Série 9

Exercice 1. Vous voulez connaître le pourcentage de femmes parmi les étudiant-e-s de l'EPFL/UNIL, mais vous n'avez pas accès à la liste d'étudiant-e-s. Vous décidez donc d'utiliser ce que vous avez appris dans le cours afin d'estimer ce pourcentage.

(a) Quel sera votre modèle statistique ?

(b) Quel sera le paramètre d'intérêt ?

(c) Comment choisissez-vous votre échantillon ?

(d) Quel sera votre estimateur ? (un choix intuitif est suffisant ici)

(e) L'estimateur est une variable aléatoire. Comment voit-on cela dans votre situation ?

(f) Etant une variable aléatoire, votre estimateur a une espérance et une variance. Supposez que le vrai pourcentage est de 40 % (valeur fictive !), et que vous avez un échantillon de taille 100. Calculez l'espérance et la variance de votre estimateur. Comment ces valeurs changent-elles en fonction de la taille de l'échantillon ? Est-ce que votre estimateur est biaisé (c'est-à-dire, est-ce que l'espérance de votre estimateur est différente de la vraie valeur 40%) ?

(g) Si le vrai pourcentage est de 40 %, quelle est la taille de l'échantillon dont vous avez besoin afin d'être certain à 95 % que votre estimateur sera plus petit que 50 % ? Indication : utilisez une loi approximative.

Exercice 2. Soit X_1 une variable aléatoire dont la fonction de densité est $f(x) = cx^{\theta-1}$ pour $x \in (0, 1)$ et $f(x) = 0$ autrement. Ici c est une constante et $\theta > 0$ est un paramètre inconnu.

(a) Calculez c.

(b) Calculez l'espérance de X_1.

Maintenant, considérons un échantillon $X_1, X_2, \ldots, X_n$ de variables de cette distribution.

(c) Trouvez $\hat{\theta}_{ML}$, l'estimateur du maximum de vraisemblance de θ.

(d) Trouvez $\widehat{\theta}_{MOM}$, l'estimateur par la méthode des moments.

Exercice 3. Soient $X_1, X_2, \ldots, X_n \overset{iid}{\sim} \mathcal{B}(p)$ avec $p \in (0, 1)$.

(a) Soit $x_1, x_2, \ldots, x_n$ une réalisation de cet échantillon. Ecrivez la fonction de vraisemblance de p basée sur cette réalisation.
Rappel : pour une variable $X \sim \mathcal{B}(p)$ on a $P(X = x) = p^x(1 - p)^{1-x}$ pour $x = 0, 1$.

(b) Trouvez $\hat{p}_{MOM}$, l'estimateur des moments de p.

(c) Trouvez $\hat{p}_{ML}$, l'estimateur du maximum de vraisemblance de p.

(d) Les estimateurs définis ci-dessus sont-ils biaisés (c'est-à-dire, est-ce que leur espérance est égale à ou différente de p) ? Trouvez leur variance.

Exercice 4. On considère une variable aléatoire Y suivant une loi exponentielle de paramètre $\lambda > 0$. On veut estimer la valeur de λ à partir d'un échantillon de n observations $Y_1, \ldots, Y_n$ indépendantes et identiquement distribuées.

(a) Calculer l'espérance et la variance de Y.

(b) Quelle est la distribution approximative de $\overline{Y}_n = \sum_{i=1}^n Y_i/n$ pour n très grand ?

(c) Trouver l'estimateur du maximum de vraisemblance de λ.

(d) Quelle est la distribution approximative de l'estimateur $\hat{\lambda}_{ML}$? *Indice* : méthode delta.