Cours de probabilités discrètes

par Pierre Veuillez

1 Probabilités

1.1 Vocabulaire

Evénements Dans une expérience aléatoire, on obtient des *résultats élémentaires.*

Ces résultats élémentaires sont intéressants à connaître quand ils jouent tous le même rôle dans l'expérience : quand ils sont *équiprobables.*

L'ensemble des résultats élémentaires est appelé *univers* et noté la plupart du temps Ω.

Avec plusieurs résultats élémentaires on forme des *événements.*

Ω est appelé *événement certain.*

Son contraire est noté $\emptyset$ et appelé événement *impossible.*

Des événements sont *incompatibles* (*deux à deux*) ou *disjoints* s'ils ne peuvent se produire simultanément. Si l'un est réalisé, aucun des autres ne peut l'être.

Attention Ne pas confondre incompatibles (l'un empêche l'autre) et indépendants (l'un n'influe pas sur l'autre)

Opérations On peut faire certaines opérations sur les événements :

La *négation* (ou *contraire*) $\bar{A}$ est réalisé si et seulement si A ne l'est pas. Le contraire de noir n'est pas blanc.

La *réunion* symbole $\cup$ qui est la traduction du *ou* inclusif (l'un ou l'autre ou les deux), de *au moins* voir même de certains *et* comme : on peut gagner si l'on a blanc et si l'on a noir ...

L'*intersection* symbole $\cap$ qui est la traduction de *et,* de *tous*, de *jamais*, de *aucun*, de *à la fois...*

La *différence* $A \backslash B$ est formé des résultats qui sont dans A sans être dans B : $A \backslash B = A \cap \left(\bar{B}\right)$

Un *système complet d'événements* ou *partition de l'univers* est une famille d'événements, finie ou dénombrable (indiciable par les entiers) $(A_i)_{i \in I}$ telle qu'un et un seul de ces événements est réalisé à chaque fois (ils sont incompatibles (deux à deux) et leur réunion est l'événement certain (on est certain que l'un au moins est réalisé)

Un événement est *presque sûr* si sa probabilité vaut 1.

Un événement est *négligeable* si sa probabilité vaut 0.

Attention Un événement ne peut pas être conditionné : on a envie de conditionner parce que la réalisation de A dépend de de celle de B.

C'est probablement $(A \cap B)$ qu'il faut écrire. Le conditionnement apparaît alors tout naturellement lorsque l'on calcule la probabilité via la formule des probabilité composées : $P(A \cap B) = P_B(A)\,P(B)$

1.2 Cadre théorique

Pour aller (beaucoup) plus loin dans l'étude des probabilité on doit préciser la nature mathématique des objets sur lesquels on travaillait naïvement.
L'univers modélise l'ensemble des résultats possibles. C'est un ensemble et c'est tout. On le note Ω. C'est l'événement certain
Une tribu de Ω sera l'ensemble des événements. On devra pouvoir faire avec ces événements les opérations nécessaires et trouver comme résultat un événement : réunion, contraire et intersection.

Tribu $\mathcal{T}$ est une tribu de Ω si

- $\mathcal{T}$ est un ensemble de parties de Ω. (ensemble d'ensembles) Les éléments de la tribu sont appelés événements.
- qui contient Ω
- si $A \in \mathcal{T}$ alors $\bar{A} \in \mathcal{T}$ (le contraire d'un événement en est un)
- si I est un ensemble fini ou dénombrable (que l'on peut compter avec les entiers) et $(A_i)_{i\in I}$ une famille d'événements (finie ou dénombrable) alors $\bigcup\limits_{i\in I} A_i$ est un événement. (appartient à $\mathcal{T}$)

$(\mathcal{T}, \Omega\ )$ est appelé espace probabilisable

On définit alors la probabilité qui respecte les propriétés de la fréquence statistique.

Probabilité Soit $(\mathcal{T}, \Omega\ )$ un espace probabilisable et P une application de $\mathcal{T}$ dans $\mathbb{R}^+$ ($P(A)$ doit être défini pour tout événement A et être positive). P est une probabilité sur $(\mathcal{T}, \Omega\ )$ si
- $P(\Omega) = 1$
- et si pour toute famille finie ou dénombrable $(A_i)_{i\in I}$ d'événements incompatibles, la somme de la série $\sum\limits_{i\in I} P(A_i)$ vaut $P\left(\bigcup\limits_{i\in I} A_i\right)$

Conséquences On en déduit :
- Le contraire de Ω qui est l'événement impossible $\emptyset$ est un événement est porte donc bien son nom.
- le contraire d'une réunion est une intersection donc un intersection finie ou dénombrable d'événements est encore un événement.

1.3 Les propriétés

Evénements Il y a quelques traductions à savoir faire :

Le contraire de "tous sont" c'est "au moins un n'est pas" ($\overline{\bigcap\limits_{i\in I} A_i} = \bigcup\limits_{i\in I} \overline{A_i}$)

Le contraire de "au moins un est " c'est "aucun n'est"

Si A est réalisé **alors** B est réalisé se traduit par $A \subset B$; on a alors $P(A) \leq P(B)$

Probabilité une probabilité est toujours comprise entre 0 et 1

$P(\emptyset) = 0, \quad P(\Omega) = 1, \quad P(\bar{A}) = 1 - P(A)$

Si A et B sont incompatibles alors $P(A \cup B) = P(A) + P(B)$

Si $(A_i)_{i\in\mathbb{N}}$ est une famille d'événements incompatibles alors la série $\sum_{i\in\mathbb{N}} P(A_i)$ converge et

$$P\left(\bigcup_{i=0}^{+\infty} A_i\right) = \sum_{i=0}^{+\infty} P(A_i)$$

Si $(A_i)_{i\in\mathbb{N}}$ est une famille croissante (i.e. que $\forall i \in \mathbb{N} \quad A_i \subset A_{i+1}$; par exemple $A_i =$"avoir au moins un pile lors des i premiers lancers") alors :

$$P\left(\bigcup_{i=0}^{+\infty} A_i\right) = \lim_{n\to+\infty} P(A_n)$$

Et si $(A_i)_{i\in\mathbb{N}}$ est une famille décroissante (i.e. que $\forall i \in \mathbb{N} \quad A_{i+1} \subset A_i$; par exemple $A_i =$"n'avoir que des piles lors des i premiers lancers") alors :

$$P\left(\bigcap_{i=0}^{+\infty} A_i\right) = \lim_{n\to+\infty} P(A_n)$$

Si $(A_i)_{i\in\mathbb{N}}$ est une famille quelconque, $\boxed{P\left(\bigcup_{i=0}^{+\infty} A_i\right) = \lim_{n\to+\infty} P\left(\bigcup_{i=0}^{n} A_i\right)}$

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

Si $(A_i)_{i\in\mathbb{N}}$ est une famille quelconque, $\boxed{P\left(\bigcap_{i=0}^{+\infty} A_i\right) = \lim_{n\to+\infty} P\left(\bigcap_{i=0}^{n} A_i\right)}$

Equiprobabilité Quand tous les résultats élémentaires jouent le même rôle, sont équiprobables, on modélise par la probabilité équiprobable :

$$P(A) = \frac{|A|}{|\Omega|}$$

Par exemple, on fait 10 tirages sans remise dans une urne. Si l'on ne sait rien sur les résultats précédents, à chaque tirage toutes les boules seront équiprobables (bien que le contenu de l'urne change à chaque fois)

Au contraire si l'on tire dans une urne ou une autre, sachant dans quelle urne on tire, les boules de cette urne seront équiprobables. C'est ici la probabilité conditionnelle qui sera la probabilité équiprobable.

1.4 Probabilité conditionnelle, indépendance

Probabilité conditionnelle On définit la probabilité conditionnelle de A sachant B et on note $P_B(A)$ ou $P_B(A) = \dfrac{P(A\cap B)}{P(B)}$ (notée autrefois $P(A/B)$). Cela conduit à la formule de Bayes.

Mais pour la calculer, c'est la plupart du temps : la probabilité que l'on a quand on sait que B est réalisé (si la réalisation de B permet de connaître les *conditions expérimentales*)

La *probabilité conditionnelle* est la probabilité qui correspond à la fréquence statistique quand on restreint a priori l'ensemble des possibles.

- Une probabilité conditionnelle est une probabilité. On retrouve donc les mêmes règles de calcul en remplaçant partout la probabilité par la probabilité conditonnelle :
  $P_B(\Omega) = 1, \quad P_B(\emptyset) = 0, \quad P_B\left(\bar{A}\right) = 1 - P_B(A)\ldots$

La probabilité conditionnelle intervient naturellement dans la calcul de la probabilité d'une intersection et dans la formule des probabilités totales :

- formule des probabilité composées : $P(A\cap B) = P(A)\cdot P_A(B)$ ou plus généralement :

$$P(A_1\cap A_2\cap\cdots\cap A_n) = P(A_1)\cdot P_{A_1}(A_2)\cdots\cdot P_{A_1\cap\cdots\cap A_{n-1}}(A_n)$$

  que l'on a intérêt à écrire en respectant l'ordre chronologique de l'expérience pour pouvoir ensuite interpréter le contitonnement.
- Formule des probabilités totales. On l'utilise quand la réalisation d'un événement A dépend des résultats $(B_i)_{i\in I}$ d'une étape précédentes.
  Si $(B_i)_{i\in I}$ est un système complet d'événements alors $P(A) = \sum_{i\in I} P_{B_i}(A)\, P(B_i)$
  N.B. : la série converge et la somme de la série vaut $P(A)$ dans le cas où l'on a un système complet d'événement infini
  N.B. : il faut que $P(B_i) \neq 0$ pour que les $P_{B_i}(A)$ soit défini ... ce que l'on ne vérifie jamais. Mais comme $P_{B_i}(A)$ est alors multiplié par $P(B_i)$ qui est nulle, les formules restent exploitables.

- Formule de Bayes enfin qui permet de calculer une probabilité conditionnelle en inversant l'événement et le conditionnement :
$$P_B(A) = \frac{P(A \cap B)}{P(B)} = \frac{P_A(B)\,P(A)}{P(B)}$$
le $P(B)$ du dénominateur étant souvent calculé par la formule des probabilités totales

Indépendance : Deux événements A et B sont indépendants si $P(A \cap B) = P(A)\,P(B)$. (ce qui peut s'écrire $P_B(A) = P(A)$ si $P(B) \neq 0$ ou $P_A(B) = P(B)$ si $P(A) \neq 0$).

Une famille $(A_i)_{i \in I}$ d'événements est indépendante si, quand on en prend n'importe quelle sous-famille finie, la probabilité de leur intersection est le produit de leurs probabilités.

Cette définition est la modélisation du fait que la réalisation de l'un des événements ne change pas, n'influe pas, sur la probabilité de réalisation de l'autre.

La plupart du temps, l'indépendance viendra des conditions expérimentales. (soit explicitement, soit implicitement)

Exemples les exemple type sont :

- une suite de lancers de pièces. Le fait qu'un lancer donne pile ou face ne change pas la probabilité d'avoir pile ou face à un autre lancer. Les lancers seront modélisés comme étant indépendants.
- Des tirages *avec remise* de boules dans une urne seront modélisés indépendants
- Des tirages sans remise seront dépendants. Dans le calcul de probabilité d'une intersection, (formule des probabilités composées) la probabilité conditonnelle apparaîtra... le conditionnement qui donnera la composition de l'urne.
- Par contre si on lance la pièce *jusqu'à* obtenir pile, on arrête les lancers dès que l'on a pile. Le fait d'avoir pile à un lancer implique que l'on en a pas eu avant . Et que l'on aura plus ni pile ni face après. Les lancers ne sont plus indépendants.
- De même si l'on choisi d'abord une pièce (truquée ou équilibrée) puis qu'on fait une suite de lancers, les résultats d'un lancé est lié à la pièce choisie. Donc via la pièce choisie, les résultats des lancers sont liés les uns aux autres et ne sont pas indépendants. Par contre si la pièce est connue (sachant que la pièce est truquée) indépendants.

2 Variables aléatoires

Les variables aléatoires donnent une valeur pour chaque résultat d'une expérience aléatoire.
C'est une fonction de l'ensemble des résultats (Ω) dans $\mathbb{R}$ définie presque sûrement. (”presque” exemple ?)
On note $(X = a)$ l'ensemble des résultas ω pour lesquels $X(\omega) = a$, de même $b < X \leq a$ ...
Et on doit pouvoir définir la probabilité de $X < a$ qui doit donc être un événement. (ce qui n'est pas toujours vrai pour des univers infinis)

Attention Une variable aléatoire n'est pas un événement. Si on écrit $P(X)$ on est certain de se tromper

Pour fabriquer un événement, il faut comparer la variable à autre chose ($X < x; \quad X = Y \ldots$)

Par contre, on peut faire des opérations avec des variables aléatoires pour en fabriquer de nouvelles.

2.1 Cadre théorique

Définition X est une variable aléatoire sur un espace probabilisé $(\Omega, \mathcal{T}, \mathrm{P})$ si X est une fonction définie presque sûrement sur Ω telle que pour tout réel a, $(X \leq a)$ appartient à la tribu (est un événement)

2.2 Loi et fonction de répartition

Définition pour une variable aléatoire discrète, la loi de X est la donnée de l'ensemble des valeurs possibles de X, $X(\Omega)$, et la probabilité que X prenne chacune de ces valeurs.

Cela peut être donné par une ou plusieurs formules, ou bien les valeurs peuvent être énumérées dans un tableau.

Caractérisation La loi d'une variable aléatoire discrète est caractérisée par le fait que chacune des probabilités est positive ou nulle et la somme des probabilités (ou la somme de la série des probabilité dans le cas discret infini) est égal à 1.

Fonction de répartition La fonction de répartition F d'une variable aléatoire X est définie par $F(x) = P(X \leq x)$ pour tout réel x.

– On peut retrouver la loi à partir de la fonction de répartition : si X prend les valeurs $x_1 < x_2 < \cdots < x_n$ alors

$$\begin{aligned} P(X = x_k) &= P(X \leq x_k) - P(X < x_k) \\ &= P(X \leq x_k) - P(X \leq x_{k-1}) \text{ si } k-1 \geq 1 \end{aligned}$$

La première valeur $X = x_1$ joue un rôle particulier et est à traiter a-priori à part. On peut souvent la réintégrer a-posteriori dans la formule générale.

– Et on peut retrouver la fonction de répartition à partir de la loi : si $x_k \leq x < x_{k+1}$ alors

$$F(x) = P(X \leq x) = P(X \leq x_k) = \sum_{i=1}^{k} P(X = x_i)$$

– La fonction de répartition est plus facile à obtenir que la loi quand on cherche la loi du *maximum* de plusieurs variables aléatoires.
Dire que "le plus grand est inférieur à x" signifie que "tous sont inférieurs à x".

2.3 Espérance et variance

Définition L'espérance d'une variable aléatoire qui ne prend qu'un nombre fini de valeur est : $E(X) = \sum_{k \in X(\Omega)} k \cdot P(X = k)$

Si X prend un nombre infini dénombrable de valeur, X n'a une espérance que si la série $\sum_{k \in X(\Omega)} k \cdot P(X = k)$ est absolument convergente. L'espérance de X est alors la somme de la série (sans valeur absolue)

La variance de X est $V(X) = E\left[(X - E(X))^2\right] = E(X^2) - E(X)^2$

Attention Quand la loi de X est donnée par plusieurs formules, il faut décomposer la somme pour pouvoir substituer la formule à $P(X = k)$

Calculs Pour calculer l'espérance d'une variable Y "fabriquée" à partir d'une autre X, il n'est pas utile de chercher d'abord la loi de Y :

– si $Y = f(X)$

$$E(Y) = E(f(X)) = \sum_{k \in X(\Omega)} f(k) P(X = k)$$

la somme étant celle de la série qui doit être absolument convergente dans le cas d'un nombre infini dénombrable de valeurs.

– En particulier, on en déduit la *linéarité* de l'espérance : si a et b sont des réels (et non pas des variables aléatoires) et X une variable aléatoire $E(aX + b) = aE(X) + b$.

- On trouve aussi (pour le calcul de la variance) que $E\left(X^2\right) = \sum_{k\in X(\Omega)} k^2 P\left(X = k\right)$ (somme de la série si elle est absolument convergente dans le cas discret infini)
- La variance se calcule plus facilement par $V\left(X\right) = E\left(X^2\right) - E\left(X\right)^2$ et on a pour des réels (pas des variables aléatoires) a et b : $V\left(aX + b\right) = a^2 V\left(X\right)$

2.4 Couple de variables

On peut considérer simultanément deux -ou plus- variables aléatoires.

Loi la loi du couple (X, Y) est la donnée des valeurs possibles de X et de Y et pour chacune des valeurs x de X et y de Y la donnée de la probabilité $P\left(X = x \cap Y = y\right)$ (qui peut se calculer par dénombrement dans le cas équiprobable, par transformation de l'écriture, par les probabilités composées...)

Elle peut être donnée par une (ou des) formules ou par l'énumération (dans un tableau à double entrée) de toutes les valeurs possibles.

Opérations Des événements (ou des variables aléatoires) sont souvent définis à partir de deux (ou plus) variables aléatoires. Pour en calculer la probabilité, il faut alors se ramener à la loi de chacune des variable, en décomposant les événements.

Les classiques $(X = Y) = \bigcup_{i\in I} (X = i \cap Y = i)$ l'ensemble I étant déterminé par la double contrainte que $(X = i)$ et $(Y = i)$ soient simultanément possibles.

$(X + Y = k) = \bigcup_{i\in I} (X = i \cap Y = k - i)$ avec comme contraintes : $i \in X\left(\Omega\right)$ et $k - i \in Y\left(\Omega\right)$

$(X \leq Y) = \bigcup_{i\in I} (X = i \cap Y \geq i)$ avec $i \in X\left(\Omega\right)$ et $k - i \in Y\left(\Omega\right)$

N.B. Une autre démarche consiste à passer par la formule des probabilités totales en analysant ainsi :

La réalisation dépend de la valeur prise par X d'où l'utilisation de la formule des probabilités totales avec $(X = i)_{i\in I}$ comme système complet d'événements :

$P\left(X = Y\right) = \sum_{i\in I} P_{X=i}\left(X = Y\right) P\left(X = i\right)$

le conditionnement étant ensuite interprété $P\left(X = Y\right) = \sum_{i\in I} P_{X=i}\left(Y = i\right) P\left(X = i\right)$

et enfin, si X et Y sont indépendantes, on peut supprimer le conditionnement $P\left(X = Y\right) = \sum_{i\in I} P\left(Y = i\right) P\left(X = i\right)$

Lois marginales On peut retrouver les lois de X et de Y à partir de la loi du couple :

$$P\left(X = x\right) = \sum_{y\in Y(\Omega)} P\left(X = x \cap Y = y\right)$$

dans le cas d'un nombre infini de valeurs pour Y, la série est convergente et $P\left(X = x\right)$ est la somme de la série.

Attention Quand la loi du couple est donnée par plusieurs formules, il faut décomposer la somme pour pouvoir substituer.

Espérance Si f est une application qui à un couple de réel (x, y) associe $f\left(x, y\right)$, définie sur $X\left(\Omega\right) \times Y\left(\Omega\right)$ alors

$$E\left(f\left(X, Y\right)\right) = \sum_{x\in X(\Omega)} \sum_{y\in Y(\Omega)} f\left(x, y\right) P\left(X = x \cap Y = y\right)$$

On a en particulier pour le calcul de la covariance :

$$E\left(X \cdot Y\right) = \sum_{x\in X(\Omega)} \sum_{y\in Y(\Omega)} x \cdot y \cdot P\left(X = x \cap Y = y\right)$$

Elle permet aussi d'obtenir, si X et Y sont deux variables aléatoires $E(X+Y) = E(X)+E(Y)$
Et si X et Y sont indépendantes, $E(X \cdot Y) = E(X) \cdot E(Y)$

Covariance Outre les lois marginales, l'intérêt des couples de variables aléatoires et de pouvoir calculer la variance d'une somme via la covariance.
La covariance est $\mathrm{Cov}(X,Y) = E[(X - E(X))(Y - E(Y))] = E(X \cdot Y) - E(X) \cdot E(Y)$ (dans le cas discret infini, elle existe si les espérances existent et donc si les séries sont absolument convergentes)
Cette covariance se comporte bien avec la somme et le produit par des constantes :
$\mathrm{Cov}(aX+b, cY+d) = ac\mathrm{Cov}(X,Y)$
Si X et Y sont indépendantes alors $\mathrm{Cov}(X,Y) = 0$. la réciproque est fausse.

Variance C'est LA formule pour la variance d'une somme $V(X+Y) = V(X)+V(Y)+2\mathrm{Cov}(X,Y)$ (c'est la formule du binôme ou le carré est remplacé par la variance et le produit par la covariance)
et plus généralement pour une somme de n variables aléatoires :

$$V\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} V(X_i) + \sum_{i=1}^{n} \sum_{j \neq i}^{n} \mathrm{Cov}(X_i, X_j)$$

Coefficient de corrélation linéaire Soient X et Y deux variables aléatoires discrètes ayant une variance non nulle et une covariance.
Le coefficient de corrélation linéaire est $\rho(X,Y) = \dfrac{\mathrm{Cov}(X,Y)}{\sqrt{V(X) \cdot V(Y)}}$.

Propriétés On a $|\rho(X,Y)| \leq 1$ et $|\rho(X,Y)| = 1$ si et seulement si Y est fonction affine de X presque sûrement
(Si il existe a et b réels tels que $\mathrm{P}(Y = aX+b) = 1$)
Plus $|\rho(X,Y)|$ est proche de 1 et meilleur sera l'approximation de Y par une fonction affine de X.

N.B. La proximité entre Y et $aX+b$ se mesure par $E\left((Y - aX - b)^2\right)$ (écart quadratique moyen) et non par $E(Y - aX - b)$ (écart moyen)

Preuve de $|\rho(X,Y)| \leq 1$
Pour Z variable aléatoire $V(Z) = E\left((Z - E(Z))^2\right) \geq 0$ et nulle si et seulement si $\mathrm{P}(Z = E(Z)) = 1$ (si Z est constante presque sûrement)
Et avec $Z = Y - aX$ on a

$$V(Y - aX) = V(Y) - 2a\mathrm{Cov}(X,Y) + a^2 V(X)$$

polynôme du second degré en a et positif ou nul donc de discriminant $\Delta = 4\mathrm{Cov}(X,Y) - 4V(X)V(Y) \leq 0$
Or

$$\Delta = 4V(X)V(Y)\left[\rho(X,Y)^2 - 1\right] \leq 0$$

Donc $\rho(X,Y)^2 \leq 1$ et $|\rho(X,Y)| \leq 1$

Preuve de l'ajustement linéaire (affine)
$E\left((Y - aX - b)^2\right) = E(Y - aX - b)^2 + V(Y - aX - b) = E(Y - aX - b)^2 + V(Y - aX)$
Comment choisir b pour que ma première espérance soit nulle ?

b étant ainsi fixé, reste à réduire

$$\begin{aligned} V(Y-aX) &= V(Y) - 2a\mathrm{Cov}(X,Y) + a^2 V(X) \text{ canonique} \\ &= \left[a\sqrt{V(X)} - \frac{\mathrm{Cov}(X,Y)}{\sqrt{V(X)}}\right]^2 + V(Y) - \frac{\mathrm{Cov}(X,Y)^2}{V(X)} \\ &= \left[a\sqrt{V(X)} - \frac{\mathrm{Cov}(X,Y)}{\sqrt{V(X)}}\right]^2 + V(Y)\left(1-\rho(X,Y)^2\right) \end{aligned}$$

somme de deux quantités positive, minimale pour $a = \frac{\mathrm{Cov}(X,Y)}{V(X)}$ et dont le minimum vaut

$$V(Y) - \frac{\mathrm{Cov}(X,Y)^2}{V(X)} = V(Y)\left(1-\rho(X,Y)^2\right)$$

Donc plus $\rho(X,Y)^2$ est proche de 1 et meilleur sera l'ajustement.
dans le cas où $\rho(X,Y)^2 = 1$, pour les valeurs de a et b susdites, l'écart quadratique moyen sera nul.

3 Lois usuelles

Quand on a envie d'utiliser une loi usuelle mais que le paramètre qui devrait être un réel est une variable aléatoire,
(par exemple N le rang du premier pile. On relance N fois la pièce. X le nombre de pile obtenus)
on obtient d'abord la loi conditionnelle et ensuite la loi par les probabilités totales.

3.1 loi uniforme sur [[1,n]]

Modèle on tire au hasard un nombre entier dans l'intervalle $[[1,n]]$. Ces nombres sont donc équiprobables.
On note X le résultat (résultat d'un dé ou d'une boule numérotée dans une urne)

Définition X suit une loi uniforme sur $[[1,n]]$ si $X(\Omega) = [[a,b]]$ et si pour tout $k \in [[1,n]]$, $p(X=k) = \dfrac{1}{n}$
L'espérance de X est alors $E(X) = \dfrac{n+1}{2}$ et $V(X) = \frac{n^2-1}{12}$

3.2 Loi de Bernouilli

Modèle Elle *compte* le nombre de succès en *une seule* expérience (donc 0 ou 1).
Elle vaut 1 en cas de succès et 0 en cas d'échec.
Elles sont utiles à plusieurs : si X_k indique le succès lors de la $k^{ième}$ expérience, $\sum_k X_k$ compte le nombre total de succès.

Définition X suit une loi de Bernouilli de paramètre p si $X(\Omega) = \{0,1\}$ avec $P(X=1) = p$
On a alors $E(X) = p$ et $V(X) = p(1-p)$

3.3 Loi hypergéométrique

Modèle tirages successifs *sans remises* ou tirages simultanés parmi des bons et des mauvais. N le nombre total d'éléments, n le nombre d'éléments prélevés et p la proportion de bons. On note X le nombre de bons éléments prélevés.

Définition X suit une loi hypergéométrique de paramètres N, n et p si : on note $a = pN$ (nombre de bons éléments) et $b = (1-p)\,N$ (nombre de mauvais éléments)

$$P(X=k) = \frac{\binom{a}{k}\binom{b}{n-k}}{\binom{N}{n}}$$

avec $X(\Omega) = [[\max(0, n-b)\,;\min(a,n)]]$
(La formule reste vraie en dehors de cet intervalle, les probabilités étant simplement nulles)
On a $E(X) = np$

3.4 Loi binomiale

Modèle C'est la loi du *nombre de succès* en n expériences *indépendantes* qui ont toutes la même probabilité p de succès.

Définition X suit une loi binomiale si $X(\Omega) = [[0;n]]$ et pour tout entier $k \in [[0;n]]$: $P(X=k) = \binom{n}{k} p^k (1-p)^{n-k}$
On a alors $E(X) = np$ et $V(X) = np(1-p)$

Somme une somme de variable aléatoires indépendantes suivants des lois binomiales de même probabilité de succès en est encore une de même paramètre de succès et de premier paramètre (nombre d'expérience) la somme de leurs premiers paramètres.

3.5 Loi géométrique

Modèle C'est la loi du rang du premier succès dans une suite (infinie) d'expériences indépendantes qui ont toutes la même probabilité p de succès.

Sans mémoire Dans le cas où, tant que l'on a un échec la probabilité de succès reste la même p (variable sans mémoire) alors le rang X du premier succès suit une loi géométrique de paramètre p.

Définition X suit une loi géométrique si $X(\Omega) = [[1;+\infty[[$ et pour tout entier $k \in [[1;+\infty[[$: $P(X=k) = (1-p)^{k-1}\,p$
On a alors $E(X) = \dfrac{1}{p}$ et $V(X) = \dfrac{(1-p)}{p^2}$

Propriété Si $X(\Omega) = [[1,+\infty[[$ et que $\mathrm{P}(X=1) = p$ alors
$X \hookrightarrow \mathcal{G}(p) \Longleftrightarrow X$ est sans mémoire ($\mathrm{P}_{X\geq n}(X \geq m+n) = \mathrm{P}(X \geq m)$ pour tout m et $n \in \mathbb{N}^*$)

3.6 Loi de Poisson

Modèle Ce n'est pas une loi que l'on rencontre directement. Mais c'est une loi qui approche la loi Binomiale $B(n,p)$ quand n tend vers $+\infty$ mais que le produit $n \cdot p$ reste constant $= \alpha$ (ou tend vers cette constante)
C'est la loi qui (empiriquement) modélise bien les fréquentations (nombre de clients à une caisse dans une journée, nombre d'élèves en ECE1 une année donnée ...)

Définition X suit une loi de Poisson de paramètre α si $X(\Omega) = \mathbb{N}$ et pour tout entier $k \in \mathbb{N}$: $P(X=k) = \dfrac{\alpha^k e^{-\alpha}}{k!}$
On a alors $E(X) = \alpha$ et $V(X) = \alpha$

Somme une somme de variable aléatoires indépendantes suivants des lois de Poisson en est encore une de paramètre la somme des paramètres