UNIVERSITÉ PAUL SABATIER **F.S.I./UPSSITECH**

EXAMEN D'ESTIMATION STOCHASTIQUE – 2ASRI

1° session – Mercredi 16 Décembre 2015 – Durée 1h15

Tous documents de Cours, TD, TP autorisés – Tablettes et objets communicants interdits

*les questions **I/**, **II/**, **III/**-8, **III/**-9, **III/**-11, **III/**-12 sont indépendantes.*

I/ Questions de cours. Répondre *en trois phrases maximum convenablement construites, sans nécessairement invoquer des formules mathématiques* à chacune des questions suivantes.

1. Indiquer en quoi diffèrent les cadres théoriques de l'estimation dite « classique » et de l'estimation Bayésienne.
2. Expliquer en langage simple ce que sont le biais et la covariance d'un estimateur.
3. À quoi sert l'inégalité de Cramér-Rao ?
4. Soient Θ et Z deux variables aléatoires. Quelle signification peut-on accorder à la loi *a priori* $p_\Theta(\theta)$ de Θ et à sa loi *a posteriori* $p_{\Theta|Z}(\theta|z)$?

II/ Soit X une variable aléatoire vectorielle, et $X_1, \ldots, X_N$ ses composantes.

5. À quelle condition $X_1, \ldots, X_N$ sont-elles mutuellement indépendantes ?
6. On suppose que $X \sim \mathcal{N}(\bar{x}, P)$, c.-à-d. que X suit la loi Gaussienne réelle multidimensionnelle de moyenne $\bar{x}$ et de covariance P. Répondre aux questions suivantes.
   (a) Comment s'écrit la densité de probabilité $p_X(x)$ de X ?
   (b) Si les variables aléatoires scalaires $X_1, \ldots, X_N$ sont indépendantes identiquement distribuées, comment $p_X(x)$ se réécrit-elle ?
   (c) À quelle condition sur $\bar{x}$ et/ou P les variables aléatoires scalaires $X_1, \ldots, X_N$ sont-elles (a) centrées ; (b) mutuellement indépendantes ; (c) corrélées sans être linéairement dépendantes ; (d) linéairement dépendantes ?

III/ Soit $\mathcal{N}(m, \sigma^2)$ la loi Gaussienne scalaire réelle de moyenne m et de variance $v = \sigma^2$. On suppose que m est une constante connue. Sur la base d'un vecteur z constitué de n échantillons $z_1, \ldots, z_N$ i.i.d. selon $\mathcal{N}(m, \sigma^2)$, on construit l'estimé

$$\hat{v} = g(z) = \frac{1}{N} \sum_{n=1}^{N} (z_n - m)^2 \tag{1}$$

de v. Cette question se propose d'analyser l'estimateur $\hat{V} = g(Z)$ associé, où $Z = (Z_1, \ldots, Z_N)^T$ désigne la variable vectorielle aléatoire qui s'est réalisée en $z = (z_1, \ldots, z_N)^T$.

7. Exprimer $\hat{V}$ en fonction de $Z_1, \ldots, Z_N$.
8. On souhaite montrer que $\hat{V}$ n'est pas biaisé.
   (a) Écrire quelle égalité mathématique doit être vérifiée pour justifier ce résultat.
   (b) Comme indiqué ci-dessus, $\forall n,\ Z_n \sim \mathcal{N}(m, v)$. Comment s'écrit alors l'espérance $\mathbb{E}\{(Z_n - m)^2\}$?
   (c) En déduire la propriété de non-biais de $\hat{V}$.
9. On souhaite exprimer la variance de $\hat{V}$. Une manière de procéder est de constater que $X = \frac{N}{\sigma^2}\hat{V}$ est équivalent à la somme des carrés de N variables Gaussiennes mutuellement indépendantes, de moyenne nulle et de variance unité.
   (a) Quelle est la distribution de probabilité (que l'on notera $X \sim \chi^2_N$) de X ?

(b) Exprimer le lien de proportionnalité qui unit la variance $\mathrm{Var}(\hat{V})$ de $\hat{V}$ à la variance $\mathrm{Var}(X)$ de X.

(c) Sachant que $\mathrm{Var}(X) = 2N$ (propriété de la loi χ^2_N), en déduire $\mathrm{Var}(\hat{V})$.

10. Déduire des propriétés démontrées ci-dessus le comportement de $\hat{V}$ lorsque le nombre d'échantillons N croît indéfiniment.

11. On se propose de démontrer en outre que $\hat{V}$ est l'estimateur du maximum de vraisemblance de v.

    (a) Écrire la loi $p_{Z|v}(z|v)$ et simplifier son expression en exploitant l'indépendance des composantes de Z.

    (b) On rappelle que $p_{Z|v}(z|v)$, considérée comme une fonction de v, exprime la vraisemblance $\mathrm{L}(v; z)$ de v étant donné un vecteur d'observation z. Écrire l'anti log-vraisemblance $\mathrm{NLL}(v; z) = -\ln \mathrm{L}(v; z)$. Rappeler quels problèmes d'optimisation permettent l'obtention de l'estimé du maximum de vraisemblance $\hat{v}_{\mathrm{MLE}}$ de v à partir de $p_{Z|v}(z|v)$, puis de $\mathrm{NLL}(v; z)$.

    (c) Écrire la condition nécessaire de stationnarité sur $\mathrm{NLL}(v; z)$ que doit satisfaire $\hat{v}_{\mathrm{MLE}}$, et la développer. En déduire l'égalité de $\hat{V}_{\mathrm{MLE}}$ et de $\hat{V}$ introduit en (1).

12. $\hat{V}$ est-il l'estimateur efficace ? Justifier soigneusement le résultat.