MIPS:
Probabilités

Pr. Lacroix

SEATECH
Université de Toulon

Promo 2018

- Probabilité = **mesure positive de masse totale égale à** 1 ;
- on la note $\boxed{\mathbb{P}}$ plutôt que μ ;
- $\int_\Omega f d\mathbb{P}$ est notée $\boxed{\mathbb{E}(f)}$, et appelée l'**espérance** de f ;
- on parle souvent d'une **loi** de probabilité au lieu de probabilité tout court.

Exemple

- $\mathbb{P}(\{i\}) = \frac{1}{6}$, $i = 1 \dots 6$; le dé équilibré à six faces ;
- pour $\lambda > 0$, $\mathbb{P}(\{n\}) = e^{-\lambda}\frac{\lambda^n}{n!}$, loi de Poisson de paramètre λ ;
- $\mathbb{P}(A) = \lambda_1(A \cap [0,1])$, $A \in \mathcal{B}_1$, loi dite *uniforme* sur $[0,1]$.

Cas discret : si $\Omega = \{\omega_1, \ldots, \omega_n, \ldots\}$, alors on assigne $\mathbb{P}(\{\omega_i\}) \geq 0$ de sorte que

$$\boxed{\sum_i \mathbb{P}(\{\omega_i\}) = 1}.$$

Les exemples les plus connus sont les lois de Poisson, les lois exponentielles ($0 < q < 1$ et $\boxed{\mathbb{P}(\{i\}) = (1-q)q^i}$) etc.

Cas absolument continu : on se trouve dans le cas où $\mathbb{P} << \lambda_n$, et on appelle alors **densité de la loi de probabilité** $\mathbb{P}$ sa **dérivée de Radon-Nikodym** relativement à λ_n.

- loi uniforme sur $[a, b]$: $\boxed{\frac{d\mathbb{P}}{d\lambda_1} = \frac{1}{b-a} 1_{[a,b]}}$;
- loi normale centrée réduite : $\boxed{\frac{d\mathbb{P}}{d\lambda_1}(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}}$;
- loi exponentielle positive de paramètre $p > 0$: $\boxed{\frac{d\mathbb{P}}{d\lambda_1}(x) = pe^{-px} 1_{[0,+\infty[}(x)}$;
- il existe des lois Gaussiennes sur $\mathbb{R}^n$, les lois de Student, du Chi-deux (χ^2) ou de Fisher utilisées en statistiques....

Pour construire d'autres lois, on peut aussi

- considérer la loi image d'une loi de probabilité : c'est une probabilité ;
- considérer le produit de lois de probabilité : idem ;
- prendre des combinaisons barycentriques de lois $\mathbb{P}_1, \ldots, \mathbb{P}_n$ sur un même espace mesuré $(\Omega, \mathcal{B})$: on considère $\alpha_i \geq 0$ avec $\sum_{i=1}^{n} \alpha_i = 1$, alors $\mathbb{P} := \sum_i \alpha_i \mathbb{P}_i$ est une probabilité

Caractériser les lois permet de les reconnaître à l'aide d'objets associés qui sont parfois plus simples à manipuler :

Théorème (Fonction de répartition)

La fonction de répartition d'une loi de probabilité $\mathbb{P}$ sur $(\mathbb{R}^n, \mathcal{B}_n)$ est l'application F définie sur $\mathbb{R}^n$ par

$$F(t_1, \dots, t_n) := \mathbb{P}(\prod_{i=1}^{n}]-\infty, t_i]).$$

*Alors **deux lois de fonctions de répartitions identiques sont nécessairement identiques**.*
Lorsque $n = 1$, toute fonction $F : \mathbb{R} \to [0, 1]$ croissante, continue à droite, ayant une limite nulle en $-\infty$ et 1 en $+\infty$, est la fonction de répartition d'une loi, et réciproquement.

Exemple (Lois classiques)

- Calculer les FR des lois de Bernoulli de paramètre p, binomiale de paramètres n et p, de Poisson de paramètre λ, géométrique de paramètres q.
- Idem pour les lois uniforme sur $[0, 1]$, géométrique positive de paramètre λ.
- Trouver les valeurs de la FR d'une loi normale centrée réduite dans excel.

Proposition

Soit $\mathbb{P}$ une loi de probabilité sur $(\mathbb{R}, \mathcal{B}_1)$ et F sa FR.

- *Si F a un saut de discontinuité de valeur $F(t_0^+) - F(t_0^-)$ en t_0, alors $\mathbb{P}$ charge t_0 avec la masse $F(t_0^+) - F(t_0^-)$;*
- *Si F a une dérivée à droite F'^+ sauf peut être en un ensemble au plus dénombrable de points D où elle présente des sauts de discontinuité, alors $\mathbb{P}$ est somme de la mesure positive discrète chargeant les points de D avec les sauts respectifs de F en ces points, et de la mesure positive absolument continue de densité F'^+ relativement à λ_1.*

Théorème (Fonction caractéristique)

La fonction caractéristique d'une loi de probabilité $\mathbb{P}$ *sur* $(\mathbb{R}^n, \mathcal{B}_n)$ *est l'application* Φ *définie sur* $\mathbb{R}^n$ *par*

$$\Phi(t_1, \ldots, t_n) := \int_{\mathbb{R}^n} e^{i(t_1 x_1 + \ldots + t_n x_n)} d\mathbb{P}(x_1, \ldots, x_n).$$

Alors deux lois de fonctions caractéristiques identiques sont nécessairement identiques.

La fonction caractéristique d'une loi est uniformément continue, a une limite nulle en l'infini (Lemme de Riemann-Lebesgue), et vaut 1 à l'origine.

Exemple (Lois classiques)

- Calculer les FC des lois de Bernoulli de paramètre p, binomiale de paramètres n et p, de Poisson de paramètre λ, géométrique de paramètres q.
- Idem pour les lois uniforme sur $[0, 1]$, géométrique positive de paramètre λ.

Proposition

Soit Φ la FC d'une loi $\mathbb{P}$ sur $(\mathbb{R}, \mathcal{B}_1)$. Soit $n \geq 1$ et supposons que $\mathbb{E}(|x|^n) < +\infty$. Alors Φ est de classe $\mathcal{C}^n$ et

$$\boxed{\Phi^{(k)}(0) = i^k \mathbb{E}(x^k),\ 0 \leq k \leq n}.$$

On appelle $\mathbb{E}(x^k)$ le moment d'ordre k de la loi $\mathbb{P}$.

Exemple (Moments des lois classiques)

- Calculer les moments d'ordre 1 et 2 des lois de Bernoulli de paramètre p, binomiale de paramètres n et p, de Poisson de paramètre λ, géométrique de paramètres q.
- Idem pour les lois uniforme sur $[0, 1]$, géométrique positive de paramètre λ.

Théorème (Loi produit)

Si $(\mathbb{R}^{n_i}, \mathcal{B}_{n_i}, \mathbb{P}_i)$, $i = 1, 2$, *sont deux espaces probabilisés, de fonctions de répartition* F_i *et de fonctions caractéristiques* Φ_i, *alors si* F *(reps.* Φ*) désigne la fonction de répartition de la loi produit* $\mathbb{P}_1 \otimes \mathbb{P}_2$ *(reps. sa fonction caractéristique),*

$$\boxed{F(t,s) = F_1(t)F_2(s)} \text{ et } \boxed{\Phi(t,s) = \Phi(t)\Phi(s)}, \quad t \in \mathbb{R}^{n_1}, \ s \in \mathbb{R}^{n_2}.$$

On reconnaît donc qu'une loi est le produit de deux autres s'il en est ainsi de sa fonction de répartition ou caractéristique.

Supposons que $\mathbb{P}$ soit une loi de probabilités sur $(\mathbb{R}^d, \mathcal{B}_d)$. Notons $\Phi_{\mathbb{P}}$ sa fonction caractéristique.

Théorème (Inversion de Fourier)

Si $\Phi_{\mathbb{P}}$ est λ_d intégrable, alors $\mathbb{P}$ a une densité f relativement à λ_d donnée par

$$f(x_1, \ldots, x_d) = \frac{1}{(2\pi)^d} \int_{\mathbb{R}^d} e^{-i(t_1 x_1 + \ldots + t_d x_d)} \Phi_{\mathbb{P}}(t_1, \ldots, t_d) d\lambda_d(t_1, \ldots, t_d), \quad \lambda_d - p.p..$$

Définition

Une **variable aléatoire** *est une application* $(\mathcal{T}, \mathcal{B}_n)$*-mesurable* $X : \Omega \to \mathbb{R}^n$*, où* $(\Omega, \mathcal{T}, \mathbb{P})$ *est un espace probabilisé. Lorsque* $n = 1$*, on parle de variable aléatoire réelle (VAR), sinon de vecteur aléatoire.*

La loi de la variable aléatoire est notée $\mathbb{P}_X$*, et c'est bien la loi image de* $\mathbb{P}$ *par* X*.*

Sa fonction de répartition, notée F_X*, est celle de sa loi.*

Sa fonction caractéristique, notée Φ_X*, aussi.*

En probabilités et encore plus en statistiques, les variables aléatoires modélisent des caractères aléatoires, que l'on connaît au travers de leurs lois, ou dont on cherche à déterminer les lois.

Définition

On dit qu'une suite finie ou dénombrable de VAs $(X_i : \Omega \to \mathbb{R}^{n_i})$ *est* **deux à deux indépendante** *ssi quels que soient* $i \neq j$,

$$\boxed{\mathbb{P}_{(X_i, X_j)} = \mathbb{P}_{X_i} \otimes \mathbb{P}_{X_j}}.$$

On dit d'une telle suite qu'elle est **indépendante** *ssi quel que soit* $J \subset I$ *fini,*

$$\boxed{\mathbb{P}_{(X_i)_{i \in J}} = \bigotimes_{i \in J} \mathbb{P}_{X_i}}.$$

L'indépendance deux à deux est moins forte que l'indépendance tout court.

On dit qu'une suite d'évènements $(A_n)_{n\geq 1}$ d'une tribu $\mathcal{B}$ d'une espace probabilisé $(\Omega, \mathcal{B}, \mathbb{P})$ est indépendante (resp. deux à deux indépendante) ssi la suite de var $(\mathbf{1}_{A_n})_{n\geq 1}$ l'est.
Cette définition coïncide avec la définition habituelle de l'indépendance d'évènements.

On appelle limite supérieure (reps. limite inférieure) d'une suite d'évènements $(A_n)_{n\geq 1}$ d'un espace probabilisé $(\Omega, \mathcal{B}, \mathbb{P})$ l'ensemble noté $\limsup A_n$ (reps. $\liminf A_n$) et défini par

$$\limsup A_n = \cap_n \cup_{m\geq n} A_m \text{ (resp. } \liminf A_n = \cup_n \cap_{m\geq n} A_m).$$

Lemme (BC 1)

Si $\sum_n \mathbb{P}(A_n) < +\infty$, *alors* $\mathbb{P}(\limsup A_n) = 0$.

Lemme (BC 2)

Si $(A_n)_{n\geq 1}$ *est indépendante, et si* $\sum_n \mathbb{P}(A_n) = +\infty$, *alors* $\mathbb{P}(\limsup A_n) = 1$.

Soit $(X_n : \Omega \to \mathbb{R}^d)_{n\geq 1}$ une suite de va indépendante. On appelle **tribu queue** la tribu

$$\mathcal{T} = \cap_{n\geq 1}\sigma(\cup_{k\geq n}X_k^{-1}(\mathcal{B}_d)).$$

Théorème (Loi $0-1$ de Kolmogorov)

Quel que soit $A \in \mathcal{T}$, $\mathbb{P}(A) \in \{0, 1\}$.

Théorème (Existence de VAs de lois données - Kolmogorov)

Étant donnée une suite finie ou infinie dénombrable de lois $(\mathbb{P}_i)_{i\in I}$ *sur* $(\mathbb{R}^{n_i}, \mathcal{B}_{n_i})$, *il existe* $(\Omega, \mathcal{B}, \mathbb{P})$ *et une suite de VAs* $(X_i : \Omega \to \mathbb{R}^{n_i})_{i\in I}$ *indépendantes telles que*

$$\mathbb{P}_{X_i} = \mathbb{P}_i, \; i \in I.$$

En particulier il existe des suites **indépendantes identiquement distribuées** (iid), une telle suite étant une suite indépendante de VAs dont on suppose en plus que toutes sont de même loi, appelée la **loi commune**.
On notera $X \coprod Y$ pour signifier que X et Y sont deux VAs indépendantes.

Proposition

Soit X une VAR. On dit qu'elle est dans $\mathcal{L}^p$ si $\mathbb{E}(|X|^p) < +\infty$. Si $1 \leq p \leq q$ alors $\mathcal{L}^q \subset \mathcal{L}^p$.

Le **moment d'ordre** p **d'une VAR** X *est* $\boxed{\mathbb{E}(X^p)}$ *si il existe.*

Si $X \in \mathcal{L}^2$, alors sa **variance**, *notée $Var(X)$, est définie par*

$$\boxed{Var(X) = \mathbb{E}(X^2) - \mathbb{E}(X)^2 = \mathbb{E}((X - \mathbb{E}(X))^2)}.$$

Son **écart type** *noté $\sigma(X)$ est défini par* $\boxed{\sigma(X) = \sqrt{Var(X)}}$.

Variance et écart type sont des paramètres de dispersion :
$Var(a + X) = Var(X), \ Var(aX) = a^2 Var(X)$.

Si $X \coprod Y$ et si X, Y et $XY \in \mathcal{L}^1$, alors $\mathbb{E}(XY) = \mathbb{E}(X)\mathbb{E}(Y)$. En particulier, $Var(X + Y) = Var(X) + Var(Y)$ si $X \coprod Y$.

On appelle **covariance** *de deux VARs $X, Y \in \mathcal{L}^2$ la quantité notée $Cov(X, Y)$ et définie par*

$$\boxed{Cov(X, Y) = \mathbb{E}((X - \mathbb{E}(X))(Y - \mathbb{E}(Y))) = \mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y)}.$$

Définition

Si $X = (X_1, \ldots, X_n)$ est un n-uplet de VARs toutes dans $\mathcal{L}^2$, alors sa **matrice des covariances** *est définie par*

$$\boxed{Cov(X) = (Cov(X_i, X_j))_{1 \leq i,j \leq n}}.$$

Ses éléments diagonaux sont les $Var(X_i)$.

Si les X_i sont indépendantes alors $Cov(X) = Diag(Var(X_i))$. Mais la réciproque est fausse.

D'après la partie mesure et intégration, on peut calculer $\mathbb{E}(g(X))$ pour une VA X sans connaître explicitement l'application X, car il suffit pour cela de connaître $\mathbb{P}_X$:

Proposition (Théorème de transfert)

Soit $X : \Omega \to \mathbb{R}^n$ un VA et $g : \mathbb{R}^n \to \mathbb{R}$ mesurable. Alors $g(X)$ est $\mathbb{P}$-intégrable ssi g est $\mathbb{P}_X$-intégrable, et si tel est le cas, alors

$$\boxed{\mathbb{E}(g(X)) = \int_{\mathbb{R}^n} g(x_1, \dots, x_n) d\mathbb{P}_X(x_1, \dots, x_n)}.$$

Exemple (Calculs de variance)

Soit X une VAR.

- Calculer son espérance et sa variance si sa loi est Bernoulli de paramètre p, binomiale de paramètres n et p, de Poisson de paramètre λ, géométrique de paramètres q.
- Idem pour une loi uniforme sur $[0, 1]$, géométrique positive de paramètre λ.

Proposition (Indépendance et paquets, fonctions, espérance)

- *Soit* $(X_n)_{n\geq 1}$ *des VA indépendantes, et soit* $(I_p)_{p\geq 1}$ *une partition de* $\mathbb{N}^*$*. Alors la famille de VAs* $((X_n)_{n\in I_p})_{p\geq 1}$ *est indépendante.*
- *Si les* I_p *sont finis et pour chaque* p, $f_p : \mathbb{R}^{\mathrm{Card}\, I_p} \to \mathbb{R}^{q_p}$ *est mesurable, alors la famille* $((f_p(X_n))_{n\in I_p})_{p\geq 1}$ *est indépendante.*
- *Si* $(Y_1, \ldots, Y_n)$ *sont des VAs indépendantes sur* $\mathbb{R}^{q_k}$ *respectivement, et pour chaque* k, $f_k : \mathbb{R}^{q_k} \to \mathbb{R}$ *est mesurable, alors si* $\prod_k f_k(Y_k)$ *et les* $f_k(Y_k)$ *sont intégrables,*

$$\boxed{\mathbb{E}(\prod_k f_k(Y_k)) = \prod_k \mathbb{E}(f_k(Y_k))}.$$

- *En particulier, si* $(Y_1, \ldots, Y_n)$ *sont des VAR indépendantes,* $\boxed{\Phi_{Y_1+\ldots+Y_n} = \Phi_{Y_1}\ldots\Phi_{Y_n}}$.

Exemple (Calculs de fonctions caractéristiques)

Soient $X_1, \ldots, X_n$ des VARs iid. Calculer la FC de $X_1 + \ldots + X_n$ si la loi commune est

- Bernoulli de paramètre p,
- binomiale de paramètres n et p,
- Poisson de paramètre λ,
- géométrique de paramètres q,
- une loi uniforme sur $[0, 1]$,
- géométrique positive de paramètre λ,

En déduire dans chacun de ces cas la loi de la somme.

Proposition (Indépendance et densité)

Soient $X_1, \ldots, X_n$ des VARs indépendantes de lois de densités respectives f_k relativement à λ_1. Alors la loi de $(X_1, \ldots, X_n)$ a une densité relativement à λ_n et

$$\boxed{\frac{d\mathbb{P}_{(X_1,\ldots,X_n)}}{d\lambda_n}(x_1, \ldots, x_n) = f_1(x_1) \ldots f_n(x_n)}.$$

Exemple (Calculs d'espérances)

Soient $X_1, \ldots, X_n$ des VARs iid de loi commune la loi uniforme sur $[0, 1]$. Calculer $\mathbb{E}(\min_{1 \le i \le n} X_i)$.

Définition

La convolée de deux mesures positives μ et ν sur $(\mathbb{R}, \mathcal{B}_1)$ est la mesure image de $\mu \otimes \nu$ par $s : (x, y) \mapsto x + y$: on la note $\mu \star \nu$. Soient f et g deux applications mesurables sur $\mathbb{R}$ et intégrables relativement à λ_1. Alors la convolée de f par g, notée $f \star g$, est définie par

$$f \star g(x) = \int_{\mathbb{R}} f(y)g(x-y)d\lambda_1(y).$$

La convolution de fonctions ainsi définie est associative et commutative. Par ailleurs

$$\int_{\mathbb{R}} f \star g(x)dx = \left(\int_{\mathbb{R}} f(u)du\right)\left(\int_{\mathbb{R}} g(v)dv\right).$$

Proposition (Indépendance, densité, somme et convolution)

Si X et Y sont deux VARs de lois de densités respectives f_X et f_Y, alors la VAR $X+Y$ a une densité égale à $f_X \star f_Y$:

$$\boxed{\frac{d\mathbb{P}_{X+Y}}{d\lambda_1}(x) = \frac{d\mathbb{P}_X}{d\lambda_1} \star \frac{d\mathbb{P}_Y}{d\lambda_1}(x)}.$$

Définition

Une suite $(X_n : \Omega \to \mathbb{R}^p)_{n \geq 1}$ *de VAs est dite* **converger presque sûrement vers la VA** $X : \Omega \to \mathbb{R}^p$ *ssi* $\mathbb{P}(\{X_n \nrightarrow X\}) = 0$. *On note* $\boxed{X_n \underset{p.s.}{\longrightarrow} X}$.

On dit que (X_n) **converge en probabilité** *vers* X *ssi quel que soit* $\varepsilon > 0$, $\mathbb{P}(\| X_n - X \| > \varepsilon) \to 0$. *On note* $\boxed{X_n \underset{\mathbb{P}}{\to} X}$.

On dit que (X_n) **converge faiblement** *vers* X *ssi quelle que soit* $f : \mathbb{R}^n \to \mathbb{R}$ *continue bornée,* $\mathbb{E}(f(X_n)) \to \mathbb{E}(f(X))$. *On note* $\boxed{X_n \underset{w^\star}{\longrightarrow} X}$.

Si $p = 1$, $q \geq 1$ *et* $X_n, X \in \mathcal{L}^q$, *on dit que* (X_n) **converge dans** $\mathcal{L}^q$ *ssi* $\mathbb{E}(|X_n - X|^q) \to 0$. *On note* $\boxed{X_N \underset{\mathcal{L}^q}{\longrightarrow} X}$.

Proposition

Les deux conditions suivantes sont équivalentes :

- $X_n \underset{w^\star}{\longrightarrow} X$;
- $F_{X_n} \to F_X$ *simplement en tout point de continuité de* F_X.

On suppose $p = 1$ si nécessaire :

$$
\begin{tikzcd}
X_n \xrightarrow[p.s.]{} X \arrow[dr] & & X_n \xrightarrow[\mathcal{L}^q]{} X \arrow[ll, "\text{le long d'une sous-suite}"'] \arrow[dl] \\
& X_n \xrightarrow[\mathbb{P}]{} X \arrow[d] & \\
& X_n \xrightarrow[w^\star]{} X &
\end{tikzcd}
$$

Théorème (Loi faible des grands nombres)

Si (X_n) est iid de loi commune dans $\mathcal{L}^1$, alors

$$\boxed{\bar{X}_n := \frac{1}{n}\sum_{i=1}^{n} X_i \underset{\mathbb{P}}{\rightarrow} \mathbb{E}(X_1)}.$$

$$\Uparrow$$

Théorème (Loi forte des grands nombres)

Si (X_n) est iid de loi commune dans $\mathcal{L}^1$, alors

$$\boxed{\bar{X}_n \xrightarrow[p.s.]{} \mathbb{E}(X_1)}.$$

Exemple

- La fonction *rand* appelle en informatique le résultat d'un tirage aléatoire d'une valeur dans $[0, 1]$ selon la loi uniforme. Dans excel, taper "*rand*" dans une case, puis tirer la colonne. Enfin, calculer la moyenne arithmétique des valeurs obtenues dans la colonne. Conclusion ?
- Si $f : [0, 1] \to \mathbb{R}$ est Riemann intégrable, et si X est une VAR de loi $\mathcal{U}([0, 1])$, calculer $\mathbb{E}(f(U))$. Dans le tableau excel précédent, taper à droite de chaque cellule "*rand*" sa valeur par f, et faire la moyenne arithmétique en bas de la seconde colonne. Conclusion ?

Théorème (Théorème de Continuité de Lévy)

Soit (X_n) une suite de VA, et X un VA. Alors

$$\boxed{X_n \underset{w^\star}{\longrightarrow} X \iff \Phi_{X_n} \to \Phi_X \textit{ simplement}}.$$

Exemple

- Montrer que si X_n est une VAR de loi $\mathcal{B}(n, p_n)$ et que $np_n \to \lambda$ avec $\lambda > 0$ alors $X_n \underset{w^\star}{\longrightarrow} \mathcal{P}(\lambda)$. C'est pour cette raison que l'on appelle $\mathcal{P}(\lambda)$ la loi des évènements rares.
- Si Y_n suit une loi géométrique de paramètre p_n et si $p_n \to 0$ et $p_n/a_n \to \lambda > 0$, alors $a_n Y_n \underset{w^\star}{\longrightarrow} \mathcal{E}(\lambda)$.

Définition

On appelle **loi normale ou Gaussienne** *de paramètres* $\mu \in \mathbb{R}$ *et* $\sigma^2 > 0$*, la loi absolument continue par rapport à* λ_1*, et de densité*

$$\boxed{f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\dfrac{(x-\mu)^2}{2\sigma^2}}}.$$

On y réfère par la notation $\boxed{\mathcal{N}(\mu, \sigma^2)}$. *Si* $\mu = 0$ *et* $\sigma^2 = 1$*, on parle de loi normale centrée rénduite, ou loi de Gauss.*

Théorème (Théorème Centrale Limite)

Si $(X_n)_{n \geq 0}$ *est une suite de VAR iid dans* $\mathcal{L}^2$*, telle que* $Var(X_1) > 0$*, alors*

$$\boxed{\frac{X_1 + \ldots + X_n - n\mathbb{E}(X_1)}{\sigma(X_1 + \ldots + X_n)} = \frac{\bar{X}_n - \mathbb{E}(\bar{X}_n)}{\sigma(\bar{X}_n)} \xrightarrow[w^\star]{} \mathcal{N}(0,1)}.$$

Exemple

- Notons F la FR de $\mathcal{N}(0,1)$. Montrer que $F(-t) = 1 - F(t)$.
- Rechercher la fonction donnant les valeurs de F dans excel.
- Expliquer pourquoi le TCL entraîne que si (X_n) satisfait ses hypothèses et si $a < b$ alors

$$\mathbb{P}(a \leq \frac{\bar{X}_n - \mathbb{E}(\bar{X}_n)}{\sigma(\bar{X}_n)} \leq b) \to F(b) - F(a).$$

- En supposant que n soit suffisamment grand pour que l'on puisse supposer dans la convergence ci-dessus qu'il y ait égalité, en déduire un intervalle I tel que

$$\mathbb{P}(\bar{X}_n \in I) = F(b) - F(a).$$

- Trouver b et $a = -b$ tels que $F(b) - F(a) = 95\%, = 99\%$.

Théorème (Espérance conditionnelle)

Soit $(\Omega, \mathcal{B}, \mathbb{P})$ un espace probabilisé, et soit $\mathcal{C} \subset \mathcal{B}$ une sous-tribu. Soit X une VAR intégrable définie sur Ω. Alors il existe une unique VAR Y, intégrable, telle que

- *Y est $(\mathcal{C}, \mathcal{B}_\infty)$-mesurable ;*
- *$\forall C \in \mathcal{C}, \int_C X d\mathbb{P} = \int_C Y d\mathbb{P}$.*

*On appelle Y l'**espérance conditionnelle de X sachant $\mathcal{C}$**. On la note $\mathbb{E}(X|\mathcal{C})$. C'est une variable aléatoire.*

Exemple

Si $(\Omega_i)_{i\in I}$ est une partition mesurable finie ou dénombrable de ω telle que $\mathbb{P}(\Omega_i) > 0$, et si $\mathcal{C} = \sigma(\{\Omega_i : i \in I\})$, alors

$$\mathbb{E}(X|\mathcal{C}) = \sum_{i\in I} \frac{\int_{\Omega_i} X d\mathbb{P}}{\mathbb{P}(\Omega_i)} \mathbf{1}_{\Omega_i}.$$

Définition

Si $A \in \mathcal{B}$, on appelle **probabilité de A sachant $\mathcal{C}$**, *notée* $\mathbb{P}(A|\mathcal{C})$, *la quantité définie par*

$$\mathbb{P}(A|\mathcal{C}) = \mathbb{E}(\mathbf{1}_A|\mathcal{C}).$$

Théorème (Loi conditionnelle et espérance conditionnelle)

Soit X une VAR. Alors il existe $\mu : \mathcal{B}_1 \times \Omega \to [0,1]$ telle que

- *pour chaque $\omega \in \Omega$, $\mu(\cdot, \omega)$ est une probabilité sur $(\mathbb{R}, \mathcal{B}_1)$;*
- *pour chaque $A \in \mathcal{B}_1$, $\mu(A, \cdot) = \mathbb{E}(\mathbf{1}_A(X)|\mathcal{C})(\cdot)$.*

On appelle $\mu(\cdot, \omega)$ la loi conditionnelle de X sachant $\mathcal{C}$, parfois notée $\mathbb{P}_X(\cdot|\mathcal{C})(\omega)$.

Alors si ϕ est mesurable et si $\phi(X)$ est intégrable,

$$\mathbb{E}(\phi(X)|\mathcal{C})(\omega) = \int_{\mathbb{R}} \phi(x) d\mathbb{P}_X(x|\mathcal{C})(\omega).$$

Proposition (Conditionnement et densité)

Soient X et Y deux VAR. On appelle **espérance conditionnelle de X sachant Y**, *et on note* $\mathbb{E}(X|Y)$, *l'espérance conditionnelle de X relativement à la tribu $Y^{-1}(\mathcal{B}_1)$.*

On note $\mathbb{P}_X(\cdot|Y)(\omega)$ *la loi conditionnelle de X sachant Y en ω, définie par $\mathbb{P}_X(\cdot|Y)(\omega) = \mathbb{P}_X(\cdot|Y^{-1}(\mathcal{B}_1)(\omega)$.*

Supposons que le couple (X, Y) ait une loi de densité $f(x, y)$ relativement à λ_2. Alors chacune des variables X et Y a une densité f_X et f_Y, données par

$$f_X(x) = \int_{\mathbb{R}} f(x, y)dy, \; f_Y(y) = \int_{\mathbb{R}} f(x, y)dx.$$

Par ailleurs, dans ce cas, la loi de X sachant Y admet une densité sur $\{f_Y(y) > 0\}$, donnée par

$$\frac{d\mathbb{P}_X(\cdot|Y)(\omega)}{d\lambda_1}(x) = \frac{f(x, Y(\omega))}{f_Y(Y(\omega))}.$$

Proposition (Principe de l'inversion)

- *Si X est une VAR de FR F_X, alors l'inverse généralisé de F_X, noté F_X^{-1}, est défini par*

$$F_X^{-1}(y) = \inf\{x \in \mathbb{R} : F(x) \geq y\}.$$

- $F_X^{-1}(U)$ *suit la même loi que X si U suit une loi $\mathcal{U}([0,1])$.*
- *Si F_X est continue, alors $F_X(X)$ suit une loi $\mathcal{U}([0,1])$.*

Proposition

Soit X un VA prenant ses valeurs dans $\{v_n : n \in I\} \subset R^n$ fini ou dénombrable, I étant un intervalle de $\mathbb{N}$. Notons $p_i = \mathbb{P}(X = v_i)$, $i \in I$. Soit U une VAR de loi uniforme sur $[0,1]$. Alors le VA

$$Y = \sum_{i \in I} v_i \mathbf{1}_{[\sum_{j<i} p_j, \sum_{j \leq i} p_j[}(U)$$

est tel que $\boxed{\mathbb{P}_X = \mathbb{P}_Y}$.

Simulation à l'aide d'une suite iid :

Proposition

Soit $(U_i)_{i\geq 0}$ *une suite iid de VARs de loi commune* $\mathcal{U}[0,1]$. *Soit* $0 < p < 1$. *Alors*

- $\mathbf{1}_{U\leq p}$ *suit une loi* $\mathcal{B}(p)$ *;*
- $\sum_{i=1}^{n} \mathbf{1}_{U_i\leq p}$ *suit une loi* $\mathcal{B}(n,p)$ *;*
- $\min\{i \geq 0 : U_i \geq p\}$ *suit une loi géométrique de paramètre* p *;*
- $\min\{n : U_0 \cdots U_n < e^{-\lambda}\}$ *suit une loi* $\mathcal{P}(\lambda)$.

Proposition

- *Si* U *suit la loi* $\mathcal{U}([0,1])$, *alors* $X = a + (b-a)U$ *suit la loi* $\mathcal{U}([a,b])$.
- *Soit* $p : \mathbb{R} \to \mathbb{R}^+$ *la densité d'une loi sur* $\mathbb{R}$ *telle que* $p > 0$. *Soit* F *la FR associée. Alors* $F^{-1} :]0,1[\to \mathbb{R}$ *est bien définie et* $F^{-1}(U)$ *suit une loi de densité* p.

Exemple

- Si X suit une loi de Cauchy de paramètre $a > 0$ de densité $p(x) = \frac{a}{\pi(x^2+a^2)}$, alors $a \tan(\pi U)$ a la même loi que X.
- $-\frac{1}{\lambda}\ln(1-U)$ suit une loi $\mathcal{E}(\lambda)$.

Proposition (Méthode polaire pour la loi normale centrée réduite)

Soient R de loi $\mathcal{E}(\frac{1}{2})$ et θ de loi $\mathcal{U}([0, 2\pi])$. Alors

$$X = \sqrt{R}\cos\theta \text{ et } Y = \sqrt{R}\sin\theta$$

sont indépendantes de lois $\mathcal{N}(0, 1)$.

On revient à la simulation : un petit résultat préliminaire :

Proposition

Soit $d \geq 1$, $D \subset D' \in \mathcal{B}_d$ *deux boréliens tels que* $0 < \lambda_d(D) \leq \lambda_d(D') < +\infty$. *Soit* Y *un VA de loi uniforme sur* D' ($\mathbb{P}(Y \in C) := \lambda_d(C \cap D')/\lambda_d(D')$). *Alors la loi de* Y *sachant que* $Y \in D$ *est la loi uniforme sur* D.

Proposition (Méthode du rejet pour la simulation)

Soit X une VAR de densité f.

- *Si f est continue à support compact $[a, b]$ bornée par M, si $((X_i, Y_i))_{i \geq 0}$ est une suite iid de loi uniforme sur $[a, b] \times [0, M]$, et soit $N = \inf\{i : f(X_i) \geq Y_i\}$. Alors p.s., $N < +\infty$, X_N et N sont indépendantes et suivent respectivement la loi de X et la loi géométrique de paramètre $1/M(b-a)$.*
- *Supposons qu'il existe une densité g telle que $f \leq cg$ pour une constante $c \geq 1$. Soit alors $((W_i, U_i))_{i \geq 0}$ une suite iid de loi commune la loi produit $\mathbb{P}_{\{g\}} \otimes \mathcal{U}([0,1])$ où $\mathbb{P}_{\{g\}}$ désigne la loi de densité g. On note $\tilde{N} = \inf\{i \geq 0 : f(W_i) \geq cU_i g(W_i)\}$. Alors $W_{\tilde{N}}$ a une loi de densité f.*

Exemple

- Prendre pour f la densité de $\mathcal{N}(0,1)$ et $g(x) = \frac{1}{4}e^{-|x|}$, $c = 2\sqrt{\frac{2e}{/}\pi}$.