ENSAE

ECOLE NATIONALE
DE LA STATISTIQUE
ET DE L'ADMINISTRATION
ECONOMIQUE

Cursus Intégré
2004-2005

Rappels de statistique mathématique
Corrigé des travaux dirigés n°7

Guillaume Lacôte
Bureau **E03**
✉ Guillaume.Lacote@ensae.fr
☞ http://ensae.no-ip.com/SE222/

Corrigé de l'exercice 1

☞ Q1 (a) *Rappel :*

Essentiellement, si $(X_i)_{i\in[\![1,n]\!]} \underset{iid}{\rightsquigarrow} \mathcal{L}$ loi discrète sur $[\![1,m]\!]$, et si $p_k = \mathbb{P}(X_i = k)$ et N_k $|\{i/X_i = k\}|$ est le nombre de réalisations de $k \in [\![1,m]\!]$, alors (en appliquant le Théorè Central Limite centré réduit à $(N_1, \ldots, N_m)$ puis en l'"élevant au carré") on a

$$\xi_n^{\text{adéq}} = \sum_{k=1}^{m} \frac{(N_k - np_k)^2}{np_k} \xrightarrow[n\to+\infty]{\mathcal{L}} \chi^2_{m-1}$$

(cf cours de Paul Doukhan, théorème 9.2) On en déduit une région de test asymptotique niveau $\alpha \in]0,1[$: $W_n^{\text{adéq}} = \{\xi_n^{\text{adéq}} > q_{1-\alpha}^{\chi^2_{m-1}}\}$. A distance finie on s'assurera avant d'appliq ce test asymptotique que $\forall k \in [\![1,m]\!]$, $np_k > 5$, quitte à fusionner plusieurs classe cette condition est requise pour que l'approximation d'une binômiale en normale s raisonnable. Enfin, dans le cas d'une loi continue on se ramènera au cas précédent discrétisant l'espace des réalisations en m classes.

Pourquoi ne pas utilise le test de Kolmogorov-Smirnov ? (cf théorème 9.3)

D'une certaine façon ce test se fonde sur la distance en $\|\|_\infty$ tandis que le test du χ^2 fonde sur la distance en $\|\|_2$ entre la loi théorique et la loi empirique. De plus, ce t utilise une statistique dont on ne sait pas exprimer la loi limite (mais on sait la tabuler) l'inverse le test du χ^2 utilise une statistique dont la loi limite est bien connue mais nécess de discrétiser l'univers continu en classes ce qui introduit une erreur. De toute façon test est inapplicable ici car la loi de Poisson est discrète, tandis que le test du χ^2 est b adapté.

Pour tester l'adéquation à une famille de lois paramétrée, une idée naturelle est de tes l'adéquation simple à la loi la plus vraisemblable, c'est-à-dire à la loi dont le paramè est l'estimateur du maximum de vraisemblance.

Considérons donc le modèle statistique $\left(\mathbb{N}^{200}, \tau, \left(\mathcal{P}(\lambda)^{\otimes 200}\right)_{\lambda\in\mathbb{R}^{+*}}\right)$. Dans ce modèle vraisemblance s'écrit $L_{X_1,\ldots,X_n}(x_1,\ldots,x_n;\lambda) = \prod_{i=1}^n e^{-\lambda}\frac{\lambda^{x_i}}{x_i!}$ et l'estimateur du maxim de vraisemblance est donc $\widehat{\lambda} = \bar{X}$; on a numériquement

$$\begin{aligned}\bar{x} &= \frac{6\cdot 0 + 15\cdot 1 + 40\cdot 2 + 42\cdot 3 + 37\cdot 4 + 30\cdot 5 + 10\cdot 6 + 9\cdot 7 + 5\cdot 8 + 9\cdot 3 + 2\cdot 1}{200} \\ &\simeq 3,7\end{aligned}$$

En conséquence, si la famille $(X_i)_i$ suivait bien une loi de Poisson, celle-ci serait de pa mètre $\widehat{\lambda} \simeq 3,7$ et par conséquent on aurait

$$\begin{aligned}\forall k \in [\![0,11]\!],\ \widehat{p_k} &= \mathbb{P}_{\widehat{\lambda}}(X_i = k) \\ &= e^{-\widehat{\lambda}}\frac{\widehat{\lambda}^k}{k!} \\ &\simeq e^{-3,7}\frac{3,7^k}{k!}\end{aligned}$$

et il ne reste plus qu'à pratiquer une test d'adéquation simple à la loi $\mathcal{P}\left(\widehat{\lambda}\right)$.

Pour ce faire, afin de préserver la légitimité du test asymptotique à distance finie on s'assure que $\forall k \in [\![0, 11]\!]$, $N_k < 5$, ce qui conduit à fusionner les quatre dernières modalités selon le tableau suivant :

Modalité	Effectif observé
$X_i = 0$	$N_0 = 6$
$X_i = 1$	$N_1 = 15$
$X_i = 2$	$N_2 = 40$
$X_i = 3$	$N_3 = 42$
$X_i = 4$	$N_4 = 37$
$X_i = 5$	$N_5 = 30$
$X_i = 6$	$N_6 = 10$
$X_i = 7$	$N_7 = 9$
$X_i \geq 8$	$N_8 = 11$

Par conséquent $\xi_{200}^{\text{adéq}} = \sum_{k=0}^{8} \frac{(N_k - n\widehat{p_k})^2}{n\widehat{p_k}} \simeq 7,58$. Or si $(X_i)_i$ suit bien une certaine loi de poisson, alors $\xi_{200}^{\text{adéq}}$ suit une loi du χ^2 à $(9-1) - \dim(\mathbb{R}^{+*}) = 7$ degrés de liberté (puisqu'on n'a conservé que 9 classes). Le fractile de niveau 5% valant $q_{95\%}^{\chi_7^2} \simeq 14,1$, on **accepte** finalement au seuil 5% l'hypothèse d'adéquation de la vrai loi inconnue de $(X_i)_i$ à la famille des lois de Poisson ; on peut en outre affirmer que $(X_i)_i$ suit vraisemblablement une loi de Poisson de paramètre $3,7$.

(b) Le niveau limite entre l'acceptation et le rejet de l'hypothèse pour l'échantillon observé est par définition la *p-value*, qui est telle que $\xi_{200}^{\text{adéq}} = q_{p-value}^{\chi_7^2}$; on a ici $p-value \simeq 19\%$, ce qui est assez élevé : l'hypothèse est rejetée au seuil 20%, mais acceptée à tout niveau $\alpha \leq 18\%$.

☞ Q2 On cherche cette fois à tester l'adéquation simple de la loi des $(X_i)_i$ à la loi de Poisson de paramètre 4.
De la même façon que dans le cas précédent les quatre dernières classes doivent être regroupées, et on calcule successivement $p_k = e^{-4} \cdot \frac{4^k}{k!}$ pour $k \in [\![0, 8]\!]$, puis $\xi_{200}^{\text{adéq}} = \sum_{k=0}^{8} \frac{(N_k - np_k)^2}{np_k} \simeq 16,8$, qui est censée suivre une loi du χ^2 à $(9-1) = 8$ degrés de liberté (il n'y a pas de paramètre à estimer ici, donc $\dim(\Theta) = 0$). Comme $q_{99\%}^{\chi_8^2} \simeq 20,09$ on **accepte** l'hypothèse de stabilité au niveau 1%.

☞ Q3 De façon similaire, on reprend $p_k = e^{-3,7} \frac{3,7^k}{k!}$ pour $k \in [\![0, 8]\!]$, ce dont on tire $\xi_{200}^{\text{adéq}} \simeq 7,58$, qui est censée suivre une loi du χ^2 à 8 degrés de liberté (là encore, $\lambda = 3,7$ est donné a priori et il n'y a pas de paramètre à estimer). Bien évidemment, $3,7$ étant la valeur de l'estimateur du maximum de vraisemblance dans le modèle statistique d'une famille de lois de Poisson de paramètre inconnu, à laquelle l'adéquation a déjà été acceptée au niveau α, l'adéquation simple à la loi de paramètre $\widehat{\lambda} = 3,7$ est **nécessairement** acceptée au même niveau α puisque $1 - F_{\chi_8^2}(x) > 1 - F_{\chi_7^2}(x)$ pour tout x.

En définitive, en notant $p^{\chi^2}(k)$ la p-value de la loi du χ^2 à k degrés de liberté on a :

$\xi_{200}^{\text{adéq}}$	Adéquation à la famile $\mathcal{P}(\lambda)_\lambda$	Adéquation simple à la loi $\mathcal{P}(3,7)$
$[0, q^{\chi^2}(7)[$	acceptée	acceptée
$]q^{\chi^2}(7), q^{\chi^2}(8)[$	rejetée	acceptée
$]q^{\chi^2}(8), +\infty[$	rejetée	rejetée

⋆

⋆ ⋆

Corrigé de l'exercice 2

☞ Q1 Soit p la vraie probabilité qu'un ménage soit équipé en magnétoscope (on identifiera proportion de ménages équipés à p). On a immédiatement $\widehat{p}(y_1, \ldots, y_n) = \bar{y}$, pour leq
$\sqrt{n}\,(\widehat{p} - p) \xrightarrow[n \to +\infty]{\mathcal{L}} \mathcal{N}(0, p(1-p))$
En appliquant la même méthode que pour l'exercice exercice 3
on montre qu'on accepte H_0 : "$p \leq p_0 = 20\%$" *ssi* $\bar{y} < k_\alpha$ avec

$$k_\alpha = p_0 + q_{1-\alpha}^{\mathcal{N}(0,1)} \sqrt{\frac{\bar{y}((1-\bar{y})}{n}}$$

☞ Q2 (a) La vraisemblance du modèle s'écrit pour $(y_1, \ldots, y_n) \in \mathbb{R}^{+n}$ et $(y_1, \ldots, y_n) \in \mathbb{R}^n$

$$L_{Y_1,\ldots,Y_n|X_1=x_1,\ldots,X_n=x_n}(y_1, \ldots, y_n; a, b) = \prod_{i=1}^{n} \frac{\left(e^{a+bx_i}\right)^{y_i}}{1 + e^{a+bx_i}}$$

Le score s'écrit donc

$$\frac{\partial L_{Y_1,\ldots,Y_n|X_1=x_1,\ldots,X_n=x_n}}{\partial \begin{pmatrix} a \\ b \end{pmatrix}}(y_1, \ldots, y_n; a, b) = \begin{pmatrix} \sum_{i=1}^{n} \left(y_i - \frac{e^{a+bx_i}}{1+e^{a+bx_i}}\right) \\ \sum_{i=1}^{n} \left(y_i - \frac{e^{a+bx_i}}{1+e^{a+bx_i}}\right) x_i \end{pmatrix}$$

et l'information de Fisher est donc

$$I_n(a, b) = \sum_{i=1}^{n} \mathbb{E}\left(\frac{e^{a+bx_i}}{(1 + e^{a+bx_i})^2} \begin{pmatrix} 1 & x_i \\ x_i & {x_i}^2 \end{pmatrix}\right)$$

qu'on ne peut expliciter davantage sans hypothèse sur la distribution de $(X_i)_i$.

(b) Sous l'hypothèse H_0 : "$b = 0$" il vient

$$\forall x_i \in \mathbb{R},\ \mathbb{P}_{H_0}\left(Y_i = 1 | X_i = x_i\right) = \frac{e^a}{1+e^a}$$

c'est-à-dire que $(Y_i|X_i)$ suit une binômiale de paramètre $p_0 = \frac{e^a}{1+e^a}$ indépendant de x_i.
La vraisemblance conditionnelle sous H_0 est donc

$$L^{H_0}_{Y_1,\dots,Y_n|X_1=x_1,\dots,X_n=x_n}\left(y_1,\dots,y_n;a,b\right) = \frac{e^{a\sum_{i=1}^n y_i}}{\left(1+e^a\right)^n}$$

et un estimateur $\widehat{a^0}$ de a vérifie donc, sous H_0, $\frac{e^{\widehat{a^0}}}{1+e^{\widehat{a^0}}} = \bar{y}$, soit $\widehat{a^0} = \ln \frac{\bar{y}}{1-\bar{y}}$. On a bien entendu $\widehat{b^0} = 0$.
Il vient alors

$$\frac{\partial L^{H_0}_{Y_1,\dots,Y_n|X_1=x_1,\dots,X_n=x_n}}{\partial \begin{pmatrix} a \\ b \end{pmatrix}}\left(y_1,\dots,y_n;\widehat{a^0},\widehat{b^0}\right) = \begin{pmatrix} 0 \\ \sum_{i=1}^n \left(y_i - \bar{y}\right) x_i \end{pmatrix}$$

(c) La matrice d'information de Fisher sous H_0 est

$$I_n(a,b) =_{|H_0} \frac{e^a}{\left(1+e^a\right)^2}\mathbb{E}\left(\sum_{i=1}^n \begin{pmatrix} 1 & x_i \\ x_i & {x_i}^2 \end{pmatrix}\right)$$

Or $\frac{e^{\widehat{a^0}}}{1+e^{\widehat{a^0}}} = \bar{y}$, donc $\frac{e^{\widehat{a^0}}}{\left(1+e^{\widehat{a^0}}\right)^2} = \bar{y}(1-\bar{y})$ et par conséquent (en notant $\bar{x^2} = \frac{1}{n}\sum_{i=1}^n {x_i}^2$)

$$\widehat{I_1^{H_0}} = \bar{y}(1-\bar{y})\begin{pmatrix} 1 & \bar{x} \\ \bar{x} & \bar{x^2} \end{pmatrix}$$

est un estimateur convergent de I_1 si H_0 est vraie.

(d) On a par définition

$$\xi_n^S = \frac{1}{n}\widehat{S^0}'\left(\widehat{I_1^{H_O}}\right)^{-1}\widehat{S^0}$$

en notant $\widehat{S^0} = \frac{\partial L_{Y_1,\dots,Y_n|X_1=x_1,\dots,X_n=x_n}}{\partial (a,b)'}\left(y_1,\dots,y_n;\widehat{a^0},\widehat{b^0}\right)$.
Or

$$\begin{aligned}\left(\widehat{I_1^{H_O}}\right)^{-1} &=_{|H_0} \frac{1}{\bar{y}(1-\bar{y})}\cdot\left(\begin{pmatrix} 1 & \bar{x} \\ \bar{x} & \bar{x^2} \end{pmatrix}\right)^{-1} \\ &=_{|H_0} \frac{1}{\bar{y}(1-\bar{y})}\cdot\frac{1}{\bar{x^2}-(\bar{x})^2}\cdot\begin{pmatrix} \bar{x^2} & -\bar{x} \\ -\bar{x} & 1 \end{pmatrix}\end{aligned}$$

Donc

$$\begin{aligned}&\xi_n^S(y_1,\dots,y_n) \\ =_{|H_0}\ & \frac{1}{n}\left(0, \sum_{i=1}^n \left(y_i-\bar{y}\right)x_i\right)\times\left(\frac{1}{\bar{y}(1-\bar{y})}\frac{1}{\bar{x^2}-(\bar{x})^2}\begin{pmatrix} \bar{x^2} & -\bar{x} \\ -\bar{x} & 1 \end{pmatrix}\right)\times\Bigg(\begin{matrix} 0 \\ \sum_{i=1}^n (y_i - \end{matrix} \\ =_{|H_0}\ & \frac{1}{n}\left(\sum_{i=1}^n \left(y_i-\bar{y}\right)x_i\right)\cdot\frac{1}{\bar{y}(1-\bar{y})}\frac{1}{\bar{x^2}-(\bar{x})^2}\cdot\left(\sum_{i=1}^n \left(y_i-\bar{y}\right)x_i\right) \\ =_{|H_0}\ & \frac{1}{\bar{y}(1-\bar{y})}\cdot\frac{\left(\sum_{i=1}^n \left(y_i-\bar{y}\right)x_i\right)^2}{\sum_{i=1}^n \left(x_i-\bar{x}\right)^2}\end{aligned}$$

Par construction, si H_0 est vraie alors

$$\xi_n^S \xrightarrow[n\to+\infty]{\mathcal{L}} \chi_1^2$$

ce qui permet de tester effectivement H_0.
Remarquons que ξ_n^S est une fonction de $(X_i)_{i\in[\![1,n]\!]}$ uniquement à travers sa variance
sa covariance avec $(Y_i)_{i\in[\![1,n]\!]}$; en outre elle est croissante avec cette dernière, comme
suggère l'intuition : dire que la loi de Y_i conditionnellement à X_i dépend effectivement
X_i (ce qui revient à dire "$\neg H_0$"), c'est dire que les observations Y_i sont significativem
corrélées avec celles de X_i, et inversement.

⋆
⋆ ⋆

Corrigé de l'exercice 3

☞ Q1 Dans la pratique la variable Y_k est tronquée, c'est-à-dire l'opérateur n'attend pas plus de N
tirages successifs qu'une pièce soit défectueuse. Les calculs pourraient être menés directem
dans le cas où $N = +\infty$, ce qui revient à supprimer tous les termes en facteur de " x^N " da
les lignes qui suivent.
On a pour $k \in [\![1,K]\!]$ et $n \in [\![1,N]\!]$, en notant $M = \sum_{n=N+1}^{+\infty} p(1-p)^{n-1}$

$$\begin{aligned}\mathbb{P}(Y_k = n) &= \mathbb{P}(\text{“Les } n-1 \text{ premières pièces tirées sont conformes et la } n\text{-ième est défectu} \\ &= \frac{1}{M}(1-p)^{n-1}\cdot p \quad \text{(le tirage s'effectue avec remise)}\end{aligned}$$

Le tirage s'effectuant avec remise, le nombre de tirages nécessaires n pour obtenir une pi
défectueuse n'est pas borné (avec toutefois $\mathbb{P}(n = +\infty) = 0$).

$$
\begin{aligned}
\mathbb{E}\left(Y_k\right) &= \sum_{n=0}^{N} n\mathbb{P}(Y_k = n) \\
&= \frac{1}{M}\sum_{n=0}^{N} np(1-p)^{n-1} \\
&= \frac{1}{M}p\sum_{n=0}^{N}\left(\frac{\partial}{\partial x}(x \mapsto x^n)\right)_{(1-p)} \\
&= \frac{1}{M}p\left(\left(\frac{\partial}{\partial x}(x \mapsto \sum_{n=0}^{N} x^n)\right)_{(1-p)}\right) \\
&= \frac{1}{M}p\left(\left(\frac{\partial}{\partial x}(x \mapsto \frac{1-x^{N+1}}{1-x})\right)_{(1-p)}\right) \\
&= \frac{1}{M}p\left(x \mapsto \frac{-(N+1)x^N(1-x)+(1-x^{N+1})}{(1-x)^2}\right)_{(1-p)} \\
&= \frac{1}{M}p\left(x \mapsto \frac{1+Nx^{N+1}-(N+1)x^N}{(1-x)^2}\right)_{(1-p)} \\
&= \frac{1}{M}\frac{1}{p}\left(1-(1-p)^N\left((N+1)p+1-p\right)\right) \\
&\underset{N\infty}{\longrightarrow} \frac{1}{p}
\end{aligned}
$$

et d'autre part que

$$
\begin{aligned}
\mathbb{E}\left({Y_k}^2\right) &= \sum_{n=0}^{N} n^2\mathbb{P}(Y_k = n) \\
&= \frac{1}{M}\sum_{n=0}^{N} n^2p(1-p)^{n-1} \\
&= \frac{1}{M}p\left((1-p)\sum_{n=2}^{N} n(n-1)(1-p)^{n-2} + \sum_{n=1}^{N} n(1-p)^{n-1}\right) \\
&= \frac{1}{M}p\left((1-p)\left(\sum_{n=2}^{N}\frac{\partial^2}{\partial x^2}(x \mapsto x^n)\right)_{(1-p)} + \left(\sum_{n=1}^{N}\frac{\partial}{\partial x}(x \mapsto x^n)\right)_{(1-p)}\right) \\
&= \frac{1}{M}p\left((1-p)\left(\frac{\partial^2}{\partial x^2}(x \mapsto \frac{1-x^{N+1}}{1-x})\right)_{(1-p)} + \left(\frac{\partial}{\partial x}(x \mapsto \frac{1-x^{N+1}}{1-x})\right)_{(1-p)}\right) \\
&= \frac{1}{M}\left\{\begin{array}{l} p(1-p)\left(x \mapsto \frac{(N(N+1)x^N-(N+1)Nx^{N-1})(1-x)+2(1+Nx^{N+1}-(N+1)x^N)}{(1-x)^3}\right)_{(1-p)} \\ +p\left(x \mapsto \frac{1+Nx^{N+1}-(N+1)x^N}{(1-x)^2}\right)_{(1-p)} \end{array}\right. \\
&= \frac{1}{M}\left\{\begin{array}{l} p(1-p)\left(x \mapsto \frac{2-N(N+1)x^{N-1}+2N^2X^N-N(N-1)x^{N+1}}{(1-x)^3}\right)_{(1-p)} \\ +p\left(x \mapsto \frac{1+Nx^{N+1}-(N+1)x^N}{(1-x)^2}\right)_{(1-p)} \end{array}\right. \\
&= \frac{1}{p^2}\left\{\begin{array}{l} 2(1-p)-N(N+1)(1-p)^N+2N^2(1-p)^{N+1}-N(N+1)(1-p)^{N+\cdots} \\ \ldots+p-(N+1)p(1-p)^N-(1-p)^{N+1} \end{array}\right. \\
&= \frac{1}{M}\frac{2-p-(1-p)^N\left(N(N+1)-2N^2(1-p)+N(N-1)(1-p)^2+(N+1\cdots\right.}{p^2} \\
&\underset{N\infty}{\longrightarrow} \frac{2-p}{p^2}
\end{aligned}
$$

de sorte que

$$
\begin{aligned}
\mathbb{V}\left(Y_k\right) &= \mathbb{E}\left({Y_k}^2\right) - \left(\mathbb{E}\left(Y_k\right)\right)^2 \\
&= \frac{1}{M}\frac{1}{p^2}\left\{\begin{array}{l} 1-p \\ -(1-p)^N\left[N(N+1)-2N^2(1-p)+N(N-1)(1-p)^2-((N+1)\cdots\right. \\ -(1-p)^{2N}\left[N(N+1)-2N^2(1-p)+N(N-1)(1-p)^2+(N+1)\cdots\right. \end{array}\right. \\
&\underset{N\infty}{\longrightarrow} \frac{1-p}{p^2}
\end{aligned}
$$

On se placera dans toute la suite dans le cas où $N = +\infty$

☞ Q2 Le test de Wald se fonde sur la normalité asymptotique de l'estimateur du maximum de vr semblance, qui vaut ici $\widehat{p} = \frac{1}{\bar{y}}$. Ecrire le Théorème Central Limite implique donc de l'écrire

$\bar{y}$ puis d'appliquer le théorème de Slutsky à $g : (x \mapsto \frac{1}{x})$. On en tirerait alors la statistique de Wald habituelle, dont on saurait qu'elle suit asymptotiquement une loi du χ^2 à un degré de liberté.

Pour autant, comme g est bijective, nous préférons appliquer le Théorème Central Limite à $\bar{y}$ (comparé à $g^{-1}(p_0) = \frac{1}{p_0}$), et reconstruire de façon analogue à la construction de la statistique de Wald une statistique qui suit un χ^2 à un degré de liberté.

On a en effet d'après le Théorème Central Limite

$$\sqrt{K}\left(\bar{y} - \frac{1}{p}\right) \xrightarrow[n\to+\infty]{\mathcal{L}} \mathcal{N}\left(0, \frac{1-p}{p^2}\right)$$

donc

$$\sqrt{K}\frac{\bar{y} - \frac{1}{p}}{\sqrt{\frac{1-p}{p^2}}} \xrightarrow[K\infty]{\mathcal{N}} (0,1)$$

et donc

$$K\frac{\left(\bar{y} - \frac{1}{p}\right)^2}{\frac{1-p}{p^2}} \xrightarrow[K\infty]{} \chi^2_1$$

Enfin, en approchant la variance $\frac{1-p}{p^2}$ par $\frac{1-\frac{1}{\bar{y}}}{\frac{1}{\bar{y}^2}} = \bar{y}(\bar{y}-1)$ il vient

$$\xi_n^W(y_1, \ldots, y_n) = K\frac{\left(\bar{y} - \frac{1}{p_0}\right)^2}{\bar{y}(\bar{y}-1)} \xrightarrow[K\infty]{H_0} \chi^2_1$$

☞ Q3 On cherche à tester l'hypothèse H_0 : "$p \leq p_0 = 5\%$" au seuil $\alpha \in [0,1[$ en comparant $\widehat{p} = \bar{y}$ à un certain fractile k_α. L'idée est donc d'exploiter la normalité asymptotique de l'estimateur $\widehat{p}$ pour en déduire la forme explicite (et tabulable) de k_α

Considérons donc $k_\alpha \in \mathbb{R}$; on a

$$\begin{aligned}\mathbb{P}_p(A_k) &= \mathbb{P}_p(\bar{y} < k_\alpha) \\ &= \mathbb{P}_p\left(\sqrt{K}\frac{\bar{y} - \frac{1}{p}}{\sqrt{\bar{y}(\bar{y}-1)}} < \sqrt{K}\frac{k_\alpha - \frac{1}{p}}{\sqrt{\bar{y}(\bar{y}-1)}}\right)\end{aligned}$$

<u>*Remarque :*</u> Même si cela ne change rien à la méthode, rien n'implique que k_α ne dépende pas de la taille K de l'échantillon. On est donc amené à exprimer k_α en tant que fonction de K, supposée sufissamment régulière.

Ainsi, si $k_\alpha : (\mathbb{N} \mapsto \mathbb{R})$ est une fonction de K qui admet une limite finie lorsque $K \to +\infty$, on a

$$\mathbb{P}_p(A_k) \xrightarrow[K\infty]{} \Phi\left(\lim_{K\infty}\sqrt{K}\frac{k_\alpha(K) - \frac{1}{p}}{\sqrt{\bar{y}(\bar{y}-1)}}\right)$$

où Φ est la fonction de répartition de la loi normale centrée réduite ; il reste déterminer $\sup_{p\leq p_0}\lim_{K\infty}\mathbb{P}_p(A_k)$.

Or si $p_2 \leq p_1 \leq p_0$, on a

$$\sqrt{K}\frac{k_\alpha(K) - \frac{1}{p_2}}{\sqrt{\bar{y}(\bar{y}-1)}} \leq \sqrt{K}\frac{k_\alpha(K) - \frac{1}{p_1}}{\sqrt{\bar{y}(\bar{y}-1)}}$$

et par conséquent

$$\Phi\left(\lim_{K\infty}\sqrt{K}\frac{k_\alpha(K) - \frac{1}{p_2}}{\sqrt{\bar{y}(\bar{y}-1)}}\right) \leq \Phi\left(\lim_{K\infty}\sqrt{K}\frac{k_\alpha(K) - \frac{1}{p_1}}{\sqrt{\bar{y}(\bar{y}-1)}}\right)$$

soit encore

$$\mathbb{P}_{p_2}(A_k) \leq \mathbb{P}_{p_1}(A_k)$$

Il s'ensuit que

$$\begin{aligned}\sup_{p\leq p_0}\lim_{K\infty}\mathbb{P}_p(A_k) &= \lim_{K\infty}\sup_{p\leq p_0}\mathbb{P}_p(A_k) \quad \text{par uniforme continuité} \ldots \\ &= \lim_{K\infty}\mathbb{P}_{p=p_0}(A_k) \\ &= \Phi\left(\lim_{K\infty}\sqrt{K}\frac{k_\alpha(K) - \frac{1}{p}}{\sqrt{\bar{y}(\bar{y}-1)}}\right)\end{aligned}$$

Pour K asymptotiquement grand et $\alpha \in [0,1[$ on accepte donc H_0 *ssi*

$$\sqrt{K}\frac{k_\alpha(K) - \frac{1}{p_0}}{\sqrt{\bar{y}(\bar{y}-1)}} > q_\alpha^{\mathcal{N}(0,1)}$$

où $q_\alpha^{\mathcal{N}(0,1)}$ désigne le fractile de niveau α de la loi normale centrée réduite.

L'idée est alors de définir k_α de façon à ce que $\Phi(T(k_\alpha)) = q_\alpha^{\mathcal{N}(0,1)}$, où $T(x) = \sqrt{K}\frac{x-\frac{1}{p}}{\sqrt{\bar{y}(\bar{y}-1)}}$; cette façon, tester si $p \leq p_0$ au seuil α, c'est-à-dire comparer $\sup_{p\leq p_0}\mathbb{P}_{p,k\infty}(A_k)$ à α, revi juste à comparer $\bar{y}$ à k_α.

Plus précisément, H_0 est acceptée *ssi* $\bar{y} > k_\alpha(K)$ avec

$$\boxed{k_\alpha(K) = \frac{1}{p_0} + q_\alpha^{\mathcal{N}(0,1)}\sqrt{\frac{\bar{y}(\bar{y}-1)}{K}}}$$

(et on vérifie a posteriori que $\lim_{K\infty} k_\alpha(K) = \frac{1}{p_0}$ existe).

⋆

⋆ ⋆

Corrigé de l'exercice 4

☞ Q1 (a) On a

$$\begin{aligned}
m &= \mathbb{E}(W) \\
&= \int_{\mathbb{R}} \frac{bx}{W_0}\left(\frac{W_0}{W}\right)^{\alpha} \mathbb{1}_{W \geq W_0} W dW \\
&= (\alpha-1)(\alpha-1){W_0}^{\alpha-1}\int_{W_0}^{+\infty} W^{1-\alpha} dW \\
&= (\alpha-1){W_0}^{\alpha-1}\left[\frac{W^{2-\alpha}}{2-\alpha}\right]_{W_0}^{+\infty} \\
&= \boxed{\tfrac{\alpha-1}{\alpha-2} W_0}
\end{aligned}$$

aussitôt que $\alpha > 2$, et par ailleurs

$$\begin{aligned}
\mathbb{E}\left(W^2\right) &= \int_{\mathbb{R}} \frac{bx}{W_0}\left(\frac{W_0}{W}\right)^{\alpha} \mathbb{1}_{W \geq W_0} W^2 dW \\
&= (\alpha-1){W_0}^{\alpha-1}\int_{W_0}^{+\infty} W^{2-\alpha} dW \\
&= (\alpha-1){W_0}^{\alpha-1}\left[\frac{W^{3-\alpha}}{3-\alpha}\right]_{W_0}^{+\infty} \\
&= \frac{\alpha-1}{\alpha-3}{W_0}^2
\end{aligned}$$

d'où

$$\begin{aligned}
v &= \mathbb{E}\left(W^2\right) - \left(\mathbb{E}(W)\right)^2 \\
&= \left(\frac{\alpha-1}{\alpha-3} - \left(\frac{\alpha-1}{\alpha-2}\right)^2\right){W_0}^2 \\
&= \boxed{\tfrac{\alpha-1}{(\alpha-2)^2(\alpha-3)}{W_0}^2}
\end{aligned}$$

aussitôt que $\alpha > 3$.
En toute rigueur, il faudrait encore prouver que lorsque $\alpha < 2$,

$\mathbb{E}(W^2) - (\mathbb{E}(W))^2 = (+\infty) - (+\infty)$ diverge effectivement, en écrivant

$$\begin{aligned}
\mathbb{E}\left(W^2\right) - \left(\mathbb{E}(W)\right)^2 &= \int_{\mathbb{R}} W^2 f(W|\alpha)dW - \left(\int_{\mathbb{R}} W f(W|\alpha)dW\right)^2 \\
&= \lim_{M\infty}\left(\left(\int_{W_0}^{M} W^2 f(W|\alpha)dW\right) - \left(\int_{W_0}^{M} W f(W|\alpha)dW\right)^2\right) \\
&= \lim_{M\infty}\left(\left(\int_{W_0}^{M} W^2 f(W|\alpha)dW\right) - \left(\int_{W_0}^{M} W f(W|\alpha)dW\right)\left(\int_{W_0}^{M} \right.\right. \\
&= \lim_{M\infty}\left(\int_{W_0}^{M}\left(W^2 - W\left(\int_{W_0}^{M} \omega f(\omega|\alpha)d\omega\right)\right) f(W|\alpha)dW\right) \\
&= \lim_{M\infty}\left(\int_{W_0}^{M}\left(W^2 - W\left[\frac{W^{2-\alpha}}{2-\alpha}\right]_{W_0}^{M}\right) f(W|\alpha)dW\right) \\
&= \lim_{M\infty}\left(\int_{W_0}^{M} W\left(W - \frac{M^{2-\alpha}}{2-\alpha} + \frac{{W_0}^{2-\alpha}}{2-\alpha}\right)(\alpha-1){W_0}^{\alpha-1}W^{-\alpha}dW\right. \\
&= (\alpha-1){W_0}^{\alpha-1}\lim_{M\infty}\left(\int_{W_0}^{M} W^{2-\alpha} - \left(\frac{M^{2-\alpha}}{2-\alpha} + \frac{{W_0}^{2-\alpha}}{2-\alpha}\right)W^{1-\alpha}\right. \\
&\underset{M\infty}{\approx} (\alpha-1){W_0}^{\alpha-1}\left(\frac{M^{3-\alpha}}{3-\alpha} - \frac{M^{4-2\alpha}}{(2-\alpha)^2}\right) \\
&\underset{M\infty}{\longrightarrow} \pm\infty \quad \text{selon le signe de } \alpha - 1
\end{aligned}$$

Ainsi les deux premiers moments existent *ssi* $\alpha = Xb + 1 > 3$, condition supposée vérif
par la suite.

(b) On a pour $(w_1, \ldots, w_n) \in [W_0, +\infty[\ ^n$

$$\begin{aligned}
L_{W_1,\ldots,W_n}(w_1,\ldots,w_n;\alpha) &= \prod_{i=1}^{n} \frac{\alpha-1}{W_0}\left(\frac{W_0}{w_i}\right)^{\alpha} \\
&= (\alpha-1)^n {W_0}^{n(\alpha-1)} \frac{1}{\prod_{i=1}^{n} {w_i}^{\alpha}}
\end{aligned}$$

Par suite l'estimateur du maximisation de vraisemblance $\widehat{\alpha}$ de α vérifie

$$\frac{n}{\widehat{\alpha}-1} - \sum_{i=1}^{n} \ln\frac{w_i}{W_0} = 0$$

de sorte que

$$\boxed{\widehat{\alpha} = 1 + \frac{n}{\sum_{i=1}^{n} \ln \frac{w_i}{W_0}}}$$

(et on vérifie que cette condition est bien suffisante)

(c) On a pour une observation $w \in [W_0, +\infty[$

$$\begin{aligned} S_1(w;\alpha) &= \frac{\partial \ln L_{W_1}}{\partial \alpha}(w;\alpha) \\ &= \frac{1}{\alpha-1} - \ln \frac{w}{W_0} \end{aligned}$$

Or le score est centré, ce qui s'écrit $\mathbb{E}\left(\frac{1}{\alpha-1} - \ln \frac{W}{W_0}\right) = 0$, soit

$$\boxed{\mathbb{E}(\ln W) = \frac{1}{\alpha-1} + \ln W_0}$$

De façon similaire on a

$$\begin{aligned} I_1(\alpha) &= -E\frac{\partial^2 \ln L_{W_1}}{\partial \alpha^2}(w;\alpha) \\ &= \frac{1}{(\alpha-1)^2} \end{aligned}$$

Or $I_1(\alpha) = \mathbb{V}(S_1(W;\alpha)) = \mathbb{E}(S_1(W;\alpha)^2)$, donc finalement

$$\mathbb{V}(\ln W) = \frac{1}{(\alpha-1)^2}$$

et

$$\boxed{\mathbb{E}\left((\ln W)^2\right) = \frac{1}{(\alpha-1)^2} + \left(\frac{1}{\alpha-1} + \ln W_0\right)^2}$$

(d) Notons $Y_i = \ln W_i$ pour $i \in [\![1,n]\!]$; alors

$$s = \frac{1}{n}\sum_{i=1}^{n}\left(Y_i - \bar{Y}\right)^2$$

est la variance empirique de $(\ln W_1, \ldots, \ln W_n)$.
Or si H_0 est vraie, alors $\mathbb{V}(\ln W_i) = \frac{1}{(\alpha-1)^2}$ dont un estimateur convergent est $\frac{1}{(\widehat{\alpha}-1)^2}$, d'où l'idée de construire une statistique de test fondée sur $s - \frac{1}{(\widehat{\alpha}-1)^2}$.
Un tel test est habituellement appelé *test d'indépendance de Fisher*.

(e) On a

$$\sqrt{n}\left(\begin{pmatrix} \overline{\ln W} \\ \overline{(\ln W)^2} \end{pmatrix} - \begin{pmatrix} \mathbb{E}(\ln W) \\ \mathbb{E}\left((\ln W)^2\right) \end{pmatrix}\right) \xrightarrow[n\to+\infty]{\mathcal{L}} \mathcal{N}\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \Omega\right)$$

avec $\Omega = \begin{pmatrix} \mathbb{V}(\ln W) & \mathbb{C}ov(\ln W, (\ln W)^2) \\ \mathbb{C}ov(\ln W, (\ln W)^2) & \mathbb{V}((\ln W)^2) \end{pmatrix}$ que l'on n'explicitera pas davantage.

Soit alors $g : \begin{pmatrix} \mathbb{R}^2 & \to & \mathbb{R} \\ (m,v) & \mapsto & v - m^2 \end{pmatrix}$; alors d'après le théorème de Slutsky

$$\sqrt{n}\left(s - \frac{1}{(\alpha-1)^2}\right) \xrightarrow[n\to+\infty]{\mathcal{L}} \mathcal{N}(0, M(\alpha))$$

avec $M(\alpha) = \frac{\partial g}{\partial \alpha}(\alpha)' \times \Omega \times \frac{\partial g}{\partial \alpha}(\alpha) \in \mathbb{R}^+$.
Enfin, en considérant l'estimateur convergent $\widehat{M(\alpha)} = M(\widehat{\alpha})$ de $M(\alpha)$ on en tire que

$$\sqrt{n}\left(\frac{s - \frac{1}{(\alpha-1)^2}}{\sqrt{M(\widehat{\alpha})}}\right) \xrightarrow[n\to+\infty]{\mathcal{L}} \mathcal{N}(0,1)$$

de sorte que finalement si H_0 est vraie

$$\boxed{\xi_n^F = n\frac{\left(s - \frac{1}{(\alpha-1)^2}\right)^2}{M(\widehat{\alpha})} \xrightarrow[n\to+\infty]{\mathcal{L}} \chi_1^2}$$

☞ Q2 (a) On a pour $(w_1, \ldots, w_n, x_1, \ldots, x_n) \in [W_0, +\infty[^{2n}$

$$\begin{aligned} L_{W_1,\ldots,W_n|X_1=x_1,\ldots,X_n=x_n}(w_1,\ldots,w_n,x_1,\ldots,x_n;b) &= \prod_{i=1}^{n}\frac{bx_i}{W_0}\left(\frac{W_0}{w_i}\right)^{bx_i+1} \\ &= b^n\left(\prod_{i=1}^{n}x_i\right){W_0}^{b\sum_{i=1}^{n}x_i}\frac{1}{\prod_{i=1}^{n}{w_i}^{bx_i}} \end{aligned}$$

Par suite l'estimateur du maximum de vraisemblance $\widehat{b}$ de b vérifie

$$\frac{n}{\widehat{b}} + \sum_{i=1}^{n}x_i \ln W_0 - \sum_{i=1}^{n}x_i \ln w_i = 0$$

ce dont on tire

$$\boxed{\widehat{b} = \frac{n}{\sum_{i=1}^{n}x_i \ln \frac{w_i}{W_0}}}$$

(et on vérifie que cette condition est bien suffisante)

(b) Le score s'écrit cette fois pour une observation

$$\begin{aligned} S_1(w|X=x;b) &= \frac{\partial \ln L_{W_1|X=x}}{\partial b}(w;b) \\ &= \frac{1}{b} - x\ln\frac{w}{W_0} \end{aligned}$$

Il est centré, et donc

$$\boxed{\mathbb{E}\left(X \ln \tfrac{W}{W_0}\right) = \tfrac{1}{b}}$$

Par ailleurs l'information de Fisher s'écrit

$$\begin{aligned} I_1(b) &= -E\frac{\partial^2 \ln L_{W_1|X_1}}{\partial b^2}(w;b) \\ &= \frac{1}{b^2} \end{aligned}$$

Or

$$\begin{aligned} I_1(b) &= \mathbb{V}\left(S_1(W|X;b)\right) \\ &= \mathbb{E}\left(S_1(W|X;b)^2\right) \\ &= \mathbb{E}\left(\left(\frac{1}{b} - X\ln\frac{W}{W_0}\right)^2\right) \\ &= \mathbb{E}\left(\frac{1}{b^2} - 2\frac{X}{b}\ln\frac{W}{W_0} + X^2\left(\ln\frac{W}{W_0}\right)^2\right) \\ &= \frac{1}{b^2} + \mathbb{E}\left(X^2\left(\ln\frac{W}{W_0}\right)^2 - 2\frac{X}{b}\ln\frac{W}{W_0}\right) \end{aligned}$$

Par conséquent,

$$\boxed{\mathbb{E}\left(X^2\left(\ln\tfrac{W}{W_0}\right)^2 - 2\tfrac{X}{b}\ln\tfrac{W}{W_0}\right) = 0}$$

De façon similaire au cas où les X_i étaient constants, on cherche alors à tester H_0 en considérant l'estimateur empirique de cette expression, à savoir

$$\frac{1}{n}\sum_{i=1}^{n} {x_i}^2\left(\ln\frac{w_i}{W_0}\right)^2 - \frac{1}{b}\frac{2}{n}\sum_{i=1}^{n} x_i \ln\frac{w_i}{W_0}$$

Il suffit alors, en s'appuyant sur la loi asymptotique de $\begin{pmatrix} X\ln\frac{W}{W_0} \\ X^2\left(\ln\frac{W}{W_0}\right)^2 \end{pmatrix}$ obtenue par le Théorème Central Limite, d'appliquer le théorème de Slutsky en considérant $g : \begin{pmatrix} \mathbb{R}^2 & \to & \mathbb{R} \\ (m,v) & \mapsto & v - \frac{2}{b}m \end{pmatrix}$, ce qui conduit à

$$\sqrt{n}\,(t-0) \xrightarrow[n\to+\infty]{\mathcal{L}} \mathcal{N}\left(0, N(b)\right)$$

où $t = \frac{1}{n}\sum_{i=1}^{n} {x_i}^2\left(\ln\frac{w_i}{W_0}\right)^2 - \frac{2}{n^2}\left(\sum_{i=1}^{n} x_i\ln\frac{w_i}{W_0}\right)^2$ (en substituant $\widehat{b}$ à b)
et $N(b) = \frac{\partial g}{\partial b}(b)' \times \mathbb{V}\left(\begin{pmatrix} X\ln\frac{W}{W_0} \\ X^2\left(\ln\frac{W}{W_0}\right)^2 \end{pmatrix}\right) \times \frac{\partial g}{\partial b}(b)$.

On en tire en définitive la statistique de test

$$\xi_n^F = n\frac{\left(t - \frac{1}{b^2}\right)^2}{N(\widehat{b})} \xrightarrow[n\to+\infty]{\mathcal{L}} \chi_1^2$$

si H_0 est vraie.

☞ Q3 (a) On a pour $(w_1,\ldots,w_n,x_1,\ldots,x_n) \in \mathbb{R}^{2n}$

$$L_{W_1,\ldots,W_n|X_1=x_1,\ldots,X_n=x_n}(w_1,\ldots,w_n;W_0,b) = b^n\left(\prod_{i=1}^{n}\frac{x_i}{(w_i)^{bx_i+1}}\right){W_0}^{b\sum_{i=1}^n x_i}\mathbb{1}_{\min_i w_i \geq W_0}$$

En particulier l'estimateur du maximum de vraisemblance $\widehat{W_0}$ de W_0 est

$$\boxed{\widehat{W_0} = \min_i w_i}$$

(b) La difficulté est que le modèle statistique n'est plus régulier ; en particulier, la l
vraisemblance n'est plus dérivable, donc ni le score ni la matrice d'information de Fis
ne sont définis. La procédure de test utilisée précédemment, fondées sur leurs proprié
statistiques, n'est donc plus applicable.

⋆
⋆ ⋆