Devoir de méthode de rééchantillonnage

Ranujan KATHIR

03 Mai 2022

1 Modèle de Statistiques

On considère un échantillon i.i.d. $X_1, X_2, ..., X_{140} \sim P$ ou P est une loi inconnue définie sur R de moment d'ordre 2.

2 Procédure d'Inférence

On souhaite estimer l'écart-type : σ.
On considère l'estimateur empirique :
$\hat{\sigma} = \sqrt{\frac{1}{139} \cdot \sum_{i=1}^{140} (x - \bar{x})^2}$
Notre objectif est d'estimer l'écart-type de la distribution $\hat{\sigma}$. Considérons $R = \hat{\sigma} - E_p(\hat{\sigma})$.

3 Méthode de rééchantillonnage

L'échantillon bootstrap est obtenu par tirage avec remise $X_1^*, .., X_{140}^* | X \sim P_n = \frac{1}{n} \sum_{i=1}^{140} \delta_{X_i}$. La réplication bootstrap de R : $R^* = \hat{\sigma}^* - E_p(\hat{\sigma}^*|X)$ où $\hat{\sigma}^*$ est l'écart-type de X^*. Si l'heuristique d'Efron est vérifiée $L_p(R) \approx L_p(R^*|X)$ alors un estimateur bootstrap de $E_p(R^2) = \hat{\sigma}$ est : $E_p(R^{*2}|X) = Var_p(\hat{\sigma}^*|X)$.
En pratique, on ne peut calculer cet estimateur.

4 Procédure de Monte-Carlo

On effectue une procédure de Monte-Carlo. On fixe B le nombre de répétitions de Monte-Carlo. Pour b=1,...,B faire :

1. générer x^{*^b} un échantillon bootstrap de X
2. calculer $\hat{\sigma^{*^b}}$ la réplication bootstrap de X

On retrouve comme estimateur de Monte-Carlo la valeur suivante :
$\hat{I_B} = \frac{1}{B-1} \sum_{i=1}^{B} (\hat{\sigma}^{*i} - \frac{1}{B} \sum_{j=1}^{B} \hat{\sigma}^{*j})^2$.

5 Contrôle de convergence

Avant de faire la procédure de Monte-Carlo, il faut choisir B de telle sorte que :

1. les moyennes ergodiques sont stabilisés;
2. l'erreur de Monte-Carlo est assez faible.

5.1 Moyenne ergodiques

L'estimateur de Monte-Carlo est consistent. En effet, on a : $\lim_{B\to\infty} \hat{I_B} = .$ Ci-dessous, un graphique permettant de vérifier que les moyennes ergodiques converge vers I_B.

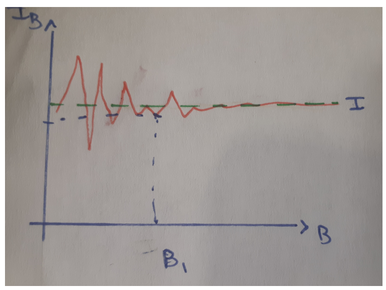

Figure 1: Convergence d'une moyenne ergodique

5.2 Erreur de Monte-Carlo

Développons la formule de $\hat{I_B}$ pour déterminer l'erreur de Monte-Carlo. On a :

$$\begin{aligned}
\hat{I_B} &= \frac{1}{B-1}\sum_{i=1}^{B}(\hat{\sigma}^{*i} - \frac{1}{B}\sum_{j=1}^{B}\hat{\sigma}^{*j})^2 \\
&= \frac{1}{B-1}[\sum_{i=1}^{B}(\sigma^{*^i})^2 - 2\sigma^{*^i}\frac{1}{B}\sum_{j=1}^{B}\sigma^{*^j} + \frac{1}{B^2}(\sum_{j=1}^{B}\sigma^{*^j})^2)] \\
&= \frac{1}{B-1}[\sum_{i=1}^{B}(\sigma^{*^i})^2 - 2\frac{1}{B^2}(\sum_{j=1}^{B}\sigma^{*^j})^2 + \frac{1}{B^2}(\sum_{j=1}^{B}\sigma^{*^j})^2)] \\
&= \frac{1}{B-1}[\sum_{i=1}^{B}(\sigma^{*^i})^2 - \frac{1}{B^2}(\sum_{j=1}^{B}\sigma^{*^j})^2)] \\
&= \frac{B}{B-1}[\frac{1}{B}\sum_{i=1}^{B}(\sigma^{*^i})^2 - \frac{1}{B}(\frac{1}{B}\sum_{j=1}^{B}\sigma^{*^j})^2)]
\end{aligned}$$

Cette décomposition nous sera utile pour déterminer l'erreur de Monte-Carlo. Posons :
$U_B = \frac{1}{B}\sum_{i=1}^{B}(\sigma^{*^i})^2$;
$V_B = \frac{1}{B}(\frac{1}{B}\sum_{j=1}^{B}\sigma^{*^j})^2$;

Déterminer l'erreur de Monte-Carlo pour U_B : il faut chercher un estimateur de l'écart-type ici :

$\sqrt{Var(U_B)} = \frac{\sqrt{Var((\sigma^{*^1})^2)}}{\sqrt{B}} \longrightarrow \frac{S_1}{\sqrt{B}}$ où S_1 est la variance empirique de $(\sigma^{*^1})^2$ lorsque n tends vers $+\infty$.

Concernant V_B, on détermine d'abord l'erreur de Monte-Carlo pour $\frac{1}{B}\sum_{j=1}^{B} \sigma^{*^j}$. Encore une fois, nous cherchons à déterminer l'écart-type, et nous trouvons : $\frac{S_2}{\sqrt{B}}$ où S_2 est la variance empirique de σ^{*^1}.

Par combinaisons, l'erreur de Monte-Carlo peut être déduit de la valeur suivante : $\sqrt{\frac{B}{B-1}[\frac{S_1}{\sqrt{B}} + \frac{1}{B}\frac{S_2^2}{B}]}$.