STATISTIQUE MATHÉMATIQUE - MS ENSAI
ÉLEMENTS DE CORRECTION (chapitre 3, sept. 2025)

Densité des statistiques d'ordre

La densité jointe de toutes les statistiques d'ordre est

$$f(x_{(1)},...,x_{(n)}) = n!\prod_{i=1}^{n} f(x_{(i)})\mathbb{1}_{[x_{(1)}<x_{(2)}<...<x_{(n)}]}$$

La densité marginale la statistique d'ordre k est donnée par

$$f_k(x) = \frac{n!}{(k-1)!(n-k)!} f(x)\mathrm{F}(x)^{k-1}(1-\mathrm{F}(x))^{n-k}$$

Pour le démontrer, on calcule la fonction de répartition $\mathbb{P}[\mathrm{X}_{(k)} \le x]$. Cet évènement signifie « au moins k des n valeurs de l'échantillon sont plus petites que x. D'après la formule de la loi binomiale, cela signifie :

$$\mathbb{P}[\mathrm{X}_{(k)} \le x] = \sum_{i=k}^{n} \binom{n}{i} \mathrm{F}(x)^i (1-\mathrm{F}(x))^{n-i}$$

on dérive alors par rapport à x cette expression (c'est assez calculatoire...).

Somme de n variables exponentielles indépendantes de paramètre λ

La transformée de Laplace de la densité est :

$$\begin{aligned}\psi(t) = \mathbb{E}[e^{-t\mathrm{X}}] &= \lambda\int_0^{+\infty} e^{-(t+\lambda)x}dx \\ &= \frac{\lambda}{\lambda+t}\end{aligned}$$

La transformée de Laplace de la somme est, d'après le caractère i.i.d. de l'échantillon, le produit des n transformées :

$$\psi_n(t) = \mathbb{E}[e^{-t\mathrm{S}}] = \left(\frac{\lambda}{\lambda+t}\right)^n$$

Enfin, si l'on calcule la transformée de Laplace d'une variable Y de loi Gamma de paramètres (n,λ), on trouve :

$$\begin{aligned}\phi_n(t) = \mathbb{E}[e^{-t\mathrm{Y}}] &= \int_0^{+\infty} e^{-ty}\frac{\lambda^n y^{n-1}e^{-\lambda y}}{(n-1)!}dy \\ &= \left(\frac{\lambda}{\lambda+t}\right)^n\end{aligned}$$

Par identification, on en déduit que la somme de n variables exponentielles indépendantes de même paramètre λ est la loi Gamma précédente.

Minimalité du couple min max sur $[\theta-1/2,\theta+1/2]$ pour un échantillon uniforme

La vraisemblance d'un échantillon de cette loi s'écrit

$$\begin{aligned}\mathrm{L}(x,\theta) &= \mathbb{1}_{[\theta-1/2 \le x_{(1)} \le x_{(n)} \le \theta+1/2]} &(1)\\ &= \mathbb{1}_{[x_{(n)}-1/2 \le \theta \le x_{(1)}+1/2]} &(2)\end{aligned}$$

Pour une observation x fixée, notons $\mathbb{J}_x$ l'intervalle $[x_{(n)}-1/2\ ;\ x_{(1)}+1/2]$. Fixons x et y deux observations.

$$\mathrm{LR}(x,y,\theta) = \mathbb{1}_{\mathbb{J}_x(\theta)}/\mathbb{1}_{\mathbb{J}_y(\theta)} \qquad (3)$$

Lorsque cette fonction (de θ) est définie, elle vaut 1 ou 0. Plus précisément,

$$\mathrm{LR}(x,y,\theta) = \begin{cases} 1 & \text{si } \theta\in\mathbb{J}_x \text{ et } \theta\in\mathbb{J}_y \\ 0 & \text{si } \theta\notin\mathbb{J}_x \text{ et } \theta\in\mathbb{J}_y\end{cases} \qquad (4)$$

et par conséquent il est clair que sa valeur dépend de θ dès lors que $\mathbb{J}_x \ne \mathbb{J}_y$.

En effet, si par exemple $y_{(n)} < x_{(n)}$, $y_{(n)}-1/2 < x_{(n)}-1/2$. Considérons alors $\theta \in]y_{(n)}-1/2\ ;\ x_{(n)}-1/2[$ tout en étant plus petit que $x_{(1)}+1/2$ et $y_{(1)}+1/2$ (θ peut varier dans $]0,\infty[$). $\mathrm{L}(y,\theta)$ vaut 1 tandis que $\mathrm{L}(x,\theta)=0$, donc $\mathrm{LR}(x,y,\theta)=0$. Si maintenant $\theta > x_{(n)}-1/2$, $\mathrm{LR}(x,y,\theta)=0$. On peut raisonner de la même façon avec $x_{(1)}$ et $y_{(1)}$. En clair, LR ne dépend pas de θ si et seulement si $x_{(n)} = y_{(n)}$ et $x_{(1)} = y_{(1)}$.

Loi de l'étendue sur $[\theta-1/2,\theta+1/2]$ pour un échantillon uniforme

Nous restons sur le même échantillon uniforme avec θ inconnu. Notons X une va générique, de même loi que $\mathrm{X}_1,...,\mathrm{X}_n$. Nous allons déterminer la loi du couple $\mathrm{S} = (\mathrm{X}_{(1)},\mathrm{X}_{(n)})$ puis la loi de l'étendue $\mathrm{E}_n = \mathrm{X}_{(n)} - \mathrm{X}_{(1)}$. Déterminons la fonction de répartition $\mathrm{F}(x,y)$ du couple :

$$\begin{aligned}\mathrm{F}(x,y) &= \mathbb{P}\left([\mathrm{X}_{(1)} \le x] \cap [\mathrm{X}_{(n)} \le y]\right) &(5)\\ &= \mathbb{P}[\mathrm{X}_{(n)} \le y] - \mathbb{P}\left([\mathrm{X}_{(1)} \ge x] \cap [\mathrm{X}_{(n)} \le y]\right) &(6)\\ &= \mathbb{P}[\mathrm{X} \le y]^n - \mathbb{P}[x \le \mathrm{X} \le y]^n &(7)\end{aligned}$$

Le passage de la première à la seconde ligne vient de la partition suivante :

$$\begin{aligned}[\mathrm{X}_{(n)} \le y] = \left([\mathrm{X}_{(1)} \le x] \cap [\mathrm{X}_{(n)} \le y]\right) \cup ...\\ ...\left([\mathrm{X}_{(1)} > x] \cap [\mathrm{X}_{(n)} \le y]\right)\end{aligned}$$

Le passage de la seconde à la troisième ligne vient du caractère i.i.d. de l'échantillon. En dérivant en x puis en y, on obtient la densité du couple :

$$g(x,y) = \frac{\partial \mathrm{F}(x,y)}{\partial y \partial x} = n(n-1)f(x)f(y)(\mathrm{F}(y)-\mathrm{F}(x))^{n-2}$$

et la loi du couple est portée par le sous-ensemble de $\mathbb{R}^2$ donné par $\Delta = \{(x,y) \in [\theta-1/2,\theta+1/2]^2 : x \le y\}$. Puisque $f(x) = f(y) = 1$ sur cet ensemble, on a :

$$g(x,y) = n(n-1)(y-x)^{n-2}\mathbb{1}_{[\theta-1/2\le x\le y\le\theta+1/2]}$$

Maintenant pour obtenir la loi de l'étendue E_n partant de la loi du couple, le plus simple est d'utiliser la densité du couple et d'effectuer un changement de variables de $\mathbb{R}^2$ dans $\mathbb{R}^2$. Posons $(\mathrm{R},\mathrm{S}) = (\mathrm{X}_{(n)} - \mathrm{X}_{(1)},\mathrm{X}_{(1)})$. Ce changement de variables est linéaire, par conséquent la valeur absolue du déterminant jacobien vaut 1 et l'inverse de la transformation est donnée par $(X_{(1)}, X_{(n)}) = (S, S+R)$. Notons (x, y) une réalisation du couple initial et (r, s) les variables sous-jacentes à (R, S). Lorsque x et y varient dans $[\theta - 1/2, \theta + 1/2]$ avec $x < y$, alors s varie dans ce même intervalle et r varie dans $[0, 1]$ puisque $0 < y - x < 1$. Enfin, la densité de (R, S) est donnée par

$$f(r, s) = g(s, r+s)$$

La densité de l'étendue est la marginale de cette densité par rapport à la première variable. Ainsi,

$$\begin{aligned} f_R(r) &= \int_{1-\theta/2}^{1+\theta/2-r} ds \mathbb{1}_{[0,1]}(r) \\ &= n(n-1)r^{n-2}\mathbb{1}_{[0,1]}(r)\,(\theta + 1/2 - r - \theta + 1/2) \\ &= n(n-1)r^{n-2}(1-r)\mathbb{1}_{[0,1]}(r) \end{aligned}$$

qui est la densité d'une loi Béta $B(n-1, 2)$. Par ailleurs, on constate que la densité ne dépend pas de θ, ce qui prouve que la loi est libre de θ.