbf{1}\{x < -\tau\} + x\mathbf{1}\{|x| \leq \tau\} + \tau\mathbf{1}\{x > \tau\}$ et $\mathcal{K}_\eta$ est un noyau multivarié de fenêtre η.

Estimateur simple

- (H-5) p.p. pour $s \in \mathbb{S}^L$: $s^T(0,\dots,0,1) \geq 0$ et p.s. pour $x \in supp(X)$, $supp\left(f_{\widetilde{V}|\widetilde{S},X,D=1}(\cdot|s,x)\right) = \mathbb{R}$; Pour la fonction ϕ considérée, p.p. en $s \in \mathbb{S}^L$: $s^T(0,\dots,0,1) \geq 0$ et p.s. en $x \in supp(f_{X|D=1})$, $v \mapsto \mathbb{E}_D[\phi(Y_0)B|(\widetilde{S},\widetilde{V}) = (s,v), X = x]$ est continue et $v \mapsto \mathbb{E}_D[\phi(Y_0)B|(\widetilde{S},\widetilde{V}) = (s,v), X = x]$ et $v \mapsto \partial_v\mathbb{E}_D[\phi(Y_0)B|(\widetilde{S},\widetilde{V}) = (s,v), X = x]$ sont bornées par des polynômes en v.
- Sous (H-5) nous pouvons utiliser l'estimateur

$$\frac{1}{n}\sum_{i=1}^n \frac{\widetilde{K}_{T_n}(\widetilde{s}_i^T\gamma - \widetilde{v}_i)T_{\tau_n}(\phi(y_{0i}))b_i}{\widehat{f_{\widetilde{S},\widetilde{V},X|D=1}}(\widetilde{s}_i,\widetilde{v}_i,x)}\mathbf{1}\left\{\left|\widehat{f_{\widetilde{S},\widetilde{V},X|D=1}}(\widetilde{s}_i,\widetilde{v}_i,x)\right| > t_n\right\}\mathcal{K}_{\eta_n}(x_i - x)$$

où $\widetilde{K}_T(u) := 2(2\pi)^{-L}\int_0^T \sin(tu)t^L\Psi(t/T)dt$, $\widehat{f_{\widetilde{S},\widetilde{V},X|D=1}}$ est un estimateur plug-in de $f_{\widetilde{S},\widetilde{V},X|D=1}$, $T_\tau(x) = -\tau\mathbf{1}\{x < -\tau\} + x\mathbf{1}\{|x| \leq \tau\} + \tau\mathbf{1}\{x > \tau\}$ et $\mathcal{K}_\eta$ est un noyau multivarié de fenêtre η.

Résultats asymptotiques

- $g(\gamma) = R^{-1}\left[\overline{\partial_v \mathbb{E}_D\left[\phi(Y_0)B \,\middle|\, \left(\widetilde{S}, \widetilde{V}\right) = \cdot\right]}\right](\gamma)$, est estimé par

$$\hat{g}(\gamma) = \frac{1}{n}\sum_{i=1}^{n} \frac{\widetilde{K}_{T_n}(\widetilde{s}_i'\gamma - \widetilde{v}_i)\, T_{\tau_n}(\phi(y_{0i}))b_i}{\max\left(\widehat{f_{\widetilde{S},\widetilde{V}|D=1}}(\widetilde{s}_i, \widetilde{v}_i), m_n\right)}.$$

- $W^{s,\infty}(\mathbb{R}^L) := \left\{f \in \mathrm{L}^\infty(\mathbb{R}^L) :\ \forall|\alpha| \le s,\ \partial^\alpha f \in \mathrm{L}^\infty(\mathbb{R}^L)\right\}$ où $s \in \mathbb{N}\setminus\{0\}$, $\alpha \in \mathbb{N}^L$, $|\alpha| := \sum_{l=1}^{L}\alpha_l$ and $\partial^\alpha f := \prod_{l=1}^{L}\partial_l^{\alpha_l} f$, $\|f\|_{s,\infty} := \sum_{\alpha:\ |\alpha|\le s}\|\partial^\alpha f\|_\infty$.
  Nous considérons les ellipsoides de Sobolev

$$W^{s,\infty}(M) := \left\{f \in W^{s,\infty}(\mathbb{R}^L) :\ \|f\|_{s,\infty} \le M\right\}$$

- $\mathcal{B}_n$ un fermé de $\mathbb{R}^L$ et $d(\mathcal{B}_n)$ son diamètre pour la norme Euclidiène.

Résultats asymptotiques

- $g(\gamma) = R^{-1}\left[\overline{\partial_v \mathbb{E}_D\left[\phi(Y_0)B \middle| \left(\widetilde{S}, \widetilde{V}\right) = \cdot\right]}\right](\gamma)$, est estimé par

$$\hat{g}(\gamma) = \frac{1}{n}\sum_{i=1}^{n} \frac{\widetilde{K}_{T_n}(\widetilde{s}_i'\gamma - \widetilde{v}_i)\,T_{\tau_n}(\phi(y_{0i}))b_i}{\max\left(\widehat{f_{\widetilde{S},\widetilde{V}|D=1}}(\widetilde{s}_i, \widetilde{v}_i), m_n\right)}.$$

- $W^{s,\infty}(\mathbb{R}^L) := \left\{f \in \mathrm{L}^\infty(\mathbb{R}^L) : \ \forall |\alpha| \le s,\ \partial^\alpha f \in \mathrm{L}^\infty(\mathbb{R}^L)\right\}$ où $s \in \mathbb{N}\setminus\{0\}$, $\alpha \in \mathbb{N}^L$, $|\alpha| := \sum_{l=1}^L \alpha_l$ and $\partial^\alpha f := \prod_{l=1}^L \partial_l^{\alpha_l} f$,
  $\|f\|_{s,\infty} := \sum_{\alpha:\ |\alpha|\le s} \|\partial^\alpha f\|_\infty$.
  Nous considérons les ellipsoides de Sobolev

$$W^{s,\infty}(M) := \left\{f \in W^{s,\infty}(\mathbb{R}^L) : \ \|f\|_{s,\infty} \le M\right\}$$

- $\mathcal{B}_n$ un fermé de $\mathbb{R}^L$ et $d(\mathcal{B}_n)$ son diamètre pour la norme Euclidiène.

Résultats asymptotiques

- $g(\gamma) = R^{-1}\left[\overline{\partial_v \mathbb{E}_D\left[\phi(Y_0)B\,\middle|\left(\widetilde{S},\widetilde{V}\right) = \cdot\,\right]}\right](\gamma)$, est estimé par

$$\hat{g}(\gamma) = \frac{1}{n}\sum_{i=1}^{n}\frac{\widetilde{K}_{T_n}(\widetilde{s}_i'\gamma - \widetilde{v}_i)T_{\tau_n}(\phi(y_{0i}))b_i}{\max\left(\widehat{f_{\widetilde{S},\widetilde{V}|D=1}}(\widetilde{s}_i,\widetilde{v}_i), m_n\right)}.$$

- $W^{s,\infty}(\mathbb{R}^L) := \{f \in \mathrm{L}^\infty(\mathbb{R}^L) :\ \forall |\alpha| \le s,\ \partial^\alpha f \in \mathrm{L}^\infty(\mathbb{R}^L)\}$ où $s \in \mathbb{N}\setminus\{0\}$, $\alpha \in \mathbb{N}^L$, $|\alpha| := \sum_{l=1}^L \alpha_l$ and $\partial^\alpha f := \prod_{l=1}^L \partial_l^{\alpha_l} f$,
  $\|f\|_{s,\infty} := \sum_{\alpha:\ |\alpha|\le s} \|\partial^\alpha f\|_\infty$.
  Nous considérons les ellipsoides de Sobolev

$$W^{s,\infty}(M) := \left\{f \in W^{s,\infty}(\mathbb{R}^L) :\ \|f\|_{s,\infty} \le M\right\}$$

- $\mathcal{B}_n$ un fermé de $\mathbb{R}^L$ et $d(\mathcal{B}_n)$ son diamètre pour la norme Euclidiène.

Résultats asymptotiques 2

Proposition

Supposons (H-5) et $g \in \mathrm{W}^s_\infty(M)$, $\exists \alpha > 0 : \log(T_n^3/m_n) + L\log(d(\mathcal{B}_n)) \leq \alpha$, $\exists r_{IV,n} \underset{n\to\infty}{\to} 0$ *et* M_{IV} *tq*

$$\overline{\lim}_{n\to\infty} r_{IV,n}^{-1} \max_{i=1,\dots,n} \left| f_{\widetilde{S},\widetilde{V}|D=1}(\widetilde{s}_i,\widetilde{v}_i) - \widehat{f_{\widetilde{S},\widetilde{V}|D=1}}(\widetilde{s}_i,\widetilde{v}_i) \right| \leq M_{IV} \quad a.s.$$

pour $M(\alpha)$, $C(s)$ *avec proba 1,* $\forall \epsilon > 0$, $\exists N_\epsilon > 0 : \ \forall N > N_\epsilon$

$$\begin{aligned}
\|(\hat{g} - g)\mathbf{1}\{\mathcal{B}_n\}\|_\infty \leq\ & (M_{IV} + \epsilon)\min\left(\tau_n, \|\phi\|_\infty\right) r_{IV,r} m_n^{-1} \left\| \mathbb{E}\left[\frac{\left|\widetilde{K}_{T_n}\left(\widetilde{S}'\gamma - \widetilde{V}\right)\right|}{\max\left(\widehat{f_{\widetilde{S},\widetilde{V}|D=1}}\left(\widetilde{S},\widetilde{V}\right), m_n\right)} \right] \right\|_\infty \\
& + (M_{IV} + \epsilon)\min\left(\tau_n, \|\phi\|_\infty\right) r_{IV,n} m_n^{-3/2}(M(\alpha) + \epsilon)\left(\frac{\log n}{n}\right)^{1/2} T_n^{L+1/2} \\
& + m_n^{-1/2}(M(\alpha) + \epsilon)\min\left(\tau_n, \|\phi\|_\infty\right)\left(\frac{\log n}{n}\right)^{1/2} T_n^{L+1/2} \\
& + \min\left(\tau_n, \|\phi\|_\infty\right) \sup_{\gamma\in\mathcal{B}_n} \int_{\left\{(s,v):\ f_{\widetilde{S},\widetilde{V}|D=1}(s,v) < m_n\right\}} \left|\widetilde{K}_{T_n}(s'\gamma - v)\right| d\sigma(s)dv \\
& + MC(s)T_n^{-s} \\
& + \frac{1}{(2\pi)^L} T_n^{L+2} \||t|^L\psi\|_1 \mathbb{E}[|\phi(Y)|\mathbf{1}\{|\phi(Y)| > \tau_r\}].
\end{aligned}$$

Résultats asymptotiques 3

Dans le cas idéal où : (1) $f_{\widetilde{S},\widetilde{V}|D=1}$ est minoré, (2) sa densité est suffisamment régulière pour que le premier terme soit négligéable et (3) le biais lié à la troncature est négligeable (*ex.* lorsque ϕ est borné), nous obtenons, $\exists M_l > 0$

$$\overline{\lim}_{n\to\infty}\left(\frac{\log n}{n}\right)^{-\frac{s}{2s+2L+1}}\|\hat{g}-g\|_\infty \leq M_l \quad p.s.$$

lorsque $T_n \asymp (n/\log(n))^{1/(2s+2L+1)}$. La vitesse d'un problème direct est $(n/\log(n))^{s/(2s+L)}$.

Extensions - Approfondissements

- Il est possible de considérer des instruments binaires.
- Modèles d'attrition dans les panel et modèle de censure.
- Inférence sur la distribution de Y sans imputation mais en utilisant le plan pour l'estimation de la loi de Y.
- Effets hétérogènes de X sur Y (modèle à coefficients aléatoires) en présence de non-réponse non-MAR.
- avec J. Heckman : distribution des gains ex-ante et ex-post, avec H. Broome et S. Hoderlein : application du papier d'évaluation des politiques publiques à l'effet des études supérieures sur le salaire (à partir de NLSY 79).

Extensions - Approfondissements

- Il est possible de considérer des instruments binaires.
- Modèles d'attrition dans les panel et modèle de censure.
- Inférence sur la distribution de Y sans imputation mais en utilisant le plan pour l'estimation de la loi de Y.
- Effets hétérogènes de X sur Y (modèle à coefficients aléatoires) en présence de non-réponse non-MAR.
- avec J. Heckman : distribution des gains ex-ante et ex-post, avec H. Broome et S. Hoderlein : application du papier d'évaluation des politiques publiques à l'effet des études supérieures sur le salaire (à partir de NLSY 79).

Extensions - Approfondissements

- Il est possible de considérer des instruments binaires.
- Modèles d'attrition dans les panel et modèle de censure.
- Inférence sur la distribution de Y sans imputation mais en utilisant le plan pour l'estimation de la loi de Y.
- Effets hétérogènes de X sur Y (modèle à coefficients aléatoires) en présence de non-réponse non-MAR.
- avec J. Heckman : distribution des gains ex-ante et ex-post, avec H. Broome et S. Hoderlein : application du papier d'évaluation des politiques publiques à l'effet des études supérieures sur le salaire (à partir de NLSY 79).

Extensions - Approfondissements

- Il est possible de considérer des instruments binaires.
- Modèles d'attrition dans les panel et modèle de censure.
- Inférence sur la distribution de Y sans imputation mais en utilisant le plan pour l'estimation de la loi de Y.
- Effets hétérogènes de X sur Y (modèle à coefficients aléatoires) en présence de non-réponse non-MAR.
- avec J. Heckman : distribution des gains ex-ante et ex-post, avec H. Broome et S. Hoderlein : application du papier d'évaluation des politiques publiques à l'effet des études supérieures sur le salaire (à partir de NLSY 79).

Extensions - Approfondissements

- Il est possible de considérer des instruments binaires.
- Modèles d'attrition dans les panel et modèle de censure.
- Inférence sur la distribution de Y sans imputation mais en utilisant le plan pour l'estimation de la loi de Y.
- Effets hétérogènes de X sur Y (modèle à coefficients aléatoires) en présence de non-réponse non-MAR.
- avec J. Heckman : distribution des gains ex-ante et ex-post, avec H. Broome et S. Hoderlein : application du papier d'évaluation des politiques publiques à l'effet des études supérieures sur le salaire (à partir de NLSY 79).

Quelques éléments de bibliographie

- Gautier, E. (2005) : " Eléments sur la sélection dans les enquêtes et sur la nonréponse non ignorable", Actes des Journées de Méthodologie Statistique 2005.
- Heckman, J. J., et E. Vytlacil (2005) : "Structural Equations, Treatment Effects, and Econometric Policy Evaluation". *Econometrica*.
- Helgason, S. (1999) : *The Radon Transform*. Birkhauser.
- Imbens, G. W., et J. D. Angrist (1994) : "Identification and Estimation of Local Average Treatment Effects". *Econometrica*.
- Little R.J.A. et Rubin D.B. (2002) : *Statistical analysis with missing data*. Wiley.
- Vytlacil, E. (2002) : "Independence, Monotonicity, and Latent Index Models : An Equivalence Result". *Econometrica*.

Quelques éléments de bibliographie

- Gautier, E. (2005) : " Eléments sur la sélection dans les enquêtes et sur la nonréponse non ignorable", Actes des Journées de Méthodologie Statistique 2005.
- Heckman, J. J., et E. Vytlacil (2005) : "Structural Equations, Treatment Effects, and Econometric Policy Evaluation". *Econometrica*.
- Helgason, S. (1999) : *The Radon Transform*. Birkhauser.
- Imbens, G. W., et J. D. Angrist (1994) : "Identification and Estimation of Local Average Treatment Effects". *Econometrica*.
- Little R.J.A. et Rubin D.B. (2002) : *Statistical analysis with missing data*. Wiley.
- Vytlacil, E. (2002) : "Independence, Monotonicity, and Latent Index Models : An Equivalence Result". *Econometrica*.

Quelques éléments de bibliographie

- Gautier, E. (2005) : “ Eléments sur la sélection dans les enquêtes et sur la nonréponse non ignorable”, Actes des Journées de Méthodologie Statistique 2005.
- Heckman, J. J., et E. Vytlacil (2005) : “Structural Equations, Treatment Effects, and Econometric Policy Evaluation”. *Econometrica*.
- Helgason, S. (1999) : *The Radon Transform*. Birkhauser.
- Imbens, G. W., et J. D. Angrist (1994) : “Identification and Estimation of Local Average Treatment Effects”. *Econometrica*.
- Little R.J.A. et Rubin D.B. (2002) : *Statistical analysis with missing data*. Wiley.
- Vytlacil, E. (2002) : “Independence, Monotonicity, and Latent Index Models : An Equivalence Result”. *Econometrica*.

QUELQUES ÉLÉMENTS DE BIBLIOGRAPHIE

GAUTIER, E. (2005) : " Eléments sur la sélection dans les enquêtes et sur la nonréponse non ignorable", Actes des Journées de Méthodologie Statistique 2005.

HECKMAN, J. J., ET E. VYTLACIL (2005) : "Structural Equations, Treatment Effects, and Econometric Policy Evaluation". *Econometrica*.

HELGASON, S. (1999) : *The Radon Transform*. Birkhauser.

IMBENS, G. W., ET J. D. ANGRIST (1994) : "Identification and Estimation of Local Average Treatment Effects". *Econometrica*.

LITTLE R.J.A. ET RUBIN D.B. (2002) : *Statistical analysis with missing data*. Wiley.

VYTLACIL, E. (2002) : "Independence, Monotonicity, and Latent Index Models : An Equivalence Result". *Econometrica*.

Quelques éléments de bibliographie

Gautier, E. (2005) : " Eléments sur la sélection dans les enquêtes et sur la nonréponse non ignorable", Actes des Journées de Méthodologie Statistique 2005.

Heckman, J. J., et E. Vytlacil (2005) : "Structural Equations, Treatment Effects, and Econometric Policy Evaluation". *Econometrica*.

Helgason, S. (1999) : *The Radon Transform*. Birkhauser.

Imbens, G. W., et J. D. Angrist (1994) : "Identification and Estimation of Local Average Treatment Effects". *Econometrica*.

Little R.J.A. et Rubin D.B. (2002) : *Statistical analysis with missing data*. Wiley.

Vytlacil, E. (2002) : "Independence, Monotonicity, and Latent Index Models : An Equivalence Result". *Econometrica*.

Quelques éléments de bibliographie

GAUTIER, E. (2005) : " Eléments sur la sélection dans les enquêtes et sur la nonréponse non ignorable", Actes des Journées de Méthodologie Statistique 2005.

HECKMAN, J. J., ET E. VYTLACIL (2005) : "Structural Equations, Treatment Effects, and Econometric Policy Evaluation". *Econometrica*.

HELGASON, S. (1999) : *The Radon Transform*. Birkhauser.

IMBENS, G. W., ET J. D. ANGRIST (1994) : "Identification and Estimation of Local Average Treatment Effects". *Econometrica*.

LITTLE R.J.A. ET RUBIN D.B. (2002) : *Statistical analysis with missing data*. Wiley.

VYTLACIL, E. (2002) : "Independence, Monotonicity, and Latent Index Models : An Equivalence Result". *Econometrica*.

MERCI POUR VOTRE ATTENTION !