Chapitre14

FEUILLE D'EXERCICES

ESTIMATION

Estimation ponctuelle

EXERCICE 14.1. Comparaison de quelques estimateurs
Soit X de loi $\mathscr{U}([0;a])$ et $(X_1,\cdots,X_n)$ un n-échantillon de X. On cherche à estimer a. On note $\overline{X_n}$ la moyenne empirique de X.

1. Soit $T_n = 2\overline{X_n}$. Montrer que T_n est un estimateur sans biais de a et calculer son risque quadratique.
2. Soit $T'_n = \max(X_1,\cdots,X_n)$.

   Donner la fonction de répartition de X, puis déterminer celle de T'_n.

   En déduire une densité de T'_n puis son biais et son risque quadratique.
3. Soit $T''_n = \dfrac{n+1}{n}T'_n$. Déterminer son biais et son risque quadratique.
4. Pour de grandes valeurs de n, quelle est le meilleur estimateur de a ?

EXERCICE 14.2. D'après ESC 2008
Un joueur A dispose d'une pièce qui a la propriété de faire PILE avec la probabilité $\dfrac{1}{3}$.
Un joueur B dispose d'une pièce qui a la propriété de faire PILE avec la probabilité $p \in]0,1[$.
Les résultats des lancers de ces pièces seront toujours supposés indépendants.

1. Montrer que la probabilité que les lancers de A et de B soient différents est $\dfrac{1+p}{3}$.
   On procède alors au jeu suivant (N est un entier naturel fixé non nul) :
   Les joueurs A et B lancent leur pièce N fois de suite.
   Le joueur B paye un euro à A à chaque fois que les pièces n'affichent pas le même résultat.
   On note H_N la variable aléatoire égale à la somme payée par le joueur B au joueur A.
2. Montrer que H_N suit une loi classique que l'on détaillera.
3. Montrer que $\left(\dfrac{3H_N}{N}-1\right)$ est un estimateur sans biais du réel p et déterminer son risque quadratique.

EXERCICE 14.3. Maximum de vraisemblance pour la loi de Poisson
On considère un n-échantillon d'une loi de Poisson de paramètre λ que l'on cherche à estimer, ainsi que $(x_1,\dots,x_n) \in \mathbb{R}^n$ un n-uplet fixé.
On note :

$$s_n = \sum_{i=1}^{n} x_i \quad \text{et} \quad p_n = \prod_{i=1}^{n}(x_i)!.$$

Enfin pour tout $\theta > 0$, on pose

$$L_n(\theta) = \prod_{i=1}^{n} P_\theta(X_i = x_i).$$

1. Exprimer L_n en fonction de θ, s_n et p_n.

2. Montrer que L_n est dérivable sur $\mathbb{R}_+^*$ et vérifier que L_n est maximale en $\theta^* = \frac{s_n}{n}$.
3. En déduire l'expression du maximum de vraisemblance pour la loi de Poisson.

EXERCICE 14.4. Maximum de vraisemblance pour la loi géométrique
Reprendre les démarches précédentes pour montrer que l'estimateur du maximum de vraisemblance de la loi $\mathscr{G}(p)$ pour un n-échantillon $(X_1,...,X_n)$ est donné par la formule

$$\hat{\theta}_n = \frac{n}{X_1 + ... + X_n}.$$

EXERCICE 14.5. Maximum de vraisemblance d'une loi à densité
On considère un n-échantillon de la loi $\mathscr{U}([0;\theta])$ dont on cherche à estimer le paramètre $\theta > 0$.
Soit $(x_1,...,x_n) \in (\mathbb{R}_+^*)^n$ fixé. On note f_θ la d.d.p. de la loi $\mathscr{U}([0;\theta])$.
Pour tout $\theta > 0$, on note :

$$L_n(\theta) = \prod_{i=1}^{n} f_\theta(x_i).$$

1. Montrer que, pour tout $\theta > 0$, on a $L_n(\theta) = \begin{cases} \theta^{-n} & \text{si } \theta \geqslant \max(x_1,...,x_n) \\ 0 & \text{sinon} \end{cases}$
2. En déduire que l'estimateur du maximum de vraisemblance pour la loi $\mathscr{U}([0;\theta])$ est donnée par :

$$\hat{\theta}_n = \max(X_1,...,X_n).$$

Estimation par intervalle de confiance

EXERCICE 14.6. Intervalle de confiance avec IBT
On suppose que le paramètre p, qui exprime la probabilité qu'un individu contagieux transmette le virus à un individu sain, est inconnu, et on cherche à l'estimer. On rappelle que : $q = 1 - p$.
Pour m entier supérieur ou égal à m, on considère un m-échantillon $(Y_1, Y_2, \ldots, Y_m)$ de variables aléatoires indépendantes, de même loi de Bernoulli de paramètre p, définies sur un espace probabilisé $(\Omega, \mathscr{A}, \mathrm{P})$ et on considère $\overline{Y_m}$ l'estimateur moyenne empirique.

1. (a) Rappeler l'expression de $\overline{Y_m}$ puis montrer que $\overline{Y_m}$ est un estimateur sans biais de p;
   (b) Déterminer son risque quadratique.
2. À l'aide de l'inégalité de Bienaymé-Tchebycheff, montrer que l'intervalle $\left[\overline{Y_m} - \sqrt{\frac{5}{m}}, \overline{Y_m} + \sqrt{\frac{5}{m}}\right]$ est un intervalle de confiance de p au niveau de confiance $0,95$.

EXERCICE 14.7. Intervalle de confiance asymptotique avec le TLC
Lors d'un sondage sur 100 personnes interrogées, 60 pensent voter pour A.
On modélise le choix par un échantillon $(X_1, \cdots, X_{100})$ de variables indépendantes de même loi de Bernoulli de paramètre p.
On cherche à déterminer un intervalle de confiance pour p au niveau de confiance 99 %.

1. Déterminer l'espérance et la variance de la moyenne empirique $F = \dfrac{1}{100}\displaystyle\sum_{i=1}^{100} X_i$.
2. On note F^* la variable centrée réduite associée à F.
   Par quelle loi peut-on approcher celle de F^* ? On suppose désormais que F^* suit la loi $\mathscr{N}(0,1)$.
3. Déterminer à l'aide d'une table de la fonction Φ un réel t tel que $P(-t \leqslant F^* \leqslant t) \geqslant 0,99$ et en déduire que

$$P\left(F - t\frac{\sqrt{p(1-p)}}{10} \leqslant p \leqslant F + t\frac{\sqrt{p(1-p)}}{10}\right) \geqslant 0,99$$

4. Montrer que pour tout $p \in [0;1]$, $p(1-p) \leqslant \dfrac{1}{4}$ et en déduire un intervalle de confiance pour p au niveau de confiance $0,99$ puis en donner une estimation.

EXERCICE 14.8. Type concours : oral HEC 2012

Soit X une variable aléatoire définie sur un espace probabilisé $(\Omega, \mathscr{A}, P)$, à valeurs dans $[0, \theta]$ où θ est un paramètre réel strictement positif inconnu. Une densité f de X est donnée par : $f(x) = \begin{cases} \frac{2x}{\theta^2} & \text{si } x \in]0, \theta]. \\ 0 & \text{sinon} \end{cases}$

1. Question de cours : Estimateur sans biais ; risque quadratique d'un estimateur.
2. Calculer l'espérance et la variance de X. Pour n entier de $\mathbb{N}^*$, soit $(X_1, X_2, \ldots, X_n)$ un n-échantillon de variables aléatoires indépendantes et de même loi que X. On pose pour tout $n \in \mathbb{N}^*$: $\overline{X}_n = \dfrac{1}{n} \displaystyle\sum_{i=1}^{n} X_i$
3. (a) Déterminer la fonction de répartition F de X.
   (b) Tracer dans un repère orthogonal, l'allure de la courbe représentative de F.
4. (a) Déterminer un estimateur T_n de θ, sans biais et de la forme $c\overline{X}_n$, où c est un réel que l'on précisera.
   (b) Quels sont les risques quadratiques respectifs associés aux estimateurs $\overline{X}_n$ et T_n de θ ?
5. On pose pour tout $n \in \mathbb{N}^*$: $M_n = \max(X_1, X_2, \ldots, X_n)$.
   (a) Déterminer la fonction de répartition G_n et une densité g_n de M_n.
   (b) Calculer l'espérance de M_n . En déduire un estimateur sans biais W_n de θ.
   (c) Entre T_n et W_n , quel estimateur doit-on préférer pour estimer θ ?
6. Soit α un réel donné vérifiant $0 < \alpha < 1$
   (a) Etablir l'existence de deux réels a et b tels que $0 < a < 1$ et $0 < b < 1$, vérifiant $P(M_n \leqslant a\theta) = \alpha/2$ et $P(b\theta \leqslant M_n \leqslant \theta) = \alpha/2$.
   (b) En déduire un intervalle de confiance pour le paramètre θ au niveau de confiance $1 - \alpha$.