Feuille de TD 5

Exercice 1

Soit $X_1, \dots, X_n$ un échantillon d'espérance m et de variance σ^2. On note $\bar{X}_n = \frac{1}{n}\sum_{i=1}^n X_i$ l'estimateur de la moyenne.

1) Rappeler les deux expressions pour la variance de X_1.
2) En déduire un estimateur de la variance.
3) Calculer $\mathrm{Var}(\bar{X}_n)$ puis $\mathbb{E}(\bar{X}_n^2)$.
4) Calculer le biais de l'estimateur de la variance.

Exercice 2

Soit T une variable aléatoire réelle dont la densité est donnée par :

$$f(x) = \begin{cases} \dfrac{2x}{R^2} & \text{si} \quad 0 \leq x \leq R, \\ 0 & \text{sinon.} \end{cases}$$

(c'est la loi du maximum de 2 variables indépendantes de loi uniforme sur $[0, R]$). Soient aussi $T_1, \dots, T_n$ n variables aléatoires indépendantes de même loi que T. On pose

$$X_n = \frac{1}{n}\sum_{i=1}^n T_i \qquad \text{et} \qquad Y_n = \max_{1\leq i\leq n} T_i.$$

1) Calculer $\mathbb{E}[T]$ et $\mathrm{Var}(T)$.
2) Calculer $\mathbb{E}[X_n]$ et $\mathrm{Var}(X_n)$.
3) Montrer que $F_{Y_n}(x) = F_T(x)^n$. En déduire la densité de Y_n, puis son espérance.
4) Construire à partir de X_n et de Y_n respectivement deux estimateurs sans biais de R. Ces deux estimateurs sont-ils consistants ?

Exercice 3

Une entreprise automobile s'intéresse au nombre de défauts dans un de ses processus de fabrication industrielle. Pour n chaînes de production, on note $X_1, \dots, X_n$ le nombre de défauts, et on suppose que ce sont des variables aléatoires indépendantes de loi $\mathcal{P}(\lambda)$. On souhaite estimer le paramètre λ.

1) Donner la loi de l'échantillon, c'est-à-dire calculer $h(x_1, \dots, x_n) = \mathbb{P}(X_1 = x_1, \dots, X_n = x_n)$.
2) Calculer la vraisemblance $V(\lambda)$ et montrer que

$$\ln(V(\lambda)) = -n\lambda + \sum_{i=1}^n X_i \ln(\lambda) - \sum_{i=1}^n \ln(X_i!)$$

3) Calculer T_n l'estimateur du maximum de vraisemblance. Quel estimateur retrouve-t-on ?
4) Montrer que T_n est un estimateur sans biais et consistant du paramètre λ.

Exercice 4

On observe $X_1, \dots, X_n$ un échantillon de densité appartenant à la famille $(f_\theta)_{\theta>0}$ où

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

(c'est la loi du carré d'une variable de loi uniforme sur $[0, \sqrt{\theta}]$).

1) Calculer $\mathbb{E}(X_1)$ et $\mathrm{Var}(X_1)$.
2) Proposer deux estimateurs de θ, l'un par la méthode des moments et l'autre par la méthode du maximum de vraisemblance.
3) Quelles sont leurs propriétés ?

Exercice 5

On s'intéresse à la distribution de l'âge des femmes dans la population française et on la suppose gaussienne d'écart-type $\sigma = 23$ ans et de moyenne m, qu'on cherche à estimer. Pour cela, on demande leur âge à n françaises prises au hasard et on note $X_1, \dots X_n$ leur âge.

1) Quelle est la loi de la moyenne empirique $\bar{X}_n = \frac{1}{n}\sum_{i=1}^n X_i$?
2) Trouver en fonction de m, σ et n un intervalle dans lequel on a 90% de chances de trouver $\bar{X}_n$.
3) En déduire, en fonction de $\bar{X}_n$, σ et n un intervalle (aléatoire) dans lequel on a 90% de chances de trouver m, c'est-à-dire un intervalle de confiance au niveau 90% pour m.
4) Les âges des 15 personnes interrogées sont les suivants : 14 ans, 46 ans, 68 ans, 62 ans, 53 ans, 9 ans, 33 ans, 19 ans, 75 ans, 38 ans, 15 ans, 21 ans, 49 ans, 37 ans, 41 ans.
   Calculer l'intervalle de confiance dans ce cas. *Indication : le quantile d'ordre 0.95 de la loi normale standard vaut* $q = 1.65$.
5) Quelle est la précision de l'estimation fournie ? Quelle taille minimale d'échantillon faudrait-il avoir pour avoir une précision inférieure à une année ?

Exercice 6

On cherche à observer des mutations sur des cellules identiques en les soumettant à une source de radio-activité. À l'issue de chaque expérience, la cellule testée a une probabilité p d'avoir muté.

1) Si on a réalisé n expériences indépendamment les unes des autres, quelle est la loi du nombre S_n de cellules mutées parmi les n testées ?
2) Quelle est l'espérance de $\hat{p}_n = S_n/n$? Sa variance ?
3) À l'aide de l'inégalité de Bienaymé-Tchebichev, donner un intervalle de confiance pour p au niveau 95%. Application numérique : $S_n = 3$ et $n = 40$.

4) Quelle convergence en loi a-t-on pour $\hat{p}_n - p$? Peut-on déduire de cette convergence un intervalle de confiance asymptotique pour p ?

Exercice 7

On a testé le taux de cholestérol sanguin de 25 hommes et on a obtenu les valeurs suivantes

164 272 261 248 235 192 203 278 268 230 242 305 286

310 345 289 326 335 297 328 400 292 194 338 252

On suppose qu'on est dans un modèle gaussien. Donner un intervalle de confiance à 90% pour la moyenne du taux de cholestérol (on pourra utiliser que le quantile d'ordre 95% de la loi de Student $\mathcal{T}(24)$ vaut $t_{0,95} = 1.71$).

Exercice 8

On observe une population de bactéries, et on note $X_1, \ldots, X_n$ la durée de vie de chaque bactérie. On suppose qu'on peut les modéliser par des variables indépendantes de loi $\mathcal{E}(\theta)$.

1) Calculer $\mathbb{E}(X_1)$ et en déduire un estimateur $\hat{\theta}_n$ de θ ; montrer qu'il est consistant.
2) Déterminer l'estimateur du maximum de vraisemblance.
3) À l'aide de l'inégalité de Bienaymé-Tchebichev, donner un intervalle de confiance pour θ au niveau $1 - \alpha$.

Exercice 9

Un standard téléphonique reçoit N appels par heures où N suit une loi de Poisson de paramètre λ. Une étude portant sur 80 heures a donné un nombre moyen de 21 appels par heures. Déterminer l'intervalle de confiance asymptotique à 95% du nombre moyen d'appels par heure λ. *Indication : le quantile d'ordre 0.975 de la loi normale standard vaut* $q = 1.96$.

Exercice 10

Soit $(X_1, \ldots, X_n)$ un n-échantillon de loi $\mathcal{N}(0, 1/\theta)$, avec $\theta > 0$.

1) Quelle est la densité de X_1 ?
2) Avec la méthode des moments, construire un estimateur $\widehat{\theta}_n$ consistant de θ.
3) Soit $\alpha \in]0,1[$. Construire un intervalle de confiance asymptotique pour θ de niveau $1 - \alpha$, faisant intervenir $\widehat{\theta}_n$. On pourra utiliser sans démonstration que $\mathrm{Var}(X_1^2) = 2/\theta^2$.