Variables Aléatoires à Valeurs dans $\mathbb{R}$

Le concept de *variable aléatoire* introduite comme une application X de l'espace discret des événements dans $\mathbb{R}$ se généralise pour un espace probabilisé $(\Omega, \mathcal{A}, Pr)$ quelconque. Il faut toutefois qu'on puisse définir, pour tout borélien $B \in \mathcal{B}$, $Pr(B)$. Or, ceci n'est possible que si $\{X \in B\} = X^{-1}(B)$ appartient à $\mathcal{A}$:

Définition. Soit $(\Omega, \mathcal{A}, Pr)$ un espace probabilisé quelconque. Une application X de Ω dans $\mathbb{R}$ est une *variable aléatoire* (v.a.) si, pour tout $B \in \mathcal{B}$, $X^{-1}(B) \in \mathcal{A}$.

Nous admettons la proposition suivante :

Proposition. Une application X de Ω dans $\mathbb{R}$ est une v.a. sur $(\Omega, \mathcal{A}, Pr)$ ssi $X^{-1}(]-\infty, x]) = \{X \leq x\} \in \mathcal{A}$.

Définition. On appelle *fonction de répartition* (f.r.) F de le v.a. X la f.r. de la loi de probabilité de X :

$$F(t) = Pr(X \leq t), \quad \forall t \in \mathbb{R}.$$

Exemple. Soit le réel U uniformément généré dans l'intervalle $[0,2]$. On définit l'application X par $X(u) = u^2$. Étudier X et en donner la f.r., s'il s'agit d'une v.a.

Solution. L'espace probabilisé initial est $\big([0,2], \mathcal{B}_{[0,2]}, Pr\big)$, où $\mathcal{B}_{[0,2]}$ est la tribu des boréliens contenus dans $[0,2]$ et Pr la mesure de probabilité qui associe à chaque sous-intervalle de $[0,2]$ la moitié de sa longueur. La f.r. de U est donnée par :

$$F_U(t) = \begin{cases} 0 & \text{si } t < 0 \\ \frac{t}{2} & \text{si } 0 \leq t < 2 \\ 1 & \text{si } t \geq 2. \end{cases}$$

L'application X est bien une v.a., puisque l'image inverse d'un sous-intervalle $\langle a, b \rangle$ de $[0,4]$ par X est le sous-intervalle $\langle \sqrt{a}, \sqrt{b} \rangle$ de $[0,2]$, qui est un borélien.

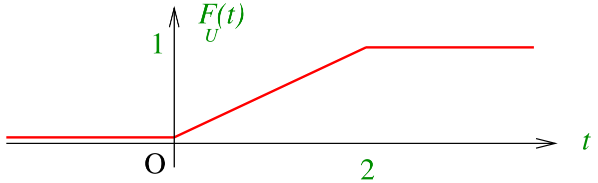
Pour calculer le f.r. de X, nous remarquons d'abord que $F_X(x)$ vaut 0 pour $x < 0$ et 1 pour $x \geq 4$. Ses valeurs non triviales sont obtenues pour $0 \leq x \leq 4$.

Soit $x \in [0,4]$. Nous avons :

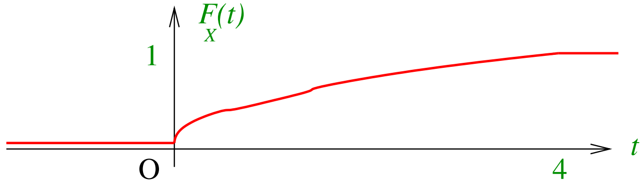
$$F_X(x) = Pr(X \leq x) = Pr(U \leq \sqrt{x}) = F_U(\sqrt{x}) = \sqrt{x}/2.$$

D'où :

$$F_X(x) = \begin{cases} 0 & \text{si } x < 0 \\ \frac{\sqrt{x}}{2} & \text{si } 0 \leq x < 4 \\ 1 & \text{si } x \geq 4. \end{cases}$$



Fonction de répartition de U tiré uniformément dans [0,2]



Fonction de répartition de $X=U^2$

Probabilité Conditionnelle

Le concept de conditionnement introduit pour les espaces discrets s'étend sans difficulté aux espaces quelconques

Définitions. Soit $(\Omega, \mathcal{A}, Pr)$ un espace probabilisé et soit $A \in \mathcal{A}$ de probabilité non nulle. On définit une application $Pr(./A)$ de $\mathcal{A}$ dans $[0,1]$, par :

$$Pr(B/A) = \frac{Pr(A \cap B)}{Pr(A)}, \quad \forall B \in \mathcal{A}.$$

Elle est appelée *probabilité conditionnelle sachant* A. Le nombre $Pr(B/A)$ est appelé *probabilité conditionnelle de* B sachant A.

On démontre facilement que le triplet $(\Omega, \mathcal{A}, Pr(./A))$ est un espace probabilisé. Il est appelé *espace probabilisé conditionné par* A.

On remarque aussi que la probabilité d'une intersection peut s'écrire en termes de probabilité conditionnelle :

$$Pr(A \cap B) = Pr(A)Pr(B/A).$$

On dit alors que l'on a effectué un *conditionnement*.

On démontre aussi facilement la proposition suivante.

Proposition. Soit $A_1, A_2, ..., A_n$ une partition de Ω dans $\mathcal{A}$, telle que $Pr(A_i) \neq 0, i = 1, ..., n$. Alors, pour tout $B \in \mathcal{A}$, on a :

$$Pr(B) = \sum_{i=1}^{n} Pr(A_i)Pr(B/A_i).$$

La définition d'*indépendance* pour les événements d'un espace généralisé est tout à fait identique à celle d'un espace discret.

Indépendance

Soit $(\Omega, \mathcal{A}, Pr)$ un espace probabilisé. On peut définir un *vecteur aléatoire* (ou une *variable aléatoire à valeurs dans* $\mathbb{R}^k$) comme une application de Ω dans $\mathbb{R}^k$ telle que, pour tout $B \in \mathcal{B}_k$, $\{X \in B\} = X^{-1}(B)$ soit dans $\mathcal{A}$.

Si X est une v.a. à valeurs dans $\mathbb{R}^k$ sur l'espace $(\Omega, \mathcal{A}, Pr)$, On peut munir $(\mathbb{R}^k, \mathcal{B}_k)$ de la mesure de probabilité qui associe à tout $B \in \mathcal{B}_k$, $Pr(\{X \in B\})$.

La proposition suivante permet de réduire la vérification de la propriété d'appartenance à $\mathcal{A}$ aux images inverses des pavés.

Proposition. Une application X de Ω dans $\mathbb{R}^k$ est une v.a. ssi, pour tout $x \in \mathbb{R}^k$, $\{X \leq x\} \in \mathcal{A}$.

Dans ce cas, on peut définir la *fonction de répartition* du vecteur X par :

$$F(x) = F(x_1, ..., x_k) = Pr\,(X \leq (x_1, ..., x_k)).$$

Si l'on pose $X = (X_1, ..., X_k)$, on aura :

$$Pr(X_1 \leq x_1, ..., X_k \leq x_k) = F(x_1, ..., x_k).$$

Définition. Soit $X = (X_1, ..., X_k)$ une v.a. définie sur $(\Omega, \mathcal{A}, Pr)$. Les v.a. $X_1, ..., X_k$ sont dites *indépendantes*, si pour tout $B_1, ..., B_n \in \mathcal{B}$:

$$Pr\,(X_1 \in B_1, ..., X_k \in B_k) = \prod_1^k Pr(X_i \in B_i).$$

Définition. Si $F(x_1, ..., x_k)$ est la f.r., la *fonction de répartition marginale* $F_i(x_i), i = 1, ..., k$, est $\lim_{x_j \to \infty, \forall j \neq i} F(x_1, ..., x_k)$.

Nous admettons les deux propositions suivantes.

Proposition. Les v.a. $X_1, ..., X_k$ sont indépendantes, ssi :

$$F(x_1, ..., x_k) = \prod_1^k F_i(x_i).$$

Proposition. Soient $X_1, ..., X_k$ des v.a. indépendantes réelles. Soient $\varphi_1, ..., \varphi_k$ des applications réelles. Alors les v.a. $\varphi_1(X_1), ..., \varphi_k(X_k)$ sont des v.a. indépendantes.

Cette proposition est importante. Elle confirme en effet l'idée intuitive que les transformations des v.a. indépendantes préservent l'indépendance des variables.

Dans la suite, nous supposerons que les v.a. non discrtètes admettent une densité. Cette hypothèse permet de simplifier les calculs en termes d'intégrales de Riemann, dont la connaissance est supposée être acquise en classe terminale. Sinon, une révision est vivement conseillée !

Espérance des v.a. Continues

La *densité* d'une v.a. $X = (X_1, ..., X_k) \in \mathbb{R}^k$, lorsqu'elle existe, est la fonction $f(t_1, ..., t_k)$ telle que, pour tout $, x_1, ..., x_k \in \mathbb{R}$:

$$F(x_1, ..., x_k) = \int_{-\infty}^{x_k} ... \int_{-\infty}^{x_1} f(t_1, ..., t_k) dt_1 ... dt_k.$$

Notons tout de suite qu'on doit avoir :

$$\int_{-\infty}^{\infty} ... \int_{-\infty}^{\infty} f(t_1, ..., t_k) dt_1 ... dt_k = 1.$$

Définitions

- **(v.a. positives)** Soit X une v.a. à valeurs dans $\mathbb{R}^+$, admettant la densité f. Son *espérance* est définie par :

  $$\mathbb{E}X = \int_{-\infty}^{\infty} x f(x) dx,$$

  si l'intégrale converge. Si elle diverge, on dit que X *n'admet pas d'espérance.*
- **(v.a. de signe quelconque)** Soit X une v.a. réelle. Décomposons X en partie positive et partie négative : $X = X^+ - X^-$, avec $X^+ = \max\{X, 0\}$ et $X^- = -\min\{X, 0\}$. Si X^+ et X^- admettent toutes deux une espérance, on dit que X *admet l'epérance* :

  $$\mathbb{E}X = \mathbb{E}X^+ - \mathbb{E}X^-.$$

  Sinon, on dit que X *n'admet pas d'espérance.*

Quelques Exemples

Exemple 1 (Loi Uniforme). Soit la v.a. X uniformément répartie dans $[a,b]$, avec $a<b$. Elle admet la densité :

$$f(x)=\begin{cases} 0 & \text{si} \quad x<a \\ \frac{1}{b-a} & \text{si} \quad a\le x<b \\ 0 & \text{si} \quad x\ge b. \end{cases}$$

Nous avons :

$$\mathbb{E}X=\int_a^b \frac{x}{b-a}dx=\left[\frac{x^2}{2(b-a)}\right]_a^b=\frac{b^2-a^2}{2(b-a)}=\frac{b+a}{2}.$$

Il est à noter qu'on peut calculer l'espérance de toute v.a. de la forme $\varphi(X)$ par une intégration similaire. Par exemple l'espérance de X^2 vaut :

$$\mathbb{E}X^2=\int_a^b \frac{x^2}{b-a}dx=\left[\frac{x^3}{3(b-a)}\right]_a^b=\frac{b^3-a^3}{3(b-a)}=\frac{b^2+a^2+ab}{3}.$$

Exemple 2 (Loi Normale). La distribution que nous allons introduire a un rôle central en statistique. Elle est appelée aussi *loi gaussienne* ou *loi de Laplace-Gauss*. Elle intervient en effet dans la modélisation probabiliste d'un nombre de grandeurs physiques observée un grand nombre de fois.

Soit X une v.a. réelle. On dit que X *suit une loi normale de paramètres* m et σ, si elle admet la densité :

$$f(x)=\frac{1}{\sigma\sqrt{2\pi}}\exp\left(-\frac{(x-m)^2}{2\sigma^2}\right), \quad x\in\mathbb{R}.$$

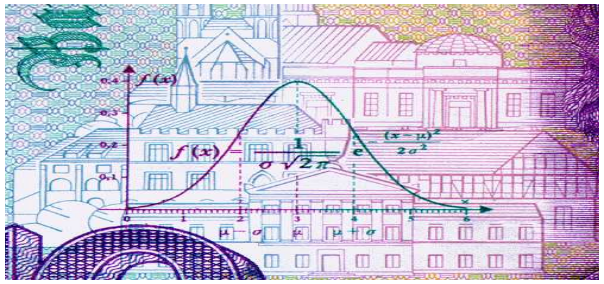
Il est d'usage de noter une telle distribution par $\mathcal{N}(m,\sigma)$.

Pour prouver qu'il s'agit d'une densité, il faut montrer que $\int_{-\infty}^{\infty} f(x)dx=1$. Pour ce faire, on introduit d'abord le changement de variable $y=\frac{x-m}{\sqrt{2}\sigma}$. Ensuite l'application de l'équation :

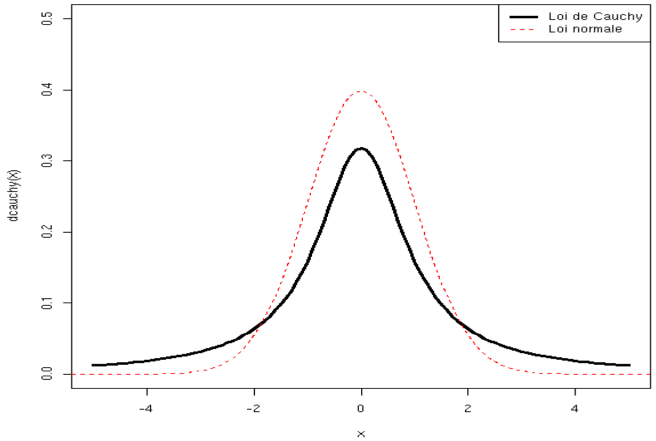
$$\int_{-\infty}^{\infty} e^{-y^2}dy=\sqrt{\pi}$$

permet de confirmer l'identité souhaitée.

Gauss et sa courbe



Comparaison de la densité de la loi de Cauchy et de la normale



Exemple 3 (Loi de Cauchy). La distribution de Cauchy est donnée par sa densité :

$$f(x) = \frac{1}{\pi} \frac{1}{1+x^2}, \; x \in \mathbb{R}.$$

On démontre facilement :

$$F(t) = Pr(X \leq t) = \frac{1}{2} + \frac{Arctan(t)}{\pi}, \; t \in \mathbb{R}.$$

Ce qui fournit comme corollaire $\int_{-\infty}^{\infty} = 1$.