
Probabilités et statistiques

2011 – 2012

Version 0.4

17 octobre 2011

Hubert PERROT

Table des matières

1	Espaces probabilisés	3
1.1	Ensemble fondamental – événement	3
1.2	Définition d’une fonction probabilité	4
1.2.1	Fonction probabilité	4
1.2.2	Méthode pratique pour définir la fonction probabilité	5
1.3	Probabilité conditionnelle – indépendance	6
1.3.1	Probabilités conditionnelles	6
1.3.2	Indépendance d’événements	6
1.3.3	Théorème de la multiplication	6
1.3.4	Formule des probabilités totales	6
1.4	Techniques de comptage – analyse combinatoire	7
1.4.1	Principe fondamental de comptage	7
1.4.2	Permutations	7
2	Variables aléatoires	10
2.1	V. A. à une dimension (définition, lois les plus usuelles, espérance, variance. . .)	10
2.1.1	Définitions et exemples	10
2.1.2	Loi de probabilité	11
2.2	Lois de probabilités les plus connues	14
2.2.1	v. a. discrètes	14
2.2.2	v. a. continues	15
2.2.3	Fonction d’une v. a.	18
3	L’espérance mathématique	21
3.1	V. A. à deux dimensions	22
4	Génération de nombres aléatoires	22

Pourquoi les probas ?

Parce que c'est une aide à la décision. C'est également une aide à l'analyse de données.

Consignes

- Ordinateurs tolérés ;
- Venir en cours et le suivre ;
- Poser des questions ;
- Devoirs écrits sans documents (et peut-être aussi sans calculatrices).

1 Espaces probabilisés

1.1 Ensemble fondamental – événement

Expérience aléatoire Une expérience qui, répétée dans les mêmes conditions ne donne pas toujours le même résultat.

- Tirer une carte dans un jeu de 52 cartes ;
- Compter le nombre de particules émises par un radio-élément entre un instant 0 et t ;
- Noter la distance entre le point d'impact d'une fléchette et le centre de la cible.

Ensemble fondamental L'ensemble fondamental, ou encore espace des épreuves, noté Ω . Il s'agit de l'ensemble des résultats de l'expérience.

- $\Omega = \{\text{As P, As K, As C, As T, } \dots, 2 \text{ P, } 2 \text{ K, } 2 \text{ C, } 2 \text{ T}\}$;
- $\Omega = \{0, 1, 2, 3, \dots, \infty\}$;
- $\Omega = \{x \in \mathbb{R}^+\} = [0, R]$ où R rayon de la cible.

Il est à noter que Ω peut être de dimension finie (cartes) ou infinie (particules ou fléchette).

Événement C'est un sous-ensemble de Ω . Notion **fondamentale** en probabilités.

- Événement $A = \text{« tirer un pique »} = \{\text{As P, R P, } \dots, 2 \text{ P}\}$;
- Événement $B = \text{« compter un nombre de particules émises »}$;
- Événement $C = \text{« approcher le centre de la cible de moins } \frac{R}{2} \text{ »}$.

Remarques :

A partir d'événements A et B , on peut construire d'autres événements.

Complémentaire noté A^c ou $\bar{A}$. Si on revient aux 52 cartes, A^c ou $\bar{A} = \{39 \text{ cartes autres que pique}\}$;

Union de deux événements $A \cup B$. L'événement A ou l'événement B se réalise, le « ou » n'étant pas exclusif. exemples : $A = \text{"tirer un pique"}$ $B = \text{"tirer un as"}$

$$A \cup B = \{AsK, AsT, AsC + 13 \text{ piques}\}$$

Intersection de deux événements $A \cap B$.

$$A \cap B = \text{l'événement } A \text{ et l'événement } B \text{ se réalisent.} = \{AsP\}$$

Événement A moins un événement B $A - B$. Donc ici, pour le jeu de 52 cartes, $A - B = \{RP, DP \dots 2P\}$. $A - B = A \cap B^c$.

2. Ω = événement certain $\emptyset$ = événement impossible
3. Un événement de type "tirer un as" est couramment assimilé au sous-ensemble $B = \{AsP, AsK, AsC, AsT\}$. Donc on peut :
- faire des dessins pour raisonner ;
 - utiliser les lois de la théorie des ensembles. $A \cup A = A$ $A \cap A = A$
- Associativité : $(A \cup B) \cup C = A \cup (B \cup C)$ Commutativité : $A \cup B = B \cup A$ Distributivité : $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$ Identité : $A \cup \emptyset = A$ $A \cup \Omega = \Omega$ $A \cap \emptyset = \emptyset$ $A \cap \Omega = A$ Complémentarité : $A \cup A^c = \Omega$ $(A^c)^c = A$ $A \cap A^c = \emptyset$ $\Omega^c = \emptyset$ $\emptyset^c = \Omega$

1.2 Définition d'une fonction probabilité

1.2.1 Fonction probabilité

Avant d'effectuer une expérience, on ne sait pas si un événement A sera réalisé ou non. La donnée d'une probabilité permet de mesurer les chances pour que l'événement A se réalise.

Définition Soit une expérience associée à un ensemble fondamental Ω et $\mathcal{a}$, l'ensemble des événements. p est une probabilité sur $(\Omega, \mathcal{a})$ si l'application p de $\begin{matrix} \mathcal{a} \rightarrow [0, 1] \\ A \rightarrow p(A) \end{matrix}$ vérifie :

$$\begin{aligned} p(\Omega) &= 1 \\ p(A \cup B) &= p(A) + p(B) \text{ si } A \text{ et } B \text{ disjoints } A \cap B = \emptyset \\ p\left(\bigcup_{i=1}^{+\infty} A_i\right) &= \sum_{i=1}^{\infty} p(A_i) \quad \text{pour des événements disjoints deux à deux} \end{aligned}$$

Propriétés

- $p(\emptyset) = 0$;
- $p(A^c) = 1 - p(A)$. Démonstration :
On a $A \cup A^c = \Omega$ donc $p(A \cup A^c) = p(\Omega) = 1 = p(A) + p(A^c)$ car A et A^c sont disjoints = $1 \Rightarrow p(A^c) = 1 - p(A)$;
- $p(A \cup B) = p(A) + p(B) - p(A \cap B)$ pour A et B quelconques. On a $A \cup B = (A - B) \cup B$ et

$$\begin{aligned} A - B \text{ et } B \text{ sont disjoints, donc } p((A - B) \cup B) &= p(A - B) + p(B) \\ &= p(A) - p(A \cap B) + p(B) \\ &= p(A) + p(B) - p(A \cap B) = p(A \cup B) \end{aligned}$$

Exemple Un jeu de 52 cartes, on veut trouver la probabilité de tirer un as.

- On va définir un événement élémentaire.
On le note A = "tirer une carte donnée". On va chercher $p(A) = ?$. On a Ω = "réunion de 52 événements élémentaires disjoints". Or $p(\Omega) = 1 = \sum$ probabilité de chaque événement élémentaire = 52 x probabilité d'une carte donnée car on suppose l'équiprobabilité des événements donc $p(A) = p(\text{"tirer une carte donnée"}) = \frac{1}{52}$
- On a un événement E = "tirer un as" et on veut $p(E) = ?$, or $E = \{AsK\} \cup \{AsT\} \cup \{AsC\} \cup \{AsP\}$

$$\begin{aligned} \text{Donc } p(E) &= p(\{AsK\} \cup \{AsT\} \cup \dots) = p(AsK) + p(AsT) + p(AsC) + p(AsP) \\ &= p(A) + p(A) + p(A) + p(A) = \frac{4}{52} \end{aligned}$$

On remarque que $p(E) = \frac{\text{nombre cas favorables}}{\text{nombre cas possibles}}$, uniquement si équiprobables.

1.2.2 Méthode pratique pour définir la fonction probabilité

Cas où Ω est de dimension finie On a $\Omega = \{\omega_1, \omega_2, \dots, \omega_n\}$ avec ω_i événements élémentaires.

Méthode : A chaque événement élémentaire ω_i , on va associer un nombre réel p_i , appelé probabilité de ω_i telle que $0 \leq p_i \leq 1$ et $\sum_{i=1}^n p_i = 1$.

La probabilité d'un événement E est la somme des probabilités des ω_i qu'il contient.

Exemple : $p_1 = p_2 = \dots = p_{52} = \frac{1}{52}$.

et $p(E) = p(\text{"tirer un as"}) = p_{AsK} + p_{AsC} \dots = \frac{4}{52}$

Cas particulier Si ω_i sont équiprobables

$$\begin{aligned} \text{Dans ce cas, comme } \sum_{i=1}^n p_i = 1 &\Rightarrow n \times p_i = 1 \\ &\Rightarrow p_i = \frac{1}{n} \end{aligned}$$

$$\text{d'où } p(E) = \frac{\text{nombre d'événements } \omega_i \text{ contenus dans } E}{\text{nombre d'événements } \omega_i \text{ contenus dans } \Omega} = \frac{\text{nombre cas favorables}}{\text{nombre cas possibles}}$$

Cas où Ω est infini mais dénombrable On a toujours $\Omega = \{\omega_1, \omega_2, \dots, \omega_\infty\}$

A chaque événement élémentaire ω_i , on associe la probabilité p_i avec $0 \leq p_i \leq 1$ et $\sum_{i=1}^\infty p_i = 1$. Pas d'événements équiprobables dans ce cas.

Exemple : On lance une pièce de monnaie et on compte le nombre de lancers avant d'avoir pile.

$$\Omega = \{1, 2, 3, \dots, \infty\}$$

On considère $\omega_i = \text{"nombre de lancers"}$.

On a $p_1 = p(\text{événement } \omega_1) = \frac{1}{2}$.

$$p_2 = \frac{1}{2} \times \frac{1}{2} = \frac{1}{2^2}$$

$$p_3 = \frac{1}{2^3}$$

$$p_n = \frac{1}{2^n}$$

$$0 \leq p_i \leq 1 \quad \text{est vérifié}$$

$$\sum_{i=1}^\infty p_i = 1? \quad p_1 + p_2 + \dots + p_\infty \stackrel{?}{=} 1$$

$$\frac{1}{2} + \frac{1}{2^2} + \frac{1}{2^3} + \dots + \frac{1}{2^\infty} = \frac{1}{2} \left(1 + \frac{1}{2} + \frac{1}{2^2} + \dots \right) = \frac{1}{2} \times \frac{1}{1 - \frac{1}{2}} = 1$$

Cas les plus compliqués Démarche en trois étapes ;

- Décrire l'expérience à l'aide d'événement(s) élémentaire(s) simples ;
- Écrire l'événement dont on cherche la probabilité en fonction du ou des événements élémentaires ;
- Calculer la probabilité en utilisant les cinq ou six formules du cours (*cf. infra*).

1.3 Probabilité conditionnelle – indépendance

1.3.1 Probabilités conditionnelles

On appelle la probabilité conditionnelle de l'événement A sachant B la quantité $p(A|B) = \frac{p(A \cap B)}{p(B)}$.

Le rapport $\frac{p(A \cap B)}{p(B)}$ donne la probabilité relative de l'événement A dans l'espace réduit B .

Exemple : A = "tirer un roi", B = "tirer une carte supérieure à 10" (valet, dame, roi, as).

$$\begin{aligned} p(A|B) &= \frac{p(A \cap B)}{p(B)} = \frac{p(A)}{p(B)} \text{ car } A \subset B \\ &= \frac{\frac{4}{52}}{\frac{16}{52}} = \frac{4}{16} = \frac{1}{4} \end{aligned}$$

1.3.2 Indépendance d'événements

Définition Un événement B est dit indépendant d'un événement A si la probabilité que B se produise ne dépend pas du fait que A se produise ou non. Dans ce cas, $p(B|A) = p(B)$ si A et B indépendants ou encore $p(A \cap B) = p(A)p(B)$ si A et B **indépendants**.

Exemple Une urne avec 4 rouges et 3 jaunes. Tirage, avec remise, de 2 billes (c'est-à-dire que chacune est remise une fois tirée). On veut calculer la probabilité de tirer rouge et jaune (dans cet ordre-là).

- R_i = "tirer rouge au i^{e} tirage". J_i = "tirer jaune au i^{e} tirage" ;
- E = "tirer rouge puis jaune" = $R_1 \cap J_2$;
- $p(E) = p(R_1 \cap J_2) = p(R_1)p(J_2)$ car R_1 et J_2 indépendants, car tirage avec remise.
 $= \frac{4}{7} \times \frac{3}{7} = \frac{12}{49}$.

1.3.3 Théorème de la multiplication

On a $p(A_1 \cap A_2) = p(A_2|A_1)p(A_1)$

Exemple Urne avec 4 rouges et 3 jaunes, tirage **sans remise**. Probabilité d'avoir deux rouges ?

- R_i = « tirer rouge au i^{e} tirage » ;
- $E = R_1 \cap R_2$;
- $p(E) = p(R_1 \cap R_2) = p(R_2|R_1)p(R_1) = \frac{3}{6} \times \frac{4}{7}$

Extension à trois événements :

$$p(A_1 \cap A_2 \cap A_3) = p(A_3|A_2 \cap A_1)p(A_2|A_1)p(A_1)$$

1.3.4 Formule des probabilités totales

Soient $A_1, A_2, \dots, A_n$ n événements réalisant une partition de $\Omega \Leftrightarrow A_i$ sont disjoints deux à deux et $\bigcup A_i = \Omega$. Soit B , un événement quelconque de Ω .

$$\begin{aligned} \text{Alors, } B &= B \cap \Omega \\ &= B \cap (A_1 \cup A_2 \dots \cup A_n) \\ &= (B \cap A_1) \cup (B \cap A_2) \cup \dots \cup (B \cap A_n). \end{aligned}$$

$$\begin{aligned} d'o\grave{u} P(B) &= p((B \cap A_1) \cup (B \cap A_2) \cup \dots (B \cap A_n)) \\ &= p(B \cap A_1) + p(B \cap A_2) + \dots + p(B \cap A_n) \end{aligned}$$

$$p(B) = p(B|A_1)p(A_1) + p(B|A_2)p(A_2) + \dots + p(B|A_n)p(A_n)$$

Intérêt : On a $p(B)$ connaissant toutes les causes A_i et les probabilités de transition de A_i .

Formule de BAYES

$$p(A_i|B) = \frac{p(B|A_i)p(A_i)}{p(B|A_1)p(A_1) + p(B|A_2)p(A_2) + \dots + p(B|A_n)p(A_n)}$$

Démonstration :

$$p(A_i|B) \stackrel{\text{p. condi.}}{=} \frac{p(A_i \cap B)}{p(B)} = \frac{p(B \cap A_i)}{p(B|A_i)} = \frac{p(B|A_i)p(A_i)}{p(B|A_i)}$$

Exemple :

On a trois machines A , B et C qui produisent 50 %, 30 %, 20 % du nombre total de vis. Le % de vis défectueuses est respectivement 3, 4 et 5 %

- On prend une vis au hasard, quelle est la probabilité quelle soit défectueuse ?
 - M = "la vis est défectueuse". A , B et C = "la vis est produite par A , B , C " ;
 - $A \cup B \cup C$ partition de Ω . Probabilité totale : $p(M) = p(M|A)p(A) + p(M|B)p(B) + p(M|C)p(C)$.
 - $p(M) = 0,03 \times 0,5 + 0,04 \times 0,3 + 0,05 \times 0,2 = 3,7 \%$
 - ;
 - Si une vis est défectueuse, probabilité qu'elle soit produite par la machine A ?

1.4 Techniques de comptage – analyse combinatoire

L'objectif est de compter rapidement le nombre d'éléments d'un ensemble et le nombre de résultats possibles d'une expérience.

1.4.1 Principe fondamental de comptage

Soit une expérience avec p étapes successives et indépendantes.

Pour la première étape, on va avoir n_1 résultats possibles ; Pour la deuxième étape, n_2 résultats possibles ; Pour la p ème étape, n_p étapes possibles.

Le nombre de résultats possibles de l'expérience vaut $N = n_1 \times n_2 \times n_3 \times \dots \times n_p$.

Exemple Soit un jeu, avec trois cartes notées A , B et C et deux cartes notées 1 et 2. On imagine que l'on va tirer une première carte dans les lettres et une seconde dans les chiffres.

Quel est le nombre de résultats possibles de l'expérience ?

1^{er} tirage

On tire soit la carte A , soit la carte B , soit la carte C

2^e tirage, A peut être associé avec 1 ou 2, B , avec 1 ou 2...

Il y a six résultats possibles, la formule est la suivante : $N = n_1 \times n_2 = 3 \times 2 = 6$

1.4.2 Permutations

Définition Une permutation de n objets, notés $a_1, a_2, a_3, \dots, a_n$ est un arrangement de ces n objets dans un **ordre donné**.

Exemple

$a_1 a_2 a_3 \dots a_n$ et $a_2 a_1 a_3 \dots a_n$ sont deux permutations.

Permutations sans répétition Objets discernables. Alors le nombre de permutations :

$$N = n!$$

Exemple On veut estimer le nombre de mots de trois lettres formés avec les lettres a, b et c. Cela équivaut au nombre de permutations.

1^{re} solution abc acb cab cba bac bca, il y a six mots donc six permutations ;

2^e solution On va utiliser le principe du paragraphe 1. On imagine une expérience où l'on va remplir un tableau avec, à l'état initial, trois cases vides. L'expérience, en trois étapes, consiste à remplir ce tableau avec les trois lettres.

$$\begin{array}{c|c|c} 3 & 2 & 1 \\ 1^{\text{er}} & 2^{\text{e}} & 3^{\text{e}} \\ n_1 & n_2 & n_3 \end{array}$$

Si on applique le principe fondamental de comptage, on aura $N = n_1 \times n_2 \times n_3 = 3 \times 2 \times 1 = 3! = 6$.

3^e solution On utilise directement la formule $N = n!$.

Permutation avec répétitions Soient n objets avec α , β et γ objets indiscernables. Alors le nombre de permutations est :

$$N = \frac{n!}{\alpha! \beta! \gamma!}$$

Exemples On va prendre trois lettres : aab, trois objets dont deux sont indiscernables, donc le nombre de permutations :

aab aba baa, donc trois permutations.

Par rapport au cas précédent : $\frac{3!}{2!} = \frac{6}{2} = 3$.

Combien peut-on former de mots avec les lettres du mot « electricite » (sans accents) ?

Si toutes ces lettres étaient discernables : $N = 11!$. Donc, dans notre cas : $\frac{11!}{3!2!2!2!}$ (car le « e » apparaît trois fois, le « t » deux fois...).

Arrangements

Définition Soient n objets, un arrangement de p objets ($p < n$) est obtenu en tirant p objets parmi n objets et en **les rangeant**. On va chercher à calculer le nombre d'arrangements ?

Exemple On considère quatre lettres a, b, c et d c'est-à-dire quatre objets, on veut trouver le nombre d'arrangements constitués de deux lettres.

1^{re} solution ab ac ad bc bd cd ba ca da cb db dc, donc douze arrangements ;

2^e solution On va réutiliser le principe fondamental de comptage en partant d'un tableau vide de deux cases.

$$\begin{array}{c|c} 4 & 3 \\ N = 4 \times 3 = 12. \end{array}$$

On peut donc généraliser :

$$p \text{ cases vides } \quad n \quad \left| \quad n-1 \quad \right| \quad \dots \quad \left| \quad n-p+1 \quad \right|$$

$$\text{d'où } N = n \times (n-1) \times (n-2) \times \dots \times (n-p+1)$$

Nombre d'arrangements de p objets parmi n vaut :

$$N = n(n-1)(n-2)\dots(n-p+1) = A_n^p$$

Combinaisons

Définition Soient n objets :

Une combinaison de p objets est obtenue en tirant p objets parmi n , sans notion d'ordre.

Exemple Quatre lettres a, b, c et d :

$ab \ ac \ ad \ bc \ bd \ cd$, donc six combinaisons.

Donc ici, le nombre de combinaisons :

$$\frac{12}{2} = \frac{\text{nombre d'arrangements}}{\text{nombre de permutations de } p \text{ objets}} = \frac{A_n^p}{p!} = C_n^p = \frac{n(n-1)\dots(n-p+1)}{p!} = \frac{n!}{p!(n-p)!}$$

Si on revient à l'exemple avec les quatre lettres :

$$C_4^2 = \frac{4 \times 3}{2} = 6$$

Quelques formules sur les combinaisons

$$C_n^0 = 1$$

$$C_n^1 = n$$

$$C_n^p = C_n^{n-p}$$

$$C_n^p = C_{n-1}^p + C_{n-1}^{p-1}$$

Formule du binôme :

$$(a+b)^n = \left(\sum_{p=0}^n C_n^p a^p b^{n-p} \right)$$

Utilisation d'un arbre L'arbre permet d'énumérer tous les résultats possibles se déroulant en plusieurs étapes.

Exemple On considère une partie de tennis et deux joueurs de tennis, M et E . Le premier qui gagne deux sets de suite ou un total de trois sets gagne la rencontre.

De combien de façons peut se dérouler la partie ?

1 ^{er} set	2 ^e set	3 ^e set	4 ^e set	5 ^e set
M	M ¹	M	M ³	M ⁴
	E		E	E ⁵
		E ²		

Donc $5 \times 2 = 10$ déroulements possibles.

On considère que M et E sont de force égale. M gagne le 1^{er} set. Quelle est la probabilité pour que M gagne la partie ?

Événement élémentaire :

$$M_i = \text{« } M \text{ gagne le } i^{\text{e}} \text{ set »}$$

Événement dont on cherche la probabilité :

$$\text{« } M \text{ gagne la partie »} = \{M_2 \cup (\bar{M}_2 \cap M_3 \cap M_4) \cup (\bar{M}_2 \cap M_3 \cap \bar{M}_4 \cap M_5)\}$$

$$\begin{aligned} P(M \text{ gagne}) &= p(M_2 \cup (\bar{M}_2 \cap M_3 \cap M_4) \cup (\bar{M}_2 \cap M_3 \cap \bar{M}_4 \cap M_5)) \\ &= p(M_2) + p(\bar{M}_2 \cap M_3 \cap M_4) + p(\bar{M}_2 \cap M_3 \cap \bar{M}_4 \cap M_5) \text{ événements sont disjoints} \\ &= p(M_2) + p(\bar{M}_2) p(M_3) p(M_4) + p(\bar{M}_2) p(M_3) p(\bar{M}_4) p(M_5) \text{ car indépendants} \end{aligned}$$

$$p(M \text{ gagne}) = \frac{1}{2} + \left(\frac{1}{2}\right)^3 + \left(\frac{1}{2}\right)^4 = \frac{11}{16}$$

2 Variables aléatoires

2.1 V. A. à une dimension (définition, lois les plus usuelles, espérance, variance...)

2.1.1 Définitions et exemples

Soit une expérience aléatoire et Ω son ensemble fondamental.

Définition Une application $X : \Omega \rightarrow \mathbb{R}$ est une v. a. discrète si l'ensemble des valeurs prises par X est fini ou infini mais dénombrables, ou encore $X(\omega) = \{x_1, x_2, x_3 \dots x_n \dots\}$.

La v. a. est dite **continue** si l'ensemble des valeurs prises par $X(\omega)$ est un intervalle de $\mathbb{R}$ ou $\mathbb{R}$ en entier.

Exemple On lance deux dés, on va définir comme v. a. $X = \text{« somme des points obtenus »}$, alors X est une application de $\Omega \rightarrow \{2, 3, 4, \dots, 12\}$ où $\Omega = \{36 \text{ couples } (i, j) \text{ avec } 1 \leq i \leq 6 \text{ et } 1 \leq j \leq 6\}$.

Dans ce cas X est une v. a. **discrète**.

Remarques A partir d'une expérience aléatoire, on peut définir une infinité de v. a. Dans l'exemple des deux dés, X peut être la somme des carrés ou des cubes des points des deux dés.

Avec les v. a. discrètes, on va calculer $P(X = x)$ ou $P(X = x_i)$, avec x et x_i la valeur numérique attribuée à X , et X le nom de la v.a.

Par exemple, on va chercher $P(X = 2) = \frac{1}{36}$ (cas des deux dés).

Exemple : taille d'un individu

Ω = l'ensemble des habitants de la planète

On prend une personne au hasard. L'application $X : \Omega \rightarrow \mathbb{R}$
 $\omega = \text{« choix de la personne »} \rightarrow X(\omega) = \text{« taille de cette personne »}$
est une v. a. discrète ou continue.

Tout va dépendre de la précision de la mesure de taille.

- Si on prend le cm : $X(\omega)$ varie de 40 à 240 cm $\Rightarrow$ 200 valeurs discrètes possibles ;
- Si on prend le μm : $X(\omega)$ varie de $0,4 \cdot 10^6 \mu\text{m}$ à $2,4 \cdot 10^6 \mu\text{m} \Rightarrow 2 \cdot 10^6$ valeurs ; a priori discret, mais assez difficile de traiter cette v. a. avec les outils classiques, car il faudrait estimer $P(X = x)$ pour $2 \cdot 10^6$ valeurs.

Alors, il paraît plus simple de considérer cette v. a. comme continue et de travailler sur des intervalles. On supposera que X est compris entre 0,4 m et 2,4 m.

Remarques Peut-on avoir $P(X) = x_0$ si x_0 vaut exactement 1,720000... En fait, $P(X = x_0)$ car personne n'a exactement cette taille.

Par contre, $P(1,7 < X < 1,72) \neq 0$ et on pourra la calculer.

Avec les v. a. continues, on va travailler sur des intervalles $[a, b]$ et calculer $P(a < X < b)$.

Exemple : Durée de vie d'un composant Soit Ω = ensemble d'un grand nombre de composants « identiques ».

L'application $X : \Omega \rightarrow \mathbb{R}$
 $\omega = \text{« choix du composant »} \rightarrow X(\omega) = \text{« durée de vie d'un composant »}$
est une v. a. continue.

2.1.2 Loi de probabilité

Cas d'une variable aléatoire discrète Soit X , une v. a. discrète qui prend comme valeurs $x_1, x_2, \dots, x_n$.

Définition La loi de probabilité de X est l'ensemble des valeurs des probabilités $P(X = x_i)$ ou $P(X = x)$ ou $P(\{\omega | X(\omega) = x_i \text{ ou } x\})$ et qui satisfait :

- $0 \leq P(X = x) \leq 1$;
- $\sum_x P(X = x) = 1$ **Condition de normalisation.**

Exemple des deux dés X = « somme des points ».

Quelle est la loi de probabilité de X ?

1. Trouver les valeurs de X ;
 2. Calculer $P(X = x)$;
 3. Vérifier la condition de normalisation.
1. $X = \{2, 3, \dots, 12\}$

2. 36 couples équiprobables

Probabilité pour un couple : $\frac{1}{6} \times \frac{1}{6} = \frac{1}{36}$, donc événements équiprobables et $p = \frac{\text{nombre de cas favorables}}{\text{nombre de cas possibles}}$

x_i	Couples	$P(X = x_i)$
2	(1, 1)	$\frac{1}{36}$
3	(1, 2) (2, 1)	$\frac{2}{36}$
4	(1, 3) (2, 2) (3, 1)	$\frac{3}{36}$
5	(1, 4) (2, 3) (3, 2) (4, 1)	$\frac{4}{36}$
6	(1, 5) (2, 4) (3, 3) (4, 2) (5, 1)	$\frac{5}{36}$
7	(1, 6) (2, 5) (3, 4) (4, 3) (5, 2) (6, 1)	$\frac{6}{36}$
8	...	$\frac{5}{36}$
9	...	$\frac{4}{36}$

3. Vérifier : $\frac{1+2+3+4+\dots}{36} = \frac{36}{36} = 1$

$A_i = \text{« tirer } i \text{ avec le dé } A \text{ »}$

$B_i = \text{« tirer } i \text{ avec le dé } B \text{ »}$

$$P(X = 2) = P(A_1 \cap B_1) = P(A_i) P(B_1) = \frac{1}{36}$$

$$\begin{aligned} P(X = 3) &= P((A_1 \cap B_2) \cup (A_2 \cap B_1)) \\ &= P(A_1 \cap B_2) + P(A_2 \cap B_1) \quad \text{car disjoints} \\ &= \frac{1}{36} + \frac{1}{36} = \frac{2}{36} \end{aligned}$$

Remarque On peut définir $P(X \leq 3) = P(X = 2) + P(X = 3) = \frac{3}{36} = \frac{1}{12}$.
On peut aussi définir $P(3 \leq X \leq 4) = P(X = 3) + P(X = 4) = \frac{2}{36} + \frac{3}{36} = \frac{5}{36}$
 $P(3 < X < 4) = 0$

Cas d'une v. a. continue On va définir une **densité de probabilité**.

Définition Soit X , une v. a. continue de $\begin{matrix} \Omega & \rightarrow & [a, b] \subset \mathbb{R} \\ \omega & \rightarrow & X(\omega) \end{matrix}$

Soit $x_0 \in [a, b]$ ($P(X = x_0) = 0$), la probabilité que X soit compris entre x_0 et $x_0 + dx$ s'écrit :

$$P(x_0 \leq X \leq x_0 + dx) = f(x_0) dx$$

La fonction $f(x_0)$ est la densité de probabilité de X .

Si on fait une représentation de $f(x)$:

cf. Marc || Sébastien

$$\begin{aligned} \text{Donc } f(x_0) dx &= \text{Aire hachurée} \\ &= P(x_0 \leq X \leq x_0 + dx) \end{aligned}$$

Propriétés

– $f(x) \geq 0$ car $f(x) dx$ est une probabilité = $P(x \leq X \leq x + dx)$

- $P(x_1 \leq X \leq x_2) = ?$
alors

$$P(x_1 \leq X \leq x_2) = \int_{x_1}^{x_2} f(x) dx$$

- Dans le cas continu, $P(X = x_0) = 0$ donc $P(x_1 \leq X \leq x_2) = P(x_1 < X < x_2)$
-

$$P(a < X < b) = 1 \text{ donc } \int_a^b f(x) dx \text{ ou } \int_{-\infty}^{+\infty} f(x) dx = 1 \quad \text{Condition de normalisation}$$

Fonction de répartition

Définition C'est la fonction F de $\mathbb{R} \rightarrow [0, 1]$
 $x \rightarrow P(X \leq x)$

Le cas discret est peu utilisé, cf. l'exemple avec les deux dés ; $P(X \leq 3) = F(3) = \frac{1}{12}$.
 Dans le cas continu, très utilisé

$$F(x) = P(X \leq x) = P(a \leq X \leq x) = \int_a^x f(t) dt$$

pour éviter la confusion avec la v. a. x qui correspond à la borne supérieure.

$$F'(x) = f(x)$$

$$\begin{aligned} P(x_1 < X < x_2) &= \int_{x_1}^{x_2} f(x) dx \\ &= \int_{x_1}^{x_2} f(t) dt \\ &= \int_{x_1}^a f(t) dt + \int_a^{x_2} f(t) dt \\ &= \int_a^{x_2} f(t) dt - \int_a^{x_1} f(t) dt \\ &= F(x_2) - F(x_1) \end{aligned}$$

$$\text{donc } P(x_1 < X < x_2) = F(x_2) - F(x_1)$$

Conclusion La loi de probabilité d'une v. a. est donnée :

- dans le cas discret, par la connaissance des $P(X = x)$
- dans le cas continu, par la donnée de la densité de probabilité, $f(x)$ ou celle de la fonction de répartition $F(x)$

2.2 Lois de probabilités les plus connues

2.2.1 v. a. discrètes

Loi uniforme Soit une v. a. X discrète qui prend les valeurs $x_1, \dots, x_n$, alors la loi de probabilité de X est uniforme si $P(X = x_i)$ est indépendant de i . Cela signifie que $P(X = x_i) = \text{constante}$.

$$\text{C. n. : } \sum_{i=1}^n P(X = x_i) = 1.$$

$$\Rightarrow \sum_{i=1}^n \alpha = 1 \Rightarrow n\alpha = 1 \Rightarrow \alpha = \frac{1}{n}$$

$$\text{Donc } P(X = x_i) = \frac{1}{n}$$

Remarque Si la v. a. X prend un nombre infini de valeurs, alors la loi ne peut être discrète.

Exemple Un dé, $X = \ll \text{le double du nombre tiré} \gg$. Loi de probabilité de X ?

X	2	4	6	8	10	12
$P(X = x)$	$\frac{1}{6}$	$\frac{1}{6}$	...	...	...	...

Loi binomiale On va partir d'un exemple. On considère une ligne de transmission qui a transmis N bits (0 ou 1). Chaque bit est mal transmis avec la probabilité p et bien transmis avec la probabilité $q = 1 - p$. On suppose l'indépendance des erreurs de transmission. Soit $X = \ll \text{nombre de bits mal transmis} \gg$. Trouver la loi de probabilité de X .

1. Trouver les valeurs de X , de 0 à N ;
2. $P(X = 0) = P(\ll \text{tous les bits bien transmis} \gg)$.

$B_i = \ll \text{le bit numéro } i \text{ est bien transmis} \gg$. $M_i = \ll \text{le bit numéro } i \text{ est mal transmis} \gg$.

$$P(X = 0) = P(B_1 \cap B_2 \cap \dots \cap B_N)$$

$$P(X = 0) = P(B_1) P(B_2) \dots P(B_N) = q^N = (1 - p)^N \text{ car événements indépendants.}$$

$$\begin{aligned} P(X = 1) &= P((M_1 \cap B_2 \cap B_3 \cap \dots \cap B_N) \cup (B_1 \cap M_2 \cap B_3 \cap \dots \cap B_N) \cup \dots \cup (B_1 \cap B_2 \cap \dots \cap B_{N-1} \cap M_N)) \\ &= P(M_1 \cap B_2 \dots) + P(B_1 \cap M_2 \dots) + \dots + P(B_1 \cap B_2 \dots \cap M_N) \text{ car disjoints} \\ &= P(M_1) P(B_2) + \dots + p(B_1) P(B_2) \dots \times P(M_N) \text{ car indépendances} \\ &= p(1 - p)^{N-1} + p(1 - p)^{N-1} + \dots \\ &= Np(1 - p)^{N-1} = C_N^1 p(1 - p)^{N-1} \end{aligned}$$

$$\begin{aligned} P(X = 2) &= P((M_1 \cap M_2 \cap B_3 \dots \cap B_N) \cup (M_1 \cap B_2 \cap M_3 \dots) \dots (B_1 \cap \dots M_{N-1} \cap M_N)) \text{ indépendants et disjoints} \\ &= P(M_1) P(M_2) P(B_3) \dots + \dots + P(B_1) P(B_2) \dots P(M_{N-1}) P(M_N) \\ &= p^2(1 - p)^{N-2} + p^2(1 - p)^{N-2} + \dots + p^2(1 - p)^{N-2} \\ &= C_N^2 p^2(1 - p)^{N-2} \end{aligned}$$

$$P(X = 3) = C_N^3 p^3(1 - p)^{N-3}$$

$$P(X = N) = p(M_1 \cap M_2 \cap \dots \cap M_N) = p^N$$

Si on généralise :

$$P(X = x) = C_N^x p^x (1 - p)^{N-x}$$

Loi binomiale de paramètres N et p . Notation : $B(N, p)$.

$$\sum P(X = x) = 1 = \sum C_N^x p^x (1 - p)^{N-x} = \sum p^x q^{N-x} = (p + q)^N = 1$$

Condition de normalisation vérifiée.

Utilisation de la loi Binomiale :

- Une expérience avec deux résultats possibles (Succès : p et échec : q) ;
- N réalisations successives ;
- N réalisations indépendantes.

Si X est le nombre de succès sur N réalisations, alors $P(X = x) = C_N^x p^x q^{N-x}$ avec $x : 0, 1, \dots, N$.

Loi de POISSON Il s'agit d'une extension de la loi binomiale pour N grand et p petit.

Alors $P(X = x) = \frac{\exp^{-\lambda} \lambda^x}{x!}$ avec $\lambda = Np$.

On peut valider cette formule théoriquement, mais c'est trop long, la flemme. Toutefois, on va montrer l'équivalence sur un exemple.

Exemple Ligne de transmission qui transmet 100 bits ($N = 100$). Probabilité de mauvaise transmission : $p = \frac{1}{100}$.

Soit $X =$ « le nombre de bits mal transmis ». Loi binomiale : $P(X = x) = C_{100}^x \left(\frac{1}{100}\right)^x \left(1 - \frac{1}{100}\right)^{100-x}$.

Loi de Poisson : $P(X = x) = \frac{\exp^{-1}}{x!}$.

x	0	1	2	...
Loi binomiale : $P(X = x)$	0,366	0,370	0,185	...
Loi de Poisson : $P(X = x)$	0,368	0,368	0,184	...

L'écart est faible entre les deux types de calculs.

Remarque La loi de Poisson est la plus utilisée.

Exemple, le nombre d'appels téléphoniques reçus par minute par un central téléphonique ; ou encore le nombre de fautes d'impression par page dans un grand texte.

2.2.2 v. a. continues

Loi uniforme On va utiliser la densité de probabilité, notée $f(x)$.

$$P(x \leq X \leq x + dx) = f(x) dx$$

$$P(x_1 \leq X \leq x_2) = \int_{x_1}^{x_2} f(x) dx$$

Densité de probabilité pour une loi uniforme. Dans ce cas, la densité de probabilité, $f(x)$ est constante. Si X est uniforme sur un intervalle $[a, b]$:

$$\int_a^b f(x) dx = 1 \quad \text{C. n.}$$

$$\Rightarrow \int_a^b k dx = 1 \Rightarrow k [x]_a^b = 1 \Rightarrow k = \frac{1}{b-a}$$

Donc $f(x) = \frac{1}{b-a}$ pour X v. a. uniforme $[a, b]$.
D'où la représentation :
cf. Sébastien

Remarque a et b sont finis ; pas de loi uniforme sur $\mathbb{R}$ tout entier.

Fonction de répartition

$$F(x) = P(X \leq x) = \int_a^x f(t) dt$$

$$\text{où } f(t) = f(x) = \frac{1}{b-a} \quad \text{si } T \text{ ou } X \text{ uniforme sur } [a, b]$$

$$\text{donc } F(x) = \int_a^x \frac{1}{b-a} dt = \frac{x-a}{b-a}$$

Donc pour une loi uniforme :

$$\begin{aligned} F(x) &= 0 & \text{si } x < a \\ F(x) &= \frac{x-a}{b-a} & \text{si } a \leq x \leq b \\ F(x) &= 1 & \text{si } x > b \end{aligned}$$

Remarque On peut aussi écrire :

$$F(x) = \frac{\text{longueur } [a, b]}{\text{longueur } [a, b]}$$

Exemple de la roue de loterie Roue de loterie avec Θ l'angle entre l'index et la verticale. On peut montrer que cet angle Θ suit une loi uniforme sur l'intervalle $[0, 2\pi]$. Cela revient à prendre au hasard une valeur entre 0 et 2π . La densité de probabilité $f(\Theta) = \frac{1}{2\pi}$.

Exemple du générateur de nombres aléatoires Celui-ci va délivrer des nombres, au hasard, sur un intervalle $[0, 1]$. On imagine que l'on va générer 100000 valeurs. Soit X valeurs au hasard entre 0 et 1, X suit-elle une loi uniforme ?

Graphique, voir Sébastien.

Cet histogramme représente la densité de probabilité à condition de normaliser l'axe des y .

Loi exponentielle

Densité de probabilité On sait que $f(x) = C \exp^{-\lambda x}$ pour $x \geq 0$, $f(x) = 0$ ailleurs. Avec C constante réelle déterminée et λ constante réelle positive caractéristique de cette loi.

Que vaut C ?

La condition de normalisation nous donne :

$$\begin{aligned} \int_{-\infty}^{+\infty} f(x) dx &= 1 \Rightarrow \int_0^{+\infty} C \exp^{-\lambda x} dx = 1 \\ &\Rightarrow C \left[-\frac{\exp^{-\lambda x}}{\lambda} \right]_0^{+\infty} = C \\ &\Rightarrow C \left[0 + \frac{1}{\lambda} \right] = 1 \Rightarrow \frac{C}{\lambda} = 1 \Rightarrow C = \lambda \end{aligned}$$

Loi exponentielle

$$f(x) = \lambda \exp^{-\lambda x} \text{ si } x \geq 0 = 0 \text{ ailleurs}$$

Représentation graphique cf. Marc & Sébastien

$$\text{Or } f(x_1) dx) P(x_1 \leq X \leq x_1 + dx)$$

On constate que cette probabilité est plus importante pour x_1 que pour x_2 . Donc, les valeurs de x faibles sont les plus fréquentes (probabilité plus grande).

Fonction de répartition

$$\begin{aligned} F(x) &= P(X \leq x) \\ &= \int_0^x f(t) dt = \int_0^x \lambda \exp^{-\lambda t} dt \\ &= X \left[\frac{-\exp^{-\lambda t}}{-\lambda} \right]_0^x = 1 - \exp^{-\lambda x} \end{aligned}$$

Donc

$$F(x) = 1 - \exp^{-\lambda x} \text{ si } x \geq 0 = 0 \text{ ailleurs}$$

Remarque Dans l'exemple précédent, loi uniforme, on a expliqué qu'on pouvait justifier cette loi en traçant un histogramme.

petit schéma

Équivalent au graphe de $f(x)$ en fonction de x pour une loi uniforme.

Que faire pour justifier une loi exponentielle.

Exemple On utilise cette loi exponentielle pour traiter un flux simple d'événements. Par exemple, l'arrivée dans une file d'attente.

Loi de Gausse ou loi Normale

Densité de probabilité On doit définir pour la v. a. X sa moyenne m et sa variance σ^2 , alors $f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp \frac{-(x-m)^2}{2\sigma^2}$ pour x réel.

Notation $N(m, \sigma^2)$

Représentation Pour ces v. a. de type Gaussien, on travaille avec une v. a. centrée et réduite. Ici, on va considérer que $m = 0$ et $\sigma = 1$. Dans ce cas, on passe à une variable notée U et $u = \frac{x-m}{\sigma}$, alors la nouvelle densité de probabilité : $f(u) = \frac{1}{\sqrt{2\pi}} \exp^{-\frac{u^2}{2}}$.

En fait, on a tabulé les valeurs de la fonction de répartition $F(u)$. (xf. TD). Cela nous permettra de calculer les probabilités de X .

Largeur à mi-hauteur cf. M. et S.

Valeur de $f(x)$ à mi-hauteur : $\frac{1}{2\sigma\sqrt{2\pi}}$, largeur à mi-hauteur $L = x_2 - x_1$.

Or $f(x) = \frac{1}{2\sigma\sqrt{2\pi}} = \frac{1}{\sigma\sqrt{2\pi}} \exp^{-\frac{(x-m)^2}{2\sigma^2}}$

$$\Rightarrow \frac{1}{2} = \exp^{-\frac{(x-m)^2}{2\sigma^2}} \Leftrightarrow \ln(2) = \frac{(x-m)^2}{2\sigma^2}$$

$$\Rightarrow (x-m)^2 = 2 \ln(2) \sigma^2$$

$$\Rightarrow |x-m| = \sqrt{2 \ln(2)} \sigma \Rightarrow x_1 = m - \sqrt{2 \ln(2)} \sigma \text{ et } x_2 = m + \sqrt{2 \ln(2)} \sigma$$

2.2.3 Fonction d'une v. a.

Soit X une v. a. avec X application de Ω dans $\mathbb{R}$. Soit h une fonction de $\mathbb{R}$ dans $\mathbb{R}$ qui a x associe $h(x)$ dans l'application :

$$\begin{array}{ccc} \Omega & \rightarrow & \mathbb{R} \\ \omega & \mapsto & h[X(\omega)] \end{array}$$

(discrète ou continue)

Exemples $h(X) = aX + b, X^2, \ln(X), \sin(X)$

Cas des v. a. discrètes On connaît X , $h(X)$, la loi de probabilité associée à X . On va chercher la loi de probabilité de la nouvelle v. a. $Y = h(X)$.

On va chercher $P(Y = y_1)$ et $P(Y = y_2)$ en fonction de $P(X = x)$?

$P(X = y_1) = P(X = x_1)$ changements bijectifs

$P(Y = y_2) = P(X = x_2) + P(X = x_3)$

Si on généralise :

$$P(Y = y_i) = \sum_{x_j \text{ avec } P(x_j)=y_i} P(X = x_j)$$

Exemple On lance un dé. Soit X la v. a. correspondant au nombre total de points obtenus.

$$\text{Soit } Y = P(X) = 2X - 1 \text{ [4]}$$

On va chercher la loi de proba. de Y .

$$P(X = x) = \frac{1}{6} \text{ loi uniforme}$$

→ valeurs de Y :

$$P(Y = 1) = P(X = 1) + P(X = 3) + P(X = 5) = \frac{1}{6} + \frac{1}{6} + \frac{1}{6} = \frac{1}{2}$$

Pour $Y = 3$:

$$P(Y = 3) = P(X = 2) + P(X = 4) + P(X = 6)$$

Fonction d'une v. a. : cas où X est une variable continue à densité

$$\Omega \xrightarrow{X} \mathbb{R} \xrightarrow{h} \mathbb{R} \quad Y = h(X)$$

Question Soit $f(x)$ la densité de la variable X . Quelle est la densité de Y ?

$$p(X \in [a, b]) = \int_a^b f(x) \, dx$$

$$p(Y \in [c, d]) = \int_c^d ? \, dy$$

Hypothèse h suffisamment régulière (au moins continue, et même dérivable). Alors soit h est croissante, soit elle est décroissante.

$$h^{-1}([c, d]) = [h^{-1}(c), h^{-1}(d)]$$

graphique, cf. Sébastien & Marc

$$\begin{aligned} p(Y \in [c, d]) &= p(h(X) \in [c, d]) = p(X \in [h^{-1}(c), h^{-1}(d)]) \\ &= \int_{h^{-1}(c)}^{h^{-1}(d)} \text{changement de module } y = h(x) \end{aligned}$$

$$x = h^{-1}(y)$$

$$dy = h'(x) \, dx \quad dx = \frac{1}{h' \circ h^{-1}(y)} \, dy$$

$$p(Y \in [c, d]) = \int_c^d \frac{f \circ h^{-1}(y)}{h' \circ h^{-1}(y)} \, dy$$

$$\text{Densité de } Y : \frac{f \circ h^{-1}}{h' \circ h^{-1}}$$

Deuxième méthode

$$f(x) dx = p(X \in [x, x + dx])$$

$$h \text{ croissante} \Rightarrow X \in [x, x + dx] \Leftrightarrow Y = h(X) \in [h(x), h(x + dx)]$$

$$p(X \in [x, x + dx]) = p(Y \in [h(x), h(x) + h'(x) dx]) = p(Y \in [y, y + dy]) = g(y) dy \text{ où } g \text{ est la densité de } Y$$

$$\text{On pose } y = h(x), dy = h'(x) dx.$$

$$\text{Donc } f(x) dx = g(y) dy \text{ donc } g(y) = f(x) \frac{dx}{dy} \text{ or } x = h^{-1}(y) \frac{dx}{dy} = \frac{1}{h' \circ h^{-1}(y)}$$

Conclusion

$$\text{On a bien } g(y) = \frac{f(h^{-1}(y))}{h' \circ h^{-1}(y)}$$

Cas où h est bijective continue décroissante dessin cf. Marc & Sébastien.

$$dy = h'(x) dx$$

$$h \text{ décroissante} \implies h'(x) dx < 0$$

$$f(x) dx = p(X \in [x, x + dx]) = p\left(Y \in \begin{matrix} \text{intervalle de longueur } |dy| \\ [y + dy, y] \end{matrix}\right) = g(y) |dy|$$

$$\text{où } dy = h'(x) dx < 0 \quad y = h(x)$$

$$\text{Donc } \boxed{f(x) dx = g(y) |dy|}$$

$$g(y) = f(x) \left| \frac{dx}{dy} \right| = \frac{f \circ h^{-1}(y)}{|h' \circ h^{-1}(y)|}$$

Exemple X variable à densité :

$$f(x) = \begin{cases} \frac{x^2}{81} & \text{si } x \in [-3, 6] \text{ « support de } X \text{ »} \\ 0 & \text{sinon} \end{cases}$$

$$\begin{aligned} h : \mathbb{R} &\implies \mathbb{R} \\ x &\implies h(x) = \frac{1}{3}(12 - x) \end{aligned}$$

$$Y = h(X) \text{ densité de } Y ?$$

$$h \text{ décroissante}$$

$$f(x)$$

$$y = \frac{1}{3}(12 - x) \Rightarrow x = 12 - 3y, \frac{dx}{dy} = -3$$

$$g(y) = f(x) \left| \frac{dx}{dy} \right| = 3f(x) = 3f(12 - 3y)$$

$$\begin{aligned} -3 &\leq 12 - 3y \leq 6 \\ \Leftrightarrow -15 &\leq -3y \leq -6 \\ \Leftrightarrow 2 &\leq y \leq 5 \end{aligned}$$

$$\text{donc } g(y) = \begin{cases} 3 \frac{(12-3y)^2}{81} & \text{si } x \in [2, 5] \text{ « support de } X \text{ »} \\ 0 & \text{sinon} \end{cases}$$

3 L'espérance mathématique

Soit X , une variable aléatoire. Espérance = « note de moyenne ».

Exemple 5 notes de maths : 7, 14, 7, 18, 16

$$\text{moyenne : } 7 \times \frac{1}{5} + 14 \times \frac{1}{5} + 7 \times \frac{1}{5} \dots = 7 \times \frac{2}{5} + 14 \times \frac{1}{5} + 18 \times \frac{1}{5} + 16 \times \frac{1}{5}$$

La moyenne des notes est pondérée par les fréquences relatives des notes. L'espérance est une moyenne des valeurs de X pondérée par leurs probabilités respectives.

Définition dans le cas où X est discrète Supposons que $X : \Omega \longrightarrow \{x_1, x_2, x_3, \dots\}$

L'espérance de X est le nombre $E(X) = \sum_i x_i p(X = x_i)$.

Cas où X est à densité $f(x)$

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

Remarque Dans le cas discret, si X a une infinité (dénombrable) de valeurs possibles, la définition de $E(X)$ n'a de sens que si la série est **absolument convergente** c'est-à-dire $\sum [x_i | p(X = x_i)] < +\infty$.

Interprétation de $E(X)$ Répéter l'expérience aléatoire un grand nombre de fois n .

X_i = le résultat de la i^{e} répétition.

$$\begin{aligned} E(X) &\approx \text{moyenne de } X_i \\ &\approx \frac{X_1 + X_2 + \dots + X_n}{n} \end{aligned}$$

graphique cf. Marc || Sébastien.

Exemple 6 tirages successifs à pile ou face.

X = nombre de « face » obtenu

X suit une loi binomiale de paramètres $(6, \frac{1}{2})$.

$$\begin{aligned} E(X) &= \sum_{i=0}^6 xp(X=x) = 0p(X=0) + 1p(X=1) + 2p(X=2) + \dots + 6p(X=6) \\ &= C_6^1 \left(\frac{1}{2}\right)^1 \left(\frac{1}{2}\right)^5 + 2C_6^2 \left(\frac{1}{2}\right)^2 \left(\frac{1}{2}\right)^4 + \dots + 6 \left(\frac{1}{2}\right)^6 \\ &= (6 + 2 \times 15 + 3 \times 20 + 4 \times 15 + 5 \times 6 + 6) \left(\frac{1}{2}\right)^6 \\ &= 3 \end{aligned}$$

Triangle de Pascal

Plus généralement, si X suit une loi binomiale de paramètres (n, p) , alors $E(X) = np$.

Vérifions-le $E(X) = \sum_{k=0}^n kp(X=k) = \sum kC_n^k p^k (1-p)^{n-k}$

Lemme

$$kC_n^k = nC_{n-1}^{k-1}$$

$$E(X) = \sum_{k=1}^n nC_{n-1}^{k-1} p^k (1-p)^{n-k} = np \sum_{k=1}^n C_{n-1}^{k-1} p^{k-1} (1-p)^{(n-1)-(k-1)}$$

On pose $q = k - 1$. On a $1 = (p + (1-p))^{n-1}$

$$E(X) = np \sum_{q=0}^{n-1} C_{n-1}^q p^q (1-p)^{(n-1)-q}$$

$$E(X) = np$$

Si X suit une loi uniforme sur $[a, b]$.

$$E(X) = \int_a^b x \frac{dx}{b-a} = \frac{1}{b-a} \left[\frac{x^2}{2} \right]_a^b = \frac{1}{2} \times \frac{b^2 - a^2}{b-a} = \frac{a+b}{2}$$

Premier cas h bijective (et continue)

3.1 V. A. à deux dimensions

4 Génération de nombres aléatoires