Examen Estimation - Détection
Mercredi 19 Novembre 2008, 8h00-10h00

$1^{ère}$ Partie

1) La vraisemblance s'écrit

$$L(x_1, ..., x_n; \lambda) = \prod_{i=1}^{n} \frac{\lambda^{x_i}}{x_i!} e^{-\lambda}$$

et donc son logarithme est

$$\ln[L(x_1, ..., x_n; \lambda)] = \left(\sum_{i=1}^{n} x_i\right) \ln\lambda - n\lambda - \sum_{i=1}^{n} \ln[x_i!]$$

La dérivée de la log-vraisemblance possède donc l'expression suivante

$$\frac{\partial \ln[L(x_1, ..., x_n; \lambda)]}{\partial \lambda} = \frac{1}{\lambda} \sum_{i=1}^{n} x_i - n$$

d'où

$$\frac{\partial \ln[L(x_1, ..., x_n; \lambda)]}{\partial \lambda} = 0 \Longleftrightarrow \lambda = \frac{1}{n} \sum_{i=1}^{n} x_i$$

On vérifie aisément que cette valeur de λ correspond au maximum global unique de la log-vraisemblance, d'où

$$\widehat{\lambda}_{\text{MV}} = \overline{X} = \frac{1}{n} \sum_{i=1}^{n} X_i$$

Cet estimateur est sans biais et de variance

$$\text{Var}\left[\widehat{\lambda}_{\text{MV}}\right] = \frac{\text{Var}[X_1]}{n} = \frac{\lambda}{n}$$

L'estimateur $\widehat{\lambda}_{\text{MV}}$ est donc convergent.

2) Des calculs élémentaires permettent d'obtenir

$$\begin{aligned} \text{BCR}(\lambda) &= -E\left[\frac{\partial^2 \ln[L(x_1, ..., x_n; \lambda)]}{\partial \lambda^2}\right]^{-1} \\ &= \frac{-1}{E\left[\frac{-1}{\lambda^2}\sum_{i=1}^{n} X_i\right]} = \frac{\lambda^2}{n\lambda} = \frac{\lambda}{n} \end{aligned}$$

L'estimateur $\widehat{\lambda}_{\text{MV}}$ est donc sans biais et de variance égale à la borne de Cramer-Rao associée à un estimateur non-biaisé de λ. L'estimateur $\widehat{\lambda}_{\text{MV}}$ est donc l'estimateur efficace de λ.

3a) La matrice d'information de Fisher a été déterminée à la question précédente

$$E\left[\frac{1}{\lambda^2} \sum_{i=1}^{n} X_i\right] = \frac{n}{\lambda}$$

La matrice d'information a priori est définie par

$$E\left[-\frac{\partial^2 \ln\left[g\left(\lambda\right)\right]}{\partial\lambda^2}\right]$$

Mais

$$\ln\left[g\left(\lambda\right)\right] = C + \left(\alpha - 1\right)\ln\lambda - \beta\lambda$$

d'où

$$\frac{\partial^2 \ln\left[g\left(\lambda\right)\right]}{\partial\lambda^2} = \frac{1-\alpha}{\lambda^2}$$

On en déduit

$$E\left[-\frac{\partial^2 \ln\left[g\left(\lambda\right)\right]}{\partial\lambda^2}\right] = \left(\alpha - 1\right) E\left[\frac{1}{\lambda^2}\right]$$

Le calcul de $E\left[\lambda^{-2}\right]$ est élémentaire

$$\begin{aligned}
E\left[\frac{1}{\lambda^2}\right] &= \int_0^\infty \frac{\beta^\alpha}{\Gamma\left(\alpha\right)}\lambda^{\alpha-3}e^{-\lambda\beta}d\lambda \\
&= \frac{\beta^\alpha}{\Gamma\left(\alpha\right)}\int_0^\infty \left(\frac{u}{\beta}\right)^{\alpha-3} e^{-u}\frac{du}{\beta} \\
&= \frac{\beta^2}{\Gamma\left(\alpha\right)}\int_0^\infty u^{\alpha-3}e^{-u}du \\
&= \beta^2\frac{\Gamma\left(\alpha-2\right)}{\Gamma\left(\alpha\right)} \\
&= \frac{\beta^2}{\left(\alpha-1\right)\left(\alpha-2\right)}
\end{aligned}$$

d'où

$$E\left[-\frac{\partial^2 \ln\left[g\left(\lambda\right)\right]}{\partial\lambda^2}\right] = \frac{\beta^2}{\left(\alpha-2\right)}$$

La borne de Cramer-Rao a posteriori s'écrit donc

$$\begin{aligned}
\mathrm{BCR}_{\mathrm{post}}\left(\lambda\right) &= \left[E\left[-\frac{\partial^2 \ln\left[f\left(x_1, ..., x_n|\lambda\right)\right]}{\partial\lambda^2}\right] + E\left[-\frac{\partial^2 \ln\left[g\left(\lambda\right)\right]}{\partial\lambda^2}\right]\right]^{-1} \\
&= \left[\frac{n}{\lambda} + \frac{\beta^2}{\left(\alpha-2\right)}\right]^{-1} \\
&= \frac{\lambda\left(\alpha-2\right)}{n\left(\alpha-2\right)+\lambda\beta^2} \\
&= \frac{\lambda}{n}\frac{1}{1+\frac{\lambda\beta^2}{n\left(\alpha-2\right)}}
\end{aligned}$$

On vérifie que

$$\frac{n}{\lambda} + \frac{\beta^2}{\left(\alpha-2\right)} \geq \frac{n}{\lambda}$$

et donc

$$\mathrm{BCR}_{\mathrm{post}}(\lambda) = \left[\frac{n}{\lambda} + \frac{\beta^2}{(\alpha-2)}\right]^{-1} \leq \frac{\lambda}{n} = \mathrm{BCR}(\lambda)$$

3b) L'estimateur du maximum a posteriori du paramètre λ noté $\widehat{\lambda}_{\mathrm{MAP}}$ s'obtient en maximisant le logarithme de la densité a posteriori, c'est-à-dire de la façon suivante

$$\frac{\partial \ln\left[f\left(\lambda \mid x_1,...,x_n\right)\right]}{\partial \lambda} = 0 \Longleftrightarrow \lambda = \frac{n\overline{x}+\alpha-1}{n+\beta} = \frac{\overline{x}+\frac{\alpha-1}{n}}{1+\frac{\beta}{n}}$$

L'estimateur du maximum a posteriori du paramètre λ est donc

$$\widehat{\lambda}_{\mathrm{MAP}} = \frac{\overline{X}+\frac{\alpha-1}{n}}{1+\frac{\beta}{n}} = \frac{\widehat{\lambda}_{MV}+\frac{\alpha-1}{n}}{1+\frac{\beta}{n}}$$

Pour n grand, l'information a priori influe peu sur l'estimation et donc l'estimateur du maximum a posteriori se réduit à l'estimateur du maximum de vraisemblance $\widehat{\lambda}_{\mathrm{MAP}} \simeq \widehat{\lambda}_{\mathrm{MV}}$.

3c) L'estimateur $\widehat{\lambda}$ qui minimise l'erreur quadratique moyenne $E\left[\left(\lambda-\widehat{\lambda}\right)^2\right]$ est l'estimateur MMSE défini par

$$\widehat{\lambda}_{\mathrm{MMSE}} = E\left[\lambda \mid X_1,...,X_n\right]$$

La loi a posteriori de $\lambda \mid X_1,...,X_n$ est telle que

$$f\left(\lambda \mid x_1,...,x_n\right) \propto \lambda^{n\overline{x}+\alpha-1} e^{-\lambda(n+\beta)}$$

C'est donc une loi gamma de paramètres $n\overline{x}+\alpha$ et $n+\beta$, i.e.

$$\lambda \mid X_1,...,X_n \sim \Gamma\left(n\overline{x}+\alpha, n+\beta\right)$$

On en déduit

$$E\left[\lambda \mid X_1,...,X_n\right] = \frac{n\overline{X}+\alpha}{n+\beta} = \frac{\overline{X}+\frac{\alpha}{n}}{1+\frac{\beta}{n}}$$

$2^{ème}$ Partie

1) Le test de Neyman-Pearson pour le problème

$$\begin{cases} H_0 : \lambda = \lambda_0 \\ H_1 : \lambda = \lambda_1 \end{cases}$$

s'écrit

$$\text{Rejet de } H_0 \text{ si } \frac{L\left(x_1,...,x_n;\lambda_1\right)}{L\left(x_1,...,x_n;\lambda_0\right)} > S_\alpha$$

c'est-à-dire qu'on rejette H_0 si

$$\ln\left[L\left(x_1,...,x_n;\lambda_1\right)\right] - \ln\left[L\left(x_1,...,x_n;\lambda_0\right)\right] > \ln\left[S_\alpha\right]$$

ou

$$\left(\sum_{i=1}^n x_i\right)\left(\ln\lambda_1 - \ln\lambda_0\right) - n\left(\lambda_1-\lambda_0\right) > \ln\left[S_\alpha\right]$$

d'où le test suivant

- Si $\lambda_1 > \lambda_0$

$$\text{Rejet de } H_0 \text{ si } T = \sum_{i=1}^{n} X_i > K_\alpha$$

- Si Si $\lambda_1 < \lambda_0$

$$\text{Rejet de } H_0 \text{ si } T = \sum_{i=1}^{n} X_i < K_\alpha$$

2) La fonction caractéristique de T est

$$E\left[e^{itT}\right] = \prod_{j=1}^{n} E\left[e^{itX_j}\right] = \exp\left[n\lambda\left(e^{it}-1\right)\right]$$

qui est la fonction caractéristique d'une loi de Poisson de paramètre $n\lambda$ donc $T \sim P(n\lambda)$. On en déduit donc

$$\begin{array}{lcl} \text{Sous } H_0 & : & T \sim P(n\lambda_0) \\ \text{Sous } H_1 & : & T \sim P(n\lambda_1) \end{array}$$

3) On a

$$\begin{aligned} \alpha &= P\left[\text{rejeter } H_0 \middle| H_0 \text{ vraie}\right] \\ &= P\left[T < K_\alpha \middle| T \sim P(n\lambda_0)\right] \\ &= F_{n\lambda_0}(K_\alpha) \end{aligned}$$

où $F_{n\lambda_0}(.)$ est la fonction de répartition de la loi de Poisson de paramètre $n\lambda_0$. On a donc

$$K_\alpha = F_{n\lambda_0}^{-1}(\alpha)$$

4a) Le théorème de la limite centrale s'écrit

$$\frac{\sum_{i=1}^{n} X_i - n\lambda}{\sqrt{n\lambda}} \sim \mathcal{N}(0,1) \text{ ou } T \underset{\text{pour } n \text{ grand}}{\sim} \mathcal{N}(n\lambda, n\lambda)$$

4b) En utilisant la loi normale pour T, on a

$$\begin{aligned} \alpha &= P\left[\text{rejeter } H_0 \middle| H_0 \text{ vraie}\right] \\ &= P\left[T < K_\alpha \middle| T \sim \mathcal{N}(n\lambda_0, n\lambda_0)\right] \\ &= P\left[U = \frac{T - n\lambda_0}{\sqrt{n\lambda_0}} < \frac{K_\alpha - n\lambda_0}{\sqrt{n\lambda_0}} \middle| U \sim \mathcal{N}(0,1)\right] \\ &= F\left(\frac{K_\alpha - n\lambda_0}{\sqrt{n\lambda_0}}\right) \end{aligned}$$

d'où

$$K_\alpha = \sqrt{n\lambda_0}F^{-1}(\alpha) + n\lambda_0$$

De la même façon, la puissance du test (ou puissance du test) s'écrit

$$\begin{aligned} PD &= P\left[\text{rejeter } H_0 | \, H_1 \text{ vraie}\right] \\ &= P\left[T < K_\alpha | \, T \sim \mathcal{N}\left(n\lambda_1, n\lambda_1\right)\right] \\ &= F\left(\frac{K_\alpha - n\lambda_1}{\sqrt{n\lambda_1}}\right) \end{aligned}$$

Les courbes COR expriment PD en fonction de α et s'écrivent donc

$$\begin{aligned} PD &= F\left[\frac{\sqrt{n\lambda_0}F^{-1}\left(\alpha\right) + n\lambda_0 - n\lambda_1}{\sqrt{n\lambda_1}}\right] \\ &= F\left[\sqrt{\frac{\lambda_0}{\lambda_1}}F^{-1}\left(\alpha\right) + \sqrt{n}\frac{\lambda_0 - \lambda_1}{\sqrt{\lambda_1}}\right] \end{aligned}$$

Les paramètres qui influent sur la performance asymptotique ($n \to \infty$) du test sont donc n, λ_0 et λ_1. On en déduit

$$\begin{aligned} \textit{Premier Cas} &: \quad n = 100, \lambda_0 = 1, \lambda_1 = 0.1 \\ PD &= F\left[\sqrt{10}F^{-1}\left(\alpha\right) + 9\sqrt{10}\right] \\ \textit{Deuxième Cas} &: \quad n = 100, \lambda_0 = 10, \lambda_1 = 1 \\ PD &= F\left[\sqrt{10}F^{-1}\left(\alpha\right) + 90\right] \end{aligned}$$

Comme F est une fonction croissante, on aura une meilleure performance dans le second cas.
5) On suppose désormais que les connaissances sur le paramètre λ sous les deux hypothèses H_0 et H_1 sont contenues dans la loi a priori $\lambda \sim \Gamma(\alpha, \beta)$ avec

$$\begin{cases} H_0 : \alpha = \alpha_0, \beta = 1 \\ H_1 : \alpha = \alpha_1, \beta = 1 \end{cases}$$

La vraisemblance s'écrit alors

$$\begin{aligned} f\left(x_1, ..., x_n\right) &= \int f(x_1, ..., x_n, \lambda) d\lambda \\ &= \int f\left(x_1, ..., x_n | \, \lambda\right) g(\lambda) d\lambda \\ &= \frac{\beta^\alpha}{\Gamma\left(\alpha\right)}\left(\prod_{i=1}^n x_i!\right)^{-1} \int_0^\infty \lambda^{n\overline{x}+\alpha-1} e^{-\lambda(n+\beta)} d\lambda \end{aligned}$$

En posant $u = \lambda\left(n + \beta\right)$, on obtient

$$\begin{aligned} f\left(x_1, ..., x_n\right) &= \frac{\beta^\alpha}{\Gamma\left(\alpha\right)}\left(\prod_{i=1}^n x_i!\right)^{-1} \frac{1}{\left(n+\beta\right)^{n\overline{x}+\alpha-1}} \int_0^\infty u^{n\overline{x}+\alpha-1} e^{-u} \frac{du}{n+\beta} \\ &= \frac{\beta^\alpha}{\Gamma\left(\alpha\right)}\left(\prod_{i=1}^n x_i!\right)^{-1} \frac{1}{\left(n+\beta\right)^{n\overline{x}+\alpha}} \int_0^\infty u^{n\overline{x}+\alpha-1} e^{-u} du \\ &= \frac{\beta^\alpha}{\Gamma\left(\alpha\right)}\left(\prod_{i=1}^n x_i!\right)^{-1} \frac{\Gamma\left(n\overline{x}+\alpha\right)}{\left(n+\beta\right)^{n\overline{x}+\alpha}} \end{aligned}$$

Le test de Bayes associé à ce problème s'écrit

$$\text{Rejet de } H_0 \text{ si} \frac{f\left(x_1, ..., x_n \middle| H_1\right)}{f\left(x_1, ..., x_n \middle| H_0\right)} > S_\alpha$$

c'est-à-dire

$$\text{Rejet de } H_0 \text{ si } \frac{\Gamma\left(n\overline{x} + \alpha_1\right)}{\Gamma\left(n\overline{x} + \alpha_0\right)} \frac{(n+1)^{\alpha_0}}{(n+1)^{\alpha_1}} \frac{\Gamma\left(\alpha_0\right)}{\Gamma\left(\alpha_1\right)} > S_\alpha$$

Finalement, on rejette H_0 si

$$\frac{\Gamma\left(n\overline{x} + \alpha_1\right)}{\Gamma\left(n\overline{x} + \alpha_0\right)} > K_\alpha$$

$\mathbf{3}^{\grave{e}me}$ Partie

1) Le paramètre λ étant inconnu, on l'estime à l'aide de la méthode du maximum de vraisemblance et on obtient

$$\widehat{\lambda}_{\text{MV}} = \overline{X}$$

On construit des classes comme par exemple

$$\begin{array}{lcl} \text{Classe 1} & : & X_i \in \{0,1,2\} \text{ de probabilité } p_1 \\ \text{Classe 2} & : & X_i \in \{3,4,5\} \text{ de probabilité } p_2 \\ \text{Classe 3} & : & X_i \geq 6 \text{ de probabilité } p_3 \end{array}$$

On calcule les effectifs des trois classes, notés n_1, n_2 et n_3 calculés à l'aide des données. On en déduit la statistique du test

$$\phi = \sum_{i=1}^{3} \frac{1}{np_i} \left(n_i - np_i\right)^2$$

On a estimé un paramètre donc la loi asymptotique de ϕ est une loi du χ_1^2. On se fixe α (par exemple $\alpha = 0.05$) qui vérifie la relation suivante

$$\begin{aligned} \alpha &= P\left[\text{rejeter } H_0 \middle| H_0 \text{ vraie}\right] \\ &= P\left[\phi > K_\alpha \middle| \phi \sim \chi_1^2\right] \\ &= 1 - G_1(K_\alpha), \end{aligned}$$

où G_1 est la fonction de répartition d'une loi du χ_1^2. On en déduit le seuil K_α comme suit

$$K_\alpha = G_1^{-1}\left(1 - \alpha\right)$$

On accepte l'hypothèse H_0 si $\phi < K_\alpha$ et on rejette H_0 dans le cas contraire

2) Ce test a un fonctionnement très similaire au test du χ^2 puisque

$$\begin{aligned} \alpha &= P\left[\text{rejeter } H_0 \middle| H_0 \text{ vraie}\right] \\ &= P\left[U = \frac{\sum_{i=1}^n \left(X_i - \overline{X}\right)^2}{\overline{X}} > S_\alpha \middle| U \sim \chi_{n-1}^2\right] \\ &= 1 - G_{n-1}(S_\alpha), \end{aligned}$$

où G_{n-1} est la fonction de répartition d'une loi du χ_{n-1}^2. On a donc

$$S_\alpha = G_{n-1}^{-1}\left(1 - \alpha\right)$$

On accepte l'hypothèse H_0 si $U < S_\alpha$ et on rejette H_0 dans le cas contraire.