

TP4 Statistiques noté : Estimation et l'intervalle de confiance

Nom1Prenom1-Nom2Prenom2

11 avril 2025

Estimateur du maximum de vraisemblance

Soit X_1, \dots, X_n un échantillon de normale $X \sim \mathcal{N}(\theta, \theta^2)$ où θ est un réel strictement positif. Nous voulons estimer θ à partir d'une réalisation de cet échantillon $x = (x_1, \dots, x_n)$ de X .

Ex1 Donner une expression explicite de l'estimateur du maximum de vraisemblance pour θ .

$$\hat{\theta}_n^{MV} = \hat{\theta}_n^{MV}(X_1, \dots, X_n) = ?$$

Ex2. Simuler un échantillon i.i.d de X de taille $n = 50$ avec $\theta = 1$ en utilisant `rnorm` et donner une estimation de θ obtenue par les calculs dans l'**Ex1**. Que remarquez-vous ?

```
n <-  
theta <-  
x <-  
  
theta_MV <-  
thetha_MV
```

Ex3. Créer une fonction de vraisemblance ou log de la vraisemblance, nommée `L_norm`, en fonction de (θ, x) , qui donne la vraisemblance d'un échantillon $x = (x_1, \dots, x_n)$ pour une valeur donnée de θ .

```
L_norm <- function(theta, x){  
  ...  
  return(...)  
}
```

Ex4.. Pour l'échantillon généré dans l'**Ex2**, calculer la vraisemblance de cet échantillon des normales de paramètre θ (le range des valeurs de θ à votre choix). Tracer la courbe des valeurs calculées en fonction de θ . Que remarquez-vous?

```
thetavec = seq(...,...,by=...)  
L = sapply(thetavec, L_norm, x=...)  
L  
plot(...)
```

Ex5.. En utilisant la fonction `optim` de R, trouvez la valeur de θ la plus probable d'avoir généré cet échantillon.

```
mL_norm <- function(theta,x){ -L_norm(theta,x) }  
## optimization standard (minimization)  
p0 = ... # valeur initiale pour l'algorithme  
res = optim(p0, mL_norm, x=x, method = "...",...)  
res
```

Ex6. Faire varier les échantillons de taille n allant de $n = 10$ à $n = 2000$ et comparer l'écart entre la valeur théorique attendue, l'estimation obtenue dans l'**Ex1** et la valeur obtenue pqr **optim**. Que remarquez-vous? Sont-ils consistents ?

Information de Fisher et l'Intervalle de confiance

On se rappelle que l'Information de Fisher associée au modèle au point θ est

$$I(\theta) = \mathbb{E}_{\theta} \left[(\nabla_{\theta} \mathcal{L}(X; \theta))^2 \right] = -\mathbb{E}_{\theta} \left[\nabla_{\theta}^2 \mathcal{L}(X; \theta) \right].$$

Ex7. Montrer théoriquement que $I(\theta) = \frac{3}{\theta^2}$.

Ex8. Faire un histogramme de $\sqrt{nI(\theta)}(\hat{\theta}_n^{MV} - \theta)$ et superposer la densité théorique de la loi normale standard sur l'histogramme. Expliquer les plots obtenus.

Ex9 Donner les intervalles de confiance de niveau 0.90 pour le paramètre θ .

Ex10. Pour n donné, simuler 500 échantillons et obtenir des intervalles de confiance. Compter le nombre de fois où l'intervalle contient le vrai paramètre. Quelle est votre conclusion?