Probabilités

I Concept

Un premier exemple

Considérons un dé cubique à six faces numérotées de 1 à 6, et équilibré, c'est-à-dire pour lequel la masse est régulièrement répartie. Nous avons simulé sur tableur l'expérience qui consisterait à le lancer 10 000 fois. Les résultats sont consignés dans le tableau statistique qui suit. L'effectif donne le nombre de fois où une face est apparue et la fréquence est le rapport de ce nombre au nombre de lancers. Il est arrondi à 0,1 %

faces	1	2	3	4	5	6
effectif	1638	1632	1694	1654	1671	1711
fréquence	16,4%	16,3%	16,9%	16,5%	16,7%	17,1%

Un tableur possède une fonction générant un nombre aléatoire dans le segment $[0;1[$. La formule que nous avons utilisé sous Excel est : =ALEA(). Afin d'obtenir un nombre aléatoire donnant une des faces possibles d'un dé, nous l'avons modifiée en notant que :

$$0 \leq x < 1 \Rightarrow 0 \leq 6\,x < 6 \Rightarrow 0 \leq ent(6\,x) \leq 5 \Rightarrow 1 \leq ent(6\,x) + 1 \leq 6$$

où $ent(6\,x)$ désigne la partie entière de $6\,x$.

Nous avons donc entré dans une cellule la formule : = ENT(ALEA()*6)+1

puis recopié 9999 fois vers le bas cette formule et extrait à l'aide d'autres formules les informations d'effectif et de fréquence.

Au vu des résultats, on constate que les fréquences d'apparition des six faces sont très voisines, l'écart maximal étant de 0,4% et en relançant plusieurs fois cette simulation avec la touche F9 du clavier, il apparaît que l'écart maximal reste du même ordre.

Un second exemple :

Nous avons simulé sur tableur le lancer de deux dés équilibrés puis noté la somme des faces obtenues.

somme des dés	2	3	4	5	6	7	8	9	10	11	12
fréquence empirique	2,7%	5,4%	8,6%	11,1%	13,4%	16,4%	14,3%	11,1%	8,7%	5,3%	2,9%
fréquence théorique	2,8%	5,6%	8,3%	11,1%	13,9%	16,7%	13,9%	11,1%	8,3%	5,6%	2,8%

La fréquence empirique correspond à la fréquence observée dans la simulation. La fréquence théorique va être expliquée ci-après.

Là encore, en relançant la simulation plusieurs fois, on constate que les fréquences observées varient peu, de l'ordre de quelques dixièmes de pourcent. Voyons alors comment, sans faire d'expérience, les fréquences observées peuvent être prédites avec une bonne précision, pourvu que le nombre de lancers soit suffisant, ce qui formera ce qu'on appelle la **loi des grands nombres**.

Supposons pour fixer les idées qu'un dé est vert et l'autre rouge et présentons dans un tableau les différents résultats donnés par les différents couples de faces possibles.

	1	2	3	4	5	6
1	2	3	4	5	6	7
2	3	4	5	6	7	8
3	4	6	6	7	8	9
4	5	6	7	8	9	10
5	6	7	8	9	10	11
6	7	8	9	10	11	12

Si l'on conçoit que pour un grand nombre de lancers chaque couple de faces apparaît un même nombre de fois, la fréquence d'apparition d'une somme égalé à 2 sera celle d'apparition du couple (1 ;1) soit :

$$\frac{1}{36} \approx 2{,}8\%$$

La fréquence d'apparition d'une somme égale à 3 sera celle d'apparition des couples (1 ;2) et (2 ;1) soit :

$$\frac{2}{36} \approx 5{,}6\%$$

Et ainsi de suite. Nous pouvons donc rassembler dans un tableau ces fréquences dites théoriques.

Somme des dés	2	3	4	5	6	7	8	9	10	11	12
Fréquence théorique	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{3}{36}$	$\frac{4}{36}$	$\frac{5}{36}$	$\frac{6}{36}$	$\frac{5}{36}$	$\frac{4}{36}$	$\frac{3}{36}$	$\frac{2}{36}$	$\frac{1}{36}$
Valeur en %	2,8%	5,6%	8,3%	11,1%	13,9%	16,7%	13,9%	11,1%	8,3%	5,6%	2,8%

Nous observons que pour un nombre de lancers de l'ordre de plusieurs milliers, les fréquences empiriques (donc observées) sont très voisines (à quelques dixièmes de pourcent) des fréquences théoriques. En revanche, si le nombre de lancers n'est pas très grand (inférieur à la centaine par exemple), les fréquences observées fluctuent plus notablement autour des fréquences théoriques.

II Vocabulaire

Une **expérience aléatoire** est une expérience pouvant être reproduite (à l'infini pourrait on dire mais l'infini est une vue de l'esprit) et dans laquelle on s'intéresse à un résultat généralement représenté par un ou plusieurs nombres et appelé **issue.** L'ensemble des issues est appelé **univers** et est noté Ω. Le nombre d'issues possibles quand il est fini est appelé cardinal de l'univers et noté $Card(\Omega)$.

Dans le premier exemple, l'expérience aléatoire est le lancer d'un dé, les issues sont les numéros des faces, ainsi l'univers est :

$$\Omega = \{1,2,3,4,5,6\}$$

Dans le second exemple, l'expérience aléatoire est le lancer de deux dés, les issues sont les sommes pouvant être obtenues, ainsi l'univers est :

$$\Omega = \{2,3,4,5,6,7,8,9,10,11,12\}$$

Les fréquences observées pour les différentes issues quand le nombre de répétitions de l'expérience aléatoire devient grand (se fixer en référence 5000 répétitions) se rapprochent de valeurs fixes appelées fréquences théoriques (prévisibles par calcul bien souvent comme dans les deux exemples exposés) ou encore **probabilités élémentaires**.

On appelle **évènement**, un sous-ensemble de l'univers et on le désigne généralement par une lettre majuscule ou par un énoncé littéraire caractérisant les issues qui le composent.

Exemples :

Dans le premier exemple, nous pouvons considérer le sous ensemble des issues paires :

$$A = \{2,4,6\}$$

Cet évènement est caractérisé par l'énoncé : « Obtenir une face paire ». On pourra alors écrire :

$$A = \{2,4,6\} = \text{« Obtenir une face paire »}$$

Dans le second exemple, le même évènement sera :

$$A = \{2,4,6,8,10,12\} = \text{« Obtenir une face paire »}$$

On appelle **probabilité d'un évènement**, la somme des probabilités élémentaires de ses issues.

Ainsi, dans le premier exemple :

$$P(A) = \frac{1}{6} + \frac{1}{6} + \frac{1}{6} = \frac{3}{6} = \frac{1}{2}$$

Dans le second exemple :

$$P(A) = \frac{1}{36} + \frac{3}{36} + \frac{5}{36} + \frac{5}{36} + \frac{3}{36} + \frac{1}{36} = \frac{18}{36} = \frac{1}{2}$$

Lorsque, dans une expérience aléatoire, les issues ont toute la même probabilité, on dit qu'elles sont **équiprobables**. C'est le cas du premier exemple mais pas celui du second.

Dans le cas où toutes les issues sont équiprobables, la probabilité d'un évènement A est donnée par :

$$P(A) = \frac{Card(A)}{Card(\Omega)} = \frac{Nombre\ d'issues\ de\ A}{Nombre\ d'issues\ totales}$$

On définit les **évènements intersection et union** de deux évènements A et B. Reprenant les exemples précédents et ajoutant l'évènement :

$$B = "Obtenir\ un\ multiple\ de\ 3"$$

Nous avons pour le premier exemple :

$$A = \{2,4,6\} = \text{"Obtenir une face paire"}$$

$$B = \{6\} = \text{"Obtenir un multiple de 3"}$$

$$A \cap B = \{6\} = \text{"Obtenir une face paire et un mutiple de 3"}$$

$$A \cup B = \{2,4,6\} = \text{"Obtenir une face paire ou un mutiple de 3"}$$

Pour le second exemple :

$$A = \{2,4,6,8,10,12\} = \text{"Obtenir une face paire"}$$

$$B = \{3,6,9,12\} = \text{"Obtenir un multiple de 3"}$$

$$A \cap B = \{6,12\} = \text{"Obtenir une face paire et un mutiple de 3"}$$

$$A \cup B = \{2,3,4,6,8,9,10,12\} = \text{"Obtenir une face paire ou un mutiple de 3"}$$

Nous avons dans tous les cas la propriété suivante :

$$\boxed{P(A \cup B) = P(A) + P(B) - P(A \cap B)}$$

On définit l'évènement contraire d'un évènement *A* comme étant son complémentaire dans l'univers.

Nous avons ainsi pour le premier exemple :

$$A = \{2,4,6\} = \text{"Obtenir une face paire"}$$

$$\overline{A} = \{1,3,5\} = \text{"Obtenir une face impaire"}$$

Nous avons dans tous les cas la propriété suivante :

$$P(\overline{A}) = 1 - P(A)$$

On définit l'évènement contraire d'un évènement *A* dans un évènement *B* par le complémentaire de *A* dans *B*.

Ainsi, pour le second exemple :

$$A = \{2,4,6,8,10,12\} = \text{"Obtenir une face paire"}$$

$$B = \{3,6,9,12\} = \text{Obtenir un multiple de 3}$$

$$A \backslash B = \{2,4,8,10\} = \text{"Obtenir une face paire non multiple de 3"}$$

Nous avons dans tous les cas la propriété suivante :

$$P(A \backslash B) = P(A) - P(A \cap B)$$

On définit enfin les **évènements certains et impossibles par** respectivement l'univers et l'ensemble vide.

Ainsi, pour le premier exemple :

$$\Omega = \{1,2,3,4,5,6\} = \text{"obtenir une face paire ou impaire"}$$

$$\Omega = \emptyset = \text{"obtenir une face de numéro négatif"}$$

III <u>Fréquences conditionnelles, probabilités conditionnelles</u>

Reprenons l'expérience du lancer de deux dés équilibrés pour laquelle on s'intéresse à la somme des faces obtenues.

Supposons que nous ayons obtenu, sur 20 lancers, les sommes suivantes :

3 – 5 – 11 – 7 – 3 – 9 – 8 – 7 – 6 – 9 – 8 – 9 – 3 – 6 – 12 – 10 – 8 – 7 – 9 – 7

Reprenons les évènements A et B cités précédemment :

$$A = \{2,4,6,8,10,12\} = \text{"Obtenir une face paire"}$$

$$B = \{3,6,9,12\} = \text{Obtenir un multiple de 3}$$

A s'est réalisé 7 fois sur 20 donc la fréquence de réalisation de A est :

$$f(A) = \frac{7}{20}$$

B s'est réalisé 10 fois sur 20 donc la fréquence de réalisation de B est :

$$f(B) = \frac{10}{20}$$

A et B se sont réalisés simultanément 3 fois sur 20 donc la fréquence de réalisation de $A \cap B$ est :

$$f(A \cap B) = \frac{3}{20}$$

Parmi les 7 lancers où A s'est réalisé, B s'est réalisé 3 fois. On définit alors la fréquence de réalisation de B sachant A par la fréquence des réalisations de B dans celles de A soit :

$$f_A(B) = \frac{3}{7}$$

Reportant sur un schéma ces données :

$$\xrightarrow{f(A)} A \xrightarrow{f_A(B)} B$$

Nous voyons que nous avons la relation :

$$f(A) \times f_A(B) = f(A \cap B)$$

ou encore :

$$f_A(B) = \frac{f(A \cap B)}{f(A)}$$

Cela amène à définir la probabilité dite conditionnelle d'un événement B sachant un autre évènement A réalisé par :

$$\boxed{P_A(B) = \frac{P(A \cap B)}{P(A)}}$$

Elle s'interprète comme étant la fréquence des réalisations de B parmi les réalisations de A pour un grand nombre de répétitions de l'expérience aléatoire.

Elle possède la propriété facile à déduire :

$$\boxed{P_A(B \cup C) = P_A(B) + P_A(C) - P_A(B \cap C)}$$

IV Indépendance de deux évènements-indépendance mutuelle

1) Indépendance de deux évènements

On dit qu'un évènement B est indépendant d'un évènement A si on :

$$P_A(B) = P(B)$$

Dans ce cas on a aussi :

$$P_B(A) = P(A)$$

autrement dit B est indépendant de A

On dit alors que les évènements A et B sont indépendants.

Une caractérisation de l'indépendance résulte de sa définition :

Deux évènements A et B sont indépendants si et seulement si ils vérifient :

$$\boldsymbol{P(A \cap B) = P(A)\, P(B)}$$

Exemple :

Soit un univers formé par les 200 élèves d'un établissement dans lequel il y a 120 filles et 80 garçons et 60 fumeurs filles et garçons. On sélectionne un élève au hasard. Combien faut il qu'il y ait de garçons fumeurs et de filles fumeuses pour que les évènements $F =$ "$Obtenir\ un\ élève\ fumeur$" et $G =$ "$Obtenir\ un\ garçon$" soient indépendants ?

Réponse :

$$p(F) = \frac{60}{200} = \frac{3}{10} = 0{,}3$$
$$p(G) = \frac{80}{200} = \frac{4}{10} = 0{,}4$$

La condition est :

$$p(F \cap G) = \frac{3}{10} \times \frac{4}{10} = \frac{12}{100} = 0{,}12$$

Il doit donc y avoir 12% de garçons fumeurs soit $0{,}12 \times 200 = 24$

2) Indépendance mutuelle de trois évènements

On dit que trois évènements A, B et C sont indépendants s'ils vérifient :

$$\boldsymbol{P(A \cap B) = P(A)\,P(B)}$$

$$\boldsymbol{P(A \cap C) = P(A)\,P(C)}$$

$$\boldsymbol{P(B \cap C) = P(B)\,P(C)}$$

$$\boldsymbol{P(A \cap B \cap C) = P(A)\,P(B)\,P(B)}$$

autrement dit si les évènements sont indépendants deux à deux et si chacun est indépendant avec l'intersection des deux autres.

Exemple :

Dans une urne se trouvent une boule verte et deux boules rouges indiscernables au toucher. On tire une boule au hasard et on la remet. On recommence le procédé deux fois de suite. On s'intéresse aux trois boules tirées avec remise. Soit alors les évènements :

$$A = \text{"}la\ première\ boule\ tirée\ est\ verte\text{"}$$

$$B = \text{"}la\ deuxième\ boule\ tirée\ est\ verte\text{"}$$

$$C = \text{"}la\ troisième\ boule\ tirée\ est\ verte\text{"}$$

Une issue peut être représentée par un triplet formé avec trois boules, deux boules pouvant être identique, par exemple (V, R_1, R_1) qui signifie qu'on a tiré la boule verte en premier puis deux fois de suite la même boule rouge notée R_1. Montrons par exemple que A est indépendant de $B \cap C$. Tous les triplets étant équiprobables, nous avons :

$$p(A) = \frac{Card(A)}{Card(\Omega)} = \frac{1 \times 3 \times 3}{3 \times 3 \times 3} = \frac{1}{3}$$

$$p(B \cap C) = \frac{Card(B \cap C)}{Card(\Omega)} = \frac{3 \times 1 \times 1}{3 \times 3 \times 3} = \frac{1}{9}$$

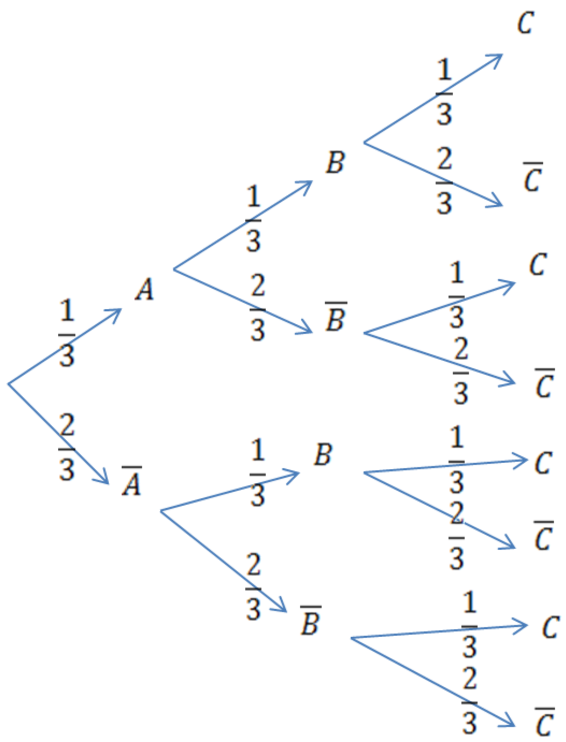
$$p(A \cap B \cap C) = \frac{Card(A \cap B \cap C)}{Card(\Omega)} = \frac{1 \times 1 \times 1}{3 \times 3 \times 3} = \frac{1}{27}$$

On a donc bien :

$$p(A \cap B \cap C) = p(A)\, p(B \cap C)$$

On vérifierait toutes les autres relations de la même façon ce qui montre que les évènements A, B, C sont mutuellement indépendants.

La situation peut alors être résumée par un arbre, qualifié de schéma de Bernoulli.



3) Indépendance mutuelle de n évènements

On dit que n évènements $A_1, A_2, \ldots, A_n$ sont indépendants s'ils vérifient pour tous $i, j, k, \ldots$ entiers compris entre 1 et n :

$$\boldsymbol{P(A_i \cap A_j) = P(A_i)\, P(A_j)}$$

$$\boldsymbol{P(A_i \cap A_j \cap A_k) = P(A_i)\, P(A_j)\, P(A_k)}$$

$$\vdots$$

$$\boldsymbol{P(A_1 \cap A_2 \cap A_n) = P(A_1)\, P(A_2) \ldots P(A_n)}$$

autrement dit si chacun des évènements est indépendant avec l'intersection d'un nombre quelconque d'autres.

V Variable aléatoire réelle

1) Définition

Reprenons l'exemple du lancer de deux dés équilibrés et considérons cette fois-ci l'univers des couples de faces possibles, dé rouge en premier, dé vert en second. Par exemple, si le dé rouge a donné 5 et le dé vert 6, le couple est (5,6). Cet univers Ω comporte 36 éléments. On peut le désigner ainsi :

$$\Omega = \{(i,j) : \ 1 \leq i \leq 6, 1 \leq j \leq 6 \} = [\![1;6]\!]^2$$

A chaque issue (i,j) de Ω on peut associer le résultat $i+j$ qui correspond à la somme des faces obtenues. On définit ainsi une fonction notée par exemple X de Ω dans un nouvel univers $X(\Omega) = [\![2;12]\!]$ appelé univers image et dont la probabilité d'une issue k est définie et notée comme suit :

$$P(X = k) = P\{(i,j) : \ i + j = k \}$$

Ainsi :

$$P(X = 2) = P\{(1,1)\} = \frac{1}{36}$$

$$P(X = 3) = P\{(1,2),(2,1)\} = \frac{2}{36}$$

2) Loi conjointe de deux variables aléatoires

Si X, Y sont deux variables aléatoires réelles prenant des valeurs dans les univers respectifs :

$$\Omega_X = \{x_1, x_2, \dots x_n\}, \qquad \Omega_Y = \{y_1, y_2, \dots y_p\}$$

Alors la loi conjointe est la donnée des probabilités suivantes pour tous les couples (i,j) de $\{1,2, \dots n\} \times \{1,2, \dots p\}$

$$P\left((X = x_i) \cap (Y = y_j)\right) = p_{ij}$$

Il en résulte les lois dites marginales :

$$P(X = x_i) = \sum_{j=1}^{p} P\left((X = x_i) \cap (Y = y_j)\right) = \sum_{j=1}^{p} p_{ij} = p_{i*}$$

$$P(Y = y_j) = \sum_{i=1}^{n} P\left((X = x_i) \cap (Y = y_j)\right) = \sum_{i=1}^{n} p_{ij} = p_{*j}$$

Nous voyons que la définition d'une loi conjointe se prête à l'usage d'un tableau pour la représenter, ainsi dans l'exemple :

X\Y	y_1	y_2	y_3	total
x_1	0,10	0,40	0,10	0,60
x_2	0,05	0,10	0,05	0,20
x_3	0,01	0,02	0,10	0,13
x_4	0,04	0,01	0,02	0,07
total	0,20	0,53	0,27	1

Les lois marginales apparaissent précisément aux marges du tableau à la ligne total et la colonne total.

3) Variables aléatoires indépendantes

Deux variables aléatoires X, Y sont dites indépendantes si pour tous les couples (i, j) de $\{1,2, \dots n\} \times \{1,2, \dots p\}$

$$P\left((X = x_i) \cap (Y = y_j)\right) = P(X = x_i)\, P(Y = y_j)$$

Soit avec les notations adoptées précédemment :

$$p_{ij} = p_{i*}\, p_{*j}$$

Cela se traduit dans le tableau de la loi conjointe par le fait que le contenu de chaque cellule du tableau est le produit des nombres figurant dans la ligne total et la colonne total de cette cellule.

n variables aléatoires $X_1, X_2, \dots, X_n$sont dites mutuellement indépendantes si pour tous les n-uplets $(x_{11}, x_{21}, \dots, x_{n1})$ de valeurs respectives de ces variables on a :

$$P\big((X_1 = x_{11}) \cap (X_2 = x_{21}) \dots \cap (X_n = x_{n1})\big) = P(X_1 = x_{11})\, P(X_2 = x_{21}) \dots P(X_n = x_{n1})$$

4) Espérance

L'espérance d'une variable aléatoire est la valeur moyenne des valeurs prises par la variable sur un très grand nombre de répétition de l'expérience aléatoire (en théorie une infinité, en pratique plusieurs milliers). On la définit donc par la même formule qu'en statistiques :

$$E(X) = \sum_{i=1}^{N} x_i \ P(X = x_i)$$

Propriétés de l'espérance :

a) Variable aléatoire constante

Si X est une variable aléatoire réelle ne prenant qu'une seule valeur b, on dit que X est constante et on a :

$$E(X) = b$$

b) Linéarité de l'espérance :

Si a est un réel et X, Y deux variables aléatoires réelles, on a :

$$E(a\,X) = a\,E(X)$$

$$E(X + Y) = E(X) + E(Y)$$

En particulier :

$$E(a\,X + b) = a\,E(X) + b$$

Preuve :

$$E(a\,X) = \sum_{i=1}^{N} a\,x_i \ P(X = x_i) = a\sum_{i=1}^{N} x_i \ P(X = x_i) = a\,E(X)$$

$$E(X + Y) = \sum_{(i,j)\in[\![1,n]\!]\times[\![1,p]\!]} \left(x_i + y_j\right) \ P\left((X = x_i) \cap \left(Y = y_j\right)\right)$$

$$= \sum_{(i,j)\in[\![1,n]\!]\times[\![1,p]\!]} x_i \ p_{ij} + \sum_{(i,j)\in[\![1,n]\!]\times[\![1,p]\!]} y_j \ p_{ij}$$

$$= \sum_{i=1}^{n}\sum_{j=1}^{p} x_i\,p_{ij} + \sum_{j=1}^{p}\sum_{i=1}^{n} y_j\,p_{ij}$$

$$= \sum_{i=1}^{n}\left(x_i \sum_{j=1}^{p} p_{ij}\right) + \sum_{j=1}^{p}\left(y_j \sum_{i=1}^{n} p_{ij}\right)$$

$$= \sum_{i=1}^{n}(x_i p_{i*}) + \sum_{j=1}^{p}(y_j p_{*j})$$

$$= E(X) + E(Y)$$

c) Espérance d'un produit de variables aléatoires

Si X, Y sont deux variables aléatoires réelles indépendantes alors :

$$\boxed{E(X\,Y) = E(X)\,E(Y)}$$

Preuve :

$$E(X\,Y) = \sum_{(i,j)\in[\![1,n]\!]\times[\![1,p]\!]} (x_i\, y_j)\; P\left((X = x_i) \cap (Y = y_j)\right)$$

$$= \sum_{(i,j)\in[\![1,n]\!]\times[\![1,p]\!]} \left(x_i\, y_j\, P(X = x_i)\, P(Y = y_j)\right)$$

$$= \sum_{(i,j)\in[\![1,n]\!]\times[\![1,p]\!]} \left(x_i\, P(X = x_i)\right)\left(y_j\, P(Y = y_j)\right)$$

$$= \left(\sum_{i=1}^{n}\left(x_i\, P(X = x_i)\right)\right)\;\left(\sum_{j=1}^{p}\left(y_j\, P(Y = y_j)\right)\right)$$

$$= E(X)E(Y)$$

5) Variance

Par une démarche analogue à celle utilisée en statistiques, on définit la variance d'une variable aléatoire X par :

$$V(X) = E(X^2) - E(X)^2 = \sum_{i=1}^{N} {x_i}^2 \; P(X = x_i) - \left(\sum_{i=1}^{N} x_i \; P(X = x_i) \right)^2$$

<u>Propriétés de la variance</u> :

Si a, b sont deux réels et X, Y deux variables aléatoires réelles, alors :

$$V(a\,X + b) = \; a^2\, V(X)$$

$$V(X + Y) = V(X) + V(Y) + \; 2\, cov(X, Y)$$

où :

$$cov(X, Y) = E(X\,Y) - E(X)E(Y)$$

En particulier, si X et Y sont indépendantes alors :

$$V(X + Y) = V(X) + V(Y)$$

Si $X_1, X_2, \dots, X_n$ sont n variables aléatoire réelles alors :

$$V(X_1 + X_2 + \dots + X_n) = V(X_1) + V(X_2) + \dots + V(X_n) + 2 \sum_{1 \le i < j \le n} cov\left(X_i, X_j\right)$$

En particulier, si les variables sont indépendantes deux à deux (donc en particulier si elles sont mutuellement indépendantes) on a :

$$V(X_1 + X_2 + \dots + X_n) = V(X_1) + V(X_2) + \dots + V(X_n)$$

<u>Preuves</u> :

$$V(a\,X + b) = E((a\,X + b)^2) - \left(E(a\,X + b)\right)^2$$

$$= E(a^2\, X^2 + 2\, a\, X + b^2) - (a\, E(X) + b)^2$$

$$= a^2\, E(\,X^2) + 2\, a\, E(X) + b^2 - a^2 (E(X))^2 - 2\, a\, b\, E(X) - b^2$$

$$= a^2\, (E(\,X^2) - (E(X))^2)$$

$$= \; a^2\, V(X)$$

$$V(X + Y) = E((X + Y)^2) - \left(E(X + Y)\right)^2$$

$$= E(\,X^2 + 2\,X\,Y + Y^2) - (E(X) + E(Y))^2$$

$$= E(\,X^2) + 2\,E(X\,Y) \; + E(Y)^2 - \left(E(X)\right)^2 - 2\,E(X)E(Y) - E(Y)^2$$

$$= \left(E(X^2) - \left(E(X)\right)^2\right) + \left(E(Y^2) - \left(E(Y)\right)^2\right) + 2\left(E(X\,Y) - E(X)E(Y)\right)$$

$$= V(X) + V(Y) + 2\,cov(X,Y)$$

$$V(X_1 + X_2 + \cdots + X_n) = E\left(\left(\sum_{i=1}^{n} X_i\right)^2\right) - \left(E\left(\sum_{i=1}^{n} X_i\right)\right)^2$$

$$= E\left(\sum_{(i,j)\in[\![1,n]\!]} \left(X_i\,X_j\right)\right) - \left(\sum_{i=1}^{n} E(X_i)\right)^2$$

$$= \sum_{(i,j)\in[\![1,n]\!]} E\left(X_i\,X_j\right) - \sum_{(i,j)\in[\![1,n]\!]} \left(E(X_i)\,E\left(X_j\right)\right)$$

$$= \sum_{(i,j)\in[\![1,n]\!]} \left(E\left(X_i\,X_j\right) - E(X_i)\,E\left(X_j\right)\right)$$

$$= \sum_{(i,j)\in[\![1,n]\!]} cov\left(X_i, X_j\right)$$

$$= \sum_{i=1}^{n} cov(X_i, X_i) + 2\sum_{1\le i<j\le n} cov\left(X_i, X_j\right)$$

$$= \sum_{i=1}^{n} V(X_i) + 2\sum_{1\le i<j\le n} cov\left(X_i, X_j\right)$$

6) Ecart type

L'écart type est la racine carrée de la variance soit :

$$\sigma(X) = \sqrt{V(X)}$$

Propriétés :

Si a, b sont deux réels et X, Y deux variables aléatoires réelles, alors :

$$\sigma(a\,X + b) = |a|\,\sigma(X)$$

Si X et Y sont indépendantes alors :

$$\sigma(X+Y) = \sqrt{\left(\sigma(X)\right)^2 + \left(\sigma(Y)\right)^2}$$

Si $X_1, X_2, \dots, X_n$ sont n variables aléatoire réelles indépendantes deux à deux alors :

$$\sigma(X_1 + X_2 + \cdots + X_n) = \sqrt{\left(\sigma(X_1)\right)^2 + \left(\sigma(X_2)\right)^2 + \cdots + \left(\sigma(X_n)\right)^2}$$

7) Variable aléatoire discrète

Prenons l'exemple d'une pièce équilibrée que l'on lance un très nombre de fois (qu'on peut par la pensée imaginer infini). Intéressons nous à la séquence de faces obtenues. Une issue est donc une suite de piles ou faces par exemple : (P,P,F,P,F,F,F,…). Considérons alors la variable aléatoire X prenant la valeur du rang pour lequel on a obtenu pour la première fois un pile. Dans l'issue donnée en exemple, X prend la valeur 1. Pour l'issue (F,F,F,F, F, P, P,F,…) X prend la valeur 6. Et pour une issue où il n'y aurait que des faces nous conviendrons que X prend la valeur +∞.

Nous sommes donc en présence d'une variable aléatoire prenant un nombre infini et discret de valeurs qui est $[\![1; +\infty]\!]$. La loi de X sera étudiée un peu plus loin et qualifiée de loi géométrique.

L'espérance, la variance et l'écart type d'une variable aléatoire discrète se définissent alors naturellement ainsi :

$$E(X) = \sum_{i=1}^{+\infty} x_i \; P(X = x_i)$$

$$V(X) = E(X^2) - E(X)^2 = \sum_{i=1}^{+\infty} {x_i}^2 \; P(X = x_i) - \left(\sum_{i=1}^{+\infty} x_i \; P(X = x_i)\right)^2$$

$$\sigma(X) = \sqrt{V(X)}$$

Mais se pose à priori un problème, la convergence des séries. Et nous pouvons exhiber un exemple dans lequel l'espérance et donc la variance ne peuvent pas être déterminés.

Soit la variable X prenant les valeurs entières de $[\![1; +\infty]\!]$ et telle que :

$$P(X = i) = \frac{6}{\pi^2} \times \frac{1}{i^2}$$

X est bien définie car :

$$\sum_{i=1}^{+\infty} P(X=i) = \frac{6}{\pi^2} \sum_{i=1}^{+\infty} \frac{1}{i^2} = \frac{6}{\pi^2} \times \frac{\pi^2}{6} = 1$$

Et pourtant elle n'a pas d'espérance car :

$$\sum_{i=1}^{N} i\, P(X=i) = \frac{6}{\pi^2} \sum_{i=1}^{N} \frac{1}{i} \sim \frac{6}{\pi^2}\, Ln(N)$$

Donc

$$\lim_{N\to+\infty} \sum_{i=1}^{N} i\, P(X=i) = +\infty$$

8) <u>Variable aléatoire continue</u>

Prenons l'exemple de la population européenne. On sélectionne au hasard un individu et on considère la variable aléatoire X qui prend pour valeur la taille en centimètres de l'individu.

Il ne serait pas commode de décrire la distribution des tailles même arrondies au centimètre à l'aide d'une variable aléatoire prenant un nombre fini de valeurs car beaucoup de valeurs seraient très proches les unes des autres. On imagine alors ce que donnerait un histogramme statistique à pas constant et très petit et on lisse l'histogramme pour le rendre plus continu. La fonction $f(x)$ ainsi obtenue est appelée densité de probabilité et doit vérifier les conditions :

$$\begin{cases} \forall\, x \in \mathbb{R} :\ f(x) \geq 0 \\ \displaystyle\int_{-\infty}^{+\infty} f(x)dx = 1 \end{cases}$$

L'espérance, la variance et l'écart type d'une variable aléatoire continue se définissent alors naturellement ainsi :

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

$$V(X) = E(X^2) - E(X)^2 = \int_{-\infty}^{+\infty} x^{\,2} f(x)dx - \left(\int_{-\infty}^{+\infty} x\, f(x)dx \right)^2$$

$$\sigma(X) = \sqrt{V(X)}$$

Mais se pose à priori un problème, la convergence des intégrales. Et nous pouvons exhiber un exemple dans lequel l'espérance et donc la variance ne peuvent pas être déterminés.

Soit la variable X de densité :

$$f(x) = \frac{1}{x^2}\ si\ x \geq 1,\ \ 0\ sinon$$

X est bien définie car :

$$\int_{-\infty}^{+\infty} f(x)dx = \int_{1}^{+\infty} \frac{1}{x^2} dx = \left[-\frac{1}{x}\right]_1^{+\infty} = 1$$

Et pourtant elle n'a pas d'espérance car :

$$\int_{-\infty}^{+\infty} x\, f(x)dx = \int_{1}^{+\infty} \frac{1}{x} dx = [Ln(x)]_1^{+\infty} = +\infty$$