Université Paris Dauphine
Département MIDO
L3 - Statistique Mathématique

Année 2009-2010

# Exercices supplémentaires

Source : Examen de rattrapage 2009 et TD M1 MLG 2009

**Exercice 1 :** Nous observons $n$ terminaux $X$ qui ont été mis hors service. Pour le terminal $i$, nous observons :
– $y_i^1$ sa durée de vie.
– $y_i^2$ le nombre de révisions qu'il a subi.
– $y_i^3$ le nombre de révisions avec changement de pièces.
Nous supposons que :
– $Y_i^1$ suit une loi exponentielle de paramètre $\gamma$.
– $Y_i^2$ suit une loi de Poisson de paramètre $\lambda y_i^1$. Il s'agit de la loi conditionnelle de $Y_i^2/Y_i^1$.
– $Y_i^3$ suit une loi binomiale de paramètre $(y_i^2, p)$. Il s'agit de la loi conditionnelle de $Y_i^3/Y_i^2, Y_i^1$
Les variable $Y_i^j$, $i = 1, \dots, n$, $j = 1, 2, 3$, sont supposées indépendantes et nous posons $\Theta = (\gamma, \lambda, p)$.

1. Donner la log-vraisemblance $\ell_n(Y, \theta)$ de $Y = (Y_1, \dots, Y_n)$ où $Y_i = (Y_i^1, Y_i^2, Y_i^3)$ et vérifier que $\mathbb{E}_\theta[\frac{\partial}{\partial\theta}\ell_n(Y, \theta)] = 0$.
2. Déterminer les estimateurs du maximum de vraisemblance des paramètres $\gamma$, $\lambda$ et $p$ ainsi que leur distribution asymptotique. En déduire des intervalles de confiance de niveau asymptotique $1 - \alpha$ avec $\alpha \in ]0, 1[$ pour les paramètres $\gamma$, $\lambda$ et $p$.
3. Determiner la statistique de Wald et celle du rapport de vraisemblance pour l'hypothèse $H_0 : \gamma = \lambda$ et proposer un test consistant de niveau asymptotique $\alpha$.

**Exercice 2 :** On observe $X_1$, de loi $U[0, 1]$ sous $H_0$, ou $U[2, 3]$ sous $H_1$. Proposer un test de l'hypothèse $H_0$ contre l'alternative $H_1$ et calculer ses risques de première et seconde espèce.

**Exercice 3 :** On suppose que l'on observe $X_1, \dots, X_n$, i.i.d. de loi $\mathcal{N}(\mu, 1)$. On veut tester $H_0 : \mu = 0$ contre $H_1 : \mu = m < 0$.

1. Donner la forme du test de Neyman-Pearson de niveau $\alpha \in (0, 1)$ pour ce problème.
2. Calculer la puissance $m \in \mathbb{R}_- \mapsto \pi_n(m)$ de ce test et tracer son graphe. Etudier la convergence simple de $\pi_n$ lorsque $n$ tend vers $+\infty$. Peut–on parler de convergence uniforme sur $\mathbb{R}_-$ ?
3. On considère l'alternative dépendant de $n$

$$H_1 : \mu = -Cn^{-\gamma}$$

avec $C > 0$ et $\gamma \in \mathbb{R}$. Etudier le comportement de la puissance du test sur $H_1$ en fonction de $\gamma$ lorsque $n$ tend vers $+\infty$.

**Exercice 4 :** Un marchand de graines a l'habitude de fournir à un ingénieur agronome un mélange de 6 types de graines qui entrent toutes dans le mélange avec les même proportions (1/6 chacune). Quel que soit le type, les graines ont sensiblement le même poids et la même taille. Elles diffèrent d'un type à l'autre par leurs formes. On désigne ces 6 types par A, B, C, D, E, F.

Un jour, l'ingénieur lui demande de préparer un mélange spécial dans lequel il y aura 1/4 de graines A, 1/2 de B, les autres graines restant en proportions inchangées. Le jour de la livraison, un employé un peu trop zélé a déplacé les sacs. Le grainetier, fort perplexe, croit que le premier sac est bon, mais il n'en est pas certain. Comment faire ? Évidemment, on pourrait trier les graines de ce sac, mais ceci serait fastidieux. On propose ici un test pour apporter une solution à ce problème : il s'agit, à partir d'un échantillon de graines du sac, de décider entre deux hypothèses :

"ce sac est bon" contre "c'est un sac ordinaire".

1. A partir des observations de la forme "la $i$-ième graine tirée est de type A ou non", formuler un modèle statistique et poser le problème de test d'hypothèses.
2. Quelle est, sous chacune des ces hypothèses, la loi du nombre $X$ de graines de type A dans un échantillon de taille $n$ ? Donner les valeurs de $\mathbb{E}_\theta(X)$ et $\operatorname{Var}_\theta(X)$ correspondant à chaque cas.
3. Donner la forme du test Neyman–Pearson de niveau $\alpha$ pour tester l'hypothèse "ce sac est bon" contre l'alternative "c'est un sac ordinaire".
4. En utilisant une approximation de la loi de la variable aléatoire $X$ par une loi normale (sous chacune des hypothèses), déterminer la taille minimale $n_0$ de l'échantillon et la région critique pour avoir un test de niveau de confiance $\alpha$ et de puissance plus grande que $1-\beta$ avec $0 < \alpha = \beta < 1$. Que vaut la région de rejet et $n_0$ pour $\alpha = \beta = 0.05$ ? Conclure.
   (On rappelle que pour $Z \sim \mathcal{N}(0,1)$, on a $P_\theta(Z > 1,64) \approx 0,05$).

**Exercice 5 :** Soit $X_1, \ldots, X_n$ un $n$-échantillon distribué suivant une loi de Bernoulli de paramètre $0 < \theta < 1$. Soit

$$\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i.$$

1. Montrer qu'un estimateur $T = T(X_1, \ldots, X_n)$ est un estimateur de variance minimale de son espérance si et seulement si c'est une fonction de $\overline{X}$.
2. Montrer qu'une fonction $g(\cdot)$ de $\theta$ admet un estimateur sans biais si et seulement si $g(\cdot)$ est un polynôme de degré inférieur ou égal à $n$.
3. Donner la formule explicite d'un estimateur de variance minimale $T_k$ de $\theta^k$ pour $1 \leq k \leq n$.
4. Ces estimateurs de $\theta^k$ sont ils efficaces ?