

Rappels de statistique mathématique
Corrigé des travaux dirigés n° 8

Guillaume Lacôte
Bureau E03

✉ Guillaume.Lacote@ensae.fr

☞ <http://ensae.no-ip.com/SE222/>

Corrigé de l'exercice 1

☞ Q1 Posons $\mu = \delta_a + \lambda$ et $\mu_n = \mu^{\otimes n}$.

Remarque : nous restreindrons les démonstrations au cas où une seule observation X_1 est disponible, dans la mesure où car la généralisation à n observations i.i.d. se réduit à un passage au produit immédiat :

$$dP_{X_1, \dots, X_n} = \prod_{i=1}^n dP_{X_i}$$

$$\frac{dP_{X_1, \dots, X_n}}{d\mu_n} = \prod_{i=1}^n \frac{dP_{X_i}}{d\mu}$$

Remarque préliminaire

Pour \mathcal{A} ensemble mesurable, on a

$$\begin{aligned} \mathbb{P}(X_1 \in \mathcal{A}) &= \mathbb{P}(X_1 \in \mathcal{A} | X_1 = a) \mathbb{P}(X_1 = a) + \mathbb{P}(X_1 \in \mathcal{A} | X_1 \neq a) \mathbb{P}(X_1 \neq a) \\ &= \alpha \mathbb{P}(X_1 \in \mathcal{A} | X_1 = a) + (1 - \alpha) \int_{\mathcal{A}} \varphi(x) d\lambda(x) \end{aligned}$$

où φ désigne la densité de la loi normale centrée réduite.

Or $\mathbb{P}(X_1 \in \mathcal{A} | X_1 = a)$ vaut 1 si $a \in \mathcal{A}$ et 0 sinon, de sorte que

$$\mathbb{P}(X_1 \in \mathcal{A}) = \alpha \mathbf{1}_{a \in \mathcal{A}} + (1 - \alpha) \int_{\mathcal{A}} \varphi(x) d\lambda(x)$$

Or μ est une mesure dominante d'une variable X_1 ssi :

$$\forall \mathcal{A} \text{ mesurable, } \mu(\mathcal{A}) = 0 : P_{X_1}(\mathcal{A}) = 0$$

En l'occurrence pour tout \mathcal{A} mesurable :

$$- \mu(\mathcal{A}) = 0 : \delta_a(\mathcal{A}) = 0$$

$$- \delta_a(\mathcal{A}) = 0 : a \notin \mathcal{A}, \text{ donc d'après 1 } \lambda(\mathcal{A}) = 0 : P_{X_1}(\mathcal{A}) = 0$$

Donc $(\delta_a + \lambda)^{\otimes n}$ est une mesure dominante.

☞ Q2 On a $\mathbf{1}_{a \in \mathcal{A}} = \int_{\mathcal{A}} \delta_a(x)$,

donc en reportant dans l'équation 1 il vient

$$P(X_1 \in \mathcal{A}) = \int_{\mathcal{A}} (\alpha \delta_a(x) + (1 - \alpha) \varphi(x)) d\lambda(x)$$

d'où on tire facilement

$$P(X_1 \in \mathcal{A}) = \int_{\mathcal{A}} (\alpha \mathbf{1}_a(x) + (1 - \alpha) \varphi(x) (1 - \mathbf{1}_a(x))) d\mu(x)$$

de sorte qu'en définitive

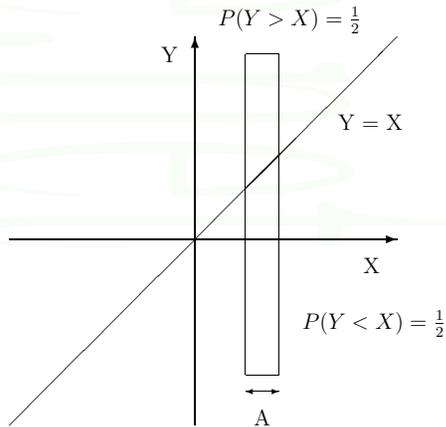
$$\frac{dP_{X_1}(x_1)}{d(\delta_a + \lambda)} = \alpha \mathbf{1}_a(x_1) + (1 - \alpha) (1 - \mathbf{1}_a(x_1)) \varphi(x_1)$$

d'où le résultat par indépendance de (X_1, \dots, X_n) .

*
* *

Corrigé de l'exercice 2

☞ Q1 Si un couple de variables aléatoires réelles (X, Y) vérifie \mathcal{P} , alors au voisinage de tout X , la probabilité est la même pour Y d'être au-dessus ou au-dessous de X , comme illustré sur la figure suivante :



Pour simplifier la présentation, supposons que vous gagnez +1 si vous avez deviné juste et -1 sinon ; ce problème revient à savoir si vous pouvez vous garantir un gain espéré strictement positif.

☞ Q1 Soit $\sigma : \mathbb{R} \rightarrow \{-1, +1\}$ votre stratégie, et soient (x, y) une réalisation du tirage.

- Si x vous est présenté :
Si $x > y$, votre gain est $\sigma(x)$; si $x < y$, votre gain est $-\sigma(x)$. Votre gain espéré, sachant que x vous est présenté, est donc nul.
- Si y vous est présenté, il en va de même.

Le choix de l'enveloppe étant indépendant du tirage de X et de Y , votre gain espéré est toujours nul, quelle que soit votre stratégie.

Comme l'hypothèse \mathcal{P} semble "raisonnable", l'intuition suggère qu'à ce jeu vous ne pourrez jamais obtenir un gain espéré strictement positif. Cette intuition est cependant erronée.

☞ Q2 (a) On a dans tous les cas $\sigma(x) = \sigma(y) = -1$ car $x < s$ et $y < s$.
Or si $x < y$, le gain espéré est

$$\mathbb{P}(\text{le nombre présenté est } x) \cdot 1 + \mathbb{P}(\text{le nombre présenté est } y) \cdot (-1) = 0$$

puisque ces deux alternatives sont équiprobables et que le choix de l'enveloppe est indépendant du tirage de X et de Y .

Il en va de même lorsque $x > y$, de sorte que le gain espéré, sachant que $x < s$ et $y < s$ est nul.

(b) Si le nombre présenté est x , on a $\sigma(x) = -1$ car $x < s$; or $x < y$, donc la réponse est juste et le gain est +1.
Si le nombre présenté est y , on a $\sigma(y) = +1$ car $y > s$; or $y > x$, donc la réponse est encore juste et le gain est toujours +1.
Ainsi dans tous les cas le gain espéré, sachant que $x < s < y$, est +1. Il en va de même bien-sûr sachant que $y < s < x$.

(c) D'après les résultats précédents le gain espéré en jouant σ_s est donc

$$g(s) = \mathbb{P}(X < s < Y) + \mathbb{P}(Y < s < X)$$

Or ces deux probabilités ne peuvent être simultanément nulles pour tout s : en effet

$$\begin{aligned} \int_{\mathbb{R}} \mathbb{P}(X < s < Y) + \mathbb{P}(Y < s < X) ds &= \int_{\mathbb{R}} \left(\int_{\mathbb{R}^2} (\mathbb{1}_{x < s < y} + \mathbb{1}_{y < s < x}) f_{X,Y}(x, y) dx dy \right) ds \\ &= \int_{\mathbb{R}^2} \left(\int_{\mathbb{R}} \mathbb{1}_{x < s < y} ds + \int_{\mathbb{R}} \mathbb{1}_{y < s < x} ds \right) f_{X,Y}(x, y) dx dy \\ &= \int_{\mathbb{R}^2} (\mathbb{1}_{x < y} (y - x) + \mathbb{1}_{y < x} (x - y)) f_{X,Y}(x, y) dx dy \\ &= \int_{\mathbb{R}^2} |x - y| \mathbb{1}_{x \neq y} f_{X,Y}(x, y) dx dy \\ &= \mathbb{E}(|X - Y|) \end{aligned}$$

Si $\mathbb{P}(X \neq Y) = 0$ alors $\mathbb{P}(X = Y) = 1$ et le jeu devient trivial : la stratégie qui présente le nombre tiré est le plus grand est presque sûrement toujours gagnante ; en particulier lorsque $s = 0$ $g(0) > 0$.

Et si $\mathbb{P}(X \neq Y) \neq 0$ alors $\mathbb{E}(|X - Y|) > 0$, et donc $\int_{\mathbb{R}} g(s) ds > 0$.

Donc dans tous les cas

$$\exists s \in \mathbb{R} / g(s) > 0$$

En fait, si (X, Y) est distribué de telle sorte que toute partie de mesure non-nulle a une probabilité non-nulle, alors n'importe quel seuil s assure un gain espéré strictement positif.

☞ Q3 Si une telle distribution sur \mathbb{R}^2 existait, alors d'après la question (1) dans le jeu de Blackwell associé nul ne saurait gagner plus que zéro en moyenne, ce qui est pourtant possible d'après la question (2).

☞ Q1 La formulation du paradoxe suggère que l'on se donne une distribution de probabilité a priori sur les deux sommes X et Y ; cette distribution est nécessairement telle que $\mathbb{P}(X = 2Y \text{ ou } Y = 2X) = 1$.

Etant donné $x \in \mathbb{R}$ et à supposer que $X = x$, l'hypothèse \mathcal{H}_x s'exprime en termes mathématiques sous la forme

$$\mathbb{P}(Y = 2x | X = x) = \mathbb{P}\left(Y = \frac{x}{2} | X = x\right) = \frac{1}{2}$$

☞ Q2 L'hypothèse \mathcal{H}_y analogue à la précédente en $Y = y \in \mathbb{R}$ s'écrit

$$\mathbb{P}(X = 2y | Y = y) = \mathbb{P}\left(X = \frac{y}{2} | Y = y\right) = \frac{1}{2}$$

Dans l'énoncé du paradoxe, l'hypothèse selon laquelle "chacune a autant de chances que l'autre de contenir la plus grosse somme" se traduit alors en \mathcal{H}_x quelle que soit la somme x de la première enveloppe, et \mathcal{H}_y quelle que soit la somme y de la deuxième enveloppe.

En d'autres termes

$$\begin{aligned} \forall x \in \mathbb{R}, \mathbb{P}(Y = 2x | X = x) &= \mathbb{P}\left(Y = \frac{x}{2} | X = x\right) = \frac{1}{2} \\ \forall y \in \mathbb{R}, \mathbb{P}(X = 2y | Y = y) &= \mathbb{P}\left(X = \frac{y}{2} | Y = y\right) = \frac{1}{2} \end{aligned}$$

Or pour tout $x \in \mathbb{R}$ l'événement $Y = 2x$ est égal à l'événement $Y > x$, et l'événement $Y = \frac{x}{2}$ à $Y < x$. Par suite

$$\begin{aligned} \forall x \in \mathbb{R}, \mathbb{P}(Y > x | X = x) &= \mathbb{P}(Y < x | X = x) = \frac{1}{2} \\ \forall y \in \mathbb{R}, \mathbb{P}(X > y | Y = y) &= \mathbb{P}(X < y | Y = y) = \frac{1}{2} \end{aligned}$$

Par conséquent pour toute partie mesurable \mathcal{A} de \mathbb{R} telle que $\mathbb{P}(X \in \mathcal{A}) \neq 0$ et $\mathbb{P}(Y \in \mathcal{A}) \neq 0$ il vient

$$\begin{aligned} \mathbb{P}(Y > X | X \in \mathcal{A}) &= \mathbb{P}(Y < X | X \in \mathcal{A}) = \frac{1}{2} \\ \mathbb{P}(X > Y | Y \in \mathcal{A}) &= \mathbb{P}(X < Y | Y \in \mathcal{A}) = \frac{1}{2} \end{aligned}$$

de sorte que (X, Y) vérifie la propriété \mathcal{P} .

☞ Q3 Ainsi, à supposer que l'on se donne une distribution a priori sur les sommes d'argent de chacune des deux enveloppes, il n'est pas possible de traduire l'hypothèse "chacune a autant de chances que l'autre de contenir la plus grosse somme" comme il est suggéré, car la distribution vérifierait alors la propriété \mathcal{P} ce qui n'est pas possible.

La démarche proposée, bien qu'intuitive, n'est donc pas raisonnable : en particulier il est faux qu'après un nombre quelconque de changements d'enveloppe on ait "à nouveau intérêt à changer d'enveloppe".

*
* * *

Corrigé de l'exercice 3

☞ Q1 Il suffit de poser $c = -x_1$ dans la définition de l'équivariance.

☞ Q2 Le risque associé vérifie :

$$\begin{aligned} R_w(T, \theta) &= E_\theta[w(T - \theta)] \\ &= \int_{(\mathbb{R}^k)^n} w(T(x_1, \dots, x_n) - \theta) \prod_{i=1}^n f(x_i - \theta) dx_i \\ &= \int_{(\mathbb{R}^k)^n} w(T(x_1 - \theta, \dots, x_n - \theta)) \prod_{i=1}^n f(x_i - \theta) dx_i \\ &= \int_{(\mathbb{R}^k)^n} w(T(u_1, \dots, u_n)) \prod_{i=1}^n f(u_i) du_i \\ &= R_w(T, 0) \end{aligned}$$

en effectuant un changement de variable $u_i = x_i - \theta$.

☞ Q3 (a) Soit $c \in \mathbb{R}^k$; on a

$$\begin{aligned} \psi(x + c) &= \int w(x + c - u) \prod_{i=1}^n f(X_i - u) du \\ &= \int w(x - v) \prod_{i=1}^n f(X_i - c - v) dv \end{aligned}$$

en effectuant le changement de variable $v = u - c$.

En particulier $\psi(x + X_1) = \int w(x - v) f(-v) \prod_{i=2}^n f((X_i - X_1) - v) dv$

Donc $\psi(x + X_1)$ ne dépend que de x et des variables aléatoires $X_2 - X_1, \dots, X_n - X_1$ soit donc T_1 tel que $\arg \min_{x \in \mathbb{R}^k} \psi(x + X_1) = T_1(X_2 - X_1, \dots, X_n - X_1)$.

Or $\forall \phi, \forall c \in \mathbb{R}^k, \arg \inf (x \mapsto \phi(x + c)) = (\arg \inf (x \mapsto \phi(x))) - c$.

Donc en définitive

$$T_1(X_2 - X_1, \dots, X_n - X_1) = (\arg \inf \psi) - X_1 = T^* - X_1$$

Autrement dit T^* est équivariant d'après la question précédente.

(b)

$$\begin{aligned} &E_\theta[\varphi(X_1, \dots, X_n) | X_2 - X_1 = y_2, \dots, X_n - X_1 = y_n] \\ &= \int \varphi(x_1, x_1 + y_2, \dots, x_1 + y_n) d\mathbb{P}(X_1 = x_1 | X_2 - X_1 = y_2, \dots, X_n - X_1 = y_n) \\ &= \int \varphi(x_1, x_1 + y_2, \dots, x_1 + y_n) \frac{f(x_1 - \theta) \prod_{i=2}^n f(x_1 + y_i - \theta)}{[\int f(x_1 - \theta) \prod_{i=2}^n f(x_1 + y_i - \theta) dx_1]} dx_1 \end{aligned}$$

Effectuons le changement de variables $u = \theta - x_1 + c$ (c constante quelconque) dans deux intégrales :

$$\begin{aligned} &E_\theta[\varphi(X_1, \dots, X_n) | X_2 - X_1 = y_2, \dots, X_n - X_1 = y_n] \\ &= \int \varphi(c + \theta - u, c + y_2 + \theta - u, \dots) \frac{f(c - u) \prod_{i=2}^n f(c + y_i - u)}{[\int f(c - u) \prod_{i=2}^n f(c + y_i - u) du]} du \end{aligned}$$

La relation suivante est alors vraie pour tout $c \in \mathbb{R}^k$:

$$\begin{aligned} &E_\theta[\varphi(X_1, \dots, X_n) | X_2 - X_1, \dots, X_n - X_1] \\ &= \int \varphi(c + \theta - u, c + X_2 - X_1 + \theta - u, \dots) \frac{f(c - u) \prod_{i=2}^n f(c + X_i - X_1 - u)}{[\int f(c - u) \prod_{i=2}^n f(c + X_i - X_1 - u) du]} du \end{aligned}$$

elle est donc vraie aussi pour $c = X_1$:

$$\begin{aligned} &E_\theta[\varphi(X_1, \dots, X_n) | X_2 - X_1, \dots, X_n - X_1] \\ &= \int \varphi(X_1 + \theta - u, X_2 + \theta - u, \dots) \frac{\prod_{i=1}^n f(X_i - u)}{[\int f(X_1 - u) \prod_{i=2}^n f(X_i - u) du]} du \end{aligned}$$

(c) Soit T un estimateur équivariant quelconque. On va montrer que T^* domine T . Pour cela, il suffit de prouver que $R_w(T, 0) \geq R_w(T^*, 0)$, selon le résultat de la question (2).

$$\begin{aligned} R_w(T, 0) &= \mathbb{E}_0[w(T)] \\ &= \mathbb{E}_0[\mathbb{E}_0[w(T)|X_2 - X_1, \dots, X_n - X_1]] \\ &= \mathbb{E}_0 \left[\int_{\mathbb{R}} w(T - u) \frac{\prod_{i=1}^n f(X_i - u)}{\int_{\mathbb{R}} \prod_{i=1}^n f(X_i - u) du} du \right] \end{aligned}$$

Or par définition de T^* , pour tous $x \in \mathbb{R}^n$ et $t \in \mathbb{R}$

$$\int_{\mathbb{R}} w(T^*(x) - u) \prod_{i=1}^n f(x_i - u) du \leq \int_{\mathbb{R}} w(t - u) \prod_{i=1}^n f(x_i - u) du$$

donc

$$\begin{aligned} R_w(T, 0) &\geq \mathbb{E}_0 \left[\int_{\mathbb{R}} w(T^*(X) - u) \frac{\prod_{i=1}^n f(X_i - u)}{\int_{\mathbb{R}} \prod_{i=1}^n f(X_i - u) du} du \right] \\ &= R_w(T^*, 0) \end{aligned}$$

Donc T^* domine T .

☞ Q4 Si $w(x - y) = \|x - y\|^2$, ψ vaut :

$$\psi(x) = \int_{\mathbb{R}^k} \|x - u\|^2 \prod_{i=1}^n f(X_i - u) du$$

et le gradient $\psi'(x)$ vaut : $\psi'(x) = \int_{\mathbb{R}^k} 2(x - u) \prod_{i=1}^n f(X_i - u) du$. T^* est défini par l'équation $\psi'(T^*) = 0$, qui se réduit ici :

$$T^* = \frac{\int_{\mathbb{R}^k} u \prod_{i=1}^n f(X_i - u) du}{\int_{\mathbb{R}^k} \prod_{i=1}^n f(X_i - u) du}$$

☞ Q5 Lorsque $X_i \sim \mathcal{N}(0, 1)$, on a

$$T^* = \frac{1}{n} \sum_{i=1}^n X_i$$

Si $X_i \sim \mathcal{U}(0, 1)$, il vient

$$T^* = \frac{\min(X_i) + \max(X_i)}{2}$$

*
* *

Corrigé de l'exercice 4

☞ Q1 Les observations étant les effectifs N_k de chaque classe $k \in \llbracket 1, K \rrbracket$, la vraisemblance s'écrit

$$L_{N_1, \dots, N_K}(n_1, \dots, n_K; p_1, \dots, p_K) = \frac{n!}{n_1! \dots n_K!} \prod_{k=1}^K p_k^{n_k} \quad \text{en notant } n = n_1 + \dots + n_K$$

La maximisation sans hypothèse conduit aux conditions nécessaires du premier ordre, en n'omettant la contrainte $p_1 + \dots + p_K = 1$

$$\begin{aligned} &\max_{p_1, \dots, p_K} \left(\sum_{k=1}^K n_k \ln p_k \right) \\ &\text{s.c. } \{ p_1 + \dots + p_K = 1 \} \end{aligned}$$

Le lagrangien associé s'écrit

$$\mathcal{L}(p_1, \dots, p_K, \lambda) = \sum_{k=1}^K n_k \ln p_k + \lambda \left(1 - \sum_{k=1}^K p_k \right)$$

et il vient

$$\forall k \in \llbracket 1, K \rrbracket, \frac{n_k}{p_k} - \lambda = 0$$

Or en sommant de 1 à K il vient

$$n = \sum_{k=1}^K n_k = \lambda \sum_{k=1}^K p_k = \lambda$$

de sorte que finalement

$$\forall k \in \llbracket 1, K - 1 \rrbracket, \hat{p}_k = \frac{n_k}{n}$$

(et on vérifie que cette condition est également suffisante).

La maximisation sous l'hypothèse H_0 correspond au programme de maximisation sous contrainte

$$\begin{aligned} &\max_{p_1, \dots, p_K} \left(\sum_{k=1}^K n_k \ln p_k \right) \\ &\text{s.c. } \begin{cases} p_1 + \dots + p_K = 1 \\ p_1 + p_2 = \frac{1}{2} \end{cases} \end{aligned}$$

Le lagrangien associé s'écrit

$$\mathcal{L}(p_1, \dots, p_K, \lambda, \mu) = \sum_{k=1}^K n_k \ln p_k + \lambda \left(1 - \sum_{k=1}^K p_k \right) + \mu \left(\frac{1}{2} - p_1 - p_2 \right)$$

et il vient

$$\begin{cases} \frac{\partial \mathcal{L}}{\partial p_1} = \frac{n_1}{p_1} - \lambda - \mu \\ \frac{\partial \mathcal{L}}{\partial p_2} = \frac{n_2}{p_2} - \lambda - \mu \\ \frac{\partial \mathcal{L}}{\partial p_k} = \frac{n_k}{p_k} - \lambda, \quad \forall k \in \llbracket 3, K \rrbracket \\ \frac{\partial \mathcal{L}}{\partial \lambda} = 1 - \sum_{k=1}^K p_k \\ \frac{\partial \mathcal{L}}{\partial \mu} = \frac{1}{2} - p_1 - p_2 \end{cases}$$

d'où finalement

$$\begin{cases} \widehat{p}_1^0 = \frac{n_1}{2(n_1+n_2)} \\ \widehat{p}_2^0 = \frac{n_2}{2(n_1+n_2)} \\ \widehat{p}_k^0 = \frac{n_k}{2(n_3+\dots+n_K)}, \quad \forall k \in \llbracket 3, K \rrbracket \end{cases}$$

Q2 Pour déterminer la statistique de Wald, appliquons le Théorème Central Limite à $\begin{pmatrix} \widehat{p}_1 \\ \vdots \\ \widehat{p}_K \end{pmatrix}$;

pour ce faire, calculons tout d'abord sa variance.

On a en effet pour $k \in \llbracket 1, K \rrbracket$, $\widehat{p}_k = \frac{n_k}{n} = \frac{1}{n} \sum_{i=1}^n \mathbf{1}_{X_i \in C_k}$ où C_k désigne la k -ième classe ; or

$$\mathbb{V}(\mathbf{1}_{X_i \in C_k}) = p_k \cdot 1 - (p_k)^2 = p_k(1 - p_k)$$

et

$$\text{Cov}(\mathbf{1}_{X_i \in C_k}, \mathbf{1}_{X_j \in C_l}) = \mathbb{E}(\mathbf{1}_{X_i \in C_k} \cdot \mathbf{1}_{X_j \in C_l}) - \mathbb{E}(\mathbf{1}_{X_i \in C_k})\mathbb{E}(\mathbf{1}_{X_j \in C_l}) = -p_k p_l \mathbf{1}_{i=j}$$

de sorte que

$$\sqrt{n} \begin{pmatrix} \widehat{p}_1 - p_1 \\ \vdots \\ \widehat{p}_K - p_K \end{pmatrix} \xrightarrow[n \rightarrow +\infty]{\mathcal{L}} \mathcal{N} \left(0, \begin{pmatrix} p_1(1-p_1) & -p_1 p_2 & \cdots & -p_1 p_K \\ -p_2 p_1 & p_2(1-p_2) & \cdots & -p_2 p_K \\ \vdots & \vdots & \ddots & \vdots \\ -p_K p_1 & -p_K p_2 & \cdots & p_K(1-p_K) \end{pmatrix} \right)$$

Posons finalement $g : \begin{pmatrix} \mathbb{R}^K & \rightarrow & \mathbb{R} \\ (x_1, \dots, x_K) & \mapsto & x_1 + x_2 \end{pmatrix}$; alors d'après le théorème de Slutsky, si H_0 est vraie on a

$$\begin{aligned} \sqrt{n} \left(\widehat{p}_1 + \widehat{p}_2 - \frac{1}{2} \right) &\xrightarrow[n \rightarrow +\infty]{\mathcal{L}} \mathcal{N} \left(0, \left(\frac{\partial g}{\partial p_1, \dots, p_K} \right)' \mathbb{V} \left(\begin{pmatrix} \widehat{p}_1 \\ \vdots \\ \widehat{p}_K \end{pmatrix} \right) \left(\frac{\partial g}{\partial p_1, \dots, p_K} \right) \right) \\ &= \mathcal{N}(0, (p_1 + p_2)(1 - p_1 - p_2)) \end{aligned}$$

d'où en centrant et en réduisant, puis en élevant au carré

$$\xi_n^W = n \frac{(\widehat{p}_1 + \widehat{p}_2 - \frac{1}{2})^2}{(\widehat{p}_1 + \widehat{p}_2)(1 - \widehat{p}_1 - \widehat{p}_2)} \xrightarrow[n \rightarrow +\infty]{\mathcal{L}} \chi_1^2$$

Q3 (a) On sait que $\widehat{p}_3^0 \xrightarrow[n \rightarrow +\infty]{p.s.} p_3$ et que $\widehat{p}_3 \xrightarrow[n \rightarrow +\infty]{p.s.} p_3$.

Soit donc pour $\lambda \in \mathbb{R}$ $p_3(\lambda) = \lambda \widehat{p}_3^0 + (1 - \lambda) \widehat{p}_3 \xrightarrow[n \rightarrow +\infty]{p.s.} p_3$.

Alors

$$\begin{aligned} \mathbb{V}_{as}(p_3(\lambda)) &= \mathbb{V}_{as} \left(\widehat{p}_3^0 + \lambda \left(\widehat{p}_3^0 - \widehat{p}_3 \right) \right) \\ &= \mathbb{V}_{as} \left(\widehat{p}_3^0 \right) + 2\lambda \text{Cov}_{as} \left(\widehat{p}_3^0, \widehat{p}_3^0 - \widehat{p}_3 \right) + \lambda^2 \mathbb{V}_{as} \left(\widehat{p}_3^0 - \widehat{p}_3 \right) \\ &\geq \mathbb{V}_{as} \left(\widehat{p}_3^0 \right) \end{aligned}$$

car \widehat{p}_3^0 est asymptotiquement efficace, car sa variance est égale à la borne FDCR $I(p_3)$. Par conséquent le polynôme du second degré $P(\mathbb{X}) = \mathbb{V}_{as} \left(\widehat{p}_3^0 - \widehat{p}_3 \right) \mathbb{X}^2 + 2\text{Cov}_{as} \left(\widehat{p}_3^0, \widehat{p}_3 - \widehat{p}_3^0 \right) \mathbb{X} + \mathbb{V}_{as} \left(\widehat{p}_3 - \widehat{p}_3^0 \right)$ est de signe constant. En particulier,

$$\text{Cov}_{as} \left(\widehat{p}_3^0, \widehat{p}_3^0 - \widehat{p}_3 \right) = 0$$

et donc

$$\begin{aligned} \mathbb{V}_{as}(\widehat{p}_3) &= \mathbb{V}_{as} \left(\widehat{p}_3^0 - \left(\widehat{p}_3^0 - \widehat{p}_3 \right) \right) \\ &= \mathbb{V}_{as} \left(\widehat{p}_3^0 \right) + \mathbb{V}_{as} \left(\widehat{p}_3 - \widehat{p}_3^0 \right) \end{aligned}$$

ce qui s'écrit encore

$$\mathbb{V}_{as} \left(\widehat{p}_3 - \widehat{p}_3^0 \right) = \mathbb{V}_{as}(\widehat{p}_3) - \mathbb{V}_{as} \left(\widehat{p}_3^0 \right)$$

(b) La loi asymptotique de $(\widehat{p}_3 - \widehat{p}_3^0)$ se déduit de celle de $\begin{pmatrix} \widehat{p}_1 \\ \vdots \\ \widehat{p}_K \end{pmatrix}$ déterminée précédemment.

Soit en effet $g : \begin{pmatrix} \mathbb{R}^K & \rightarrow & \mathbb{R} \\ (x_1, \dots, x_K) & \mapsto & \frac{x_3}{2(1-x_1-x_2)} \end{pmatrix}$; ainsi $g(\widehat{p}_3) = \widehat{p}_3^0$. Alors

$$\frac{\partial g}{\partial x_1}(p_1, \dots, p_K) = \frac{p_3}{2(1-p_1-p_2)^2} = 2p_3$$

$$\frac{\partial g}{\partial x_2}(p_1, \dots, p_K) = \frac{p_3}{2(1-p_1-p_2)^2} = 2p_3$$

$$\frac{\partial g}{\partial x_3}(p_1, \dots, p_K) = 1$$

$$\frac{\partial g}{\partial x_k}(p_1, \dots, p_K) = 0, \quad \forall k \in \llbracket 3, K \rrbracket$$

Donc d'après le théorème de Slutsky

$$\begin{aligned}
\sqrt{n}(\widehat{p}_3^0 - p_3) &\xrightarrow[n \rightarrow +\infty]{\mathcal{L}} \mathcal{N}\left(0, \begin{pmatrix} 2p_3 \\ 2p_3 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}' \begin{pmatrix} p_1(1-p_1) & -p_1p_2 & \cdots & -p_1p_K \\ -p_2p_1 & p_2(1-p_2) & \cdots & -p_2p_K \\ \vdots & \vdots & \ddots & \vdots \\ -p_Kp_1 & -p_Kp_2 & \cdots & p_K(1-p_K) \end{pmatrix} \begin{pmatrix} 2p_3 \\ 2p_3 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}\right) \\
&= \mathcal{N}\left(0, \begin{pmatrix} 2p_3 \\ 2p_3 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}' \times \begin{pmatrix} 2p_3p_1(1-p_1) - 2p_1p_2p_3 - p_1p_3 \\ -2p_1p_2p_3 + 2p_3p_2(1-p_2) - p_2p_3 \\ -2p_3^2p_1 - 2p_3^2p_2 + p_3(1-p_3) \\ 0 \\ \vdots \\ 0 \end{pmatrix}\right) \\
&= \mathcal{N}\left(0, \begin{cases} 4p_3^2p_1(1-p_1) - 4p_1p_2p_3^2 - 2p_1p_3^2 \\ -4p_1p_2p_3^2 + 4p_3^2p_2(1-p_2) - 2p_2p_3^2 \\ -2p_3^2p_1 - 2p_3^2p_2 + p_3(1-p_3) \end{cases}\right) \\
&= \mathcal{N}\left(0, \begin{cases} [4p_1(1-p_1) - 4p_1p_2 - 2p_1 - 4p_1p_2 + \\ 4p_2(1-p_2) - 2p_2 - 2p_1 - 2p_2]p_3^2 \\ -p_3^2 + p_3 \end{cases}\right) \\
&= \mathcal{N}\left(0, \{-4(p_1+p_2)^2p_3^2 - p_3^2 + p_3\}\right) \\
&= \boxed{\mathcal{N}(0, p_3(1-2p_3))} \text{ sous } H_0 : "p_1 + p_2 = \frac{1}{2}"
\end{aligned}$$

Ainsi $\mathbb{V}_{as}(\widehat{p}_3^0) = p_3(1-2p_3)$.

(c) On a

$$\begin{aligned}
\mathbb{V}_{as}(\widehat{p}_3 - \widehat{p}_3^0) &= \mathbb{V}_{as}(\widehat{p}_3) - \mathbb{V}_{as}(\widehat{p}_3^0) \\
&= p_3(1-p_3) - p_3(1-2p_3) \\
&= \boxed{p_3^2}
\end{aligned}$$

Posons donc $\mathbb{V}_{as}(\widehat{p}_3 - \widehat{p}_3^0) = \widehat{p}_3^2 \xrightarrow[n \rightarrow +\infty]{p.s.} \mathbb{V}_{as}(\widehat{p}_3 - \widehat{p}_3^0)$. On a par ailleurs

$$\begin{aligned}
\widehat{p}_3 - \widehat{p}_3^0 &= \frac{n_3}{n} - \frac{n_3}{2(n-n_1-n_2)} \\
&= \frac{n_3}{n} \left(1 - \frac{1}{2\left(1 - \frac{n_1}{n} - \frac{n_2}{n}\right)}\right) \\
&= \widehat{p}_3 \left(1 - \frac{1}{2(1 - \widehat{p}_1 - \widehat{p}_2)}\right) \\
&= \widehat{p}_3 \frac{1 - 2\widehat{p}_1 - 2\widehat{p}_2}{2 - 2\widehat{p}_1 - 2\widehat{p}_2}
\end{aligned}$$

de sorte que

$$\begin{aligned}
\xi_n^H &= \frac{\widehat{p}_3^2 \left(\frac{1-2\widehat{p}_1-2\widehat{p}_2}{2-2\widehat{p}_1-2\widehat{p}_2}\right)^2}{\widehat{p}_3^2} \\
&= \boxed{n \left(\frac{1-2\widehat{p}_1-2\widehat{p}_2}{2-2\widehat{p}_1-2\widehat{p}_2}\right)^2} \\
&\xrightarrow[n \rightarrow +\infty]{\mathcal{L}} \chi_1^2
\end{aligned}$$

(d) Si H_0 est fautive, alors $p_1 + p_2 \neq \frac{1}{2}$ et donc

$$\begin{aligned}
\frac{1}{n} \xi_n^H &= \left(\frac{1-2\widehat{p}_1-2\widehat{p}_2}{2-2\widehat{p}_1-2\widehat{p}_2}\right)^2 \\
&\xrightarrow[n \rightarrow +\infty]{p.s.} \left(\frac{p_1+p_2-\frac{1}{2}}{1-p_1-p_2}\right) \\
&> 0
\end{aligned}$$

de sorte que

$$\xi_n^H \xrightarrow[n \rightarrow +\infty]{\mathcal{L}} (+\infty)$$

Ainsi, si H_0 est fautive alors $\mathbb{P}(\xi_n^H < q_{1-\alpha}^{\chi_1^2}) \xrightarrow[n \rightarrow +\infty]{} \mathbb{P}((+\infty) < q_{1-\alpha}^{\chi_1^2}) = 0$: le test d'Hausman est donc convergent.

(e) On a enfin

$$\begin{aligned}
\frac{\xi_n^H}{\xi_n^W} &= \frac{n \left(\frac{1-2\widehat{p}_1-2\widehat{p}_2}{2-2\widehat{p}_1-2\widehat{p}_2}\right)^2}{n \frac{(\widehat{p}_1+\widehat{p}_2-\frac{1}{2})^2}{(\widehat{p}_1+\widehat{p}_2)(1-\widehat{p}_1-\widehat{p}_2)}} \\
&= \frac{\widehat{p}_1 + \widehat{p}_2}{1 - \widehat{p}_1 - \widehat{p}_2} \\
&\xrightarrow[n \rightarrow +\infty]{p.s.} (1)
\end{aligned}$$

Ainsi, les deux statistiques de test ξ_n^H et ξ_n^W sont asymptotiquement équivalentes.

*
* *

Corrigé de l'exercice 5

- ☞ Q1 (a) Comme $\llbracket 0, N-1 \rrbracket \subsetneq \mathbb{N}$, $f_{\phi,s}$ n'est pas injective, soient $i < j$ les deux premiers entiers tels que $f_{\phi,s}(i) = f_{\phi,s}(j)$, et posons $T = j - i \geq 1$. Alors $f_{\phi,s}(i+T) = f_{\phi,s}(i)$, donc $f_{\phi,s}(i+1+T) = f_{\phi,s}(i+1)$ et par récurrence $\forall k \geq i, f_{\phi,s}(k+T) = f_{\phi,s}(k)$ et donc $f_{\phi,s}$ est périodique au-delà de i .

Remarquons que la suite $(f_{\phi,s}(k))_{k \in \mathbb{N}}$ est *déterministe* et pas aléatoire; il s'agit de déterminer si elle peut être la suite des *réalisations* d'une variable aléatoire uniformément distribuée sur $\llbracket 0, N-1 \rrbracket$.

Définissons $g(k) = \lim_{N \rightarrow +\infty} \frac{1}{N} |\{i \in \mathbb{N}/y(i) = k\}|$ la fréquence empirique de $k \in \llbracket 0, N-1 \rrbracket$; alors $f_{\phi,s}$ "génère" une variable distribuée selon la loi \mathcal{L} sur $\llbracket 0, N-1 \rrbracket$ si la fréquence empirique de tout k est égale à sa probabilité selon \mathcal{L} .

En particulier, $f_{\phi,s}$ est uniformément distribuée ssi $\forall k \in \llbracket 0, N-1 \rrbracket f_{\phi,s}(k) = \frac{1}{N-1}$.

- (b) On a $T \leq N$: en effet pour tout $a \in \mathbb{N}$, $t \mapsto f_{\phi,s}(a+t)$ ne peut pas être injective de $\llbracket 0, N \rrbracket$ vers $\llbracket 0, N-1 \rrbracket$, donc il existe $i < j \in \llbracket 0, N \rrbracket$ tel que $f_{\phi,s}(a+j) = f_{\phi,s}(a+i)$, et par suite $T_{\phi,s} \leq N$.

Or si $T_{\phi,s} < N$, alors $f_{\phi,s}([i, +\infty[) = \{f_{\phi,s}(i), f_{\phi,s}(i+1), \dots, f_{\phi,s}(i+T_{\phi,s}-1)\}$, donc $|f_{\phi,s}([i, +\infty[)| \leq T_{\phi,s} < N$, et donc $f_{\phi,s}([i, +\infty[) \subsetneq \llbracket 0, N-1 \rrbracket$. Ainsi l'un au moins des entiers de $\llbracket 0, N-1 \rrbracket$ n'est plus jamais atteint au-delà de i , et sa fréquence empirique est donc nulle: $f_{\phi,s}$ ne "génère" pas une loi uniforme sur $\llbracket 0, N-1 \rrbracket$.

Remarquons réciproquement que si $f_{\phi,s}$ est périodique de période exactement N , alors elle est uniforme.

Ainsi, un "bon générateur" de nombres pseudo-aléatoires sur $\llbracket 0, N-1 \rrbracket$ est périodique de période N .

- (c) Par définition de la division euclidienne, $f_{\psi_{a,c,p},s}$ est à valeur dans $\llbracket 0, p-1 \rrbracket$ donc pour qu'elle soit un bon générateur sur $\llbracket 0, N-1 \rrbracket$ il est nécessaire que $p = N$.

Par ailleurs dans $\mathbb{Z}/p\mathbb{Z}$ la suite $(a^n)_{n \in \mathbb{N}}$ est périodique et sa période divise p . Donc si $c = 0$ et si p est premier, $f_{\psi_{a,c,p},s}$ est de période 1 ou p , donc p (car $a \geq 2$), et est donc un bon générateur uniforme sur $\llbracket 0, N-1 \rrbracket$ (en fait il suffit que p soit premier avec a , et il n'est pas nécessaire que $c = 0$).

Ces générateurs ont été introduits par Lehmer et sont encore souvent utilisés¹. On sait notamment que $2^{31} - 1 = 2147483647$ est premier (vérifié par Euler en 1750), ainsi que $2^{61} - 1$ (Pervouchine 1883) et $2^{127} - 1$ (Lucas, 1876).²

- ☞ Q2 (a) Il s'agit d'une sorte de critère ergodique: la suite x est suffisamment "équi-répartie" et "équilibrée" pour que sommer aux points chargés par x ou sommer en tout point de $[0, 1]$ soit équivalent.

- (b) Supposons que r est rationnel.

¹cf Knuth, D.E., 1981; The Art of Computer Programming, Volume 2 Seminumerical Algorithms, Addison-Wesley, Reading Mass., 688 pages, ISBN 0-201-03822-6

²Parmi les nombres sous la forme $2^n - 1$ seuls ceux où $n \in \{2, 3, 5, 7, 13, 17, 19, 31, 61, 89, 107, 127\}$ sont premiers, et non pas seulement pour $n \in \{2, 3, 5, 7, 13, 17, 19, 31, 67, 127, 257\}$ comme prétendu incorrectement par Marin Mersenne dans la préface de Cogitata Physica-Mathematica (1644). Parmi les $2^{2^n-1} - 1$ courants en informatique, ceux pour $n \in \{2, 3, 5, 7\}$ sont premiers.

- Supposons tout d'abord que $f_p = x \mapsto e^{2\pi i p x}$, $p \in \mathbb{Z}$. Par définition $rk = \text{frac}(rk) = E(rk)$ et donc

$$\begin{aligned} S_n^x(f_p) &= \frac{1}{n} \sum_{k=1}^n e^{2\pi i p x_k} \\ &= \frac{1}{n} \sum_{k=1}^n e^{(2p\pi i) rk} \text{ car } e^{(2p\pi i)E(rk)} = 1 \\ &= \frac{1}{n} \sum_{k=1}^n (e^{(2p\pi i)r})^k \end{aligned}$$

Comme r est irrationnel, si $p \neq 0$, alors $e^{(2p\pi i)r} \neq 1$ et donc

$$\begin{aligned} S_n^x(f_p) &= \frac{1}{n} \frac{e^{(2p\pi i)r} - (e^{(2p\pi i)r})^{n+1}}{1 - e^{(2p\pi i)r}} \\ &= \frac{1}{n} \frac{e^{(2p\pi i)r} \frac{\sin(\pi nr)}{\sin(\pi r)}}{1 - e^{(2p\pi i)r}} \end{aligned}$$

et donc

$$|S_n^x(f_p)| \leq \frac{1}{n} \frac{1}{|\sin(\pi r)|} \xrightarrow{n \rightarrow \infty} 0$$

soit

$$S_n^x(f_p) \xrightarrow{n \rightarrow \infty} \int_{[0,1]} e^{2\pi i p u} du$$

- Lorsque $p = 0$, on a bien-sûr $S_n^x(f_0) = 1 = \int_{[0,1]} f_0$.
 – Le résultat reste donc vrai par linéarité de l'intégrale pour tout polynôme trigonométrique.
 – Soit enfin f continue sur $[0, 1]$; d'après le théorème de Dirichlet f est limite *uniforme* d'une suite $(P_l)_{l \in \mathbb{N}}$ de polynômes trigonométriques. Donc

$$\begin{aligned} S_n^x(f) &= \frac{1}{n} \sum_{k=1}^n f(x_k) \\ &= \frac{1}{n} \sum_{k=1}^n \lim_{l \rightarrow \infty} P_l(x_k) \\ &= \lim_{k \rightarrow \infty} S_{nb}^x(P_k) \text{ car la somme } \sum_{k=1}^n \text{ est finie} \\ &= \lim_{k \rightarrow \infty} \int_{[0,1]} P_l \text{ d'après le résultat précédent} \\ &= \int_{[0,1]} \lim_{k \rightarrow \infty} P_l \text{ car la convergence de } (P_l)_{l \in \mathbb{N}} \text{ est uniforme} \\ &= \int_{[0,1]} f \end{aligned}$$

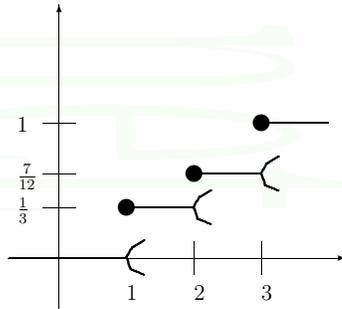
ce qui achève la preuve.

Dans le cas en revanche où $r = \frac{a}{b}$ est rationnel, alors $(rk)_{k \in \mathbb{N}} = \{0, \frac{1}{b}, \dots, \frac{b-1}{b}\}$: la suite x ne prend que b valeurs distinctes et le résultat est faux, par exemple lorsque f est la fonction affine par morceaux telle que

$$\begin{cases} f\left(\frac{k}{b}\right) = 1 \\ f\left(\frac{k+\frac{1}{2}}{b}\right) = 2 \end{cases}$$

car alors $S_n^x(f) = 1 < \frac{3}{2} = \int_{[0,1]} f$.

Q3 (a) Le graphe de $F_{\mathcal{L}}$ est le suivant :



On a donc

$$F_{\mathcal{L}}^{-1}(t) = \begin{cases} -\infty & \text{si } 0 = t \\ 1 & \text{si } 0 < t < \frac{1}{3} \\ 2 & \text{si } \frac{1}{3} \leq t < \frac{7}{12} \\ 3 & \text{si } \frac{7}{12} \leq t \leq 1 \end{cases}$$

Par conséquent lorsque $U \sim \mathcal{U}_{[0,1]}$, $X = F_{\mathcal{L}}^{-1}(U)$ ne charge positivement que $\{1, 2, 3\}$ et on a

$$\begin{aligned} \mathbb{P}(X = 1) &= \mathbb{P}\left(0 \leq U < \frac{1}{3}\right) = \frac{1}{3} \\ \mathbb{P}(X = 2) &= \mathbb{P}\left(\frac{1}{3} \leq U < \frac{7}{12}\right) = \frac{1}{4} \\ \mathbb{P}(X = 3) &= \mathbb{P}\left(\frac{7}{12} \leq U < 1\right) = \frac{5}{12} \end{aligned}$$

de sorte que $X \sim \mathcal{L}$.

(b) On a pour tout $t \in [0, 1]$

$$\begin{aligned} \mathbb{P}(X \leq t) &= \mathbb{P}(F_{\mathcal{L}}^{-1}(U) \leq t) \\ &= \mathbb{P}(U \leq F_{\mathcal{L}}(t)) \quad \text{car } F_{\mathcal{L}} \text{ est croissante} \\ &= F_{\mathcal{L}}(t) \quad \text{car } U \sim \mathcal{U}_{[0,1]} \end{aligned}$$

Par conséquent $X \sim \mathcal{L}$.

Remarquons que $F_{\mathcal{L}}$ étant strictement croissante, c'est sa continuité qui nous assure qu'elle est bijective.

(c)

$$F_{\mathcal{E}(\lambda)}(u) = \int_0^u \lambda e^{-\lambda x} dx = (1 - e^{-\lambda u})$$

Soit donc $U \sim \mathcal{U}_{[0,1]}$; alors $-\frac{1}{\lambda} \ln(1 - U) \sim \mathcal{E}(\lambda)$; or par ailleurs $(1 - U) \sim \mathcal{U}_{[0,1]}$, donc $-\frac{1}{\lambda} \ln(U) \sim \mathcal{E}(\lambda)$.

On génère donc u_n réalisation de $U \sim \mathcal{U}_{[0,1]}$ et on renvoie $x_n = -\frac{1}{\lambda} \ln(u_n)$.

(d) Soit $G : \mathbb{R}^2 \rightarrow \mathbb{R}^m$ absolument continue; alors

$$\begin{aligned} \mathbb{E}(G(R^2, \theta)) &= \int_{\mathbb{R}^2} G\left(x^2 + y^2, \arctan \frac{y}{x}\right) \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{x^2+y^2}{2\sigma^2}} dy dx \\ &= \int_{\mathbb{R}^+ \times [0, 2\pi]} G(\rho^2, \theta) \frac{1}{2\pi\sigma^2} e^{-\frac{\rho^2}{2\sigma^2}} \rho d\rho d\theta \end{aligned}$$

par le changement de variables en coordonnées polaires³. En particulier si G est un produit $G(a, b) = G_1(a)G_2(b)$ alors

$$\mathbb{E}(G_1(R^2)G_2(\theta)) = \left(\frac{1}{\sigma^2} \int_{\mathbb{R}^+} \underbrace{G_1(\rho^2) e^{-\frac{\rho^2}{2\sigma^2}} \rho}_{\text{ne dépend que de } \rho} d\rho \right) \left(\frac{1}{2\pi} \int_{[0, 2\pi]} \underbrace{G_2(\theta)}_{\text{que de } \theta} d\theta \right)$$

Par conséquent, les variables R^2 et θ sont indépendantes.

En outre on a pour toute G_1 absolument continue

$$\begin{aligned} \mathbb{E}(G_1(R^2)) &= \int_{\mathbb{R}^+} \frac{1}{\sigma^2} G_1(\rho^2) e^{-\frac{\rho^2}{2\sigma^2}} \rho d\rho \\ &= \int_{\mathbb{R}^+} \frac{1}{2\sigma^2} G_1(u) e^{-\frac{1}{2\sigma^2}u} du \end{aligned}$$

et donc $R^2 \sim \mathcal{E}(\frac{1}{2\sigma^2})$.

³Soit $f : \mathbb{R}^2 \rightarrow E$ mesurable absolument continue, et soit à calculer $\int_{\mathcal{A} \subset \mathbb{R}^2} f(x, y) dx dy$.
Posons

$$\phi : \begin{pmatrix} \mathbb{R}^* \times \mathbb{R} \\ \begin{pmatrix} x \\ y \end{pmatrix} \end{pmatrix} \mapsto \begin{pmatrix} \mathbb{R}^+ \times \mathbb{R} \\ \begin{cases} x^2 + y^2 & \text{si } x > 0 \\ \arctan\left(\frac{y}{x}\right) & \text{si } x > 0 \\ \arctan\left(\frac{y}{x}\right) + \pi & \text{si } x < 0 \end{cases} \end{pmatrix}$$

ϕ est différentiable et de jacobienne

$$Jac(\phi)_{(x,y)} = \begin{pmatrix} 2x & 2y \\ -\frac{y}{x^2} \frac{1}{1+(\frac{y}{x})^2} & \frac{1}{x} \frac{1}{1+(\frac{y}{x})^2} \end{pmatrix}$$

donc de jacobien

$$|Jac(\phi)_{(x,y)}| = 2x \frac{1}{x} \frac{1}{1+(\frac{y}{x})^2} - 2y \left(-\frac{y}{x^2} \frac{1}{1+(\frac{y}{x})^2} \right) = 2$$

de sorte que

$$\int_{\mathcal{A} \subset \mathbb{R}^2} f(x, y) dx dy = \int_{\phi(\mathcal{A})} f(\phi^{-1}(\rho^2, \theta)) 2 \left(\frac{1}{2}\rho d\rho\right) d\theta$$

Par ailleurs pour toute G_2 absolument continue

$$\mathbb{E}(G_2(\theta)) = \int_{]0, 2\pi[} G_2(\theta) \frac{1}{2\pi} d\theta$$

et donc $\theta \sim \mathcal{U}_{[0, 2\pi]}$

Pour générer une réalisation d'une variable qui suivrait la loi $\mathcal{N}(m, \sigma^2)$, on génère donc $\frac{\theta_n}{2\pi}$ uniformément sur $[0, 1]$ et r_n^2 selon la loi $\mathcal{E}(\frac{1}{\sigma^2})$ et on renvoie le nombre

$$x_n = m + \sqrt{r_n^2} \cos\left(2\pi \left(\frac{\theta_n}{2\pi}\right)\right)$$

(e)

$$F_{\mathcal{W}(\alpha, \beta)}(y) = \int_0^y f_{\mathcal{W}(\alpha, \beta)}(x) dx = 1 - e^{-\alpha y^\beta}$$

et donc on génère u_n réalisation de $U \sim \mathcal{U}_{[0, 1]}$ et on renvoie

$$x_n = \left(-\frac{1}{\alpha} \ln(u_n)\right)^{\frac{1}{\beta}}$$

(f) La densité $f_{\mathcal{P}(\lambda)}$ s'écrit sur \mathbb{R} comme une somme infinie de masses de Dirac $f_{\mathcal{P}(\lambda)}(s) = \sum_{k=0}^{+\infty} \frac{\lambda^k}{k!} e^{-\lambda} \delta_k(s)$ donc un algorithme consiste à générer u uniformément sur $[0, 1]$ et à le placer dans l'un des intervalles $[S_n, S_{n+1}[$ où $S_n = \sum_{k=0}^n \frac{\lambda^k}{k!} e^{-\lambda}$.

Le calcul des $(S_n)_{n \in \mathbb{N}}$ peut néanmoins être dissuasif en pratique; il est possible en fait de s'en passer.

Soient en effet $(Y_n)_{n \in \mathbb{N}} \stackrel{iid}{\sim} \mathcal{E}(\mu)$; alors $\forall n \in \mathbb{N}^*, Y_1 + \dots + Y_n \sim \Gamma(n, \mu)$ (voir TD 1, exercice 2), et donc

$$\mathbb{P}(Y_1 + \dots + Y_n > 1) = \int_{\frac{1}{\mu}}^{+\infty} \frac{t^{n-1}}{(n-1)!} e^{-t} dt$$

Définissons alors $\nu(Y) = \min\{n / Y_1 + \dots + Y_n > 1\} - 1$; alors

$$\begin{aligned} \mathbb{P}(\nu(Y) = n) &= \mathbb{P}(Y_1 + \dots + Y_{n+1} > 1) - \mathbb{P}(Y_1 + \dots + Y_n > 1) \\ &= \int_{\frac{1}{\mu}}^{+\infty} \frac{t^n}{n!} e^{-t} dt - \int_{\frac{1}{\mu}}^{+\infty} \frac{t^{n-1}}{(n-1)!} e^{-t} dt \\ &= \frac{1}{\mu^n n!} e^{-\frac{1}{\mu}} \end{aligned}$$

et donc $\nu(Y) \sim \mathcal{P}\left(\frac{1}{\mu}\right)$.

Ainsi, engendrer une loi de Poisson de paramètre λ revient à engendrer une succession de lois exponentielles d'espérance $\frac{1}{\lambda}$. Or $\mathcal{E}\left(\frac{1}{\lambda}\right) = \mathcal{W}\left(\frac{1}{\lambda}, 1\right)$ donc d'après la question précédente

si $U \sim \mathcal{U}_{[0, 1]}$, alors $-\frac{1}{\lambda} \ln U \sim \mathcal{E}\left(\frac{1}{\lambda}\right)$. Par conséquent posons $U_i = 1 - e^{-\frac{1}{\lambda} Y_i}$; U_i uniforme, et en outre

$$\begin{aligned} Y_1 + \dots + Y_n > 1 &\Leftrightarrow -\frac{1}{\lambda} \sum_{k=1}^n \ln U_k > 1 \\ &\Leftrightarrow U_1 \cdots U_n > e^{-\lambda} \\ &\Leftrightarrow X_n > e^{-\lambda} \end{aligned}$$

où $X_n = U_1 \cdots U_n$

Définissons donc finalement $N(X) = \min\{n / X_n < e^{-\lambda}\} - 1$; alors $N(X) = \nu(X)$, donc $N(X) \sim \mathcal{P}(\lambda)$.

Pour générer une loi $\mathcal{P}(\lambda)$, il suffit donc de générer une suite u_1, \dots, u_n, \dots uniformes $[0, 1]$, et de renvoyer le premier entier $n - 1$ tel que $u_1 \cdots u_n > e^{-\lambda}$.

Un tel algorithme s'écrit :⁴

- 1 $n \leftarrow 0$ et $x \leftarrow 1$
- 2 Tant que $x > e^{-\lambda}$
- 3 Générer u_n uniformément sur $[0, 1]$
- 4 $x \leftarrow x \times u_n$ et $n \leftarrow n + 1$
- 5 Renvoyer n

*
* *

⁴Pour un exposé plus complet, voir D. Knuth, *The Art of Computer Programming*, vol. 2, p 132.