Réduction de la Variance de l'estimateur d'indice des valeurs extrêmes sous des données censurées

Etabli par : **Bekhdidja Chaima**

Encadré par : **Meddi Fatima**

Département des mathématiques

Université Kasdi Merbah Ouargla 30000, Algérie

Résumé

On va appliquer la méthode de **Monte-Carlo pour réduction de variance** sur l'estimateur d'indice des valeurs extrêmes sous des données censurées. Nous allons présenter l'algorithme correspondant.

Mots clés : réduction de la variance, indice de l'estimateur des valeurs extrêmes

introduction

Dans ce travail nous allons réduire la variance d'indice d'estimateur des valeurs extrêmes sous des données censurées $\hat{\gamma}$.

$$\hat{\gamma}_{z,k,n}^{(c,.)} = \frac{\hat{\gamma}_{z,k,n}^{(.)}}{\frac{1}{k}\sum_{i=1}^{k}\delta_{n-i+1,n}}$$

Par la méthode de monte-carlo.

1ere chapitre : on va expliquer la méthode

2eme chapitre : on réduit la variance

2eme chapitre : simulation

Estimateur de l'indice des valeurs extrêmes :

On considère l'espace probabilisé (Ω,A,P) et soit l'échantillon $X_1,\dots,X_n$ de variables aléatoires positives définie sur l'espace de probabilité (Ω,A,P) avec sa fonction de répartition F , supposons que la queue de distribution (1-F) est a variations régulières a l'infini d'indice $(\frac{-1}{\gamma_1})$, et deuxièmes échantillon $Y_1,\dots,Y_n$ la queue de distribution 1-G d'indice $(\frac{-1}{\gamma_2})$, avec Z_i définie par : $Z_i=X_i\wedge Y_i$

Et d'autre part les indicateurs de censure $\delta_i = \mathbb{I}_{\{X_i\leq Y_i\}}$, les Z_i sont des variables aléatoires indépendantes de loi H liée a F et G par la relation : $1-H(x)=(1-F(x))(1-G(x))$

Donc H est une fonction de répartition , la queue de distribution 1-H est a variation régulière d'indice $(\frac{-1}{\gamma})$ avec : $\gamma = \gamma_1\gamma_2/(\gamma_1+\gamma_2)$

$\widehat{P}$ estime $P = \frac{\gamma_2}{\gamma_1+\gamma_2}$ par conséquent $\hat{\gamma}_{z,k,n}^{(.)}$ estimateur de γ.

$\hat{\gamma}_{z,k,n}^{(.)}$ peut être n'importe quel estimateur non adapté a la censuré, en particulier l'estimateur de Hill $\hat{\gamma}_{z,k,n}^{(H)}$. pour adapter ce dernier dans le cas censuré nous allons diviser cet estimateur par la proportion de données non censurées des k plus grandes valeurs de z, Alors l'estimateur de Hill adaptée du indice de queue γ_1 est définie par : $\hat{\gamma}_1^{(c,H)} = \frac{\hat{\gamma}^H}{\widehat{P}}$

Où $\hat{\gamma}^H = \frac{1}{k}\sum_{i=1}^{k}\log Z_{n-i+1,n} - \log Z_{n-k,n}$

Et $\widehat{P} = \frac{1}{k}\sum_{i=1}^{k}\delta_{n-i+1,n}$

Alor $$\hat{\gamma}_1^{(c,H)} = \frac{\frac{1}{k}\sum_{i=1}^{k}\log Z_{n-i+1,n} - \log Z_{n-k,n}}{\frac{1}{k}\sum_{i=1}^{k}\delta_{n-i+1,n}}$$

Définition de la méthode de Monte-Carlo

La méthode de Monte-carlo, est un ensemble des techniques (04 techniques)probabilistes qui permettent de réduire la variance d'un échantillon de taille plus grandes que 1. __Principe__ : pour estimer f(X) on peut estimer $I = \mathbb{E}(f(X)) = \int_{\mathbb{R}^d} f(X)dx$. et nous pouvons approcher I par : $I \approx \frac{X_1+\dots+X_n}{n}$

1ere technique : Variable antithétique : $\hat{I} = \frac{1}{n}\sum_{i=1}^{n}\frac{f(X_i)+f(1-X_i)}{2}$

2eme technique : échantillonnage préférée : $\hat{I} = \int_{\mathbb{R}^d}\frac{g(x)f(x)}{h(x)}h(x)dx$

3eme technique : Variable de contrôle : $\hat{I} = \mathbb{E}(f(X)-h(X)) + \mathbb{E}(h(X))$

4eme technique : stratification : $\hat{I} = \sum_{k=1}^{K}\mathbb{P}[Z\in D_k]\mathbb{E}[h(X) \mid Z\in D_k]$

Simulation

Algorithme :

1) Simuler un échantillon
2) Calculer la variance de l'estimateur par Monte-carlo

On exécute le programme suivant sous R

Email : chaima9maths@gmail.com