C

Fonction de répartition et quantile

Définition C.1. *Soit X une variable aléatoire réelle. On appelle fonction de répartition de X la fonction $F : \mathbb{R} \to [0,1]$ définie par*

$$\forall x \in \mathbb{R},\ F(x) = \mathbb{P}(X \leq x).$$

La proposition suivante regroupe des propriétés classiques de la fonction de répartition que nous ne démontrerons pas.

Proposition C.2. *La fonction de répartition F d'une variable aléatoire réelle X est croissante, continue à droite et vérifie*

$$\lim_{x\to-\infty} F(x) = 0 \quad \textit{et} \quad \lim_{x\to+\infty} F(x) = 1.$$

Ensuite, la fonction de répartition caractérise la loi : si X et Y sont deux variables aléatoires réelles qui ont même fonction de répartition, alors

$$\forall f : \mathbb{R} \to \mathbb{R} \textit{ mesurable bornée },\ \mathbb{E}[f(X)] = \mathbb{E}[f(Y)].$$

Enfin, une suite $(X_n, n \in \mathbb{N}^*)$ de variables aléatoires réelles converge en loi vers X si et seulement si, pour tout point de continuité $x \in \mathbb{R}$ de la fonction de répartition F de X, on a

$$F_n(x) \underset{n\to\infty}{\longrightarrow} F(x),$$

où F_n désigne la fonction de répartition de X_n.

Soit $x \in \mathbb{R}$ et $(X_n, n \in \mathbb{N}^*)$ une suite de variables aléatoires réelles indépendantes et de même loi. D'après la loi forte des grands nombres, on a p.s.

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} \mathbf{1}_{\{X_k \leq x\}} = \mathbb{P}(X_1 \leq x).$$

Le théorème suivant assure en fait que cette convergence est uniforme en x.

Théorème C.3 (Glivenko-Cantelli). *Soit $(X_n, n \in \mathbb{N}^*)$ une suite de variables aléatoires réelles indépendantes et de même loi. On a p.s.*

$$\lim_{n\to\infty} \sup_{x\in\mathbb{R}} |F_n(x) - \mathbb{P}(X_1 \leq x)| = 0,$$

où F_n, définie par

$$F_n(x) = \frac{1}{n}\sum_{k=1}^{n} \mathbf{1}_{\{X_k \leq x\}} = \frac{1}{n} \text{Card } \{k \in \{1, \ldots, n\} : X_k \leq x\},$$

est la fonction de répartition empirique de l'échantillon $X_1, \ldots, X_n$.

Pour la démonstration de ce résultat, nous renvoyons par exemple à la référence [1] de l'appendice A.

Définition C.4.

- *Soit X une variable aléatoire réelle de fonction de répartition F. Pour $p \in \left]0,1\right]$, on appelle quantile ou fractile d'ordre p de X le nombre* $\boxed{x_p = \inf\{x \in \mathbb{R},\ F(x) \geq p\}}$ *où par convention* $\inf \emptyset = +\infty$.
- *L'application $p \in \left]0,1\right[\to x_p \in \mathbb{R}$ s'appelle l'inverse généralisé de F. On la note F^{-1} : $\forall p \in \left]0,1\right[,\ F^{-1}(p) = x_p$.*

Le résultat suivant est à la base de la méthode d'inversion de la fonction de répartition destinée à simuler des variables aléatoires réelles de fonction de répartition F.

Proposition C.5. *Soit F une fonction de répartition et F^{-1} son inverse généralisé. Alors on a l'équivalence*

$$F(x) \geq p \Leftrightarrow x \geq F^{-1}(p). \tag{C.1}$$

En outre, si F est continue, alors

$$\forall p \in \left]0,1\right[,\ F(F^{-1}(p)) = p. \tag{C.2}$$

Soit X une variable aléatoire réelle de fonction de répartition F et U une variable aléatoire de loi uniforme sur $[0,1]$. Alors, la variable $F^{-1}(U)$ a même loi que X. Et si F est continue, alors $F(X)$ suit la loi uniforme sur $[0,1]$.

Si la fonction F est inversible de $\mathbb{R}$ dans $\left]0,1\right[$, alors elle est continue sur $\mathbb{R}$ et l'égalité (C.2) entraîne que l'inverse généralisé et l'inverse coïncident. L'intérêt de l'inverse généralisé est qu'il reste défini même lorsque F n'est pas inversible soit parce que cette fonction est discontinue soit parce qu'elle est constante sur des intervalles non vides.

Démonstration. Soit $p \in]0,1[$. Nous allons d'abord vérifier (C.1). Par définition de $x_p = F^{-1}(p)$, il est clair que si $F(x) \geq p$ alors $x \geq F^{-1}(p)$. En outre, pour tout $n \in \mathbb{N}^*$, il existe $y_n \leq F^{-1}(p) + \frac{1}{n}$ tel que $F(y_n) \geq p$. Par croissance de F, on a $F\left(F^{-1}(p)+\frac{1}{n}\right) \geq p$ pour tout $n \in \mathbb{N}^*$. Par continuité à droite de F, on en déduit que

$$\forall p \in]0,1[,\ F(F^{-1}(p)) \geq p. \tag{C.3}$$

Avec la croissance de F, cela implique que si $x \geq F^{-1}(p)$, alors $F(x) \geq p$, ce qui achève la démonstration de (C.1).

L'équivalence (C.1) implique que pour tout $x < F^{-1}(p)$, on a $F(x) < p$. Avec (C.3), on en déduit que si F est continue au point $F^{-1}(p)$ alors $F(F^{-1}(p)) = p$, ce qui entraîne (C.2).

Enfin, si X a pour fonction de répartition F et U est une variable aléatoire uniforme sur $[0,1]$, on a d'après (C.1)

$$\forall x \in \mathbb{R},\ \mathbb{P}(F^{-1}(U) \leq x) = \mathbb{P}(U \leq F(x)) = F(x).$$

Les variables X et $F^{-1}(U)$ ont même fonction de répartition. Elles ont donc même loi. Par conséquent, $F(X)$ a même loi que $F(F^{-1}(U))$, variable aléatoire qui est égale à U avec probabilité 1 lorsque F est continue d'après (C.2). $\square$