Statistique/Modèle linéaire

Corrigé du partiel du 9 novembre 2004

Exercice 1

Soient $Y_1, \ldots, Y_n$ des variables aléatoires indépendantes associées à des variables explicatives unidimensionnelles $x_1, \ldots, x_n$ telles que

$$Y|x \sim \mathcal{P}\text{oi}(\exp\{a + bx\}) .$$

1. La logvraisemblance étant

$$\ell(a, b|\mathbf{x}) = \sum_{i=1}^{n} \{y_i(a + bx_i) - \exp[a + bx_i]\} ,$$

la matrice d'information de Fisher associée au paramètre (a, b) est donnée par les dérivées secondes (où les y_i disparaissent)

$$-\begin{pmatrix} \frac{\partial^2}{\partial a^2}\ell(a,b|\mathbf{x}) = -\sum_{i=1}^n \exp[a+bx_i] & \frac{\partial^2}{\partial a b}\ell(a,b|\mathbf{x}) = -\sum_{i=1}^n x_i \exp[a+bx_i] \\ \frac{\partial^2}{\partial a b}\ell(a,b|\mathbf{x}) = -\sum_{i=1}^n x_i \exp[a+bx_i] & \frac{\partial^2}{\partial b^2}\ell(a,b|\mathbf{x}) = -\sum_{i=1}^n x_i^2 \exp[a+bx_i] \end{pmatrix} .$$

2. Sous les hypothèses de régularité usuelles, la loi asymptotique de l'estimateur du maximum de vraisemblance $(\hat{a}, \hat{b})$ est une loi normale,

$$(\hat{a}, \hat{b})^T - (a, b)^T \approx \mathscr{N}\left(0, \begin{bmatrix} \sum_{i=1}^n \exp[a+bx_i] & \sum_{i=1}^n x_i \exp[a+bx_i] \\ \sum_{i=1}^n x_i \exp[a+bx_i] & \sum_{i=1}^n x_i^2 \exp[a+bx_i] \end{bmatrix}^{-1}\right) .$$

3. Pour l'hypothèse nulle $H_0 : b = 0$, l'estimateur du maximum de vraisemblance contraint est $(\hat{a}_0, \hat{b}_0) = (\bar{y}, 0)$.
Comme l'inverse de la matrice d'information de Fisher est

$$I(a,b)^{-1} = \begin{pmatrix} \sum_{i=1}^n x_i^2 \exp[a+bx_i] & -\sum_{i=1}^n x_i \exp[a+bx_i] \\ -\sum_{i=1}^n x_i \exp[a+bx_i] & \sum_{i=1}^n \exp[a+bx_i] \end{pmatrix}$$
$$\Bigg/ \left\{ \sum_{i=1}^n \exp[a+bx_i] \sum_{i=1}^n x_i^2 \exp[a+bx_i] - \left(\sum_{i=1}^n x_i \exp[a+bx_i]\right)^2 \right\} ,$$

la statistique de Wald est fondée sur

$$\hat{b} \overset{H_0}{\approx} \mathscr{N}\left(0, \frac{n\bar{y}}{n\bar{y}\sum_{i=1}^n x_i^2 \exp[\hat{a}+\hat{b}x_i] - \left(\sum_{i=1}^n x_i y_i\right)^2}\right) ,$$

d'après les équations de vraisemblance, et elle vaut donc

$$\xi^W = \hat{b}^2 \left[n\bar{y} \sum_{i=1}^n x_i^2 \exp[a+bx_i] - \left(\sum_{i=1}^n x_i y_i\right)^2 \right] \Big/ n\bar{y} .$$

La statistique du score fait intervenir $\partial\ell(a,b|\mathbf{x})/\partial b = \sum_{i=1}^n x_i(y_i - \exp[a+bx_i])$ uniquement, dont la valeur en $(\hat{a}_0, \hat{b}_0)$ a une variance donnée par l'élément $(2,2)$ de la matrice d'information, $\sum_{i=1}^n x_i^2 \exp[a+bx_i]$, égal à $\bar{y}\sum_{i=1}^n x_i^2$ pour $(\hat{a}_0, \hat{b}_0)$. Elle vaut donc

$$\xi^S = \left[\sum_{i=1}^n x_i(y_i - \bar{y})\right]^2 \Big/ \left[\bar{y}\sum_{i=1}^n x_i^2\right] .$$

Enfin la statistique du rapport de vraisemblance est fournie par

$$\xi^R = 2\sum_{i=1}^n \left\{ y_i(\hat{a} + \hat{b}x_i - \hat{a}_0) \right\} ,$$

puisque $\sum_i \exp[a+bx_i] = n\bar{y}$ dans les deux cas.

4. Evidemment, la statistique de test du score est explicite et vaut 0.39 pour les valeurs numériques fournies. Cette valeur est bien sûr acceptable comme réalisation d'un χ^2_1.
5. Comme calculé ci-dessus, on a bien

$$\xi^R = 2\left[n\bar{y}(\hat{a} - \hat{a}_0) + \hat{b}\sum_{i=1}^n x_i y_i\right] = 2\left[n\bar{y}(\hat{a} - \log\bar{y}) + \hat{b}\sum_{i=1}^n x_i y_i\right]$$

et sa valeur numérique est 4.56, tandis que le quantile à 95% de la loi du χ^2 à 1 degré de liberté est 3.84. On rejette donc H_0 au niveau $\alpha = .05$.

Exercice 2

Soit $X_1, \dots, X_n$ un échantillon de loi de densité sur $\mathbb{R}$

$$f(x|\theta) = \frac{1}{4\theta\sqrt{|x|}}\, e^{-\sqrt{|x|}/\theta}, \quad \theta > 0\,.$$

1. Comme dans l'exercice correspondant du TD, on remarque que $Y = \sqrt{|X|}$ suit une loi de densité

$$f(y|\theta) = \frac{1}{2\theta y}\, e^{-y/\theta} \left|\frac{dx}{dy}\right| = e^{-y/\theta}/\theta\,,$$

c'est à dire une loi exponentielle $\mathscr{E}\mathrm{xp}(1/\theta)$. Comme travailler sur les y_i est équivalent à travailler sur les x_i, l'estimateur du maximum de vraisemblance de θ est donc $\hat\theta_n = 1/\bar y$ et l'information de Fisher est $1/\theta^2$.

2. De la loi asymptotique $\hat\theta_n \approx \mathscr{N}(\theta, \theta^2/n)$, on en déduit

$$\xi^W = n(\bar y - 1)^2/\bar y^2$$

de loi asymptotique χ^2_1 sous H_0. Comme la vraisemblance est $-n(\{\bar y/\theta\} + \log\theta)$, la statistique du rapport de vraisemblance vaut

$$\xi^R = 2n(\bar y - 1 - \log \bar y)\,,$$

aussi de loi asymptotique χ^2_1 sous H_0.

3. La loi a posteriori associée à $(x_1, \dots, x_n)$ est la même que celle associée à la transformation $(y_1, \dots, y_n)$,

$$\begin{aligned}\pi(\theta|y_1,\dots,y_n) = \pi(\theta|\bar y) &\propto \theta^{-3}\exp(-2/\theta) \times \theta^{-n}\exp(-n\bar y/\theta)\\ &= \theta^{-n-3}\exp(-(2+n\bar y)/\theta)\end{aligned}$$

qui est donc une loi $\mathscr{IG}a(2+n, 2+n\bar y)$. Ainsi

$$\begin{aligned}\mathbb{E}[\theta|\bar y] &= \int \theta^{-n-2}\exp(-(2+n\bar y)/\theta)dy \Big/ \int \theta^{-n-3}\exp(-(2+n\bar y)/\theta)dy\\ &= (2+n\bar y)/(n+1)\,.\end{aligned}$$

4. La loi marginale de $(x_1, \dots, x_n)$, $m(x_1, \dots, x_n)$, est

$$\begin{aligned}&\frac{1}{4^n\sqrt{|x_1\cdots x_n|}} \int_0^\infty \frac{\theta^{-n-3}}{4} \exp\left(-\left\{2 + \sum_i \sqrt{|x_i|}\right\}/\theta\right) d\theta\\ &= \frac{1}{4^n\sqrt{|x_1\cdots x_n|}} \int_0^\infty \frac{\omega^{n+1}}{4} \exp\left(-\omega\left\{2 + \sum_i \sqrt{|x_i|}\right\}\right) d\omega\\ &= \frac{1}{4^{n+1}\sqrt{|x_1\cdots x_n|}} (n+1)! \left(\sum_i \sqrt{|x_i|}\right)^{-n-1}.\end{aligned}$$

Par conséquent, le facteur de Bayes associée pour le test de $H_0 : \theta = 1$ est

$$\begin{aligned}B^\pi_{01} &= \frac{\frac{1}{4^n\sqrt{|x_1\cdots x_n|}} \exp\left(-\sum_i \sqrt{|x_i|}\right)}{\frac{1}{4^{n+1}\sqrt{|x_1\cdots x_n|}} (n+1)! \left(\sum_i \sqrt{|x_i|}\right)^{-n-1}}\\ &= \frac{4}{(n+1)!} \exp\left(-\sum_i \sqrt{|x_i|}\right)\left(\sum_i \sqrt{|x_i|}\right)^{n+1}\end{aligned}$$

Exercice 3

On considère deux échantillons normaux de taille n,

$$x_{11}, \dots, x_{1n} \sim \mathcal{N}(\mu_1, \sigma^2), \quad x_{21}, \dots, x_{2n} \sim \mathcal{N}(\mu_2, \sigma^2)\,.$$

a. Sous la loi $\pi(\mu_1, \mu_2, \sigma) = \sigma^{-p}$, la loi a posteriori est

$$\pi(\mu_1, \mu_2, \sigma|\bar x_1, \bar x_2, s) \propto \sigma^{-n-p} \exp\left\{-n(\bar x_1 - \mu_1)^2 - n(\bar x_2 - \mu_2)^2 - s^2\right\}/2\sigma^2$$

et, en utilisant le changement de variable $\xi = (\xi_1, \xi_2) = (\mu_1/\sigma, \mu_2/\sigma)$,

$$\pi(\xi, \sigma|\bar x_1, \bar x_2, s) \propto \sigma^{-n-p+2} \exp\left\{-n(\bar x_1 - \sigma\xi_1)^2 - n(\bar x_2 - \sigma\xi_2)^2 - s^2\right\}/2\sigma^2\,.$$

La loi a posteriori marginale sur ξ est donc

$$\begin{aligned}
&\int_0^\infty \sigma^{-n-p+2} \exp\Big[-\Big\{\sigma^{-2}\left[n\bar{x}_1^2+n\bar{x}_2^2+s^2\right] \\
&\qquad\qquad -2\sigma^{-1}\left[n\bar{x}_1\xi_1+n\bar{x}_2\xi_1\right]+n[\xi_1^2+\xi_2^2]\Big\}/2\Big]\ d\sigma \\
&= \int_0^\infty \sigma^{-n-p+2} \exp\Big[-\Big\{s^2\sigma^{-2}\left[n\bar{x}_1^2/s^2+n\bar{x}_2^2/s^2+1\right] \\
&\qquad\qquad -2s\sigma^{-1}\left[n\bar{x}_1\xi_1/s+n\bar{x}_2\xi_1/s\right]+n[\xi_1^2+\xi_2^2]\Big\}/2\Big]\ d\sigma \\
&\propto \int_0^\infty \omega^{n+p-4} \exp\Big[-\Big\{\omega^2\left[n\bar{x}_1^2/s^2+n\bar{x}_2^2/s^2+1\right] \\
&\qquad\qquad -2\omega\left[n\bar{x}_1\xi_1/s+n\bar{x}_2\xi_1/s\right]+n[\xi_1^2+\xi_2^2]\Big\}/2\Big]\ d\omega
\end{aligned}$$

qui ne dépend que de $z=(z_1,z_2)=(\bar{x}_1/s,\bar{x}_2/s)$.

b. La loi jointe de (z,s) est la transformée de la loi de $(\bar{x}_1,\bar{x}_2,s^2)$, soit $\mathscr{N}(\mu_1,\sigma^2/n)\times\mathscr{N}(\mu_1,\sigma^2/n)\times\mathscr{G}a(n-1,1/2\sigma^2)$:

$$g(\bar{x}_1,\bar{x}_2,s^2)\propto (s^2)^{n-2}\sigma^{-2n+2}\,e^{-s^2/2\sigma^2}\,\sigma^{-2}\,e^{-\left\{n(\bar{x}_1-\mu_1)^2+n(\bar{x}_2-\mu_2)^2\right\}/2\sigma^2}\,.$$

Donc

$$\begin{aligned}
f(z,s^2) &\propto (s^2)^{n-2+1}\sigma^{-2n}\exp\left[-\left\{s^2+n(sz_1-\sigma\xi_1)^2+n(sz_2-\sigma\xi_2)^2\right\}/2\right]\sigma^2 \\
&= (s^2/\sigma^2)^{n-1}\sigma^{-2}\exp\left[-\left\{(s/\sigma)^2+n(z_1(s/\sigma)-\xi_1)^2+n(z_2(s/\sigma)-\xi_1)^2\right\}/2\right]
\end{aligned}$$

et l'intégrale en s fait disparaître σ qui n'est alors qu'un facteur d'échelle pour s.