Master math-info 2012-2013
Travaux dirigés de Statistique 3
Vecteurs gaussiens, intervalles de confiance

Exercice 1 *Soient X une variable aléatoire réelle de loi $\mathcal{N}(0,1)$ et ε une variable aléatoire réelle de loi $\mathbb{P}_\varepsilon\{-1\} = \mathbb{P}_\varepsilon\{1\} = \frac{1}{2}$. On suppose que X et ε sont indépendants.*

1. *Montrer que la variable $Y = \varepsilon X$ est gaussienne et que les variables X, Y sont décorrélées.*

2. *Le couple (X, Y) est-il indépendant? Gaussien?*

Exercice 2 *Soit X un vecteur aléatoire de dimension 3. On suppose que la loi de X est $\mathcal{N}(0, \Gamma)$, où Γ est la matrice définie par*

$$\begin{pmatrix} 3 & -1 & 0 \\ -1 & 3 & 0 \\ 0 & 0 & 2 \end{pmatrix}$$

Peut-on trouver une matrice carrée A d'ordre 3 telle que les composantes du vecteur AX soient indépendantes et non dégénérées? On veut maintenant que AX suit la loi $\mathcal{N}(0, I)$, où I est la matrice identité.

Exercice 3 *Soit θ un paramètre strictement positif, et X une variable aléatoire suivant la loi normale $\mathcal{N}(\theta, \theta)$ dont on possède un $n-$échantillon $(X_1, \cdots, X_n)$.*

1. *Compte tenu de la double interprétation du paramètre θ, déterminer deux estimateurs sans biais de θ par la méthode des moments appliquée à $E_\theta(X)$ et $Var_\theta(X)$. Comparer au sens du risque quadratique ces deux estimateurs.*

2. *On considère la classe des estimateurs $T(\lambda)$ de la forme :*

$$T(\lambda) = \lambda \overline{X}_n + (1-\lambda) S_n'^2, 0 \leq \lambda \leq 1.$$

Montrer que, $\forall \lambda \in [0; 1]$, $T(\lambda)$ est un estimateur sans biais et fortement consistant de θ.

3. *Calculer le risque quadratique $R_\lambda(\theta)$ de l'estimateur $T(\lambda)$. Vérifier qu'il existe un unique réel λ^* appartenant à $[0; 1]$ tel que*

$$R_{\lambda^*}(\theta) \leq R_\lambda(\theta), \ \forall \theta > 0, \ \forall \lambda \in [0; 1].$$

4. *Etudier les convergences en loi de $\sqrt{n}(\overline{X}_n - \theta, S_n'^2 - \theta)$ quand $n \to +\infty$.*

5. *En déduire la convergence en loi de $\sqrt{n}(T(\lambda) - \theta)$ quand n tend vers $+\infty$.*

Exercice 4 *Soit $(X_1, \cdots, X_n)$ un n-échantillon de loi $\mathcal{N}(\mu, \sigma^2)$. On suppose que $n = 20$, et $\sigma^2 = 9$. Sur une réalisation $(x_1, \cdots, x_{20})$, on a calculé que $\overline{x}_{20} = 2.09$.*

1. *Déterminer un intervalle de confiance de niveau de confiance 90% pour μ.*

2. *Quelle devrait être la taille de l'échantillon pour que l'intervalle de confiance soit moitié moins long?*

3. *On suppose que la variance n'est pas connue. Sachant que* $\sum_{i=1}^{20}(x_i - \overline{x}_{20})^2 = 140.6$, *donner un intervalle de confiance de niveau de confiance 90% pour μ.*

4. *On suppose ici que μ est connue et vaut* 2. *Donner un intervalle de confiance de niveau de confiance 90% pour σ^2.*

5. *Même question si l'on ne suppose plus μ connue.*

Exercice 5 *Dans l'atmosphère, le taux d'un gaz nocif, pour un volume donné, suit une loi normale d'espérance μ et de variance σ^2 inconnues. On effectue n prélèvements conduisant aux valeurs* $x_1, x_2, \cdots, x_n$.

1. *Sur un échantillon de taille* $n = 10$, *on observe que* $\overline{x}_{10} = 50$ *et* ${s'}^2_{10} = 100$ *où* ${s'}^2_{10} = \sum_{i=1}^{10}(x_i - \overline{x}_{10})^2/9$. *Quel est l'intervalle de confiance de niveau de confiance 95% du taux moyen μ de gaz dans l'atmosphère ?*

2. *Quel serait cet intervalle si la variance σ^2 était connue (on la prendra égale à 100)?*

3. *On dispose maintenant d'un échantillon 10 fois plus grand conduisant aux résultats suivants:* $\overline{x}_{100} = 48$ *et* ${s'}^2_{100} = 90$. *Quel est alors l'intervalle de confiance de niveau de confiance 90% pour μ?*

Exercice 6 *Soit $(X_1, \cdots, X_n)$ un n-échantillon de loi de Poisson Π_λ.*

1. *Construire un intervalle de confiance de niveau de confiance* 90% *pour le paramètre λ lorsque* $n = 30$ *et* $\sum_{i=1}^{30} X_i = 240$ *en utilisant une approximation gaussienne de $\overline{X}$ et en résolvant une équation du second degré.*

2. *Comparer avec les résultats obtenus lorsque l'on remplace la variance par un estimateur sans biais.*

Exercice 7 *Soit $(X_1, \cdots, X_n)$ un n-échantillon de loi de Bernoulli $\mathcal{B}_p$.*

1. *Construire un intervalle de confiance de niveau de confiance* 95% *pour le paramètre p lorsque* $n = 50$ *et* $\sum_{i=1}^{50} X_i = 15$ *en utilisant une approximation gaussienne de $\overline{X}$ et en résolvant une équation du second degré.*

2. *Comparer avec les résultats obtenus lorsque l'on remplace la variance par une estimation.*