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 :

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

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.(on s'inspirera fortement du calcul vu dans le cours "premier estimateur de la variance"[1])

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

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

1. (a) Montrer que Y_k suit une loi de Bernouilli de paramètre $\mathrm{e}^{-\lambda}$.
   (b) Calculer $E(Y_k)$ puis $E(\overline{Y}_n)$. En déduire que $\overline{Y}_n$ est un estimateur sans biais de θ.
   (c) Calculer la variance de $\overline{Y}_n$ et en déduire que cette 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 $E(T_n) = \mathrm{e}^{-\lambda}$.
   (b) Montrer que T_n est un estimateur sans biais de θ.
   (c) Montrer $V(T_n) = \mathrm{e}^{-2\lambda}\left(\mathrm{e}^{\lambda/n} - 1\right)$ et en déduire que T_n est un estimateur convergent.
5. Comparer les risques quadratiques de T_n et $\overline{Y}_n$.

1. on pourrait même ne pas refaire tous ces calculs

Estimation par intervalle de confiance

Exercice 3 (Bernouilli en utilisant l'inégalité de Bienaymé Tchebichev)**.**
On suppose que $X \hookrightarrow \mathcal{B}(p)$ et que $(X_1, X_2, \ldots, X_n)$ est un échantillon. On pose

$$\overline{X_n} = \frac{1}{n}\sum_{k=1}^{n} X_k$$

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

Exercice 4 (Bernouilli : 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

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

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

$$\lim_{n\to+\infty} \mathbb{P}\left(\overline{X_n} - \frac{t_\alpha p(1-p)}{\sqrt{n}} \leqslant p \leqslant \overline{X_n} + \frac{t_\alpha p(1-p)}{\sqrt{n}}\right) = 1 - \alpha$$

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 le 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

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

   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.