# Variables discrètes

25 mars 2020

# Ce que l'on cherche à faire

La théorie des probabilités de première année se concentre sur les v.a. pouvant prendre un nombre *fini* de valeurs. De telles variables sont issues d'*une* expérience simple ou de la répétition d'un nombre fini de telles expériences.
Se limiter à telles variables est une restriction empêchant des poser de questions très naturelles sur les expériences faites.

**Proverbe (Grand Shadok/J. Rouxel)**

*Ce n'est qu'en essayant continuellement que l'on finit par réussir.... En d'autres termes... Plus ça rate et plus on a de chances que ça marche...*

Si je fais une expérience qui a une chance sur deux de « marcher » :

- Combien faut-il d'expériences, en moyenne, pour avoir le premier succès ?
- Combien faut-il d'expériences, en moyenne, pour avoir deux, trois,... succès ?

Pour répondre à ce type de questions et d'autres, nous devons nous autoriser la possibilité de faire mentalement un nombre *infini* d'expériences et considérer des variables pouvant prendre *toutes* les valeurs entières.
Par exemple, si $T$ est le nombre d'expériences à faire pour avoir le premier succès, $T$ peut prendre toutes valeurs entières, et même, en cas d'échec total, la valeur « infini », $\infty$. La moyenne cherchée est l'*espérance mathématique* de $T$, $\mathbb{E}(T)$.

Nous avons esquissé le cadre théorique permettant l'extension du calcul des probabilités à des variables pouvant prendre une infinité de valeurs, en l'espèce, les variables aléatoires réelles *à densité*.

Ce même cadre, va nous permettre de traiter le cas de variables aléatoires $X$ prenant leurs valeurs dans un ensemble *dénombrable*[1] $\mathscr{N}$. On se focalisera notamment sur les cas où

1. $\mathscr{N} = \mathbb{N}$, l'ensemble $\mathbb{N}$ des entiers naturels
2. $\mathscr{N} = \mathbb{Z}$, l'ensemble $\mathbb{Z}$ des entiers relatifs
3. $\mathscr{N} = \mathbb{N}^2$ ou $\mathscr{N} = \mathbb{Z}^2$ (Cas des couples de v.a discrètes entières)

Des exemples importants seront

1. $X$, compteur d'événements aléatoires,
2. $X$, l'instant où un événement survient dans une suite d'expériences
3. $X$, sommes, différences, couples de telles variables...

Une variable aléatoire $X$ à valeurs dans $\mathbb{N}$ ou $\mathbb{Z}$ est une variable aléatoire réelle. En particulier, les questions de l'existence de l'espérance, de la valeur de son espérance, de sa variance sont des questions importantes.
Plus généralement, si $\mathscr{N}$ est dénombrable, $X$ une v.a. à valeurs dans $\mathscr{N}$ et $f : \mathscr{N} \to \mathbb{R}$ est une fonction alors $f(X)$ est une v.a réelle prenant ses valeurs dans l'ensemble dénombrable $f(\mathscr{N})$.

# Le cadre abstrait

### Définition

*Un espace probabilisé* $(\Omega, \mathscr{A}, \mathbb{P})$ *est composé de*

1. *Un ensemble* $\Omega$,
2. *une famille* $\mathscr{A}$ *de parties de* $\Omega$ *appelée la tribu des événements.*
3. *Une application* $\mathbb{P} : \mathscr{A} \to [0,1]$, *qui, à chaque événement,* $A \in \mathscr{A}$ *associe la probabilité de* $A$.

La famille $\mathscr{A}$ et l'application $\mathbb{P}$ *doivent* satisfaire des *axiomes* que lecteur est prié d'aller retrouver dans le chapitre sur les généralités de probabilités.

⋆⋆⋆

On fixe pour tout le chapitre un espace probabilisé quelconque $(\Omega, \mathscr{A}, \mathbb{P})$. Par *événement*, on entend un élément de la tribu des événements $\mathscr{A}$. Ce sont les seules parties de $\Omega$ dont on peut calculer la probabilité.

⋆⋆⋆

Rappelons la formule des probabilités totales, *i.e.* l'axiome d'additivité dénombrable d'une probabilité $\mathbb{P}$ :
*Si $(A_n)_{n\in\mathbb{N}}$ est une suite d'événements incompatibles, ( i.e. deux à deux disjoints) alors*

$$\mathbb{P}(\cup_{n\in\mathbb{N}}A_n) = \sum_{n=0}^{+\infty}\mathbb{P}(A_n)$$

La somme est bien évidemment la somme de la série de terme général $\mathbb{P}(A_n)$.
Du fait de l'invariance d'une union infinie sous-certaines opérations (renumérotation des événements, groupements par paquets), cette règle ne peut avoir de sens si l'on n'a pas à l'esprit les résultats suivants de théories des séries.

# Trois lemmes de théorie des séries

NB : Ces trois lemmes ne sont pas au programme officiel. Ils structurent cependant toute la théorie.

**Lemme**

*Soit $(u_n)_{n\in\mathbb{N}}$ une famille de nombres positifs. Que la série $\sum_n u_n$ soit convergente ou pas, on a*

$$\sum_{n=0}^{+\infty} u_n = \sup_{I\subset\mathbb{N},\, I \text{ fini}} \left( \sum_{i\in I} u_i \right)$$

# Trois lemmes de théorie des séries

## Lemme (Réordonnancement)

*Soit $(u_n)_{n\in\mathbb{N}}$ une suite de nombres réels. $\sigma : \mathbb{N} \to \mathbb{N}$ une bijection.*

1. *(Cas positif) Si la série $\sum_n u_n$ est à termes positifs OU*
2. *(Cas ACV) si la série $\sum_n |u_n|$ est convergente, on a*

$$\sum_{n=0}^{+\infty} u_n = \sum_{k=0}^{+\infty} u_{\sigma(k)}$$

*avec, en guise de précisions :*

1. *(Cas positif) les valeurs étant soient finies, soit $\infty$ ;*
2. *(Cas ACV) les deux séries étant absolument convergentes.*

# Trois lemmes de théorie des séries

La valeur de la somme d'une série absolument convergente ne dépend pas de l'ordre des termes.
En conséquence, si $(u_n)_{n\in\mathscr{N}}$ est une suite numérique indicée par $\mathscr{N}$ dénombrable, si $\sigma : \mathbb{N} \to \mathscr{N}$ est une bijection, et $\sum_{k\in\mathbb{N}} u_{\sigma(k)}$ une série à termes positifs ou ACV, on note

$$\sum_{n\in\mathscr{N}} u_n = \sum_{k=0}^{+\infty} u_{\sigma(k)}$$

Dans ce cas, le symbole $\sum_{n\in\mathscr{N}} u_n$ est possiblement le symbole $\infty$ (cas positif, divergent) ou un nombre. Il ne dépend pas du choix de la bijection $\sigma$. Il s'agit d'une extension au cas infini dénombrable du symbole $\sum_{n\in\mathscr{N}} u_n$ bien connu lorsque $\mathscr{N}$ est un ensemble fini. On rappelle que si $\mathscr{N}$ est vide, la convention est $\sum_{n\in\mathscr{N}} u_n = 0$

# Trois lemmes de théorie des séries

## Lemme (Paquets/super-Chasles)

*Si* $(\mathscr{N}_k)_{k\in\mathbb{N}}$ *est une partition*[a] *de* $\mathscr{N}$ *et*

1. *(Cas positif) Si* $\sum_{n\in\mathscr{N}} u_n$ *est à termes positifs OU*
2. *(Cas ACV) si* $\sum_{n\in\mathscr{N}} |u_n|$ *est finie* $(\neq \infty)$,

*alors*

$$\sum_{n\in\mathscr{N}} u_n = \sum_{k=0}^{+\infty} \left( \sum_{n\in\mathscr{N}_k} u_n \right).$$

---

a. Les paquets $\mathscr{N}_k$ peuvent être formés d'un nombre fini ou infini dénombrable d'indices. Formellement

$$\mathscr{N} = \cup_{k\in\mathbb{N}} \mathscr{N}_k \text{ et } \forall k,\ell \in \mathbb{N}, k \neq \ell \Rightarrow \mathscr{N}_k \cap \mathscr{N}_\ell = \emptyset$$

# Trois lemmes de théorie des séries

Remarques :

1. (Cas positif) Que les sommes de séries soient finies soit $\infty$. Si l'une des sommes $\sum_{n\in\mathcal{N}_k} u_n = \infty$, la somme de droite vaut $\infty$.
2. (Cas ACV) Toutes les séries de la formule sont absolument convergentes

Les lemmes 4 et 5 donnent sens à l'axiome d'additivité dénombrable d'une probabilité $\mathbb{P}$ : Si $(A_n)_{n\in\mathbb{N}}$ est une famille disjointe d'événements et $\sigma : \mathbb{N} \to \mathbb{N}$ est une bijection alors $(A_{\sigma(k)})_{k\in\mathbb{N}}$ est aussi une famille disjointe d'événements et on a $\cup_{n\in\mathbb{N}} A_n = \cup_{k\in\mathbb{N}} A_{\sigma(k)}$. La probabilité de cette union ne dépend pas de l'ordre des termes, ce qui est cohérent avec le lemme 4 qui affirme que,

$$\sum_{n=0}^{+\infty} \mathbb{P}(A_n) = \sum_{k=0}^{+\infty} \mathbb{P}(A_{\sigma(k)})$$

Les lemmes 4 et 5 donnent sens à l'axiome d'additivité dénombrable d'une probabilité $\mathbb{P}$ : Si $(A_n)_{n\in\mathbb{N}}$ est une famille disjointe d'événements et si $(\mathscr{N}_k)_{k\in\mathbb{N}}$ est une partition de $\mathbb{N}$, alors

- chaque $B_k = \cup_{n\in\mathscr{N}_k} A_n$ est un événement avec
$$\mathbb{P}(B_k) = \sum_{n\in\mathscr{N}_k} \mathbb{P}(A_n)$$
- La famille $(B_k)_{k\in\mathbb{N}}$ est une famille disjointe d'événements et on doit avoir
$$\mathbb{P}(\cup_{k\in\mathbb{N}} B_k) = \sum_{k\in\mathbb{N}} \mathbb{P}(B_k) = \sum_{k=0}^{+\infty} \sum_{n\in\mathscr{N}_k} \mathbb{P}(A_n)$$
- On a $\cup_{k\in\mathbb{N}} B_k = \cup_{k\in\mathbb{N}} \cup_{n\in\mathscr{N}_k} A_n = \cup_{n\in\mathbb{N}} A_n$, ce qui est cohérent avec le lemme 5 qui affirme que,
$$\sum_{n=0}^{+\infty} \mathbb{P}(A_n) = \sum_{k=0}^{+\infty} \sum_{n\in\mathscr{N}_k} \mathbb{P}(A_n) = \sum_{k=0}^{+\infty} \mathbb{P}(B_k)$$

### Définition

*Soit* $(\Omega, \mathscr{T}, \mathbb{P})$ *un espace probabilisé. Si* $\mathscr{N}$ *est un ensemble au plus dénombrable, une variable aléatoire (sur* $\Omega$*) à valeurs dans* $\mathscr{N}$ *est une application* $X : \Omega \to \mathscr{N}$ *telle que, pour tout* $n \in \mathscr{N}$*,*

$$\{X = n\} = \{\omega \in \Omega, X(\omega) = n\}$$

*est un événement.*

- Au plus dénombrable signifie *fini* ou en bijection avec $\mathbb{N}$. Par exemple, on peut prendre $\mathscr{N} = \mathbb{N}$, $\mathscr{N} = \mathbb{Z}$, $\mathscr{N} = \mathbb{N}^2$, etc..
- $\mathbb{R}$ est « trop gros » et n'est pas au plus dénombrable. (Cantor)
- Si $X$ est une v.a à valeurs dans $\mathscr{N}$, $\mathscr{N}' \subset \mathscr{N}$ alors $\{X \in \mathscr{N}'\}$ est un événement

$$\{X \in \mathscr{N}'\} = \{\omega \in \Omega, X(\omega) \in \mathscr{N}'\} = \cup_{n \in \mathscr{N}'} \{X = n\}$$

Par exemple, si $X$ est à valeurs entières, $\{X \text{ est pair}\}$ est un événement. Il s'agit de l'évenement $\{X \in 2\mathbb{N}\}$.

- Si $f : \mathscr{N} \to \mathbb{R}$ et $X$ est une v.a à valeurs $\mathscr{N}$ alors $f(X)$ est une v.a à valeurs réelles à valeurs dans $f(\mathscr{N}) \subset \mathbb{R}$ qui est au plus dénombrable.
- Une variable aléatoire $X$ à valeurs dans un ensemble $\mathscr{R}$ est dite *discrète* si l'ensemble de ses valeurs potentielles est au plus dénombrable. Dit de façon plus précise, s'il existe $\mathscr{N}$, une partie dénombrable de $\mathscr{R}$ telle que $\{X \in \mathscr{N}\}$ est un événement et

  $$\mathbb{P}(X \in \mathscr{N}) = 1$$

  c'est à dire, si l'évenement $X \in \mathscr{N}$ est *quasi-certain* (terminologie BCPST only) ou *presque-sûr* (terminologie universelle de New-York à Pekin)

# Décomposer un événement

## Définition

*Une famille d'événements* $(A_n)_{n\in\mathcal{N}}$ *est appelée un système complet d'événements incompatibles*[a] *si*

$$\cup_{n\in\mathcal{N}} A_n = \Omega \text{ et } \forall n, m \in \mathcal{N}, m \neq n \Rightarrow A_n \cap A_m = \emptyset$$

a. Un *s.c.e.i.* On peut alléger cette définition et demander seulement que

$$\mathbb{P}(\cup_{n\in\mathcal{N}} A_n) = 1 \text{ et } \forall n, m \in \mathcal{N}, m \neq n \Rightarrow \mathbb{P}(A_n \cap A_m) = 0$$

Si $X$ est une v.a. discrète à valeurs dans $\mathcal{N}$ dénombrable, alors $(A_n)_{n\in\mathcal{N}} = (\{X = n\})_{n\in\mathcal{N}}$ est un s.c.e.i dénombrable. Ecrire, en utilisant l'axiome d'additivité de $\mathbb{P}$ que

$$\mathbb{P}(B) = \sum_{n\in\mathbb{N}} \mathbb{P}(B \cap \{X = n\})$$

s'appelle *décomposer l'événement* $B$ *suivant les valeurs de* $X$ ou *calculer* $\mathbb{P}(B)$ *en conditionnant sur les valeurs de* $X$.
Cette formule s'écrit aussi à l'aide des probabilités conditionnelles (*formule des probabilités totales* )

$$\mathbb{P}(B) = \sum_{n\in\mathbb{N}} \mathbb{P}(B|X = n)\mathbb{P}(X = n)$$

Rq : si $\mathbb{P}(X = n) = 0$ alors $\mathbb{P}(B|X = n)$ n'est pas défini. On prend comme convention dans cette formule que le produit

$$\mathbb{P}(B|X = n)\mathbb{P}(X = n) = 0$$

# La formule des probabilités totales dénombrable.

Plus généralement, l'axiome d'additivité dénombrable implique que

### Théorème

*Si* $A_1, \ldots, A_\ell, \ldots$ *est un système complet d'événements incompatibles, dénombrable alors, pour tout événement* $B$*,*

$$\mathbb{P}(B) = \sum_{\ell=1}^{+\infty} \mathbb{P}(B \cap A_\ell) = \sum_{\ell=1}^{+\infty} \mathbb{P}(B|A_\ell)\mathbb{P}(A_\ell)$$

Rq : là encore, si $\mathbb{P}(A_\ell) = 0$ alors $\mathbb{P}(B|A_\ell)$ n'est pas défini. On prend comme convention dans cette formule que le produit

$$\mathbb{P}(B|A_\ell)\mathbb{P}(A_\ell) = 0.$$

On remarque aussi l'écriture alternative pour dénoter une probabilité conditionnelle $\mathbb{P}_{A_\ell}(B) = \mathbb{P}(B|A_\ell)$.

# Distribution/Loi d'une v.a à valeurs $\mathscr{N}$

## Définition

*Soit* $\mathscr{N}$ *un ensemble au plus dénombrable,* $X$ *une variable aléatoire à valeurs dans* $\mathscr{N}$.
*La distribution ou loi de* $X$ *est la donnée de la famille des nombres positifs* $(p_n)_{n\in\mathscr{N}}$ *où*

$$\forall n \in \mathscr{N},\, p_n = \mathbb{P}(X = n)$$

1. Cette définition étend celle connue pour la distribution d'une variable prenant un nombre fini de valeurs.
2. Pour tout $n \in \mathcal{N}$, $0 \leq p_n \leq 1$ et $\sum_{n \in \mathcal{N}} p_n = 1$.
3. Si $\mathcal{N}' \subset \mathcal{N}$,
$$\mathbb{P}(X \in \mathcal{N}') = \sum_{n \in \mathcal{N}'} p_n$$
4. Si $\mathcal{N} \subset \mathbb{R}$, il est équivalent de donner la famille $(\mathbb{P}(X = n))_{n \in \mathcal{N}}$ ou la famille, plus large $\{\mathbb{P}(X \in I), I \text{ intervalle de } \mathbb{R}\}$. On a, si $I$ est un intervalle de $\mathbb{R}$ quelconque, en décomposant suivant les valeurs de $X$,
$$\mathbb{P}(X \in I) = \sum_{n \in I \cap \mathcal{N}} p_n = \sum_{n \in \mathcal{N}} \mathbb{1}_{\{n \in I\}} p_n$$

# Espérance

On rappelle que la fonction $1\!\!1_A$, indicatrice de l'événément $A$ est une variable aléatoire à valeurs dans[2] $\{0,1\} \subset \mathbb{R}$. On a défini l'espérance de cette v.a par

$$\mathbb{E}(1\!\!1_A) = \mathbb{P}(A)$$

On rappelle aussi que l'espérance d'une v.a. réelle positive est, d'une manière générale, soit un nombre réel positif, soit la valeur $\infty$. Une v.a. réelle $X$ admet une espérance[3] si l'espérance de sa valeur absolue est finie.

# Espérance d'une v.a. réelle, ens. dénombrable de valeurs

## Proposition-Définition

*Soit $X$ une v.a prenant ses valeurs dans $\mathscr{N} = \{x_0, \ldots, x_k, \ldots\} \subset \mathbb{R}$. On pose $\forall k \in \mathbb{N}, p_k = \mathbb{P}(X = x_k)$.*

1. *On a $\mathbb{E}(|X|) = \sum_{k=0}^{+\infty} |x_k| p_k$, que cette quantité soit finie ou pas.*
2. *Si $\mathbb{E}(|X|) < +\infty$ alors la variable $X$ admet une espérance et on [a] a $\mathbb{E}(X) = \sum_{k=0}^{+\infty} x_k.p_k$*

a. Cette série est ACV.

Remarque : on peut étendre ces notions *verbatim* si $X$ est discrète, à valeurs complexes.

# Formule de transfert, ens. dénombrable de valeurs

### Théorème

*Soit $X$ une v.a prenant ses valeurs dans $\mathscr{N} = \{x_0, \dots, x_k, \dots\}$ dénombrable, $h : \mathscr{N} \to \mathbb{R}$ une fonction. On pose $\forall k \in \mathbb{N}$, $p_k = \mathbb{P}(X = x_k)$.*

1. *On a $\mathbb{E}(|h(X)|) = \sum_{k=0}^{+\infty} |h(x_k)| p_k$, que cette quantité soit finie ou pas.*
2. *Si $\mathbb{E}(|h(X)|) < +\infty$, la variable $Y = h(X)$ admet une espérance et on a*

$$\mathbb{E}(Y) = \mathbb{E}(h(X)) = \sum_{k=0}^{+\infty} h(x_k).p_k$$

# Espérance/Variance/Autres

**Exercice 1.—**

1. Donner la formule générale pour $\mathbb{E}(X^2)$ sachant que $X$ est à valeurs entières. Idem pour $\mathbb{E}(X^k)$ où $k \in \mathbb{N}$. Traiter les cas où $X$ ne prend que des valeurs entières positives, où $X$ est à valeurs dans $\mathbb{Z}$.
2. Donner la formule pour $\mathbb{E}(h(X))$ où $h(x) = x \ln x$ si $x > 0$, $h(0) = 0$. Donner une condition d'existence.

**Exercice 2.—**

1. Donner la formule pour $\mathbb{E}(e^{i\theta.X})$ où $\theta \in \mathbb{R}$. Existence ?
   1. On suppose que $X$, à valeurs dans $\mathbb{N}$, vérifie $\forall n \in \mathbb{N}, \mathbb{P}(X = n) = p.(1-p)^n$ pour un certain $p \in ]0,1[$. Simplifier la formule donnant $\mathbb{E}(e^{i\theta.X})$.
   2. On suppose que $X$, à valeurs dans $\mathbb{N}$, vérifie $\forall n \in \mathbb{N}, \mathbb{P}(X = n) = e^{-\lambda}.\frac{\lambda^n}{n!}$ pour un certain $\lambda \in ]0,+\infty[$. Simplifier la formule donnant $\mathbb{E}(e^{i\theta.X})$.

Nous allons revenir longuement sur ces deux exemples : lois géométriques et de Poisson.

# Suites d'expériences

On va s'intéresser dans cette partie à des *processus* aléatoires décrivant l'évolution, aléatoire, d'un système. Il s'agit de suites de variables aléatoires $(X_n)_{n\in\mathbb{N}}$.
La variable $X_n$ donne l'*état*, à l'instant $n$ du système. $X_n$ peut être à valeurs numériques mais aussi à valeurs plus composites : vecteur numérique, liste, caractère qualitatif, etc...
Ce qui est crucial, c'est de bien poser la nature des objets $X_n$ car c'est ce qui permet de clarifier les calculs.
Les variables à valeurs numériques jouent un rôle prépondérant. Pour elles, tenter de calculer une espérance, une variance a un sens.
On va traiter un certain nombre d'exemples typiques : à chaque fois on spécifie l'expérience élémentaire, la façon de répéter l'expérience et une ou plusieurs variables aléatoires à valeurs entières dont on détermine la distribution.

# Epreuve de Bernoulli

Supposons que l'on dispose d'une pièce biaisée où la probabilité de sortir $P$ est $p$, un nombre entre 0 et 1 et la probabilité de sortir $F$ est[4] $q := 1 - p$. Le résultat de l'expérience est complètement décrit par un élément de $\{P, F\}$ et donc une v.a décrivant cette expérience est une v.a. $X$ à valeurs dans cet ensemble satisfaisant

$$\mathbb{P}(X = P) = p \text{ et } \mathbb{P}(X = F) = 1 - p = q$$

# Epreuve de Bernoulli

Une interprétation de cette expérience fondamentale est la suivante :
Supposons une expérience (difficile) effectuée et $A$ un événement portant sur son résultat :
Soit $A$ (prob. $p$), soit $\overline{A}$ (prob. $1-p$) est réalisé. Si on pose $X = 1\!\!1_A$, $X$ est une variable aléatoire, à valeurs dans $\{0,1\}$ telle que

$$\mathbb{P}(X=1) = p \text{ et } \mathbb{P}(X=0) = 1-p = q$$

Une telle variable est appelée une variable de Bernoulli avec probabilité de succés $p$.
Rappelons que au niveau des espérances,

$$\mathbb{E}(1\!\!1_A) = \mathbb{P}(A) = p$$

# Jouer jusqu'a la fin des temps

Donnons nous une suite $(X_n)_{n\in\mathbb{N}^*}$ de variables de Bernoulli, de proba. de succès $p$, *indépendantes*. Posons, à l'habitude, $q = 1 - p$.
Cela signifie simplement que, pour toute famille finie d'indices distincts $n_1, \ldots, n_k$, toute famille $\varepsilon_1, \ldots, \varepsilon_k$ de valeurs 0 ou 1,

$$\mathbb{P}(X_{n_1} = \varepsilon_1, \ldots, X_{n_k} = \varepsilon_k) = \mathbb{P}(X_{n_1} = \varepsilon_1). \ldots . \mathbb{P}(X_{n_k} = \varepsilon_k)$$

Une telle suite modélise une suite de lancers indépendants d'une pièce biaisée. $X_n$ est le résultat au $n$-ième lancer. Elle modélise aussi la suite des résultats (échec/succès) lors d'une suite d'expériences identiques réputées indépendantes.

# Attendre le succès, compter les échecs

Supposons $0 < p \leq 1$. Appelons $T$ le *rang du premier succès* :

$$T = \min\{n \in \mathbb{N}^*, X_n = 1\}$$

Montrons d'abord que $T$ est bien définie presque sûrement. Par cela on entend le fait, que

$$\mathbb{P}(\exists n \in \mathbb{N}^*, X_n = 1) = 1$$

# Attendre le succès, compter les échecs

Posons $A = \{\exists n \in \mathbb{N}^*, X_n = 1\} = \bigcup_{n\in\mathbb{N}}\{X_n = 1\}$ et donc $\overline{A} = \{\forall n \in \mathbb{N}^*, X_n = 0\} = \cap_{n\in\mathbb{N}^*}\{X_n = 0\}$. Ces deux ensembles sont des événéments. Par ailleurs, pour $N$ fixé, destiné à tendre vers $+\infty$,

$$\begin{aligned} \overline{A} &\subset \cap_{n=1}^{N}\{X_n = 0\}, \\ 0 \leq \mathbb{P}(\overline{A}) &\leq \mathbb{P}(\cap_{n=1}^{N}\{X_n = 0\}) = \prod_{n=1}^{N}\mathbb{P}(X_n = 0) = q^N \overset{N\to+\infty}{\to} 0 \end{aligned}$$

On a donc $\mathbb{P}(\overline{A}) = 0$ et $\mathbb{P}(A) = 1$. On définit précisémment $T$ par

$$T = \begin{cases} \min\{n \in \mathbb{N}^*, X_n = 1\} & \text{sur } A \\ \infty & \text{sur } \overline{A} \end{cases}$$

$T$ est une variable prenant ses valeurs dans l'ensemble dénombrable $\mathbb{N}^* \cup \{\infty\}$. Elle est *presque sûrement* à valeurs entières. $\mathbb{P}(T = +\infty) = 0$. Calculons sa distribution. Pour $n \in \mathbb{N}^*$,

$$\begin{array}{rcl} \mathbb{P}(T = n) & = & \mathbb{P}(X_1 = 0, \ldots, X_{n-1} = 0, X_n = 1) \\ \text{(indépendance !)} & = & \mathbb{P}(X_1 = 0). \ldots .\mathbb{P}(X_{n-1} = 0).\mathbb{P}(X_n = 1) \\ & = & q^{n-1} p \end{array}$$

# La distribution géométrique de paramètre $p$

**Définition**

*Soit* $p \in ]0,1]$. *Une v.a.* $X$ *suit la loi géométrique de paramètre* $p$ *sur* $\mathbb{N}^*$, *noté* $X \sim \mathscr{G}(\mathbb{N}^*, p)$, *si*

$$\forall n \in \mathbb{N}^*, \mathbb{P}(X = n) = (1-p)^{n-1}.p = p.q^{n-1}$$

Si $T$ est le rang d'arrivée du premier succès dans un processus de Bernoulli de paramètre $0 < p \leq 1$ alors $T$ suit une loi géométrique de paramètre $p$ sur $\mathbb{N}^*$.
Le nombre $S$ d'échecs avant le premier succès suit, lui, une loi « géométrique », hors programme, de paramètre $p$ « sur $\mathbb{N}$ ». On a $T = S + 1$ et

$$\forall n \in \mathbb{N}, \mathbb{P}(S = n) = p.q^n$$

# Espérance et variance

**Proposition**

*Si $T$ suit la loi géométrique de paramètre $p \in ]0,1]$ sur $\mathbb{N}^*$ alors, $T$ est intégrable, de carré intégrable et*

$$\begin{aligned} \mathbb{E}(T) &= \frac{1}{p} \\ \mathbb{V}(T) &= \frac{1-p}{p^2} = \frac{q}{p^2} \end{aligned}$$

**Exercice 3.**—Donner l'espérance et la variance de $S = T - 1$ si $T \sim \mathscr{G}(\mathbb{N}^*, p)$.

# Simulation

Utiliser le script `simulation-geom.py`.

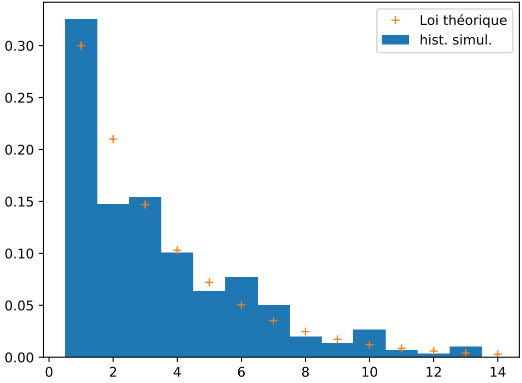



Figure – Une simulation et la loi théorique $NS = 300$ tirages, $p = 0.3$

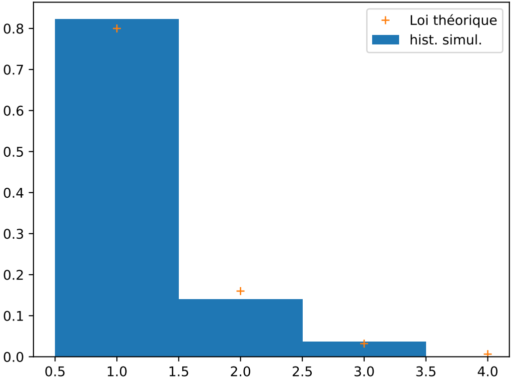



Figure – Une simulation et la loi théorique $NS = 300$ tirages, $p = 0.8$

# Compter les gains, attendre la ruine

**Exercice 4.**—Au cours d'une partie de pile ou face, on peut décider d'affecter un gain $+1$ en cas de succès, une perte $-1$ en cas d'échec. Posons $G_n = 2X_n - 1$. Il s'agit du gain à la partie $n$. $\mathbb{P}(G_n = 1) = p$, $\mathbb{P}(G_n = -1) = 1 - p = q$. Si on démarre avec un capital $k_0$, $C_0 = k_0$, le capital à l'issue de la partie $n$ sera

$$C_n = k_0 + \sum_{\ell=1}^{n} G_\ell$$

**1.** Quel est le lien entre $C_n$ et $S_n := \sum_{\ell=1}^{n} X_\ell$ ?
**2.** Quelle est la loi de $C_n$ ? Quelle est son espérance ? Quelle est sa variance ?

Une question intéressante est la suivante : supposons que l'on se donne un seuil de perte et de gain maximal $K \in \mathbb{N}^*$. On décide de quitter la partie au rang $n$ dès que $C_n = \pm K$. Quelle est la durée moyenne d'une telle partie ? Quelle est la probabilité de quitter la partie en gagnant ? en perdant ?
Pour la durée de jeu, le type de raisonnement qui menait à la loi géométrique ne peut s'appliquer : les variables $C_n$ ne sont plus indépendantes !
Ce question est un problème classique nommé « Le problème de la ruine du joueur. » Dans la partie suivante, qui ne fait pas partie des fondamentaux du cours, on développe les outils pour attaquer ce problème. Tous ces outils sont à notre programme et la partie suivante pourrait faire l'objet d'un joli problème d'écrit.

# Recompter les gains, un autre jeu

On utilise les probabilités conditionnelles pour traiter le problème de la ruine du joueur et de la durée d'une partie.
Le point de départ est la relation fondamentale :

$$\mathbb{P}(C_{n+1} = k) = \mathbb{P}(C_n = k-1)\mathbb{P}(X_{n+1} = 1) + \mathbb{P}(C_n = k+1)\mathbb{P}(X_{n+1} = 0)$$

Il s'agit de la formule des probabilités conditionnelles.(Détail)

On va jouer au jeu suivant : un damier rectiligne contient $2K+1$ cases numérotées $-K,\ldots,0,\ldots,K$. A l'instant 0, le pion est placé en $k_0$ et à chaque étape $n$, si on est sur une des cases centrales $-K+1,\ldots,0,\ldots,K-1$, suivant que $G_n = 2X_n - 1 = \pm 1$, on se déplace à droite ou à gauche sur le damier. Tant qu'on ne sort pas du damier, $C_n$ est le numéro de la case sur laquelle on se trouve. L'instant d'arrêt est le moment où l'on saute sur l'une des cases $\pm K$.
On peut décider qu'à partir de ce moment, le pion reste bloqué sur cette case. On note $(C_n^s)$ la suite des positions.

Ce jeu se symbolise de la façon suivante et on reconnait là, le schéma typique symbolisant une chaine de Markov : Dessin !
Si on note $U_n = (u_{n,k})_{k\in\{-K,\dots,K\}}$ le vecteur tel que $u_{n,k} = \mathbb{P}(C_n^s = k)$. La position de départ donne $U_0$ tel que $u_{0,k_0} = 1$, $u_{0,k} = 0$ si $k \neq k_0$ et la relation fondamentale donne une relation de récurrence du type

$$\forall n \in \mathbb{N}, U_{n+1} = M.U_n$$

et donc

$$\forall n \in \mathbb{N}, U_n = M^n.U_0$$

La matrice $M$, la *matrice de transition* ou *matrice de* Markov, vaut

$$M = \begin{pmatrix} 1 & (1-p) & 0 & 0 & 0 & & 0 \\ 0 & 0 & (1-p) & 0 & 0 & & 0 \\ 0 & p & 0 & (1-p) & 0 & & 0 \\ 0 & 0 & p & 0 & \ddots & & \vdots \\ 0 & 0 & 0 & \ddots & \ddots & (1-p) & 0 \\ 0 & & & & p & 0 & 0 \\ 0 & & & & 0 & p & 1 \end{pmatrix}$$

En d'autres termes, $M_{k\ell} = \mathbb{P}(C^s_{n+1} = k | C^s_n = \ell)$.

Peut-on calculer la matrice $M^n$ ?

# Espérance du gain à la sortie

Trouver la loi de l'instant de sortie $T$,

$$T = \begin{cases} +\infty & \text{si } \forall n \in \mathbb{N}, C_n \neq \pm K \\ \min\{n \in \mathbb{N}, C_n = \pm K\} & \text{sinon} \end{cases}$$

voire calculer simplement son espérance, est relativement compliqué. On peut montrer que $\mathbb{P}(T = +\infty) = 0$ et en tracer un histogramme empirique issu de simulation. L'instant de sortie dépend de la position initiale.

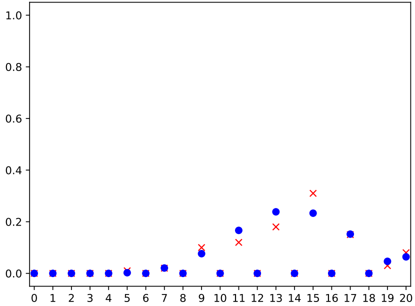



Figure – Ruine du joueur : $p = 0.45$, $K = 10$, $k_0 = 5$, 10 tours de jeu, 100 simulations.

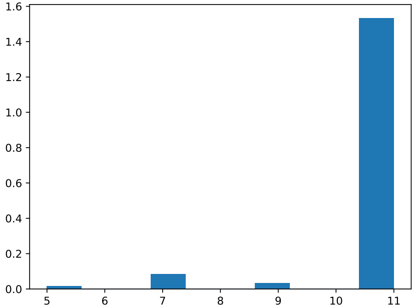



Figure – Ruine du joueur, distribution temps de sortie : $p = 0.45$, $K = 10$, $k_0 = 5$, 10 tours de jeu, 100 simulations. Au bout du 10 tours, le jeu n'est dans la plupart des cas pas fini

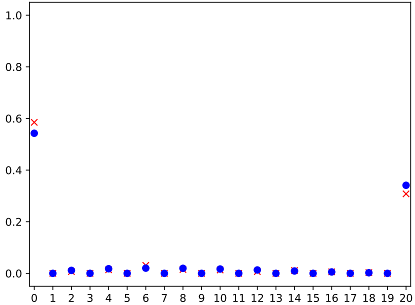



Figure – Ruine du joueur : $p = 0.45$, $K = 10$, $k_0 = 5$, 151 tours de jeu, 1000 simulations.

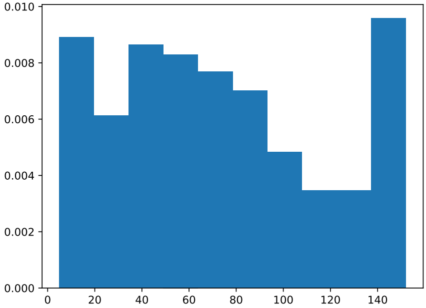



Figure – Ruine du joueur, distribution temps de sortie : $p = 0.45$, $K = 10$, $k_0 = 5$, 151 tours de jeu, 1000 simulations. Le jeu est fini en moins de 151 tours dans 90% des cas.

Assez curieusement, il est « assez » facile de calculer $\mathbb{E}(f(C_T))$ où $f : \{-K, +K\} \to \mathbb{R}$ est donnée à l'avance.

En effet, pour $k_0 \in \{-K, \ldots, +K\}$, posons $\phi_{k_0} = \mathbb{E}(f(C_T))$ où, rappelons le, $C_0 = k_0$. On a $\phi_K = f(K)$, $\phi_{-K} = f(-K)$ et, pour $|k_0| < K$,

$$\begin{aligned}\phi_{k_0} &= \mathbb{E}(f(C_T)|G_1 = +1)\mathbb{P}(G_1 = +1) \\ &\quad + \underbrace{\mathbb{E}(f(C_T)|G_1 = -1)}_{\phi_{k_0+1}\,?}\underbrace{\mathbb{P}(G_1 = -1)}_{1-p} = p.\phi_{k_0-1} + (1-p)\phi_{k_0+1}\end{aligned}$$

La suite (indexée par $k \in \{-K, \ldots, +K\}$) vérifie une récurrence linéaire d'ordre 2 d'équation caractéristique $X = p + (1-p)X^2$ dont les deux racines sont 1 et $\frac{p}{1-p} = \frac{p}{q}$.
Si $p \neq \frac{1}{2}$, il existe donc deux constantes réelles $\lambda$, $\mu$ telles que

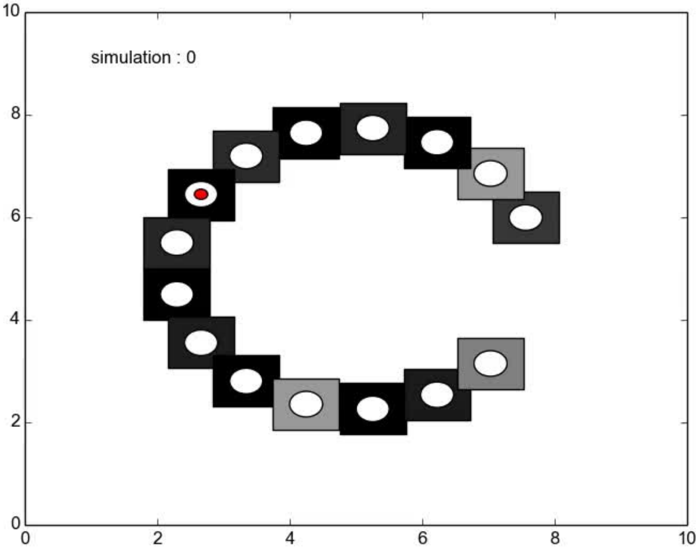
$$\phi_k = \lambda + \mu \frac{p^k}{q^k}$$

En utilisant les valeurs extrémales, on a

$$\lambda = \frac{q^{2K} f(K) - p^{2K} f(-K)}{q^{2K} - p^{2K}}, \mu = -\frac{p^K q^K (f(K) - f(-K))}{q^{2K} - p^{2K}}$$

Si $p = q = \frac{1}{2}$, je vous laisse écrire ce qui se passe.
La probabilité de gagner, ou de perdre se lit sur cette formule, pourquoi ?

Utiliser le script `ruine-joueur.py`. Ce script traite de la ruine du joueur et a permis de tracer les histogrammes présentés précédemment. Le code de simulation de chaine de Markov est universel et peut-être adapté pour toute chaine dont la matrice de transition a une taille « raisonnable ». (Eviter de dépasser $10000 \times 10000$). **Savoir programmer la simulation d'une chaine de** Markov **sera, à mon avis, un exercice typique de l'oral maths-info.**


simulation : 0
10
8
6
4
2
0
0
2
4
6
8
10

# Loi de Poisson

## Proposition-Définition

*Une variable aléatoire $N$ à valeurs entières $\geq 0$ est dite suivre la loi de* Poisson *de paramètre $\lambda > 0$, ce qu'on note*

$$N \sim \mathscr{P}(\lambda)$$

*si, pour tout $n \in \mathbb{N}$,*

$$\mathbb{P}(N = n) = e^{-\lambda}\frac{\lambda^n}{n!}$$

*On a*

$$\mathbb{E}(N) = \lambda, \mathbb{V}(N) = \lambda$$

# Approximation binômiale–Poisson

La loi de Poisson est aussi appelée la *loi du nombre d'événements rares* sur un intervalle de temps donné.
Imaginons que nous sommes sur une chaine de fabrication de circuits éléctroniques.

# Approximation binômiale–Poisson

1. La fabrication est suffisamment bien contrôlée pour que la probabilité qu'un circuit donné soit défectueux est très petite (un événement rare, donc). Le caractère défectueux d'un circuit est indépendant des autres circuits (poussières, micro-coupures,....)
2. Par contre, la chaine fabrique énormément de circuits par unité de temps $\Delta T$
3. et on remarque que la moyenne du nombre du circuits défectueux par unité de temps $\Delta T$ est proportionnelle à la durée de cette unité de temps, *i.e.* vaut $\sim \lambda.\Delta T$.

# Approximation binômiale–Poisson

Sous ces conditions, j'affirme que le nombre de circuits défectueux fabriqués sur une durée de $\Delta T$ suit approximativement une loi de Poisson de paramètre $\lambda.\Delta T$.

En effet, on modélise cette situation de la façon suivante : On se donne un intervalle de temps $I = [T_0, T_0 + \Delta T]$ pendant lequel on fabrique $\mathscr{N}$ circuits. On numérote les circuits de 0 à $\mathscr{N} - 1$ et on suppose que le circuit numéro $n$ tombe de la chaine pendant l'intervalle de temps

$$I_n := \left[T_0 + \frac{n}{\mathscr{N}}\Delta T, T_0 + \frac{n+1}{\mathscr{N}}\Delta T\right], n \in \{0, \dots, \mathscr{N} - 1\}$$

On note $X_n = 1$ si le circuit $n$ est défectueux, $X_n = 0$ sinon.

Les hypothèses se traduisent par :

1. Les v.a. de Bernoulli $X_0, \ldots, X_{\mathscr{N}-1}$ sont indépendantes, identiquement distribuées et suivant une loi $\mathscr{B}(p_{\mathscr{N}})$ où $0 < p_{\mathscr{N}} < 1$ est « petit ».
2. Le nombre $N(\Delta T)$ de circuits défectueux a pour espérance

$$\mathbb{E}(N(\Delta.T)) \simeq \lambda.\Delta T$$

3. $\mathscr{N}$ est très grand

Comme $N(\Delta T) = \sum_{n=0}^{\mathscr{N}-1} X_n$ alors $N(\Delta.T)$ suit une loi binomiale $\mathscr{B}(\mathscr{N}, p_{\mathscr{N}})$. Son espérance est donc

$$\mathbb{E}(N(\Delta T)) = \mathscr{N}.p_{\mathscr{N}} \text{ et } p_{\mathscr{N}} \sim \lambda.\frac{\Delta T}{\mathscr{N}}$$

On a donc, pour tout $k \in \{0, \dots, \mathscr{N}\}$,

$$\mathbb{P}(N(\Delta.T) = k) = \binom{\mathscr{N}}{k} p_{\mathscr{N}}^k (1 - p_{\mathscr{N}})^{\mathscr{N}-k}$$

et on démontre ce qui donne un sens rigoureux à l'affirmation faite en posant $\mu = \lambda.\Delta T$ :

**Proposition**

*Soit $\mu > 0$, et pour $\mathcal{N} \in \mathbb{N}$, $p_{\mathcal{N}} = \frac{\mu}{\mathcal{N}} + o\left(\frac{1}{\mathcal{N}}\right)$. Pour tout $k \in \mathbb{N}$,*

$$\binom{\mathcal{N}}{k} p_{\mathcal{N}}^k (1 - p_{\mathcal{N}})^{\mathcal{N}-k} \underset{\mathcal{N} \to +\infty}{\to} e^{-\mu} \frac{\mu^k}{k!}$$

# La pluie qui tombe

Imaginons que pour un TIPE, on enregistre la pluie qui tombe afin de déterminer l'intensité de l'averse.
Qu'entend-on ? Des plics, des plics et des plocs. Chacun de ces sons donne un "top horaire" sur la bande son. On compte le nombre de tops $N(\Delta t)$ sur une certaine durée $\Delta t$ à différents moments.
Comment se distribue cette quantité ?

Modélisons la situation des plics-plocs de la pluie : la durée entre le "top horaire" de rang $n$ (le plic) et celui de rang $n+1$ (le ploc) est notée $E_{n+1}$. La suite $(E_n)_{n\in\mathbb{N}^*}$ est une suite de variables *indépendantes*, suivant la loi $\mathscr{E}(\lambda)$. Pourquoi ? ceci est du à la propriété d'*absence de mémoire* de cette loi.

$$\forall s,t>0, \mathbb{P}(E>t+s|E>t)=\mathbb{P}(E>s)$$

Modélisons la situation des plics-plocs de la pluie : la durée entre le "top horaire" de rang $n$ (le plic) et celui de rang $n+1$ (le ploc) est notée $E_{n+1}$. La suite $(E_n)_{n\in\mathbb{N}^*}$ est une suite de variables *indépendantes*, suivant la loi $\mathscr{E}(\lambda)$. On pose $T_0 = 0$ et le top horaire de rang $n > 0$ a lieu à l'instant $T_n$. On a $T_{n+1} - T_n = E_{n+1}$ et donc

$$T_n = \sum_{k=1}^{n} E_n$$

Pour $t$ fixé, le nombre de tops $N(t)$ dans l'intervalle $[0, t]$ est donc caractérisé par les conditions suivantes

$$T_{N(t)} \leq t \text{ et } T_{N(t)+1} > t$$

# Calcul de la loi de $N(t)$

Rappelons que $T_n \sim \gamma(n, \lambda)$ et que cette loi a pour densité la fonction $\gamma_{n,\lambda}$ définie par

$$\forall t \in \mathbb{R}, \gamma_{n,\lambda}(t) = \frac{\lambda^n}{\Gamma(n)} t^{n-1} e^{-\lambda . t} \mathbb{1}_{\{t \geq 0\}} = \frac{\lambda^n}{(n-1)!} t^{n-1} e^{-\lambda . t} \mathbb{1}_{\{t \geq 0\}}$$

Pour $n \in \mathbb{N}$, on a les égalités d'événements, prouvant que $N(t)$ est une v.a. à valeurs dans $\mathbb{N} \cup \{+\infty\}$.

$$\{N(t) \geq n\} = \{T_n \leq t\} \text{ et } \{N(t) = n\} = \{N(t) \geq n\} \setminus \{N(t) \geq n+1\}$$

Au niveau des probabilités, pour $n \geq 1$

$$\begin{aligned}
\mathbb{P}(N(t) \geq n) &= \frac{\lambda^n}{(n-1)!} \int_0^t \tau^{n-1} e^{-\lambda.\tau} \, d\tau \\
\mathbb{P}(N(t) = n) &= \mathbb{P}(N(t) \geq n) - \mathbb{P}(N(t) \geq n+1) \\
&= \frac{\lambda^n}{n!} \int_0^t \left(n.\tau^{n-1} - \lambda.\tau^n\right) e^{-\lambda.\tau} \, d\tau \\
&= \frac{\lambda^n}{n!} \left[\tau^n.e^{-\lambda.\tau}\right]_0^t = \frac{(\lambda.t)^n}{n!} e^{-\lambda.t}
\end{aligned}$$

et une formule similaire est vraie pour $n = 0$.

On a donc $N(t) \sim \mathscr{P}(\lambda.t)$ Comme

$$\mathbb{P}(N(t) \in \mathbb{N}) = \sum_{n=0}^{+\infty} \mathbb{P}(N(t) = n) = \sum_{n=0}^{+\infty} \frac{(\lambda.t)^n}{n!} e^{-\lambda.t} = 1$$

Il vient que $\mathbb{P}(N(t) = +\infty) = 0$ et $N(t)$ est presque sûrement à valeurs entières.

# Processus de Poisson : compléménts culturels.

Le processus décrit pour le son de la pluie qui tombe s'appelle un processus de Poisson.
La modélisation adoptée montre que l'espérance du nombre de tops sur une durée $\Delta t$ est proportionnelle à $\Delta t$. Le coefficient de proportionalité est $\lambda$, ce qui donne une méthode pour tenter de l'évaluer.
Le processus de Poisson est le processus de comptage le plus élémentaire : il intervient dans de nombreuses modélisations réelles : nombres d'accidents sur une autoroute, arrivée d'autobus à un arrêt, arrivée de clients à un guichet, ampoules qui explosent, radioactivité, etc...
C'est une source inépuisable de problèmes.

Compter des algues d'une solution « homogène » par cellules de Malassez, c'est considérer que la solution est régie par un processus de Poisson *spatial* : cela implique que la quantité d'algues contenue dans un volume $V$ prélevé suit une loi de Poisson de paramètre $\lambda = c.V$, *i.e.* proportionnel au volume prélevé où $c$ est la concentration de la solution.

Utiliser le script `processus-poisson.py`. Ce script montre comment simuler un processus de Poisson et exhiber quelques statistiques issues de cette simulation. Les histogrammes présentés en figures 7 et 8 en sont issus. Noter que pour simuler une variable de Poisson $X \sim \mathscr{P}(\lambda)$, on peut d'un point de vue pratique, simuler $X \sim \mathscr{B}\left(\mathscr{N}, \frac{\lambda}{\mathscr{N}}\right)$ avec $\mathscr{N}$ « grand »..

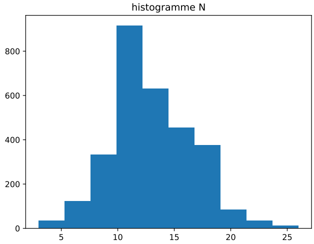



Figure – Simulation d'un processus de Poisson d'intensité $\lambda = 13$. Histogramme du nombre de points dans l'intervalle $[0,1]$, suivant une loi de Poisson de paramètre $\lambda$.

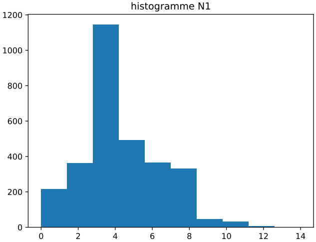



Figure – Simulation d'un processus de Poisson d'intensité $\lambda = 13$. Histogramme du nombre de points dans l'intervalle $\left[\frac{1}{3}, \frac{2}{3}\right]$, suivant une loi de Poisson de paramètre $\frac{\lambda}{3}$.

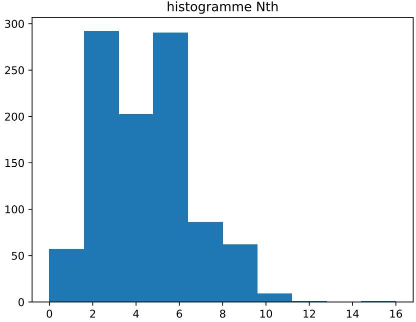



Figure – Simulation directe d'une v.a. de Poisson d'intensité $\lambda = 13/3$.

# Loi d'une v.a discrète

Rappelons le contenu de la définition 6 :
La loi d'une variable discrète $X$ prenant ses valeurs dans un ensemble dénombrable

$$\mathcal{N} = \{x_0, \ldots, x_k, \ldots\} = \{x_k, k \in \mathbb{N}\}$$

est la suite des nombres $p_k = \mathbb{P}(X = x_k), k \in \mathbb{N}$. Elle vérifie :

$$\forall k \in \mathbb{N}, p_k \in [0,1], \sum_{k=0}^{+\infty} p_k = 1.$$

1. On a $\mathbb{P}(X \notin \mathcal{N}) = 0$.
2. Réciproquement, si $\mathcal{N}$ est un ensemble dénombrable, $X$ une v.a. telle que $\sum_{x\in\mathcal{N}} \mathbb{P}(X = x) = 1$ alors $\mathbb{P}(X \notin \mathcal{N}) = 0$ et on peut considérer que $X$ est à valeurs dans $\mathcal{N}$.

# Cas particuliers

1. Si $X$ est à valeurs dans $\mathbb{N}$, sa loi est donnée par la famille de nombres $p_n$ définie par $\forall n \in \mathbb{N}, p_n = \mathbb{P}(X = n)$.
2. Si $X$ est à valeurs dans $\mathbb{Z}$, sa loi est donnée par la famille de nombres $p_n$ définie par $\forall n \in \mathbb{Z}, p_n = \mathbb{P}(X = n)$.
3. Si $Z = (X, Y)$ est à valeurs dans $\mathbb{N}^2$, sa loi est donnée par la famille de nombres

$$p_{mn} = \mathbb{P}(Z = (m, n)) = \mathbb{P}(X = m, Y = n), (m, n) \in \mathbb{N}^2$$

On a (commenter !)

$$\sum_{(m,n)\in\mathbb{N}^2} p_{mn} = \sum_{m=0}^{+\infty} \left( \sum_{n=0}^{+\infty} p_{mn} \right) = 1$$

## Proposition

*Supposons qu'il existe une variable $U$ uniformément distribuée sur $[0,1]$. Soit*

- $\mathscr{N} = \{x_k, k \in \mathbb{N}\}$ *un ensemble dénombrable et*
- $(p_k)_{k\in\mathbb{N}}$ *telle que* $\forall k \in \mathbb{N}, p_k \in [0,1]$, $\sum_{k=0}^{+\infty} p_k = 1$,

*alors il existe une v.a $X$ à valeurs dans $\mathscr{N}$ telle que pour tout $k \in \mathbb{N}$,*

$$\mathbb{P}(X = x_k) = p_k$$

# Un exemple à valeurs $\mathbb{N}^2$

**Exercice 16.—** Montrer qu'en choisissant correctement $\alpha > 0$, il existe une v.a. $X$ à valeurs $\mathbb{N}^2$ telle que pour tout $(m, n) \in \mathbb{N}^2$,
$\mathbb{P}(X = (m, n)) = \alpha \frac{1}{(m+n+1)(m+n+2)(n+2)}$
On donne les décompositions :

1. Pour $m \geq 0$, $n \geq 0$, $\frac{1}{(m+n+1)(m+n+2)(n+2)} = \frac{1}{n+2}\left(\frac{1}{m+n+1} - \frac{1}{m+n+2}\right)$
2. $\frac{1}{(n+1)(n+2)} = \frac{1}{n+1} - \frac{1}{n+2}$

La méthode de construction d'une v.a. discrète à loi imposée est une proche variante des méthodes que nous avons utilisées pour simuler des variables discrètes prenant un nombre fini de valeurs ou des variables à densité. On se base sur l'existence d'une v.a. $U$ uniforme sur $[0,1]$ que l'on corrige à l'aide des fonctions de répartition Cette méthode s'implémente en machine. On peut aussi utiliser les fonctions présentes dans le module `scipy.stats` qui donne directement distribution, fonction de répartition et variables aléatoire suivant une loi imposée classique. Utiliser le script `simulation-discretes-infinies.py`.

# Fonction de répartition

Supposons que $X$ soit à valeurs $\mathbb{N} \subset \mathbb{R}$, rappelons que la fonction de répartition d'une v.a. réelle $X$ est la fonction $F_X$ définie par

$$\forall x \in \mathbb{R}, F_X(x) = \mathbb{P}(X \leq x)$$

Comme $1\!\!1_{\{X \leq x\}} = h(X)$ avec $h(.) = 1\!\!1_{\{. \leq x\}}$, on a donc

$$F_X(x) = \sum_{n \in \mathbb{N}} 1\!\!1_{\{n \leq x\}} \mathbb{P}(X = n) = \sum_{n \in \mathbb{N}, n \leq x} \mathbb{P}(X = n)$$

$F_X$ est donc nul sur $]-\infty, 0[$ et est constante sur chacun des intervalles $I_k = ]k, k+1[,\ k \in \mathbb{N}$.

# Exercices

1. Dessiner le graphe de la fonction de répartition d'une variable géométrique de paramètre $p \in ]0,1[$. Idem pour une loi de Poisson. Savoir faire ces choses en Python rapidement.
2. Que vaut la dérivée d'une telle fonction de répartition ? Quelle est la différence avec la f.r. d'une variable à densité ?
3. Quelle formule si $X$ à valeurs $\mathbb{Z}$ ?

**Exercice 6.**—On considère $(C_n)_{n\in\mathbb{N}}$ la suite des capitaux dans le problème de la ruine du joueur décrit dans la partie 2.2.1, et, en se donnant $K$ comme seuil de gain ou perte maximal, la suite $(C_n^s)_{n\in\mathbb{N}}$ décrite par la chaine de Markov de matrice $M$.
Soit $T$ l'instant d'arrêt du jeu, *i.e.*

$$T = \begin{cases} +\infty & \text{si } \forall n \in \mathbb{N},\, C_n \notin \{-K, +K\} \\ \min\{n \in \mathbb{N}^*,\, C_n \in \{-K, +K\}\} & \text{sinon} \end{cases}$$

On suppose que $C_0 = k_0$ un capital donné.
**1.** Montrer que $\mathbb{P}(T \leq n) = \mathbb{P}(C_n^s = K) + \mathbb{P}(C_n^s = -K)$ et exprimer cette quantité en fonction de $M^n$.
**2.** Proposer une procédure permettant de calculer informatiquement la fonction de répartition de $T$. Pouvez vous l'implémenter dans le script `ruine-joueur.py` ?

# Représentations graphiques : Géométrique

Pour simuler une v.a. géométrique sur $\mathbb{N}$ de paramètre $p$, on peut

- utiliser une variable exponentielle. On a, si $X \sim \mathscr{E}(\lambda)$ alors

$$G = \lfloor X \rfloor \sim \mathscr{G}(\mathbb{N}, p) \text{ et } G' = \lfloor X \rfloor + 1 \sim \mathscr{G}(\mathbb{N}^*, p)$$

  où $p = 1 - e^{-\lambda}$, *i.e.* $\lambda = \ln \frac{1}{1-p}$ et $\lfloor . \rfloor$ désigne la fonction partie entière.

- Soit simuler le Processus d'attente du premier succès, *i.e.* tirer des v.a. de ancetreBernoulli indépendantes de paramètre $p$ et arrêter le décompte dès que l'on obtient la valeur 1.

Remarquons que la première méthode est essentiellement une variation de la méthode présentée dans le script. La fonction partie entière « trouve » le plus grand entier inférieur à son argument.
Voir les graphiques en Figs. 5a, 5b, 5c et 5d.

# Représentations graphiques : Géométrique

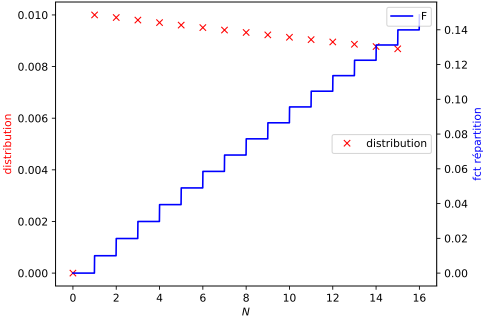



Figure – Distribution géometrique, $p = 0.01$

# Représentations graphiques : Géométrique

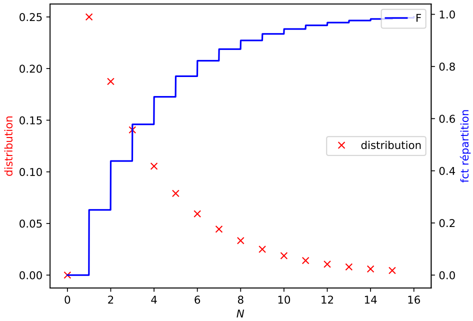



Figure – Distribution géometrique, $p = 0.25$

# Représentations graphiques : Géométrique

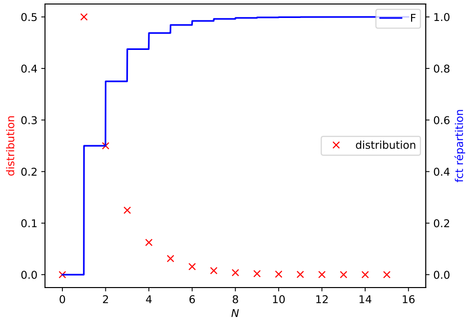



Figure – Distribution géometrique, $p = 0.5$

# Représentations graphiques : Poisson

Pour simuler une v.a. de Poisson de paramètre $\lambda > 0$, on peut

- Soit calculer et « inverser » la fonction de répartition ;
- Soit simuler un Processus de Poisson, *i.e.* sommer des variables exponentielles de paramètre $\lambda$ et arrêter le décompte dès que l'on dépasse la valeur 1.
- Soit utiliser l'approximation Binomiale/Poisson, *i.e.* sommer $N$ variables de Bernoulli indépendantes de paramètre $p = \frac{\lambda}{N}$ avec $N$ grand.

Voir les graphiques en Figs. 6a, 6b, 6c et 6d.

# Représentations graphiques : Poisson

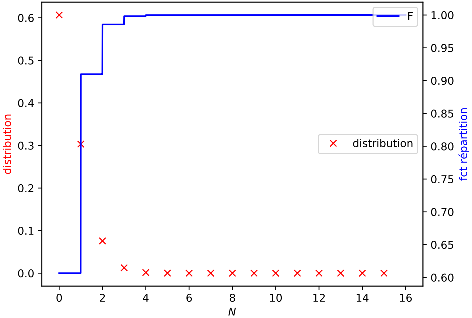



Figure – Distribution de Poisson,$\lambda = \frac{1}{2}$

# Représentations graphiques : Poisson

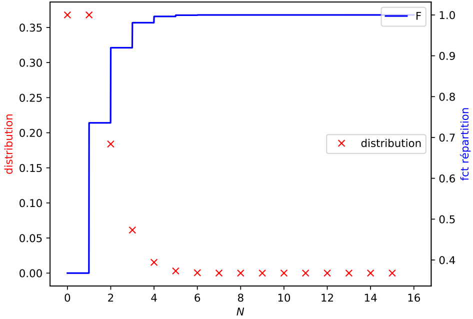



Figure – Distribution de Poisson,$\lambda = 1$

# Représentations graphiques : Poisson

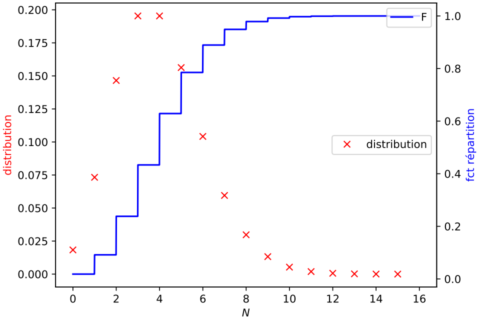



Figure – Distribution de Poisson, $\lambda = 4$

# Position du problème

Dans la partie 2, on a considéré des suites de variables indépendantes suivant soit chacune une loi de Bernoulli (expérience de Bernoulli), soit chacune une loi exponentielle (processus de Poisson), soit des lois imposées par une certaine matrice de transition (pour une chaine de Markov, on a besoin d'une suite infinie indexée par $\mathbb{N}$ de variables en chaque sommet du graphe.).
Une telle chose est-elle possible ?
Tourné autrement, peut-on trouver un $(\Omega, \mathscr{A}, \mathbb{P})$ sur lequel de telles familles de v.a définies existent ? Le point philosophique est d'affirmer que tous les calculs effectués jusqu'à présent s'inscrivent dans le cadre théorique imposé et sont donc corrects.

# Existence de familles infinies de v.a. indépendantes

### Théorème

1. *S'il existe une v.a $U$ uniformément distribuée sur $[0,1]$, il existe une famille dénombrable de variables de* Bernoulli *de paramètre $p = \frac{1}{2}$.*
2. *S'il existe une famille dénombrable de variables de* Bernoulli *de paramètre $p = \frac{1}{2}$, il existe une famille dénombrable de v.a $(U_n)$ indépendantes, uniformément distribuées sur $[0,1]$.*
3. *S'il existe une v.a $U$ uniformément distribuée sur $[0,1]$, et si $(\mathscr{L}_n)_{n\in\mathbb{N}}$ est une famille de lois de v.a réelles alors il existe une famille de v.a. réelles $(Y_n)_{n\in\mathbb{N}}$ indépendante(s) telle que*

$$\forall n \in \mathbb{N}, Y_n \sim \mathscr{L}_n$$

# Eléments de construction

1. Connaissant une v.a $U \sim \mathscr{U}_{[0,1]}$, on peut lui associer la suite de v.a. $(X_n)_{n\geq 1}$ où $X_n$ désigne le $n$-ième chiffre dans le développement binaire de $U$. Cette suite $(X_n)_{n\in\mathbb{N}}$ est indépendante et chaque $X_n \sim \mathscr{B}(\frac{1}{2})$.
2. $(\mathbb{N}^*)^2$ est dénombrable infini. On peut supposer que la suite de v.a donnée est numérotée par cet ensemble. $(X_{n,m})_{(n,m)\in(\mathbb{N}^*)^2}$. Pour chaque $n \in \mathbb{N}^*$, on pose

$$U_n = \sum_{m=1}^{+\infty} \frac{X_{nm}}{2^m}$$

$(U_n)_{n\in\mathbb{N}^*}$ est indépendante et uniformément distribuée.
3. Pour le dernier point, partant de $U$, on construit $X_{nm}$, puis $U_n$ puis $Y_n$ en suivant une vieille recette[4].

# La formule de transfert, espérances

Nous avons déjà vu en exemple (exercice 16) de tels couples à valeurs dans $\mathbb{N}^2$. Nous avons notamment montré l'existence d'une v.a. $X$ à valeurs $\mathbb{N}^2$ telle que pour tout $(m,n)\in\mathbb{N}^2$, $\mathbb{P}(X=(m,n))=\frac{1}{(m+n+1)(m+n+2)(n+2)}$.
Résumons ce qui a été dit à ce moment, étant donné une suite, doublement indexée $(p_{mn})_{(m,n)\in\mathbb{N}^2}$ telle que

$$\forall m,n\in\mathbb{N}, p_{mn}\geq 0 \text{ et } \sum_{(m,n)\in\mathbb{N}^2} p_{mn}=1,$$

on est en mesure de construire une v.a. $X$ à valeurs dans $\mathbb{N}^2$ telle que $\forall (m,n)\in\mathbb{N}^2, \mathbb{P}(X=(m,n))=p_{mn}$.
On peut poser $X=(X_1,X_2)$ où $X_1$, $X_2$ sont deux v.a. à valeurs entières.

On peut tenter de calculer les espérances de chacune de ces variables à valeurs entières : on a, par la formule de transfert, sous réserve de convergence absolue de chacune des séries

$$\mathbb{E}(X_1) = \sum_{(m,n)\in\mathbb{N}^2} m.p_{mn},\ \mathbb{E}(X_2) = \sum_{(m,n)\in\mathbb{N}^2} n.p_{mn}$$

Dans le cas de notre exemple, on a, pourvu que la série converge

$$\begin{aligned}\mathbb{E}(X_2+2) &= \sum_{(m,n)\in\mathbb{N}^2} \frac{n+2}{(m+n+1)(m+n+2)(n+2)} \\ &= \sum_{(m,n)\in\mathbb{N}^2} \frac{1}{(m+n+1)(m+n+2)}\end{aligned}$$

et on montre que ceci diverge : $X_2$ n'est pas intégrable. De façon plus élémentaire $X_1$ n'est pas intégrable non plus.

# Variance/covariance d'un couple discret

De manière analogue, on peut calculer les moments d'ordre 2 de $X_1$ et $X_2$. On a

$$\mathbb{E}(X_1^2) = \sum_{(m,n)\in\mathbb{N}^2} m^2.p_{mn},\ \mathbb{E}(X_2^2) = \sum_{(m,n)\in\mathbb{N}^2} n^2.p_{mn}$$

Exemples.

Si $\mathbb{E}(X_1^2)$ est fini, on dit que $X_1$ est de carré intégrable. $X_1$ est intégrable (pourquoi ?) et on définit sa variance par

$$\mathbb{V}(X_1) = \mathbb{E}((X_1 - \mathbb{E}(X_1))^2) = \mathbb{E}(X_1^2) - \mathbb{E}(X_1)^2$$

Si $\mathbb{E}(X_1^2)$ et $\mathbb{E}(X_2^2)$ sont finis, $X_1$ et $X_2$ sont de carré intégrable et leur covariance est bien définie. On a

$$\mathbb{E}(X_1.X_2) = \sum_{(m,n)\in\mathbb{N}^2} m.n.p_{mn}$$

et

$$\mathbb{C}ov(X_1, X_2) = \mathbb{E}(X_1.X_2) - \mathbb{E}(X_1).\mathbb{E}(X_2)$$

Exemples.

Etant donné un couple aléatoire à valeurs $\mathbb{N}^2$, $X = (X_1, X_2)$ peut-on retrouver les lois de $X_1$ de $X_2$ à partir de celle du couple ?

### Proposition-Définition

*Si $X$ est à valeurs dans $\mathbb{N}^2$, $p_{mn} = \mathbb{P}(X = (m,n))$ alors*

1. *La loi de $X_1$ est donnée par $\forall m \in \mathbb{N}, \mathbb{P}(X_1 = m) = \sum_{n=0}^{+\infty} p_{mn}$*
2. *La loi de $X_2$ est donnée par $\forall n \in \mathbb{N}, \mathbb{P}(X_2 = n) = \sum_{m=0}^{+\infty} p_{mn}$*

*Ces deux lois sont appelées les lois marginales de $X$.*

Exemple. Reprenons le vecteur aléatoire à valeurs $\mathbb{N}^2$ dont la loi est donnée dans l'exercice 16 :

$$\mathbb{P}((X_1, X_2) = (m, n)) = p_{mn} = \frac{1}{(m+n+1)(m+n+2)(n+2)}$$

On a, pour $n \in \mathbb{N}$,

$$\begin{aligned}
\mathbb{P}(X_2 = n) &= \sum_{m=0}^{+\infty} \mathbb{P}(X_1 = m, X_2 = n) \\
&= \sum_{m=0}^{+\infty} \frac{1}{(m+n+1)(m+n+2)(n+2)} = \frac{1}{(n+1)(n+2)}
\end{aligned}$$

Pour $m \in \mathbb{N}$, on a

$$\mathbb{P}(X_1 = m) = \sum_{n=0}^{+\infty} \frac{1}{(m+n+1)(m+n+2)(n+2)}$$

Cette série est convergente, peut-on simplifier son expression ?

## Proposition

*Soit $X = (X_1, X_2)$ un vecteur aléatoire à valeurs dans $\mathbb{N}^2$, $X_1$ et $X_2$ sont indépendantes si et seulement si*

$$\forall (m,n) \in \mathbb{N}^2, \mathbb{P}(X = (m,n)) = \mathbb{P}(X_1 = m)\mathbb{P}(X_2 = n)$$

Remarque : Etant donnée la loi de $X$, $(p_{mn})_{(m,n)\in\mathbb{N}^2}$, on peut calculer la première et la deuxième marginales : $(p_m^1)_{m\in\mathbb{N}}$, $(p_n^2)_{n\in\mathbb{N}}$. $X_1$ et $X_2$ sont indépendantes ssi

$$\forall (m,n) \in \mathbb{N}^2, p_{mn} = p_m^1 p_n^2$$

Dans notre exemple, celui de l'exercice 16, les deux composantes de $X$ ne sont pas indépendantes.

**Proposition**

*Si $X_1, X_2$, à valeurs $\mathbb{N}$ sont indépendantes, alors la loi de $Y = X_1 + X_2$ est donnée par*

$$\forall k \in \mathbb{N}, \mathbb{P}(Y = k) = \sum_{\ell=0}^{k} \mathbb{P}(X_1 = \ell).\mathbb{P}(X_2 = k - \ell)$$

Exemples. Cas à valeurs $\mathbb{Z}$.

Soient $X$, $Y$ deux v.a indépendantes, géométriques sur $\mathbb{N}$ de paramétres $p$ et $r$. On pose $q = 1-p$, $s = 1-r$.
Calculons la loi de $X+Y$. On a, pour $n \in \mathbb{N}$, si $q \neq s$.

$$\begin{aligned}
\mathbb{P}(X+Y=n) &= \sum_{k=0}^{+\infty} \mathbb{P}(X=k \text{ et } Y=n-k) = p.r.\sum_{k=0}^{n} q^k.s^{n-k} \\
&= p.r.\frac{q^n - s^n}{q-s}
\end{aligned}$$

Ce n'est pas un résultat à retenir !

Soient $X$, $Y$ deux v.a indépendantes, géométriques sur $\mathbb{N}$ de paramétres $p$ et $r$. On pose $q = 1 - p$, $s = 1 - r$.
Calculons la loi de $X + Y$. Si par contre $q = s$, on obtient

$$\begin{aligned}\mathbb{P}(X+Y=n) &= \sum_{k=0}^{+\infty} \mathbb{P}(X=k \text{ et } Y=n-k) = p^2.\sum_{k=0}^{n} q^k.q^{n-k} \\ &= p^2.(n+1)q^n\end{aligned}$$

Cette loi porte un nom, il s'agit de la loi Binômiale Négative sur $\mathbb{N}$ de paramètres 2 et $p$.

On s'intéresse brièvement à la loi de $D = X - Y$. $D$ est à valeurs dans $\mathbb{Z}$, la formule de convolution vue ne fonctionne pas directement et il faut refaire le raisonnement. On a, pour $\delta \in \mathbb{Z}$,

$$\mathbb{P}(D = \delta) = \sum_{n \in \mathbb{Z}} \mathbb{P}(X = n \text{ et } Y = n - \delta) = p^2 \sum_{n=0}^{+\infty} q^n . q^{n-\delta} \mathbb{1}_{\{n-\delta \geq 0\}}$$

1. Si $\delta \geq 0$, on a $\mathbb{P}(D = \delta) = p^2 \sum_{n=\delta}^{+\infty} q^{2n} . q^{-\delta} = \frac{p^2}{1-q^2} q^{\delta}$
2. Si $\delta \leq 0$, on a $\mathbb{P}(D = \delta) = p^2 \sum_{n=0}^{+\infty} q^{2n} . q^{-\delta} = \frac{p^2}{1-q^2} q^{-\delta}$

Graphe ?

### Proposition

*Si* $X \sim \mathscr{P}(\lambda)$, $Y \sim \mathscr{P}(\mu)$, $X$ *et* $Y$ *sont indépendantes alors* $X + Y \sim \mathscr{P}(\lambda + \mu)$,

La démonstration est une simple application de la formule précédente et de la formule du binôme.

# Un exemple où les deux v.a ne sont pas indépendantes

Supposons $X = (M, N)$ à valeurs dans $\mathbb{N}^2$ dont la loi est donnée par

$$\mathbb{P}(X = (m,n)) = \frac{1}{(m+n+1)(m+n+2)(n+2)(n+3)}$$

Quelle est la loi de $S = M + N$ ? On a, pour $s \in \mathbb{N}$.

$$\begin{aligned}\mathbb{P}(S = s) &= \sum_{m+n=s} \frac{1}{(m+n+1)(m+n+2)(n+2)(n+3)} \\ &= \frac{1}{(s+1)(s+2)} \sum_{n=0}^{s} \frac{1}{(n+2)(n+3)} = \frac{1}{(s+1)(s+2)}\left(\frac{1}{2} - \frac{1}{s+3}\right)\end{aligned}$$

# Somme de $k$ v.a indépendantes géométriques

On montre par récurrence que la loi de $S_k = X_1 + \cdots + X_k$ où $(X_\ell)_{1 \leq \ell \leq k}$ sont indépendantes, géométriques de paramètre $p$ sur $\mathbb{N}$ vérifie, pour tout $n \in \mathbb{N}$,

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

$S_k$ suit la loi Binomiale négative de paramètres $k$ et $p$.
Comment démontrer cette formule ? Une récurrence ! (Il faut utiliser, pour gérer les coefficients binomiaux, la formule de Vandermonde qui utilise la formule du triangle Pascal).
Comment trouver cette loi ? Fonction génératrice ! (c'est hors programme)
$S_1 + 1$ suit une loi géométrique sur $\mathbb{N}^*$. C'est la loi du rang du premier succès lors d'une suite d'épreuves de Bernoulli.

# Loi du rang du $k$-ième succès

On montre par récurrence sur $k$ que la loi de $R_k$, le rang du $k$-ième succès est la loi de $S_k + k$. Posons $\Sigma_k = R_k - k$. On a, pour $n \geq \ell \geq 0$,

$$\begin{aligned} \mathbb{P}(R_{k+1} = n+k+1 | R_k = \ell + k) &= p.q^{n-\ell} \text{ pourquoi si } n < \ell\,? \\ \mathbb{P}(\Sigma_{k+1} = n | \Sigma_k = \ell) &= p.q^{n-\ell} \\ \mathbb{P}(\Sigma_{k+1} = n) &= \sum_{\ell=0}^{n} p.q^{n-\ell}\mathbb{P}(\Sigma_k = \ell) \end{aligned}$$

$\Sigma_{k+1}$ a donc pour loi la loi de la somme de $\Sigma_k$ et d'une v.a. géométrique sur $\mathbb{N}$ indépendante de paramètre $p$.
Par hyp. de rec $\Sigma_k$ a même loi que $S_k$ et donc $\Sigma_{k+1}$ a même loi que $S_{k+1}$.

# Lois conditionnelles

Si $X = (X_1, X_2)$, on peut calculer la première marginale et ensuite calculer la la loi de $X_2$ sachant $X_1$.
Résumons : Si $(p_{mn})_{(m,n)\in\mathbb{N}^2}$ est la loi de $X$, $p^1_m = \mathbb{P}(X_1 = m) = \sum_n p_{mn}$.
On a alors, pour $m$ tel que $\mathbb{P}(X_1 = m) > 0$ que

$$\mathbb{P}(X_2 = n|X_1 = m) = \frac{\mathbb{P}(X_1 = m \text{ et } X_2 = n)}{\mathbb{P}(X_1 = m)} = \frac{p_{mn}}{\sum_n p_{mn}}$$

A $m$ fixé, la suite $(\mathbb{P}(X_2 = n|X_1 = m))_{n\in\mathbb{N}}$ s'appelle la loi conditionnelle de $X_2$ sachant $X_1 = m$.

Il est clair que l'on peut spécifier la loi d'un couple $X = (X_1, X_2)$ en donnant

1. La loi de $X_1$ : $(\mathbb{P}(X_1 = m))_{m\in\mathbb{N}} = (q_m)_{m\in\mathbb{N}}$,
2. La loi de $X_2$ sachant $X_1$ : $(\mathbb{P}(X_2 = n|X_1 = m))_{m,n\in\mathbb{N}} = (r_{mn})_{m,n\in\mathbb{N}}$,

La loi de $X$ est donnée par

$$\mathbb{P}(X = (m, n)) = q_m . r_{mn}$$

Exemples.