Université de Strasbourg — Année 2010/2011
L3 - Statistique Mathématique
Suivi pédagogique de Gilles Stupfler d'après les
Cours effectués par Ségolen Geffray et les
TD effectués par Gilles Stupfler

---

**Eléments de réponses – Examen de la session d'automne 2010/2011**
Les résultats doivent tous êtres démontrés, les réponses doivent toutes être rédigées.

---

**Exercice 1.**
Soit $X$ la variable aléatoire représentant le poids après remplissage d'un contenant pris au hasard. On suppose que $X$ suit une loi normale $\mathcal{N}(m, \sigma^2)$. Soit $(X_1, ..., X_n)$ un échantillon i.i.d. de variable parente $X$.

1. Dans le cas de la loi normale, l'estimateur de $\theta = (m, \sigma^2)$ de la méthode des moments coïncide avec l'estimateur de la méthode du maximum de vraisemblance et vaut $\widehat{\theta} = (\widehat{m}, \widehat{\sigma^2})$ avec $\widehat{m} = \frac{1}{n}\sum_{i=1}^{n} X_i$ et $\widehat{\sigma^2} = \frac{1}{n}\sum_{i=1}^{n}(X_i - \overline{X}_n)^2$.
   $\widehat{m}$ est sans biais pour $m$, $\widehat{\sigma^2}$ est biaisé pour $\sigma^2$ mais asymptotiquement non biaisé (ce que l'on obtient en calculant l'espérance de l'estimateur).
   De plus, $\widehat{\theta}$ est fortement consistant pour $\theta$ (appliquer la loi forte des grands nombres et le théorème de la transformation continue).
   Enfin, $\widehat{\theta}$ est l'EMV de $\theta$ calculé à partir d'observations i.i.d. dans un modèle régulier (puisque la loi normale appartient à la famille exponentielle en dimension 2) et est fortement consistant donc le TCL pour l'EMV s'applique, ce qui permet d'affirmer que $\widehat{\theta}$ est asymptotiquement efficace.
   Ici, les observations fournissent les estimations $\widehat{m} = 360g$ et $\widehat{\sigma^2} = 100g^2$.
2. On veut estimer
   $$p_1 = \mathbb{P}[340 < X < 360] = \mathbb{P}[\frac{340 - m}{\sqrt{\sigma^2}} < \frac{X - m}{\sqrt{\sigma^2}} < \frac{360 - m}{\sqrt{\sigma^2}}] = \phi\left(\frac{360 - m}{\sqrt{\sigma^2}}\right) - \phi\left(\frac{340 - m}{\sqrt{\sigma^2}}\right)$$
   car $\phi$ est continue. On propose l'estimateur
   $$\widehat{p_1} = \phi\left(\frac{360 - \widehat{m}}{\sqrt{\widehat{\sigma^2}}}\right) - \phi\left(\frac{340 - \widehat{m}}{\sqrt{\widehat{\sigma^2}}}\right)$$
   ce qui donne l'estimation $\widehat{p_1} = \phi(0) - \phi(-2) = \phi(2) - 0.5$
3. On veut estimer
   $$N = 1000\mathbb{P}[X < 330] = 1000\mathbb{P}[\frac{X - m}{\sqrt{\sigma^2}} < \frac{330 - m}{\sqrt{\sigma^2}}] = 1000\phi\left(\frac{330 - m}{\sqrt{\sigma^2}}\right).$$
   On propose l'estimateur
   $$\widehat{N} = 1000\phi\left(\frac{330 - \widehat{m}}{\sqrt{\widehat{\sigma^2}}}\right)$$
   ce qui donne l'estimation $\widehat{N} = 1000(1 - \phi(3))$.

4. On veut estimer

$$p_2 = \mathbb{P}[\max(X_1, ..., X_n) < 360] = \mathbb{P}[X_1 < 360, ..., X_n < 360] = \mathbb{P}[X < 360]^{100} \text{ par indépendance}$$
$$= \mathbb{P}[\frac{X - m}{\sqrt{\sigma^2}} < \frac{360 - m}{\sqrt{\sigma^2}}]^{100} = \phi\left(\frac{360 - m}{\sqrt{\sigma^2}}\right)^{100}.$$

On propose l'estimateur
$$\widehat{p_2} = \phi\left(\frac{360 - \widehat{m}}{\sqrt{\widehat{\sigma^2}}}\right)^{100}$$
ce qui donne l'estimation $\widehat{p_2} = \phi(0)^{100} = 0.5^{100}$.

5. On veut estimer

$$p_3 = \mathbb{P}[\min(X_1, ..., X_n) > 330] = \mathbb{P}[X_1 > 330, ..., X_n > 330] = \mathbb{P}[X > 330]^{100} \text{ par indépendance}$$
$$= \mathbb{P}[\frac{X - m}{\sqrt{\sigma^2}} > \frac{330 - m}{\sqrt{\sigma^2}}]^{100} = \left(1 - \phi\left(\frac{330 - m}{\sqrt{\sigma^2}}\right)\right)^{100}.$$

On propose l'estimateur
$$\widehat{p_3} = \left(1 - \phi\left(\frac{330 - \widehat{m}}{\sqrt{\widehat{\sigma^2}}}\right)\right)^{100}$$
ce qui donne l'estimation $\widehat{p_3} = \phi(3)^{100}$.

**Exercice 2.**

1. $f_\mu$ est une densité de probabilité ssi $f_\mu(x) \geq 0$ pour tout $x \in \mathbb{R}$ et $\displaystyle\int f_\mu(x)dx = 1$ puisque le modèle est dominé par la mesure de Lebesgue.
2. $\mathbb{E}_\mu[X] = \mu$, $\mathrm{Var}_\mu(X) = 2$ et $q_{1/2}(X) = \mu$.
3. Utilisons la première définition de l'information de Fisher valable sous les hypothèses $(H_1)$ et $(H_2)$ à rappeler et à vérifier. On obtient $I(\mu) = 1$.
4. Le modèle n'appartient pas à la famille exponentielle car on ne peut pas l'écrire sous la forme
$$f_\mu(x) = \exp(Q(\mu)\psi(x) - \widetilde{B}(\mu))h(x)$$
5. L'estimateur de la méthode des moments de $\mu$ est $\widetilde{\mu}_n = \overline{X}_n$. Montrer que $\widetilde{\mu}_n$ est sans biais pour $\mu$, fortement consistant et que $\sqrt{n}(\widetilde{\mu}_n - \mu) \xrightarrow{\mathcal{D}} \mathcal{N}(0, 2)$.
6. L'estimateur de la méthode du maximum de vraisemblance de $\mu$ est $\widehat{\mu}_n = \mathrm{med}(X_1, ..., X_n)=$ médiane empirique.

**Exercice 3.**

Soit $X$ la variable aléatoire représentant la durée de vie des non-fumeurs ($X(\Omega) = \mathbb{R}^{*+}$) et $Y$ la variable aléatoire représentant la durée de vie des fumeurs ($Y(\Omega) = \mathbb{R}^{*+}$).
Soit $(X_1, ..., X_{n_1})$ un échantillon i.i.d. distribué comme $X$ et soit $(Y_1, ..., Y_{n_2})$ un échantillon i.i.d. distribué comme $Y$, indépendant de $(X_1, ..., X_{n_1})$.

1. On propose d'estimer $\theta := \dfrac{\mathbb{E}[X]}{\mathbb{E}[Y]}$ par $\widehat{\theta} = \dfrac{\overline{X}_{n_1}}{\overline{Y}_{n_2}}$ avec $\overline{X}_{n_1} = \dfrac{1}{n_1}\sum_{i=1}^{n_1} X_i$ et $\overline{Y}_{n_2} = \dfrac{1}{n_2}\sum_{i=1}^{n_2} Y_i$.
En effet, la LFGN permet d'affirmer que $\overline{X}_{n_1}$ et $\overline{Y}_{n_2}$ sont fortement consistants respectivement pour $\mathbb{E}[X]$ et $\mathbb{E}[Y]$. Comme la fonction $g : (x, y) \to \dfrac{x}{y}$ est continue sur $\mathbb{R}^{*+} \times \mathbb{R}^{*+}$, on en déduit que $\widehat{\theta}$ est fortement consistant pour $\theta$.
2. Le TCL pour v.a.i.i.d. admettant des moments d'ordre 1 et 2 permet d'affirmer d'une part que :
$$\sqrt{n}(\overline{X}_n - \mathbb{E}[X]) \xrightarrow{\mathcal{D}} U \sim \mathcal{N}(0, \mathrm{Var}(X))$$
et d'autre part que
$$\sqrt{n}(\overline{Y}_n - \mathbb{E}[Y]) \xrightarrow{\mathcal{D}} V \sim \mathcal{N}(0, \mathrm{Var}(Y))$$
Comme les $X_i$ sont indépendants des $Y_i$, on en déduit que les deux convergences précédentes ont lieu de façon jointe de sorte que l'on peut écrire :
$$\sqrt{n}\begin{pmatrix}\overline{X}_n - \mathbb{E}[X]\\ \overline{Y}_n - \mathbb{E}[Y]\end{pmatrix} \xrightarrow{\mathcal{D}} \begin{pmatrix}U\\ V\end{pmatrix} \sim \mathcal{N}_2\left(\begin{pmatrix}0\\0\end{pmatrix}, \begin{pmatrix}\mathrm{Var}(X) & 0\\ 0 & \mathrm{Var}(Y)\end{pmatrix}\right)$$
Appliquons la delta-méthode bivariée avec $\psi : (u, v) \to \dfrac{u}{v}$ qui est de classe $c^1$ sur $\mathbb{R}^{*+} \times \mathbb{R}^{*+}$. Déterminons la matrice de l'application linéaire tangente à $\psi$ : $D\psi(u, v) = (\dfrac{1}{v}, -\dfrac{u}{v^2})$.
On obtient alors
$$\sqrt{n}(\psi(\overline{X}_n, \overline{Y}_n) - \psi(\mathbb{E}[X], \mathbb{E}[Y])) \xrightarrow{\mathcal{D}} D\psi(\mathbb{E}[X], \mathbb{E}[Y]).\begin{pmatrix}U\\ V\end{pmatrix} = \frac{U}{\mathbb{E}[Y]} - \frac{\mathbb{E}[X]V}{\mathbb{E}[Y]^2}$$
Déterminons la loi de $W$ : c'est l'image par une application linéaire d'un vecteur gaussien donc $W$ suit la loi normale en dimension 1. Il reste à déterminer la moyenne et la variance de la loi de $W$. On conclut
$$\sqrt{n}(\widehat{\theta} - \theta) \xrightarrow{\mathcal{D}} \mathcal{N}(0, \sigma_W^2)$$
avec $\sigma_W^2 = \dfrac{\mathrm{Var}(X)}{\mathbb{E}[Y]^2} + \dfrac{\mathbb{E}[X]^2}{\mathbb{E}[Y]^4}\mathrm{Var}(Y)$ par indépendance des deux composantes $U$ et $V$ du vecteur gaussien limite.
3. Estimons $\sigma_W^2$ par
$$\widehat{\sigma_W^2} = \frac{\widehat{\sigma_{n,X}^2}}{(\overline{Y}_n)^2} + \frac{(\overline{X}_n)^2}{(\overline{Y}_n)^4}\widehat{\sigma_{n,X}^2}$$
avec
$$\widehat{\sigma_{n,X}^2} = \frac{1}{n}\sum_{i=1}^{n}(X_i - \overline{X}_n)^2, \quad \widehat{\sigma_{n,Y}^2} = \frac{1}{n}\sum_{i=1}^{n}(Y_i - \overline{Y}_n)^2.$$
D'après la LFGN, $\widehat{\sigma_{n,X}^2} \to \sigma_{n,X}^2$ presque-sûrement. On déduit de ce qui précède et du théorème de Slutsky que
$$\sqrt{n}\frac{(\widehat{\theta} - \theta)}{\sqrt{\widehat{\sigma_{n,X}^2}}} \xrightarrow{\mathcal{D}} \mathcal{N}(0, 1)$$
On peut alors construire l'intervalle de confiance suivant pour $\theta$ au niveau de confiance $1 - \alpha$ (donner le détail de la construction) en notant $\phi$ la fonction de répartition de la loi

normale centrée réduite :

$$\mathbb{P}\left[\widehat{\theta} - \phi(1-\alpha/2)^{-1}\sqrt{\frac{\widehat{\sigma^2_W}}{n}} \leq \theta \leq \widehat{\theta} + \phi(1-\alpha/2)^{-1}\sqrt{\frac{\widehat{\sigma^2_W}}{n}}\right] = 1-\alpha.$$

**Exercice 4.**

1. $\mathbb{E}[Y|X] = a\log(X) + b$, $\mathrm{Var}(Y|X) = \sigma^2$, la covariable $X$ influe sur la moyenne de $Y$ de façon logarithmique avec une erreur aléatoire centrée.
2. La détermination de $(\widehat{a}, \widehat{b})$ estimateur des moindres carrés de $(a,b)$ peut se faire ou bien en annulant le vecteur gradient de $Q(a,b) = \sum_{i=1}^{n}(Y_i - a\log(X_i) - b)^2$ et en vérifiant que la matrice hessienne en la solution obtenue est définie positive ou bien en appliquant le théorème de projection orthogonale sur un sous-espace vectoriel de $\mathbb{R}^n$.
   Rappel du théorème de projection sur un sous espace vectoriel fermé d'un espace de Hilbert :
   *Soit $F$ un sous espace vectoriel fermé d'un espace de Hilbert muni d'un produit scalaire $\langle .,. \rangle$ et d'une norme $\|.\|$. Alors, pour tout $\mathbb{Y}$ de $H$, il existe un unique point noté $p_F(\mathbb{Y})$ tel que $\|\mathbb{Y} - p_F(\mathbb{Y})\| = \inf_{\mathbb{U}\in F}\|\mathbb{Y} - \mathbb{U}\|$ et qui se caractérise par deux conditions à savoir*
   $$\begin{cases} p_F(\mathbb{Y}) \in F, \\ \mathbb{Y} - p_F(\mathbb{Y}) \in F^{\perp}. \end{cases}$$
   Appliquons ce théorème avec $H = \mathbb{R}^n$ muni de $\langle x, y\rangle = \sum_{i=1}^{n} x_i y_i$ et de $\|x\|^2 = \sum_{i=1}^{n} x_i^2$ et avec $F = \mathrm{Vect}\left(\begin{pmatrix}1\\ \vdots \\ 1\end{pmatrix}, \begin{pmatrix}\log(X_1)\\ \vdots \\ \log(X_n)\end{pmatrix}\right)$ de sorte que déterminer $\mathrm{argmin}_{a,b} Q(a,b)$ revient à déterminer $\mathrm{argmin}_{\mathbb{U}\in F}\|\mathbb{Y} - \mathbb{U}\|$ avec $\mathbb{Y} = \begin{pmatrix}Y_1\\ \vdots \\ Y_n\end{pmatrix}$ et $\mathbb{U} = a\begin{pmatrix}\log(X_1)\\ \vdots \\ \log(X_n)\end{pmatrix} + b\begin{pmatrix}1\\ \vdots \\ 1\end{pmatrix}$. Ainsi $(\widehat{a}, \widehat{b})$ est tel que $p_F(\mathbb{Y}) = \widehat{a}\begin{pmatrix}1\\ \vdots \\ 1\end{pmatrix} + \widehat{b}\begin{pmatrix}\log(X_1)\\ \vdots \\ \log(X_n)\end{pmatrix}$. Il reste donc à déterminer $p_F(\mathbb{Y})$.
   Au bout des calculs, on obtient
   $$\widehat{a} = \frac{\overline{(Y\log(X))}_n - \overline{Y}_n\overline{\log(X)}_n}{\overline{(\log(X)^2)}_n - (\overline{\log(X)}_n)^2}, \quad \widehat{b} = \frac{\overline{Y}_n\overline{(\log(X)^2)}_n - \overline{\log(X)}_n\overline{(Y\log(X))}_n}{\overline{(\log(X)^2)}_n - (\overline{\log(X)}_n)^2}$$
   en notant
   $$\overline{Y}_n = \frac{1}{n}\sum_{i=1}^{n} Y_i, \quad \overline{\log(X)}_n = \frac{1}{n}\sum_{i=1}^{n}\log(X_i)$$
   $$\overline{(Y\log(X))}_n = \frac{1}{n}\sum_{i=1}^{n} Y_i\log(X_i), \quad \overline{(\log(X)^2)}_n = \frac{1}{n}\sum_{i=1}^{n}\log(X_i)^2.$$

**Exercice 5.**

1. $\begin{pmatrix} 6 & -3 & -2 \\ -3 & 3 & 2 \\ -2 & -2 & 3 \end{pmatrix}$ n'est pas une matrice de variance car elle n'est pas symétrique.

   $\begin{pmatrix} 4 & 1 & 2 \\ 1 & 9 & -1 \\ 2 & -1 & -3 \end{pmatrix}$ n'est pas une matrice de variance car l'un des coefficients diagonaux est négatifs.

   $M = \begin{pmatrix} 6 & -3 & 0 \\ -3 & 3 & 2 \\ 0 & 2 & 3 \end{pmatrix}$ est une matrice symétrique, tous ces coefficients diagonaux sont strictement positifs, il reste à montrer q'elle est définie positive ie que son spectre est inclus dans $\mathbb{R}^{*+}$ Or $\det(M - xI_3) = -x^3 + 8x^2 - 4x - 9$. En étudiant les variations de cette fonction et en utilisant le fait que $\det(M) = 15$, on obtient le résultat.

2. $\mathbb{E}[X - Z|Y - X] = \dfrac{11}{15}(X - Y) + \dfrac{19}{15}$ donc $\mathcal{L}(\mathbb{E}[X - Z|Y - X]) = \mathcal{N}(2, 121/15)$.