Série 10

Exercice 1. Soit l'échantillon aléatoire $Y_1, \dots, Y_n \overset{\text{iid}}{\sim} \mathcal{U}(0, \theta)$, où $\theta > 0$.

(a) On propose d'estimer θ par la valeur maximale de l'échantillon aléatoire :

$$M_n = \max(Y_1, \dots, Y_n).$$

Remarque : on peut montrer que c'est l'estimateur du maximum de vraisemblance.

(i) Trouver la fonction de densité de M_n.

(ii) Calculer le biais de l'estimateur M_n.

(iii) Proposer un estimateur non-biaisé de θ, de la forme $\hat{\theta}_1 = aM_n$ pour $a \in \mathbb{R}$. Quelle est sa variance ?

(b) On propose un deuxième estimateur de θ, noté $\hat{\theta}_2$, basé sur la méthode des moments.

(i) Trouver $\hat{\theta}_2$.

(ii) Calculer son erreur quadratique moyenne.

(c) Parmi les estimateurs $\hat{\theta}_1$ et $\hat{\theta}_2$, lequel choisiriez-vous et pourquoi ?

(d) Montrer que $I_n = [M_n, \alpha^{-1/n} \times M_n)$ est un intervalle de confiance de niveau $1 - \alpha$ pour θ.

Exercice 2. Une entreprise d'automobiles veut publier des informations sur la consommation d'un nouveau modèle de voiture. Pour 12 voitures de ce modèle on mesure la consommation d'essence, en litres, pour parcourir 100 km, et on obtient les résultats suivants :

$$14.60, 11.21, 11.56, 11.37, 13.68, 15.07, 11.06, 16.58, 13.37, 15.98, 12.07, 13.22.$$

(a) Supposez que la consommation d'essence pour une voiture de ce modèle est une variable aléatoire $X \sim \mathcal{N}(\mu, 3.5)$ avec μ inconnu, et que les données récoltées sont indépendantes. Donnez un intervalle de confiance à 95 % pour μ (la moyenne empirique de l'échantillon est $\bar{x} = 13.31$).

On va maintenant abandonner l'hypothèse de variance connue (ce qui est plus réaliste). On considère toujours un modèle normal, mais cette fois on suppose que $X \sim \mathcal{N}(\mu, \sigma^2)$ avec μ et σ^2 inconnus. Les données récoltées sont toujours considérées indépendantes.

(b) Donnez un intervalle de confiance à 95 % pour μ (la variance empirique de l'échantillon est $\frac{1}{11} \sum_{i=1}^{12} (x_i - \bar{x})^2 = 3.69$).

(c) Donnez un intervalle de confiance à 90 % pour μ. Commentez la différence avec la partie (b).
Le seuil à utiliser est : $t_{11} = 1.796$.

(d) Que l'entreprise peut-elle faire pour obtenir un intervalle de confiance à 95 % plus étroit que celui obtenu dans la partie (b) ?

Exercice 3. Un contrôle de qualité est effectué sur la production de pailles d'une usine. Les pailles produites dans cette usine sont vendues dans des paquets de 100 pièces, et on voudrait estimer le pourcentage de pièces défectueuses dans un paquet.

On prélève 18 paquets aléatoirement et on compte le nombre de pièces défectueuses dans chaque paquet. Soit Y_i le nombre de pièces défectueuses dans le $i^{\text{ième}}$ paquet. On a observé les résultats suivants :

y_1	y_2	y_3	y_4	y_5	y_6	y_7	y_8	y_9	y_{10}	y_{11}	y_{12}	y_{13}	y_{14}	y_{15}	y_{16}	y_{17}	y_{18}
0	1	1	1	1	1	0	1	2	1	0	1	1	2	0	0	0	2

(a) Proposez un modèle statistique pour résoudre ce problème d'estimation.

(b) Donnez l'estimateur du maximum de vraisemblance pour le pourcentage cherché.

(c) Donnez un intervalle de confiance approximatif à 95 % pour le pourcentage cherché, basé sur l'approximation de la distribution de l'estimateur du maximum de vraisemblance pour un échantillon de grande taille.

Exercice 4. Soient $X_1, \ldots, X_n \overset{\text{iid}}{\sim} \text{Exp}(\lambda)$ pour $\lambda > 0$, et $X_{(1)} = \min(X_1, \ldots, X_n)$. Montrer que $\lambda n X_{(1)} \sim \text{Exp}(1)$. Expliquer pourquoi il s'agit d'un pivot et donner un intervalle de confiance de niveau $1 - \alpha$ pour λ basé sur ce pivot.