T. D. n° 6
Information de Fisher et maximum de vraisemblance

Exercice 1. Information, efficacité et loi de Gauss.

Soit X une variable aléatoire suivant une loi normale $\mathcal{N}(\mu, \sigma)$. Soit $(X_1, \ldots, X_n)$ un échantillon aléatoire de taille n de loi parente X.

1. Calculer l'information de Fisher $I(\mu)$ pour le paramètre μ.
2. Calculer l'information de Fisher $I(\sigma^2)$ pour le paramètre σ^2.
3. La statistique $\overline{X}$ est-elle efficace pour μ ?

Exercice 2. Loi de Bernoulli et efficacité.

Soit X une variable aléatoire suivant une loi de Bernoulli $\mathcal{B}(1, p)$ de paramètre $0 < p < 1$. Nous disposons de $(X_1, \ldots, X_n)$, un échantillon aléatoire de taille n de loi parente X.

1. Déterminer un estimateur T sans biais pour p. Calculer $\mathrm{Var}\,[T]$ et conclure.
2. Étudier l'efficacité de l'estimateur T.

Exercice 3. Loi de Poisson, information et efficacité.

Soit X une variable aléatoire suivant une loi de Poisson de paramètre λ. Nous disposons de $(X_1, \ldots, X_n)$, un échantillon aléatoire de taille n de loi parente X.

1. Déterminer l'information de Fisher pour λ fournie par $(X_1, \ldots, X_n)$.
2. Proposer un estimateur efficace de λ.

Exercice 4. Loi gamma $\gamma(p, 1)$ et efficacité.

Soit X une variable aléatoire suivant une loi gamma $\gamma(p, 1)$. Nous disposons de $(X_1, \ldots, X_n)$, un échantillon aléatoire de taille n de loi parente X.
Montrer qu'il n'existe qu'une seule fonction du paramètre p qui puisse être estimée efficacement puis calculer-la.

Exercice 5. Loi gamma $\gamma(p, r)$ et information de Fisher.

Soit X une variable aléatoire suivant une loi gamma $\gamma(p, r)$. Nous disposons de $(X_1, \ldots, X_n)$, un échantillon aléatoire de taille n de loi parente X.
Déterminer l'information de Fisher pour r fournie par $(X_1, \ldots, X_n)$.

Exercice 6. Efficacité et loi normale.

Soit X une variable aléatoire suivant une loi normale $\gamma(\mu, \sigma)$. Nous disposons de $(X_1, \ldots, X_n)$, un échantillon aléatoire de taille n de loi parente X.
Les statistiques $\overline{X} = \widehat{\mu}_n$ et $S^2_{n,c}$ sont-elles efficaces pour m et σ^2 ?

Exercice 7. Loi uniforme et efficacité.

Soit X une variable aléatoire suivant une loi une loi uniforme $U_[0, \theta]$. Nous disposons de $(X_1, \ldots, X_n)$, un échantillon aléatoire de taille n de loi parente X.
Soit U et T les statistiques :

$$U(X_1, \ldots, X_n) = 2\overline{X} = \frac{2}{n}\sum_{i=1}^{n} X_i \quad \text{et} \quad T(X_1, \ldots, X_n) = \sup_{1 \geqslant i \geqslant n} X_i.$$

1. Montrer que U et T sont deux estimateurs de θ.
2. Comparer les deux estimateurs.

Exercice 8. Loi de Poisson et efficacité.

Soit X une variable aléatoire suivant une loi de Poisson de paramètre λ. Nous disposons de $(X_1, \ldots, X_n)$, un échantillon aléatoire de taille n de loi parente X.
Montrer que l'estimateur $\overline{X}$ est efficace pour λ en utilisant le fait que la loi de Poisson appartient à la famille exponentielle.

Exercice 9. Maximum de vraisemblance et loi de Poisson.

Soit X une variable aléatoire suivant une loi de Poisson de paramètre λ. Nous disposons de $(X_1, \ldots, X_n)$, un échantillon aléatoire de taille n de loi parente X.
Déterminer l'estimateur du Maximum de vraisemblance $\widehat{\theta}_{MV}$ de θ.

Exercice 10. Maximum de vraisemblance et loi de Bernoulli.

Soit X une variable aléatoire suivant une loi de Bernoulli de paramètre p. Nous disposons de $(X_1, \ldots, X_n)$, un échantillon aléatoire de taille n de loi parente X.
Déterminer l'estimateur du Maximum de vraisemblance $\widehat{p}_{MV}$ de p.

Exercice 11. Maximum de vraisemblance et loi normale.

Soit X une variable aléatoire suivant une loi normale $\mathcal{N}(\mu, \sigma)$. Nous disposons de $(X_1, \ldots, X_n)$, un échantillon aléatoire de taille n de loi parente X.

1. Supposons σ connu et μ inconnue. Déterminer l'estimateur du Maximum de vraisemblance $\widehat{\mu}_{MV}$ de μ.

2. Supposons σ inconnu et μ connue. Déterminer l'estimateur du Maximum de vraisemblance $\widehat{\sigma^2}_{MV}$ de σ^2.
3. Supposons σ inconnu et μ connue. Déterminer l'estimateur du Maximum de vraisemblance $\widehat{\sigma}_{MV}$ de σ. Comparer avec le résultat de la question 3..
4. Supposons σ et μ inconnus. Déterminer les estimateurs du Maximum de vraisemblance $\widehat{\mu}_{MV}$ de μ et $\widehat{\sigma}_{MV}$ de σ.

Exercice 12. Efficacité.

Soit X une variable aléatoire de densité $f(x,\theta) = \frac{A}{x^{1+\frac{1}{\theta}}}$ pour $x \geqslant 1$ et $f(x,\theta) = 0$ sinon, avec $\theta > 0$.
Nous disposons de $(X_1, \ldots, X_n)$ un échantillon aléatoire de taille n de loi parente X.

1. Déterminer A.

2. Déterminer l'estimateur du Maximum de vraisemblance $\widehat{\theta}_{MV}$ de θ.

3. Calculer, après avoir justifié pourquoi ce calcul est possible, la borne de Fréchet. L'estimateur $\widehat{\theta}_{MV}$ de θ est-il efficace ?

Exercice 13. Contrôle de qualité.

Une chaîne de production dans une usine fabrique des pièces métalliques dont une proportion θ inconnue est constituée de pièces défectueuses.
Pour estimer cette proportion il y a deux possibilités.

1. Un échantillon de pièces est tiré dans la production, de façon indépendante, puis en observant le nombre de pièces défectueuses dans l'échantillon, on élabore une estimation de la proportion de pièces défectueuses dans la production totale.
2. On tire un échantillon de pièces dans la production, de façon indépendante, jusqu'à obtenir r pièces défectueuses dans l'échantillon.
   À l'aide de la taille de l'échantillon nécessaire pour obtenir r pièces défectueuses, on élabore une autre estimation de la proportion de pièces défectueuses dans la production totale.

1. Déterminer, pour les deux schémas de tirage, la vraisemblance de l'échantillon, ainsi que les estimateurs du maximum de vraisemblance pour le paramètre θ.
2. Que dire des deux estimateurs obtenus ?

Exercice 14. Estimation de paramètres délicats.

Nous souhaitons connaître la proportion p_1 de français fraudant le fisc.
Une méthode naturelle serait d'effectuer un tirage sans remise dans l'ensemble des français soumis à l'impôt et de poser à chaque individu de l'échantillon la question « Fraudez-vous le fisc ? ».

Bien entendu il est fort probable que les individus interrogés ne répondent pas franchement à cette question. C'est pourquoi le protocole alternatif suivant est suggéré. La personne interrogée par un enquêteur tire au hasard une boule dans une urne contenant une boule blanche et une boule noire. Sans révéler la couleur de la boule tirée la personne donne une réponse de la façon suivante :
Si la boule est noire, la personne interrogée réponde à la question « Fraudez-vous le fisc ? ».
Si la boule est blanche, la personne interrogée réponde à la question « Aimez-vous Johnny Halliday ? ».
Le secret de son comportement vis-à-vis du fisc est donc conservé puisque l'enquêteur est dans l'incapacité de savoir à laquelle des deux questions la personne interrogée a répondu.

1. Soit p_2 la proportion de la population soumise à l'impôt appréciant Johnny Halliday et Y_i la variable valant 1 si la réponse de la personne est « Oui » et 0 sinon.
   a. Déterminez la loi de Y_i puis calculer son espérance et sa variance.
   b. En déduire un estimateur efficace $\widehat{p}$ de $p_1 + p_2$. Calculer alors la variance de cet estimateur.

2. Nous effectuons, de façon indépendante, une enquête exhaustive (tirage sans remise), pour estimer p_2.
Déterminer un estimateur $\widehat{p_2}$ pour estimer p_2.

3. En déduire alors un estimateur sans biais $\widehat{p_1}$ de p_1. Déterminer la variance de cette estimateur.

4. Existe-t-il une valeur de p_2 qui minimise la variance de l'estimateur trouvé.

Exercice 15. Expérience tronquée en fiabilité.

La durée de vie moyenne d'un phénomène physique est une variable aléatoire X de densité :

$$f(x) = \frac{1}{\lambda} \exp\left(-\frac{x-\alpha}{\lambda}\right), \quad \text{pour} \quad x \geqslant \alpha \quad \text{et 0 sinon.}$$

Pour tester la durée de vie de ce système physique, il est décidé tout d'abord d'observer, de manière indépendante, les durées de vie $X_1, \ldots, X_n$ de n systèmes physiques. Néanmoins, pour des raisons d'économies, les observations s'arrêtent lorsque le $r-$ème tombe en panne. Nous observons donc en fait $Y_1, \ldots, Y_r$.

1. Déterminer la fonction de répartition $F(x)$ de la variable X. Calculer l'espérance de X.

2. Montrer que la loi de $(Y_1, \ldots, Y_r)$ a pour densité :

$$f(y_1, \ldots, y_r) = \frac{n!}{(n-r)!} \frac{1}{\lambda^r} \exp\left(\frac{-1}{\lambda}\left[\sum_{i=1}^{r}(y_i - \alpha) + (n-r)(y_r - \alpha)\right]\right)$$

pour $\alpha < y_1 < \ldots < y_r$.

3. Nous posons Z la statistique définie par :

$$Z = \sum_{i=1}^{r} (Y_i - Y_1) + (n-r)(Y_r - Y_1).$$

Montrer que (Y_1, Z) est une statistique exhaustive. Quelle interprétation pouvez-vous donner à cette statistique ?