Fiche 4 - Statistique - M1 Ingé Math

Tran Viet Chi, `chi.tran@math.univ-lille1.fr`, bureau 316 (bâtiment M3).

Exercice 1 (EMV pour une loi normale)
Soit $X_1, \ldots, X_n$ des v.a. i.i.d. de loi $\mathcal{N}(\mu, \sigma^2)$. On pose $\theta = (\mu, \sigma^2)$.

1. Ecrire le modèle. Est-ce un modèle exponentiel ?

2. Donner la matrice d'information de Fisher de θ.

3. Calculer l'EMV $\widehat{\theta}_n$ de θ. Quelle est sa loi asymptotique ?

4. Aurait-on pu donner très simplement la matrice d'information de Fisher sans avoir à effectuer les calculs de la question 2 ?

5. Donner la loi asymptotique marginale de l'EMV de μ. Cette loi est-elle réellement asymptotique ? Comment appelle-t-on ce type d'estimateur ?

6. Donner la loi asymptotique marginale de l'EMV de σ^2. Cette loi est-elle réellement asymptotique ? Comment appelle-t-on ce type d'estimateur ?

7. Est-ce que $\widehat{\theta}_n$ est un estimateur efficace de θ ?

8. Donner l'EMV de l'écart-type σ ainsi que sa loi asymptotique.

9. Si la constante b est définie par l'équation $\mathbb{P}(X \leq b) = 0.9$, donner l'EMV de b ainsi que sa loi asymptotique.

Exercice 2 (EMV pour une loi uniforme)
Soit $X_1, \ldots, X_n$ des v.a. i.i.d. de loi $\mathcal{U}[0, \theta]$ et on se propose d'estimer le paramètre inconnu θ. Pour cela, on se propose d'étudier les deux estimateurs $\widehat{\theta}_1 = 2\bar{X}_n$ est $\widehat{\theta}_2 = \max\{X_i, i \in \{1, \ldots, n\}\}$.

1. Ecrire le modèle statistique. Est-ce un modèle exponentiel ?

2. Calculer l'espérance et la variance de $\widehat{\theta}_1$.

3. Même question pour $\widehat{\theta}_2$. En déduire un estimateur sans biais $\widehat{\theta}_2^*$ de θ.

4. Quelle est la loi asymptotique de $\widehat{\theta}_2$?

5. Comparer les deux estimateurs $\widehat{\theta}_1$ et $\widehat{\theta}_2^*$. Lequel des deux semble être le meilleur ? Cet estimateur est-il alors le meilleur possible ?

6. Existe-il un estimateur efficace de θ ?

Exercice 3 (Modèle de durée du chômage)
On étudie entre les dates 0 et T un groupe de n individus sans emploi à la date 0, et on cherche à modéliser les durées de chômage $(T_i)_{i \in \{1,\ldots,n\}}$. En pratique, on observe les durées de chômage en mois : on ne dispose pas de la variable T_i, mais seulement de la variable discrète T_i^* donnée par :

$$T_i^* = [T_i] + 1$$

où [.] désigne la partie entière.

Entre les dates $t \in \mathbb{N}$ et $t + 1$, on suppose que :

- l'individu i reçoit N_t^i offres d'emploi, où les N_t^i sont des v.a. i.i.d. de loi de Poisson $\mathcal{P}(\lambda)$.
- si l'individu i est toujours au chômage à la date t, et si parmi les N_t^i offres reçues, l'une au moins offre un salaire supérieure à une constante $\xi_i > 0$ propre à l'individu (salaire de réserve de l'individu i), alors l'individu n'est plus au chômage.
- les salaires des offres d'emploi sont tirés indépendamment des dates d'arrivée des offres et de leur nombre dans une loi de fonction de répartition F.

Dans un premier temps, on suppose que $T = +\infty$.

1. Que vaut T_i^* si l'individu i trouve du travail au bout de deux mois et demi ?

2. Ecrire le modèle statistique associé à l'observation des T_i^*.

3. Calculer $\mathbb{P}(T_i^* = t+1 | T_i > t)$ en fonction de F, ξ_i et λ.

4. En déduire la vraisemblance de $(T_1^*, \ldots, T_2^*)$.

Jusqu'à la fin, nous supposerons maintenant que tous les individus ont le même salaire de réserve $\xi_i = \xi$, et que ce salaire de réserve est connu, ainsi que F.

5. Trouver l'estimateur du maximum de vraisemblance de λ.

6. Quand $n \to +\infty$, cet estimateur est-il convergent ? asymptotiquement normal ?

En pratique, l'enquête se termine à la fin du T-ième mois, $T < +\infty$. A cette date, certains individus peuvent encore être au chômage. On n'observe donc que :

$$\begin{aligned} T_i^{**} =& T_i^* \qquad \text{si } T_i^* \leq T \\ =& T+1 \qquad \text{si } T_i^* > T. \end{aligned}$$

7. Donner le nouveau modèle statistique.

8. Ecrire la vraisemblance des observations $(T_1^{**}, \ldots, T_n^{**})$.

9. Donner l'estimateur du maximum de vraisemblance de λ.

10. On suppose maintenant que

$$F(x) = \left(1 - e^{-\gamma(x-\xi_0)}\right) \mathbf{1}_{x \geq \xi_0},$$

où γ est un paramètre inconnu à estimer et ξ_0 est connu. Le couple (λ, γ) est-il indentifiable ?