Statistique

Examen du 22 juin 2001

Sans document (durée deux heures). Trois exercices indépendants.

Exercice 1

Soit

$$\mathcal{Q} = \left\{(x, y);\ 0 \le y,\ 0 \le x \le e^{-(y/2x)^2}\right\}$$

1. Tracer le contour de cet ensemble.
2. On veut montrer que la surface de $\mathcal{Q}$ est $|\mathcal{Q}| = \sqrt{\pi/8}$.

   (a) Montrer que, si (X, Y) est de loi uniforme sur $\mathcal{Q}$, le couple $(X, Z = Y/X)$ suit la loi de densité

   $$\frac{1}{|\mathcal{Q}|}\, x\, \mathbb{I}_{0\le z}\, \mathbb{I}_{0\le x\le \exp(-z^2/4)}\,.$$

   (b) En déduire que $|\mathcal{Q}| = \sqrt{\pi/8}$ par intégration de cette densité.

3. Soit la fonction réelle à valeurs réelles $r(z) = \exp(-z^2/2)$. On pose

   $$x^\star = \sup_{z\ge 0} r^{1/2}(z), \quad y_\star = \inf_{z\ge 0}\left[z r^{1/2}(z)\right], \quad y^\star = \sup_{z\ge 0}\left[z r^{1/2}(z)\right]\,.$$

   Montrer que

   $$x^\star = 1, \quad y_\star = 0, \quad y^\star = \sqrt{2/e}\,.$$

4. Soit le pavé

   $$\mathcal{P} = [0, x^\star] \times [y_\star, y^\star]\,.$$

   Montrer que $\mathcal{Q} \subset \mathcal{P}$ et représenter $\mathcal{P}$ sur le graphe de $\mathcal{Q}$.

T.S.V.P/

5. On considère l'algorithme suivant :

(a) Générer (X, Y) suivant la loi uniforme sur $\mathcal{P}$.

(b) Prendre $Z = Y/X$ si $(X, Y) \in \mathcal{Q}$;
sinon, recommencer en (a)

Déduire de la question 2.(a) que Z est distribué suivant

$$f(z) = \sqrt{2/\pi}\, e^{-z^2/2}\, \mathbb{I}_{z \geq 0}\,.$$

6. En utilisant le rapport des surfaces, $|\mathcal{P}|/|\mathcal{Q}|$, déterminer le nombre moyen d'uniformes utilisées pour cette génération et comparer avec l'algorithme de Box–Müller.

Exercice 2

Une technique de dépistage d'une maladie rare utilisée durant la Seconde Guerre Mondiale est la suivante : on mélange les échantillons sanguins de n soldats testés et on effectue le test sur le mélange. Si le test est négatif, les n soldats sont tous sains; si le test est positif, au moins un soldat est atteint et on doit ensuite effectuer n tests.

1. Soient $X_1, \ldots, X_n$ des variables de Bernoulli iid, de paramètre $\theta = P(X_1 = 1)$.

   (a) Donner la probabilité que *tous* les X_i soient nuls.

   (b) En déduire que la technique de mélange ci-dessus est plus économique *en moyenne* que de tester directement les soldats un par un si

   $$1 + [1 - (1-\theta)^n]\, n < n$$

2. Dans ce contexte, on recueille m mélanges indépendants portant chacun sur n soldats. Les résultats des tests sur ces mélanges sont $y_1, \ldots, y_m$, qui prennent les valeurs 0 ou 1 suivant que les tests soient négatifs ou positifs, respectivement.

   (a) Rappeler la loi de Y_1. (On pourra utiliser la question 1 (a) ci-dessus.)

(b) Montrer que la moyenne empirique

$$\bar{Y}_m = \frac{1}{m} \sum_{i=1}^{m} Y_i$$

est exhaustive pour θ.

(c) Donner la vraisemblance du modèle et en déduire que

$$\hat{\theta}_m = 1 - \sqrt[n]{1 - \bar{Y}_m}$$

est l'estimateur du maximum de vraisemblance de θ. vers θ.

Exercice 3

Soit $X_1, \ldots, X_n$ un échantillon iid d'une loi uniforme $\mathcal{U}([0, \theta])$.

1. Donner la vraisemblance et montrer qu'elle ne dépend que de la statistique

$$X_{(n)} = \max\{X_1, \ldots, X_n\}$$

2. En déduire que $X_{(n)}$ est exhaustive.

3. Donner la fonction de répartition et la densité de $X_{(n)}$ et vérifier que cette densité est proportionnelle à la vraisemblance précédente.

4. Montrer que

$$\mathbb{E}_\theta[X_{(n)}] = \frac{n}{n+1}\,\theta\,, \quad \mathrm{var}_\theta X_{(n)} = \frac{n}{(n+2)\,(n+1)^2}\,\theta^2$$

5. En déduire que $X_{(n)}$ converge en probabilité vers θ. (*Indice :* Utiliser l'inégalité de Cebycev.)

6. On cherche à trouver la loi limite de

$$n\left(\theta - X_{(n)}\right) .$$

(a) Remarquer que n est la bonne vitesse pour une convergence en loi en montrant que

$$\mathrm{var}_\theta\left[n\left(\theta - X_{(n)}\right)\right]$$

converge vers la constante non nulle θ^2.

T.S.V.P/

(b) Calculer, pour $\delta > 0$ quelconque,

$$P_\theta\left(n\left\{\theta - X_{(n)}\right\} < \delta\right)$$

(c) En utilisant *sans le démontrer* le résultat d'analyse

$$\lim_{n\to\infty}\left(1 - \frac{a}{n}\right)^n = e^{-a}\,,$$

montrer que la loi limite de $n(\theta - X_{(n)})$ est la loi exponentielle $\mathcal{E}xp(1/\theta)$.