Maximum de Vraisemblance (MLE)

Synthèse des Informations

1. Définitions (Principe et Formules)

- **Principe :** Estimer les paramètres d'un modèle probabiliste en cherchant les valeurs qui rendent l'échantillon observé le plus probable.
- **Vraisemblance (L) :** Probabilité d'observer conjointement l'échantillon (pour des variables indépendantes et identiquement distribuées).

$$L(\theta) = \prod_{i=1}^{n} f(x_i|\theta)$$

- **Log-vraisemblance (ℓ) :** Le logarithme transforme le produit en somme (simplifie les dérivées et évite les erreurs numériques).

$$\ell(\theta) = \ln(L(\theta)) = \sum_{i=1}^{n} \ln(f(x_i|\theta))$$

- **Estimateur ($\hat{\theta}$) :** Valeur qui maximise $\ell(\theta)$, trouvée en annulant sa dérivée.

$$\frac{\partial \ell(\theta)}{\partial \theta} = 0$$

2. Confiance dans l'estimation

La confiance repose sur ses fortes **propriétés asymptotiques** (quand la taille de l'échantillon n est grande) :

- **Convergence (Consistance) :** Plus on a de données, plus l'estimation se rapproche de la vraie valeur de façon certaine.
- **Efficacité :** L'estimateur utilise l'information au maximum. Il a la variance la plus faible possible (atteint la borne de Cramér-Rao).
- **Normalité asymptotique :** L'estimateur tend à suivre une loi Normale, ce qui facilite grandement le calcul d'intervalles de confiance.

3. Exemple : Loi Exponentielle $\mathcal{E}(\lambda)$

On cherche le taux de défaillance λ à partir de n durées de vie observées ($x_1, \ldots, x_n$).

- **1. Densité :** $f(x|\lambda) = \lambda e^{-\lambda x}$
  Signification : Elle modélise la probabilité d'observer une défaillance à un instant x précis. C'est la brique de base de notre modèle de défaillance.
- **2. Vraisemblance :** $L(\lambda) = \prod_{i=1}^{n} \left(\lambda e^{-\lambda x_i}\right) = \lambda^n e^{-\lambda \sum x_i}$
  *Signification : Elle représente la probabilité globale d'avoir observé **exactement** notre échantillon de n pannes. L'objectif est de trouver le λ qui rend ce scénario le plus crédible possible.*

- **3. Log-vraisemblance :** $\ell(\lambda) = n \ln(\lambda) - \lambda \sum x_i$
  Signification : C'est une astuce mathématique. Le logarithme transforme un produit complexe en une somme facile à dériver, tout en conservant la même position pour le maximum.

- **4. Maximisation ($\partial\ell/\partial\lambda = 0$) :** $\frac{n}{\lambda} - \sum x_i = 0$
  Signification : Pour trouver le « sommet » de la vraisemblance, on calcule la dérivée par rapport à λ et on cherche l'endroit où la pente est parfaitement plate (nulle).

- **5. Résultat :** $\hat{\lambda} = \frac{n}{\sum x_i} = \frac{1}{\bar{x}}$
  Signification : Le taux de défaillance estimé ($\hat{\lambda}$) est tout simplement l'inverse de la moyenne des durées de vie observées ($\bar{x}$). C'est un résultat très intuitif et directement applicable !