

Université d'Orléans - Licence Economie et Gestion

Statistique Mathématique

C. Hurlin. Correction de l'Examen Terminal Janvier 2010

Exercice 1 *Maximum de Vraisemblance et loi exponentielle. Barème : 5 points*

Question 1 (1 point) D'après la définition de la densité, celle-ci n'est définie que si $\theta > 0$. Autre justification possible : θ étant la durée moyenne des appels, on a $\theta \geq$, par ailleurs le cas $\theta = 0$ est exclu car sinon $f_D(d; \theta)$ n'est pas défini.

Question 2 (1.5 points) La vraisemblance associée au N -échantillon aléatoire $\{D_1, \dots, D_N\}$ s'écrit :

$$\begin{aligned} L(D_1, \dots, D_N; \theta) &= \prod_{i=1}^N f_D(D_i; \theta) \\ &= \frac{1}{\theta^N} \exp\left(-\frac{1}{\theta} \sum_{i=1}^N D_i\right) \end{aligned} \quad (1)$$

Dès lors, la log-vraisemblance associée au N -échantillon aléatoire $\{D_1, \dots, D_N\}$ s'écrit :

$$\ln L(D_1, \dots, D_N; \theta) = -N \ln(\theta) - \frac{1}{\theta} \sum_{i=1}^N D_i \quad (2)$$

Question 3 (1.5 points) Soit $\hat{\theta}$ l'estimateur du maximum de vraisemblance du paramètre θ . Ce dernier vérifie :

$$\hat{\theta} = \underset{\theta > 0}{\text{Arg max}} \ln L(D_1, \dots, D_N; \theta) \quad (3)$$

La condition nécessaire est :

$$\left. \frac{\partial \ln L(D_1, \dots, D_N; \theta)}{\partial \theta} \right|_{\theta = \hat{\theta}} = -\frac{N}{\hat{\theta}} + \frac{1}{\hat{\theta}^2} \sum_{i=1}^N D_i = 0 \quad (4)$$

Si $\hat{\theta} > 0$, on a alors :

$$\hat{\theta} = \frac{1}{N} \sum_{i=1}^N D_i = \bar{D}_N \quad (0.5 \text{ point}) \quad (5)$$

La condition suffisante est alors la suivante (1 point) :

$$\begin{aligned} \left. \frac{\partial^2 \ln L(D_1, \dots, D_N; \theta)}{\partial \theta^2} \right|_{\theta = \hat{\theta}} &= \frac{N}{\hat{\theta}^2} - \frac{2}{\hat{\theta}^3} \sum_{i=1}^N D_i \\ &= \frac{N}{\bar{D}_N^2} - \frac{2}{\bar{D}_N^3} \times N \times \bar{D}_N \\ &= -\frac{N}{\bar{D}_N^2} < 0 \end{aligned} \quad (6)$$

Question 4 (1 point) On a :

$$E(\hat{\theta}) = \frac{1}{N} \sum_{i=1}^N E(D_i) = \theta \quad (7)$$

L'estimateur $\hat{\theta}$ est sans biais (0.5 point).

$$V(\hat{\theta}) = \frac{1}{N^2} V\left(\sum_{i=1}^N D_i\right) = \frac{1}{N^2} \sum_{i=1}^N V(D_i) = \frac{1}{N^2} \sum_{i=1}^N \theta^2 = \frac{\theta^2}{N} \quad (8)$$

car les variables D_i sont *i.i.d.* de même loi que D . Par conséquent :

$$E(\hat{\theta}) = \theta \quad (9)$$

$$\lim_{N \rightarrow \infty} V(\hat{\theta}) = 0 \quad (10)$$

L'estimateur du maximum de vraisemblance $\hat{\theta}$ est convergent (en probabilité) :

$$\hat{\theta} \xrightarrow[N \rightarrow \infty]{p} \theta \quad (0.5 \text{ point}) \quad (11)$$

Exercice 2 Tests paramétriques: Barème : 11 points

Question 1 (1 point) Par définition :

$$\alpha = \Pr(W \mid H_0 \text{ vraie}) \quad (12)$$

Donc :

$$\alpha = \Pr\left(t_b \geq A \mid t_b \underset{H_0}{\rightsquigarrow} t(T-1)\right) \quad 0.5 \text{ point} \quad (13)$$

Soit $G_{T-1}(x)$ la fonction de répartition de la loi de Student à $T-1$ degrés de liberté. Il vient :

$$\alpha = 1 - G_{T-1}(A) \quad (14)$$

D'où l'on tire au final :

$$A = G_{T-1}^{-1}(1 - \alpha) \quad 0.5 \text{ point} \quad (15)$$

Question 2 (1 point) Dans ce cas, on a :

$$A = G_{29}^{-1}(0.95) \quad (16)$$

où $G_{29}(x)$ désigne la fonction de répartition de la loi de Student à 29 degrés de liberté. Il vient :

$$A = 1.699 \quad 0.5 \text{ point} \quad (17)$$

Donc si la réalisation de t_b est égale à 2.34, pour un risque de première espèce de 5%, la société doit rejeter l'hypothèse nulle selon laquelle l'élasticité des ventes est égale à 1.

Question 3 (1 point) La puissance du test est définie par :

$$\begin{aligned} \text{puissance} &= \Pr(W \mid H_1 \text{ vraie}) \\ &= \Pr(t_b \geq A \mid b > 1) \quad 0.5 \text{ point} \end{aligned} \quad (18)$$

Mais sous H_1 on ne connaît pas la vraie loi (0.5 point) de la statistique de Student (lorsque $b > 1$). Donc on ne peut pas établir la fonction puissance à partir des seules hypothèses de l'énoncé.

Question 4 (1 point) La courbe de puissance dans le repère $(b, \text{puissance})$ passe par les points $(1, 0.10)$; $(1.5, 0.17)$; $(2, 0.36)$ et $(3, 0.95)$.

Question 5 (2 points) On considère finalement le test bilatéral :

$$H_0 : b = 1 \quad (19)$$

$$H_1 : b \neq 1 \quad (20)$$

La région d'acceptation de ce test est fondée sur l'intersection de régions d'acceptation des tests unilatéraux de niveau $\alpha/2$ (0.5 point)

$$H_0 : b = 1 \quad (21)$$

$$H_1 : b > 1 \quad (22)$$

et

$$H_0 : b = 1 \quad (23)$$

$$H_1 : b < 1 \quad (24)$$

Les régions critiques associées à ces deux tests sont les suivantes (0.5 point):

$$W = \left\{ q_{A,1}, \dots, q_{B,T} \mid t_b \geq G_{T-1}^{-1} \left(\frac{\alpha}{2} \right) \right\} \quad (25)$$

$$W = \left\{ q_{A,1}, \dots, q_{B,T} \mid t_b < G_{T-1}^{-1} \left(1 - \frac{\alpha}{2} \right) \right\} \quad (26)$$

Dès lors, la région d'acceptation du test unilatéral est de la forme :

$$\bar{W} = \left\{ q_{A,1}, \dots, q_{B,T} \mid G_{T-1}^{-1} \left(\frac{\alpha}{2} \right) < t_b < G_{T-1}^{-1} \left(1 - \frac{\alpha}{2} \right) \right\} \quad 1 \text{ point}$$

où $G_{T-1}(x)$ désigne la fonction de répartition de la loi de Student à $T - 1$ degrés de liberté.

Question 6 (1 point) Dans ce cas, on a :

$$\bar{W} = \left\{ q_{A,1}, \dots, q_{B,T} \mid G_{\infty}^{-1} \left(\frac{0.05}{2} \right) < t_b < G_{\infty}^{-1} \left(1 - \frac{0.05}{2} \right) \right\} \quad 0.5 \text{ point} \quad (27)$$

et donc :

$$\bar{W} = \{ q_{A,1}, \dots, q_{B,T} \mid -1.96 < t_b < 1.96 \} \quad (28)$$

Autrement dit si $|t_b| > 1.96$, au seuil de risque de 5%, on rejette l'hypothèse nulle selon laquelle $b = 1$ (1 point)

Question 7 (2 points) La région d'acceptation du test unilatéral est de la forme :

$$\bar{W} = \left\{ q_{A,1}, \dots, q_{B,T} \mid G_{\infty}^{-1} \left(\frac{\alpha}{2} \right) < t_b < G_{\infty}^{-1} \left(1 - \frac{\alpha}{2} \right) \right\} \quad (29)$$

ou encore :

$$\bar{W} = \left\{ q_{A,1}, \dots, q_{B,T} \mid |t_b| \leq G_{\infty}^{-1} \left(1 - \frac{\alpha}{2} \right) \right\} \quad (30)$$

Ici on a :

$$|t_b| = 1.645 \quad (31)$$

Donc le seuil α minimum est égal à :

$$G_{\infty}(|t_b|) = 1 - \frac{\alpha}{2} \quad (32)$$

ou encore :

$$\alpha = 2 \times [1 - G_{\infty}(|t_b|)] = 0.10 \quad (33)$$

Exercice 3 *Test d'Adéquation : Barème : 5 points.*

Tableau des effectifs théoriques (1 point) :

X_i	0	1	2	3	4	5	6	7
$\Pr(X_i = k)$	0.1250	0.1875	0.1875	0.1562	0.1172	0.0820	0.0547	0.0352
Effectifs théoriques	12.5	18.75	18.75	15.62	11.72	8.20	5.47	3.52

Après regroupement de classe, les effectifs théoriques et empiriques sont donc les suivants (2 points) :

X_i	0	1	2	3	4	5	6 et plus
Effectifs empiriques	16	18	19	16	11	9	11
Effectifs théoriques	12.5	18.75	18.75	15.62	11.72	8.20	8.99

Par conséquent, la statistique du chi-deux vaut :

$$D_{(F_N, F)} = \sum_{i=1}^7 \frac{(N_i - N * p_i)^2}{N * p_i} = 1.596 \quad (0.5 \text{ point})$$

où N_i désigne les effectifs empiriques de chaque classe et p_i la probabilité théorique associée à chaque classe. Sous H_0 , on a :

$$D_{(F_N, F)} \sim \chi^2(6) \quad 0.5 \text{ point} \quad (34)$$

Le seuil critique à 95% étant égal à 12.59, pour un risque de première espèce de 5%, on ne peut pas rejeter l'hypothèse nulle selon laquelle la loi de X est une loi binomiale négative $\text{NegBin}(3, 0.5)$. (1 point)