Examen Lundi 6 Décembre 2004

Partie 1

1) La vraisemblance de l'échantillon $(X_1, ..., X_n)$ s'écrit :

$$\begin{aligned} L(x_1, ..., x_n; p) &= \prod_{i=1}^{n} P[X_i = x_i] \\ &= \prod_{i=1}^{n} p^{x_i} (1-p)^{1-x_i} \\ &= p^{\sum_{i=1}^{n} x_i} (1-p)^{n-\sum_{i=1}^{n} x_i} \end{aligned}$$

On a alors :

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

En étudiant le signe de $\frac{\partial \ln L(x_1,...,x_n;p)}{\partial \lambda}$, on obtient :

$$\begin{aligned} \frac{\partial \ln L(x_1, ..., x_n; p)}{\partial p} &\geq 0 \Longleftrightarrow \frac{1}{p}\sum_{i=1}^{n} x_i - \left[n - \sum_{i=1}^{n} x_i\right] \frac{1}{1-p} \geq 0 \\ &\Longleftrightarrow p \leq \frac{1}{n}\sum_{i=1}^{n} x_i \end{aligned}$$

La vraisemblance possède donc un maximum global unique obtenu pour $p = \frac{1}{n}\sum_{i=1}^{n} x_i$ d'où

$$\boxed{\hat{p}_{MV} = \frac{1}{n}\sum_{i=1}^{n} X_i}$$

Puisque les variables aléatoires X_i sont indépendantes, on sait que $n\hat{p}_{MV} = \sum_{i=1}^{n} X_i$ suit une loi binomiale $\mathcal{B}(n, p)$. On en déduit

$$\begin{aligned} E[n\hat{p}_{MV}] &= np \Rightarrow E[\hat{p}_{MV}] = p \\ Var[n\hat{p}_{MV}] &= np(1-p) \Rightarrow Var[\hat{p}_{MV}] = \frac{p(1-p)}{n} \end{aligned}$$

L'estimateur $\hat{p}_{MV}$ est donc un estimateur sans biais et convergent du paramètre p.

2) La borne de Cramer-Rao d'un estimateur non biaisé du paramètre p est définie par

$$\begin{aligned} BCR(p) &= \frac{-1}{E\left[\frac{\partial^2 \ln L(X_1,...,X_n;p)}{\partial p^2}\right]} \\ &= \frac{-1}{E\left[-\frac{1}{p^2}\sum_{i=1}^{n} X_i - \left[n - \sum_{i=1}^{n} X_i\right] \frac{1}{(1-p)^2}\right]} \\ &= \frac{1}{\frac{n}{p} + \frac{n}{(1-p)}} \end{aligned}$$

d'où

$$\boxed{BCR\left(p\right) = \frac{p(1-p)}{n}}$$

L'estimateur $\hat{p}_{MV}$ est donc l'estimateur efficace de p.

3) La loi a posteriori du paramètre p vérifie

$$f(p|x_1, ..., x_n) \propto p^{a-1}\left(1-p\right)^{b-1} p^{\sum_{i=1}^n x_i}\left(1-p\right)^{n-\sum_{i=1}^n x_i}$$

c'est-à-dire

$$f(p|x_1, ..., x_n) \propto p^{a-1+\sum_{i=1}^n x_i}\left(1-p\right)^{b-1+n-\sum_{i=1}^n x_i}$$

On voit donc que la loi a posteriori de p est une loi Beta de paramètres $a+\sum_{i=1}^n x_i$ et $b+n-\sum_{i=1}^n x_i$:

$$p|x_1, ..., x_n \sim \mathcal{B}e\left(a-1+\sum_{i=1}^n x_i, b-1+n-\sum_{i=1}^n x_i\right)$$

La maximisation de $\ln f(p|x_1, ..., x_n)$ conduit à

$$\begin{aligned}\frac{\partial \ln f(p|x_1, ..., x_n)}{\partial p} &\geq 0 \Longleftrightarrow \frac{a+\sum_{i=1}^n x_i - 1}{p} - \frac{b+n-\sum_{i=1}^n x_i - 1}{1-p} \geq 0 \\ &\Longleftrightarrow p \leq \frac{a-1+\sum_{i=1}^n x_i}{a+b-2+n}\end{aligned}$$

On en déduit que l'estimateur MAP de p est :

$$\boxed{\hat{p}_{MAP} = \frac{a-1+\sum_{i=1}^n X_i}{a+b-2+n}}$$

On remarque que

$$\hat{p}_{MAP} = \frac{a-1}{a+b-2+n} + \frac{1}{1+\frac{a+b-2}{n}}\hat{p}_{MV}$$

et donc asymptotiquement (n grand), l'estimateur $\hat{p}_{MAP}$ se comporte comme $\hat{p}_{MV}$. Pour n "petit", on a

$$\hat{p}_{MAP} \approx \frac{a-1}{a+b-2},$$

qui est le maximum de la loi a priori de p. Lorsque n grand, l'estimateur a tendance à oublier l'information a priori. Pour n petit, l'estimateur MAP a tendance à faire confiance aux informations a priori et à oublier le terme lié aux données.

Partie 2

1) Le test de Neyman-Pearson est défini par

$$\text{Rejet de } H_0 \text{ si} \quad \frac{L(x_1, ..., x_n; p_1)}{L(x_1, ..., x_n; p_0)} > S_\alpha$$

Mais

$$\begin{aligned}\frac{L(x_1, ..., x_n; p_1)}{L(x_1, ..., x_n; p_0)} &> S_\alpha \Leftrightarrow \frac{p_1^{\sum_{i=1}^n x_i}\left(1-p_1\right)^{n-\sum_{i=1}^n x_i}}{p_0^{\sum_{i=1}^n x_i}\left(1-p_0\right)^{n-\sum_{i=1}^n x_i}} > S_\alpha \\ &\Leftrightarrow \ln\left(\frac{p_1\left(1-p_0\right)}{p_0(1-p_1)}\right)\sum_{i=1}^n x_i > \nu_\alpha\end{aligned}$$

Pour $p_0 < p_1$, on a $\frac{p_1}{p_0} > 1$ et $\frac{1-p_0}{1-p_1} > 1$. Un test équivalent est donc

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

La moyenne de X_i étant p, puisque le paramètre p est "petit" sous l'hypothèse H_0 et "grand" sous l'hypothèse H_1, il est logique d'accepter H_1 (et donc de rejeter H_0) lorsque $\sum_{i=1}^n X_i$ est "grand". Pour $p_0 > p_1$, on a $\frac{p_1}{p_0} < 1$ et $\frac{1-p_0}{1-p_1} < 1$. Un test équivalent est donc

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

La moyenne de X_i étant p, puisque le paramètre p est "grand" sous l'hypothèse H_0 et "petit" sous l'hypothèse H_1, il est logique d'accepter H_1 (et donc de rejeter H_0) lorsque $\sum_{i=1}^n X_i$ est "petit".

2) A partir de la définition de α, on obtient

$$\begin{aligned} \alpha &= P\left[\text{Rejeter } H_0 \mid H_0 \text{ vraie}\right] \\ &= P\left[T < K_\alpha \mid p = p_0\right] \\ &= P\left[T < K_\alpha \mid T \sim B(n, p_0)\right] \end{aligned}$$

On pose

$$\begin{aligned} p_i &= P\left[T = i \mid T \sim B(n, p_0)\right] \\ &= C_n^i \left(\frac{1}{2}\right)^i \left(\frac{1}{2}\right)^{n-i} \\ &= \frac{C_n^i}{2^n} \end{aligned}$$

On en déduit

$$\begin{aligned} \text{si } K_\alpha &\leq 0 \text{ alors } \alpha = 0 \\ \text{si } K_\alpha &\in \left]0, 1\right] \text{ alors } \alpha = p_0 = \frac{1}{1024} \simeq 9.76\ 10^{-4} \\ \text{si } K_\alpha &\in \left]1, 2\right] \text{ alors } \alpha = p_0 + p_1 = \frac{1+10}{1024} \simeq 0.0107 \\ \text{si } K_\alpha &\in \left]2, 3\right] \text{ alors } \alpha = p_0 + p_1 + p_2 = \frac{1+10+45}{1024} \simeq 0.0547 \end{aligned}$$

Donc, le test le plus puissant vérifiant $\alpha \leq 0.05$ est obtenu pour

$$\boxed{K_\alpha = 2}$$

Il permet d'obtenir un test de risque $\alpha = 0.0107$.

3) Le théorème de la limite centrale permet d'obtenir

$$\frac{\sum_{i=1}^n X_i - np}{\sqrt{np(1-p)}} \underset{n \to +\infty}{\overset{L}{\to}} \mathcal{N}(0, 1)$$

Pour n grand, on pourra donc approcher la loi de T par une loi normale $\mathcal{N}(np, np(1-p))$. En utilisant cette loi approchée, on obtient

$$\begin{aligned}
\alpha &= P\left[\text{Rejeter } H_0 \mid H_0 \text{ vraie}\right] \\
&= P\left[T < K_\alpha \mid p = p_0\right] \\
&= P\left[T < K_\alpha \mid T \sim \mathcal{N}(np_0, np_0(1-p_0))\right] \\
&= P\left[U = \frac{T - np_0}{\sqrt{np_0(1-p_0)}} < \frac{K_\alpha - np_0}{\sqrt{np_0(1-p_0)}} \mid U \sim \mathcal{N}(0,1)\right] \\
&= F\left(\frac{K_\alpha - np_0}{\sqrt{np_0(1-p_0)}}\right)
\end{aligned}$$

où $F(x)$ est la fonction de répartition de la loi normale $\mathcal{N}(0,1)$. Ceci conduit à

$$K_\alpha = F^{-1}(\alpha)\sqrt{np_0(1-p_0)} + np_0$$

Pour $\alpha = 0.05$ et $n = 10$, les tables de la loi normale $\mathcal{N}(0,1)$ permettent d'obtenir

$$\frac{-K_\alpha + np_0}{\sqrt{np_0(1-p_0)}} = 1.645 \Rightarrow K_\alpha = -1.645\sqrt{np_0(1-p_0)} + np_0 \simeq 2.399$$

Même pour $n = 10$, la valeur du seuil obtenue avec la loi approchée (i.e. $K_\alpha \simeq 2.399$) est proche de la valeur obtenue avec la loi binomiale (i.e. $K_\alpha = 2$).

4) La puissance du test de Neyman-Pearson est définie par

$$\begin{aligned}
\pi &= 1 - \beta = P\left[\text{Rejeter } H_0 \mid H_1 \text{ vraie}\right] \\
&= P\left[T < K_\alpha \mid \lambda = \lambda_1\right] \\
&= P\left[T < K_\alpha \mid T \sim \mathcal{N}(np_1, np_1(1-p_1))\right] \\
&= P\left[U = \frac{T - np_1}{\sqrt{np_1(1-p_1)}} < \frac{K_\alpha - np_1}{\sqrt{np_1(1-p_1)}} \mid U \sim \mathcal{N}(0,1)\right] \\
&= F\left(\frac{K_\alpha - np_1}{\sqrt{np_1(1-p_1)}}\right)
\end{aligned}$$

L'application numérique conduit à

$$\pi \simeq F(1.47) \simeq 0.93$$

5) Les courbes COR sont définies par

$$\begin{aligned}
\pi &= F\left(\frac{K_\alpha - np_1}{\sqrt{np_1(1-p_1)}}\right) \\
&= F\left(\frac{F^{-1}(\alpha)\sqrt{np_0(1-p_0)} + np_0 - np_1}{\sqrt{np_1(1-p_1)}}\right) \\
&= F\left(\sqrt{n}\frac{p_0 - p_1}{\sqrt{p_1(1-p_1)}} + F^{-1}(\alpha)\sqrt{\frac{p_0(1-p_0)}{p_1(1-p_1)}}\right)
\end{aligned}$$

Les paramètres qui ont une influence sur la performance du test de Neyman-Pearson sont donc le nombre de points n et

$$\frac{p_0 - p_1}{\sqrt{p_1(1-p_1)}} \text{ et } \sqrt{\frac{p_0(1-p_0)}{p_1(1-p_1)}}$$

6) Le test du χ^2 ne peut se faire dans ce cas qu'avec deux classes

$$\begin{array}{lcl} \text{Classe } C_1 & : & \{0\} \\ \text{Classe } C_2 & : & \{1\} \end{array}$$

Les nombres d'observations de ces deux classes sont

$$n_1 = 4, n_2 = 6$$

et donc la statistique de test du χ^2 peut se calculer comme suit

$$\begin{aligned} \phi &= \sum_{i=1}^{2} \frac{(n_i - np_i)^2}{np_i} \\ &= \frac{1}{5}(1+1) = 0.4 \end{aligned}$$

Le seuil se calcule à partir du risque α :

$$\begin{aligned} \alpha &= P\left[\text{Rejeter } H_0 \mid H_0 \text{ vraie}\right] \\ &= P\left[\phi > S_\alpha \mid \phi \sim \chi_1^2\right] \end{aligned}$$

Les tables donnent

$$\begin{aligned} \alpha &= 0.01 \Rightarrow S_\alpha = 6.635 \\ \alpha &= 0.05 \Rightarrow S_\alpha = 3.841 \end{aligned}$$

et donc on accepte l'hypothèse H_0 (les observations sont distribuées suivant la loi de Bernoulli avec $p = \frac{1}{2}$) avec les riques $\alpha = 0.01$ et $\alpha = 0.05$.

7) Le risque α est défini comme suit

$$\begin{aligned} \alpha &= P\left[\text{Rejeter } H_0 \mid H_0 \text{ vraie}\right] \\ &= P\left[|N_0 - N_1| > 4 \mid p = \frac{1}{2}\right] \\ &= 1 - P\left[|N_0 - N_1| \leq 4 \mid p = \frac{1}{2}\right] \end{aligned}$$

Sachant $p = 0.5$, le nombre de 0 du tableau suit la loi Binomiale $\mathcal{B}\left(10, \frac{1}{2}\right)$, d'où

$$\begin{cases} N_0 - N_1 = -4 \\ N_0 + N_1 = 10 \end{cases} \Leftrightarrow N_0 = 3 \text{ d'où } P\left[|N_0 - N_1| = -4 \mid p = \frac{1}{2}\right] = \frac{C_{10}^3}{2^{10}}$$

$$\begin{cases} N_0 - N_1 = -2 \\ N_0 + N_1 = 10 \end{cases} \Leftrightarrow N_0 = 4 \text{ d'où } P\left[|N_0 - N_1| = -2 \mid p = \frac{1}{2}\right] = \frac{C_{10}^4}{2^{10}}$$

$$\begin{cases} N_0 - N_1 = 0 \\ N_0 + N_1 = 10 \end{cases} \Leftrightarrow N_0 = 5 \text{ d'où } P\left[|N_0 - N_1| = 0 \mid p = \frac{1}{2}\right] = \frac{C_{10}^5}{2^{10}}$$

$$\begin{cases} N_0 - N_1 = 2 \\ N_0 + N_1 = 10 \end{cases} \Leftrightarrow N_0 = 6 \text{ d'où } P\left[|N_0 - N_1| = 2 \mid p = \frac{1}{2}\right] = \frac{C_{10}^6}{2^{10}}$$

$$\begin{cases} N_0 - N_1 = 4 \\ N_0 + N_1 = 10 \end{cases} \Leftrightarrow N_0 = 7 \text{ d'où } P\left[|N_0 - N_1| = 4 \mid p = \frac{1}{2}\right] = \frac{C_{10}^7}{2^{10}}$$

On en déduit

$$\begin{aligned} \alpha &= 1 - \frac{C_{10}^3 + C_{10}^4 + C_{10}^5 + C_{10}^6 + C_{10}^7}{2^{10}} \\ &= 1 - \frac{120 + 210 + 252 + 210 + 120}{1024} \simeq 0.109 \end{aligned}$$

Barème

1ère partie : 10 pts

1) 2 pts

2) 1 pt

3) 3.1) 1pt, 3.2) 1pt, 3.3) calcul de moyenne 1pt + calcul de variance 1pt + λ^* sans biais, cv, non efficace : 1pt

4) 1 + 1 (MAP + analyse)

2ème partie :

1) Test de NP : 1pt + explications moyenne : 1pt

2) K_α : 1pt

3) π : 1pt, SNR + analyse : 1pt

4) Seuil + puissance 1pt + 1pt

5) Test du chi2 : classes 1pt, calcul de ϕ 1pt, calcul du seuil 1pt, conclusions 1pt.