Contrôle continu 3

Calcul des probabilités ; L3 physique ; durée : 2h

9 janvier 2017

Préambule

— Il est nécessaire de justifier chacune des réponses et de soigner la présentation.

— Une loi Bêta (α, β) avec $\alpha > 0$ et $\beta > 0$ a pour densité

$$f(x) = \begin{cases} \frac{1}{B(\alpha,\beta)} x^{\alpha-1}(1-x)^{\beta-1} & \text{si } x \in [0,1] \\ 0 & \text{sinon.} \end{cases}$$

— Une loi Gamma (k, θ) avec $k > 0$ et $\theta > 0$ a pour densité

$$f(x) = \begin{cases} \frac{1}{\Gamma(k)\theta^k} x^{k-1} \mathrm{e}^{-x/\theta} & \text{si x>0} \\ 0 & \text{sinon.} \end{cases}$$

— $\Gamma(\alpha) = \int_0^{+\infty} t^{\alpha-1}\mathrm{e}^{-t} dt$ et $\forall z > 0,\ \Gamma(z+1) = z \times \Gamma(z)$.

— $B(\alpha, \beta) = \dfrac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)} = \int_0^1 x^{\alpha-1}(1-x)^{\beta-1} dx$.

— Si $X \sim \mathcal{N}(0,1)$, on a les quantiles suivants :

x	$\frac{11}{50}\sqrt{34} \simeq 1,3$	$\frac{2}{5}\sqrt{17} \simeq 1,6$	$\frac{\sqrt{17}}{2} \simeq 2,0$	$\frac{14}{25}\sqrt{17} \simeq 2,3$	$\sqrt{\frac{34}{5}} \simeq 2,6$
$P(X \leq q)$	$0,9$	$0,95$	$0,975$	$0,99$	$0,995$

Selon, le contexte, on utilisera soit les valeurs en écriture décimale, soit les valeurs contenant des racines carrées.

Exercice 1. Soit X_1, ..., X_n un échantillon i.i.d. de loi $\mathrm{Exp}(\lambda)$ avec $n = 100$ et $n^{-1}\sum_{i=1}^{100} X_i = 4,0$.

1. Quelle est la loi de $Y = \exp\{-X_1\}$?
2. Quel est l'estimateur du maximum de vraisemblance pour λ?
3. En déduire un estimateur $\hat{\theta}$ pour $\theta = \lambda^{-1}$ et calculer son biais.
4. Expliquer pourquoi on a $\sqrt{n}\left(\dfrac{\hat{\theta}-\theta}{\theta}\right) \xrightarrow{\mathcal{L}} \mathcal{N}(0,1)$ et $\sqrt{n}\left(\dfrac{\hat{\theta}-\theta}{\hat{\theta}}\right) \xrightarrow{\mathcal{L}} \mathcal{N}(0,1)$.
5. En déduire un intervalle de confiance asymptotique à 95% pour θ.

Exercice 2. Soit Y_1, ..., Y_n un échantillon i.i.d. de densité

$$f_Y(y) = \begin{cases} \frac{y}{\sigma^2} \exp\left\{-\frac{y^2}{2\sigma^2}\right\}, & \text{si } y \geq 0 \\ 0 & \text{sinon.} \end{cases}$$

1. Proposer un estimateur de σ par la méthode des moments, que l'on notera $\hat{\sigma}_{MM}$.
   Indication : $\forall a > 0, \quad \int_0^{+\infty} \exp\left\{-\dfrac{x^2}{a}\right\} dx = \dfrac{\sqrt{\pi a}}{2}$.

2. Quel est l'estimateur du maximum de vraisemblance de σ, que l'on notera $\hat{\sigma}_{MV}$?
3. Calculer le biais de $\hat{\sigma}^2_{MV}$, estimateur de σ^2.
4. Trouver la loi limite de $\hat{\sigma}^2_{MV}$.
5. Soit $X \sim \mathcal{N}\left(0, \sigma^2\right)$ et Y de densité f_Y. Montrer que la densité de $Z = X/Y$ est, $\forall z \in \mathbb{R}$, $f_Z(z) = \left(z^2+1\right)^{-3/2}/2$.
6. Calculer l'espérance et la variance de Z, si elles existent.

Exercice 3. Soit le modèle de régression linéraire multiple dans le cas homoscédastique :

$$Y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i,2} + \varepsilon_i, \qquad \forall i \in \{1, \dots, n\}.$$

Si on note

$$\mathbb{X} = \begin{pmatrix} 1 & x_{11} & x_{12} \\ \vdots & \vdots & \vdots \\ 1 & x_{n1} & x_{n2} \end{pmatrix}, \quad \mathbf{Y} = \begin{pmatrix} Y_1 \\ \vdots \\ Y_n \end{pmatrix}, \quad \boldsymbol{\varepsilon} = \begin{pmatrix} \varepsilon_1 \\ \vdots \\ \varepsilon_n \end{pmatrix} \quad \text{et} \quad \boldsymbol{\beta} = \begin{pmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \end{pmatrix},$$

on a

$$\mathbb{X}^\top\mathbb{X} = \begin{pmatrix} 20 & 0 & 0 \\ 0 & 10 & -5 \\ 0 & -5 & 5 \end{pmatrix}, \quad \left(\mathbb{X}^\top\mathbb{X}\right)^{-1} = \frac{1}{20}\begin{pmatrix} 1 & 0 & 0 \\ 0 & 4 & 4 \\ 0 & 4 & 8 \end{pmatrix}, \quad \mathbb{X}^\top\mathbf{Y} = \begin{pmatrix} 20 \\ 5 \\ -10 \end{pmatrix} \quad \text{et} \quad \mathbf{Y}^\top\mathbf{Y} = 70.$$

1. Rappeler les hypothèses du modèle et donner le nombre d'observations n.
2. Donner l'expression de $\hat{\boldsymbol{\beta}}$ et faire l'application numérique.
3. Écrire l'équation d'analyse de la variance et en calculer les termes.
4. Calculer le coefficient de détermination R^2.
5. Donner l'expression d'un estimateur sans biais de σ^2 que l'on notera $\hat{\sigma}^2$ et faire l'application numérique.
6. On sait que de manière asymptotique,

$$\sqrt{n} \times \frac{\hat{\beta}_0 - \beta_0}{\hat{\sigma}\sqrt{\left[(\mathbb{X}^\top\mathbb{X})^{-1}\right]_{11}}} \xrightarrow{\mathcal{L}} \mathcal{N}(0,1).$$

   En déduire un intervalle de confiance asymptotique à 90% pour β_0.
7. On observe $x_{n+1,1} = 1$ et $x_{n+1}, 2 = 1$. Calculer une prévision pour Y_{n+1}.
8. On sait que de manière asymptotique,

$$\sqrt{n} \times \frac{\hat{Y}_{n+1} - Y_{n+1}}{\hat{\sigma}\sqrt{\left(1 + \mathbf{x}_{n+1}^\top (\mathbf{X}^\top\mathbf{X})^{-1}\mathbf{x}_{n+1}\right)}} \xrightarrow{\mathcal{L}} \mathcal{N}(0,1)$$

   avec $\mathbf{x}_{n+1}^\top = \begin{pmatrix} 1 & x_{n+1,1} & x_{n+1,2} \end{pmatrix}^\top$. En déduire un intervalle de confiance asymptotique à 98% pour Y_{n+1}.