Variables aléatoires continues

À l'intention des étudiants de STT-1500



Jérôme Soucy

Département de mathématiques et de statistique
Université Laval

Hiver 2025

Résumé

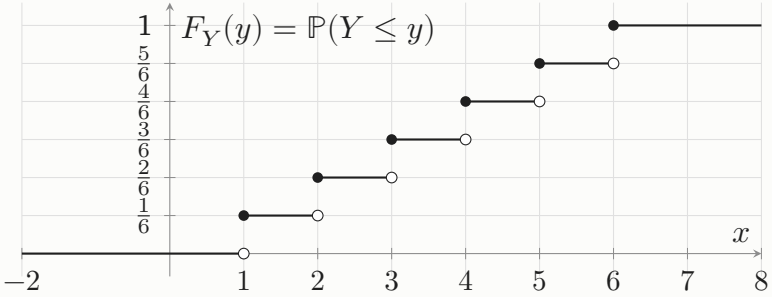
1. Introduction
2. Fonction de répartition
3. Espérance d'une v.a. abs. continue
4. Deux distributions
5. Fonction d'une variable aléatoire
6. Simulation
7. Conditionnement

1

Introduction

Résultats du lancer d'un dé

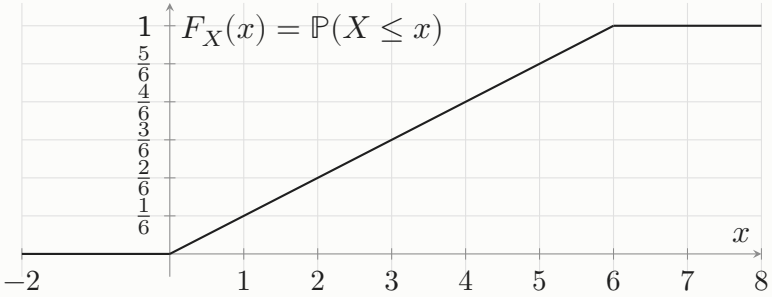
Supposons que Y est la variable aléatoire correspondant au résultat du lancer d'un dé régulier.



Vers Idée de base

On souhaite modéliser des phénomènes aléatoires ou l'ensemble des valeurs possibles est non dénombrables. Commençons par un exemple simple : choisir au hasard un nombre réel dans l'intervalle $]0, 6[$. Intuitivement, si X désigne une variable aléatoire correspondant au nombre choisi, nous voudrions que $\mathbb{P}(X \leq 0) = 0, \mathbb{P}(X \leq 1) = 1/6, \mathbb{P}(X \leq 2) = 2/6 \ldots$ et $\mathbb{P}(X \leq 6) = 1$. Si on note $F_X(x)$ la fonction correspondant à $\mathbb{P}(X \leq x)$, à quoi devrait ressembler le graphique de F_X ?

Choisir un nombre entre 0 et 6



Écrire F_X comme une intégrale

Dans le cas où on s'intéresse à choisir un nombre t dans $]0, 6[$, $\mathbb{P}(\{X = t\}) = 0$. Par contre, si on s'intéresse à la réalisation d'un événement du genre $\mathbb{P}(1 < X \leq 2)$, on constate que cela correspond à

$$F_X(2) - F_X(1).$$

C'était aussi le cas dans le cas discret (lancer d'un dé). Le graphique tracé précédemment suggère que F_X est continue et que sa dérivée est constante. On peut donc écrire $F_X(x) = \int_0^x f(t)\,dt$, où

$$f(t) = \begin{cases} 0 & \text{si } x \leq 0, \\ \frac{1}{6} & \text{si } 0 < x \leq 6. \end{cases}$$

2

Fonction de répartition

Définition

Définition. Soit X une variable aléatoire. La **fonction de répartition** de X est la fonction $F(x)$ définie par

$$F(x) = \mathbb{P}(X \leq x).$$

Les variables aléatoires discrètes, de même que les variables aléatoires absolument continues, possèdent une fonction de répartition.

Théorème. Soit X une variable aléatoire, et soit a et b des nombres réels vérifiant $a < b$. Alors

$$\mathbb{P}(a < X \leq b) = F(b) - F(a).$$

Lien entre fonction de masse et fonction de répartition

Soit X une v.a. discrète dont la fonction de masse est $p(x)$. On a donc que

$$F(x) = \sum_{u \in (-\infty, x]} p(u).$$

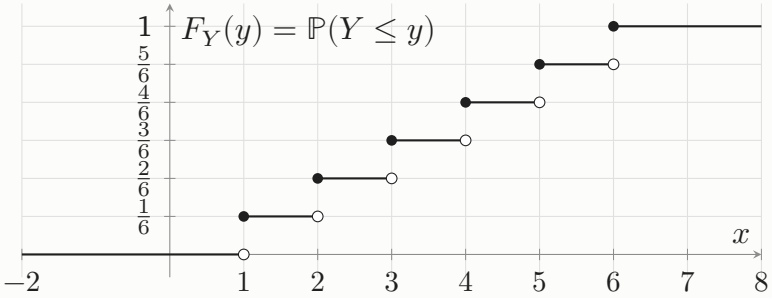
Si on pose $F(x^-) = \lim_{\Delta x \to 0^-} F(x + \Delta x)$, on trouve que

$$p(x) = F(x) - F(x^-).$$

Cette quantité correspond à la grandeur du saut de F en x.

Fonction de répartition associée au lancer d'un dé

On lance un dé régulier à 6 faces. On note Y le nombre obtenu. Tracez la fonction de répartition de Y.



Lien entre densité et fonction de répartition

Définition. Soit X une v.a. telle que pour tout sous-ensemble B de $\mathbb{R}$ on ait

$$\mathbb{P}(X \in B) = \int_B f(x)\,dx.$$

X est alors une v.a. (absolument) continue de **densité de probabilité** $f(x)$.

On a donc que

$$F(x) = \mathbb{P}(X \leq x) = \mathbb{P}(X \in (-\infty, x]) = \int_{-\infty}^{x} f(u)du.$$

En utilisant le théorème fondamental du calcul différentiel et intégral, on obtient aussi que $F' = f$.

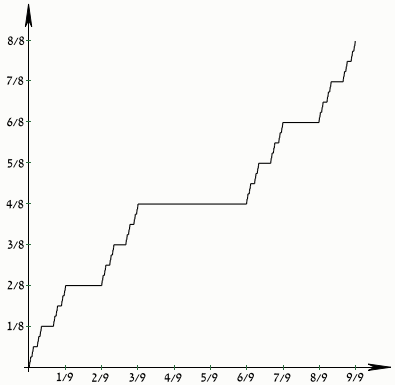
Fonction de répartition sur $\mathbb{R}$

Une fonction $F : \mathbb{R} \to \mathbb{R}$ est appelée **fonction de répartition** si elle satisfait les conditions ci-dessous :

1. Si $x < y$, alors $F(x) \leq F(y)$;
2. Pour tout $x \in \mathbb{R}$, $\lim_{\Delta x \to 0^+} F(x + \Delta x) = F(x)$;
3. $\lim_{x \to -\infty} F(x) = 0$;
4. $\lim_{x \to \infty} F(x) = 1$.

Continuité vs continuité absolue

L'escalier de Cantor est un exemple d'une fonction continue monotone croissante de dérivée nulle presque partout (la dérivée n'est pas définie pour les $x \in \mathbb{Q}$). On ne peut pas l'écrire comme l'intégrale d'une fonction continue, elle n'est donc pas absolument continue. La fonction illustrée ci-contre est un exemple de fonction de répartition continue qui n'est pas absolument continue. Comme ces cas sont plutôt rares en pratique, nous ne distinguerons pas ces types de continuité.



Escalier de Cantor (voir [1])

3

Espérance d’une v.a. abs. continue

Espérance d'une v.a. absolument continue

Définition. Soit X une v.a. avec densité $f(x)$. Si $\int_{-\infty}^{\infty} |x| f(x) dx < \infty$, **l'espérance** de X est définie par

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

Remarques. On écrit souvent que $\mu_X = \mathbb{E}(X)$, et on appelle parfois **moyenne** l'espérance.

4

Deux distributions

La loi uniforme

Définition. On dira qu'une v.a. X suit une **loi uniforme** sur l'intervalle ouvert (a, b), si sa densité de probabilité est donnée par

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

On écrira alors que

$$X \sim \text{uniforme}\,(a, b)\,.$$

On trouve facilement que si $X \sim \text{uniforme}\,(a, b)$, alors $\mathbb{E}(X) = \frac{a+b}{2}$.

La loi exponentielle

Définition. Soit $\lambda > 0$. On dira qu'une v.a. X suit une **loi exponentielle** de paramètre λ si sa densité de probabilité est donnée par

$$f(x) = \begin{cases} \lambda \mathrm{e}^{-\lambda x} & \text{si } x > 0, \\ 0 & \text{sinon.} \end{cases}$$

On écrira alors que

$$X \sim \text{exponentielle}\,(\lambda)\,.$$

On trouve facilement que si $X \sim \text{exponentielle}\,(\lambda)$, alors

$$\mathbb{E}(X) = \frac{1}{\lambda}.$$

5

Fonction d’une variable aléatoire

Motivation - Fonction d'une v.a.

Soit X une variable aléatoire suivant une loi uniforme sur l'intervalle $(0, 1)$. Quelle est la distribution de la variable aléatoire Y correspondant à X^2 ? Plus généralement, si X est une variable aléatoire, quelle sera la distribution de la variable aléatoire Y définie par $Y = g(X)$, où g est une fonction réelle d'une variable réelle ?

Cas discret

Soit X une v.a. discrète dont la fonction de masse est $p_X(x)$. On veut trouver $p_Y(y)$, la fonction de masse de la variable aléatoire Y définie par $Y = g(X)$. Pour y arriver, on :

1. détermine l'ensemble des valeurs possibles pour Y ;
2. calcule, pour chaque valeur possible y, la probabilité $p_Y(y)$ en remarquant que

$$\begin{aligned} p_Y(y) &= \mathbb{P}(Y = y) = \mathbb{P}(g(X) = y) \\ &= \mathbb{P}(X \in \{x \in \mathbb{R} : g(x) = y\}) \\ &= \sum_{\{x \in \mathbb{R} : g(x) = y\}} p_X(x). \end{aligned}$$

Pour obtenir la fonction de répartition de Y, on remarque que $F_Y(y) = \sum_{u \in (-\infty, y]} p_Y(u)$.

Cas absolument continu

Soit X une v.a. absolument continue dont la densité de probabilité est $f_X(x)$, et la fonction de répartition est $F_X(x)$. On veut trouver $f_Y(y)$, la densité de probabilité de la variable aléatoire Y définie par $Y = g(X)$. Pour y arriver, on

1. détermine l'ensemble des valeurs possibles pour Y ;
2. calcule la fonction de répartition $F_Y(y)$ en observant que

$$\begin{aligned} F_Y(y) &= \mathbb{P}(Y \leq y) \\ &= \mathbb{P}(g(X) \leq y) \\ &= \mathbb{P}(X \in \{x \in \mathbb{R} : g(x) \leq y\}) \\ &= \int_{\{x \in \mathbb{R} : g(x) \leq y\}} f_X(x) dx; \end{aligned}$$

3. dérive la fonction $F_Y(y)$ par rapport à y pour obtenir $f_Y(y)$.

Deux cas fréquents

Soit X une v.a. absolument continu dont la densité de probabilité est $f_X(x)$. Supposons que l'ensemble des valeurs de x telles que $f_X(x) > 0$ est un intervalle ouvert I. Supposons que $Y = g(X)$, où g est une fonction strictement croissante sur I, avec dérivée strictement positive partout sur I, ou bien strictement décroissante sur I, avec dérivée strictement négative sur I. Appelons J l'image de I par g. Alors la densité de Y est donnée par

$$f_Y(y) = \begin{cases} f_X\left(g^{-1}(y)\right) \left| \frac{d}{dy} g^{-1}(y) \right| & \text{si } y \in J; \\ 0 & \text{sinon.} \end{cases}$$

6 Simulation

Cas discret

Faisons l'hypothèse que nous avons la possibilité de générer un nombre aléatoire dans l'intervalle $(0, 1)$.
On veut simuler, par exemple, une variable aléatoire dont la fonction de masse est $p(x)$. On suppose que $p(x)$ prend des valeurs positives seulement pour $x_1, x_2, \ldots$. L'idée est de partitionner l'intervalle $(0, 1)$ de telle sorte que pour chaque intervalle I_j de la partition choisie, nous avons que $p(x_j) =$ longueur de l'intervalle I_j.
Si $U \sim$ uniforme $(0, 1)$, alors

$$\mathbb{P}\big(X = x_j\big) = \mathbb{P}\big(U \in I_j\big) = p(x_j).$$

Cas absolument continu

Faisons à nouveau l'hypothèse que nous avons la possibilité de générer un nombre aléatoire dans l'intervalle $(0,1)$. C'est donc dire qu'on peut simuler la variable aléatoire U, où $U \sim$ uniforme $(0,1)$.
On veut simuler une variable aléatoire X dont la fonction de densité est $f(x)$ et la fonction de répartition $F(x)$. On fait l'hypothèse supplémentaire que F est strictement croissante sur un intervalle I vérifiant $F(I) = (0,1)$. Ce n'est pas une hypothèse trop exigeante. Il existe donc une fonction $F^{-1} : (0,1) \to I$ qui est une bijection. Définissons la variable aléatoire $Y = F^{-1}(U)$. Nous avons que

$$\begin{aligned}\mathbb{P}(Y \leq y) &= \mathbb{P}(F^{-1}(U) \leq y) = \mathbb{P}(U \leq F(y)) \qquad (F \text{ est strictement croissante}) \\ &= \frac{\text{longueur de l'intervalle } (0, F(y))}{\text{longueur de l'intervalle } (0,1)} \\ &= F(y).\end{aligned}$$

Fonction quantile

Définition. Soit X une v.a. de fonction de répartition continue F. La fonction inverse de F se nomme la **fonction quantile** de X.

Si on souhaite déterminer la médiane de X, il suffit de calculer $F^{-1}(0{,}5)$. De même, si on souhaite trouver le premier quartile, on peut calculer $F^{-1}(0{,}25)$.

7

Conditionnement

Cas discret

Soit E_1, E_2, ... E_n des ensembles qui forment une partition de Ω, l'ensemble fondamental d'une expérience aléatoire. D'après la loi des probabilités totales, nous avons que

$$\mathbb{P}(A) = \sum_{i=1}^{n} \mathbb{P}(A|E_i)\mathbb{P}(E_i).$$

Imaginons maintenant une variables aléatoire X telle que $X = x_i \iff E_i$ se réalise. On peut donc écrire que

$$\mathbb{P}(A) = \sum_{i=1}^{n} \mathbb{P}(A|X = x_i)\mathbb{P}(X = x_i).$$

Cas absolument continu

Soit X une variable aléatoire dont la densité de probabilité est $f_X(x)$. Supposons qu'on souhaite calculer la probabilité qu'un événement A se réalise, et que la probabilité de réalisation de A est affectée par la valeur que prend la variable aléatoire X. Faisons l'hypothèse que $\mathbb{P}(A|X=x)$ est une fonction de x connue, où facilement calculable. Alors

$$\mathbb{P}(A) = \int_{\mathbb{R}} \mathbb{P}(A|X=x) f_X(x) dx.$$

Notons que l'expression $\mathbb{P}(A|X=x)$ ne correspond pas à la définition du chapitre 3, puisque l'événement $\{X=x\}$ est de probabilité nulle.

Bibliographie

[1] Theon. *A graph of the Cantor function*. CC BY-SA 3.0, via Wikimedia Commons. Document consulté le 18 février 2025. URL : https://commons.wikimedia.org/wiki/File:CantorEscalier.svg.

Remarques.

- Ce document est rédigé avec Beamer, une classe de LaTeX.

Pour signaler une erreur dans ce document, veuillez écrire à jerome.soucy@mat.ulaval.ca.