Variables aléatoires discrètes

On considère dans ce chapitre un univers Ω dénombrable et une probabilité P définie sur Ω.

1. Variable aléatoire réelle

Définition d'une variable aléatoire réelle

Une variable aléatoire réelle X est une application définie sur un univers Ω et à valeurs dans $\mathbb{R}$.

$$X:\ \Omega \to \mathbb{R}$$
$$\omega \mapsto X(\omega)$$

On note $X(\Omega)$ l'ensemble des valeurs réelles prisent par $X(\omega)$ lorsque ω décrit l'univers Ω.

$$X(\Omega)=\{X(\omega)|\omega\in\Omega\}\subset\mathbb{R}$$

Si Ω est dénombrable alors $X(\Omega)$ est dénombrable, X est dite discrète et on peut noter $X(\Omega)=(x_i)_{i\in\mathbb{N}}$

Exemple : l'expérience consiste à lancer une pièce jusqu'à obtention de « Pile ». On note X le nombre de lancers nécessaires jusqu'à l'arrêt de l'expérience.

$\Omega=\{P,FP,FFP,FFFP,\dots\}$

$X(\Omega)=\dots$

```python
import numpy as np

def experience():
    mot=np.random.choice(['P','F'],1,p=[1/4,3/4])[0]
    while mot[-1]!='P':
        mot=mot+np.random.choice(['P','F'],1,p=[1/4,3/4])[0]
    return mot

def frequences(L,X):
    LX=[X(evenement) for evenement in L]
    return [[x,LX.count(x)/len(LX)] for x in set(LX)]

L=[experience() for k in range(1000)]
print(frequences(L,lambda eve:len(eve)))
```

[[1, 0.249], [2, 0.182], [3, 0.153], [4, 0.092], [5, 0.086], [6, 0.06], [7, 0.041], [8, 0.034], [9, 0.023], [10, 0.016], [11, 0.009], [12, 0.016], [13, 0.008], [14, 0.006], [15, 0.006], [16, 0.005], [17, 0.003], [18, 0.002], [19, 0.001], [20, 0.002], [21, 0.001], [22, 0.001], [23, 0.001], [24, 0.002], [27, 0.001]]

Événements définis par une variable aléatoire réelle

Soit X une variable aléatoire réelle discrète, A une partie de $\mathbb{R}$ et x un réel.

L'événement $X \in A$ est constitué des issues dont l'image par X appartient à la partie A.

i.e $(X \in A) = \{\omega \in \Omega | X(\omega) \in A\}$

L'événement $X = x$ est constitué des issues dont l'image par X est égale à x.

i.e. $(X = x) = \{\omega \in \Omega | X(\omega) = x\}$

Les événements $(X = x)_{x \in X(\Omega)}$ forment un système complet d'événements.

Remarque : Dans les deux cas, les événements sont dénombrables.

si $A \cap X(\Omega) = \emptyset$ alors $P(X \in A) = \dots$

si $x \notin X(\Omega)$ alors $P(X = x) = \dots$

Exemple : l'expérience consiste à lancer une pièce équilibrée jusqu'à obtention de « pile ». On note X le nombre de lancers nécessaires jusqu'à l'arrêt de l'expérience.

$(X = 3) = ...$

$(X \leqslant 3) = \dots$

$(2 < X \leqslant 5) =$

Image d'une variable aléatoire par une fonction

Soit X une variable aléatoire réelle discrète sur un univers Ω et une fonction $f : X(\Omega) \to \mathbb{R}$.

L'image de la variable aléatoire X par la fonction f est la variable aléatoire discrète notée $f(X)$ qui à tout événement $\omega \in \Omega$ associe le réel $f(X(\omega))$.

$$f(X) : \begin{array}{ccc} \Omega & \to & \mathbb{R} \\ \omega & \mapsto & f(X(\omega)) \end{array}$$

Soient A une partie de $\mathbb{R}$ et x un réel.

L'événement $f(X) \in A$ est constitué des issues dont l'image par $f \circ X$ appartient à la partie A, c'est-à-dire dont l'image par X appartient à l'image réciproque de A par f :

i.e $(f(X) \in A) = (X \in f^{-1}(A))$

L'événement $f(X = x)$ est constitué des issues dont l'image par $f \circ X$ est égale à x, c'est-à-dire dont l'image par X appartient à l'image réciproque de $\{x\}$ par f :

i.e. $(f(X) = x) = (X \in f^{-1}(\{x\}))$

Exemple : l'expérience consiste à lancer une pièce équilibrée jusqu'à obtention de « pile ». On note X le nombre de lancers nécessaires jusqu'à l'arrêt de l'expérience.

Soit $f : x \mapsto x - 2\left\lfloor \frac{x}{2} \right\rfloor$ alors $f(X(\Omega)) = \dots$

Et l'événement $(f(X) = 0) = \dots$

Loi de probabilité d'une variable aléatoire

Soit X une variable aléatoire réelle sur un univers Ω.

La loi de probabilité de X est l'application notée P_X qui à toute partie A de $\mathbb{R}$ associe la probabilité de l'événement $X \in A$.

$$P_X : \begin{array}{ccc} \mathscr{P}(\mathbb{R}) & \to & [0;1] \\ A & \mapsto & P(X \in A) \end{array}$$

Données suffisantes pour définir une loi de probabilité

Soit X une variable aléatoire réelle discrète sur un univers Ω.

L'application P_X est entièrement déterminée par la donnée de $P(X = x)$ pour tout réel x de $X(\Omega)$.

$$P_X : \begin{array}{ccc} X(\Omega) & \to & [0;1] \\ x & \mapsto & P(X = x) \end{array}$$

Dans ce cas la série $\sum_{x \in X(\Omega)} P(X = x)$ est une série à termes positifs convergente dont la somme est égale à 1.

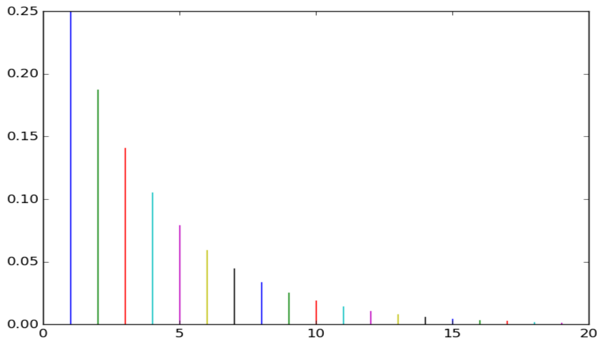
Exemple : l'expérience précédente est réalisée avec une pièce truquée ayant 3 fois plus de chance de donner « Face » que de donner « Pile ». Exemple de simulation de 1000 expériences :

[[1, 0.25], [2, 0.201], [3, 0.144], [4, 0.101], [5, 0.08], [6, 0.063], [7, 0.054], [8, 0.017], [9, 0.03], [10, 0.023], [11, 0.01], [12, 0.006], [13, 0.005], [14, 0.004], [15, 0.005], [16, 0.003], [19, 0.001], [21, 0.001], [24, 0.001], [34, 0.001]]
Les lancers successifs étant mutuellement indépendants on a :

$\forall k \in \mathbb{N}^*,\ \mathrm{P}(\mathrm{X}=k)=\ldots$

Ainsi on a bien : $\sum_{k=1}^{+\infty} \mathrm{P}(\mathrm{X}=k)=$ ...

```python
def P(x):
    if int(x)==x and x>=1:
        return (3/4)**(x-1)*(1/4)
    return 0

import matplotlib.pyplot as plt
for k in range(1,20):
    plt.plot([k,k],[0,P(k)])
plt.show()
```



Formule des probabilités totales utilisant une variable aléatoire

Soient X une variable aléatoire réelle sur un univers Ω et un événement $\mathrm{A} \in \Omega$.

$$\mathrm{P}(\mathrm{A})=\sum_{x \in \mathrm{X}(\Omega)} \mathrm{P}(\mathrm{A} \cap (\mathrm{X}=x))$$

Démonstration : la famille d'événements $(\mathrm{X}=x)_{x \in \mathrm{X}(\Omega)}$ forme un système complet d'événements.

Application à la somme de deux variables aléatoires

Soient X et Y deux variables aléatoires réelles sur un univers Ω .

$$\forall k \in \mathrm{X}+\mathrm{Y}(\Omega),\ \mathrm{P}(\mathrm{X}+\mathrm{Y}=k)=\sum_{x \in \mathrm{X}(\Omega)} \mathrm{P}((\mathrm{X}=x) \cap (\mathrm{Y}=k-x))$$

Démonstration : $\mathrm{P}(\mathrm{X}+\mathrm{Y}=k)=\sum_{x \in \mathrm{X}(\Omega)} \mathrm{P}((\mathrm{X}=x) \cap (\mathrm{X}+\mathrm{Y}=k))$ =...

Définition de la fonction de répartition d'une variable aléatoire

Soit X une variable aléatoire réelle sur un univers Ω .
La fonction de répartition de X,notée F_{X} , est la fonction qui à un réel x associe la probabilité de l'événement $(\mathrm{X} \leqslant x)$

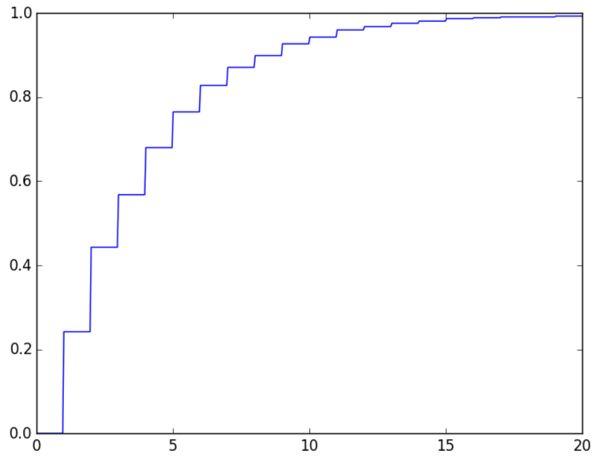
$$\begin{array}{llll} \mathrm{F}_{\mathrm{X}}: & \mathbb{R} & \rightarrow & [0;1] \\ & x & \mapsto & \mathrm{P}(\mathrm{X} \leqslant x) \end{array}$$

Si X est une variable aléatoire discrète alors $\forall x \in \mathbb{R}$, $\mathrm{F}_{\mathrm{X}}(x)=\sum_{x_i \in \mathrm{X}(\Omega) \cap]-\infty;x]} \mathrm{P}(\mathrm{X}=x_i)$.

Si $\mathrm{X}(\Omega) \subset \mathbb{Z}$ alors $\forall x \in \mathbb{R}$, $\mathrm{F}_{\mathrm{X}}(x)=\mathrm{F}_{\mathrm{X}}(\lfloor x \rfloor)$

Exemple : l'expérience précédente est réalisée avec une pièce truquée ayant 3 fois plus de chance de donner « Face » que de donner « Pile ». On note X le rang du premier « Pile » obtenu.
Fréquences cumulées croissantes obtenues sur 1000 expériences.

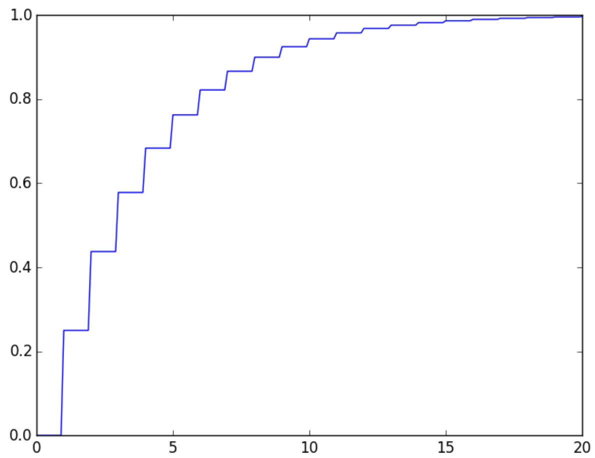
```python
import numpy as np

def experience():
    mot=np.random.choice(['P','F'],1,p=[1/4,3/4])[0]
    while mot[-1]!='P':
        mot=mot+np.random.choice(['P','F'],1,p=[1/4,3/4])[0]
    return mot

def frequences_cumulees_croissantes(LX,k):
    return sum([1 for x in LX if x<=k])/len(LX)

X=lambda eve:len(eve)
LX=[X(experience()) for k in range(1000)]

import matplotlib.pyplot as plt
I=np.linspace(0,20,500)
plt.plot(I,[frequences_cumulees_croissantes(LX,x) for x in I])
plt.show()
```



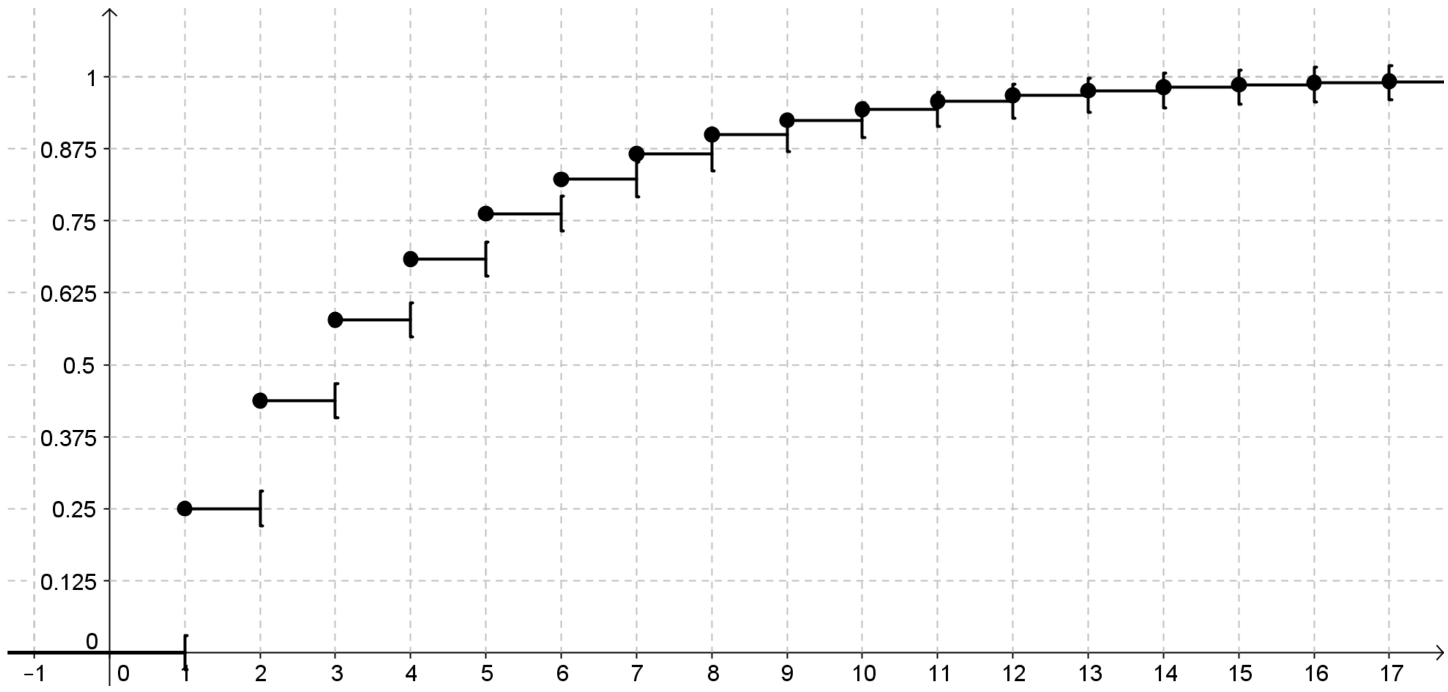
Si $x\in]-\infty;1[$ alors $F_X(x)=\dots$

Si $x\in[1;+\infty[$ alors $F_X(x)=\dots$

```python
def P(x):
    if int(x)==x and x>=1:
        return (3/4)**(x-1)*(1/4)
    return 0

def F(x):
    return sum(P(k) for k in range(1,int(x)+1))

import numpy as np
LX=np.linspace(0,20,20*10+1)
LY=[F(x) for x in LX]

import matplotlib.pyplot as plt
plt.plot(LX,LY)
plt.show()
```



⚠ Les segments semblant parallèles à l'axe des ordonnées n'ont pas de sens pour la représentation graphique d'une fonction.



Propriété de croissance de la fonction de répartition

Soit X une variable aléatoire réelle sur un univers Ω et F_X sa fonction de répartition.

F_X est croissante sur $\mathbb{R}$

i.e. $\forall (a\,;b) \in \mathbb{R}^2 \,,\ a<b \Rightarrow F_X(a) \leqslant F_X(b)$

Démonstration : si $a<b$ alors $(X\leqslant a) \subset (X \leqslant b)$ donc ...

□

Remarques : Soit un réel a, $(X>a) = \overline{(X\leqslant a)}$ donc $\boxed{P(X>a)=1-F_X(a)}$

Soit deux réels $a<b$, $\begin{cases}(X\leqslant b)=(X\leqslant a)\cup(a<X\leqslant b)\\(X\leqslant a)\cap(a<X\leqslant b)=\varnothing\end{cases}$ donc $P(X\leqslant b)=P(X\leqslant a)+P(a<X\leqslant b)$

Ainsi $\boxed{P(a<X\leqslant b)=F_X(b)-F_X(a)}$

Exemple : $P(1{,}5<X\leqslant 5{,}2)=\ldots$

Détermination d'une loi de probabilité à l'aide de sa fonction de répartition

Soit X une variable aléatoire réelle discrète sur un univers Ω, $X(\Omega)=(x_i)_{i\in\mathbb{N}}$ et F_X sa fonction de répartition.

La fonction F_X est constante par morceaux et continue à droite en tout point.

Plus précisément, pour tout $(i\,;j)\in\mathbb{N}^2$ tel que $\begin{cases}x_i<x_j\\ ]x_i\,;x_j[\cap X(\Omega)=\varnothing\end{cases}$, on a, $\forall x\in[x_i\,;x_j[\,,\ F_X(x)=F_X(x_i)$

Soit un réel k. La fonction de répartition F_X est discontinue à gauche en k si et seulement si $P(X=k)>0$.

Plus précisément, si $\lim\limits_{\substack{x\to k\\x<k}} F_X(x)<F_X(k)$ alors $F_X(k)-\lim\limits_{\substack{x\to k\\x<k}} F_X(x)=P(X=k)$

Remarque : on peut retenir que :

les points de discontinuité de F_X sont les éléments de $X(\Omega)$

un « saut » dans les valeurs de $F_X(x)$ au point $x=x_i$ correspond à $P(X=x_i)$

Démonstration : $\forall x\in[x_i\,;x_j[\,,\ F_X(x)=F_X(x_i)+P(x_i<X\leqslant x)$ or $]x_i\,;x[\subset]x_i\,;x_j[$ donc $]x_i\,;x[\cap X(\Omega)=\varnothing$

ainsi $P(x_i<X\leqslant x)=0$ d'où : $F_X(x)=F_X(x_i)$

Soit $k\in\mathbb{R}$, puisque X est discrète, il existe deux réels $\begin{cases}M=min(X(\Omega)\cap]k\,;+\infty[)\\m=max(X(\Omega)\cap]-\infty\,;k[)\end{cases}$.

Si $k\notin X(\Omega)$ alors $\forall x\in]m\,,k[\,,\ (X(\Omega)\cap]-\infty\,;k])=(X(\Omega)\cap]-\infty\,;x])$ donc $F_X(x)=F_X(k)$ donc $\lim\limits_{\substack{x\to k\\x<k}} F_X(x)=F_X(k)$

Si $k\in X(\Omega)$ alors $\forall x\in]m\,,k[\,,\ (X(\Omega)\cap]-\infty\,;k])=(X(\Omega)\cap]-\infty\,;x])\cup\{k\}$ donc

$$\mathrm{F_X}(k)=\sum_{x_i\in \mathrm{X}(\Omega)\cap]-\infty;k]} \mathrm{P}(\mathrm{X}=x_i)$$

$$\mathrm{F_X}(k)=\mathrm{P}(\mathrm{X}=k)+\sum_{x_i\in \mathrm{X}(\Omega)\cap]-\infty;x]} \mathrm{P}(\mathrm{X}=x_i)$$

$$\mathrm{F_X}(k)=\mathrm{P}(\mathrm{X}=k)+\mathrm{F_X}(x)$$

Ainsi $\forall x\in]m,k[$, $\mathrm{F_X}(k)-\mathrm{F_X}(x)=\mathrm{P}(\mathrm{X}=k)$ donc $\mathrm{F_X}(k)-\lim\limits_{\substack{x\to k\\ x<k}}\mathrm{F_X}(x)=\mathrm{P}(\mathrm{X}=k)$. □

Exemple de détermination d'une loi de probabilité à partir d'une fonction de répartition :

```python
LX=[-1+k/1000 for k in range(20001)]
LY=[1-(3/4)**max(0,int(x)) for x in LX]

Loi_de_X=[]
for i in range(len(LX)-1):
    if LY[i]<LY[i+1]:
        Loi_de_X.append([LX[i+1],LY[i+1]-LY[i]])

print(Loi_de_X)
```

[[1.0, 0.25], [2.0, 0.1875], [3.0, 0.140625], [4.0, 0.10546875], [5.0, 0.0791015625], [6.0, 0.059326171875], [7.0, 0.04449462890625], [8.0, 0.0333709716796875], [9.0, 0.025028228759765625], [10.0, 0.018771171569824222], [11.0, 0.014078378677368164], [12.0, 0.010558784008026123], [13.0, 0.007919088006019592], [14.0, 0.005939316004514694], [15.0, 0.004454487003386021], [16.0, 0.0033408652525395155], [17.0, 0.0025056489394046366], [18.0, 0.0018792367045534775], [19.0, 0.001409427528415108]]

2. Espérance

Variable aléatoire d'espérance finie

Soit X une variable aléatoire réelle discrète et $\mathrm{X}(\Omega)=(x_i)_{i\in\mathbb{N}}$.

X est dite d'espérance finie si et seulement si la série $\sum\limits_{i\in\mathbb{N}} x_i\mathrm{P}(\mathrm{X}=x_i)$ est absolument convergente.

Si X est d'espérance finie alors l'espérance de X est $\mathrm{E}(\mathrm{X})\overset{\text{def}}{=}\sum\limits_{i=0}^{+\infty} x_i\mathrm{P}(\mathrm{X}=x_i)$.

Remarque : on admet que l'ordre d'énumération de $\mathrm{X}(\Omega)$ n'a pas d'influence.

Exemple : On lance successivement une pièce truquée ayant 3 fois plus de chance de donner « Face » que de donner « Pile ». On note X le rang du premier « Pile » obtenu.

Donc $\mathrm{E}(\mathrm{X})=\ldots$

Contre-exemple : en considérant la fonction $f:x\mapsto 2^x$, la variable aléatoire $f(\mathrm{X})$ n'est pas d'espérance finie car...

Linéarité de l'espérance

Soient X et Y deux variables aléatoires réelles discrètes, $\mathrm{X}(\Omega)=(x_i)_{i\in\mathbb{N}}$ et $\mathrm{Y}(\Omega)=(y_i)_{i\in\mathbb{N}}$ et un réel λ .

Si X est d'espérance finie alors la variable aléatoire $\lambda\mathrm{X}$ est d'espérance finie et $\mathrm{E}(\lambda\mathrm{X})=\lambda\mathrm{E}(\mathrm{X})$

Si X et Y sont d'espérances finies alors la variable aléatoire $\mathrm{X}+\mathrm{Y}$ est d'espérance finie et $\mathrm{E}(\mathrm{X}+\mathrm{Y})=\mathrm{E}(\mathrm{X})+\mathrm{E}(\mathrm{Y})$

Remarque : L'ensemble des variables aléatoires discrètes d'espérance finie est un sous espace vectoriel des fonctions définies sur Ω à valeurs réelles.

L'application qui à une variable aléatoire d'espérance finie associe son espérance est une application linéaire.

Démonstration : par linéarité de la somme d'une série convergente :

$$\mathrm{E}(\lambda\mathrm{X})=\sum_{i=0}^{+\infty}\lambda x_i\mathrm{P}(\mathrm{X}=x_i)=\lambda\sum_{i=0}^{+\infty}x_i\mathrm{P}(\mathrm{X}=x_i)=\lambda\mathrm{E}(\mathrm{X})$$

$$\mathrm{E}(\mathrm{X}+\mathrm{Y})=\sum_{\omega\in\Omega}(\mathrm{X}(\omega)+\mathrm{Y}(\omega))\mathrm{P}(\omega)=\sum_{\omega\in\Omega}\mathrm{X}(\omega)\mathrm{P}(\omega)+\sum_{\omega\in\Omega}\mathrm{Y}(\omega)\mathrm{P}(\omega)=\mathrm{E}(\mathrm{X})+\mathrm{E}(\mathrm{Y})$$

En utilisant la loi du couple de variables aléatoires (X;Y) on a aussi (démonstration hors programme mais instructive):

$\mathrm{E}(\mathrm{X}+\mathrm{Y})=\sum_{i=0}^{+\infty}\sum_{j=0}^{+\infty}(x_i+y_j)\mathrm{P}((\mathrm{X};\mathrm{Y})=(x_i;y_j))$ (en admettant pouvoir sommer dans n'importe quel ordre)

$\mathrm{E}(\mathrm{X}+\mathrm{Y})=\sum_{i=0}^{+\infty}\left(x_i\sum_{j=0}^{+\infty}\mathrm{P}((\mathrm{X};\mathrm{Y})=(x_i;y_j))\right)+\sum_{j=0}^{+\infty}\left(y_j\sum_{i=0}^{+\infty}\mathrm{P}((\mathrm{X};\mathrm{Y})=(x_i;y_j))\right)$

Or d'après la formule des probabilités totales : $\sum_{j=0}^{+\infty}\mathrm{P}((\mathrm{X};\mathrm{Y})=(x_i;y_j))=\mathrm{P}(\mathrm{X}=x_i)$ et $\sum_{i=0}^{+\infty}\mathrm{P}((\mathrm{X};\mathrm{Y})=(x_i;y_j))=\mathrm{P}(\mathrm{Y}=y_j)$

D'où, $\mathrm{E}(\mathrm{X}+\mathrm{Y})=\sum_{i=0}^{+\infty}x_i\mathrm{P}(\mathrm{X}=x_i)+\sum_{j=0}^{+\infty}y_j\mathrm{P}(\mathrm{Y}=y_j)=\mathrm{E}(\mathrm{X})+\mathrm{E}(\mathrm{Y})$ □

Théorème du transfert : espérance de $f(\mathrm{X})$ en fonction des événements élémentaires ou bien en fonction des valeurs de la variables aléatoire.

Soit X une variable aléatoire discrète, $\mathrm{X}(\Omega)=(x_i)_{i\in\mathbb{N}}$, et une fonction $f:\mathrm{X}(\Omega)\to\mathbb{R}$.

La variable aléatoire $f(\mathrm{X})$ est d'espérance finie si et seulement si la série $\sum_{i\geqslant 0}f(x_i)\mathrm{P}(\mathrm{X}=x_i)$ est absolument convergente.

Si $f(\mathrm{X})$ est d'espérance finie alors son espérance est définie par $\mathrm{E}(f(\mathrm{X}))\stackrel{\text{def}}{=}\sum_{i=0}^{+\infty}f(x_i)\mathrm{P}(\mathrm{X}=x_i)$.

Dans le cas où les séries sont absolument convergentes on a : $\mathrm{E}(f(\mathrm{X}))=\sum_{\omega\in\Omega}f(\mathrm{X}(\omega))\mathrm{P}(\omega)=\sum_{i=0}^{+\infty}f(x_i)\mathrm{P}(\mathrm{X}=x_i)$

⚠ En général, si f n'est pas affine, $\mathrm{E}(f(\mathrm{X}))\neq f(\mathrm{E}(\mathrm{X}))$

Exemple : On lance successivement une pièce truquée ayant 3 fois plus de chance de donner « Face » que de donner « Pile ». On note X le rang du premier « Pile » obtenu et $f:x\mapsto x-2\left\lfloor\frac{x}{2}\right\rfloor$.

$\mathrm{E}(f(\mathrm{X}))=\ldots$

3. Variance et écart type d'une variable aléatoire

Espérance de X^2

Soit X une variable aléatoire discrète et $\mathrm{X}(\Omega)=(x_i)_{i\in\mathbb{N}}$.

Si la variable aléatoire X^2 est d'espérance finie alors la variable aléatoire X est d'espérance finie.

i.e. si la série $\sum_{i\geqslant 0}(x_i)^2\mathrm{P}(\mathrm{X}=x_i)$ est convergente alors la série $\sum_{i\geqslant 0}x_i\mathrm{P}(\mathrm{X}=x_i)$ est absolument convergente.

Démonstration : $\forall x\in\mathbb{R}$, $|x|\leqslant x^2+1$, ainsi $\forall i\in\mathbb{N}$, $|x_i|\mathrm{P}(\mathrm{X}=x_i)\leqslant(x_i)^2\mathrm{P}(\mathrm{X}=x_i)+\mathrm{P}(\mathrm{X}=x_i)$

La série $\sum_{i\geqslant 0}\mathrm{P}(\mathrm{X}=x_i)$ est convergente (sa somme vaut 1).

Si la série $\sum_{i\geqslant 0}(x_i)^2\mathrm{P}(\mathrm{X}=x_i)$ est convergente alors par somme, la série $\sum_{i\geqslant 0}((x_i)^2+1)\mathrm{P}(\mathrm{X}=x_i)$ est convergente donc,

d'après le critère de majoration pour les séries à termes positifs, la série $\sum_{i\geqslant 0}x_i\mathrm{P}(\mathrm{X}=x_i)$ est absolument convergente. □

Définition de la variance d'une variable aléatoire discrète

Soit X une variable aléatoire discrète et $\mathrm{X}(\Omega)=(x_i)_{i\in\mathbb{N}}$.

Si X^2 est d'espérance finie alors la variance de X est définie par $\mathrm{V}(\mathrm{X})\stackrel{\text{def}}{=}\mathrm{E}(\mathrm{X}^2)-(\mathrm{E}(\mathrm{X}))^2$

i.e. $\mathrm{V}(\mathrm{X})\stackrel{\text{def}}{=}\sum_{i=0}^{+\infty}(x_i)^2\mathrm{P}(\mathrm{X}=x_i)-\left(\sum_{i=0}^{+\infty}x_i\mathrm{P}(\mathrm{X}=x_i)\right)^2$

et l'écart type de X est défini par $\sigma(\mathrm{X})=\sqrt{\mathrm{V}(\mathrm{X})}$

Remarque : par linéarité de l'espérance :

$\mathrm{E}((\mathrm{X}-\mathrm{E}(\mathrm{X}))^2)=\mathrm{E}(\mathrm{X}^2-2\mathrm{X}\mathrm{E}(\mathrm{X})-(\mathrm{E}(\mathrm{X}))^2)=\mathrm{E}(\mathrm{X}^2)-2\mathrm{E}(\mathrm{X})\times\mathrm{E}(\mathrm{X})+(\mathrm{E}(\mathrm{X}))^2\mathrm{E}(1)=\mathrm{E}(\mathrm{X}^2)-(\mathrm{E}(\mathrm{X}))^2$

Exemple : On lance successivement une pièce truquée ayant 3 fois plus de chance de donner « Face » que de donner « Pile ». On note X le rang du premier « Pile » obtenu.

La série $\sum_{n\geqslant 0} n^2 P(X=n)$ est ...

Donc $V(X)=\ldots$

Variance de $aX+b$

Soit X une variable aléatoire discrète, $X(\Omega)=(x_i)_{i\in\mathbb{N}}$ et deux réels a et b.
Si X^2 est d'espérance finie alors la variable aléatoire $(aX+b)^2$ est d'espérance finie et $V(aX+b)=a^2V(X)$

Démonstration : $(aX+b)^2=a^2X^2+2abX+b^2$

b^2 est d'espérance finie car $\sum_{i=0}^{+\infty} b^2 P(X=x_i)=\ldots$

Si X^2 est d'espérance finie alors X est d'espérance finie donc, par linéarité, $2abX$ est d'espérance finie.
Par linéarité, si X^2 est d'espérance finie alors $a^2X^2+2abX+b^2$ est d'espérance finie.

$V(aX+b)=E((aX+b)^2)-(E(aX+b))^2$ =...

□

Inégalité de Bienaymé-Tchebychev

Soit X une variable aléatoire discrète et $X(\Omega)=(x_i)_{i\in\mathbb{N}}$.
Si X^2 est d'espérance finie alors $\forall\alpha>0,\ P(|X-E(X)|\geqslant\alpha)\leqslant\frac{V(X)}{\alpha^2}$

Remarques : on peut retenir que la probabilité que la distance entre $X(\omega)$ et E(X) soit supérieure à α est majorée par $\frac{V(X)}{\alpha^2}$. Cette majoration n'a d'intérêt que pour $\alpha^2>V(X)$ donc $\alpha>\sigma(X)$.

Démonstration : soit $\alpha>0$, la démonstration repose sur l'égalité des événements : $(|X-E(X)|\geqslant\alpha)=((X-E(X))^2\geqslant\alpha^2)$

$$V(X)=E((X-E(X))^2)=\sum_{i=0}^{+\infty}(x_i-E(X))^2P(X=x_i)$$

$$V(X)=\sum_{\substack{i=0\\(x_i-E(X))^2<\alpha^2}}^{+\infty}(x_i-E(X))^2P(X=x_i)+\sum_{\substack{i=0\\(x_i-E(X))^2\geqslant\alpha^2}}^{+\infty}(x_i-E(X))^2P(X=x_i)$$

Or, par somme de réels positifs ou nuls, $\sum_{\substack{i=0\\(x_i-E(X))^2<\alpha^2}}^{+\infty}(x_i-E(X))^2P(X=x_i)\geqslant 0$

Donc $V(X)\geqslant\sum_{\substack{i=0\\(x_i-E(X))^2\geqslant\alpha^2}}^{+\infty}(x_i-E(X))^2P(X=x_i)\geqslant\alpha^2\sum_{\substack{i=0\\(x_i-E(X))^2\geqslant\alpha^2}}^{+\infty}P(X=x_i)=\alpha^2P((X-E(X))^2\geqslant\alpha^2)=\alpha^2P(|X-E(X)|\geqslant\alpha)$

Or $\alpha^2>0$ donc $\frac{V(X)}{\alpha^2}\geqslant P(|X-E(X)|\geqslant\alpha)$ □

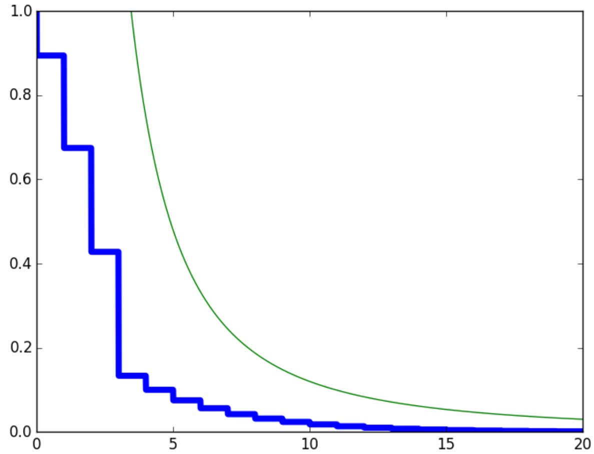
Remarque : $|X-E(X)|\geqslant\alpha \Leftrightarrow X-E(X)\geqslant\alpha$ ou $X-E(X)\leqslant-\alpha \Leftrightarrow X\geqslant E(X)+\alpha$ ou $X\leqslant E(X)-\alpha$

Ainsi $\forall\alpha>0$, $P(|X-E(X)|\geqslant\alpha)=P(X\leqslant E(X)-\alpha)+P(X\geqslant E(X)+\alpha)$
$=P(X\leqslant E(X)-\alpha)+P(X>E(X)+\alpha)+P(X=E(X)+\alpha)$
$=F_X(E(X)-\alpha)+1-F_X(E(X)+\alpha)+P(X=E(X)+\alpha)$

Exemple : l'expérience consiste à lancer une pièce équilibrée jusqu'à obtention de « pile ». On note X le nombre de lancers nécessaires jusqu'à l'arrêt de l'expérience.

On a $E(X)=4$ et $V(X)=12$ donc, d'après l'inégalité de Bienaymé-Tchebychev, $\forall\alpha>0$, $P(|X-4|\geqslant\alpha)\leqslant\frac{12}{\alpha^2}$

```python
def P(x):
    if int(x)==x and x>=1:
        return (3/4)**(x-1)*(1/4)
    return 0

def F(x):
    return sum(P(k) for k in range(1,int(x)+1))

import numpy as np
L_alpha=np.linspace(0,20,100*20+1)
E=4
LY=[F(E-al)+1-F(E+al)+P(E+al) for al in L_alpha]

import matplotlib.pyplot as plt
plt.plot(L_alpha,LY,linewidth=5)
plt.plot(L_alpha,[12/al**2 for al in L_alpha])
plt.axis([0,20,0,1])
plt.show()
```



4. Lois usuelles

Définition de la loi géométrique de paramètre $p \in]0;1[$

Soit un réel $p \in]0;1[$ et X une variable aléatoire.

X suit la loi géométrique de paramètre p notée $\mathcal{G}(p)$ si et seulement si $\begin{cases} X(\Omega)=\mathbb{N}^* \\ \forall k \in \mathbb{N}^*,\ P(X=k)=p(1-p)^{k-1} \end{cases}$

On note alors $X \hookrightarrow \mathcal{G}(p)$.

Démonstration : démontrons que la loi géométrique est une loi de probabilité.
Soit $p \in]0;1[$ alors $-1<-p<0$ donc $0<1-p<1$
ainsi $\forall k \in \mathbb{N}^*,\ 0<(1-p)^{k-1}$ donc par multiplication $0<p(1-p)^{k-1}$.

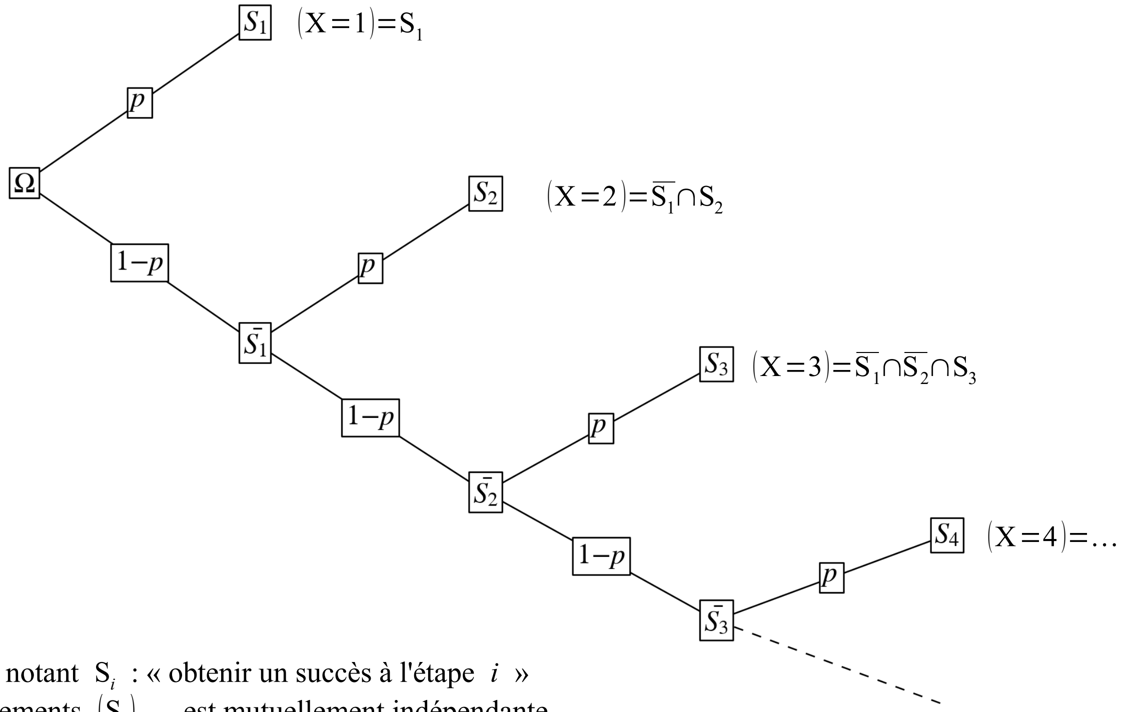
Par ailleurs puisque $0<1-p<1$, la série géométrique $\sum_{k \geqslant 1} (1-p)^{k-1}$ est convergente et par linéarité on a :

$\sum_{k=1}^{+\infty} p(1-p)^{k-1} = \ldots$

□

Champs d'application de la loi géométrique

En considérant une suite infinie d'épreuves de Bernoulli mutuellement indépendantes et de même paramètre p, si X est le rang du premier succès alors X suit la loi $\mathcal{G}(p)$.



Démonstration : En notant S_i : « obtenir un succès à l'étape i »

La famille des événements $(S_i)_{i\in\mathbb{N}}$ est mutuellement indépendante.

Ainsi, $P(X=k)=P\left(S_k\cap\left(\bigcap_{i=1}^{k-1}\overline{S_i}\right)\right)=P(S_k)\times\prod_{i=1}^{k-1}P(\overline{S_i})=p(1-p)^{k-1}$

Exemple de simulation utilisant des expériences de Bernoulli de paramètre $\frac{1}{4}$ indépendantes :

```python
import numpy as np

def Y():
    y=1
    while np.random.binomial(1,0.25)==0:
        y=y+1
    return y

print(Y())
```

Espérance et variance d'une variable aléatoire suivant une loi géométrique

Soit un réel $p\in]0;1[$ et X une variable aléatoire suivant la loi $\mathcal{G}(p)$.

$$E(X)=\frac{1}{p}$$
$$V(X)=\frac{1-p}{p^2}$$

Démonstration :

La série entière de variable x, $\sum_{k\geqslant 1} k\,x^k$ a le même rayon de convergence que celui de la série géométrique $\sum_{k\geqslant 1} x^k$.

Ainsi $\forall x\in]-1;1[$, la série $\sum_{k\geqslant 1} k\,x^{k-1}$ est convergente et sa somme est égale à :

$\sum_{k=1}^{+\infty} k\, x^{k-1} = \dots$

Soit $p \in]0;1[$ alors $0 < 1-p < 1$ donc la série $\sum_{k \geqslant 1} k\, p(1-p)^{k-1}$ est convergente et sa somme est égale à :

$\mathrm{E}(\mathrm{X}) = \sum_{k=1}^{+\infty} k\, p(1-p)^{k-1} = \dots$

La série entière de variable x, $\sum_{k \geqslant 1} k^2 x^k$ a le même rayon de convergence que celui de la série géométrique $\sum_{k \geqslant 1} k\, x^k$.

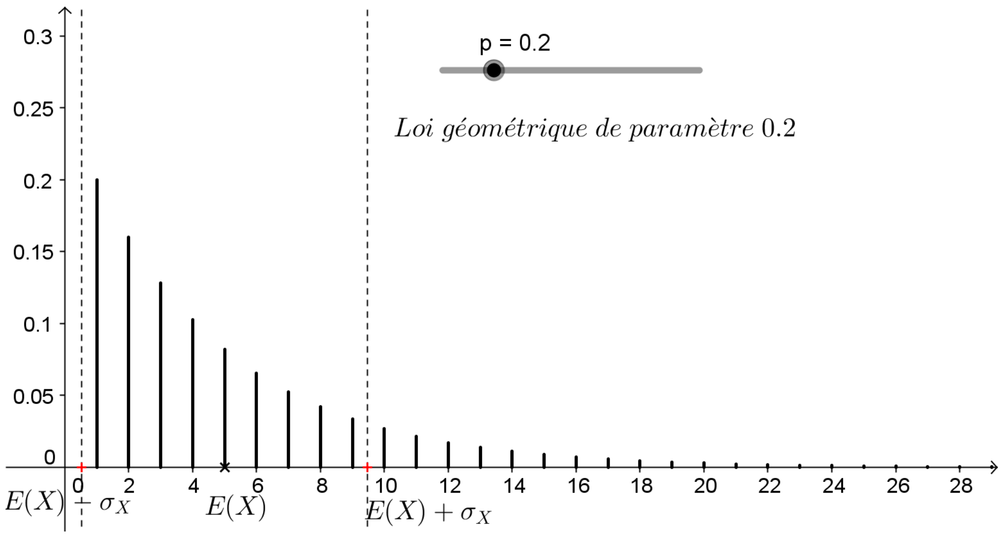
Ainsi $\forall x \in]-1;1[$, la série $\sum_{k \geqslant 1} k^2 x^{k-1}$ est convergente et sa somme est égale à :

$\sum_{k=1}^{+\infty} k^2 x^{k-1} = \dots$

Soit $p \in]0;1[$ alors $0 < 1-p < 1$ donc la série $\sum_{k \geqslant 1} k^2 p(1-p)^{k-1}$ est convergente et sa somme est égale à :

$\mathrm{E}(\mathrm{X}^2) = \sum_{k=1}^{+\infty} k^2 p(1-p)^{k-1} = \dots$

Donc $\mathrm{V}(\mathrm{X}) = \dots$ □



Simulation de la loi géométrique à l'aide du module numpy :

```python
import numpy as np
print(np.random.geometric(0.2))
print(np.random.geometric(0.2,10))
```

Définition de la loi de Poisson (1781-1840) de paramètre λ

Soit un réel λ strictement positif et X une variable aléatoire.

X suit la loi de Poisson de paramètre λ notée $\mathscr{P}(\lambda)$ si et seulement si $\begin{cases} \mathrm{X}(\Omega) = \mathbb{N} \\ \forall k \in \mathbb{N},\ \mathrm{P}(\mathrm{X} = k) = \mathrm{e}^{-\lambda} \dfrac{\lambda^k}{k!} \end{cases}$

On note alors $\mathrm{X} \hookrightarrow \mathscr{P}(\lambda)$.

Démonstration : démontrons que la loi de Poisson est une loi de probabilité.

Soit un réel $\lambda > 0$, $\forall k \in \mathbb{N}$, $\left.\begin{matrix} \mathrm{e}^{-\lambda} > 0 \\ \dfrac{\lambda^k}{k!} > 0 \end{matrix}\right\}$ donc $\mathrm{e}^{-\lambda} \dfrac{\lambda^k}{k!} > 0$

La série $\sum_{k \geqslant 0} \dfrac{\lambda^k}{k!}$ est convergente car...

Ainsi, par linéarité, $\sum_{k=0}^{+\infty} \mathrm{e}^{-\lambda} \dfrac{\lambda^k}{k!} = \dots$ □

Champs d'application de la loi de Poisson

Sur un laps de temps fixé, un phénomène aléatoire, pour lequel le futur est indépendant du passé, se produit en moyenne λ fois. Soit X la variable aléatoire égale au nombre de réalisation du phénomène sur ce laps de temps. La variable aléatoire suit alors la loi $\mathscr{P}(\lambda)$.

Espérance et variance d'une variable aléatoire suivant une loi de Poisson

Soit un réel λ et X une variable aléatoire suivant la loi $\mathscr{P}(\lambda)$.

$$\mathrm{E}(\mathrm{X})=\lambda$$
$$\mathrm{V}(\mathrm{X})=\lambda$$

Démonstration : soit $\lambda\in\mathbb{R}$

$\forall k\in\mathbb{N}^*,\ k\,\mathrm{e}^{-\lambda}\dfrac{\lambda^k}{k!}=\lambda\,\mathrm{e}^{-\lambda}\dfrac{\lambda^{k-1}}{(k-1)!}$ donc $\forall\lambda\in\mathbb{R}$, la série $\sum\limits_{k\geqslant1} k\,\mathrm{e}^{-\lambda}\dfrac{\lambda^k}{k!}$ est convergente et

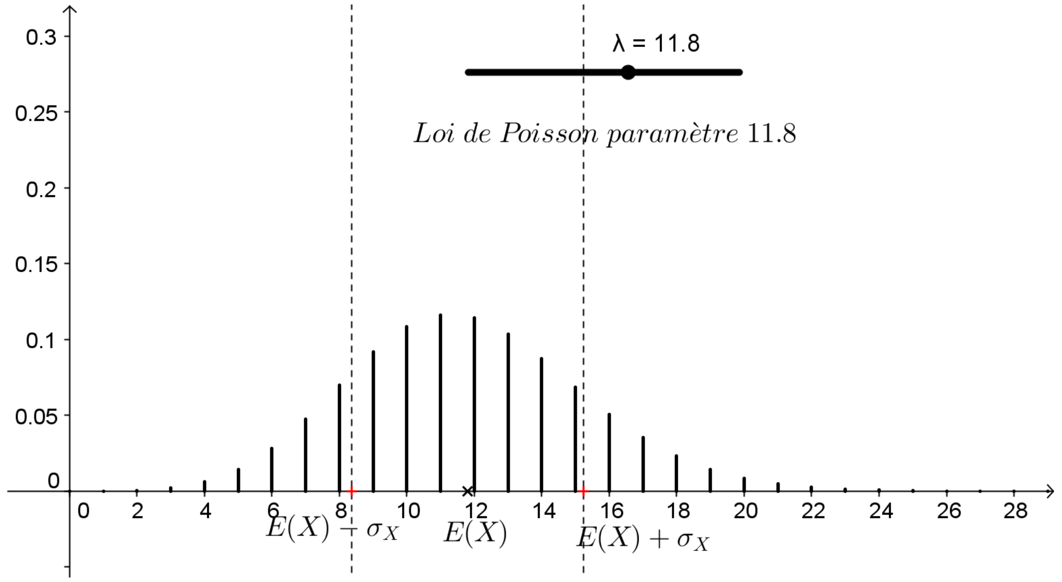
$\mathrm{E}(\mathrm{X})=\sum\limits_{k=1}^{+\infty} k\,\mathrm{e}^{-\lambda}\dfrac{\lambda^k}{k!}=\ldots$

$\forall k\in\mathbb{N}$, $k^2=k(k-1)+k$ ainsi pour tout entier naturel $k\geqslant2$

$k^2\,\mathrm{e}^{-\lambda}\dfrac{\lambda^k}{k!}=(k(k-1)+k)\,\mathrm{e}^{-\lambda}\dfrac{\lambda^k}{k!}=\mathrm{e}^{-\lambda}\dfrac{\lambda^k}{(k-2)!}+\mathrm{e}^{-\lambda}\dfrac{\lambda^k}{(k-1)!}=\ldots$

Ainsi $\mathrm{E}(\mathrm{X}^2)=\ldots$

Donc $\mathrm{V}(\mathrm{X})=\ldots$ □



Simulation de la loi de Poisson à l'aide du module numpy :

```python
import numpy as np
print(np.random.poisson(5))
print(np.random.poisson(5,10))
```

Approximation de la loi binomiale par la loi de Poisson

Soit $(\mathrm{X}_n)_{n\in\mathbb{N}}$ une suite de variables aléatoires.

$$\text{Si } \left.\begin{array}{r}\forall n\in\mathbb{N},\ \mathrm{X}_n\hookrightarrow\mathscr{B}(n,p_n)\\ \lim\limits_{n\to+\infty} n\,p_n=\lambda>0\end{array}\right\}\ \text{alors } \forall k\in\mathbb{N},\ \lim_{n\to+\infty}\mathrm{P}(\mathrm{X}_n=k)=\mathrm{e}^{-\lambda}\frac{\lambda^k}{k!}$$

Démonstration hors programme mais instructive :

Soit $k\in[\![0;n]\!]$, $\mathrm{P}(\mathrm{X}_n=k)=\dbinom{n}{k}(p_n)^k(1-p_n)^{n-k}=\dfrac{n!}{k!(n-k)!}(p_n)^k(1-p_n)^{n-k}=\dfrac{\overbrace{n(n-1)\ldots(n-k+1)}^{k\text{ facteurs}}}{k!}(p_n)^k(1-p_n)^{n-k}$

$$\mathrm{P}(\mathrm{X}_n=k)=\frac{n^k\left(1-\frac{1}{n}\right)\left(1-\frac{2}{n}\right)\dots\left(1-\frac{k-1}{n}\right)}{k!}(p_n)^k(1-p_n)^{n-k}=\frac{(n\,p_n)^k}{k!}\left(1-\frac{1}{n}\right)\left(1-\frac{2}{n}\right)\dots\left(1-\frac{k-1}{n}\right)\mathrm{e}^{(n-k)\ln(1-p_n)}$$

Si $\lim\limits_{n\to+\infty} n\,p_n=\lambda$ alors par produit de k limites on a : $\lim\limits_{n\to\infty}(n\,p_n)^k=\lambda^k$

Par ailleurs, pour tout entier $j\in[\![1\,;k-1]\!]$, $\lim\limits_{n\to+\infty}\frac{j}{n}=0$ donc $\lim\limits_{n\to+\infty}\left(1-\frac{j}{n}\right)=1$

Enfin, si $\lim\limits_{n\to+\infty} n\,p_n=\lambda$ alors $p_n \underset{n\to+\infty}{\sim} \frac{\lambda}{n}$ c'est-à-dire $p_n \underset{n\to+\infty}{=} \frac{\lambda}{n}+o\left(\frac{1}{n}\right)$ donc $\ln(1-p_n) \underset{n\to+\infty}{=} -\frac{\lambda}{n}+o\left(\frac{1}{n}\right)$

Donc $(n-k)\ln(1-p_n) \underset{n\to+\infty}{=} -\lambda+\frac{k\lambda}{n}+o(1) \underset{n\to+\infty}{\to} -\lambda$

Ainsi $\lim\limits_{n\to+\infty}\mathrm{P}(\mathrm{X}_n=k)=\frac{\lambda^k}{k!}\mathrm{e}^{-\lambda}$ □

Remarque : cette propriété, aussi appelée loi des événements rares, s'interprète dans le cas où n (pour n « grand ») expériences de Bernoulli indépendantes ont une probabilité faible de succès alors la somme du nombre de succès (X_n) suit une loi proche d'une loi de Poisson.

Dans la pratique cette approximation est utilisée pour $n\geqslant 50$ et ($n\,p<10$ ou $n(1-p)<10$)