Exercices, 1

EXERCICE 1 (Lemme de Slutsky et Delta-méthode). — On se donne $(X_n)_{n\geq 1}$ et $(Y_n)_{n\geq 1}$ deux suites de variables aléatoires réelles, telles que $X_n \underset{n\to\infty}{\Rightarrow} X$ et $Y_n \underset{n\to\infty}{\Rightarrow} c$, où X est une variable aléatoire réelle et c est une constante.

1. Montrer que $Y_n \xrightarrow[n\to\infty]{\mathbb{P}} c$.
2. Montrer que $(X_n, Y_n) \underset{n\to\infty}{\Rightarrow} (X, c)$, en déduire que $X_n + Y_n$, $X_n \times Y_n$ convergent respectivement en loi vers $X + c$ et cX.
3. Trouver (X_n, Y_n) tel que $X_n \underset{n\to\infty}{\Rightarrow} X$ et $Y_n \underset{n\to\infty}{\Rightarrow} Y$, mais $((X_n, Y_n))_{n\geq 1}$ ne converge pas en loi.
4. On suppose maintenant que $(X_n)_{n\geq 1}$ est une suite de variables aléatoires i.i.d, de moyenne μ et de variance σ^2. Soit g une fonction dérivable en μ, telle que $g'(\mu) \neq 0$. Montrer que

$$\sqrt{n}\left(g(\bar{X}_n) - g(\mu)\right) \underset{n\to\infty}{\Rightarrow} \mathscr{N}(0, \sigma^2 g'(\mu)^2),$$

   et en déduire un intervalle de confiance asymptotique au niveau 95% pour $g(\mu)$.

EXERCICE 2. — On se donne $Y_1, \ldots, Y_n$, i.i.d de moyenne μ et variance σ^2.

1. On suppose μ connu. Donner un estimateur non biaisé de σ^2.
2. On suppose μ inconnu. Calculer l'espérance de $\sum_{i=1}^n (Y_i - \bar{Y}_n)^2$. En déduire un estimateur non biaisé de σ^2.

EXERCICE 3. — Un actif financier rapporte R_n% au jour n; le rendement moyen de cet actif après n époques est donc $\varrho_n = \prod_{i=1}^n (1 + X_i)^{1/n} - 1$. On suppose pour simplifier que les R_i sont des variables aléatoires iid à valeurs dans $[-\varepsilon, \varepsilon]$ pour un certain $\varepsilon < 1$.

1. Montrer que ϱ_n converge en probabilité vers une constante r.
2. Trouver un intervalle de confiance à 95% pour r.

EXERCICE 4. — On suppose que l'on observe $X_1, \ldots, X_n$ i.i.d de loi $\mathscr{P}(\theta)$.

1. Montrer que $\bar{X}_n$ est un estimateur non-biaisé de θ. Est-il consistant? Quel est son risque quadratique?
2. Montrer que $\sqrt{\bar{X}_n} \xrightarrow[n\to\infty]{\mathbb{P}} \sqrt{\theta}$.
3. Donner deux intervalles de confiance au niveau 98% pour $\sqrt{\theta}$, et les comparer.

EXERCICE 5. — Soient $X_1, \ldots, X_n$, i.i.d. de loi

$$\mu_\theta(dx) = \exp(\theta - x)\mathbb{1}_{[\theta, +\infty[}(x)dx, \quad \theta > 0.$$

1. Calculer $\mathbb{E}_\theta[X_1]$ et en déduire un estimateur de θ que l'on notera $\hat{\theta}_n$.
2. Étudier le risque quadratique de l'estimateur $\hat{\theta}_n$. $\hat{\theta}_n$ est-il consistant ?
3. Donner un intervalle de confiance $I_1(\alpha)$ non-asymptotique pour θ au niveau de risque $0 < \alpha < 1$.

4. Construire un intervalle de confiance asymptotique $I_2(\alpha)$ pour θ à partir de $\hat{\theta}_n$.

5. Montrer que, pour $x > 1/\sqrt{2\pi}$, $\int_x^\infty \frac{e^{-\frac{t^2}{2}}}{\sqrt{2\pi}} dt \leq e^{-x^2/2}$, puis en déduire une majoration du quantile d'ordre $1-\beta$ de la loi $\mathscr{N}(0,1)$.

6. Comparer les longueurs des intervalles de confiance $I_1(\alpha)$ et $I_2(\alpha)$, pour $\alpha \to 0$.

7. Montrer que l'estimateur $\theta_n^\star := \min_{1\leq i\leq n} X_i$ est meilleur que $\hat{\theta}_n$, au sens du risque quadratique.

8. Donner un intervalle de confiance $I_3(\alpha)$ pour θ au niveau de risque α, basé sur $\theta_n^\star$.

9. Comparer les longueurs de $I_2(\alpha)$ et $I_3(\alpha)$ lorsque $n \to \infty$.

Exercice 6. — Au cours de la seconde guerre mondiale, l'armée alliée notait les numéros de série $X_1, \ldots, X_n$ de tous les tanks nazis capturés ou détruits, afin d'obtenir un estimateur du nombre total N de tanks produits.

1. Proposer un modèle pour le tirage de $X_1, \ldots, X_n$.

2. Calculer l'espérance de $\bar{X}_n$. En déduire un estimateur non biaisé de N. Indication: la loi de n tirages sans remise est échangeable.

3. Étudier la loi de $X_{(n)}$ et en déduire un estimateur non biaisé de N.

4. Proposer deux intervalles de confiance de niveau α. On pourra utiliser le fait que l'inégalité de Hoeffding s'applique également aux tirages sans remise.

Remarque d'ordre historique — Selon Ruggles et Broodie (1947, JASA), la méthode statistique a fourni comme estimation une production moyenne de 246 tanks/mois entre juin 1940 et septembre 1942. Des méthodes d'espionnage traditionnelles donnaient une estimation de 1400 tanks/mois. Les chiffres officiels du ministère nazi des Armements ont montré après la guerre que la production moyenne était de 245 tanks/mois.

Exercice 7. — 1. On considère $(X_1, \ldots, X_n)$ un échantillon de loi uniforme sur $]\theta, \theta+1[$.

   (a) Donner la densité de la loi de la variable

   $$R_n = X_{(n)} - X_{(1)},$$

   où $X_{(1)} = \min(X_1, \ldots, X_n)$ et $X_{(n)} = \max(X_1, \ldots, X_n)$.

   (b) Étudier les différents modes de convergence de R_n quand $n \to \infty$.

   (c) Étudier le comportement en loi de $n(1-R_n)$ quand $n \to \infty$.

2. Soit $X_1, \ldots, X_n$ un échantillon de loi $\mathscr{U}([0,\theta])$, on veut estimer $\theta > 0$.

   (a) Déterminer un estimateur de θ à partir de $\bar{X}_n$.
   On considère l'estimateur suivant : $X_{(n)} = \max_{1\leq i\leq n} X_i$.

   (b) Déterminer les lois limites de ces estimateurs. Que pouvez-vous dire des propriétés de ces estimateurs.

   (c) Comparer les performances des deux estimateurs.

   (d) Donner un intervalle de confiance non asymptotique de niveau $1-\alpha$ pour θ.