

M1 – Master Recherche Science et Technologie
MATH 2M06, Probabilités et statistiques

Feuille de TD n° 6

2006

1. (*Le pouvoir d'une minorité décidée*) 1.000.000 d'électeurs participent à une élection entre deux candidats A et B . Parmi eux, 2.000 électeurs connaissent le candidat A à cause de l'organisation de la campagne électorale, et votent en bloc pour lui. Les 998.000 électeurs restants sont plus ou moins indécis et prennent leur décision indépendamment les uns des autres en jouant à pile ou face avec une pièce non faussée. Quelle est la probabilité p_A d'une victoire du candidat A ?
2. (*Estimateur de variance minimum*)
Soit X_1, \dots, X_n un échantillon i.i.d. d'une loi de moyenne μ et de variance σ^2 .
 - (a) Donner la condition sur les constantes réelles a_1, \dots, a_n pour que $\sum_{k=1}^n a_k X_k$ soit un estimateur sans biais de μ .
 - (b) Parmi les estimateurs sans biais de cette forme déterminer celui qui est de variance minimum. Calculer sa variance.
3. (*Estimateur du maximum de vraisemblance*) Soit un modèle statistique donné par la famille de lois $(P_\vartheta : \vartheta \in \Theta)$ sur $(\mathbb{R}^k, \mathcal{B}(\mathbb{R}^k))$ de densités f_ϑ , c.à.d. $P_\vartheta(dx) = f_\vartheta(x)dx$, où Θ est l'espace des paramètres. Si l'observation est constitué d'un n -échantillon (X_1, \dots, X_n) de variables de loi P_ϑ , on appelle *vraisemblance du modèle* la variable aléatoire

$$I_\vartheta(X_1, \dots, X_n) = f_\vartheta(X_1) \cdot \dots \cdot f_\vartheta(X_n).$$

On appelle *statistique* toute fonction mesurable $T(X_1, \dots, X_n)$ de (X_1, \dots, X_n) et *estimateur du maximum de vraisemblance* toute statistique $\hat{\vartheta} = T(X_1, \dots, X_n)$ à valeurs dans Θ , qui maximise la fonction $\vartheta \mapsto I_\vartheta(X_1, \dots, X_n)$.

- (a) Soit (X_1, \dots, X_n) un n -échantillon des v.a. indépendantes, chacune suivant une loi exponentielle de paramètre $\vartheta > 0$. Déterminer l'estimateur de maximum de vraisemblance de ϑ .
 - (b) Soit (X_1, \dots, X_n) un n -échantillon des v.a. de loi uniforme $\mathcal{U}([0, \vartheta])$. Déterminer l'estimateur de maximum de vraisemblance de ϑ .
 - (c) Soit (X_1, \dots, X_n) un n -échantillon des v.a. de loi $\mathcal{N}(m, \sigma^2)$. Suivant que m ou σ^2 ou les deux sont inconnus, déterminer les estimateurs de maximum de vraisemblance.
4. (*Inégalité de Cramer-Rao*)
Soient X_1, \dots, X_n de v.a. i.i.d., chacune suivant une loi de densité $f_\vartheta > 0$ où le paramètre $\vartheta \in \Theta$ est inconnu. On suppose que $\Theta \subset \mathbb{R}$ est un ouvert et que $\vartheta \mapsto f_\vartheta(x)$ est de classe C^2 . Dans cette situation l'*information de Fisher* est définie par

$$I(\vartheta) = \mathbb{E}_\vartheta \left[\left(\frac{d}{d\vartheta} \log f_\vartheta(X_1) \right)^2 \right], \quad \vartheta \in \Theta,$$

où l'espérance est calculée par rapport à $P_\vartheta(dx) = f_\vartheta(x)dx$. On suppose que $0 < I(\vartheta) < \infty$ pour tout $\vartheta \in \Theta$.

(a) Montrer que

$$I(\vartheta) = -\mathbb{E}_{\vartheta} \left[\frac{d^2}{d\vartheta^2} \log f_{\vartheta}(X_1) \right].$$

(b) Montrer que si $\hat{\vartheta} = T(X_1, \dots, X_n)$ est un estimateur non biaisé de ϑ , c.à.d. $\mathbb{E}_{\vartheta}[\hat{\vartheta}] = \vartheta$ pour tout $\vartheta \in \Theta$, alors

$$\text{var}_{\vartheta}(\hat{\vartheta}) \geq \frac{1}{nI(\vartheta)}.$$

(c) Un estimateur non biaisé $\hat{\vartheta} = T(X_1, \dots, X_n)$ du paramètre ϑ sera dit *efficace* si $\text{var}_{\vartheta}(\hat{\vartheta}) = 1/(nI(\vartheta))$. Montrer que si (X_1, \dots, X_n) est un n -échantillon de loi $\mathcal{N}(m, \sigma^2)$, alors $\bar{X} = (X_1 + \dots + X_n)/n$ est un estimateur efficace du paramètre m , et $S^2 = \frac{1}{n} \sum_{i=1}^n (X_i - m)^2$ est un estimateur efficace du paramètre σ^2 . Est-ce que $S_1^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2$ est un estimateur efficace du paramètre σ^2 ?

5. (*Intervalles de confiance asymptotiquement*)

Soit X_1, X_2, \dots une suite de v.a. i.i.d. de loi commune de Poisson de paramètre $\lambda > 0$. Pour $\alpha \in]0, 1[$ donné, construire une suite d'intervalles $I_n = I_n(X_1, \dots, X_n)$ tels que

$$P\{\lambda \in I_n\} \rightarrow 1 - \alpha, \quad \text{quand } n \rightarrow \infty.$$

6. (*Fille ou garçon ?*) En Suisse, entre 1871 et 1900, sont nés 1.359.670 garçons et 1.286.086 filles. Qu'est-ce que vous pensez de l'hypothèse que les filles et les garçons naissent avec la même probabilité ?

7. (*Faiseurs de pluie*) Des relevés effectués pendant de nombreuses années ont permis d'établir que le niveau naturel des pluies dans la Beauce en millimètres par an suit une loi normale de moyenne 600 et de variance 50^2 . Des entrepreneurs, surnommés faiseurs de pluie, prétendaient pouvoir augmenter de 50 mm le niveau moyen de pluie, ceci par insémination des nuages au moyen d'iodure d'argent. Leur procédé fut mis à l'essai entre 1991 et 1999 et on releva les hauteurs de pluies suivantes :

année	1991	1992	1993	1994	1995	1996	1997	1998	1999
mm	510	614	780	512	501	534	603	788	650

Que pouvait-on en conclure ?

8. (*Un peu de génétique*) Un croisement entre roses rouges et blanches a donné en seconde génération des roses rouges, roses et blanches. Sur un échantillon de taille 600, on a trouvé les résultats suivants :

couleur	rouges	roses	blanches
effectifs	141	315	144

Ces résultats expérimentaux sont-ils conformes au modèle mendélien au niveau 0.05 ?