ESTIMATION ULTIME DES INDICATEURS DE PAUVRETÉ PAR RAPPORT AU REVENU ET A LA DEPENSE DANS UNE BASE DE DONNÉES DE LA PAUVRETÉ.

GANE SAMB LO

ABSTRACT. Les indicateurs de pauvreté dépendent sûrement de la loi des revenus et des dépenses. Nos travaux actuels ont permis d'estimer les indicateurs et leurs variances asymptotiques. Il semble indiqué alors de déterminer la loi des revenus et des dépenses dans la base ESAM I (1996) du Sénégal. Les résultats concernent l'ensemble des dix régions existantes dans la période de l'enquête.

1. INTRODUCTION

L'objectif de cette note est l'estimation des indicateurs exacts de pauvreté, suivant les résultats de normalité asymptotique de [6], à l'ordre de grandeur de la taille des échantillons disponibles. La présente application concerne les données de l'Enquête Sénégalaise Auprès des Ménages du Sénégal de 1996 (ESAM). Ces revenus jouent un rôle important dans l'appréciation quantitative de la pauvreté. Celle-ci est un phénomène complexe dont l'exploration et l'analyse requièrent l'utilisation simultanée de plusieurs disciplines : l'économie, la sociologie, les statistiques, la médecine, etc. Les aspects quantitatifs, à coté des aspects qualitatifs, jouent un rôle important. L'appréciation quantitative se réalise par plusieurs indicateurs de pauvreté dont naturellement la prévalence de pauvreté, c'est-à-dire la proportion de pauvres dans la population étudiée.

Ces mesures de pauvreté sont basées sur le revenu ou la dépense des ménages. Dans la suite, la variable revenu ou dépense de la population étudiée de N ménages sera notée Y prenant les valeurs ordonnées : $\mathrm{Y}_1, Y_2, ..., Y_N$. Pour définir un ménage pauvre, on fait appel à un seuil ou ligne de pauvreté Z de sorte qu'un ménage j est déclaré pauvre si son revenu est inférieur à Z, c'est-à-dire

$$(j \text{ pauvre}) \Longleftrightarrow (Y_j < Z) \tag{1.1}$$

Le nombre de ménages pauvres est le nombre Q tel que $\mathrm{Y}_Q < Z \leq Y_{Q+1}$. Les revenus pauvres sont $\{\mathrm{Y}_1, Y_2, ..., Y_Q\}$. La détermination du seuil est aussi une question complexe, ayant fait l'objet de larges discussions, suivant les approches utilisées : en termes de conditions de vie minimales, en termes de capacités, en termes de survie, etc. Le lecteur intéressé peut trouver une bonne base de discussion dans Ravallion [1]. Les indicateurs de pauvreté qui nous pré-occupent dans cet article sont des fonctions mathématiques des revenus des Q pauvres sous la forme globale

$$P(Y,N,Q)=\frac{1}{\delta(H)\rho(L)}\sum_{j=1}^{Q}\beta\left(M,N,j\right)\gamma\left(e_{j}\right) \tag{1.2}$$

où $e_j = (Z - Y_j)/Z$ est le déficit relaif de pauvreté du $j^{éme}$ pauvre, δ, β, ρ et γ étant des fonctions données; H, L et M étant pris dans {N, Q}.

La statistique (1.2) est la forme empirique de l'indicateur dans l'échantillon choisi. On pourra alors supposer que les revenus ou les dépenses observées $Y_1, Y_2, ..., Y_n$ sont des observations indépendantes d'une même variable aléatoire de Loi G(y)=$P(Y_i \leq y)$. Et (1.2) devient

$$J_n = \frac{1}{n}\sum_{j=1}^{q}c(n,q,j)\ d\left(\frac{Z-Y_{j,n}}{Z}\right), \tag{1.3}$$

où $Y_{1,n} \leq Y_{2,n} \leq ... \leq Y_{n,n}$ est la statistique d'ordre associée à l'échantillon $Y_1, Y_2, ..., Y_n$, $\{h,l,m\} \subset \{q,n\}$, et q est le nombre de pauvres dans l'échantillon, i.e., $q/n = G_n(Z)$ la valeur en Z pour la fonction de répartition empirique de l'échantillon, avec bien entendu, par le théorème de Glivenko-Cantelli, $G_n(Z) \to G(Z)$, quand n tend vers l'infini.

Il devient alors important de faire une théorie asymptotique de la convergence de J_n et de sa loi limite. Lô et al.[6] et G. Dia [4] se sont attelés à cette tâche. Le rôle de la loi de la variable, c'est à dire, de G est très important, contrairement à ce que pensent Boccanfuso et al.[3].

Une théorie asymptotique élargie a été faite dans [6], concernant les principaux indicateurs de la littérature pour une très vaste classe de fonctions de répartition G. Néanmoins, (1.3) montre que les statistiques de la pauvreté concernent la queue inférieure de la distribution et la partie centrale. Or, d'une manière générale, c'est la queue supérieure qui est davantage étudiée en statistiques puisqu'on peut toujours se ramener un problème de queue inférieure en un problème de queue supérieure. Nous allons suivre cette tendance en considérant la suite de variables aléatoires

$$X_i = 1/(Y_i - y_0),\ \text{i=1,...,n} \tag{1.4}$$

avec $y_0 = \inf\{y \geq 0, G(y) > 1\}$, le revenu ou la dépense théorique minimum. La fonction de répartition F des X_i possède un support infini à droite, i.e., $x_0 = \sup\{x, F(x) < 1\} = +\infty$. Elle est liée à G par

$$G(y) = 1 - F(1/(y - y_0),\ y \geq y_0. \tag{1.5}$$

et

$$G^{-1}(u) = y_0 + 1/F^{-1}(u),\ 0 \leq u \leq 1. \tag{1.6}$$

De manière concrète, nous estimerons y_0 par $Y_{1,n}$ la plus petite observation et la série $X_i = 1/(Y_i - Y_{1,n})$, avec l'exclusion des variables qui réalise ce minimum, sera considérée dans la suite. J_n devient

$$J_n = \frac{1}{n}\sum_{j=1}^{q}c(n,q,j)\ d\left(\frac{Z-y_0-X_{n-j+1,n}^{-1}}{Z}\right) \tag{1.7}$$

Il nous faut alors exposer les résultats de [6] en ignorant les conditions rigoureuses qu'on peut toujours retrouver dans l'article cité. Précisément, nous avons

Theorem 1. *$\sqrt{n}(J_n - D_n) \rightarrow N(0, \theta^2)$ avec*

$$D_n = \int_0^{G_n(Z)} L(s)\ d\left(\frac{Z - y_0 - 1/F^{-1}(1-s)}{Z}\right) ds \tag{1.8}$$

et

$$\theta^2 = Z^{-2} \int_0^{G(Z)} \int_0^{G(Z)} L(u)\ L(v)\ h(u)h(v)(u \wedge v - uv)\ du\ dv. \tag{1.9}$$

avec $h(s)$=$a(s)$ $d'(\frac{Z-y_0-1/F^{-1}(1-s)}{Z})$. De plus, si $lim_{u\downarrow 0}\ d(u) = 0$, alors D_n peut être remplacé par

$$D = \int_0^{G(Z)} L(s)\ d\left(\frac{Z - y_0 - 1/F^{-1}(1-s)}{Z}\right) ds. \tag{1.10}$$

où $a(s)$=$(1/F^{-1}(1-u))'$ et $L(s)$=$\lim_{n\rightarrow\infty} (c(n,q,j)_{s=j/n}$.

Ce résultat nous donne l'intervalle de confiance au seuil α de D, selon la formule

$$D = J_n \pm \theta \times z_{1-\alpha/2}/\sqrt{n}$$

C'est ce résultat que nous allons appliquer. Il nous faut alors avoir une estimation de F. Dans [5], les techniques d'ajustage utilisant le test de Kolmogorov-Smirnov et la régression simple ont été appliquées aux variables revenus et dépenses de la base ESAM. Le modèle généralement accepté et fortement recommandé, pour toutes les régions et les grandeurs nationales, est le modèle lognormal pour F, i.e.,

$$F(x) = \begin{cases} \Phi((\log\ x - m)/\sigma), & si\ \ x \geq c \\ \text{0 sinon} \end{cases}, \tag{1.11}$$

où

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} \exp(-t^2/2)\ dt \tag{1.12}$$

Les paramètres σ et m sont estimés selon la méthode du maximum de vraisemblance par

$$\widehat{m} = \overline{\log X} = \frac{1}{n} \sum_{j=1}^{j=n} \log X_j \tag{1.13}$$

et

$$\widehat{\sigma}^2 = \frac{1}{n-1} \sum_{j=1}^{n} (\log\ \ X_j - \overline{\log X})^2. \tag{1.14}$$

2. Applications

La série X_i, $(i = 1, ..., n)$ est la transformation du revenu ou de la dépense Y selon (1.4) pour une région donnée ou pour tout le Sénégal. Nous adoptons le modèle lognormal avec les estimations de paramètres (1.13) et (1.14).

Nous aurons alors

$$h(s) = \sigma\sqrt{2\pi}\exp(-\Phi^{-1}(1-s)^2/2)\exp(-\sigma\Phi^{-1}(1-s)-m) \tag{2.1}$$

Pour les différents indicateurs, nous avons pour la famille FGT de paramètre α, L(s)=1 et d(u)=u$^{\alpha}$; pour la mesure de Sen,

$$L(s) = (1 - s/G(Z)) \tag{2.2}$$

et d(u)=u; et pour la mesure de Shorrocks◦

$$L(s) = 2(1 - s) \tag{2.3}$$

et d(u)=u.

Rappelons que la mesure de pauveté est relative dans le temps et dans l'espace. Dans le temps, il s'agit avant tout de voir si l'indicateur utilisé a varié suite à une politique de réduction et de mesurer l'impact des actions de développement. Dans l'espace, il s'agit de voir la répartition spatiale de la pauvreté afin de bien cibler les groupes cibles et d'aider les plus pauvres.

Sur le plan statistique, les différences entre les indicateurs dans le temps et dans l'espace peuvent simplement provenir de la variabilité normale mesurée par la variance et n'être pas significatives. Il faut donc se fier à l'intervalle de confiance et non la valeur scalaire.

Ici, nous pourrons reclasser les régions du Sénégal, avec les mesures et les dépenses, et voir les classifications significatives au seuil standard de 5%, avec une valeur critique $z_{1-\alpha/2} = 1.95$. Pour le revenu et pour les dépenses, nous avons utilisé le seuil de pauvreté de 392. L'interprétation des comparaisons devient réellement significative, sur le plan statistique, par la considération des intervalles de confiance. Nous allons donner les résultats des estimations, obtenues en utilisant des formules d'intégrations numériques des formules (1.9) et (1.10). Dans les tableaux pour la mesure FGT de paramètres unité, nous rappelons les prévalences des différentes régions.

2.1. **Variable revenu.** FGT 1

régions/paramètres	m	σ	prévalence	Borne inf	Borne sup	D
Kolda	-10.73	1.20	81.31	46.42	55.55	52.43
Tamba	-10.94	1.09	78.57	41.24	52.68	48.15
Diourbel	-11.08	1.39	68.83	34.94	43.62	40.56
Fatick	-11.20	1.29	68.33	37.21	47.42	41.30
Thies	-11.19	1.39	67.33	36.43	43.37	41.33
Kaolack	-11.35	1.18	65.82	33.35	40.99	37.33
St-Louis	-11.30	1.37	65.29	32.91	40.86	40.08
Louga	-11.25	1.31	64.37	29.35	38.92	35.12
Zigu	-11.45	1.44	60.65	30.25	39.93	37.13
Dakar	-12.39	1.32	35.38	14.70	18.04	17.24
Senegal	-11.64	1.30	59.92	30.57	32.96	32.57

FGT 2

régions/paramètres	m	σ	Borne inf	Borne sup	D
Kolda	-10.73	1.20	32.61	40.87	38.22
Tamba	-10.94	1.09	27.51	37.58	33.88
Diourbel	-11.08	1.39	22.99	30.11	27.75
Fatick	-11.20	1.29	25.89	34.61	29.21
Kaolack	-11.19	1.39	25.60	31.55	29.64
Thiès	-11.35	1.18	21.76	27.93	25.33
St-Louis	-11.30	1.37	21.65	28.52	28.94
Louga	-11.25	1.31	18.05	25.46	22.64
Zigu	-11.45	1.44	20.86	29.17	26.88
Dakar	-12.39	1.32	9.25	11.79	10.98
Senegal	-11.64	1.30	20.75	22.73	22.62

SEN

régions/paramètres	m	σ	Borne inf	Borne sup	D
Kolda	-10.73	1.20	59.20	64.38	61.79
Tamba	-10.94	1.09	60.22	70.36	66.53
Diourbel	-11.08	1.39	46.15	51.25	49.67
Fatick	-11.20	1.29	48.38	54.64	50.94
Thies	-11.19	1.39	48.06	52.18	50.78
Kaolack	-11.35	1.18	43.71	48.59	46.64
St-Louis	-11.30	1.37	43.86	48.70	49.51
Louga	-11.25	1.31	39.69	45.69	43.62
Zigu	-11.45	1.44	41.71	47.65	46.06
Dakar	-12.39	1.32	20.79	23.26	22.43
Senegal	-11.64	1.30	39.52	41.10	40.98

SHORROCKS

régions/paramètres	m	σ	Borne inf	Borne sup	D
Kolda	-10.73	1.20	65.64	73.06	70.61
Tamba	-10.94	1.09	60.22	70.36	66.53
Diourbel	-11.08	1.39	53.68	62.43	59.39
Fatick	-11.20	1.29	56.79	67.32	60.85
Thies	-11.19	1.39	52.25	63.43	61.15
Kaolack	-11.35	1.18	51.81	60	56.20
St-Louis	-11.30	1.37	51.59	60.12	60.09
Louga	-11.25	1.31	46.61	57.10	53.01
Zigu	-11.45	1.44	49.24	60.29	57.10
Dakar	-12.39	1.32	26.44	31.63	30.30
Senegal	-11.64	1.30	48.90	51.82	51.38

2.2. **Variable dépense.** FGT 1

régions/paramètres	m	σ	prévalence	Borne inf	Borne sup	D
Kolda	-11.10	0.81	76,26	28.68	35.82	33.89
Tamba	-11.24	0.88	62,7	16.55	24.28	22.34
Diourbel	-11.29	0.86	54,07	19.60	25.68	24.29
Fatick	-11.23	0.93	67,76	27.04	35.47	34.29
Kaolack	-11.19	0.95	61,41	23.72	29.29	28.14
Thiès	-11.36	0.95	62,09	20.37	25.38	25.65
St-Louis	-11.56	0.67	52,24	16.31	21.28	20.55
Louga	-11.21	0.95	54,6	10.70	15.95	15.38
Zigu	-11.37	1.07	61,57	22.84	30.29	28.16
Dakar	-12.29	0.92	19,16	3.76	5.12	4.55
Senegal	-11.78	0.88	48,72	16.34	18.01	18.94

FGT 2

régions/paramètres	m	σ	Borne inf	Borne sup	D
Kolda	-11.10	0.81	13.90	18.82	18.05
Tamba	-11.24	0.88	7.06	11.36	9.8
Diourbel	-11.29	0.86	8.48	12.10	11.44
Fatick	-11.23	0.93	14.22	20.43	20.05
Kaolack	-11.19	0.95	11.20	14.75	14.30
Thiès	-11.36	0.95	9.62	12.82	13.11
St-Louis	-11.56	0.67	6.58	9.48	9.25
Louga	-11.21	0.95	3.32	5.64	5.46
Zigu	-11.37	1.07	12.03	17.14	15.70
Dakar	-12.29	0.92	1.16	1.73	1.52
Senegal	-11.78	0.88	7.57	8.64	9.58

SEN

régions/paramètres	m	σ	Borne inf	Borne sup	D
Kolda	-11.10	0.81	37.38	43.01	42.55
Tamba	-11.24	0.88	24.70	30.52	28.67
Diourbel	-11.29	0.86	27.26	31.86	31.16
Fatick	-11.23	0.93	37.19	43.14	43.17
Kaolack	-11.19	0.95	32	35.98	35.63
Thiès	-11.36	0.95	28.56	32.25	32.79
St-Louis	-11.56	0.67	22.97	27.12	26.99
Louga	-11.21	0.95	15.98	19.97	19.90
Zigu	-11.37	1.07	31.88	37.10	35.71
Dakar	-12.29	0.92	5.43	6.64	6.17
Senegal	-11.78	0.88	22.22	23.57	24.85

SHORROCKS

régions/paramètres	m	σ	Borne inf	Borne sup	D
Kolda	-11.10	0.81	42.59	49.76	48.44
Tamba	-11.24	0.88	27.98	37.22	34.57
Diourbel	-11.29	0.86	31.64	38.83	37.38
Fatick	-11.23	0.93	42.31	51.28	50.55
Kaolack	-11.19	0.95	37.11	43.32	42.33
Thiès	-11.36	0.95	33.24	39.24	39.78
St-Louis	-11.56	0.67	26.97	33.39	32.72
Louga	-11.21	0.95	18.52	24.37	24.80
Zigu	-11.37	1.07	37.29	46.09	43.58
Dakar	-12.29	0.92	7.13	9.55	8.55
Senegal	-11.78	0.88	27.60	29.96	31.54

CONCLUSION

Les estimations étant consistentes, il sera question dans un article à venir d'estimer les indicateurs de pauvreté pour ces modèles. Mais il faudra au paravent simuler les lois limites qui seront utilisées pour les lois décrites ci-haut avec des paramètres dans l'ordre de grandeur des estimations trouvées ci-haut. Il sera aussi interessant d'appliquer les méthodes décritres ici à des bases de même nature Rapport technique, Université Gaston Berger de Saint-Louis, Lerstad n°4.

References

[1] Ravallion M.(1992). Poverty Comparisons. A Guide to Concepts and Methods. Lsms, Working Paper, n°88, *WorldBank.*

[2] Barbut M.(1989). Distribution de type parétien et représentation des inégalités. In : Mathématiques informatique et sciences humaines. 106, pp.53-69

[3] Boccanfuso D., Decaluwé B et Savard L.(2003). Poverty, Income Distribution and CGE modeling : Does the Functionnal Form Matter?. Priliminary Draft. Dakar

[4] Dia G.(2004). Répartition aléatoire des revenus et estimation de l'indice de pauvreté. Publication de l'Ufr SAT, Université Gaston Berger de Saint-louis, Lerstad n°4.

[5] Lô. G. S.et al.(2004). Estimation des lois du revenu et de la dépense dans une base de données de la pauvreté. Université de Saint-Louis.

[6] Lô, G.S, Sall S.T. et Seck C.T.(2004). The asymptotic theory of the poverty measures under the influence of the extremes. Publications de l'Ufr SAT, Université Gaston Berger de Saint-louis, Lerstad n°7.

[7] Lô, G. S. et Diop. A (2004). On a continuous Hill statistic process and its asymptotic normality theory.

LERSTAD, Université Gaston Berger de Saint-Louis, Sénégal
E-mail address: ganesamblo@ufrsat.org, ganesamblo@yahoo.com