§4.3 Espérance d'une variable aléatoire discrète

Déf Soit X une variable aléatoire discrète et p sa fonction de masse. Si

$$\sum_{x\in\mathbb{R}} |x|p(x) \text{ converge}, \qquad (*)$$

alors on définit l'<u>espérance mathématique</u> de X par

$$\mu_X := \mathbb{E}(X) = \sum_{x\in\mathbb{R}} x\, p(x).$$

La condition $(*)$ dit que la série de l'espérance est absolument convergente. Pour le cours, on travaillera toujours avec des variables discrètes qui vérifient $(*)$, donc on ne demandera jamais de vérifier que cette condition est satisfaite.

Exemples sur les diapo semaine 4.

Interprétations

- **Cas équiprobable** : Disons que $\Omega = \{x_1, \dots, x_n\}$ et X est uniforme, c'est-à-dire $p(x_j) = 1/n \ \forall j$. Alors

$$\mathbb{E}(X) = \sum_{j=1}^{n} x_j\, p(x_j) = \sum_{j=1}^{n} x_j \frac{1}{n} = \frac{x_1 + x_2 + \cdots + x_n}{n} = \text{moyenne arithmétique}$$

Dans le cas général, $E(X)$ est une moyenne pondérée

$$\mathbb{E}(X) = x_1\, p(x_1) + x_2\, p(x_2) + \cdots + x_n\, p(x_n) + \cdots$$

pondération
(ou poids)

- **Statistique** : On répète l'expérience aléatoire un très grand nombre de fois et à chaque fois, on note le résultat

(on y reviendra plus tard)

de X, par exemple $\overbrace{x_1, x_2, x_3, x_4, \dots}^{\text{données statistiques}}$ $(\operatorname{Im} X = \{y_1, \dots, y_k\})$

Soit $y_1, \dots, y_k$ les valeurs possibles de X,

n_j le nombre de fois que l'expérience donne y_j

et $n = n_1 + \dots + n_k$ le nombre total d'expériences effectuées.

Alors $\mathbb{P}(X = y_j) = p(y_j) \approx \dfrac{n_j}{n}$ $\left(\text{on verra que } \lim\limits_{n\to\infty} \dfrac{n_j}{n} = p(y_j)\right)$

et $\mathbb{E}(X) = y_1 p(y_1) + \dots + y_k p(y_k)$

$$\approx y_1 \frac{n_1}{n} + \dots + y_k \frac{n_k}{n}$$

$$= \frac{\overbrace{y_1 + y_1 + \dots + y_1}^{n_1\text{-fois}} + \overbrace{y_2 + \dots + y_2}^{n_2\text{-fois}} + \dots + \overbrace{y_k + \dots + y_k}^{n_k\text{-fois}}}{n}$$

C'est donc la moyenne arithmétique des données observées sur un nombre arbitrairement grand d'observations $(n \to \infty)$.

Ex : On lance deux dés 45 fois et on obtient les observations suivantes de la variable X = somme des dés et on obtient les données suivantes

Valeur de la somme	Nombre d'observation
2	3
3	1
4	5
5	3
6	6
7	11
8	4
9	4
10	5
11	3

12	0

$$E(X) \approx \frac{2\times 3+3\times 1+4\times 5+5\times 3+6\times 6+7\times 11+8\times 4+9\times 4+10\times 5+11\times 3+12\times 0}{45}$$

$$= 6,8\bar{4}$$

Vrai espérance

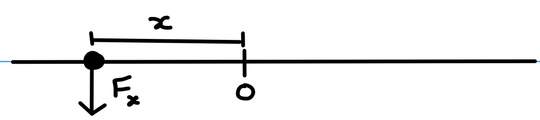
$$\mathbb{E}(X) = \frac{2\times 1+3\times 2+4\times 3+5\times 4+6\times 5+7\times 6+8\times 5+9\times 4+10\times 3+11\times 2+12\times 1}{36}$$

$$=7$$

C'est pour cela qu'on appelle l'espérance « la moyenne ».
En moyenne, on s'attend que la somme des dés vaut 7.

• **Physique.**

Si on voit $p(x)$ comme la quantité de masse au point x, alors $x\,p(x)$ est le moment de force par rapport à 0 (on prend la constante gravitationnelle $=1$) :



Le centre de masse μ est le point qui satisfait $\sum_x (x-\mu)\,p(x) = 0$.
(Les moments de force s'annulent autour de μ)

On a

$$\mathbb{E}(X-\mu) = \sum_x (x-\mu)\,p(x) \quad \text{et} \quad \mathbb{E}(X-\mu) = \mathbb{E}(X) - \mathbb{E}(\mu)$$
$$= \mu - \mu = 0$$

Donc $\mu = \mathbb{E}(X)$ est bien le centre de masse.

On reviendra aux propriétés de l'espérance au chapitre 5.

§4.4 Trois distributions discrètes classiques

Loi binomiale

Soit $\mathcal{E}$ une expérience aléatoire où il est seulement d'avoir un succès ou un échec avec $P(\text{succès}) = p$ et $P(\text{Échec}) = 1-p$.
On appelle une telle expérience une <u>épreuve de Bernoulli</u>.

On répète l'expérience n fois de façon indépendante et on pose
$X =$ nb de succès.

S = succès
E = echec.

Si on observe SSE, alors $X=3$.
$\mathbb{P}(SSE) = pp(1-p) = p^2(1-p)$
De même, si on observe SES, alors $X=3$ et
$\mathbb{P}(SES) = p(1-p)p = p^2(1-p)$.

De façon général, s'il y a k succès, chaque configuration aura une proba $p^k(1-p)^{n-k}$ de se realiser. Il y a $\binom{n}{k}$ configurations, donc

$$p(k) = \mathbb{P}(X=k) = \mathbb{P}(k \text{ succès}) = \binom{n}{k} p^k (1-p^k) \quad \text{pour } k \in \{0,1,\dots,n\}.$$

Def On dit que $X \sim \text{Binom}(p,n)$ si sa fonction de masse est

(*) $p_X(k) = \binom{n}{k} p^k (1-p)^{n-k}$ si $k \in \{0,1,2,\dots,n\}$ et $p_X(x) = 0$ sinon.

Ici, n = nb d'épreuve et $n \in \mathbb{N}$
p = proba. d'un succès et $0 \le p \le 1$.

Attention
p est le paramètre
et p_X est la fonction de masse

Vérifions que p dans (*) est bien une fonction de masse :

$p(x) \geq 0$ (clair)

et

$$\sum_{k=0}^{n} p(k) = \sum_{j=0}^{n} \binom{n}{k} p^k (1-p)^{n-k} \underset{\text{binôme de Newton}}{=} (p+1-p)^n = 1^n = 1$$

Proposition Si $X \sim \text{binom}(n,p)$, alors
$$\mathbb{E}(X) = np.$$

Dém.

$$\mathbb{E}(X) = \sum_{x\in\mathbb{R}} x\,p(x) = \sum_{k=0}^{n} k\,p(k)$$

$$= \sum_{k=0}^{n} k \binom{n}{k} p^k (1-p)^{n-k}$$

$$= p \sum_{k=1}^{n} k \binom{n}{k} p^{k-1} (1-p)^{n-k}$$

$(x+y)^n = \sum_{k=0}^{n} \binom{n}{k} x^k y^{n-k}$

On dérive en x des deux côtés

$n(x+y)^{n-1} = \sum_{k=1}^{n} k \binom{n}{k} x^{k-1} y^{n-k}$

$$= pn(p+1-p)^{n-1}$$

$$= pn. \qquad \square$$

En moyenne, il y a aura np succès.

Loi géométrique

On considère une série d'épreuves de Bernoulli, avec une proba de succès égale à p et indépedantes les unes des autres.

Soit X = nombre d'épreuve pour obtenir le premier succès.

P.ex. si on observe EEESESE..., alors $X=4$.

$X = k \iff$ on observe $k-1$ échec suivi de 1 succès, donc

$$\mathbb{P}(X=k) = \mathbb{P}(\underbrace{EE\cdots E}_{k-1 \text{ fois}}S) = \underbrace{(1-p)\cdots(1-p)}_{k-1 \text{ fois}}\,p = (1-p)^{k-1}p$$

Déf On dit que $X \sim \text{géo}(p)$ si sa fonction de masse est
$p_X(k) = (1-p)^{k-1}p$ si $k \in \mathbb{N}^*$ et $p_X(x) = 0$ sinon.
Le paramètre p est la proba. d'un succès et $0 \le p \le 1$.

Attention
p est le paramètre et p_X est la fonction de masse

Vérifions que p_X est une fonction de masse.

$p_X(x) \geq 0$ (clair)

$$\sum_{x \in \mathbb{R}} p_X(x) = \sum_{k=1}^{\infty} p_X(k) = \sum_{k=1}^{\infty} (1-p)^{k-1} p$$

$$= p \sum_{k=1}^{\infty} (1-p)^{k-1}$$

$$= p \sum_{n=0}^{\infty} (1-p)^n$$

(changement de variable $n = k-1$)

Série géométrique : si $-1 < r < 1$:

$$\sum_{n=0}^{\infty} r^n = \frac{1}{1-r}$$

$$= p \frac{1}{1-(1-p)}$$

$$= \frac{p}{p} = 1$$

Proposition Si $X \sim geo(p)$, alors $\mathbb{E}(X) = \frac{1}{p}$.

Dém. On a

$$\mathbb{E}(X) = \sum_{x \in \mathbb{R}} x p_X(x)$$

$$= \sum_{k=1}^{\infty} k p_X(k)$$

$$= \sum_{k=1}^{\infty} k (1-p)^{k-1} p$$

$$= p \sum_{k=1}^{\infty} k (1-p)^{k-1}$$

$$= p \frac{1}{(1-(1-p))^2}$$

$$= p \frac{1}{p^2} = \frac{1}{p}. \quad \square$$

Somme géométrique :

$$\frac{r^{n+1}-1}{r-1} = \sum_{k=0}^{n} r^k$$

On dérive chaque côté par rapport à r

$$\frac{(n+1) r^n (r-1) - (r^{n+1}-1)}{(r-1)^2} = \sum_{k=1}^{n} k r^{k-1}$$

Si $n \to \infty$, alors $r^n \to 0$ et $(n+1) r^n \to 0$, donc

$$\frac{0-(0-1)}{(r-1)^2} = \frac{1}{(1-r)^2} = \sum_{k=1}^{\infty} k r^{k-1}$$

P.ex. Si on lance un dé et succès = rouler 6, alors $p = 1/6$ et $\mu = 1/{1/6} = 6$, donc il faut en moyenne 6 lancers pour obtenir son premier succès.

Loi de poisson

Déf On dit que $X \sim$ poisson(ν) si X est discrète et sa fonction de masse est

$$p_X(k) = \begin{cases} e^{-\nu} \nu^k / k! \,, & \text{si } k \in \mathbb{N} \cup \{0\}, \\ 0 \,, & \text{sinon} \end{cases}$$

Ici, ν est un paramètre réel strictement positif.

Verifions que p_X est bien une fonction de masse :

$p_X(x) \geq 0$ (clair)

$$\sum_{x \in \mathbb{R}} p_X(x) = \sum_{k=0}^{\infty} p_X(k) = \sum_{k=0}^{\infty} e^{-\nu} \nu^k / k!$$

$$= e^{-\nu} \underbrace{\sum_{k=0}^{\infty} \frac{\nu^k}{k!}}_{\text{série de l'exponentielle}}$$

$$= e^{-\nu} e^{\nu}$$

$$= 1.$$

Proposition Si $X \sim$ Poisson(ν), alors $\mathbb{E}(X) = \nu$.

Ainsi, le paramètre de X représente sa moyenne.

Dém. On a

$$\mathbb{E}(X) = \sum_{x \in \mathbb{R}} x \, p_X(x) = \sum_{k=0}^{\infty} k \, p_X(k)$$

$$= \sum_{k=0}^{\infty} k e^{-\nu} \nu^k / k!$$

$$= e^{-\nu} \sum_{k=1}^{\infty} \frac{\nu^k}{(k-1)!}$$

$$= \nu e^{-\nu} \sum_{k=1}^{\infty} \frac{\nu^{k-1}}{(k-1)!}$$

$$= \nu e^{-\nu} \underbrace{\sum_{n=0}^{\infty} \frac{\nu^n}{n!}}_{\text{(Série de l'exponentielle)}} \quad \text{(changement de variable } n = k-1\text{)}$$

$$= \nu e^{-\nu} e^{\nu}$$

$$= \nu$$

On appelle parfois la loi de poisson la loi des événements rares, car elle apparaît comme cas limite de la loi binom(n,p) lorsque n est grand et p est petit.

Supposons que $X_n \sim \text{Binom}(n,p)$ avec $\nu = np$. Alors

$\hookrightarrow p = \frac{\nu}{n}$

$$\mathbb{P}(X_n = k) = \binom{n}{k} p^k (1-p)^{n-k}$$

$$= \frac{n!}{(n-k)!\,k!} \left(\frac{\nu}{n}\right)^k \left(1 - \frac{\nu}{n}\right)^{n-k}$$

$$= \underbrace{\frac{n!}{(n-k)!\,n^k}}_{\to 1} \frac{\nu^k}{k!} \underbrace{\left(1 - \frac{\nu}{n}\right)^n}_{\to e^{-\nu}} \underbrace{\left(1 - \frac{\nu}{n}\right)^{-k}}_{\to 1}$$

$$\longrightarrow e^{-\nu} \frac{\nu^k}{k!} \quad \text{lorsque } n \to \infty.$$

Ex. On considère l'intervalle de temps entre 15h et 16h.
On considère qu'il y a un succès si quelqu'un se met en file à la caisse dans cette intervalle. $\text{Binom}(1, p_1)$

On divise l'intervalle en deux : 15h à 15h30 et 15h30 à 16h, on a maintenant deux épreuves. $\text{Binom}(2, p_2)$

On divise l'intervalle en 3, ce qui donne 3 épreuves $\text{Binom}(3, p_3)$
et ainsi de suite.
⋮

Lorsqu'on a divisé l'intervalle une infinité de fois, X = « nb de personnes qui s'est mis file » est une loi de poisson avec $\nu = \lim_{n\to\infty} np_n$.

§4.5 Distribution absolument continue.

Déf Une variable aléatoire X est dite <u>absolument continue</u> si sa distribution est absolument continue : il existe $f : \mathbb{R} \to \mathbb{R}$ intégrable telle que $\forall a \leq b$

$$\mathbb{P}(a \leq X \leq b) = \int_a^b f(x)\,dx.$$

Remarque technique : en fait, il faut $\mathbb{P}(X \in B) = \int_B f(x)dx$ pour tout borélien $B \subseteq \mathbb{R}$.

On appelle f la <u>densité de probabilité</u> de X.

Propriétés de f :

1. $f(x) \geq 0 \quad \forall x \in \mathbb{R}$
2. $\int_{\mathbb{R}} f(x)dx = 1.$

Dém. 1. (Dans le cas où f est continue)

Pour tout $x \in \mathbb{R}$ et tout $h > 0$

$$\frac{1}{h}\mathbb{P}(x \leq X \leq x+h) \geq 0 \implies \frac{1}{h}\int_x^{x+h} f(t)\,dt \geq 0$$

$$\implies \lim_{h\to 0^+} \frac{1}{h}\int_x^{x+h} f(t)\,dt = f(x) \geq 0.$$

2. $\mathbb{P}(X \in \mathbb{R}) = 1 = \int_{\mathbb{R}} f(x)dx.$