ACT2040 : actuariat IARD 2

Examen Intra - session automne

Mardi 1er Novembre 2011

Instructions générales

L'examen débute à 9 :00 et se termine à 12 :00.

Aucune feuille aide-mémoire n'est autorisée.

Sueles les modèles de calculatrices suivants sont acceptés : Texas Instruments BA-35, BA II Plus, BA II Plus Professional ,TI-30Xa, TI-30X ISS (ou IIB), TI-30XS (ou XB) MultiView.

Inscrivez vos réponses dans le cahier-réponses.

Pour chaque question, des points seront attribués à la justification et à la clarté de la solution et des calculs.

Les trois exercices sont indépendants. L'examen contient 40 points, et compte pour 20% de la note finale du cours.

Notations

Pour rappel, une loi de la famille exponentielle s'écrit

$$f(y;\theta,\varphi)=\exp\left(\frac{y\theta-b(\theta)}{\varphi}+c(y,\varphi)\right)$$

On notera $h(\cdot)$ l'inverse de la fonction de lien canonique, au sens où $\mu=\mathbb{E}(Y)=h(\theta)$, et $g(\cdot)$ la fonction de lien canonique, au sens où $g=h^{-1}$. On appelera V la fonction variance, au sens où $Var(Y)=V(\mathbb{E}(Y))$.
Dans les exercices, nous considérerons des échantillons $(\boldsymbol{Y},\boldsymbol{X})=((Y_1,\boldsymbol{X}_1),\cdots,(Y_n,\boldsymbol{X}_n))$ où Y_i a pour loi $f(y;\theta_i,\varphi)$, où $\theta_i=\boldsymbol{X}_i'\boldsymbol{\beta}$. On note $\mathcal{L}(\theta,\varphi;\boldsymbol{Y})$ la vraisemblance associée. On notera $\widehat{\theta}$ et $\widehat{\varphi}$ les estimateurs du maximum de vraisemblance des paramètres θ et φ respectivement. On note $\widehat{Y}_i$ la prédiction faite par le modèle, i.e. $\widehat{Y}_i=h(\widehat{\theta}_i)$. Enfin, on appelera déviance la quantité

$$D=2(\log[\mathcal{L}(g(\boldsymbol{Y}),\widehat{\varphi};\boldsymbol{Y})]-\log[\mathcal{L}(g(\widehat{\boldsymbol{Y}}),\widehat{\varphi};\boldsymbol{Y})]),$$

où log désigne le logarithme népérien.

Exercice 1 [12 pts]

On consière la loi exponentielle définie sur $\mathbb{R}$, par $b(\theta) = \theta^2/2$ et $c(y, \varphi) = -(y^2/\varphi + \log[2\pi\varphi])/2$, où $\theta \in \mathbb{R}$ et $\varphi > 0$.

1. Soit Y une variable aléatoire suivant une telle loi. Calculer $\mathbb{E}(Y)$ et $\mathrm{var}(Y)$.
2. On dispose d'un échantillon (Y_i, X_i) pour $i = 1, \ldots, n$, et on suppose que Y_i a pour loi $f(y; \theta_i, \varphi)$. En supposant les observations indépendantes, écrire la log-vraisemblance de (θ, φ).
3. On suppose maintenant que $\theta_i = \beta_0 + \beta_1 X_i$. Ecrire la vraisemblance du triplet $(\beta_0, \beta_1, \varphi)$. Montrez que les estimateurs du maximum de vraisemblance pour les deux premiers paramètres sont
$$\widehat{\beta}_0 = \frac{1}{n}\sum_{i=1}^{n} Y_i \text{ et } \widehat{\beta}_1 = \frac{\sum_{i=1}^{n} X_i Y_i}{\sum_{i=1}^{n} X_i^2}$$
4. Ecrire la déviance D du modèle.
5. Donner la forme des résidus de déviance et de résidus de Pearson.

Exercice 2 [6 pts]

On considère la loi de la famille exponentielle définie pour $y \in \mathbb{N}$ par $\varphi = 1$, $b(\theta) = \exp(\theta)$, et $c(y, \varphi) = -\log(y!)$. On notera $y \mapsto f(y; \theta)$ la loi associée.

1. Calculer $\mathbb{E}(Y)$ où Y a pour loi $f(y; \theta)$
2. Calculer $\text{var}(Y)$ où Y a pour loi $f(y; \theta)$

Exercice 3

Partie 1 [14 pts]

Soit Λ une loi Gamma de paramètres μ et ν, au sens où sa densité s'écrit

$$g(\lambda;\mu,\nu) = \frac{1}{\lambda\Gamma(\nu)}\left(\frac{\lambda\nu}{\mu}\right)^{\nu}\exp\left[-\frac{\lambda\nu}{\mu}\right] \text{ pour } \lambda \geq 0,$$

où la fonction $\Gamma(\cdot)$ est définie par

$$\Gamma(z) = \int_0^{+\infty} t^{z-1}\, e^{-t}dt, \text{ pour } z \in \mathbb{R}_+ \text{ et } \Gamma(z+1) = z! \text{ pour } z \in \mathbb{N}.$$

On considère une variable aléatoire N telle que conditionnellement à $\Lambda = \lambda$, N suive une loi de Poisson de paramètre λ.

1. Calculer $\int_0^{+\infty} t^{z-1}\, e^{-\alpha t}dt$ en fonction de $\Gamma(z)$, α et z.
2. Calculer $\mathbb{E}(\Lambda)$. On admettra que $\text{Var}(\Lambda) = \mu^2/\nu$
3. Calculer $\mathbb{E}(N)$ et $\text{Var}(N)$.
4. Donner la loi de N. Montrer que cette loi appartient à la famille exponentielle (on ne spécifiera que l'expression du paramètre canonique θ en fonction de μ et ν ; en particulier il n'est pas nécessaire de spécifier la forme des fonctions b et c ou du paramètre φ).
5. On supposera par la suite que ν est fixé et connu. La loi de N dépend alors simplement du paramètre μ. Donner la fonction lien et son inverse.
6. Montrer que pour cette famille, la fonction variance $V(\cdot)$ est de la forme $V(\mu) = \mu + \kappa\mu^2$. Donner la valeur de κ.

Partie 2 [8 pts]

On considère une regression log-Poisson, i.e. on suppose que $Y_i|X_i$ suit une loi de Poisson de paramètre λ_i où $\lambda_i = \exp[\beta_0 + \beta_1 X_i]$, où X_i est à valeurs dans $\mathbb{R}$.

1. Calculer $\mathbb{E}(Y_i|X_i)$ et $\text{Var}(Y_i|X_i)$.
2. Montrez que $\mathbb{E}(Y|X = x + 1)$ est proportionnel à $\mathbb{E}(Y|X = x)$ pour tout $x \in \mathbb{R}$, et donner la valeur du coefficient de proportionnalité.

On suppose à partir de maintenant que le *vrai* modèle inclus une seconde variable explicative, *malheureusement inobservée*, i.e. $Y_i|(X_i, U_i)$ suit une loi de Poisson de paramètre λ_i où $\lambda_i = \exp[\beta_0 + \beta_1 X_i + U_i]$, où X et U sont à valeurs dans $\mathbb{R}$.

3. Calculer la loi 'non-conditionnelle" de $Y_i|X_i$, en suppossant que $H_i = \exp[U_i]$ est une variable aléatoire distribuée suivant une loi Gamma d'espérance 1 et de variance γ.
4. Calculer $\mathbb{E}(Y_i|X_i)$ et $\text{Var}(Y_i|X_i)$. Comparer avec les valeurs obtenues dans la question 1 (de la partie 2 de cet exercice).