Tome 2
Les probabilités continues

Chapitre 10

Introduction aux probabilités continues

Rappel

Définition

Une variable aléatoire est une fonction d'un espace de probabilité Ω vers $\mathbb{R}$, $X : \Omega \to \mathbb{R}$

Si X peut prendre un nombre fini ou dénombrable de valeurs (cad si $X(\Omega)$ est un ensemble fini, ou infini dénombrable), on parle d'une **variable aléatoire discrête**

Définition

La **fontion de distribution d'une variable aléatoire discrête** est alors la fonction

$$\begin{aligned} \mathbb{P} \quad : \quad X(\omega) &\to \mathbb{R} \\ x_i &\mapsto \mathbb{P}(X = x_i) \end{aligned}$$

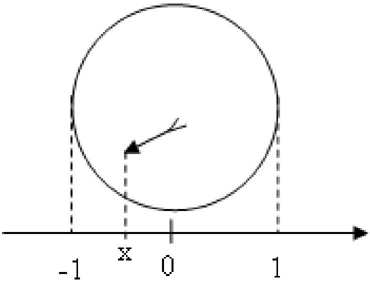
Exemple : Distribution binomiale,
$\mathbb{P}(X = k) = C_n^k p^k (1-p)^{n-k} \ \forall k = 0, 1, ..., n$
($=$ distribution de la valeur finale dans un arbre binomial)

Il existe des variables aléatoires qui peuvent prendre des valeurs dans un sous-ensemble **continue** de $\mathbb{R}$. Par exemple :

$X : \Omega \to [a, b]$	l'image de X est un intervalle [a,b]
$X : \Omega \to \mathbb{R}^+$	l'image de X est $[0, \infty[$
$X : \Omega \to \mathbb{R}$	l'image de X est $\mathbb{R}$ tout entier

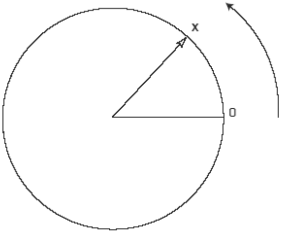
Par exemple, dans un jeu de fléchette,

- si on regarde l'absisse du point atteint par le fléchette,
- elle peut prendre n'importe quelle valeur dans $[-1, 1]$
- il n'existe pas un ensemble fini de valeurs possibles
- en fonction de l'habilité du joueur, ce point tombera plus ou moins probablement dans la région centrale de l'intervalle
- la distribution d'une v.a. continue décrit la probabilité de tomber dans tel ou tel sous-ensemble (région) de $X(\Omega)$



Autre exemple : Soit un cercle de circonférence unitaire, et un pointeur placé en son centre

- on choisit arbitrairement un pt du cercle que l'on note 0
- chaque point du cercle est ensuite définit par la distance entre 0 et ce point, mesuré dans le sens inverse des aiguille d'une montre
- l'expérience consiste à faire tourner le pointeur et noter le point du cercle indiqué par le pointeur
- soit X le résultat de l'expérience
- l'univers-image $X(\Omega)$ est l'intervalle $[0; 1[$



- on cherche à construire un modèle de probabilités dans lequel chaque réalisation a la même proba de survenance
- si on procède comme dans le cas de variables discrète, on doit assigner une probabilité nulle à chaque **réalisation**, sinon la somme des proba. ne sera pas égale à 1.
- cependant, si toute les proba sont nulles, la somme est 0 alors qu'elle devrait être 1.
- ainsi, la proba de chaque événement ne doit pas être nulle
- par exemple, la probabilité $\mathbb{P}(0 \leq X < 1)$ que le pointeur montre un point du cercle doit être égale à 1
- de même, on devrait avoir $\mathbb{P}(0 \leq X < 1/2) = \mathbb{P}(1/2 \leq X < 1) = 1/2$
- de manière plus générale, il serait normal que $\mathbb{P}(c \leq X < d) = d - c$

- Si on note $E = [c, d]$, la formule précédente peut s'écrire

$$\mathbb{P}(E) = \int_E f_X(x)dx$$

avec $f(x)$ une fonction constante de valeur $1 \ \forall x$

- cela devrait vous rappeler la formule correspondante dans le cas discret : $\mathbb{P}(E) = \sum_{\omega \in E} m(\omega)$
- La différence est que dans le cas continu, la quantité integrée : $f_X(x)$, n'est pas la probabilité de l'issue x (Cependant, si on utilise des variation infinitisémales, on peut considérer $f_X(x).dx$ comme la proba de l'issue x.)

Definition

Soit X une variable aléatoire continue (à valeurs réelles). Une **fonction de densité** de X, est une fonction à valeurs réelles qui satisfait

$$\mathbb{P}(a \leq X \leq b) = \int_a^b f_X(x)dx \ \forall [a, b] \in \mathbb{R}$$

Remarque : toutes les v.a. continues, ne possèdent pas nécessairement de fonction de densité. On ne considèrera cependant par la suite que des v.a. pour lesquelles une fonction de densité existe.

En termes de densité, si E est un sous-ensemble de $\mathbb{R}$, alors

$$\mathbb{P}(X \in E) = \int_E f_X(x)dx$$

La notation suppose ici que E est un sous-ensemble de $\mathbb{R}$ pour lequel $\int_E f_X(x)dx$ fait sens.

Dans l'exemple "du pointeur"

- l'ensemble des possibles est l'intervalle $0 \leq x < 1$,
- et la fonction de densité s'écrit

$$f(x) = \begin{cases} 1, \text{ si } 0 \leq x < 1 \\ 0, \text{ sinon} \end{cases}$$

- si E représente l'évenement "le pointeur montre la partie haute du cercle", alors $E = \{x/0 \leq x \leq \frac{1}{2}\}$, et donc

$$\mathbb{P}(E) = \int_0^{1/2} 1dx = \frac{1}{2}$$

- D'une manière plus générale, si E est l'événement "le pointeur est dans l'intervalle [a,b], alors

$$\mathbb{P}(E) = \int_a^b 1dx = b - a$$

- On voit dans cette exemple que (contrairement au cas discret), la valeur $f_X(x)$ de la fonction de densité pour l'issue x n'est pas la probabilité que x survienne (cette proba ne peut pas être égale à 1 pour tout x!)

- Cependant, la fonction de densité contient bien de l'info sur les proba puisque la proba de chaque événement peut en être déduit

- En particulier, la proba que l'issue de l'expérience soit dans l'intervalle [a,b] est donné par $\mathbb{P}([a,b]) = \int_a^b f_X(x)dx$, c'est-à-dire l'**aire sous le graphe de la fonction de densité dans l'intervalle [a,b]**

- Ainsi, il y a une forte connection entre probabilités et aire

- La probabilité d'occurrence d'un évenement de la forme $[x, x + dx]$, où dx est petit, est approximativement donnée par $\mathbb{P}([x; x + dx]) \simeq f_X(x)dx$, c'est-à-dire par l'aire du rectangle sous le graphe de f

- Notez que quand $dx \to 0$, cette probabilité tend vers 0, de telle sorte que la probabilité $\mathbb{P}(\{x\})$ d'un simple point est 0, comme dans l'exemple.

Définition

Soit X une variable aléatoire continue à valeurs réelles. Alors la **fonction de distribution** de X est définie par l'équation

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

Si X possède une fonction de densité, alors elle a aussi une fonction de distribution, et le théorème suivant montre comment ces deux fonctions sont reliées

Théorème

Soit X une v.a. continue de densité $f_X(x)$. Alors la fonction définie par $F(x) = \int_{-\infty}^{x} f_X(t)dt$ est la fonction de distribution de X. De plus, on a $\frac{d}{dx}F(x) = f(x)$

<u>Preuve</u> $P(X \leq x) = P(E)$ avec $E =]-\infty, x]$

Illustration du lien entre densité et fonction de répartition

Soit A=[a,b]

$$
\begin{aligned}
\mathbb{P}(A) &= \mathbb{P}(a \leq X \leq b) = \mathbb{P}(X \leq b) - \mathbb{P}(X \leq a) \\
&= F_X(b) - F_X(a) = \int_{-\infty}^{b} f_X(x)dx - \int_{-\infty}^{a} f_X(x)dx \\
&= \int_{a}^{b} f_X(x)dx \\
&= \int_{A} f_X(x)dx
\end{aligned}
$$

Implication : $\int_{-\infty}^{+\infty} f_X(x)dx = 1$

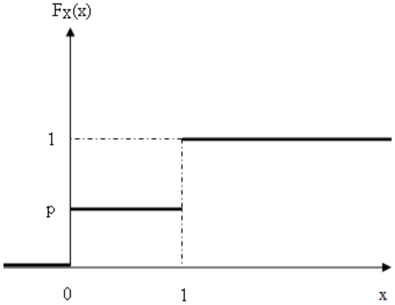
Propriétés de la fonction de distribution

1. F_X est croissante $s < t \Rightarrow F_X(s) \leq F_X(t)$
   (direct car $F(x) = \mathbb{P}(X \leq x)$)
2. F_X est continue à droite, c'est-à-dire que la limite à droite existe partout et $\lim\limits_{t \to a} F_X(t) = F_X(a)$
3. $\lim\limits_{x \to -\infty} F_X(x) = 0$
4. $\lim\limits_{x \to +\infty} F_X(x) = 1$

La fonction de distribution est définie pour toute v.a. (pas seulement continue). Si X est une variable discrète, alors F_X est discontinue aux points appartenant à l'image de X

Par exemple, si X est une variable aléatoire telle que $\mathbb{P}(X=0)=p$ et $\mathbb{P}(X=1)=1-p$, alors

- Si $x<0$, $F(x)=\mathbb{P}(X\leq x)=0$
- Si $x\in[0,1[$, $F(x)=\mathbb{P}(X\leq x)=\mathbb{P}(X=0)=p$
- Si $x\geq 1$, $F(x)=\mathbb{P}(X\leq x)=\mathbb{P}(X=0 \text{ ou } X=1)=1$

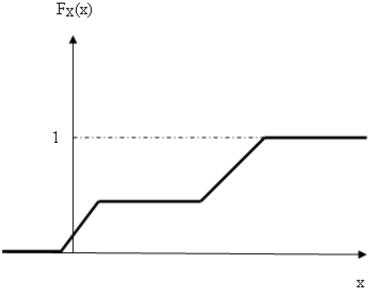


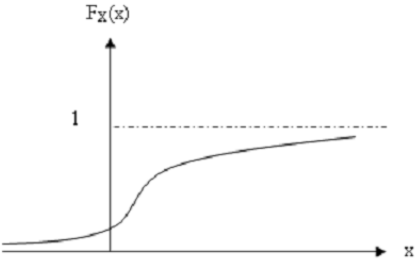
On peut maintenant donner une définition précise d'une variable aléatoire continue

Définition

Une variable aléatoire est dite continue si sa fonction de distribution est continue

Exemples





Les distributions (absolument) continues les plus importantes

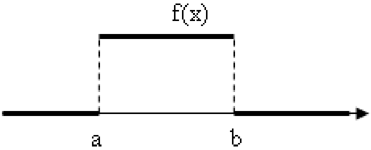
- La distribution unifome continue
- La distribution normale (ou gaussienne)

La distribution uniforme continue (1)

Comme dans l'exemple du "pointeur", il est parfois utile de considérer que chaque réalisation a la même probabilité. C'est l'intérêt de la loi uniforme continue.

On dit que X suit une loi uniforme sur l'intervalle $[a, b]$, noté $\mathcal{U}([a, b])$ quand :

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



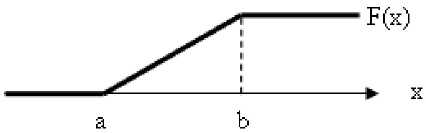
Remarque : On a bien $\int_{-\infty}^{+\infty} f(x)dx = 1$

La distribution uniforme continue (2)

Ainsi,

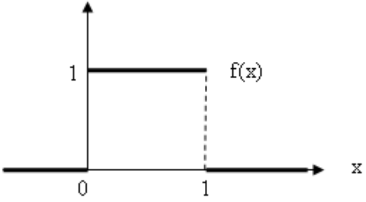
$$F(x) = \int_{-\infty}^{x} f(t)dt = \begin{cases} 0, \text{ si } x < a \\ \frac{x-a}{b-a} & \text{si } x \in [a, b] \\ 1, \text{ si } x > b \end{cases}$$

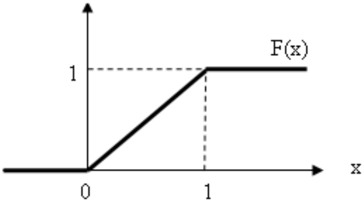
Remarque : $F(x)$ est absolument continue



La distribution uniforme continue (3)

Le cas le plus simple est celui de la loi $\mathcal{U}(0,1)$ (comme dans l'exemple du pointeur)





Dans cette exemple, on remarque bien que l'aire sous $f(x)$ est égale à 1

La distribution normale ou gaussienne (1)

Considérons maintenant la loi continue la plus importante : la loi normale. Le Théorème Central Limite (que nous aborderons plus loin) indique que sous des conditions générales, la distribution d'une somme de variables aléatoires peut être finement approximée par une densité continue specifique : la densité normale (ou gaussienne).

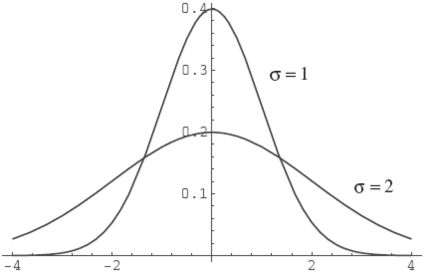
La densité normale de paramètre μ et σ s'écrit

$$f(x) = \frac{1}{\sqrt{2\Pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

On note alors $X \sim \mathcal{N}(\mu, \sigma)$,

La distribution normale ou gaussienne (2)

En représentant cette densité (pour $\mu = 0$)



on remarque que

- μ represente le "centre" de la densité (on verra plus loin qu'il s'agit de l'éspérance)
- σ mesure l'"étalement" de la densité (on verra plus loin qu'il s'agit de l'écart type)

La distribution normale ou gaussienne (3)

Ainsi, la loi normale est extrémement important car elle permet notamment d'illustrer les erreurs de mesure.

Il est par ailleurs possible de montrer que

$$\int_{-\infty}^{\infty} f(x)dx = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\Pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx = 1$$

Enfin, la fonction de distribution normale s'écrit

$$F(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\Pi\sigma^2}} e^{-\frac{(t-\mu)^2}{2\sigma^2}} dt$$

Notez qu'on ne peut pas écrire $F(x)$ plus simplement ce qui va amener certaines difficultés...