Feuille de TD 8

Estimation statistique, Intervalles de confiance

Toutes les variables aléatoires (v.a.) sont définies sur le même espace de probabilité $(\Omega, \mathcal{F}, \mathbb{P})$.

Exercice 1 (Partiel 2017)
Soit $\theta > 0$ un paramètre inconnu, et soit $(X_n)_{n\geq 1}$ une suite de variables aléatoires i.i.d. de loi de densité f_θ donnée par

$$f_\theta(x) = \frac{1}{2\sqrt{x\theta}}\mathbf{1}_{]0,\theta]}(x).$$

1. Vérifier que f_θ est une densité de probabilité.
2. On définit une suite de variables aléatoires $(\tilde{\theta}_n)_{n\geq 1}$ par $\tilde{\theta}_n = \frac{3(X_1+\cdots+X_n)}{n}$.
   (a) Calculer l'espérance et la variance de $\tilde{\theta}_n$.
   (b) Montrer que $\sqrt{n}(\tilde{\theta}_n - \theta)$ converge en loi vers une variable aléatoire dont on précisera la loi.
3. On définit un autre estimateur $\hat{\theta}$ par la formule suivante, pour $n \geq 1$,

$$\hat{\theta}_n = \max(X_1, \ldots, X_n).$$

   (a) Déterminer la fonction de répartition de $\hat{\theta}_n$.
   (b) Est-ce que $\hat{\theta}$ est sans biais ?
   (c) Est-ce que $\hat{\theta}$ est consistant ?
   (d) Quelle est la vitesse de $\hat{\theta}$?
   (e) Comparez les estimateurs $\hat{\theta}$ et $\tilde{\theta}$.
   (f) Comment modifier $\hat{\theta}$ pour en faire un estimateur sans biais ?

Exercice 2 Loi de Pareto
On considère une variable aléatoire X de loi de Pareto $\mathcal{P}(a)$ de densité

$$f_a(x) = \frac{a}{x^{a+1}}\mathbf{1}_{[1,+\infty[}(x).$$

On dispose de n observations $x_1, \ldots, x_n$ pour cette loi ; on cherche à estimer le paramètre a.

1. Déterminer l'estimateur du maximum de vraisemblance $\hat{a}_{MV}$ de a.
2. Montrer que la variable aléatoire $Z = \ln(X)$ suit une loi Gamma dont on déterminera les paramètres. En déduire la loi de $U = \sum_{i=1}^n \ln(X_i)$ lorsque $(X_1, \ldots, X_n)$ est un échantillon de loi de Pareto $\mathcal{P}(a)$.
3. Déterminer l'espérance et la variance de $\frac{1}{U}$.
4. Déterminer le biais de l'estimateur $\hat{a}_{MV}$. En déduire un estimateur non biaisé de a noté $\tilde{a}$ et déterminer la variance de cet estimateur.
5. Est-ce que $\tilde{a}$ est consistant ?
6. Quelle est la vitesse de $\tilde{a}$?

Exercice 3 Intervalles de confiance

Deux candidats sont en lice pour la prochaine élection, A et B. Un sondage donne A gagnant avec 55% des voix contre 45% pour son adversaire. On va construire un intervalle de confiance pour la proportion p de votants en faveur de A, au niveau $1-\alpha = 95\%$.

1. Modéliser ce problème et donner un estimateur consistant et sans biais $\hat{p}_n$ de cette proportion.
2. Posons $\sigma = \sqrt{p(1-p)}$. Montrer que pour tout $\epsilon > 0$ il existe $x > 0$ tel que la limite suivante existe et vérifie
$$\lim_{n\to\infty} \mathbb{P}\big(|\hat{p}_n - p| > \frac{\sigma x}{\sqrt{n}}\big) \leq \epsilon$$
3. Trouver une suite d'intervalles I_n dépendant de $\hat{p}_n$, de longueur tendant vers 0, telle que
$$\lim_{n\to\infty} \mathbb{P}\big(p \notin I_n\big) \leq 0.05.$$
Indication : Si F est la fonction de répartition de la loi normale centrée réduite, on a $2(1-F(1.96)) \leq 0.05$.
4. Supposons que 250 personnes ont été sondées. Calculer I_{250}. *N.B. : On l'appelle intervalle de confiance asymptotique au niveau 95%.*