Probabilités et lois de variables aléatoires discrètes

Espaces probabilisés

1. Tribu sur un ensemble Ω. Espace probabilisable $(\Omega, \mathcal{A})$.
2. Événements : généralisation du vocabulaire relatif aux événements introduit en première année.
3. Probabilité sur un espace probabilisable, σ-additivité. Espace probabilisé $(\Omega, \mathcal{A}, P)$.
4. Continuité croissante, continuité décroissante.
5. Pour une suite $(A_n)_{n\in\mathbb{N}}$ d'événements non nécessairement monotone,
$$\lim_{n\to\infty} P\Big(\bigcup_{k=0}^{n} A_k\Big) = P\Big(\bigcup_{k=0}^{\infty} A_k\Big) \quad \text{et} \quad \lim_{n\to\infty} P\Big(\bigcap_{k=0}^{n} A_k\Big) = P\Big(\bigcap_{k=0}^{\infty} A_k\Big).$$
6. Propriété de sous-additivité de P pour une réunion dénombrable d'événements.
7. Événements négligeables, événements presque sûrs. Réunion et intersection finie ou dénombrable.
8. Systèmes quasi-complets d'événements.

Probabilités conditionnelles et indépendance

9. Extension des résultats vus en première année : probabilité conditionnelle, formule des probabilités composées, formule des probabilités totales, formule de Bayes.
10. Indépendance de deux événements. Famille d'événements indépendants. L'indépendance deux à deux n'implique pas l'indépendance.
11. Si A et B sont indépendants, A et $\overline{B}$ le sont aussi.

Espaces probabilisés discrets

12. Si Ω est un ensemble, une distribution de probabilités discrètes sur Ω est une famille d'éléments de $\mathbb{R}^+$ indexée par Ω et de somme 1.
13. Support d'une distribution de probabilités discrète. Le support est au plus dénombrable.
14. Probabilité définie sur $\mathcal{A} = \mathcal{P}(\Omega)$ associée à une distribution de probabilités discrètes sur Ω.

Variables aléatoires discrètes

1. Définition d'une variable aléatoire discrète X définie sur l'espace probabilisé $(\Omega, \mathcal{A}, P)$ et à valeurs dans E.
2. Loi P_X d'une variable aléatoire discrète X.
   Dans ce qui suit, toutes les variables aléatoires sont supposées discrètes.
3. La probabilité P_X est déterminée par la distribution de probabilités discrète $(P(X = x))_{x\in X(\Omega)}$.
4. Notation $X \sim Y$. Variable aléatoire $f(X)$. Si $X \sim Y$ alors $f(X) \sim f(Y)$.
5. Loi conditionnelle d'une variable aléatoire X sachant un événement A.
6. Couple de variables aléatoires. Loi conjointe, lois marginales. Détermination des lois marginales à partir de la loi conjointe. Extension aux n-uplets de variables aléatoires.

Variables aléatoires indépendantes

7. Couple de variables aléatoires indépendantes, famille finie de variables aléatoires indépendantes. Extension aux n-uplets de variables aléatoires. Famille quelconque de variables aléatoires indépendantes.
8. Fonctions de variables aléatoires indépendantes : si $X \perp\!\!\!\perp Y$, alors $f(X) \perp\!\!\!\perp g(Y)$. Extension au cas de plus de deux variables. Lemme des coalitions.
9. Existence d'espaces probabilisés portant une suite de variables indépendantes de lois discrètes données. Modélisation du jeu de pile ou face infini : suite i.i.d. de variables de Bernoulli.

Lois usuelles

10. Loi géométrique. Interprétation comme rang du premier succès dans le jeu de pile ou face infini.
11. Loi de Poisson. Interprétation en termes d'événements rares.

Dans ce chapitre, nous généralisons à un univers infini le cadre de la théorie des probabilités présenté en première année. Nous commençons par des révisions de dénombrement ; les révisions de probabilité de première année se feront au fur et à mesure du chapitre.
On rappelle qu'un ensemble est *au plus dénombrable* quand il est fini ou dénombrable.

1 Révisions de dénombrement (MPSI)

1.1 cardinal

Définition 1

Un ensemble E est fini s'il est vide ou s'il existe un entier naturel non nul n et une bijection f de E dans $[\![1, n]\!]$.
n est appelé cardinal de E et noté $\operatorname{Card} E$; il représente concrètement le nombre d'éléments de E.
L'ensemble vide $\emptyset$ est fini de cardinal 0.

Les normes mathématiques internationales donnent la notation $|E|$.

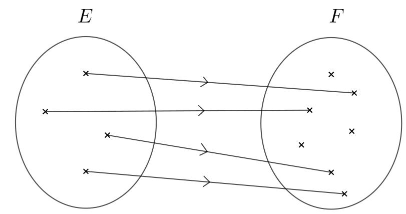
Propriété 1

- pour $A_1, \ldots, A_p$ des sous-ensembles DISJOINTS de E,
$$\operatorname{Card}(\bigcup_{i=1}^{p} A_i) = \sum_{i=1}^{p} \operatorname{Card} A_i$$
En particulier, $\operatorname{Card} A + \operatorname{Card} \overline{A} = \operatorname{Card} E$.
- pour $F_1, \ldots, F_p$ ensembles,
$$\operatorname{Card}(F_1 \times \cdots \times F_p) = (\operatorname{Card} F_1) \times \cdots \times (\operatorname{Card} F_p)$$
- $\operatorname{Card} \mathcal{P}(E) = 2^{\operatorname{Card} E}$
- pour A et B sous-ensembles de E, $\operatorname{Card} A \cup B = \operatorname{Card} A + \operatorname{Card} B - \operatorname{Card} A \cap B$

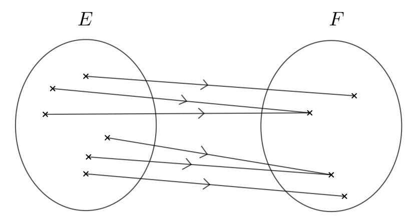
Propriété 2 – effet d'une application

Soient E et F deux ensembles et $f : E \to F$ une application.

- Si f est injective et si F est fini, alors E est aussi fini. On a $\operatorname{Card} E \leqslant \operatorname{Card} F$, avec égalité si et seulement si f est bijective.



- Si f est surjective et si E est fini, alors F est aussi fini. On a $\operatorname{Card} F \leqslant \operatorname{Card} E$, avec égalité si et seulement si f est bijective.



- Si E et F sont finis et de même cardinal :
  f bijective $\Leftrightarrow$ f injective $\Leftrightarrow$ f surjective.

Notre motivation est de savoir répondre à des questions du type :

— Combien y a-t-il d'anagrammes du mot « Mathématiques » ?

— À partir de combien d'élèves dans un lycée est-on sûr d'en avoir au moins deux ayant les mêmes initiales :

  — si on les a placées dans l'ordre prénom-nom ?

  — si on n'a pas fait attention à l'ordre entre le nom et le prénom ?

— On a n enfants. Combien y a-t-il de manières de les disposer en file indienne ? en ronde ?

Si chacune des questions requiert un minimum de théorie mathématique, il importe de rester concret (on se représente concrètement le problème !), de proposer une représentation du problème (schéma, arbre, liste...) et d'être bien ordonné. Par exemple, énumérer les mots de 3 lettres distinctes formées avec A, B et C se fait avec l'ordre lexicographique :

ABC, ACB,

Ces principes de bon sens nous aideront à nous ramener à des situations familières, s'appuyant par exemple sur l'un des trois modèles suivants :

— tirages successifs avec remise (modélisés par des listes)

— tirages successifs sans remise (modélisés par des arrangements)

— tirages simultanés (modélisés par des combinaisons).

Dans la suite, E est un ensemble fini à n éléments et p est un entier naturel.

1.2 listes et arrangements

Définition - propriété 1

- Une p-liste ou p-uplet de E est un élément de E^p, c'est-à-dire une liste $(x_1, x_2, ..., x_p)$ où pour tout $i \in [\![1, p]\!]$, $x_i \in E$. Dans une p-liste, l'ordre a de l'importance et il peut y avoir répétition. Le nombre de p-listes d'éléments de E est $(\operatorname{Card} E)^p$.
- Un p-arrangement de E est une p-liste d'éléments distincts. Dans un p-arrangement, l'ordre a de l'importance et on n'autorise pas les répétitions. En notant $n = \operatorname{Card} E$, le nombre de p-arrangements de E est

$$\begin{cases} n \times (n-1) \times \ldots \times (n-p+1) = \dfrac{n!}{(n-p)!} \text{ si } n \geqslant p \\ 0 \text{ sinon} \end{cases}$$

- Un n-arrangement de E est appelé permutation. Il y en a $n!$.

Les listes sont utilisées pour modéliser des tirages SUCCESSIFS AVEC REMISE.
Les arrangements sont utilisés pour modéliser des tirages SUCCESSIFS SANS REMISE.

Exercice 1 : Un code de carte bancaire est constitué de 4 chiffres entre 0 et 9. Combien y a-t-il :

— de tels codes ?

— de tels codes n'utilisant que des chiffres pairs ?

— de tels codes n'utilisant ni le 0 ni le 9 ?

— de tels codes commençant et finissant par un chiffre impair ?

Exercice 2 : On a n enfants. Combien y a-t-il de manières de les disposer en file indienne ? en ronde ?

Exercice 3 : E est toujours notre ensemble de cardinal n. Soit F un ensemble de cardinal p. Combien y a-t-il d'applications injectives de E dans F ?

1.3 combinaisons

Définition - propriété 2

On appelle p-combinaison de E toute partie de E de cardinal p. On note $\{x_1, x_2, ...x_p\}$ une telle partie.
Dans une partie, l'ordre n'a pas d'importance et il n'y a pas de répétition.
Le nombre de p-combinaisons de E est

$$\binom{n}{p} = \begin{cases} \dfrac{n!}{p!(n-p)!} \text{ si } p \leqslant n \\ 0 \text{ sinon} \end{cases}$$

Quand on décide de noter une combinaison, le choix de la numérotation est totalement arbitraire, la combinaison en tant que telle n'a pas de premier élément, deuxième élément, etc.
Les combinaisons sont utilisées pour modéliser des tirages SIMULTANÉS.

Méthode – tirages simultanés dans des sous-ensembles disjoints

Soient E_1, E_2,...E_p des sous-ensembles DISJOINTS d'un ensemble E, à respectivement N_1, N_2,...N_p éléments.
Le nombre de parties de E comportant n_1 éléments dans E_1, n_2 éléments dans E_2,... n_p éléments dans E_p est $\binom{N_1}{n_1} \times \binom{N_2}{n_2} \times ... \times \binom{N_p}{n_p}$.

Exercice 4 :
Une personne dispose en tas 20 vestes : 12 vestes foncées et 8 vestes claires. Les vestes sont mélangées, et la personne en prend 5 au hasard simultanément. Combien a-t-elle :

— de possibilités ?

— de choix comprenant (exactement) 3 vestes foncées ?

Exercice 5 :
On tire 3 cartes dans un jeu de 52 cartes. Combien de tirages :

— comportent un trèfle et un pique ? *(exactement un...)*

— comportent trois cartes de même hauteur ?

— comportent (en tout) un roi, une dame et un cœur ?

— comportent 1 paire et une carte d'une autre hauteur ? *(une paire = deux cartes de même hauteur...)*

Propriété 3 – propriétés des coefficients binomiaux

Soient $n \in \mathbb{N}$ et $k \in \mathbb{Z}$.

— $\binom{n}{n} = 1$, $\binom{n}{0} = 1$ et $\binom{n}{1} = n$

— Symétrie : $\binom{n}{p} = \binom{n}{n-k}$ Formule de Pascal : $\binom{n}{k} = \binom{n-1}{k} + \binom{n-1}{k-1}$

— Formule du capitaine : $n\binom{n-1}{k-1} = k\binom{n}{k}$ ou de factorisation : $\binom{n}{k} = \frac{n}{k}\binom{n-1}{k-1}$ pour $k \neq 0$.

Exercice 6 :
Soient N_1 et N_2 des entiers naturels. À l'aide d'un dénombrement, montrer la formule de Vandermonde :

$$\forall n \in \mathbb{N}, \quad \binom{N_1+N_2}{n} = \sum_{k=0}^{n} \binom{N_1}{k}\binom{N_2}{n-k}$$

En déduire $S_n = \sum_{k=0}^{n} \binom{n}{k}^2$.

1.4 approfondissements

Nombre de suites strictement croissantes de p éléments de E
Soit E un ensemble fini à n éléments. On voudrait savoir combien il y a de p-listes formées d'éléments de E rangés par ordre strictement croissant.
Par exemple, si $E = \{0, 1, 2, ...9\}$ et $p = 3$, $(1, 4, 7)$ et $(4, 5, 9)$ sont deux telles listes. On remarque que choisir une 3-liste formée de chiffres strictement croissants, c'est choisir 3 éléments distincts de E (et les positionner

dans l'ordre strictement croissant). Ainsi, le nombre cherché correspond dans l'exemple au nombre de choix de 3 éléments parmi 10, à savoir $\binom{10}{3}$.

Propriété 4 – (non exigible)

Soit E un ensemble fini à n éléments. Il y a $\binom{n}{p}$ p-listes formées d'éléments de E rangés par ordre strictement croissant, car à chaque partie à p éléments de E correspond une et une seule p-liste formée d'éléments strictement croissants de E.

Choix de positions dans une liste
Exercice 7 :
Combien y a-t-il de codes de « longueur » 8 comportant deux fois la lettre A, les six autres caractères étant des chiffres de $\{0, 1, \ldots, 9\}$?

Exercice 8 :

1. Combien y a-t-il d'anagrammes du mot VACANCES ?
2. Combien y a-t-il d'anagrammes du mot VACANCES dans lesquels les voyelles apparaissent avant les consonnes ?

2 Espace probabilisable

2.1 expérience aléatoire et univers

Une *expérience aléatoire* est une expérience qui ne dépend que du hasard (lancer d'un dé, même pipé ; tirage au sort...).
L'ensemble Ω des résultats, ou *issues*, possibles de l'expérience, est appelé *univers*.

Pour chacun des exemples suivants, donnons l'ensemble Ω des résultats possibles :

1. On lance une pièce de monnaie : $\Omega =$
2. On lance deux fois successives un dé à 6 faces : $\Omega =$
3. On lance une pièce jusqu'à ce que l'on obtienne Pile : $\Omega =$
4. Un joueur effectue une succession infinie de lancers d'une pièce de monnaie : $\Omega =$
5. On suit une molécule dans une enceinte gazeuse d'un centimètre cube ; au bout de 15 secondes, on note sa position (ses coordonnées).
   $\Omega =$

Un événement est une partie de Ω. Il y a en particulier l'événement impossible $\emptyset$, l'événement certain Ω, les événements élémentaires constitués d'un singleton $\{x\}$, $x \in \Omega$.
Des événements sont *incompatibles* lorsqu'ils ne peuvent pas être réalisés simultanément : $A \cap B = \emptyset$.

Langage probabiliste « français »	Langage ensembliste « mathématiques »	*exemple : un lancer d'un dé*
ensemble des résultats possibles	Ω	
événement	sous-ensemble de Ω	*obtenir un nombre pair :*
E_1 ou E_2 est réalisé	$E_1 \cup E_2$	*obtenir 1 ou 3 est*
les deux événements E_1 et E_2 sont réalisés	$E_1 \cap E_2$	
événement contraire	$\overline{E_1}$	
E_1 est réalisé, mais pas E_2	$E_1 - E_2 = E_1 \cap \overline{E_2}$	
au moins l'un des événements est réalisé	union $\bigcup_{i \in I} E_i$	
tous les événements sont réalisés	intersection $\bigcap_{i \in I} E_i$	
« lorsque E_1 est réalisé, alors E_2 l'est aussi »	$E_1 \subset E_2$	*si on obtient 1, a fortiori on obtient un nombre impair :*

2.2 cas d'un univers fini

Ici $\Omega = \{\omega_1, \omega_2, ...\omega_n\}$ est fini.

- Une probabilité sur l'espace $(\Omega, \mathcal{P}(\Omega))$ est une application $P : \mathcal{P}(\Omega) \longrightarrow [0,1]$ vérifiant :
  - $P(\Omega) = 1$
  - pour A et $B \in \mathcal{P}(\Omega)$ tels que $A \cap B = \emptyset$, on a $P(A \cup B) = P(A) + P(B)$.

  Le triplet $(\Omega, \mathcal{P}(\Omega), P)$ est appelé espace probabilisé fini.

- Soient p_1, p_2,...p_n des réels.
  Il existe une probabilité P sur $(\Omega, \mathcal{P}(\Omega))$ vérifiant $\forall i \in [\![1, n]\!]$, $P(\{\omega_i\}) = p_i$ si, et seulement si, les réels p_i sont positifs et de somme 1. Dans ce cas, la probabilité d'un événement A se calcule par

$$P(A) = \sum_{i/\omega_i \in A} P(\{\omega_i\}) = \sum_{i/\omega_i \in A} p_i$$

  Par exemple, dans un lancer de dé,

$$P(\{\textit{obtenir un nombre pair}\}) = P(\{2\}) + P(\{4\}) + P(\{6\})$$

- L'espace probabilisable $(\Omega, \mathcal{P}(\Omega))$ est muni de la *probabilité uniforme* ou *équiprobabilité* si tous les événements élémentaires ont même probabilité :

$$P(\{\omega_i\}) = \frac{1}{n} \quad \text{pour tout } i \in [\![1, n]\!]$$

  On calcule dans ce cas la probabilité d'un événement $A \in \mathcal{P}(\Omega)$ par :

$$P(A) = \frac{\text{nombre de cas favorables à } A}{\text{nombre de cas possibles}} = \frac{\text{Card}(A)}{\text{Card}(\Omega)}$$

Exercice 9 : On dispose d'un jeu de 52 cartes et on effectue un tirage simultané de 5 cartes.
Soit A l'événement : « on obtient (exactement) 2 rois, une dame et un valet ». Calculer sa probabilité.

Exercice 10 : Dans un jeu de 32 cartes, une main est un ensemble de 6 cartes non ordonnées, tirées au hasard. Donner la probabilité des événements :
A « on obtient 2 brelans » (un brelan est constitué de 3 cartes de même hauteur)
B « on obtient une paire et un carré »

2.3 notion de tribu

Un peu de culture générale :

Les travaux de Kolmogorov (*années 1930*) en théorie des probabilités l'ont amené à chercher les conditions requises sur l'ensemble $\mathcal{A}$ des événements liés à l'expérience aléatoire pour qu'on puisse y définir une application probabilité (satisfaisant aux propriétés intuitives d'une probabilité, notamment l'idée d'additivité). Un point délicat de la théorie des probabilités est qu'on ne considère pas toujours toutes les parties de Ω comme des événements : dans certains cas $\mathcal{A} \subset \mathcal{P}(\Omega)$ mais $\mathcal{A} \neq \mathcal{P}(\Omega)$.

Considérons l'exemple de la position aléatoire x d'une particule dans l'intervalle $\Omega = [0, 1]$. Il serait naturel de définir $P($« x appartient à $]a; b[$ »$) = b - a$ et plus généralement, pour A partie de Ω, $P([x \in A]) = \text{longueur}(A)$. Toutefois, il existe des parties de $[0, 1]$ dont on ne parvient pas à donner une définition de longueur (qu'on ne parvient pas à « mesurer »)... Ainsi, on dût se limiter à définir les probabilités non pas sur $\mathcal{P}(\Omega)$ entier, mais sur un sous-ensemble restreint (tribu $\mathcal{A}$) vérifiant certaines propriétés algébriques de stabilité...

Quelles sont les conditions requises pour $\mathcal{A}$?
On va demander de pouvoir effectuer sur les événements de $\mathcal{A}$ les opérations d'union (le *ou* en français), d'intersection (le *et* du français), de passage au complémentaire (pour pouvoir considérer les *événements contraires*) autant de fois dénombrables qu'on le désire en restant dans le même ensemble d'événements $\mathcal{A}$.

Définition 2

On appelle *tribu* sur Ω tout sous-ensemble $\mathcal{A}$ de $\mathcal{P}(\Omega)$ vérifiant :

— Ω appartient à $\mathcal{A}$

— $\mathcal{A}$ est stable par passage au complémentaire : si $A \in \mathcal{A}$, alors $\overline{A} \in \mathcal{A}$

— $\mathcal{A}$ est stable par union dénombrable :
si $(A_n)_{n\in\mathbb{N}}$ est une suite d'événements de $\mathcal{A}$, alors $\bigcup\limits_{n\in\mathbb{N}} A_n$ appartient à $\mathcal{A}$.

Un ensemble d'une tribu est appelé *événement*.
On appelle espace probabilisable associé à une expérience aléatoire la donnée de l'ensemble des résultats possibles Ω et d'une tribu $\mathcal{A}$ d'événements.

L'ensemble $\{\emptyset, \Omega\}$ est une tribu de Ω. Elle n'est pas intéressante car elle ne permet de considérer que l'événement impossible et l'événement certain.

L'ensemble $\mathcal{P}(\Omega)$ des parties de Ω est une tribu de Ω. Elle n'est pas toujours intéressante, car dans certains cadres, elle est un peu trop grosse et il est impossible de définir une probabilité sur cette tribu. Néanmoins, si Ω est fini ou dénombrable, $\mathcal{P}(\Omega)$ est une tribu.

Propriété 5

Soit $\mathcal{A}$ une tribu. On a les propriétés :

— $\mathcal{A}$ contient $\emptyset$

— $\mathcal{A}$ est stable par union finie : si A_1, A_2,...A_n sont dans $\mathcal{A}$, alors $A_1 \cup A_2 \cup ... \cup A_n$ appartient à $\mathcal{A}$

— $\mathcal{A}$ est stable par intersection finie : si A_1, A_2,...A_n sont dans $\mathcal{A}$, alors $A_1 \cap A_2 \cap ... \cap A_n$ appartient à $\mathcal{A}$

— soit $(A_i)_{i\in\mathbb{N}}$ est une famille d'événements de $\mathcal{A}$, alors $\bigcap\limits_{i\in\mathbb{N}} A_i$ appartient à $\mathcal{A}$.

Définition 3

Soit $(\Omega, \mathcal{A})$ un espace probabilisable. On appelle *système complet d'événements*, toute famille, finie ou infinie dénombrable, $(A_i)_{i\in I}$ telle que :

— les A_i sont des événements : $A_i \in \mathcal{A}$

— les A_i sont deux à deux incompatibles : $A_i \cap A_j = \emptyset$ pour $i \neq j$

— l'union des A_i est Ω : $\Omega = \bigcup\limits_{i\in I} A_i$.

Exercice 11 : Écrire à l'aide d'unions, intersections, complémentaires et des événements A, B, C, $A_1, A_2, \ldots$ les événements suivants :

1. L'un au moins des événements A, B, C est réalisé.
2. L'un et l'un seulement des événements A, B, C est réalisé.
3. A et B se réalisent mais pas C.
4. Tous les événements A_n sont réalisés à partir du rang 10.
5. Seul un nombre fini d'événements A_n, $n \geqslant 1$, se réalise.
6. Il y a une infinité des événements A_n, $n \geqslant 1$, qui se réalisent.

3 Application probabilité sur un espace probabilisable

Soit $(\Omega, \mathcal{A})$ un espace probabilisable.

Définition 4

On appelle *probabilité* toute application $P : \mathcal{A} \longrightarrow [0,1]$ vérifiant :

— $P(\Omega) = 1$

— P est $\sigma-$additive, c'est-à-dire que si $(A_n)_{n\in\mathbb{N}}$ est une suite d'événements 2 à 2 incompatibles, alors : $P(\bigcup\limits_{n=0}^{\infty} A_n) = \sum\limits_{n=0}^{\infty} P(A_n)$

L'espace probabilisable $(\Omega, \mathcal{A})$ muni d'une probabilité P est appelé espace probabilisé.

En prenant $A_n = \emptyset$ pour tout n, on obtient $P(\emptyset) = \sum\limits_{n=0}^{\infty} P(\emptyset)$ puis $P(\emptyset) = 0$.

La σ-additivité nous donne, en prenant $A_k = \emptyset$ pour $k > N$, la propriété d'additivité :

$$\text{pour } A_1, A_2, ..., A_N \text{ deux à deux incompatibles, } P(A_1 \cup A_2 \cup ... \cup A_N) = P(A_1) + P(A_2) + ... + P(A_N)$$

On retrouve en particulier, pour A et B événements incompatibles, $P(A \cup B) = P(A) + P(B)$.

Propriété 6

On a les propriétés élémentaires :

1. $P(\overline{A}) = 1 - P(A)$
2. Si $A \subset B$, $\quad P(B \setminus A) = P(B) - P(A)$
3. Si $A \subset B$, $\quad P(A) \leqslant P(B)$. Cette propriété s'appelle la croissance de P.
4. $P(A \cup B) = P(A) + P(B) - P(A \cap B)$

Théorème 1 – théorème de continuité monotone

- Soit $(A_n)_{n\in\mathbb{N}}$ une suite croissante d'événements, c'est-à-dire telle que $A_n \subset A_{n+1}$. Alors :
$$P\left(\bigcup_{n=0}^{\infty} A_n\right) = \lim_{n\to+\infty} P(A_n)$$
- Soit $(A_n)_{n\in\mathbb{N}}$ une suite décroissante d'événements, c'est-à-dire telle que $A_{n+1} \subset A_n$. Alors :
$$P\left(\bigcap_{n=0}^{\infty} A_n\right) = \lim_{n\to+\infty} P(A_n)$$

Corollaire 1 – ♡

Pour toute suite (A_n) d'événements,
$$P\left(\bigcup_{n=0}^{\infty} A_n\right) = \lim_{n\to+\infty} P\left(\bigcup_{k=0}^{n} A_k\right) \quad \text{et} \quad P\left(\bigcap_{n=0}^{\infty} A_n\right) = \lim_{n\to+\infty} P\left(\bigcap_{k=0}^{n} A_k\right)$$

Exercice 12 : On effectue une succession infinie de lancers d'une pièce équilibrée. On admet qu'il existe une tribu $\mathcal{A} \subset \mathcal{P}(\Omega)$ et une probabilité P sur $(\Omega, \mathcal{A})$ qui coïncide avec la probabilité uniforme sur l'univers correspondant aux n premiers lancers, pour tout n.
A_n désigne l'événement « les n premiers lancers donnent tous Face ».

1. Soit $n \in \mathbb{N}^*$. Calculer la probabilité de A_n.
2. Soit A l'événement « on obtient toujours face ». Calculer sa probabilité.

Exercice 13 : On lance indéfiniment une pièce donnant Pile avec probabilité $p \in]0;1[$ et Face avec probabilité $q = 1 - p$. On note S_k l'événement « on obtient Pile pour la première fois au k^{e} lancer ».

1. Calculer la probabilité de l'événement A « on obtient au moins une fois Pile » en utilisant les événements S_k.
2. Soit B l'événement « on obtient Pile pour la première fois au bout d'un nombre pair de lancers ». Montrer que $P(B) = \frac{q}{1+q}$.

Propriété 7 – sous-additivité de P

- $P(A_1 \cup A_2 \cup \cdots \cup A_n) \leqslant P(A_1) + P(A_2) + \cdots + P(A_n)$.
- $P\left(\bigcup_{k=0}^{+\infty} A_k\right) \leqslant \sum_{k=0}^{+\infty} P(A_k) \leqslant +\infty$

On peut donner un sens à l'inégalité précédente même si la série $\sum_k P(A_k)$ diverge, puisqu'étant à termes positifs, si la série diverge, sa somme est $+\infty$.
Plus généralement, pour $(A_i)_{i\in I}$ famille au plus dénombrable d'événements, on a :

$$P\left(\bigcup_{i\in I} A_i\right) \leqslant \sum_{i\in I} P(A_i) \leqslant +\infty$$

Définition 5

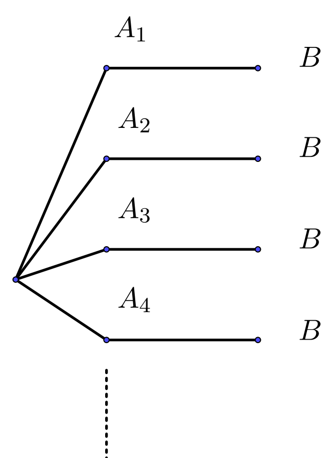
On appelle *système quasi-complet d'événements* toute famille au plus dénombrable $(A_i)_{i\in I}$ d'événements deux à deux incompatibles tels que $P\left(\bigcup_{i\in I} A_i\right) = 1$.

En particulier, un système complet d'événements est un système quasi-complet d'événements.

Propriété 8 – formule des probabilités totales

Pour tout système quasi-complet d'événements $(A_i)_{i\in I}$ et tout événement B, on a :

$$P(B) = \sum_{i\in I} P(A_i \cap B)$$



Définition 6

- Tout événement de probabilité nulle est dit *négligeable*.
- Tout événement de probabilité 1 est dit *presque sûr* ou *presque certain*.

Propriété 9

- Une réunion au plus dénombrable d'événements négligeables est un événement négligeable.
- Une intersection au plus dénombrable d'événements presque sûrs est un événement presque sûr.

4 probabilités conditionnelles et indépendance

4.1 Trois grandes formules de probabilité

Définition 7

Soit $(\Omega, \mathcal{A}, P)$ un espace probabilisé et A un événement tel que $P(A) \neq 0$. On appelle *probabilité conditionnelle sachant* A l'application P_A définie sur $\mathcal{A}$ par : $P_A(B) = \dfrac{P(A \cap B)}{P(A)}$. On peut aussi noter $P_A(B) = P(B|A)$.

Propriété 10

P_A est une probabilité, autrement dit $(\Omega, \mathcal{A}, P_A)$ est un espace probabilisé. Par conséquent, **toutes** les propriétés relatives aux espaces probabilisés sont valables pour P_A.

Propriété 11 – formule des probabilités composées

Soient A_1, A_2,..., A_n des événements tels que $P(A_1 \cap A_2 \cap ... \cap A_{n-1}) \neq 0$. On a :

$$P\left(\bigcap_{i=1}^{n} A_i\right) = P(A_1) \times P_{A_1}(A_2) \times P_{A_1 \cap A_2}(A_3) \times P_{A_1 \cap A_2 \cap A_3}(A_4) \times ... \times P_{A_1 \cap A_2 \cap ... A_{n-1}}(A_n)$$

Exercice 14 : Dans cet exercice, a, b désignent des entiers naturels non nuls. n est un entier supérieur ou égal à 2.
Un panier de fraises contient a fraises mûres et b fraises pas mûres. On tire au hasard du panier n fraises successivement, en procédant comme suit :

- si une fraise est mûre, on la mange,
- sinon, on la remet dans le panier.

On note :
M_k « la k^{e} fraise tirée est mûre »
E_k « la k^{e} fraise tirée est mûre et c'est la seule fraise mûre tirée au cours des n tirages »
A « on n'a mangé aucune fraise » et B « on a mangé exactement une fraise »

Calculer $P(A)$. Calculer $P(E_k)$. Calculer $P(B)$.

Propriété 12 – formule des probabilités totales

Soit J une partie de $\mathbb{N}$ et $(B_j)_{j \in J}$ un système complet d'événements de probabilités non nulles. Alors pour tout événement A, on a :

$$P(A) = \sum_{j \in J} P(B_j) P_{B_j}(A)$$

Cette formule reste vraie pour un système quasi-complet d'événements.

Pour éviter les complications, on convient souvent dans cette formule que $P(B_j)P_{B_j}(A) = 0$ si $P(B_j) = 0$, ce qui permet de l'utiliser même sans savoir que les événements du système complet sont de probabilités non nulles.

Exercice 15 : On veut enquêter sur la proportion (x, inconnue) de personnes qui volent dans les grandes surfaces. La question est « Avez-vous volé dans un magasin au cours des douze derniers mois ? ». On comprend que la personne interrogée hésite à répondre oui. Afin de mettre en confiance les personnes sondées pour qu'elles ne mentent pas, on met en œuvre le protocole suivant.
On appelle *question 1* la question précédente, et *question 2* la question « Avez-vous pris des vacances à plus de 100 kilomètres de chez vous au cours des douze derniers mois ? ». L'enquêteur demande à la personne interrogée de jouer à Pile ou Face, hors de sa présence, et de répondre à la question 1 si elle obtient Pile et à la question 2 sinon. La personne interrogée donne seulement sa réponse sans dire à quelle question elle a répondu.
Une enquête indépendante montre que le taux de réponses affirmatives à la question 2 est 60%.
À la fin des sondages réalisés, l'enquêteur obtient un taux de réponses affirmatives (pour les deux questions) égal à t. Trouver alors la valeur de x apportée par ce sondage.

Propriété 13 – formule de Bayes

On dispose d'un système complet d'événements $(B_j)_{j\in J}$ et d'un événement A de probabilité non nulle. On a :

$$P_A(B_i) = \frac{P(B_i)P_{B_i}(A)}{P(A)} = \frac{P(B_i)P_{B_i}(A)}{\sum\limits_{j\in J} P(B_j)P_{B_j}(A)}$$

Exercice 16 : Pour se rendre au travail, un travailleur a le choix entre 3 solutions :

— il y va en RER (solution 1, qu'il utilise avec la probabilité 0,45) ;

— il y va en vélo (solution 2, qu'il utilise avec la probabilité 0,3) ;

— il y va en bus (solution 3).

S'il utilise la solution 1, il est retardé avec une probabilité de 20%. Cette probabilité vaut 10% et 30% respectivement pour les solutions 2 et 3.
Un matin, il a été retardé. Quelle est la probabilité qu'il soit venu en bus ?

4.2 événements indépendants

Définition 8

Soit $(\Omega, \mathcal{A}, P)$ un espace probabilisé. Les événements A et B sont indépendants si

$$P(A \cap B) = P(A) \times P(B)$$

Si A est de probabilité non nulle, A et B sont indépendants si, et seulement si, $P_A(B) = P(B)$. Autrement dit, si savoir que A est réalisé n'influence pas le calcul de la probabilité de B.
Dans la pratique, il y a deux types de questions. Il y a des exercices où l'on doit démontrer que A et B sont indépendants, en calculant $P(A)$, $P(B)$ et $P(A \cap B)$ et en regardant si on a l'égalité voulue pour l'indépendance. Et il y a d'autres exercices, où l'on invoque d'emblée une indépendance classique, garantie en général dans l'énoncé (lancers de pièce, tirages avec remise...).

Définition 9

Soit $(A_i)_{i\in I}$ une famille d'événements.

- Les événements sont 2 à 2 indépendants si $\forall i \neq j$, les événements A_i et A_j sont indépendants.
- Les événements sont (mutuellement) indépendants si pour toute partie finie J de I on a :

$$P\left(\bigcap_{j\in J} A_j\right) = \prod_{j\in J} P(A_j)$$

Exercice 17 : On considère l'expérience aléatoire consistant à lancer deux dés. On munit l'univers de la probabibilité uniforme. On considère les événements suivants :

A : « le résultat du premier dé est pair »

B : « le résultat du deuxième dé est pair »

C : « la somme des résultats des deux dés est pair ».

Montrer que les événements A, B et C sont deux à deux indépendants mais pas mutuellement indépendants.

Propriété 14

Si A et B sont deux événements indépendants, alors les événements A et $\overline{B}$ sont indépendants.

Corollaire 2

Si $(A_i)_{i\in I}$ est une famille d'événements indépendants, et si pour tout indice i, $B_i = A_i$ ou $\overline{A_i}$, alors $(B_i)_{i\in I}$ est aussi une famille d'événements indépendants.

Exercice 18 : On considère deux urnes :

U_1 contenant n_1 boules noires et b_1 boules blanches ;

U_2 contenant n_2 boules noires et b_2 boules blanches.

On choisit de manière équiprobable une urne, et on y effectue deux tirages successifs d'une boule avec remise. On note A (respectivement B) l'événement « on tire une boule noire au premier (resp. au second) tirage ».

1. Calculer $P(A)$ et $P(B)$.
2. À quelle condition nécessaire et suffisante sur n_1, n_2, b_1, b_2, les événements A et B sont-ils indépendants ?

5 Retour sur les espaces probabilisés

Définition 10

Une *distribution de probabilités discrète* sur un ensemble Ω est une famille sommable $(p_\omega)_{\omega\in\Omega}$ de réels positifs, de somme 1.

Son *support* est l'ensemble $\{\omega \in \Omega \,|\, p_\omega \neq 0\}$. Cet ensemble est au plus dénombrable.

Au chapitre Familles sommables, on a vu que le support d'une famille sommable de nombres complexes était dénombrable, c'est-à-dire que si $\sum\limits_{i\in I} |a_i| < +\infty$, alors $\{i \in I \mid a_i \neq 0\}$ est au plus dénombrable. Et donc en fait, cela revient à sommer uniquement sur un ensemble dénombrable.

Propriété 15

Soit Ω un ensemble au plus dénombrable. Si $(p_\omega)_{\omega\in\Omega}$ est une distribution de probabilités discrète, il existe une unique probabilité P définie sur $(\Omega, \mathcal{P}(\Omega))$ telle que pour tout $\omega \in \Omega$, $p_\omega = P(\{\omega\})$; elle est définie par :

$$P(A) = \sum_{\omega\in A} p_\omega$$

L'espace probabilisé $(\Omega, \mathcal{P}(\Omega), P)$ est appelé *espace probabilisé discret*.

Exercice 19 :

1. On lance un dé à 6 faces numérotées de 1 à 6 pour lequel la probabilité d'apparition d'une face est proportionnelle au numéro de la face. Déterminer la distribution de probabilité discrète correspondant à cet énoncé.
2. Montrer que $((\frac{1}{2})^n)_{n\in\mathbb{N}^*}$ est une distribution de probabilité discrète. Donner l'application probabilité associée. Montrer que pour cette probabilité, il est plus probable d'obtenir un nombre impair qu'un nombre pair.

6 Variables aléatoires discrètes

6.1 généralités

Définition 11

Soient $(\Omega, \mathcal{A})$ un espace probabilisable et E un ensemble non vide. On appelle *variable aléatoire discrète* toute application $X : \Omega \to E$ vérifiant :

- $X(\Omega)$ est au plus dénombrable
- pour tout $x \in X(\Omega)$, $X^{-1}(\{x\}) \in \mathcal{A}$.

L'événement $X^{-1}(\{x\}) = \{\omega \in \Omega \mid X(\omega) = x\}$ est noté $(X = x)$ ou $\{X = x\}$.

Remarques :

— L'appellation variable aléatoire est malheureuse. En effet, X n'est pas une variable, mais une fonction et celle-ci n'est pas aléatoire, mais parfaitement déterminée. Ce sont les valeurs de X qui varient selon le résultat de l'expérience aléatoire. Par exemple, X est la variable aléatoire égale au nombre de fois où on obtient 6 quand on lance 10 fois un dé.
On lance 10 fois le dé, on obtient 2 six, $X(\omega) = 2$ pour ce résultat ω.

— Lorsque Ω est dénombrable et que $\mathcal{A} = \mathcal{P}(\Omega)$, toute fonction de Ω dans E est une variable aléatoire, il n'y a rien à vérifier de plus.

— Lorsque $E = \mathbb{R}$, on parle de variable aléatoire réelle discrète.

— Lorsque $X(\Omega)$ est un ensemble fini, on parle de variable aléatoire finie.

— Une fonction constante sur Ω est une variable aléatoire finie. En effet, il existe a tel que $X(\Omega) = \{a\}$. On a $(X = a) = \Omega \in \mathcal{A}$.

Propriété 16

Soit X une variable aléatoire discrète sur $(\Omega, \mathcal{A})$ et à valeurs dans E.
Pour toute partie A de E, l'ensemble :

$$\{X \in A\} = (X \in A) = \{\omega \in A \mid X(\omega) \in A\} = X^{-1}(A)$$

est un événement.

Dans la pratique, on ne revient pas aux ω.
Par exemple, si on lance une pièce et que X est le temps d'attente du premier face, alors

$(X \geqslant 5) =$

Autre exemple. On lance trois fois un dé. X est la somme des numéros obtenus. On a

$$\begin{aligned}
(X \leqslant 2) &= \\
(X = 3) &= \\
(X = 18) &= \\
(X = 17) &=
\end{aligned}$$

Propriété 17

Soit X une variable aléatoire discrète. La famille $\{(X = x)\}_{x \in X(\Omega)}$ est un système complet d'événements, appelé *système complet d'événements associé à X*.

6.2 loi de probabilité

Définition - propriété 3 – loi de X

Soit X une variable aléatoire discrète sur l'espace probabilisé $(\Omega, \mathcal{A}, P)$. La *loi de X* est :

$$P_X : \left(\begin{array}{lcl} \mathcal{P}(X(\Omega)) & \rightarrow & [0,1] \\ A & \mapsto & P(X \in A) \end{array} \right)$$

P_X est une probabilité sur $(X(\Omega), \mathcal{P}(X(\Omega)))$.

> Il s'ensuit que
> $$P(X \in A) = \sum_{x \in A} P(X = x)$$
> et donc que la loi de X est déterminée par la distribution de probabilité discrète $(P(X = x))_{x \in X(\Omega)}$.

Réciproquement, la propriété suivante montre que toute probabilité discrète sur un ensemble peut être considérée comme la loi d'une variable aléatoire discrète. C'est une propriété importante pour introduire de nouvelles lois théoriques.

Propriété 18 – existence d'une variable aléatoire de distribution de probabilités donnée

Soit $(p_x)_{x\in E}$ une distribution de probabilités définie sur un ensemble E au plus dénombrable.
Il existe un espace probabilisé $(\Omega, \mathcal{A}, P)$ et une variable aléatoire discrète X de loi donnée par $(p_x)_{x\in E}$.

— Si une variable X suit une certaine loi de probabilité $\mathcal{L}$, on note $X \sim \mathcal{L}$.

— Deux variables X et Y ont même loi, et on note $X \sim Y$, si $P_X = P_Y$, c'est-à-dire si $X(\Omega) = Y(\Omega)$ et pour tout $x \in X(\Omega)$, $P(X = x) = P(Y = x)$.

Dire que $X \sim Y$ ne présuppose pas que les variables X et Y sont définies sur le même univers, ni donc que $X = Y$. Par exemple :

- Soit X une variable aléatoire suivant une loi de Rademacher, c'est-à-dire que
$$P(X = 1) = P(X = -1) = \frac{1}{2}$$
On a $X \sim -X$.
- Dans n lancers d'une pièce équilibrée, le nombre X de piles obtenus et le nombre Y de faces obtenues suivent la même loi mais ne sont pas égales.

Définition 12 – loi conditionnelle sachant un événement

Soit A un événement de probabilité non nulle et X une variable aléatoire discrète.
La *loi conditionnelle de X sachant A* est la loi de X dans l'espace probabilisé $(\Omega, \mathcal{A}, P_A)$. Elle est donc déterminée par des probabilités $P_A(X = x)$ pour $x \in X(\Omega)$.

6.3 transformée d'une variable aléatoire

Définition - propriété 4

Pour X variable aléatoire discrète et f une application définie sur un ensemble contenant $X(\Omega)$, $f \circ X$ est une variable aléatoire discrète, notée $f(X)$.
Si $X \sim Y$, alors $f(X) \sim f(Y)$.

7 Lois usuelles

7.1 loi certaine

Définition 13

On dit que la variable aléatoire X suit la loi certaine ou qu'elle est presque sûrement constante s'il existe $a \in E$ tel que $P(X = a) = 1$.

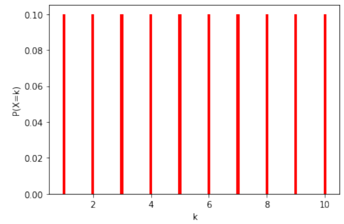
7.2 loi uniforme

Définition - propriété 5

On dit que la variable aléatoire X suit une *loi uniforme* sur un ensemble fini E si :
$$X(\Omega) = E \text{ et pour tout } x \in E,\ P(X = x) = \frac{1}{\operatorname{Card} E}$$

Pour $X \sim U(\llbracket 1, n \rrbracket)$, $X(\Omega) = \llbracket 1, n \rrbracket$ et $\quad \forall k \in X(\Omega),\ P(X = k) = \frac{1}{n}$.

Situation type : une urne contient n boules numérotées de 1 à n. On en prend une au hasard et on note X le numéro tiré.



Une simulation en Python :

```
import numpy.random as rd
X = rd.randint(1,n+1)
```

7.3 loi de Bernoulli

Définition - propriété 6

On dit que la variable aléatoire X suit la *loi de Bernoulli de paramètre* $p \in [0, 1]$ si :

$$X(\Omega) = \{0, 1\} \quad \text{et } P(X = 1) = p$$

On note $X \sim b(p)$ ou $X \sim \mathcal{B}(1, p)$.

Situation type
On tire une boule dans une urne avec n boules noires et b boules blanches. X est le nombre de boules noires obtenues en un tirage, $X \sim b(p)$ où $p = \frac{n}{b+n}$.

Une simulation en Python :

```
import numpy.random as rd
if rd.random()<p:
    X = 1
else:
    X = 0
```

Définition 14

Pour A événement, la variable aléatoire $\mathbb{1}_A$, indicatrice de l'événement A :

$$\mathbb{1}_A(\omega) = \begin{cases} 1 & \text{si } \omega \in A \\ 0 & \text{sinon} \end{cases}$$

suit la loi de Bernoulli de paramètre $p = P(A)$.

7.4 loi binomiale

Définition - propriété 7

On dit que X suit une *loi binomiale de paramètres* $n \in \mathbb{N}^*$ *et* $p \in [0, 1]$, et on note $X \sim \mathcal{B}(n, p)$, lorsque :

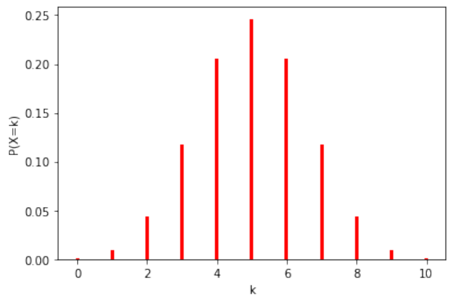
$$X(\Omega) = \llbracket 0, n \rrbracket \text{ et } P(X = k) = \binom{n}{k} p^k (1-p)^{n-k} \text{ pour } k \in \llbracket 0, n \rrbracket$$

Situation type
On réalise une succession de n épreuves de Bernoulli, de probabilité de succès $p \in [0, 1]$, indépendantes. La variable aléatoire égale au nombre de succès obtenus suit une loi binomiale de paramètres n et p.

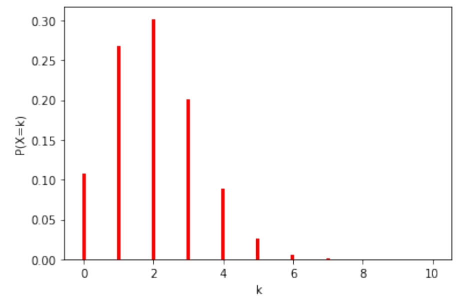
Une simulation en Python :

```
import numpy.random as rd
X = rd.binomial(n,p)
```

Loi binomiale $\mathcal{B}(10, \frac{1}{2})$



Loi binomiale $\mathcal{B}(10, \frac{1}{5})$



7.5 loi géométrique

Définition - propriété 8

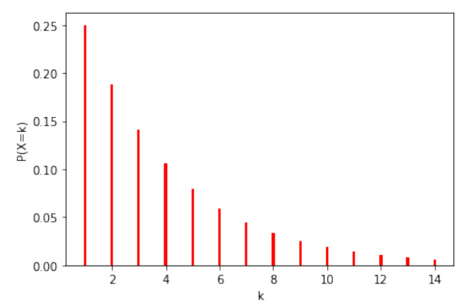
On dit que X suit une *loi géométrique de paramètre* $p \in]0, 1[$, et on note $X \sim \mathcal{G}(p)$, lorsque

$$X(\Omega) = \mathbb{N}^* \text{ et } \quad \forall k \in \mathbb{N}^*,\ P(X = k) = p(1-p)^{k-1}$$

Simulation en Python

```
import numpy.random as rd
X = rd.geometric(p)
```

Loi géométrique $\mathcal{G}(1/4)$



Propriété 19 – modèle du temps d'attente

On réalise des épreuves de Bernoulli successives indépendantes de probabilité de succès $p \in]0, 1[$.
Soit X la variable aléatoire égale au nombre d'épreuves nécessaires à l'obtention du premier succès.
On a $X \sim \mathcal{G}(p)$.

✎ Exercice 20 : Soit $X \sim \mathcal{G}(p)$.

1. Montrer que
$$\forall (m, n) \in \mathbb{N}^2,\ P_{[X>m]}(X > m + n) = P(X > n)$$

2. Si l'on pense à X comme à une durée, X ne tient pas compte du passé. On dit que la loi géométrique est « sans mémoire », « sans vieillissement ». Illustrer ce propos avec des valeurs de n et m bien choisies dans l'égalité précédente.

7.6 loi de Poisson

Définition 15

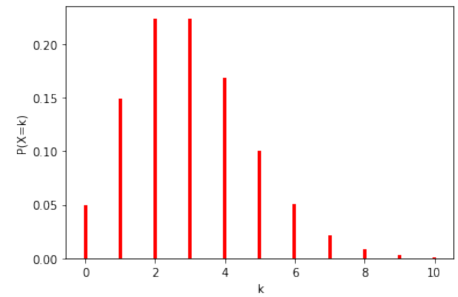
On dit qu'une variable aléatoire réelle discrète X suit *une loi de Poisson de paramètre* $\lambda > 0$, et on note $X \sim \mathcal{P}(\lambda)$, lorsque

$$X(\Omega) = \mathbb{N} \text{ et pour } k \in \mathbb{N}, P(X = k) = e^{-\lambda} \frac{\lambda^k}{k!}$$

Simulation en Python

```python
import numpy.random as rd
X = rd.poisson(a)
```

Loi de Poisson $\mathcal{P}(3)$



Exercice 21 : Soit $X \sim \mathcal{P}(\lambda)$. Déterminer, en fonction du paramètre $\lambda > 0$, le ou les modes de X, c'est-à-dire la ou les valeurs de k pour lesquelles la probabilité $P(X = k)$ est maximale.

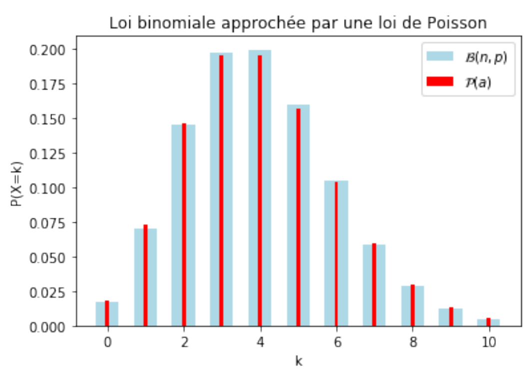
Exercice 22 : **Loi des événements rares**

1. Soit $(X_n)_{n \in \mathbb{N}^*}$ une suite de variables aléatoires discrètes telles que X_n suive la loi binomiale $\mathcal{B}(n, p_n)$, où $\lim\limits_{n \to +\infty} np_n = \lambda > 0$. Montrer que pour tout $k \in \mathbb{N}$,

$$\lim_{n \to +\infty} P(X_n = k) = \frac{\lambda^k e^{-\lambda}}{k!}$$

   Ainsi, si X suit une loi binomiale avec n grand et p proche de 0, X suit approximativement une loi de Poisson de paramètre $\lambda = np$.

2. Expliquer, dans cette optique, pourquoi on peut modéliser par une loi de Poisson :
   - le nombre de voitures se présentant à un péage pendant un laps de temps,
   - le nombre de clients se présentant dans un magasin.



```python
import numpy as np
n = 100
p = 0.04
a = n*p
def facto(k):
 return np.math.factorial(k)
def binomiale(k):
    coeff = facto(n)/facto(k)/facto(n-k)
    return coeff*p**k*(1-p)**(n-k)
def Poisson(k):
    return np.exp(-a)*a**k/facto(k)

k = np.arange(0,11)
probas_B = [binomiale(i) for i in k]
probas_P = [Poisson(i) for i in k]
g1 = plt.bar(k, probas_B, width=0.6, color='lightblue')
g2 = plt.bar(k,probas_P, width=0.1, color='red')
plt.xlabel('k')
plt.ylabel('P(X=k)')
plt.title('Loi binomiale approchée par une loi de Poisson')
plt.legend([g1,g2],['$\mathcal{B}(n,p)$','$\mathcal{P}(a)$'])
plt.show()
```

8 Couples et uplets de variables aléatoires

8.1 couples de variables aléatoires discrètes

Un couple est une variable aléatoire à valeurs dans un produit. Le produit d'espaces au plus dénombrables est au plus dénombrable, donc on reste bien dans le cadre de variables aléatoires discrètes.

Définition - propriété 9

Soient X, Y deux variables aléatoires discrètes définies sur un espace probabilisé $(\Omega, \mathcal{A}, P)$ et à valeurs respectivement dans E et F. L'application $(X, Y) : \Omega \to E \times F$ est une variable aléatoire discrète, appelée *loi conjointe* de X et Y.
Donner la loi du couple (X, Y), c'est donner $X(\Omega)$, $Y(\Omega)$, et les valeurs :

$$P(X = x, Y = y) = P((X = x) \cap (Y = y)) \text{ pour tout } (x, y) \in X(\Omega) \times Y(\Omega)$$

— Par la propriété 4, à partir d'un couple de variables aléatoires discrètes, on a beaucoup de variables aléatoires transformées : $X + Y$, XY, $\max(X, Y)$,...

— La famille $((X = x) \cap (Y = y))_{(x,y)\in X(\Omega)\times Y(\Omega)}$ est un système complet d'événements. On a

$$\sum_{(x,y)\in X(\Omega)\times Y(\Omega)} P(X = x, Y = y) = 1$$

— Réciproquement, des réels $p_{x,y}$ forment les coefficients d'une loi de couple s'ils sont positifs et de somme 1.

— Dans des exercices simples, on peut utiliser un tableau pour présenter les probabilités de la loi conjointe.

Exercice 23 : Une urne comporte 2 boules blanches, 1 rouge et 1 noire. On tire simultanément deux boules de cette urne et on note X le nombre de boules blanches et Y le nombre de boules noires tirées. Compléter le tableau donnant la loi de (X, Y).

	$X=0$	$X=1$	$X=2$	
$Y=0$				
$Y=1$				
$Y=2$				

Par la formule des probabilités totales appliquée avec le système complet d'événements $([Y=y])_{y\in Y(\Omega)}$, on peut obtenir la loi (dite loi marginale) de X :

$$P(X=x)=\sum_{y\in Y(\Omega)} P(X=x,Y=y)$$

et de même, avec le système complet d'événements $([X=x])_{x\in X(\Omega)}$, on peut obtenir la loi marginale de X :

$$P(Y=y)=\sum_{x\in X(\Omega)} P(X=x,Y=y)$$

La réciproque est fausse : avec les lois marginales, on ne peut pas reconstruire la loi conjointe.

Exercice 24 : p est un réel de $]0,1[$. On pose $q=1-p$. λ est un réel positif. On suppose qu'un jour donné, le nombre X de clients d'un marchand de fruits et légumes suit une loi de Poisson de paramètre λ. Chaque client achète : soit des fruits, avec une probabilité p, soit des légumes avec probabilité $q=1-p$. On suppose que les choix des clients sont indépendants.
Soit Y le nombre de clients achetant des fruits et Z le nombre de clients achetant des légumes.

1. (a) Donner la loi conditionnelle de Y sachant $[X=n]$, pour $n\in\mathbb{N}^*$.
   (b) En déduire la loi conjointe de (X,Y).
2. Déterminer la loi de Y.
3. Déterminer la loi de $X-Y$.

Définition 16

Soient X,Y deux variables aléatoires discrètes définies sur un espace probabilisé $(\Omega,\mathcal{A},P)$ et à valeurs respectivement dans E et F.
Les variables aléatoires X et Y sont *indépendantes* si pour toute partie A de E et toute partie B de F, les événements $(X\in A)$ et $(Y\in B)$ sont indépendants.

Propriété 20 – caractérisation de l'indépendance

Les variables aléatoires discrètes X et Y sont indépendantes si et seulement si

$$\text{pour tout } (x,y)\in X(\Omega)\times Y(\Omega),\quad P(X=x,Y=y)=P(X=x)\,P(Y=y)$$

Cela peut se reformuler en $P_{(X,Y)}=P_X\times P_Y$.

Exercice 25 :

1. On lance 100 fois une pièce. Soit X la variable aléatoire égale au nombre de Face obtenu et Y la variable aléatoire égale au nombre de Pile obtenu. X et Y sont-elles indépendantes ?
2. On lance indéfiniment une pièce. Soit X la variable aléatoire égale au rang du premier Pile obtenu, Y la variable aléatoire égale au rang du premier Face obtenu et Z la variable aléatoire égale au rang du

deuxième Pile obtenu. X et Y sont-elles indépendantes ? X et Z sont-elles indépendantes ?

Exercice 26 : Soient X et Y deux variables aléatoires indépendantes et de même loi géométrique de paramètre p. Calculer $P(X=Y)$, $P(X \geqslant Y)$, $P(X \geqslant 2Y)$.

Propriété 21 – transfert d'indépendance

Si X et Y sont deux variables aléatoires discrètes indépendantes, alors pour toute fonction f définie sur $X(\Omega)$ et pour toute fonction g définie sur $Y(\Omega)$, les variables aléatoires $f(X)$ et $f(Y)$ sont indépendantes.

Par exemple, l'indépendance de X et Y entraîne celle de X^2 et Y^2.

8.2 familles finies de variables aléatoires discrètes

On étend sans difficulté les définitions et résultats précédents aux n-uplets de variables aléatoires :

- Si pour tout $i \in [\![1,n]\!]$, X_i est une variable discrète sur $(\Omega, \mathcal{A}, P)$ à valeurs dans un ensemble E_i, alors $X=(X_1,\dots,X_n)$ est une variable discrète à valeurs dans $E_1 \times \cdots \times E_n$.
- La loi de X est appelée loi conjointe et les lois des variables X_i sont appelées lois marginales.
- Chaque loi marginale se retrouve à partir de la loi conjointe.
  En posant $F_i = E_1 \times \cdots \times E_{i-1} \times E_{i+1} \times \cdots \times E_n$,

$$\forall i \in [\![1,n]\!],\ \forall t \in X_i(\Omega), \quad P(X_i=t) = \sum_{(x_1,\dots,x_{i-1},x_{i+1},\dots,x_n)\in F_i} P(X_1=x_1,\dots,X_{i-1}=x_{i-1},X_i=t,X_{i+1}=x_{i+1},\dots,X_n=x_n)$$

Définition - propriété 10

Soit $(X_i)_{1\leqslant i\leqslant n}$ une famille finie de variables aléatoires discrètes définies sur $(\Omega, \mathcal{A}, P)$. Les variables sont dites (mutuellement) indépendantes si on a l'une des deux assertions équivalentes suivantes :

1. Pour tout $(x_1,\dots,x_n) \in X_1(\Omega)\times\cdots\times X_n(\Omega)$, $P(X_1=x_1,\dots,X_n=x_n) = \prod_{i=1}^{n} P(X_i=x_i)$.
2. Pour tout $(A_1,\dots,A_n) \in X_1(\Omega)\times\cdots\times X_n(\Omega)$, les événements $(X_1\in A_1),\dots,(X_n\in A_n)$ sont mutuellement indépendants.

L'indépendance mutuelle implique l'indépendance deux à deux, mais la réciproque est fausse.

Propriété 22 – lemme des coalitions

Si $(X_1,X_2,\dots,X_n)$ sont indépendantes, toute variable aléatoire fonction de $X_1,X_2,\dots,X_p$ est indépendante de toute variable aléatoire fonction de $X_{p+1},X_{p+2},\dots,X_n$.

On peut généraliser à plusieurs coalitions.
Pour $(X_1,X_2,\dots,X_n)$ indépendantes, $p\in[\![1,n]\!]$, $1\leqslant i_1<i_2<\cdots<i_p\leqslant n$, et $f_1,f_2,\dots,f_p$ fonctions, les variables aléatoires :

$$f_1(X_1,X_2,\dots,X_{i_1}),\ f_2(X_{i_1+1},\dots,X_{i_2}),\dots,f_p(X_{i_{p-1}+1},\dots,X_{i_p})$$

sont indépendantes.
Exemple : si (X_1,X_2,X_3,X_4) sont indépendantes, alors :

— $X_1+X_2+X_3$ et X_4 sont indépendantes

— (X_1+X_2, X_3, e^{X_4}) sont indépendantes

— le lemme des coalitions ne peut pas s'appliquer pour juger de l'indépendance de $(X_1 + X_2, X_3, X_4 - X_2)$.

Autre exemple : on considère des variables aléatoires $X_1, X_2, \dots, X_n$ indépendantes suivant la loi uniforme sur $[\![0, 9]\!]$. On pose, pour tout $k \in [\![1, n]\!]$:

$$Y_k = \begin{cases} 1 & \text{si } X_k = 1 \\ 0 & \text{si } X_k \in [\![2, 9]\!] \end{cases}$$

Par le lemme des coalitions (ou bien, là, tout simplement un transfert d'indépendance), $Y_1, Y_2, \dots, Y_n$ sont indépendantes.

8.3 suites de variables aléatoires discrètes

On considère une suite $(X_n)_{n\in\mathbb{N}}$ de variables aléatoires sur un espace probabilisé $(\Omega, \mathcal{A}, P)$.

- Les variables X_n sont *indépendantes* si toute sous-famille $(X_i)_{i\in I}$ avec I partie finie de $\mathbb{N}$ est constituée de variables aléatoires indépendantes.
- si les variables X_n suivent de plus toutes la même loi, on dira que la suite est une suite de variables indépendantes et identiquement distribuées ou, en abrégé, une suite de variables i.i.d.

Théorème 2 – théorème d'extension de Kolmogorov - admis

Soit $(\mathcal{L}_n)_{n\in\mathbb{N}}$ une suite de lois discrètes sur des ensembles E_n. Il existe un espace probabilisé $(\Omega, \mathcal{A}, P)$ et une suite $(X_n)_{n\in\mathbb{N}}$ de variables aléatoires indépendantes telles que pour tout n, $X_n \sim \mathcal{L}_n$.

Pour nous, ce théorème sert uniquement à légitimer certains énoncés. Par exemple, il permet de donner un cadre théorique au jeu de pile ou face infini. En effet, on peut affirmer l'existence d'un espace probabilisé pour lequel il existe une suite de variables aléatoires indépendantes suivant une loi de Bernoulli de paramètre p donné ($1/2$ si la pièce lancée est équilibrée).
Plus généralement, ce théorème valide, dans un énoncé, toute formulation de la forme : « On considère une suite $(X_n)_{n\in\mathbb{N}}$ de variables aléatoires indépendantes et identiquement distribuées, de loi ... ».

9 Annexe : quelques éléments de démonstrations

Définition-propriété 1 – admise

Soit E un ensemble fini de cardinal n. Procédons par récurrence sur p.

— Pour $p = 1$, une 1-liste d'éléments (nécessairement d'éléments distincts!) de E est un singleton de E : il y a bien n possiblités.

— Soit p un entier naturel tel qu'il existe n^p p-listes d'éléments de E. Pour toute p-liste d'éléments de E, on peut construire n $(p+1)$-liste d'éléments de E, en choisissant de n façons différentes le dernier élément de la liste. Par l'hypothèse de récurrence, on a en tout $n \times n^p$ $(p+1)$-listes d'éléments de E, ce qui achève la récurrence.

— Soit p un entier naturel tel qu'il existe $n(n-1)...(n-p+1)$ p-listes d'éléments distincts de E. À partir de toute p-liste d'éléments distincts de E, on peut construire $n-p$ $(p+1)$-liste d'éléments distincts de E, en choisissant de $n-p$ façons différentes le dernier élément de la liste, qui doit être différent des p premiers éléments déjà choisis. Par l'hypothèse de récurrence, on a en tout $n \times n(n-1)...(n-p+1)(n-p)$ $(p+1)$-listes d'éléments distincts de E, ce qui achève la récurrence.

Propriété 3

(i) et (ii) résultent d'un calcul facile.
(iii) Par le calcul :

$$\begin{aligned}\binom{n-1}{p} + \binom{n-1}{p-1} &= \frac{(n-1)!}{p!(n-p-1)!} + \frac{(n-1)!}{(p-1)!(n-p)!} \\ &= \frac{(n-1)!}{p!(n-p)!}(n-p+p) \text{ (bien choisir le dénominateur commun !)} \\ &= \frac{n!}{p!(n-p)!} = \binom{n}{p}\end{aligned}$$

et de manière combinatoire : intéressons-nous aux $(k+1)$-combinaisons de $[\![1, n+1]\!]$. Il en existe $\binom{n+1}{k+1}$ mais nous pouvons les dénombrer autrement en distinguant celles qui contiennent $n+1$ de celles qui ne le contiennent pas.

— Il y a autant de k-combinaisons de $[\![1, n+1]\!]$ qui contiennent $n+1$ que de $(k-1)$-combinaisons de $[\![1, n]\!]$ à savoir $\binom{n}{k-1}$.

— Les k-combinaisons de $[\![1, n+1]\!]$ qui ne contiennent pas $n+1$ sont exactement les k-combinaisons de $[\![1, n]\!]$ et il y en a $\binom{n}{k}$.

(iv) La formule du capitaine peut se montrer par le calcul. Donnons une démonstration combinatoire, qui donne son nom à la formule. De combien de manières peut-on former une équipe de k entiers de $[\![1, n]\!]$ comportant un entier nommé capitaine? Nous allons dénombrer ces équipes de deux manières et le résultat en découlera.

— On peut commencer par choisir les k membres de l'équipe ($\binom{n}{k}$ possibilités), puis désigner le capitaine après coup parmi eux (k possibilités). On crée ainsi $k\binom{n}{k}$ équipes.

— On peut procéder autrement et choisir d'abord le capitaine (n possibilités), puis compléter l'équipe en choisissant les $k-1$ autres membres ($\binom{n-1}{k-1}$ possibilités). On crée cette fois $n\binom{n-1}{k-1}$ équipes.

Propriété 6

1. $\Omega = A \cup \overline{A}$ et l'union est disjointe. Donc $1 = P(A) + P(\overline{A})$.
2. Si $A \subset B$, B est l'union disjointe de A et $B \cap \overline{A}$ (faire un schéma) donc $P(B) = P(A) + P(B \setminus A)$.
   $P(B \setminus A) = P(B) - P(A)$ et ainsi $P(A) \leqslant P(B)$.
3. $A \cup B$ est l'union disjointe de A et $B \cap \overline{A}$, donc $P(A \cup B) = P(A) + P(B \cap \overline{A})$.
   $P(A \cup B) = P(A) + P(B) - P(A \cap B)$

Théorème 1

(i) Soit (A_n) une suite croissante d'événements. On pose *(faire un dessin)* :

$$B_0 = A_0 \text{ et } B_n = A_n \cap \overline{A_{n-1}} \text{ pour } n \geqslant 1$$

• Les B_n sont 2 à 2 incompatibles :
pour $j > i \geqslant 0$, $A_i \subset A_{j-1}$ donc $\overline{A_{j-1}} \cap A_i = \emptyset$. Donc pour $i \geqslant 1$, $B_i \cap B_j = A_i \cap \overline{A_{i-1}} \cap A_j \cap \overline{A_{j-1}} = \emptyset$ et $B_0 \cap B_j = A_0 \cap A_j \cap \overline{A_{j-1}} = \emptyset$.

• On a vu plus haut que $P(B_n) = P(A_n) - P(A_{n-1})$, d'où $\sum_{n=0}^{N} P(B_n) = P(A_0) + \sum_{n=1}^{N} P(A_n) - P(A_{n-1}) = P(A_N)$.

Par ailleurs $P(\bigcup_{n=0}^{\infty} B_n) = \sum_{n=0}^{\infty} P(B_n) = \lim_{N \to +\infty} \sum_{n=0}^{N} P(B_n)$.

La conclusion vient du fait que $\bigcup\limits_{n=0}^{\infty} B_n = \bigcup\limits_{n=0}^{\infty} A_n$. En effet, $\bigcup\limits_{n=0}^{N} B_n = \bigcup\limits_{n=0}^{N} A_n$, $\forall N$.

Soit maintenant (A_n) une suite décroissante d'événements. La suite $(\overline{A_n})$ est croissante, et par ce qui précède $P(\bigcup\limits_{n=0}^{\infty} \overline{A_n}) = \lim P(\overline{A_N})$ Par les lois de Morgan, $\bigcup\limits_{n=0}^{\infty} \overline{A_n} = \overline{\bigcap\limits_{n=0}^{\infty} A_n}$ et on obtient

$$1 - P(A_N) \to 1 - P(\bigcap_{n=0}^{\infty} A_n)$$

Corollaire 1

(i) posons $B_n = \bigcup\limits_{k=0}^{n} A_k$. La suite (B_n) est croissante, montrons que $\bigcup\limits_{n=0}^{\infty} A_n = \bigcup\limits_{n=0}^{\infty} B_n$, ainsi nous aurons par la propriété précédente

$$P(\bigcup_{n=0}^{\infty} A_n) = P(\bigcup_{n=0}^{\infty} B_n) = \lim_{n\to+\infty} P(B_n) = \lim_{n\to+\infty} P(\bigcup_{k=0}^{n} A_k)$$

Comme $A_k \subset B_k$, $\forall k$, on a $\bigcup\limits_{n=0}^{\infty} A_n \subset \bigcup\limits_{n=0}^{\infty} B_n$. Comme $\forall n$, $B_n \subset \bigcup\limits_{k=0}^{\infty} A_k$, on a $\bigcup\limits_{n=0}^{\infty} B_n \subset \bigcup\limits_{n=0}^{\infty} A_n$.

(ii) Notons $C_n = \overline{A_n}$.

$P(\bigcap\limits_{n=0}^{\infty} A_n) = P(\overline{\bigcup\limits_{n=0}^{\infty} C_n}) = 1 - P(\bigcup\limits_{n=0}^{\infty} C_n) = 1 - \lim\limits_{n\to+\infty} P(\bigcup\limits_{k=0}^{n} C_k) = \lim\limits_{n\to+\infty} 1 - P(\bigcup\limits_{k=0}^{n} C_k) = \lim\limits_{n\to+\infty} P(\overline{\bigcup\limits_{k=0}^{n} C_k}) = \lim\limits_{n\to+\infty} P(\bigcap\limits_{k=0}^{n} A_k)$.

Propriété ?? de sous-additivité

- Cas $I = [\![0, n]\!]$. Par récurrence.
- Cas $I = \mathbb{N}$. Par le premier cas, pour tout $n \in \mathbb{N}$,

$$P(\bigcup_{k=0}^{n} A_k) \leqslant \sum_{k=0}^{n} P(A_k)$$

et par le corollaire du théorème de continuité monotone, quand on fait tendre n vers $+\infty$, on a le résultat.

- Dans le cas général, I au plus dénombrable est en bijection avec $[\![0, n]\!]$ ou $\mathbb{N}$. Il existe φ bijection d'un de ces deux ensembles, mettons $\mathbb{N}$, dans I. Introduire φ, c'est effectuer un changement d'indices.

$$P(\bigcup_{i\in I} A_i) = P(\bigcup_{k\in\mathbb{N}} A_{\varphi(k)}) \leqslant \sum_{k\in\mathbb{N}} P(A_{\varphi(k)}) = \sum_{i\in I} P(A_i)$$

Propriété 8

Je peux prendre $I = \mathbb{N}$ pour simplifier.
Posons $C = \bigcup\limits_{i\in I} A_i$. Par hypothèse, $P(C) = 1$.

$$\begin{aligned} \sum_{i\in I} P(A_i \cap B) &= P(\bigcup_{i\in I} A_i \cap B) \text{ par } \sigma-\text{additivité} \\ &= P(C \cap B) = P(B) - P(\overline{C} \cap B) \end{aligned}$$

Mais par croissance, $P(\overline{C} \cap B) \leqslant P(\overline{C}) = 0$.

Propriété 9

Soit $(A_i)_{i\in I}$, avec I au plus dénombrable, des ensembles négligeables. Par sous-additivité de P, on a

$$P(\bigcup_{i\in I} A_i) \leqslant \sum_{i\in I} P(A_i) = 0$$

et donc $\bigcup\limits_{i\in I} A_i$ est négligeable.

Soit $(B_i)_{i\in I}$, avec I au plus dénombrable, des ensembles presque sûrs. Les $\overline{B_i}$ sont négligeables. Par ce qui précède,

$$P(\bigcap_{i\in I} B_i) = 1 - P(\bigcup_{i\in I} \overline{B_i}) = 1 - 0 = 1$$

Propriété 10

Soit $B \in \mathcal{A}$.
On a $0 \leqslant P_A(B)$, $P_A(\Omega) = 1$ et comme $A \cap B \subset A$, on a $P(A \cap B) \leqslant P(A)$, puis $P_A(B) \leqslant 1$.
Soit (A_n) une suite d'événements deux à deux incompatibles. On a :

$$P_A(\bigcup_{n\in\mathbb{N}} A_n) = \frac{P(A \cap (\bigcup\limits_{n\in\mathbb{N}} A_n))}{P(A)} = \frac{P(\bigcup\limits_{n\in\mathbb{N}} (A_n \cap A))}{P(A)} \overset{*}{=} \frac{\sum\limits_{n\in\mathbb{N}} P(A \cap A_n)}{P(A)} = \sum_{n\in\mathbb{N}} P_A(A_n)$$

$(*)$ vient de : P est une probabilité et $(A \cap A_n)_{n\in\mathbb{N}}$ est une suite d'événements incompatibles.

Propriété 14

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

Propriété 15

Soit Ω au plus dénombrable et $(p_\omega)_{\omega\in\Omega}$ une distribution de probabilités discrète.
Unicité
Si P est une probabilité sur $\mathcal{P}(\Omega)$ vérifiant : $P(\{\omega\}) = p_\omega$ pour tout ω, déterminons P.
Soit $A \in \mathcal{P}(\Omega)$. Comme sous-ensemble d'un ensemble au plus dénombrable, A est au plus dénombrable. La σ-additivité de P s'applique :

$$P(A) = P(\bigcup_{\omega\in A}\{\omega\}) = \sum_{\omega\in A} P(\{\omega\}) = \sum_{\omega\in A} p_\omega$$

et P est unique.

Existence
Introduisons $P : \left(\begin{array}{ccc} \mathcal{P}(\Omega) & \to & \mathbb{R} \\ A & \mapsto & \sum\limits_{\omega\in A} p_\omega \end{array}\right)$ et montrons que P est une application probabilité (on a bien $P(\{\omega_0\}) = p_{\omega_0}$).

- Comme les p_ω sont positifs, $P(1) \geqslant 0$ et $P(A) \leqslant \sum\limits_{\omega\in\Omega} p_\omega = 1$. P est à valeurs dans $[0, 1]$.
- On a bien $P(\Omega) = \sum\limits_{\omega\in\Omega} p_\omega = 1$.
- Soit $(A_n)_{n\in\mathbb{N}}$ une suite d'événements deux à deux incompatibles. On pose $A = \bigcup\limits_{n\in\mathbb{N}} A_n$. Par le théorème de sommation par paquets,

$$\sum_{\omega\in A} P(\{\omega\}) = \sum_{n=0}^{+\infty}\sum_{\omega\in A_n} P(\{\omega\})$$

et donc $P(A) = \sum\limits_{n=0}^{+\infty} P(A_n)$. P est σ-additive.

Propriété 16

Remarque : l'écriture $X^{-1}(A) = \bigcup\limits_{\omega\in\Omega/\, X(\omega)\in A} (X = x)$ bloque puisque Ω n'est pas supposé dénombrable.
On remarque plutôt que :

$$X^{-1}(A) = \{\omega \in \Omega /\, X(\omega) \in A \cap X(\Omega)\} = \bigcup_{x\in A\cap X(\Omega)} (X = x)$$

Comme $X(\Omega)$ est au plus dénombrable, il s'agit là d'une union au plus dénombrable d'événements, donc c'est un événement.

Propriété 18

Soit $(p_x)_{x\in E}$ des réels positifs de somme 1.
On considère F le support de cette famille : $F = \{x \in E \mid p_x > 0\}$. On a vu que F était dénombrable. La famille $(p_x)_{x\in F}$ est sommable et $\sum\limits_{x\in F} p_x = \sum\limits_{x\in E} p_x = 1$. Donc $(p_x)_{x\in F}$ est une distribution de probabilités sur F.
On prend $\Omega = F$, $\mathcal{A} = \mathcal{P}(F)$ (car F est au plus dénombrable) et P la probabilité associée à la distribution de probabilités discrète $(p_x)_{x\in F}$, à savoir :

$$\forall A \in \mathcal{P}(F), \quad P(A) = \sum_{x\in A} p_x$$

On prend enfin $X = \mathrm{Id}_F$.

On a $X(\Omega) = F$ au plus dénombrable.
Pour $x \in F$, $X^{-1}(\{x\}) = \{x\} \in \mathcal{A}$.
Donc X est une variable aléatoire discrète telle que $X(\Omega) \subset E$. Enfin pour $x \in E$,

$$P(X = x) = \begin{cases} P(\{x\}) = p_x & \text{si } x \in F \\ P(\emptyset) = 0 = p_x & \text{si } x \notin F \end{cases}$$

Définition-propriété 4

$(f \circ X)(\Omega) = f(X(\Omega))$ est l'image par f d'un ensemble au plus dénombrable, donc est au plus dénombrable.
Pour $y \in f \circ X(\Omega)$, on a

$$(f(X) = y) = \{\omega \in \Omega \mid f(X(\omega)) = y\} = \{\omega \in \Omega \mid X(\omega) \in f^{-1}(\{y\})\} = (X \in f^{-1}(\{y\})) \in \mathcal{A}$$

Donc $f(X)$ est une variable aléatoire.
Soient maintenant X et Y définies sur les espaces probabilisés $(\Omega, \mathcal{A}, P)$ et $(\Omega', \mathcal{A}', P')$ respectivement. On suppose que $X \sim Y$,

donc $X(\Omega) = Y(\Omega)$ et pour x dans cet ensemble, $P(X = x) = P'(Y = x)$.
On a $f(X(\Omega)) = f(Y(\Omega))$ et pour y dans cet ensemble,

$$P(f(X) = y) = \sum_{x \in f^{-1}(\{y\})} P(X = x) = \sum_{x \in f^{-1}(\{y\})} P'(X = y) = P'(f(Y) = x)$$

Propriété 19

Considérons les événements S_k « on obtient un succès au k^{e} lancer » et E_k « on obtient un échec au k^{e} lancer ».
On a $(X = n) = E_1 \cap E_2 \cap ... \cap E_{n-1} \cap S_n$ et

$$\begin{array}{rcl} P(X = n) & = & P(E_1 \cap ... \cap E_{n-1} \cap S_n) \\ & = & P(E_1) \times ... \times P(E_{n-1}) \times P(S_n) \textit{ car ces événements sont indépendants} \\ & = & (1-p)^{n-1} p \end{array}$$

Dans le modèle proposé, X peut a priori prendre la valeur $+\infty$. Montrons que $P(X = +\infty) = 0$.

$$\begin{array}{rcl} (X = +\infty) & = & \bigcap_{k=1}^{+\infty} E_k \\ P(\bigcap_{k=1}^{n} E_k) & = & q^n \\ P(X = +\infty) & = & \lim q^n = 0 \text{ par continuité décroissante} \end{array}$$

Remarque : On a $P(X \in \mathbb{N}^*) = P(\bigcup_{k=1}^{\infty} (X = k)) = \sum_{k=1}^{\infty} p(1-p)^{k-1} = \frac{p}{(1-(1-p))} = 1$. X est donc presque-sûrement à valeurs dans $\mathbb{N}^*$.