**Corrigé du TD 2**
Février 2011

**Exercice 1.** Un estimateur linéaire et sans biais de $\theta$ s'écrit sous la forme

$$\hat{\theta}_n = \sum_{i=1}^{n} a_i X_i, \text{ avec } \sum_{i=1}^{n} a_i = 1.$$

Comme les variables aléatoires sont indépendantes, on a

$$\mathrm{Var}(\hat{\theta}_n) = \sum_{i=1}^{n} a_i^2 \sigma^2 = \sigma^2 \sum_{i=1}^{n} a_i^2, \text{ avec } \sigma^2 = \mathrm{Var}(X_1).$$

Notre but ici est de trouver un estimateur linéaire ayant la variance la plus faible possible, donc chercher la/les $(a_1, ..., a_n)$ qui minimise $\mathrm{Var}(\hat{\theta}_n)$.
D'après l'inégalité de Cauchy-Schwarz, on a

$$\left(\sum_{i=1}^{n} a_i^2\right)\left(\sum_{i=1}^{n} \frac{1}{n^2}\right) \geq \left(\sum_{i=1}^{n} \frac{a_i}{n}\right)^2$$

Puisque $\sum_{i=1}^n a_i = 1$, on déduit que $\sum_{i=1}^n a_i^2 \geq \frac{1}{n}$ avec égalité si et seulement si $a_i = \frac{1}{n}$ pour tout $i \in \{1, ..., n\}$. L'estimateur de variance minimale dans la classe des estimateurs linéaires et sans biais est donc la moyenne empirique.

**Remarque** : Ici, pour trouver les $(a_1, ..., a_n)$ qui minimisent $\mathrm{Var}(\hat{\theta}_n)$, on aurait pu chercher les candidats $(a_1, ..., a_n)$ parmi ceux qui annulent les dérivées partielles $\frac{\partial \mathrm{Var}(X_1)}{\partial a_i}$. Mais pour cela, on doit d'abord se débarrasser de la contrainte $\sum_{i=1}^n a_i = 1$, en exprimant $a_n = 1 - \sum_{i=1}^{n-1} a_i$ et en injectant cette expression de $a_n$ dans celle de $\mathrm{Var}(\hat{\theta}_n)$, puis minimiser cette dernière par rapport à $(a_1, ..., a_{n-1})$, ce qui nous donnerait $a_1 = ... = a_{n-1} = \frac{1}{n}$ comme seul candidat. Grâce à la contrainte, on obtiendrait également $a_n = \frac{1}{n}$ !

**Exercice 2.** On considère le modèle d'échantillonnage $X_1 ..., X_n$ de taille $n$ associé à la famille de lois exponentielles $P = \{\mathcal{E}(\lambda), \lambda > 0\}$. On veut estimer $\lambda$

**2.1. A partir de la méthode des moments, construire un estimateur convergent $\hat{\lambda}_n$ de $\lambda$.**

La loi forte des grands nombres implique que $(\bar{X}_n, n = 1)$, où $\bar{X}_n = \frac{1}{n}\sum_{i=1}^n X_i$, converge p.s. vers $\mathbb{E}[X_1] = \frac{1}{\lambda}$. Par continuité de la fonction : $x \to \frac{1}{x}$ sur $]0, \infty[$, on en déduit que $\hat{\lambda}_n = 1/\bar{X}_n$ est un estimateur convergent de $\lambda$.

**2.2. Vérifier qu'il s'agit de l'estimateur du maximum de vraisemblance.**

La log-vraisemblance est donnée par

$$l_n(x;\lambda) = \log L_n(x;\lambda) = \sum_{i=1}^{n} \log\left(\lambda e^{-\lambda x_i} 1_{x_i>0}\right) = n\log\lambda - \lambda\sum_{i=1}^{n} x_i + \log \Pi_{i=1}^n 1_{x_i>0}.$$

Pour calculer l'EMV (estimateur du maximum de vraisemblance), on cherche les zéros de la dérivée de la log-vraisemblance $L_n$ :

$$\frac{\partial L}{\partial \lambda}_n(x,\lambda) = \frac{n}{\lambda} - \sum_{i=1}^{n} x_i = 0 \Leftrightarrow \lambda = \frac{n}{\sum_{i=1}^{n} x_i}.$$

Comme $\lim_{\lambda \to 0} L_n(x,\lambda) = \lim_{\lambda \to \infty} L_n(x,\lambda) = -\infty$, on en déduit que la log-vraisemblance est maximale pour $\lambda = \frac{n}{\sum_{i=1}^{n} x_i}$. L'estimateur du maximum de vraisemblance est donc $\hat{\lambda}_n$.

**2.3. Déterminer la loi de $\sum_{i=1}^{n} X_i$ . Calculer $E_\lambda[\hat{\lambda}_n]$. L'estimateur est-il sans biais ?**

En utilisant les fonctions caractéristiques, on vérifie que la loi de $\sum_{i=1}^{n} X_i$ est la loi $\Gamma_n(\lambda)$ sous $P_\lambda$. On obtient pour $n \geq 2$

$$\begin{aligned}
E_\lambda\left[\frac{1}{\sum_{i=1}^{n} X_i}\right] &= \int_0^\infty \frac{1}{s} \frac{\lambda^n}{(n-1)!} s^{n-1} e^{-\lambda s} ds \\
&= \frac{\lambda}{(n-1)} \int_0^\infty \frac{\lambda^{n-1}}{(n-2)!} s^{n-2} e^{-\lambda s} ds \\
&= \frac{\lambda}{(n-1)}
\end{aligned}$$

où on a identifié la densité de $\Gamma(n-1,\lambda)$ dans l'intégrale. Donc on a $E_\lambda[\hat{\lambda}_n] = \frac{n}{(n-1)}\lambda$. L'estimateur $\hat{\lambda}_n$ est donc biaisé pour $n \geq 2$ (Pour $n = 1$, on vérifie que $E_\lambda[\hat{\lambda}_1] = \infty$.)

**2.4. Déterminer un estimateur $\hat{\lambda}_n^*$ sans biais et un estimateur $\hat{\lambda}_n^o$ qui minimise le risque quadratique parmi les estimateurs $\hat{\lambda}_n^{(c)} = \frac{c}{\sum_{i=1}^{n} X_i}$, où $c > 0$**

Les calculs précédents nécessitent de supposer $n \geq 2$ et donnent comme estimateur sans biais

$$\hat{\lambda}_n^* = \frac{n-1}{n}\hat{\lambda}_n = \frac{n-1}{\sum_{i=1}^{n} X_i}.$$

On calcule le deuxième moment de la même manière que le premier, pour $n \geq 3$

$$\begin{aligned}
E_\lambda\left[\frac{1}{(\sum_{i=1}^{n} X_i)^2}\right] &= \frac{\lambda^2}{(n-1)(n-2)} \int_0^\infty \frac{\lambda^{n-2}}{(n-3)!} s^{n-3} e^{-\lambda s} ds \\
&= \frac{\lambda^2}{(n-1)(n-2)}.
\end{aligned}$$

Donc pour tout $c > 0$, on a

$$\begin{aligned}
R(\hat{\lambda}_n^{(c)},\lambda) &= E_\lambda\left[\left(\frac{c}{(\sum_{i=1}^{n} X_i)} - \lambda\right)^2\right] \\
&= \frac{\lambda^2}{(n-1)(n-2)}\left[c^2 - 2(n-2)c + (n-1)(n-2)\right).
\end{aligned}$$

Le risque quadratique est minimal pour $2c - 2(n-2) = 0$ soit $c = n-2$. Parmi les estimateurs $\hat{\lambda}^{(c)}$, c'est donc $\hat{\lambda}^{(o)} = \hat{\lambda}^{(n-2)}$ qui minimise le risque quadratique. Pour $n \leq 2$ le risque quadratique est infini.

**Exercice 3.** Soit $(X_1, ..., X_n)$ un chantillon i.i.d. de loi uniforme sur $]\theta, 2\theta[$ où $\theta > 0$.
**3.1. Estimer $\theta$ par la méthode des moments**

$X$ de loi uniforme sur $]\theta, 2\theta[$, donc admet comme densité la fonction : $x \mapsto \frac{1}{\theta} 1_{]\theta,2\theta[}(x)$. Sa moyenne vaut

$$E[X] \;=\; \frac{1}{\theta}\int_{\theta}^{2\theta} x dx \;=\; \frac{1}{2\theta}[4\theta^2 - \theta^2] = \frac{3}{2}\theta$$

En résolvant $\frac{3}{2}\theta = E[X]$, on obtient l'estimateur donné par la méthode des moments, soit

$$\hat{\theta}_n = \frac{2}{3}\bar{X}_n, \text{ avec } \bar{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i$$

**3.2. Déterminer l'estimateur du maxi de vraisemblance $\phi$ de $\theta$ et calculer la constante $k$ telle que $E_\theta[k\phi] = \theta$**

Une vraisemblance $L_n$ s'écrit $\forall x = (x_1, ..., x_n) \in \mathbb{R}^n, \forall \theta > 0$,

$$\begin{aligned} L_n(x,\theta) &= \Pi_{i=1}^{n} \frac{1}{\theta} 1_{]\theta,2\theta[}(x_i) \\ &= \frac{1}{\theta^n} 1_{]\theta,\infty[}(x_{(1)}) 1_{]0,2\theta[}(x_{(n)}) \\ &= \frac{1}{\theta^n} 1_{]\frac{1}{2}x_{(n)},x_{(1)}[}(\theta). \end{aligned}$$

où $x_{(1)}$ et $x_{(n)}$ sont respectivement le min et le max de l'ensemble $\{x_1, ..., x_n\}$.
Comme $\theta \mapsto \frac{1}{\theta^n}$ est une fonction décroissante sur $\mathbb{R}_+^*$, la vraisemblance est maximisée au point $\theta = \frac{1}{2}x_{(n)}$. Ainsi, l'estimateur du maximum de vraisemblance est $\phi = \frac{1}{2}X_{(n)}$, avec $X_{(n)} = \max\{X_1, ..., X_n\}$.
Or, la statistique $X_{(n)}$ admet comme densité (exercice de proba classique pour retrouver la loi densité de $X_{(n)} = \max_{i=1}^{n} X_i$),

$$y \mapsto \frac{n}{\theta^n}(y-\theta)^{n-1} 1_{[\theta,2\theta[}(y).$$

Donc, son espérance vaut

$$E_\theta[X_{(n)}] = \int_{\theta}^{2\theta} y \frac{n}{\theta^n}(y-\theta)^{n-1} dy$$

ce qui après intégration par parties donne

$$
\begin{aligned}
E_\theta[X_{(n)}] &= \frac{n}{\theta^n}\left\{\left[y\frac{(y-\theta)^n}{n}\right]_\theta^{2\theta} - \int_\theta^{2\theta}\frac{(y-\theta)^n}{n}d\theta\right\} \\
&= 2\theta - \frac{1}{\theta^n}\left[\frac{(y-\theta)^{n+1}}{n+1}\right]_\theta^{2\theta} \\
&= 2\theta - \frac{\theta}{n+1} \\
&= \frac{2n+1}{n+1}\theta.
\end{aligned}
$$

Par conséquence, $E_\theta[\phi] = \frac{1}{2}\frac{2n+1}{n+1}\theta$ et $k = \frac{2n+2}{2n+1}$ donne $E_\theta[k\phi] = \theta$.

**Exercice 4. Pour un échantillon i.i.d. $(X_1, ..., X_n)$ d'une loi de Bernoulli de paramètre inconnu $\theta \in [0,1]$, montrez que la moyenne empirique $\bar{X}_n : (x_1, ..., x_n) \to \frac{1}{n}\sum_{i=1}^{n} x_i$ est le seul estimateur sans biais de $\theta$ fonction de la somme $\Sigma_n : (x_1, ..., x_n) \to \sum_{i=1}^{n} x_i$**

On sait déjà que $\bar{X}_n$ est un estimateur sans biais pour estimer $\theta$ ($E\left[\bar{X}_n\right] = \theta$). En effet,d'après cours de Proba, $\Sigma_n = \sum_{i=1}^{n} x_i$ suit la loi binomiale $\mathcal{B}(n,\theta)$. Ainsi, $E\left[\Sigma_n\right] = n\theta$, d'où

$$
\begin{aligned}
E\left[\bar{X}_n\right] &= \frac{1}{n}E\left[\Sigma_n\right] \\
&= \sum_{x=0}^{n}\frac{x}{n}C_n^x\theta^x(1-\theta)^{n-x} = \theta \text{ (on calcule ici l'espérance d'une loi binomiale.)} \qquad (1)
\end{aligned}
$$

Montrons maintenant que $\bar{X}_n$ le seul estimateur sans biais de $\theta$, fonction de $\Sigma_n$.
Supposons maintenant qu'il existe une fonction $\psi$ de $\{0, ..., n\}$ dans $\mathbb{R}$, telle que $\psi(\Sigma_n)$ est un estimateur sans biais de $\theta$, c'est-à-dire :

$$
E\left[\psi(\Sigma_n\right] = \sum_{x=0}^{n}\psi(x)C_n^x\theta^x(1-\theta)^{n-x} = \theta \qquad (2)
$$

En utilisant (1) et (2), on obtient $\forall\theta \in [0,1[$ :

$$
\sum_{x=0}^{n}\left(\psi(x) - \frac{x}{n}\right)C_n^x\left(\frac{\theta}{1-\theta}\right)^x = 0 \qquad (3)
$$

Lorsque $\theta$ décrit $[0,1[$, le rapport $\frac{\theta}{1-\theta}$ varie dans tout $\mathbb{R}^+$. Ainsi l'égalité (3) exprime que le polynôme de degré $n$ de coefficients $\left(\psi(x) - \frac{x}{n}\right)C_n^x$ est identiquement nul. Ceci signifie que ses coefficients sont nuls et donc que :

$$
\forall x \in \{0, 1, ..., n\}\ \psi(x) = \frac{x}{n},
$$

soit finalement que $\psi(\Sigma_n) = \frac{\Sigma_n}{n} = \bar{X}_n$.