ESTIMATION.

Estimation ponctuelle

Exercice 1.

On a vu que la variance empirique est un estimateur biaisé de la variance. Soit X une variable aléatoire d'espérance m et de variance σ^2 . On pose :

$$\bar{X}_n = \frac{1}{n} \sum_{k=1}^n X_k \quad S_n = \frac{1}{n} \sum_{k=1}^n (X_k - m)^2 \quad T_n = \frac{1}{n-1} \sum_{k=1}^n (X_k - \bar{X}_n)^2$$

1. Montrer que S_n est un estimateur sans biais de σ^2 . Quel est son inconvénient ?
2. Montrer que T_n est un estimateur sans biais de σ^2 .

Exercice 2.

D'après Romain Boillaud On dispose de $n > 2$ observations indépendantes $X_1, X_2, \dots, X_n$ de même loi de Poisson de paramètre $\lambda > 0$ inconnu. On souhaite estimer le paramètre $\theta = e^{-\lambda}$.

Pour tout $k \in \mathbb{N}^*$, on pose

$$Y_k = \begin{cases} 1 & \text{si } X_k = 0 \\ 0 & \text{sinon} \end{cases}$$

On introduit également :

$$\bar{Y}_n = \frac{1}{n} \sum_{k=1}^n Y_k \quad S_n = \sum_{k=1}^n X_k$$

1. (a) Montrer que Y_k suit une loi de Bernoulli de paramètre $e^{-\lambda}$.
(b) Calculer $E(Y_k)$ puis $E(\bar{Y}_n)$. En déduire que $\bar{Y}_n$ est un estimateur sans biais de θ .
(c) Calculer la variance de $\bar{Y}_n$ et en déduire que cet estimateur est convergent.
2. Expliquer pourquoi S_n suit une loi de Poisson et trouver son paramètre

3. On pose pour $i \in \mathbb{N}$

$$\varphi(i) = \mathbb{P}_{S_n=i}(X_1 = 0)$$

Montrer que pour $i \in \mathbb{N}$

$$\varphi(i) = \left(\frac{n-1}{n}\right)^i$$

4. On pose $T_n = \varphi(S_n)$

- (a) À l'aide du théorème de transfert, montrer que $E(T_n) = e^{-\lambda}$.
- (b) Montrer que T_n est un estimateur sans biais de θ .
- (c) Montrer $V(T_n) = e^{-2\lambda}(e^{\lambda/n} - 1)$ et en déduire que T_n est un estimateur convergent.

5. Comparer les risques quadratiques de T_n et $\bar{Y}_n$.

Estimation par intervalle de confiance

Exercice 3 (Bernoulli en utilisant l'inégalité de Bienaymé Tchebichev).

On suppose que $X \hookrightarrow \mathcal{B}(p)$ et que $(X_1, X_2, \dots, X_n)$ est un échantillon. On pose

$$\bar{X}_n = \frac{1}{n} \sum_{k=1}^n X_k$$

1. Donner l'espérance et la variance de $\bar{X}_n$. (cours)
2. Montrer que $\bar{X}_n$ est un estimateur sans biais et convergent vers p . (cours).
3. Montrer que $\mathbb{P}(|\bar{X}_n - p| \leq \varepsilon) \geq 1 - \frac{p(1-p)}{n\varepsilon^2}$
4. Soit $\alpha \in]0; 1[$ Trouver α tel que $1 - \frac{p(1-p)}{n\varepsilon^2} = 1 - \alpha$
5. Montrer que $\mathbb{P}\left(|\bar{X}_n - p| \leq \sqrt{\frac{p(1-p)}{n\alpha}}\right) \geq 1 - \alpha$.
6. Montrer que pour tout $p \in]0; 1[$ $p(1-p) \leq 1/4$. En déduire que
7. Montrer que $\mathbb{P}\left(|\bar{X}_n - p| \leq \frac{1}{2\sqrt{n\alpha}}\right) \geq 1 - \alpha$.
8. En déduire un intervalle de confiance de niveau de confiance $1 - \alpha$ de p .

Exercice 4 (Bernoulli : en utilisant le théorème limite-centrale).

On reprend les mêmes notations que dans l'exercice 3. On note Φ la fonction de répartition de la loi normale centrée réduite.

1. Pourquoi Φ est elle une fonction bijective de $]-\infty; +\infty[$ dans $]0; 1[$.
2. Exprimer $\overline{X}_n^*$ la variable réduite centrée associée à $\overline{X}_n$.
3. En citant u théorème du cours expliquer pourquoi

$$\lim_{n \rightarrow +\infty} \mathbb{P} \left(a \leq \sqrt{n} \frac{\overline{X}_n - p}{\sqrt{p(1-p)}} \leq b \right) = \Phi(b) - \Phi(a)$$

4. Soit $\alpha \in]0; 1[$, on pose $t_\alpha = \Phi^{-1}(1 - \alpha/2)$. Montrer que

$$\lim_{n \rightarrow +\infty} \mathbb{P} \left(\overline{X}_n - \frac{t_\alpha p(1-p)}{\sqrt{n}} \leq p \leq \overline{X}_n + \frac{t_\alpha p(1-p)}{\sqrt{n}} \right) = 1 - \alpha$$

5. en déduire que $\left[\overline{X}_n - \frac{t_\alpha}{2\sqrt{n}}; \overline{X}_n + \frac{t_\alpha}{2\sqrt{n}} \right]$ est un intervalle de confiance asymptotique au niveau $1 - \alpha$

Exercice 5.

On veut estimer la masse m d'un certain objet. Pour cela, on effectue des pesées successives et l'on note m la moyenne obtenue. On admet que la variable aléatoire renvoyant le résultat d'une pesée de l'objet étudié suit la loi normale $\mathcal{N}(m, \sigma^2)$ avec $\sigma = 0.1$. On effectue une suite de pesées et on note X_i le résultat de la i -ième pesée.

1. On note $\overline{X}_n = \frac{1}{n} \sum_{k=1}^n X_k$. Quelle est la loi suivie par $\overline{X}_n$?
2. Calculer l'espérance et la variance de $\overline{X}_n$.
3. On note $\overline{X}_n^* = \sqrt{n} \frac{\overline{X}_n - m}{\sigma}$. Quelle est la loi suivie par $\overline{X}_n^*$?
4. On effectue 10 pesées et on obtient une masse moyenne de 45,5 grammes.
5. Trouver un réel x tel que

$$\mathbb{P}(|\overline{X}_{10} - m| \leq \frac{\sigma}{\sqrt{10}} x) \geq 0.9$$

On donne $\Phi(1,65) \approx 0,95$

6. Montrer que $\left[45,5 - \frac{\sigma}{\sqrt{10}} x; 45,5 + \frac{\sigma}{\sqrt{10}} x \right]$ est un intervalle de confiance au niveau de confiance 90% pour la masse m .
7. Déterminer quelle valeur de n il faut choisir pour que l'intervalle de confiance de niveau de confiance 90% soit de longueur 0.05.