Inférence fondée sur les quantiles de tranformation paramétrique entre deux distributions

Laurent Gobillon/Sébastien Roux

INED,INRA-PSE CREST-INSEE,INRA-PSE

Journées de Méthodologie Statistique

Introduction

Motivation : Nous voulons comparer les distributions conditionnelles de deux groupes d'individus.
Peut-on considérer une distribution comme la transformée de l'autre par des transformations paramétriques ?

- Idée principale : nous supposons qu'une distribution est la transformée de l'autre, les transformations étant liées à des mécanismes économiques, résumés par des paramètres à estimer.
  Ces transformations peuvent correspondre à la troncation (à gauche ou à droite) d'une distribution par rapport à l'autre, ou plus généralement à un mécanisme de sélection paramétrique.
- L'inférence sur ces paramètres est conduite en comparant les quantiles des distributions de chaque groupe. Aucune hypothèse paramétrique n'est faite sur la forme de la distribution initiale.
- Aucune information n'est nécessaire sur les observations censurées.
- Il est possible de tester la spécification du modèle.

Littérature

Principalement une littérature statistique, liée à celle s'intéressant aux effets de traitements.

- S'inspire de la littérature sur la dominance stochastique (test de la position relative de deux distributions).
- Doksum (1974), Effet d'un traitement dépendant du quantile.
- Lié aux régressions quantiles : une approche complémentaire (mais différent),(Koenker, 2005).
- Abadie, Angrist et Imbens (2002), et Chernozhukov et Hansen (2004), Introduction de variables instrumentales et explicatives.
- Athey and Imbens (2006), différence-de-différences appliquée aux distributions, décomposition de la transformation de deux distributions.

Estimation des effets d'un traitement sous un jeu minimal d'hypothèses paramétriques.

Notre approche

- Estimer et tester un modèle sous une paramétrisation restrictive.
- Les distributions entre les deux groupes sont supposées être liées par
  - Une transformation directe des valeurs,$y_2 = z_{\beta_0}(y_1)$
  - Une transformation des rangs, Soit u_j le rang d'une observation dans le groupe j, $u_2 = r_{\gamma_0}(u_1)$.

  Ainsi, soit $\lambda_j(.)$ la fonction quantile dans le groupe j, nous supposons la relation :

  $$\lambda_2 [r_{\gamma_0}(u)] = z_{\beta_0} [\lambda_1 (u)] \tag{1}$$

  pour chaque rang possible u.
- Les transformations de rang et de valeur sont supposées monotones.
- Les estimation de β_0 et γ_0 minimisent la distance entre les quantiles des deux groupes, une fois les transformations en valeur et en rang appliquées.
- Test asymptotique possible de la spécification du modèle.

Applications

- On considère une situation où les agents sont hétérogènes dans une dimension donnée, et dont la distribution est *ex-ante* la même dans les deux groupes (par exemple les abilités).
- La distribution *ex-post* peut différer à cause de mécanismes affectant les valeurs (ex : primes spécifiques) ou les rangs (sélection). Exemples :
  - Distinguer les effets d'agglomération et de compétition pour interpréter les différences de productivité entre différentes zones géographiques (cf. Combes et al., 2007)
  - Tester l'hypothèse de plafond de verre contre la discrimination uniforme (application présentée dans ce papier).
- Test de la représentation présentée.

Plan

1. Introduction
2. Une Interprétation structurelle
3. Estimation
4. Simulations de Monte Carlo

Une interprétation structurelle

- On considère une population caractérisées par un résultat $\tilde{y} \in I$ tiré dans une distribution $\tilde{F}$
- Certaines personnes sont sélectionnées, en fonction de leur position dans $\tilde{F}$.
- Pour un individu i, $P(s_i = 1|\tilde{y}_i) = p_\gamma\left(\tilde{F}(\tilde{y}_i)\right) = p_\gamma(\tilde{u}_i)$ $\tilde{u}_i$ est le rang de i dans $\tilde{F}$.
- Une fois sélectionné, le résultat final est $y_i = z_\beta(\tilde{y}_i)$, $y_i \in I_z$
- On appelle F la distribution des résultats des individus sélectionnés, u_i est le rang de i dans F. On a $u_i = r_\gamma(\tilde{u}_i)$ où

$$r_\gamma(\widetilde{u}_i) = \frac{\int_0^{\widetilde{u}_i} p_\gamma(v)\,dv}{\int_0^1 p_\gamma(v)\,dv} \tag{2}$$

- Á partir de ce processus, on déduit la relation entre F et $\tilde{F}$

$$F[z_\beta(y)] = r_\gamma\left[\widetilde{F}(y)\right] \text{ for all } y \in I \tag{3}$$

Applications

- Extension de ce cas à deux groupes différents dont la variable de résultat initiale est tirée dans une distribution de base non observée.
    - L'identification des paramètres de transformation s'appuie sur la comparaison des deux groupes.
    - En fonction de la spécification paramétrique, seules des combinaisons de la transformation des paramètres sont identifiés.
    - Mais la distribution de base n'a pas à être spécifiée.
- Deux illustrations sont actuellement développées dans des papiers compagnons :
    - Distinction des des effets de sélection et de translation entre les distributions de productivité dans des zones denses et non denses. (cf. Combes, Duranton, Gobillon, Puga et Roux, 2008)
    - Examen des effets de ségrégation sur les différentiels de salaires entre hommes et femmes, S_i étant alors une indicatrice indiquant si l'emploi i est occupé par une femme (cf. Gobillon, Meurs et Roux, 2009)

Exemple : Sélection déterministe ou troncation

- Censure à droite ou à gauche : $p_{\gamma_j}(u) = 1$ pour $u \in \left[\underline{u}_j ; \overline{u}_j\right]$, 0 ailleurs.
- Donc $r_{\gamma_j}(u) = \frac{\min\left(u, \overline{u}_j\right) - \min\left(u, \underline{u}_j\right)}{\overline{u}_j - \underline{u}_j}$
- On ne peut pas observer $\overline{u}_j$ or $\underline{u}_j$, mais on a :

$$r_{\gamma_0}(u) = \underline{v} + (\overline{v} - \underline{v})\, u, \text{ where } u \in \left[\max\left(0, \frac{-\underline{v}}{\overline{v} - \underline{v}}\right) ; \min\left(1, \frac{1 - \underline{v}}{\overline{v} - \underline{v}}\right)\right] \tag{4}$$

avec $\underline{v} = \frac{\underline{u}_1 - \underline{u}_2}{\overline{u}_2 - \underline{u}_2}$, $\overline{v} = \frac{\overline{u}_1 - \underline{u}_2}{\overline{u}_2 - \underline{u}_2}$ et $\gamma_0 = (\underline{v}, \overline{v})$

- Ici, représentation symétrique : $\underline{v}$ est le paramètre de troncation à gauche relative de la distribution 2 par rapport à la distribution 1. $\underline{v} < 0$ signifie que la distribution 1 est plus tronquée à gauche que la distribution 2.

Exemple : Sélection aléatoire

- On considère le processus de sélection fondé sur la probabilité : $p_{\gamma_j}(u) = (1-u)^{\gamma_j}$, pour $u \in [0,1]$. $\gamma_j > 0$ signifie que la probabilité d'être sélectionné est décroissante avec le rang.
- Ainsi $r_{\gamma_j}(u) = 1 - (1-u)^{\gamma_j+1}$.
- Finalement,
  $r_{\gamma_0}(x) = r_{\gamma_2} \circ r_{\gamma_1}^{-1}(x) = 1 - (1-x)^{\frac{\gamma_2+1}{\gamma_1+1}} = 1 - (1-x)^{1+\gamma_0}$
  Ainsi, si $\gamma_0 > 0$ (resp. < 0), les observations du groupe 2 sont moins (resp. plus) sélectionnées dans le haut de la distribution que dans le bas par rapport au groupe 1.

Estimation : Utilisation des fonctions quantile

On considère deux groupes, $j = 1, 2$ et z_β et r_γ les transformations de 1 vers 2.

- L'équation (3) devient :

$$F_2\left[z_\beta\left(y\right)\right] = r_\gamma\left[F_1\left(y\right)\right], \forall y \in F_1^{-1}\left(\left[0,1\right]\right) \cap F_2^{-1}\left(z_\beta^{-1}\left(\left[0,1\right]\right)\right)$$

- Soit $\lambda_j = F_j^{-1}$ la fonction quantile. L'égalité devient :

$$z_\beta\left(\lambda_1\left(u\right)\right) = \lambda_2\left(r_\gamma\left(u\right)\right), \text{ for } u \in \left[\underline{u}_\gamma, \overline{u}_\gamma\right] \tag{5}$$

avec $\underline{u}_\gamma = \max\left(r_\gamma^{-1}(0), 0\right)$ et $\overline{u}_\gamma = \min\left(r_\gamma^{-1}(1), 1\right)$.

- Pour se ramener à un support uniforme $u_\gamma\left(t\right) = \underline{u}_\gamma + \left(\overline{u}_\gamma - \underline{u}_\gamma\right) t$, on obtient :

$$h_{\beta,\gamma}\left(t\right) = z_\beta\left(\lambda_1\left(u_\gamma\left(t\right)\right)\right) - \lambda_2\left(r_\gamma\left(u_\gamma\left(t\right)\right)\right) = 0, \forall\, t \in \left[0,1\right] \tag{6}$$

Critère de minimisation

- cf. Carrasco et Florens (2000) : Soit H l'espace des fonctions carré-intégrables sur $[0,1]$ $\langle \cdot, \cdot \rangle$ le produit scalaire sur H tel que : $\langle f, g \rangle = \int_0^1 \int_0^1 f(s) g(t) \, dsdt$, et $\|\cdot\|$ la norme correspondante.
- On considère un opérateur linéaire L sur H et L^* son opérateur adjoint tel que, pour tout f et g de H, on a : $\langle Lf, g \rangle = \langle f, L^* g \rangle$. Alors, $L^* L$ peut être définie $\ell(t,s)$: $(L^* Lf)(t) = \int_0^1 f(s) \ell(t,s) \, ds$ et on a : $\|L\varphi\|^2 = \int_0^1 \int_0^1 \varphi(s) \ell(t,s) \varphi(t) \, dsdt$.
- Les estimateurs des paramètres sont la solution de :

$$\underset{(\beta,\gamma)\in\Phi}{Min} \left\| L_N \widehat{h}_{\beta,\gamma} \right\|$$

où L_N est une suite bornée d'opérateurs linéaires.

- $\widehat{h}_{\beta,\gamma}(t) = z_{\beta,2}\left(\underline{\widehat{\lambda}}_{\gamma,1}(t)\right) - \widehat{\lambda}_2\left(\underline{r}_{\gamma,2}(t)\right)$ is the empirical counterpart of $h_{\beta,\gamma}$ with $\widehat{\lambda}_j$ an estimator of λ_j, and $\underline{\widehat{\lambda}}_{\gamma,1}(t) = \widehat{\lambda}_1(u_{\gamma,1}(t))$.
- Estimators of λ_j, cf. Cheng and Parzen (1997) :

$$\widehat{\lambda}_j(u) = \int_0^1 \widetilde{\lambda}_j(t)\, d_t K_n(u,t) \tag{7}$$

  where $\widetilde{\lambda}_j$ is the sample quantile function and for each u, $K_n(u,\cdot)$ is a cdf on $[0,1]$.
- In our application, L_N is such that $l(t,s) = w_N(t)\,\tilde{l}(t,s)$, where $\tilde{l}(t,s) = 0$ for $t \neq s$ and $\tilde{l}(t,t) = \delta_d$ where δ_d is a Dirach mass and $w_N(t)$ is a weighting function.

$$(\hat{\beta,\gamma}) = \underset{\beta,\gamma}{\arg\min} \int_0^1 w_N(t)\,\hat{h}_{\beta,\gamma}(t)^2\, dt = \underset{\beta,\gamma}{\arg\min}\, \hat{C}(\beta,\gamma)$$

Asymptotique

- Idée principale : Nous adaptons Carrasco et Florens (2000) au cas particulier de la convergence du processus quantile
- Nous utilisons la propriété selon laquelle, sous certaines hypothèses, $\sqrt{N_j} f_j(\lambda)(\hat{\lambda}_j - \lambda_j)$ converge vers un pont brownien standard.
- On a alors : $\left(\hat{\beta}, \hat{\gamma}\right) \xrightarrow{P} (\beta_0, \gamma_0)$ la vraie valeur des paramètres.
- $\left(\frac{1}{\hat{N}_{\gamma,1}} + \frac{1}{\hat{N}_{\gamma,2}}\right)^{-\frac{1}{2}} \begin{pmatrix} \hat{\beta} - \beta_0 \\ \hat{\gamma} - \gamma_0 \end{pmatrix}$ converge vers un Gaussienne.

Loi Asymptotique du critère

Theorem

Sous certaines hypothèses (cf. texte) :

$$\left(\frac{1}{\hat{N}_{\gamma_0,1}}+\frac{1}{\hat{N}_{\gamma_0,2}}\right)^{-1}\hat{C}\left(\beta_0,\gamma_0\right) \overset{d}{\Longrightarrow} \int_0^1 w_1\left(t\right)\mu_1'\left(t\right)^2 B_1\left(t\right)^2 \mathrm{d}t+\int_0^1 w_2\left(s\right)\mu_2'\left(s\right)^2 B_2\left(s\right)^2 \mathrm{d}s \tag{8}$$

où $\mu_1\left(t\right)=z_{\beta_0,2}\left(\underline{\lambda}_{\gamma_0,1}\left(t\right)\right)$, $\mu_2\left(s\right)=\lambda_1\left(\underline{r}_{\gamma_0,1}\left(s\right)\right)$, $\hat{N}_{\gamma_0,j}$ *est le nombre d'observations non censurées du groupe* j, B_1 *et* B_2 *sont des ponts browniens tels que pour tout* $s,t\in[0,1]$, $B_1\left(t\right)=B_2\left(s\right)$ *avec* s *et* t *tels que :* $u_{\gamma_0,2}\left(s\right)=\underline{r}_{\gamma_0,2}\left(t\right)$. *On a aussi :*

$$\left(\frac{1}{\hat{N}_{\hat{\gamma},1}}+\frac{1}{\hat{N}_{\hat{\gamma},2}}\right)^{-1}\hat{C}\left(\hat{\beta},\hat{\gamma}\right)-\left(\frac{1}{\hat{N}_{\gamma_0,1}}+\frac{1}{\hat{N}_{\gamma_0,2}}\right)^{-1}\hat{C}\left(\beta_0,\gamma_0\right)\overset{d}{\sim}-\frac{1}{2}\sum_k D_{(\beta_0,\gamma_0),k}X_k^2 \tag{9}$$

où X_k, $k=1,...,K$, *sont des lois normales indépendantes et* $D_{(\beta_0,\gamma_0),k}$ *est la* $k^{ième}$ *valeur propre (positive) de* $\Lambda'\Gamma\Lambda$ *where* $\Gamma^{-1}\Omega\Gamma^{-1}=\Lambda\Lambda'$ *(qui est une décomposition de Choleski).*

Implications

- La loi asymptotique (8) peut être réécrite

$$\begin{aligned} & \int_0^1 w_1(t)\,\mu_1'(t)^2\,B_1(t)^2\,\mathrm{d}t + \int_0^1 w_2(s)\,\mu_2'(s)^2\,B_2(s)^2\,\mathrm{d}s \\ = & \int_0^1 \left[w_1(t) + w_2(s(t))\,z_{\beta_0,1}(\mu_1(t))^2\right]\mu_1'(t)^2\,B_1^2(t)\,\mathrm{d}t \quad (10) \end{aligned}$$

- pour des poids spécifiques
$w_1(t) = w_2(s(t)) = \frac{1}{\left[1+z'_{\beta_0,1}[\mu_1(t)]^2/s'(t)\right]\mu_1'(t)^2}$, l'intégrale (10) s'écrit comme une statistique pivotale et suit une loi de Cramer Van-Mises $\int_0^1 B_1^2(t)\,\mathrm{d}t$
- On en déduit un test du modèle

$$H_0 : C_a < \int_0^1 B_1^2(t)\,\mathrm{d}t$$

mais n'est pas très puissant : la statistique est estimée en les valeurs estimées des paramètres.

Calcul du critère

- Les données consistent en des valeurs y_k^j pour chaque groupe $j = 1, 2$ et $k \in \{1, \ldots, N_j\}$.
- Les données sont ordonnées $y_1^j < \ldots < y_{N_j}^j$. Nous en déduisons l'estimation $\hat{\lambda}\left(\frac{k}{N_j}\right) = y_k^j$.
- Pour calculer le critère, il faut évaluer le critère en tous les rangs possibles, en particulier pour le rang transformé dans l'autre groupe. Pour $\frac{k}{N_1} \in \left[\underline{u}_{\gamma,1}, \overline{u}_{\gamma,1}\right]$, soit $k^* \in \{1, ..., N_2\}$ l'entier tel que : $\frac{k^*}{N_2} < r_\gamma\left(\frac{k}{N_1}\right) < \frac{k^*+1}{N_2}$. L'estimateur de $\lambda_2\left(r_\gamma\left(\frac{k}{N_1}\right)\right)$ est défini comme :

$$\hat{\lambda}_2\left(r_\gamma\left(\frac{k}{N_1}\right)\right) = \frac{\left(\frac{k^*+1}{N_2} - r_\gamma\left(\frac{k}{N_1}\right)\right)\hat{\lambda}_2\left(\frac{k^*+1}{N_2}\right) + \left(r_\gamma\left(\frac{k}{N_1}\right) - \frac{k^*}{N_2}\right)\hat{\lambda}_2\left(\frac{k^*}{N_2}\right)}{\frac{1}{N_2}}$$

- On définit R_j le vecteur des rangs observés ou construits dans le groupe j, le critère C_1 peut être calculé :

$$\begin{aligned}
\hat{C}_1(\beta,\gamma) &= \frac{1}{\overline{u}_{\gamma,1}-\underline{u}_{\gamma,1}}\int_{\underline{u}_{\gamma,1}}^{\overline{u}_{\gamma,1}} w_1\left(\frac{u-\underline{u}_{\gamma,1}}{\overline{u}_{\gamma,1}-\underline{u}_{\gamma,1}}\right)\left[z_{\beta,2}\left(\hat{\lambda}_1(u)\right)-\hat{\lambda}_2(r_{\gamma,2}(u))\right]^2 \mathrm{d}u \\
&\approx \frac{1}{2}\sum_{i=2}^{\hat{N}}\left[\begin{array}{l} w_1\left(\frac{R_1[i]-\underline{u}_1}{\overline{u}_1-\underline{u}_1}\right)\left[z_\beta\left(\hat{\lambda}_1(R_1[i])\right)-\hat{\lambda}_2(R_2[i])\right]^2 \\ +w_1\left(\frac{R_1[i-1]-\underline{u}_1}{\overline{u}_1-\underline{u}_1}\right)\left[z_\beta\left(\hat{\lambda}_1(R_1[i-1])\right)-\hat{\lambda}_2(R_2[i-1])\right]^2 \end{array}\right] \\
&\quad \times\frac{(R_1[i]-R_1[i-1])}{\overline{u}_1-\underline{u}_1}
\end{aligned}$$

- Le critère est continu en β, mais pas en γ.
- Deux étapes de la procédure d'estimation : γ étant fixé, $\beta(\gamma)$ est estimé en utilisant la procédure usuelle. γ est alors estimé par grille ou par simplexe.

Calcul de la statistique de test

- Le critère est évalué en $\left(\hat{\beta}, \hat{\gamma}\right)$ and weights $w_1\left(t\right) = w_2\left(s\left(t\right)\right) = \frac{1}{\left[1+z'_{\beta_0,1}[\mu_1(t)]^2/s'(t)\right]\mu'_1(t)^2}$
- Difficulté, estimation de μ'_1, qui dépend de λ'.
- Á partir de Bloch et Gartswirth(1968), on a :

$$\hat{\lambda}'_j\left(u\right) = N_j \frac{y_{[k_1]} - y_{[k_0]}}{k_1 - k_0}$$

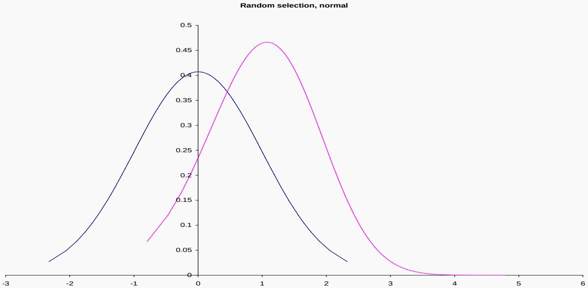
où $k_1 = [nu + m]$ (avec $[\bullet]$ la fonction partie entière) si $nu + m < n$ et $k_1 = n$ sinon, $k_0 = [nu - m]$ si $nu - m \geqslant 1$ et $k_0 = 1$ sinon. $m = \sqrt{n}$.

Monte-Carlo

- Processus générateur des données :
  1. Choix d'une distribution de base (Normale ou Pareto)
  2. Pour chaque groupe, données générées avec y tiré dans la distribution de base.
  3. Une première troncation τ pour que le support soit borné, séparément pour les deux groupes ($\tau/2$ à chaque côté). N_j est le nombre d'observations disponibles après cette opération.
  4. Transformation spécifique au groupe 2, $y_2 \rightarrow ay_2 + b$. Les valeurs du groupe 1 restent inchangées.
  5. Processus de sélection appliqué au groupe 2 :
     - Sélection aléatoire : une variable ε est tirée dans une loi uniforme, l'observation est sélectionnées si $\varepsilon \leq (1-y)^\gamma$
     - Troncation à gauche : obs. selectionnée si $\gamma \leq y$
- Chaque spécification correspond à un choix de la distribution de base et du processus de sélection

Sélection aléatoire avec une loi normale

- $y_2 \rightarrow ay_2 + b$
- sélectionnée si $\varepsilon \leq (1 - u_2)^\gamma$

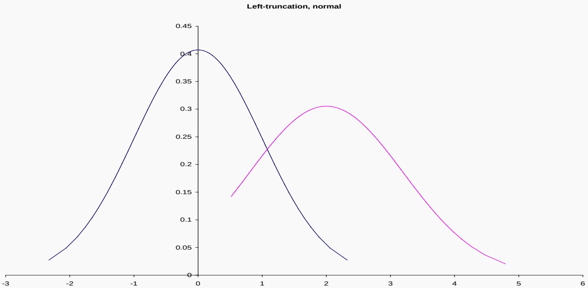


	Baseline		N=1,000		N=100,000		$\tau = 0.01$	
	Mean	*RMSE*	Mean	*RMSE*	Mean	*RMSE*	Mean	*RMSE*
$\hat{\gamma}$	0.487	*0.058*	0.395	*0.186*	0.498	*0.020*	0.460	*0.086*
$\hat{a}$	1.196	*0.014*	1.177	*0.047*	1.200	*0.005*	1.190	*0.019*
$\hat{b}$	0.991	*0.033*	0.918	*0.119*	0.999	*0.011*	0.971	*0.058*
Stat (true)	0.114	*0.086*	0.118	*0.102*	0.109	*0.081*	0.096	*0.069*
Stat (estimated)	0.041	*0.022*	0.035	*0.021*	0.039	*0.020*	0.037	*0.020*
$\hat{N}_1$	10000	*0.000*	1000	*0.000*	100000	*0.000*	10000	*0.000*
$\hat{N}_2$	6663	*47.885*	667	*15.991*	66678	*134.397*	6663	*47.885*
$\hat{C}$ x10^6	10.221	*5.456*	87.532	*51.947*	0.982	*0.499*	9.344	*5.028*
$\hat{C}_1$ x10^6	6.649	*3.572*	53.789	*29.730*	0.646	*0.332*	6.206	*3.383*
$\hat{C}_2$ x10^6	3.573	*1.940*	33.743	*23.000*	0.336	*0.170*	3.138	*1.707*

	$\tau = 0.00$		$\tau = 0.05$		γ=0.2		γ=0.7	
	Mean	*RMSE*	Mean	*RMSE*	Mean	*RMSE*	Mean	*RMSE*
$\hat{\gamma}$	0.268	*0.147*	0.495	*0.047*	0.196	*0.040*	0.654	*0.078*
$\hat{a}$	1.140	*0.040*	1.198	*0.012*	1.199	*0.013*	1.189	*0.019*
$\hat{b}$	0.818	*0.120*	0.996	*0.022*	0.998	*0.028*	0.973	*0.042*
Stat (true)	0.048	*0.032*	0.132	*0.104*	0.105	*0.085*	0.112	*0.088*
Stat (estimated)	0.026	*0.013*	0.044	*0.023*	0.037	*0.021*	0.040	*0.021*
$\hat{N}_1$	10000	*0.000*	10000	*0.000*	10000	*0.000*	10000	*0.000*
$\hat{N}_2$	6663	*47.885*	6663	*47.885*	8333	*41.749*	5878	*47.819*
$\hat{C}$ x10^6	6.489	*3.312*	10.961	*5.861*	8.111	*4.706*	10.734	*5.644*
$\hat{C}_1$ x10^6	4.352	*2.281*	6.973	*3.715*	5.047	*2.940*	6.968	*3.660*
$\hat{C}_2$ x10^6	2.137	*1.147*	3.988	*2.194*	3.064	*1.784*	3.766	*2.056*

Troncation à gauche avec une loi normale

- $y_2 \rightarrow ay_2 + b$
- selected if $\gamma \leq u_2$



	Baseline		N=1,000		N=100,000		$\tau = 0.01$	
	Mean	*RMSE*	Mean	*RMSE*	Mean	*RMSE*	Mean	*RMSE*
$\hat{\gamma}$	-0.112	*0.018*	-0.128	*0.089*	-0.111	*0.005*	-0.112	*0.015*
$\hat{a}$	1.200	*0.024*	1.220	*0.089*	1.200	*0.008*	1.200	*0.022*
$\hat{b}$	0.998	*0.038*	0.976	*0.144*	1.000	*0.012*	0.999	*0.036*
Stat (true)	0.130	*0.105*	0.112	*0.085*	0.131	*0.107*	0.121	*0.093*
Stat (estimated)	0.037	*0.019*	0.034	*0.017*	0.034	*0.015*	0.036	*0.018*
$\hat{N}_1$	8996	*140.227*	891	*53.167*	90000	*424.261*	8997	*122.112*
$\hat{N}_2$	8999	*36.048*	899	*8.793*	89999	*90.161*	8999	*36.013*
$\hat{C}$ x10^6	8.188	*4.252*	75.083	*37.171*	0.746	*0.341*	7.981	*4.024*
$\hat{C}_1$ x10^6	4.833	*2.515*	44.822	*22.248*	0.440	*0.202*	4.710	*2.379*
$\hat{C}_2$ x10^6	3.354	*1.740*	30.261	*15.453*	0.306	*0.139*	3.271	*1.648*

	$\tau = 0.00$		$\tau = 0.05$		γ=-0.250		γ=-0.053	
	Mean	*RMSE*	Mean	*RMSE*	Mean	*RMSE*	Mean	*RMSE*
$\hat{\gamma}$	-0.111	*0.014*	-0.112	*0.022*	-0.253	*0.039*	-0.053	*0.010*
$\hat{a}$	1.199	*0.022*	1.200	*0.028*	1.202	*0.035*	1.199	*0.019*
$\hat{b}$	0.999	*0.035*	0.998	*0.042*	0.995	*0.060*	0.999	*0.028*
Stat (true)	0.091	*0.061*	0.139	*0.117*	0.148	*0.110*	0.120	*0.099*
Stat (estimated)	0.035	*0.016*	0.038	*0.020*	0.038	*0.019*	0.037	*0.019*
$\hat{N}_1$	9000	*114.245*	8994	*173.351*	7988	*242.957*	9500	*85.663*
$\hat{N}_2$	8999	*36.051*	8999	*36.036*	8005	*42.277*	9498	*24.004*
$\hat{C}$ x10^6	7.828	*3.674*	8.432	*4.442*	9.468	*4.810*	7.761	*4.038*
$\hat{C}_1$ x10^6	4.617	*2.175*	4.979	*2.630*	5.591	*2.846*	4.582	*2.393*
$\hat{C}_2$ x10^6	3.211	*1.502*	3.453	*1.817*	3.878	*1.972*	3.179	*1.646*

Puissance du test

Résultats d'estimation, Pareto DGP, base normale

	Specification 1		Specification 2		Specification 3	
	Mean	*RMSE*	Mean	*RMSE*	Mean	*RMSE*
$\hat{\gamma}$	0,000	*0,000*	0,896	*0,092*	17,679	*0,092*
$\hat{a}$	1,980	*0,026*	0,522	*0,200*	110,652	*0,200*
$\hat{b}$	-2,169	*0,035*	2,122	*0,497*	-112,459	*0,497*
Stat (estimated)	57,706	*2,805*	174,380	*433,850*	19,922	*433,850*
$\hat{N}_1$	10000	*0*	9995	*70*	789	*70*
$\hat{N}_2$	6663	*47*	933	*771*	2498	*771*
$\hat{C}$ x10^6	14431,294	*695,961*	205408,529	*94108,455*	13026,401	*94108,455*
$\hat{C}_1$ x10^6	11497,727	*551,681*	39978,248	*16096,670*	13025,938	*16096,670*
$\hat{C}_2$ x10^6	2933,567	*159,804*	165430,282	*101446,770*	0,463	*101446,770*

Note : The baseline distribution is a normal. Baseline parameters for all specifications.
Data were generated with a Pareto distribution with variance one.

Estimation results, Pareto DGP, normal baseline

	Specification 1		Specification 2		Specification 3	
	Mean	*RMSE*	Mean	*RMSE*	Mean	*RMSE*
$\hat{\gamma}$	4,422	*0,205*	-1,000	*0,000*	129,161	*0,000*
$\hat{a}$	1,064	*0,029*	1,084	*0,016*	3,577	*0,016*
$\hat{b}$	3,654	*0,041*	1,977	*0,015*	9,198	*0,015*
Stat (estimated)	3,395	*0,401*	27,152	*1,664*	239,395	*1,664*
$\hat{N}_1$	10000	*0*	5001	*0*	1051	*0*
$\hat{N}_2$	6663	*47*	8999	*31*	2498	*31*
$\hat{C}$ x10^6	849,123	*100,060*	8446,571	*518,105*	143889,944	*518,105*
$\hat{C}_1$ x10^6	171,114	*23,915*	4561,615	*283,248*	125757,883	*283,248*
$\hat{C}_2$ x10^6	678,010	*78,270*	3884,956	*249,907*	18132,061	*249,907*

Note : The baseline distribution is a Pareto. Baseline parameters for all specifications.
Data were generated with a normal distribution with variance one.

Conclusion

- Méthode permettant d'estimer séparément la sélection et les paramètres de transformation. Spécification semi-paramétrique.
- Inférence fondée sur les comparaisons de quantiles : asymptotique plus complexe. Un test possible est proposé.
- Les Monte Carlo montrent que la procédure d'estimation fonctionne et que le test proposé semble fonctionner.
- Extensions :
  1. Introduction de variables explicatives dans les paramètres de transformation.
  2. Robustesse à l'erreur de mesure.