Feuille n° 4
Statistiques Inférentielles (corrigé [1])

Rappels : méthode de maximum de vraisemblance

déf : fonction de vraisemblance (loi discrète, loi à densité), log-vraisemblance
déf : estimation du maximum de vraisemblance

Exercice 1

1. On place n boules aléatoirement (indépendamment et uniformément) sur l'intervalle $[0, 1]$. Quelle est la probabilité d'avoir k boules dans l'intervalle $[0, 0.1]$?
2. Même question pour $10n$ boules dans l'intervalle $[0, 0.01]$?
3. Loi des événements rares (forme faible). Soient n et p tels que le produit $np = \lambda > 0$ est fixé. Montrer que pour $n \to \infty$ et $p \to 0$ on a :
$$\frac{n!}{k!(n-k)!} p^k (1-p)^{n-k} \to e^{-\lambda} \frac{\lambda^k}{k!}$$
4. Soit $X \sim \text{Poisson}(\lambda)$, $\lambda > 0$. Calculer $\mathbb{E}[X]$. En déduire un estimateur de λ.
5. Soient $X_1, ..., X_n$ i.i.d $\sim \text{Poisson}(\lambda)$. Montrer que la vraisemblance est donnée par :
$$L(\lambda; x_1, ...x_n) = \frac{e^{-n\lambda}}{x_1! \cdots x_n!} \lambda^{x_1 + \cdots + x_n}$$
6. Calculer la log-vraisemblance et en déduire l'estimateur du maximum de vraisemblance (EMV) de λ.
7. Quel est l'EMV de la probabilité que $X_1 = \cdots = X_n = 0$.

Corrigé Exercice 1

1. En plaçant 1 boule uniformément sur l'intervalle $[0, 1]$, on a probabilité 0.1 de l'avoir dans l'intervalle $[0, 0.1]$ (et probabilité 0.9 de l'avoir dans l'intervalle $[0.1, 1]$) ; en plaçant n boules aléatoirement (indépendamment et uniformément), la probabilité d'avoir exactement k boules dans l'intervalle $[0, 0.1]$ sera
$$\binom{n}{k} 0.1^k 0.9^{n-k}$$
2. (Dans le même esprit que la qustion précedente) En plaçant $10n$ boules aléatoirement (indépendamment et uniformément), la probabilité d'avoir exactement k boules dans l'intervalle $[0, 0.01]$ sera
$$\binom{10n}{k} 0.01^k 0.99^{10n-k}$$
3. Comme $np = \lambda$, on écrit $p = \frac{\lambda}{n}$; et on remplace cela dans l'expression, ce qui donne
$$\begin{aligned}\frac{n!}{k!(n-k)!} p^k (1-p)^{n-k} &= \frac{n!}{k!(n-k)!} \left(\frac{\lambda}{n}\right)^k \left(1 - \frac{\lambda}{n}\right)^{n-k} \\ &= \frac{\lambda^k}{k!} \frac{n(n-1)\cdots(n-k+1)}{n^k} \left(1 - \frac{\lambda}{n}\right)^n \left(1 - \frac{\lambda}{n}\right)^{-k}\end{aligned}$$

1. Yiyang Yu, yyu@lpsm.paris

Quand λ fixé, $n \to +\infty$,

$$\frac{n(n-1)\cdots(n-k+1)}{n^k} \to 1, \quad \left(1-\frac{\lambda}{n}\right)^n \to e^{-\lambda}, \quad \left(1-\frac{\lambda}{n}\right)^{-k} \to 1$$

donc pour λ fixé, $n \to +\infty$,

$$\frac{n!}{k!(n-k)!}p^k(1-p)^{n-k} \to \frac{\lambda^k}{k!}e^{-\lambda}$$

On observe la convergence de la loi binomiale vers la loi de Poisson.

4. Soit $X \sim$ Poisson(λ), $\lambda > 0$, cela veut dire que pout tout $k \in \mathbb{N}$,

$$\mathbb{P}(X = k) = \frac{\lambda^k}{k!}e^{-\lambda}$$

Le calcul de $\mathbb{E}[X]$ peut se faire en remarquant

$$\begin{aligned}\mathbb{E}[X] &= \sum_{k=0}^{+\infty} k\mathbb{P}(X = k) = \sum_{k=0}^{+\infty} k\frac{\lambda^k}{k!}e^{-\lambda} \\ &= \sum_{k=1}^{+\infty} \frac{\lambda^k}{(k-1)!}e^{-\lambda} = \lambda e^{-\lambda}\sum_{k=0}^{+\infty}\frac{\lambda^k}{k!} = \lambda\end{aligned}$$

parce que le développement en série entière de e^λ s'écrit

$$e^\lambda = \sum_{k=0}^{+\infty}\frac{\lambda^k}{k!}$$

Puisque $\mathbb{E}[X] = \lambda$, pour un n-échantillon $(X_1, \cdots, X_n)$ de la loi de Poisson(λ), un estimateur de λ sera la moyenne empirique

$$\bar{X}_n = \frac{1}{n}\sum_{i=1}^n X_i$$

5. Pour $X \sim$ Poisson(λ), et $x \in \mathbb{N}$,

$$\mathbb{P}(X = x) = \frac{\lambda^x}{x!}e^{-\lambda}.$$

Donc pour $X_1, ..., X_n$ i.i.d $\sim$ Poisson(λ), (il s'git d'une loi discrète ici) pour un échantillon $(x_1, \cdots, x_n)$ donné, la vraisemblance s'écrit

$$\begin{aligned}L(\lambda; x_1, \cdots, x_n) &= \prod_{i=1}^n \mathbb{P}(X_i = x_i) \\ &= \prod_{i=1}^n \frac{\lambda^{x_i}}{x_i!}e^{-\lambda} \\ &= \frac{e^{-n\lambda}}{x_1!\cdots x_n!}\lambda^{x_1+\cdots+x_n}\end{aligned}$$

6. La log-vraisemblance s'écrit

$$\begin{aligned}\ell(\lambda; x_1, \cdots, x_n) &= \log L(\lambda; x_1, \cdots, x_n) \\ &= -n\lambda - \log(x_1!\cdots x_n!) + (x_1 + \cdots + x_n)\log(\lambda)\end{aligned}$$

On cherche à maximiser $\ell(\lambda; x_1, \cdots, x_n)$ pour $\lambda > 0$. On remarque que $\ell \to -\infty$ quand $\lambda \to 0$, et $\ell \to -\infty$ quand $\lambda \to +\infty$. En plus

$$\frac{\partial \ell}{\partial \lambda} = 0 \implies -n + (x_1 + \cdots + x_n)\frac{1}{\lambda} = 0$$

ce qui donne l'estimation de maximum de vraisemblance

$$\hat{\lambda}^{\text{MV}} = \frac{1}{n}(x_1 + \cdots + x_n)$$

L'estimateur du maximum de vraisemblance de λ est donc donné par :

$$\hat{\lambda}^{\text{MV}} = \frac{1}{n}(X_1 + \cdots + X_n)$$

et nous retrouvons donc

$$\hat{\lambda}^{\text{MV}} = \bar{X}_n.$$

7. Si $X_1 = \cdots = X_n = 0$,

$$\hat{\lambda}^{\text{MV}} = 0$$

Exercice 2

Le 28 Janvier 1986 le réservoir de la navette Challenger explose juste après le décollage. Supposons que la probabilité d'explosion du réservoir à la température t est donnée par :

$$p(t) = \frac{e^{a+bt}}{1 + e^{a+bt}}$$

Montrer que l'application $x \mapsto \frac{e^x}{1+e^x}$ est inversible et déterminer son inverse. Déterminer la vraisemblance $L(a, b)$ ainsi que la log-vraisemblance $\ell(a, b)$.

Corrigé Exercice 2

Notons $\sigma : x \mapsto \frac{e^x}{1+e^x}$ (c'est la fonction sigmoid !). On remarque que $0 \leqslant \sigma(x) \leqslant 1$ pour tout $x \in \mathbb{R}$, et que $\sigma(x) \to 0$ pour $x \to -\infty$ et $\sigma(x) \to 1$ pour $x \to \infty$; et

$$\sigma'(x) = \frac{e^x}{(1 + e^x)^2} > 0$$

donc σ est strictement croissante ; donc σ est inversible et son inverse s'écrit

$$\sigma^{-1}(y) = \log \frac{y}{1 - y} \qquad \text{pour } y \in]0, 1[$$

Maintenant soit un n-échantillon $(T_1, \cdots, T_n)$ de T suivant la probabilité donnée par $p(t)$, et $(t_1, \cdots, t_n)$ une réalisation. Alors la vraisemblance s'écrit

$$L(a, b; t_1, \cdots, t_n) = \prod_{i=1}^{n} p(t_i) = \prod_{i=1}^{n} \frac{e^{a+bt_i}}{1 + e^{a+bt_i}}$$

et la log-vraisemblance s'écrit

$$\ell(a, b; t_1, \cdots, t_n) = \log L(a, b; t_1, \cdots, t_n) = \sum_{i=1}^{n} \log \frac{e^{a+bt_i}}{1 + e^{a+bt_i}}$$

Exercice 3

Supposons que $X \sim \mathcal{N}(0, \sigma^2)$ et $Y \sim \mathcal{N}(0, \sigma^2)$ sont des v.a. indépendantes. Alors $R = \sqrt{X^2 + Y^2}$ suit une loi de Rayleigh(σ) de densité :

$$f_\sigma(x) = \frac{x}{\sigma^2} e^{-\frac{x^2}{2\sigma^2}}, \quad \forall x \geqslant 0.$$

Quel est l'estimateur du maximum de vraisemblance de σ ?

Corrigé Exercice 3

Soit un n-échantillon $(R_1, \cdots, R_n)$ de R suivant la loi de Rayleigh(σ) ($\sigma > 0$), et $(r_1, \cdots, r_n)$ une réalisation (donc $r_i \geqslant 0, \forall i$). Alors la vraisemblance s'écrit

$$L(\sigma; r_1, \cdots, r_n) = \prod_{i=1}^{n} f_\sigma(r_i) = \prod_{i=1}^{n} \frac{r_i}{\sigma^2} e^{-\frac{r_i^2}{2\sigma^2}}$$

et la log-vraisemblance s'écrit

$$\ell(\sigma; r_1, \cdots, r_n) = \sum_{i=1}^{n} \left(\log r_i - 2\log\sigma - \frac{r_i^2}{2\sigma^2} \right) = -2n\log\sigma - \frac{1}{2\sigma^2} \sum_{i=1}^{n} r_i^2 + \sum_{i=1}^{n} \log r_i$$

où on observe que $\ell \to -\infty$ pour $\sigma \to 0$ et $\ell \to -\infty$ pour $\sigma \to +\infty$, et

$$\frac{\partial \ell}{\partial \lambda} = 0 \implies \frac{-2n}{\sigma} + \frac{1}{\sigma^3} \sum_{i=1}^{n} r_i^2 = 0 \implies \sigma = \sqrt{\frac{1}{2n} \sum_{i=1}^{n} r_i^2}$$

donc l'estimateur de maximum de vraisemblance de σ s'écrit

$$\hat{\sigma}^{\text{MV}} = \sqrt{\frac{1}{2n} \sum_{i=1}^{n} R_i^2}$$

Exercice 4

Soient $X_1, \cdots, X_n$ i.i.d $\sim \mathcal{U}([-\alpha, \alpha])$. Quel est l'estimateur du maximum de vraisemblance de α ?

Corrigé Exercice 4

Rappelons que la densité d'une v.a X suivant la loi uniforme $\mathcal{U}([-\alpha, \alpha])$ s'écrit

$$f_X(x) = \frac{1}{2\alpha} 1_{x \in [-\alpha, \alpha]}.$$

Soit $(x_1, \cdots, x_n)$ une réalisation de $(X_1, \cdots, X_n)$. Alors la vraisemblance s'écrit

$$L(\alpha; x_1, \cdots, x_n) = \prod_{i=1}^{n} f_X(x_i) = \prod_{i=1}^{n} \frac{1}{2\alpha} 1_{x_i \in [-\alpha, \alpha]} = \left(\frac{1}{2\alpha} \right)^n \prod_{i=1}^{n} 1_{x_i \in [-\alpha, \alpha]}$$

où

$$\prod_{i=1}^{n} 1_{x_i \in [-\alpha, \alpha]} = \prod_{i=1}^{n} 1_{\alpha \geqslant |x_i|} = \begin{cases} 1 & \text{si } \alpha \geqslant |x_i|, \forall i, \\ 0 & \text{sinon,} \end{cases}$$

donc

$$L(\alpha; x_1, \cdots, x_n) = \begin{cases} \left(\frac{1}{2\alpha}\right)^n & \text{si } \alpha \geqslant \max_i |x_i|, \\ 0 & \text{sinon.} \end{cases}$$

La fonction $\alpha \mapsto \left(\frac{1}{2\alpha}\right)^n$ est décroissante en α pour $\alpha > 0$, donc L prend son maximum en $\alpha = \max_i |x_i|$, donc l'estimateur de maximum de vraisemblance s'écrit

$$\hat{\alpha}^{\text{MV}} = \max_{i=1,\cdots,n} |X_i|$$