20 V.A.R. réelles finies (partie 1)

Ce qu'il faut savoir faire à l'issue de la séquence...

- ◻ *Connaître et utiliser le vocabulaire et les notations des var*
- ◻ *Calculer l'espérance, la variance d'une var*
- ◻ *Connaître et utiliser le théorème de transfert.*
- ◻ *Connaître et utiliser les lois finies usuelles, leur eséprance et leur variance.*
- ◻ *Reconnaître une situation modélisée par une var de référence.*
- ◻ *Modéliser une situation par une variable aléatoire*

Variables aléatoires réelles (v.a.r)

Dans tout ce chapitre Ω est un univers **fini**.

DÉFINITIONS

Définition 1 ***(Variable aléatoire réelle)***

Soit Ω l'univers d'une expérience aléatoire.
On appelle variable aléatoire réelle définie sur $(\Omega, \mathcal{P}(\Omega))$ toute application

$$X : \begin{array}{ccc} \Omega & \longrightarrow & \mathbb{R} \\ \omega & \longmapsto & X(\omega) \end{array}$$

Si X est une variable aléatoire, l'ensemble des valeurs prises par X est noté :

$$X(\Omega) = \{X(\omega) \mid \omega \in \Omega\}$$

Si $\Omega = \{\omega_1, ..., \omega_n\}$ est un univers fini,

$$X(\Omega) = \{X(\omega_1), ..., X(\omega_n)\} = \{x_1, ..., x_2\}$$

où pour tout $i \in [\![1; n]\!]$, $x_i = X(\omega_i)$.

L'ensemble $X(\Omega)$ est parfois appelé support de X, ou univers image de X.

On note généralement les variables aléatoires avec une lettre majuscule comme X, Y, Z mais aussi R, S, T, U, V... ou toute autre lettre majuscule.
On trouve aussi $X_1, ..., X_n$ pour désigner n v.a.r ayant un « point » commun, ou étudier des suites de v.a.r ($\rightarrow$ 2è année)

Puisque Ω est un univers fini, $X(\Omega)$ est fini également.
On parle alors de **variable aléatoire finie**.

Exemples

Exemple n° 1 Une expérience aléatoire consiste à lancer n fois une pièce équilibrée. On note 1 pour face et 0 pour pile.
Ainsi $\Omega = \{0, 1\}^n$ est l'ensemble des n-listes d'éléments de $\{0; 1\}$.

- On pose X_1 la variable aléatoire qui à un élément de Ω associe 1 si le premier lancer donne face, 0 sinon. $X_1(\Omega) = \{0, 1\}$.
- On pose X_2 la variable aléatoire qui à un élément de Ω associe le nombre de face obtenu. $X_2(\Omega) = [\![0, n]\!]$.
- On pose X_3 la variable aléatoire qui à un élément de Ω associe le rang d'apparition du premier pile, s'il existe, et 0 sinon. $X_3(\Omega) = [\![0, n]\!]$.

Exemples

Exemple n° 2 On lance 2 dés équilibrés de couleurs différentes.

- On considère Y la variable aléatoire qui à un lancer associe la somme des résultats. $Y(\Omega) = [\![2, 12]\!]$.
- On considère X la variable aléatoire qui, à un lancer associe 2 si les deux dés prennent les valeurs 5 ou 6, 1 si l'un des deux dés seulement prend la valeur 5 ou 6 et -1 sinon. $X(\Omega) = \{-1, 1, 2\}$.

Propriété 1

Soit $\left| \begin{array}{l} X \text{ une variable aléatoire} \\ f : I \longrightarrow \mathbb{R} \text{ avec } X(\Omega) \subset I \subset \mathbb{R} \end{array} \right.$

Alors
$$f(X) : \begin{array}{ccc} \Omega & \longrightarrow & \mathbb{R} \\ \omega & \longmapsto & \Big[f(X)\Big](\omega) = f(X(\omega)) \end{array}$$
est une variable aléatoire sur Ω

NB : L'hypothèse $X(\Omega) \subset \mathcal{D}_f$ permet de s'assurer que tous les $f(X(\omega))$ ont un sens.

Exemple n° 3 On reprend la variable aléatoire X de l'exemple et on pose $Y = X^2$. Déterminer la loi de Y.

$X(\Omega) = \{-1, 1, 2\}$ donc $Y(\Omega) = \{1, 4\}$.
De plus,
$$P(Y = 1) = P(X^2 = 1) = P([X = -1] \cup [X = 1]).$$
Par incompatibilité des événements $[X = -1]$ et $[X = 1]$,
$$P(Y = 1) = P(X = 1) + P(X = -1) = \frac{4}{9} + \frac{4}{9} = \frac{8}{9}$$
et de même,
$$P(Y = 4) = P(X = 2) = \frac{1}{9}.$$

Exemple n° 4 On reprend les variables aléatoires de l'exemple 1 :
• X_1 est la v.a.r qui à un $\omega \in \Omega$ associe 1 si le premier lancer donne face, 0 sinon.
• soit Y la v.a.r qui, à $\omega \in \Omega$ associe 1 si le premier lancer donne face, -1 sinon.

On a
$$Y = 2X_1 - 1, \text{ et } X = \frac{1}{2}(Y + 1)$$

Appliquer une fonction à une v.a.r permet parfois de se ramener à une loi de référence, et d'en déduire sans calcul l'espérence, la variance... *cf infra*

V.A.R ET ÉVÉNEMENTS

On utilise des notations entre crochets pour désigner les événements associés à une variable aléatoire X. Par exemple,

- $[X = x]$ désigne l'événement « X prend la valeur x ».
- $[X \leqslant x]$ désigne l'événement « X prend une valeur inférieure à x ».
- $[a \leqslant X \leqslant b]$ désigne l'événement « X prend une valeur comprise entre a et b ».

On trouve aussi les notations $(X = k)$ ou $\{X = k\}$, ou encore $X^{-1}(\{k\})$ et même parfois $X = k$ sans aucun symbole autour.

autres exemples de notation d'événements

$[\exp(X) \leqslant x]$

- si $x \leqslant 0$, $[X \leqslant x] = \varnothing$ car une exponentielle est toujours positive.
- si $x > 0$, $[\exp(X) \leqslant x] = [X \leqslant \ln x]$ (car ln est strictement croissante)

$[\,|\,X\,| \leqslant x]$

- si $x \leqslant 0$, $[X \leqslant x] = \varnothing$ car une valeur absolue est toujours positive.
- si $x > 0$, $[\,|\,X\,| \leqslant x] = [-x \leqslant X \leqslant x]$

$[\,|\,X\,| \geqslant x]$

- si $x \geqslant 0$, $[|\,X\,| \geqslant x] = \Omega$ car une valeur absolue est toujours positive.
- si $x > 0$, $[\,|\,X\,| \geqslant x] = [X \leqslant -x] \cup [X \geqslant x]$

$\longrightarrow$ on pourra faire le lien avec les raisonnements par équivalence que l'on peut mener pour la résolution d'équation.

SYSTÈME COMPLET D'ÉVÉNEMENT ASSOCIÉ À UNE V.A.R

Propriété 2 ***(Système complet d'événements lié à une variable aléatoire)***

Soit X une variable aléatoire finie, alors $([X = x])_{x \in X(\Omega)}$ est un système complet d'événements. En particulier,

$$\sum_{x \in X(\Omega)} P(X = x) = 1.$$

On l'appelle <u>système complet d'événement associé à la variable aléatoire X</u>.

Preuve : Soit $x \in X(\Omega)$. Par définition d'une variable aléatoire réelle, $[X = x]$ est un événement de $(\Omega, \mathcal{P}(\Omega))$. Si x et y sont deux éléments distincts de $X(\Omega)$, $[X = x]$ et $[X = y]$ sont incompatibles, car la variable aléatoire X ne peut pas prendre deux valeurs différentes sur un même événement élémentaire. De plus,

$$\bigcup_{x \in X(\Omega)} [X = x] = [X \in X(\Omega)] = \Omega.$$

Donc $([X = x])_{x \in X(\Omega)}$ est un système complet d'événements.
L'égalité annoncée en découle par passage aux probabilités. $\square$

L'écriture $P(X = x)$ est généralement préférée à $P([X = x])$ pour alléger les notations.

Ce système complet d'événements est très souvent associé à l'utilisation de la formule des probabilités totales.
Puisque très souvent $X(\Omega) \subset \mathbb{N}$, on rencontrera le s.c.e $\{[X = k]; k \in [\![1; n]\!]\}$ et la formule des probabilités totales s'écrira dans ce cas particulier : si $A \subset \Omega$,

$$P(A) = \sum_{k=1}^{n} P(X = k)\mathbb{P}(A \mid X = k)$$

Définition 2 ***(Loi d'une variable aléatoire)***

Soit X une variable aléatoire finie, on appelle loi de X la donnée :
- de $X(\Omega)$
- des valeurs $P(X = x_k)$ pour x_k parcourant $X(\Omega)$.

Exemple n° 5 On lance une pièce équilibrée, et on définit la variable aléatoire Z comme valant 0 si la pièce donne face, et 1 sinon. Déterminer la loi de Z.

La loi de Z est : $Z(\Omega) = \{0, 1\}$, $P(Z = 1) = \frac{1}{2}$, $P(Z = 0) = \frac{1}{2}$.

k	0	1
$P(Z = k)$	$\frac{1}{2}$	$\frac{1}{2}$

Exemple n° 6 Déterminer la loi de la variable aléatoire X qui, à un lancer de deux dés associe 2 si les deux dés prennent les valeurs 5 ou 6, 1 si l'un des deux dés seulement prend la valeur 5 ou 6 et -1 sinon.

On a déjà déterminé que $X(\Omega) = \{-1, 1, 2\}$. L'indépendance des lancers permet ensuite les calculs des probabilités :

$$P(X=2) = \frac{2}{6} \times \frac{2}{6} = \boxed{\frac{1}{9}}, \quad P(X=-1) = \frac{4}{6} \times \frac{4}{6} = \boxed{\frac{4}{9}} \quad \text{et} \quad P(X=1) = 1 - \frac{1}{9} - \frac{4}{9} = \boxed{\frac{4}{9}}.$$

Pour le calcul de $P(X = 1)$, on pouvait aussi effectuer une disjonction de cas en fonction de si c'est le premier ou second lancer qui prend la valeur 5 ou 6, ce qui permet de retrouver la dernière valeur : $P(X = 1) = \frac{2}{6} \times \frac{4}{6} + \frac{4}{6} \times \frac{2}{6} = \frac{4}{9}$.

représentation graphique de la loi d'une v.a.r.

On représente la loi d'une variable aléatoire X par un diagramme en bâtons dont les abscisses sont les $x_i \in X(\Omega)$ et les ordonnées les $P(X = x_i)$.
Ainsi, en python, cela pourrait ressembler à :

```
import matplotlib.pyplot as plt
X = [0, 1, 2, 3]
Y = [0.1, 0.2, 0.5, 0.2]
plt.bar(X,Y, width=0.1)
plt.show()
```

name : /home/jm/loibarres.png
file : /home/jm/loibarres.png
state : unknown

Pour une v.a.r X telle que $X(\Omega)$ comporte un nombre limité (moins de 10) de valeurs, on peut représenter la loi de X par un tableau.

Après avoir déterminé la loi d'une var,
<u>toujours vérifier que la somme des probabilités vaut 1</u> !

Propriété 3

Soit $(p_k)_{k\in[\![1,n]\!]} \in \mathbb{R}^n$.

Si $\left|\begin{array}{l}\text{pour tout } k \in [\![1,n]\!],\ p_k \geqslant 0\\ \sum_{k=1}^{n} p_k = 1,\end{array}\right.$

alors $[\![1,n]\!]$ et $(p_k)_{k\in[\![1,n]\!]}$ forment la loi d'une variable aléatoire finie.

Preuve : On pose $\Omega = [\![1,n]\!]$, $P(\{k\}) = p_k$, et X la fonction identité. Il est alors direct que P engendre une probabilité, et que X est une variable aléatoire. $\square$

Notons qu'**on ne s'intéresse ici qu'aux probabilités.** (dont la somme vaut 1) et on n'a pas besoin de décrire l'expérience aléatoire ni même l'univers.
On peut choisir $X(\Omega) = [\![1;n]\!]$ ou n'importe quel $X(\Omega) = \{x_1, ..., x_n\}$ tant que les x_i sont deux à deux distincts. On a alors
$P(X = k) = p_k$ ou $P(X = x_k) = p_k$ selon le choix.

Pour avoir $X(\Omega) = [\![1;n]\!]$, on n'a pas nécessairement $\mathrm{Card}(\Omega) = n$.
Par contre il est nécessaire que $\mathrm{Card}(\Omega) \geqslant n$

<u>*Exemple*</u> *n°7* Soit $n \in \mathbb{N}$ et $(p_i)_{1\leqslant i\leqslant 2n+1}$ la suite définie par : pour tout $i \in [\![1;2n+1]\!]$:

$$p_i = \begin{cases} \dfrac{i}{(n+1)^2} & \text{si } i \leqslant n+1\\ \dfrac{2n+2-i}{(n+1)^2} & \text{sinon}\end{cases}$$

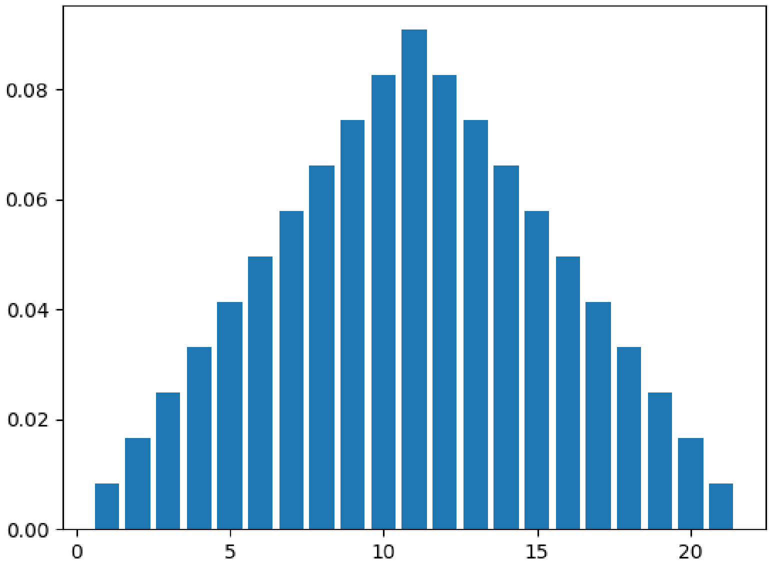
Montrer que l'on définit ainsi une loi de probabilité, la représenter graphiquement.

- On a bien $p_i > 0$ pour tout $i \in [\![1;2n+1]\!]$.

$$\begin{aligned}\sum_{i=1}^{2n+1} p_i &= \sum_{i=1}^{n+1} p_i + \sum_{i=n+2}^{2n+1} p_i\\ &= \sum_{i=1}^{n+1} \frac{i}{(n+1)^2} + \sum_{i=n+2}^{2n+1} \frac{2n+2-i}{(n+1)^2}\\ &= \frac{1}{(n+1)^2}\left(\sum_{i=1}^{n+1} i + \sum_{i=1}^{n} j\right) \text{ avec le changement d'indice } j = 2n+1-i\\ &= \frac{1}{(n+1)^2}\left(\frac{(n+1)(n+2)}{2} + \frac{n(n+1)}{2}\right)\\ &= \frac{1}{(n+1)^2} \times \frac{(n+1)(2n+2)}{2}\\ &= 1\end{aligned}$$

Avec `pyplot`, on peut obtenir la loi avec le code suivant :

```python
import matplotlib.pyplot as plt
n = 10
X = [k for k in range(1,2*n+2)]
Y = [i/(n+1)**2 for i in range(1,n+2)]
+[(2*n+2-i)/(n+1)**2 for i in range(n+2,2*n+2)]
plt.bar(X,Y)
plt.show()
```



Exemple n° 8 Soient $\alpha \in \mathbb{R}$ et $n \in \mathbb{N}^*$.
On considère la n-liste $(u_1, \ldots, u_n)$ définie pour tout $i \in [\![1, n]\!]$ par $u_i = \alpha \times i$.
Cette n-liste définit-elle une loi d'une variable aléatoire réelle ?

Pour tout $i \in [\![1, n]\!]$, on a $u_i \geqslant 0$ si et seulement si $\alpha > 0$. Par ailleurs, on a :

$$\sum_{i=1}^{n} u_i = \sum_{i=1}^{n} \alpha i = \alpha \sum_{i=1}^{n} i = \alpha \frac{n(n+1)}{2},$$

ce qui est égal à 1 ssi $\alpha = \frac{2}{n(n+1)}$. Cette valeur étant positive, la liste définit une loi de variable aléatoire finie si et seulement si $\alpha = \frac{2}{n(n+1)}$.

Fonction de répartition

Définition 3 ***(Fonction de répartition)***

Soit X une variable aléatoire finie, on appelle fonction de répartition de X la fonction F_X définie pour tout $t \in \mathbb{R}$ par

$$F_X(t) = P(X \leqslant t).$$

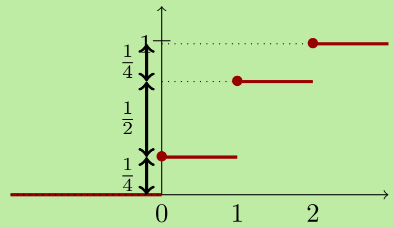
Exemple n° 9 On lance deux fois une pièce équilibrée. Soit X la variable aléatoire qui vaut 0 si on obtient deux fois pile, 2 si on obtient deux fois face, et 1 sinon.
Quelle est la loi de X ? Tracer sa fonction de répartition.

$X(\Omega) = \{0, 1, 2\}$

On a : $P(X = 0) = P(X = 2) = \frac{1}{4}$, et

$P(X = 1) = 1 - \frac{1}{4} - \frac{1}{4} = \frac{1}{2}$.

D'où le tracé :



LIEN ENTRE LOI ET FONCTION DE RÉPARTITION

Propriété 4 ***(La fonction de répartition caractérise la loi)***

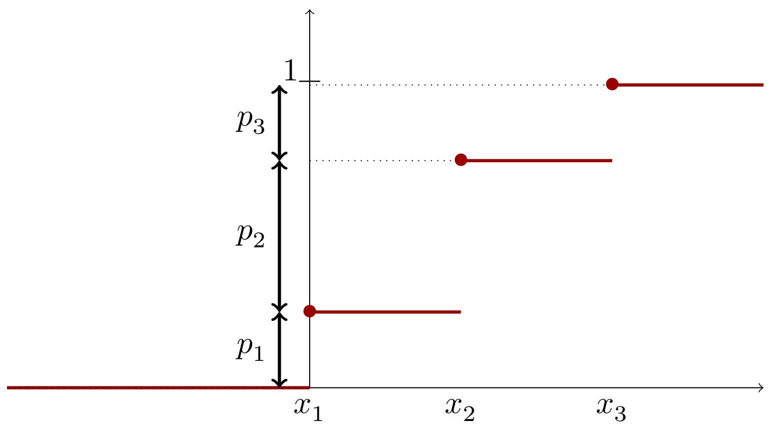
Soit X une variable aléatoire finie.
On peut supposer $X(\Omega) = \{x_1, ...x_n\}$ avec $x_1 < x_2 < ... < x_n$.

Alors
$$\forall t \in \mathbb{R}, \quad F_X(t) = \sum_{\substack{x \in X(\Omega) \\ x \leqslant t}} P(X = x) = \sum_{\substack{1 \leqslant k \leqslant n \\ x_k \leqslant t}} P(X = x_k)$$
Autrement dit, F est une fonction en escalier :

- Elle est nulle sur $]-\infty; x_1[$,
- constante sur chaque intervalle de la forme $[x_k; x_{k+1}[$ (pour $k \in [\![1; n-1]\!]$)
- égale à 1 sur $[x_n; +\infty[$.

et réciproquement, on déduire la loi de X de sa fonction de répartition :

$$P(X = x_1) = F_X(x_1)$$
$$\text{et } \forall k \in [\![2, n]\!], \quad P(X = x_k) = F_X(x_k) - F_X(x_{k-1}).$$



Preuve : Pour tout $x \in \mathbb{R}$, $[X \leqslant x] = \bigcup_{\substack{x_k \in X(\Omega) \\ x_k \leqslant x}} [X = x_k]$.

Or les événements $[X = x_k]$ sont incompatibles deux à deux. Par additivité de la probabilité, on obtient
$$\forall x \in \mathbb{R}, \quad F_X(x) = \sum_{\substack{x_k \in X(\Omega) \\ x_k \leqslant x}} P(X = x_k)$$

D'où la première relation. Si $X(\Omega) = \{x_1, \ldots x_n\}$, le cas $x = x_1$ nous donne alors directement $P(X = x_1) = F_X(x_1)$. Et si $k \in [\![2, n]\!]$,

$$F_X(x_k) = \sum_{i=1}^{k} P(X = x_i) = P(X = x_k) + F_X(x_{k-1}).$$

D'où la dernière relation. □

Pour une v.a.r à valeur dans $\mathbb{Z}$, la fonction de répartition permet de retrouver la loi !
Pour tout $k \in \mathbb{Z}$,

$$P(X = k) = P(X \leqslant k) - P(X < k) = P(X \leqslant k) - P(X \leqslant k - 1)$$

On utilise aussi la fonction d'« anti-répartition » :

$$P(X = k) = P(X \geqslant k) - P(X > k) = P(X \geqslant k) - P(X \geqslant k + 1)$$

Ces deux méthodes sont fondamentales pour tous les problèmes de min/max (cf exemple)

Exemple n° 10 Une urne contient p boules numérotées de 1 à p. On tire n boules avec remise, et on note X le plus grand des numéros tirés. Trouver la loi de X.

On remarque que $X(\Omega) = [\![1, p]\!]$: l'inclusion $X(\Omega) \subset [\![1, p]\!]$ est immédiate car X est un numéro de boule, et si $k \in [\![1, p]\!]$, on réalise $[X = k]$ par exemple en tirant $n - 1$ fois 1 puis k, d'où l'égalité des ensembles.

Si $n = 1$, on sait que X est telle que $\forall k \in [\![1, p]\!]$, $P(X = k) = \frac{1}{p}$. Le problème est de « passer au maximum » quand on a plusieurs tirages. Pour cela, on va utiliser les fonctions de répartition.

On pose X_l la variable aléatoire correspondant au numéro du l-ième tirage.
On a alors $X = \max\limits_{l \in [\![1, n]\!]} X_l$, et pour tout $k \in [\![1, p]\!]$, la définition du max et l'indépendance mutuelle des événements $[X_l \leqslant t]$ donnent :

$$F_X(k) = P(X \leqslant k) = P\left(\max_{l \in [\![1, n]\!]} X_l \leqslant k\right) = P\left(\bigcap_{l=1}^{n} [X_l \leqslant k]\right) = \prod_{l=1}^{n} P(X_l \leqslant k) = \prod_{l=1}^{n} F_{X_l}(k).$$

Or $F_{X_l}(k) = P(X_l \leqslant k) = \frac{k}{p}$ par équiprobabilité.
Donc $\forall k \in [\![1, p]\!]$, $\quad F_X(k) = \left(\frac{k}{p}\right)^n$.

La proposition précédente donne alors : $\forall k \in [\![2, p]\!]$,

$$P(X = k) = F_X(k) - F_X(k - 1) = \left(\frac{k}{p}\right)^n - \left(\frac{k-1}{p}\right)^n$$

$$\text{et} \quad P(X = 1) = F_X(1) = \left(\frac{1}{p}\right)^n.$$

Remarque : on vérifie bien par télescopage que la somme des valeurs vaut 1.

La fonction de répartition est l'équivalent probabiliste de la courbe des fréquences cumulées croissantes. Elle croît de 0 à 1.

Moments

Espérance

Définition 4 ***(Espérance)***

Soit X une variable aléatoire finie telle que $X(\Omega) = \{x_1, ..., x_n\}$.
On appelle **espérance** de X le nombre réel

$$\mathbf{E}(X) = \sum_{k=1}^{n} x_k P(X = x_k)\,.$$

Dans le cas où $X(\Omega) = [\![1, n]\!]$, alors $\mathbf{E}(X) = \sum_{k=1}^{n} kP(X = k)$.

Point rédaction

Quand la variable aléatoire est **finie**, la somme est **finie** (donc bien définie).
En deuxième année, vous verrez des variables aléatoires non finies, il faudra justifier l'existence de l'espérance (qui n'existe pas toujours). Alors prenons déjà les bonnes habitudes en rédigeant :

« X admet une espérance car $X(\Omega)$ est fini »

L'espérance est la **moyenne** des valeurs prises par une variable aléatoire, pondérées par leurs probabilités respectives.

Mais l'espérance peut ne pas être une valeur de $X(\Omega)$. Elle peut même ne pas être proche d'une valeur probable... On ne peut donc espérer obtenir cette valeur.

Par contre, si on répète un grand nombre de fois l'expérience, la moyenne des valeurs obtenues pour X sera proche de $E(X)$, d'autant plus que le nombre d'expériences est grand.

L'étude de cette convergence, des écarts... feront surtout l'objet du programme de deuxième année.

Théorème de transfert

Propriété 5 ***(Théorème de transfert)***

Soit $\left|\begin{array}{l} X \text{ une variable aléatoire finie} \\ g \text{ une fonction réelle telle que } X(\Omega) = \{x_1, ..., x_n\} \subset D_g \end{array}\right.$

on pose $Y = g(X)$. Alors :

$$\mathbf{E}(Y) = \mathbf{E}(g(X)) = \sum_{k=1}^{n} g(x_k)P(X = x_k).$$

Propriété 6 ***(Propriétés de l'espérance)***

Soient a et b deux réels et X une variable aléatoire finie.

- Si $X(\Omega) \subset [a, b]$, alors $\mathbf{E}(X) \in [a, b]$.
  En particulier, si X est à valeurs positives, $\mathbf{E}(X) \geqslant 0$.
- linéarité de l'espérance :

$$\mathbf{E}(aX + b) = a\mathbf{E}(X) + b$$

Preuve :

- Supposons $X(\Omega) = \{x_1, ..., x_n\} \subset [a, b]$. Alors
$$\forall k \in [\![1; n]\!], a \leqslant x_k \leqslant b$$
En multipliant par $P(X = x_k) \geqslant 0$, puis en sommant terme à terme, on obtient :
$$\sum_{k=1}^{n} aP(X = x_k) \leqslant \sum_{k=1}^{n} x_k P(X = x_k) \leqslant \sum_{k=1}^{n} bP(X = x_k),$$
et comme $\sum_{k=1}^{n} P(X = x_k) = 1$, le calcul des trois sommes donne bien $a \leqslant \mathbf{E}(X) \leqslant b$.

- Par théorème de transfert puis linéarité de la somme,
$$\begin{aligned}\mathbf{E}(aX + b) &= \sum_{k=1}^{n} (ax_k + b)P(X = x_k) \\ &= a\sum_{k=1}^{n} x_k P(X = x_k) + b\sum_{k=1}^{n} P(X = x_k) \\ &= a\mathbf{E}(X) + b.\end{aligned}$$

La linéarité de l'espérance reste vraie dans un cadre plus large.
Si X et Y sont deux variables aléatoires finies, on a également
$$\mathbf{E}(X + Y) = \mathbf{E}(X) + \mathbf{E}(Y).$$
Ce résultat sera vu plus tard.

Moments d'une variable aléatoire

Définition 5 ***(Moment)***

Soit X une variable aléatoire finie telle que $X(\Omega) = \{x_1, ..., x_n\}$
$r \in \mathbb{N}^*$.

On appelle **moment** d'ordre r de X le nombre réel :

$$\mathbf{E}(X^r) = \sum_{k=1}^{n} x_k^r P(X = x_k).$$

La deuxième égalité est une simple application du théorème de transfert à la variable aléatoire X et à la fonction définie sur $\mathbb{R}$ par $x \mapsto x^r$.

Définition 6 ***(Moment centré)***

Soit X une variable aléatoire finie telle que $X(\Omega) = \{x_1, ..., x_n\}$
$r \in \mathbb{N}^*$.

On appelle **moment centré** d'ordre r de X le nombre réel :

$$\mathbf{E}\Big((X - \mathbf{E}(X))^r\Big) = \sum_{k=1}^{n} (x_k - \mathbf{E}(X))^r P(X = x_k).$$

Variance

Définition 7 ***(Variance)***

Soit X une variable aléatoire finie.
On appelle **variance** de X le moment centré d'ordre deux

$$\mathbf{V}(X) = \mathbf{E}((X - \mathbf{E}(X))^2)$$

et **écart-type** de X la quantité $\sigma(X) = \sqrt{\mathbf{V}(X)}$.

Comme pour l'espérance, prenez déjà l'habitude de rédiger :

X est finie, donc admet une variance

En effet, en deuxième année, la variance n'existe pas toujours...

Propriété 7

Soient a et b deux réels et X une variable aléatoire finie, alors

$$\mathbf{V}(aX + b) = a^2\mathbf{V}(X).$$

Preuve : Par linéarité de l'espérance,

$$\begin{aligned} \mathbf{V}(aX+b) &= \mathbf{E}((aX+b-\mathbf{E}(aX+b))^2) \\ &= \mathbf{E}((aX+b-a\mathbf{E}(X)-b)^2) \\ &= a^2\mathbf{E}\Big((X-\mathbf{E}(X))^2\Big) \\ &= a^2\mathbf{V}(X) \end{aligned}$$

Propriété 8 ***(cas de la variance nulle)***

Soit X une variable aléatoire finie. Alors

$V(X) = 0 \quad \Longleftrightarrow \quad X$ est une variable aléatoire certaine égale à $E(X)$

$\rightarrow$ ces notions seront précisées dans la partie sur les variables aléatoires usuelles.

Propriété 9 ***(Formule de Kœnig-Huygens)***

Soit X une variable aléatoire finie. Alors :

$$\mathbf{V}(X) = \mathbf{E}(X^2) - \mathbf{E}(X)^2.$$

Preuve : (D.A.C)

Par linéarité de l'espérance,
$$\begin{aligned}\mathbf{V}(X) &= \mathbf{E}((X-\mathbf{E}(X))^2)\\ &= \mathbf{E}(X^2+\mathbf{E}(X)^2-2\mathbf{E}(X)X)\\ &= \mathbf{E}(X^2)+\mathbf{E}(X)^2-(2\mathbf{E}(X))\mathbf{E}(X)\\ &= \mathbf{E}(X^2)+\mathbf{E}(X)^2-2\mathbf{E}(X)^2\\ &= \mathbf{E}(X^2)-\mathbf{E}(X)^2\end{aligned}$$

Remarque : on a utilisé la version de linéarité de l'espérance pour une somme de deux variables aléatoires. □

On utilise en pratique presque toujours la formule de Kœnig Huygens pour calculer une variance.

Variable aléatoire centrée réduite

Définition 8

Soit X une variable aléatoire admettant une espérance.

- Si $E(X) = 0$, on dit que X est centrée.
- Dans le cas général, $Y = X - E(X)$ est est la variable centrée associée à X.

Définition 9 ***(Variable aléatoire centrée réduite)***

Soit X une variable aléatoire finie de variance non nulle. On appelle <u>variable aléatoire centrée réduite associée</u> à X la variable notée X^* définie par :

$$X^* = \frac{X - \mathbf{E}(X)}{\sqrt{\mathbf{V}(X)}}.$$

Comme leur nom le suggère, les variables centrées réduites ont une espérance nulle (centrée) et une variance de 1 (réduite).

En deuxième année, elles seront mises en lien avec la loi normale centrée réduite vue en Terminale.

Deux inégalités classiques

INÉGALITÉ DE MARKOV

Propriété 10 ***(Inégalité de Markov)***

Soit X une variable aléatoire <u>positive</u> finie. X admet donc une espérance
$\lambda \in \mathbb{R}_+^*$:

$$P(X \geqslant \lambda) \leqslant \frac{\mathbf{E}(X)}{\lambda}.$$

Preuve : Comme X est une v.a.r. finie, $X(\Omega) = \{x_i \mid 1 \leqslant i \leqslant n\}$. Alors :

$$\begin{aligned}
\mathbf{E}(X) &= \sum_{1\leqslant i\leqslant n} x_i P(X = x_i)\\
&\geqslant \sum_{\substack{1\leqslant i\leqslant n\\ x_i\geqslant\lambda}} x_i P(X = x_i) \quad \text{car les } x_i \text{ sont positifs}\\
&\geqslant \sum_{\substack{1\leqslant i\leqslant n\\ x_i\geqslant\lambda}} \lambda P(X = x_i)\\
&\geqslant \lambda \sum_{\substack{1\leqslant i\leqslant n\\ x_i\geqslant\lambda}} P(X = x_i)\\
&\geqslant \lambda P(X \geqslant \lambda)
\end{aligned}$$

En divisant par $\lambda > 0$, on obtient le résultat souhaité.

et en français...

l'inégalité de Markov dit que la probabilité que X prenne des valeurs plus grande que λ est contrôlée par $E(X)$ et λ.
$\rightarrow$ L'espérance est une mesure de « position ». Il est donc logique que plus l'espérance de X est grande, plus il est *possible* qu'elle prenne de grande valeurs.
$\rightarrow$ λ est la valeur à dépasser. Plus λ est grand, moins il est probable que X dépasse λ (cf division).

INÉGALITÉ DE BIENAYMÉ-TCHEBYCHEV

Propriété 11 ***(Inégalité de Bienaymé-Tchebychev)***

Soit X une variable aléatoire admettant un moment d'ordre 2 et $\varepsilon > 0$:

$$P(|\,X - \mathbf{E}(X)\,| \geqslant \varepsilon) \leqslant \frac{\mathbf{V}(X)}{\varepsilon^2}.$$

Preuve : $\varepsilon > 0$ et $x \mapsto \sqrt{x}$ est une bijection croissante de $\mathbb{R}_+$ dans $\mathbb{R}_+$ donc

$$\Big[\,|\,X - \mathbf{E}(X)\,| \geqslant \varepsilon\Big] = \Big[(X - \mathbf{E}(X))^2 \geqslant \varepsilon^2\Big]$$

La v.a.r. discrète $(X - \mathbf{E}(X))^2$ est **positive** et admet une **espérance** (car X admet une variance).
On peut donc lui appliquer l'inégalité de **Markov** avec $\lambda = \varepsilon^2 > 0$,

$$P(|\,X - \mathbf{E}(X)\,| \geqslant \varepsilon) = P((X - \mathbf{E}(X))^2 \geqslant \varepsilon^2) \leqslant \frac{\mathbf{E}((X - \mathbf{E}(X))^2)}{\varepsilon^2} = \frac{\mathbf{V}(X)}{\varepsilon^2}.$$

d'où le résultat. □

et en français...

l'inégalité de Bienaymé-Tchebychev dit que la probabilité que X s'écarte de la moyenne de plus de ε est contrôlée par $V(X)$ et ε.
$\rightarrow$ L'espérance est une moyenne. On s'attend à ce que les valeurs prises par X se répartissent autour (inférieures ou supérieures).
$\rightarrow$ La variance est une mesure de dispersion.
Plus la variance est faible, plus la probabilité de s'écarter de la moyenne diminue.
$\rightarrow$ Quant au ε, qui représente l'écart minimum à la moyenne, plus il est grand, plus la probabilité de s'écarter de plus de ε diminue. Le carré est à mettre en relation avec la variance.

Retenir aussi que l'inégalité de Bienaymé Tchebycheff parle de X « centrée ». En particulier, elle servira souvent pour des variables centrées réduites.

Exemple n° 11 X est une variable aléatoire d'espérance μ et d'écrat-type σ.

1. Justifier que pour tout entier naturel $k \geqslant 1$, $P(|X - \mu| \geqslant k\sigma) \leqslant \frac{1}{k^2}$
2. a. Déterminer une valeur de k pour laquelle $P(|X - \mu| \geqslant k\sigma) \leqslant 0,1$

   b. Interpréter la valeur obtenue.

♡♡ Lois finies usuelles ♡♡

EXTRAIT DU PROGRAMME :
Les étudiants doivent savoir reconnaitre les situations classiques de modélisation par des lois uniformes, de Bernoulli, hypergéométrique et binomiale. On fait le lien entre la loi de Bernoulli et les variables indicatrices.

J'ajouterais qu'il faut aussi reconnaître la loi par les valeurs obtenues dans un calcul !

VARIABLE ALÉATOIRE CERTAINE

Définition 10 ***(Variable aléatoire certaine)***

Une variable aléatoire finie X est <u>certaine</u> si elle ne prend qu'une seule valeur a. On a alors

$$P(X=a)=1.$$

Propriété 12 ***(Espérance et variance d'une variable aléatoire certaine)***

Si X est une variable aléatoire certaine telle que $X(\Omega)=\{a\}$, alors

$$\mathbf{E}(X)=a \quad \text{et} \quad \mathbf{V}(X)=0$$

Preuve : On trouve par calcul direct :

$$\mathbf{E}(X)=aP(X=a)=a,$$

puis par théorème de transfert appliqué à la fonction $x\mapsto(x-a)^2$:

$$\mathbf{V}(X)=(a-a)^2P(X=a)=0^2\times 1=0.$$

□

Propriété 13 ***(Variable aléatoire de variance nulle)***

Si X est une variable aléatoire finie. Alors X admet une variance et

$$V(X)=0 \iff \exists a\in\mathbb{R}, P(X=a)=1$$

Preuve : On suppose que $\mathbf{V}(X)=0$. Par définition de la variance, $(X-\mathbf{E}(X))^2$ est alors une variable aléatoire finie d'espérance nulle. Si $X(\Omega)=\{x_1,\dots,x_n\}$, on a de plus par le théorème de transfert :

$$\sum_{k=1}^{n}(x_k-\mathbf{E}(X))^2P(X=x_k)=0.$$

Or pour tout $k\in[\![1,n]\!]$, $(x_k-\mathbf{E}(X))^2P(X=x_k)\geqslant 0$, donc les termes de la somme sont tous nuls.
Donc pour tout $k\in[\![1,n]\!]$, soit $(x_k-\mathbf{E}(X))^2=0$ (càd $x_k=\mathbf{E}(X)$), soit $P(X=x_k)=0$.
Comme $\sum_{k=1}^{n}P(X=x_k)=1$, il existe au moins un k_0 tel que $P\left(X=x_{k_0}\right)\neq 0$.
Alors $x_{k_0}=\mathbf{E}(X)$. Or $\mathbf{E}(X)$ est une constante : $\forall k\neq k_0, x_k\neq\mathbf{E}(X)$, cela signifie que $(x_k-\mathbf{E}(X))^2\neq 0$ et donc $P(X=x_k)=0$. On en déduit bien :

$$P(X=x_{k_0})=1.$$

□

Loi de Bernoulli

Définition 11 ***(Loi de Bernoulli)***

Une variable aléatoire X suit la <u>loi de Bernoulli</u> de paramètre $p \in]0;1[$ lorsque $X(\Omega) = \{0,1\}$ et

$$P(X=1) = p.$$

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

On a alors $P(X=0) = 1-p$.

Propriété 14 ***(une observation bien pratique)***

$$X \hookrightarrow \mathcal{B}(p) \iff (1-X) \hookrightarrow \mathcal{B}(1-p)$$

Preuve : En remarquant que $1-(1-X) = X$, il suffit de prouver le sens direct.

Notons $Y = 1-X$. $X(\Omega) = \{0,1\}$ donc $Y(\Omega) = \{0,1\}$ car si $X=1$, $Y=0$ et si $X=0$, $Y=1$.
De plus, $P(Y=1) = P(X=0) = 1-p$ donc $Y \hookrightarrow \mathcal{B}(1-p)$.

Propriété 15 ***(Espérance et variance d'une variable aléatoire de Bernoulli)***

Si X est une variable aléatoire telle que $X \hookrightarrow \mathcal{B}(p)$, alors $\mathbf{E}(X) = p$ et $\mathbf{V}(X) = p(1-p)$.

Preuve : On trouve par calcul : $\mathbf{E}(X) = 1P(X=1) + 0P(X=0) = p$

$$\begin{aligned}\mathbf{V}(X) &= (1-E(X))^2P(X=1) + (0-\mathbf{E}(X))^2P(X=0)\\ &= (1-p)^2p + p^2(1-p)\\ &= p(1-p)(1-p+p) = p(1-p)\end{aligned}$$

<u>*Exemple*</u> *n° 12* On suppose que $X \hookrightarrow \mathcal{B}(p)$.
Quelle est la loi de la variable aléatoire X^2 ?

X^2 ne peut prendre que les valeurs 0 et 1, et $P(X^2=1) = P(X=1) = p$ comme X est à valeurs positives. Donc $X^2 \hookrightarrow \mathcal{B}(p)$.

Définition 12 ***(Variable indicatrice d'un événement)***

Soit $(\Omega, \mathcal{P}(\Omega))$ un espace probabilisable, et A un événement de Ω. La variable aléatoire finie X telle que

$$\forall \omega \in \Omega, \quad X(\omega) = \begin{cases} 1 & \text{si } \omega \in A \\ 0 & \text{si } \omega \notin A \end{cases}$$

est appelée la <u>variable indicatrice de l'événement</u> A. On la note $\mathbb{1}_A$ ou X_A.

Propriété 16

Soit A un événement et $\mathbb{1}_A$ sa variable aléatoire indicatrice.
Alors $\mathbb{1}_A \hookrightarrow \mathcal{B}(p)$ avec $p = P(A)$

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

Exemple n° 13 On lance un dé. Soit A l'événement « tirer un numéro pair ».
La variable aléatoire X qui vaut 1 si on tire un numéro pair et 0 sinon est la variable indicatrice de l'événement A.

LOI BINOMIALE

Définition 13 ***(Loi binomiale)***

Une variable aléatoire X suit la loi binomiale de paramètres $p \in [0,1]$ et $n \in \mathbb{N}^*$ lorsque $X(\Omega) = [\![0,n]\!]$ et $\forall k \in [\![0,n]\!]$,

$$P(X = k) = \binom{n}{k} p^k (1-p)^{n-k}.$$

On note alors $X \hookrightarrow \mathcal{B}(n,p)$.

Preuve : Il faut montrer que ces propriétés permettent bien de définir une variable aléatoire.

$\forall k \in [\![0,n]\!]$, $\binom{n}{k} p^k (1-p)^{n-k} \geqslant 0$ puisque $p \in [0,1]$.

De plus la formule du binôme de Newton donne :

$$\sum_{k=0}^{n} \binom{n}{k} p^k (1-p)^{n-k} = (p+1-p)^n = 1.$$

Il s'agit donc bien de la loi d'une variable aléatoire. □

Remarque : $X \hookrightarrow \mathcal{B}(1,p) \iff X \hookrightarrow \mathcal{B}(p)$.

MODÈLE : On réalise $\boxed{n}$ épreuves de Bernoulli indépendantes, de même paramètre $\boxed{p}$ et on compte le nombre de succès.

La v.a.r X comptant le nombre de succès suit une loi binômiale. $X \hookrightarrow \mathcal{B}(n;p)$

Il est immédiat que $X(\Omega) = [\![0,n]\!]$.

Soit $k \in X(\Omega)$, $[X = k]$ est l'événement « il y a exactement k succès ».
Il y a $\binom{n}{k}$ événements élémentaires qui renvoient k succès (on choisit la position des k succès dans les n épreuves), qui ont chacun une probabilité $p^k(1-p)^{n-k}$ par indépendance des épreuves.

D'où $P(X = k) = \binom{n}{k} p^k (1-p)^{n-k}$. On en déduit que $X \hookrightarrow \mathcal{B}(n,p)$. 3

Si $\binom{n}{k}$ compte les issues (les branches de l'arbre) pour réalisant k succès. Mais il n'y a pas équiprobabilité entre les différentes issues de l'univers si $p \neq \frac{1}{2}$!

Propriété 17 ***(Espérance et variance d'une variable aléatoire binomiale)***

Si X est une variable aléatoire telle que $X \hookrightarrow \mathcal{B}(n,p)$, alors
$$\mathbf{E}(X) = np \text{ et } \mathbf{V}(X) = np(1-p).$$

Preuve :

On trouve par calcul, pour $n \in \mathbb{N}^*$,

$$E(X) = \sum_{k=0}^{n} kP(X=k)$$
$$E(X) = 0 + \sum_{k=1}^{n} k\binom{n}{k}p^k(1-p)^{n-k}$$
$$= \sum_{k=1}^{n} k\frac{n}{k}\binom{n-1}{k-1}p^k(1-p)^{n-k}$$
$$E(X) = n\sum_{i=0}^{n-1}\binom{n-1}{i}p^{i+1}(1-p)^{n-1-i} \text{ en posant } i = k-1$$
$$E(X) = np\sum_{i=0}^{n-1}\binom{n-1}{i}p^{i}(1-p)^{n-1-i}$$
$$E(X) = np(p+(1-p))^n \quad \text{par formule du binôme de Newton}$$
$$E(X) = np$$

Pour la variance, il y a une astuce à retenir : $E(X(X-1))$ est simple à calculer avec la petite formule sans nom...

$$E(X(X-1)) = \sum_{k=0}^{n} k(k-1)P(X=k) \text{ avec le théorème de transfert}$$
$$= \sum_{k=2}^{n} k(k-1)\binom{n}{k}p^k(1-p)^{n-k}$$
$$= \sum_{k=2}^{n} (k-1)n\binom{n-1}{k-1}p^k(1-p)^{n-k} \text{ avec la petite formule sans nom}$$
$$= n(n-1)\sum_{k=2}^{n}\binom{n-2}{k-2}p^k(1-p)^{n-k} \text{ avec la petite formule sans nom}$$
$$= n(n-1)p^2\sum_{j=0}^{n-2}\binom{n-2}{j}p^j(1-p)^{n-2-j} \text{ avec le changement d'indice } j = k-2$$
$$= n(n-1)p^2(p+(1-p))^{n-2} \text{ avec la formule du binôme}$$
$$= n(n-1)p^2$$

Finalement

$$V(X) = E(X^2) - E(X)^2 \text{ avec la formule de Kœnig Huygens}$$
$$V(X) = E(X(X-1)) + E(X) - E(X)^2 \text{ par linéarité de l'espérance et car } X^2 = X(X-1) + X$$
$$= n(n-1)p^2 + np - (np)^2$$
$$= n^2p^2 - np^2 + np - n^2p^2$$
$$= np(1-p)$$

Une preuve plus simple sera vue plus tard et l'espérance/variance d'une somme de variables aléatoires indépendantes. $\square$

LOI UNIFORME SUR $[\![1,n]\!]$

Définition 14 ***(Loi uniforme)***

Soit $n \in \mathbb{N}^*$. Une variable aléatoire X suit la loi **uniforme** sur $[\![1,n]\!]$ lorsque $X(\Omega) = [\![1,n]\!]$ et $\forall k \in [\![1,n]\!]$,

$$P(X = k) = \frac{1}{n}.$$

On note alors $X \hookrightarrow \mathcal{U}([\![1,n]\!])$.

Preuve : Il faut montrer que ces propriétés permettent bien de définir une variable aléatoire.

- d'une part, $\frac{1}{n} \geqslant 0$, de plus :
- d'autre part, $\sum_{k=1}^{n} \frac{1}{n} = \frac{n}{n} = 1$.

Il s'agit donc bien de la loi d'une variable aléatoire. □

Propriété 18 ***(Espérance et variance d'une variable uniforme)***

Si X est une variable uniforme telle que $X \hookrightarrow \mathcal{U}([\![1,n]\!])$, alors
$$E(X) = \frac{n+1}{2} \quad \text{et} \quad V(X) = \frac{n^2-1}{12}.$$

Preuve : Il suffit de faire le calcul :

$$E(X) = \sum_{k=1}^{n} kP(X=k) = \sum_{k=1}^{n} k\frac{1}{n} = \frac{1}{n}\sum_{k=1}^{n} k = \frac{n+1}{2}.$$

On obtient ensuite avec la formule de Kœnig-Huygens :
$$\begin{aligned}
V(X) &= E(X^2) - E(X)^2 \\
&= \sum_{k=1}^{n} k^2 P(X=k) - E(X)^2 \\
&= \frac{1}{n}\frac{n(n+1)(2n+1)}{6} - \left(\frac{n+1}{2}\right)^2 \\
&= \frac{n+1}{2}\left(\frac{2n+1}{3} - \frac{n+1}{2}\right) \\
&= \frac{n+1}{2}\left(\frac{4n+2-3n-3}{6}\right) \\
&= \frac{n^2-1}{12}
\end{aligned}$$
□

On peut généraliser la définition de la loi uniforme au cas d'un intervalle $[\![a,b]\!]$ où $(a,b) \in \mathbb{Z}$.

Définition 15 ***(Loi uniforme généralisée)***

Une variable aléatoire X suit la loi uniforme sur $[\![a,b]\!]$ lorsque $X(\Omega) = [\![a,b]\!]$ et $\forall k \in [\![a,b]\!]$,

$$P(X = k) = \frac{1}{b-a+1}.$$

On note alors $X \hookrightarrow \mathcal{U}([\![a,b]\!])$.

Preuve : Il faut montrer que ces propriétés permettent bien de définir une variable aléatoire. On pourrait le faire à la main mais il est intéressant de remarquer qu'on peut se ramener par translation à une loi de type $\mathcal{U}([\![1,n]\!])$, avec $n = b-a+1$ pour que le nombre de valeurs atteignables soit identique.
En effet, on peut écrire $X = (a-1) + Y$ avec $Y \hookrightarrow \mathcal{U}([\![1, b-a+1]\!])$. C'est donc bien une variable aléatoire. □

Propriété 19 ***(Espérance et variance d'une variable uniforme généralisée)***

Si X est une variable uniforme telle que $X \hookrightarrow \mathcal{U}([\![a,b]\!])$, alors

$$E(X) = \frac{a+b}{2} \text{ et } V(X) = \frac{(b-a+1)^2-1}{12}.$$

Preuve : On peut utiliser le théorème de transfert, mais il y a plus simple en remarquant comme précédemment que $X = (a-1)+Y$ avec $Y \hookrightarrow \mathcal{U}([\![1, b-a+1]\!])$.

Les formules d'espérance et variance d'une somme donnent donc :

$$E(X) = a-1+E(Y) = a-1+\frac{b-a+2}{2} = \frac{a+b}{2},$$

$$V(X) = V(Y) = \frac{(b-a+1)^2-1}{12}$$

MODÈLE D'APPARITION : On considère une expérience aléatoire dont l'ensemble des résultats possible se partitionne en n groupes ayant la même probabilité d'apparition.

En numérotant ces groupes de 1 à n, une loi uniforme $\mathcal{U}\left([\![1;n]\!]\right)$

Remarque : Encore plus général, pour toute partie finie $\{x_1, \cdots, x_n\}$ de $\mathbb{R}$, on peut définir une v.a.r X suivant la loi uniforme sur $A = \{x_1, \cdots, x_n\}$ par :

$$P(X = x_k) = \frac{1}{\text{Card}(A)} = \frac{1}{n}$$

LOI HYPERGÉOMÉTRIQUE

introduction

Une urne contient a boules blanches et b boules noires. On tire une poignée de n boules dans l'urne. On appelle X la variable aléatoire égale au nombre de boules blanches dans la poignée.

étude de $X(\Omega)$

Si on prend plus que b boules, la poignée contient nécessairement des boules blanches, au moins $n-b$. Si on en tire moins de b, il peut y en avoir 0.

Mais on ne peut pas tirer plus de n boules blanches, ni plus de a.

Ainsi, $X(\Omega)$ $[\![\max(0, n-b); \min(n,a)]\!]$

Loi de X

Notons $N = a+b$. L'univers est l'ensemble des $\binom{N}{n}$ poignées de n boules. Il y a équiprobabilité.

Soit $k \in X(\Omega)$. Il y a $\binom{a}{k}$ choix des k boules blanches et $\binom{b}{n-k}$ choix pour les boules noires. Finalement il y a $\binom{a}{k}\binom{b}{n-k}$ issues réalisant $[X=k]$ et

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

Si on note $p = \frac{a}{a+b} = \frac{a}{N}$ la proportion de boules blanches, alors $(1-p) = \frac{b}{N}$ est la proportion de boules noires. On peut reformules les choses de la façon suivante...

Définition 16

Soit $N \in \mathbb{N}^*$, $n \in [\![0; N]\!]$ et $p \in]0; 1[$. Une variable aléatoire suit la loi hypergéométrique de paramètres N, n et p *si, et seulement si*

- $X(\Omega) = [\![\max(0, n - N(1-p)); \min(n, Np)]\!]$
- $\forall k \in X(\Omega), p(X = k) = \dfrac{\dbinom{Np}{k}\dbinom{N(1-p)}{n-k}}{\dbinom{N}{n}}$

On note alors $X \hookrightarrow \mathcal{H}(N, n, p)$

MODÈLE d'APPARITION : On prélève au hasard (simultanément ou successivement sans remise) n objets dans un ensemble de N éléments comportant deux types distincts. On connaît soit les effectifs a et b des deux espèces (et $N = a + b$), soit la proportion $p = \frac{a}{N}$ de la première espèce.
On compte le nombre d'éléments de première espèce dans le prélèvement.

Propriété 20 ***(Formule de Van der Monde (hors programme, mais à connaître))***

Soit $(a, b) \in \mathbb{N}^*$ et $n \in [\![1; a+b]\!]$ et $k \in [\![1; n]\!]$. Alors

$$\sum_{k=0}^{n} \binom{a}{k}\binom{b}{n-k} = \binom{a+b}{n}$$

Le programme dit que la formule est hors programme. Mais dans le programme il y a la loi hypergéométrique et on sait que, si $X \hookrightarrow \mathcal{H}(a+b, n, \frac{a}{a+b})$, on a avec les conventions usuelles sur les coefficients binômiaux :

$$1 = \sum_{k=0}^{n} P(X = k)$$

$$1 = \sum_{k=0}^{n} \frac{\binom{a}{k}\binom{b}{n-k}}{\binom{a+b}{n}}$$

$$\text{donc} \quad \binom{a+b}{n} = \sum_{k=0}^{n} \binom{a}{k}\binom{b}{n-k}$$

Et on peut obtenir la formule de Van der Monde « juste » en écrivant que la somme des probabilités d'obtenir les différentes valeurs pour une loi hypergéométrique vaut 1.
Alors... to be or not to be au programme, that's the question !

Propriété 21

Si $X \hookrightarrow \mathcal{H}(N, n, p)$, alors

$$E(X) = np \qquad \text{et} \qquad V(X) = \frac{npq(N-n)}{N-1}$$

Preuve : Avec les conventions usuelles sur les coefficients binômiaux, on a $\binom{a}{k} = 0$ si $k > a$ et $\binom{b}{n-k} = 0$ si $k > n$ ou $k > n - b$ et on peut écrire

$$E(X) = \sum_{k=0}^{n} kP(X = k)$$

$$E(X) = \sum_{k=1}^{n} k\frac{\binom{a}{k}\binom{b}{n-k}}{\binom{N}{n}}$$

$$E(X) = \frac{1}{\binom{N}{n}} \sum_{k=1}^{n} a\binom{a-1}{k-1}\binom{b}{n-k}$$

$$E(X) = \frac{a}{\binom{N}{n}}\binom{N-1}{n-1} \text{ avec la formule de Van der Monde}$$

$$E(X) = \frac{a\binom{N-1}{n-1}}{\frac{N}{n}\binom{N-1}{n-1}} \text{ avec la petite formule}$$

$$E(X) = \frac{an}{N} = np$$

Exemple n° 14

Approximation d'une loi hypergéométrique par une loi binômiale

Soit $X \hookrightarrow \mathcal{H}(N, n, p)$

On s'intéresse ici à ce qui se passe pour de grandes valeurs de N, n et p étant fixés.

Soit $k \in [\![0; n]\!]$.

Comme $a = Np$, pour N suffisamment grand, on a $a \geqslant n$ et $b \geqslant n$.

Notons $q = 1 - p$.

Dans toute la suite, on considère des valeurs de N pour lesquelles Np est entier.

$$P(X = k) = \frac{\binom{Np}{k}\binom{Nq}{n-k}}{\binom{N}{n}}$$

$$P(X = k) = \frac{(Np)!(Nq)!n!(N-n)!}{N!k!(Np-k)!(n-k)!(Nq-(n-k))!}$$

$$P(X = k) = \binom{n}{k}\frac{(Np)!}{(Np-k)!} \times \frac{(Nq)!}{(Nq-(n-k))!} \times \frac{(N-n)!}{N!}$$

$$P(X = k) = \binom{n}{k}\frac{(Np)(Np-1)(Np-2)...(Np-k+1) \times (Nq)(Nq-1)...(Nq-n+k+1)}{N(N-1)...(N-n+1)}$$

or, $Np+Nq = N(p+q) = N$ donc il y a autant de facteurs au numérateur et au dénominateur

$$P(X = k) = \binom{n}{k}\frac{Np}{N} \times \frac{Np-1}{N-1} \times ... \times \frac{Np-k+1}{N-k+1} \times \frac{Nq}{N-k} \times ... \times \frac{Nq-n+k+1}{N-n+1}$$

$$P(X = k) = \binom{n}{k}\prod_{j=0}^{k-1}\frac{Np-j}{N-j} \times \prod_{i=0}^{n-k-1}\frac{Nq-i}{N-(k+i)}$$

Or, k et p sont des constantes et pour tout $j \in [\![0; k-1]\!]$, $\frac{Np-j}{N-j} \underset{N\to+\infty}{\longrightarrow} p$ et,

par produit fini de limites, $$\lim_{N\to+\infty} \prod_{j=0}^{k-1} \frac{Np-j}{N-j} = p^k$$

de même, $$\lim_{N\to+\infty} \prod_{i=0}^{n-k-1} \frac{Nq-i}{N-(k+i)} = q^{n-k}$$

Finalement $$\boxed{\forall k \in [\![0; n]\!], \lim_{N\to+\infty} P(X=k) = \binom{n}{k} p^k q^k}$$

Supposons que l'on tire successivement les n boules. Quand N devient très grand par rapport à n, il ne peut y avoir de « rupture de stock » de boules blanches ou noires et la probabilité de tirer une boule blanche est quasiment inchangée. C'est pourquoi on peut approcher X par une loi binômiale $\mathcal{B}(n,p)$ qui compte le nombre de boules blanches tirées lors de n tirages successifs.

vu dans les rapports

ECRICOME 2017 : certains candidats pensent que toute variable aléatoire ne prenant que deux valeurs suit une loi de Bernoulli.

ECRICOME 2017 : On attendait ici une courte justification des valeurs extrêmes de $X_n(\Omega)$

ECRICOME 2017 : Lorsque le résultat est donné, certains candidats se contentent d'expliquer les termes qu'ils voient dans la formule fonale. On attendait ici un raisonnement rigoureux avec des événements et application de la formule des probabilités composées. Dans une question délicate de ce type, les points sont décomposés : l'écriture en événements rapporte des points, l'écriture correcte des probabilités composées avec les événements rapporte des points,...

ECRICOME 2018 : Une grande majorité des candidats montrent correctement que N suit une loi de Bernoulli, mais beaucoup peinent à déterminer le paramètre de cette loi ou le confondent avec P (N = 0).

ECRICOME 2018 : Dans cette question comme dans la suivante, la formule des probabilités totales doit être utilisée. Lors de son emploi, le système complet d'événements choisi doit être très clairement indiqué.
Elle est peu abordée et trop souvent bâclée.
[... question suivante ...]
Comme à la question précédente, la formule des probabilités totales est indispensable. Le raison- nement est sensiblement le même qu'à la question précédente. **Si la question précédente est traitée correctement**, une rédaction succincte est acceptée ici.

ECRICOME 2018 : La loi de X_1 est en général bien déterminée, même si peu parlent de la loi uniforme ; nombreux sont ceux qui pensent reconnaitre en X_1 une variable aléatoire de loi de Bernoulli. Trop de candidats s'engagent dans des calculs longs et maladroits pour déterminer l'espérance et la variance de X_1 .

ECRICOME 2018 : Si l'égalité demandée est en général obtenue, les candidats, en trop grande majorité, n'ont pas toujours précisé la formule utilisée, ni donné un système complet d'événements qui en soit vraiment un. Ainsi, $(X_k = i), (X_k = i - 1)$ n'est pas un système complet d'événements.

Exercices

exercice 1 ★★☆☆☆

Une urne contient deux boules blanches et quatre boules noires. On tire les boules une à une et sans remise jusqu'à ce qu'il ne reste dans l'urne que des boules de la même couleur. Soit X la variable aléatoire associée au nombre de tirages nécessaires.

1. Déterminer la loi de X.
2. Écrire une fonction Python sans entrée simulant l'expérience et renvoyant la valeur prise par X.

Proposition de corrigé :$X(\Omega) = [\![2;5]\!]$. En effet, il faut tirer au moins deux boules pour épuiser une couleur. Le nombre maximum de boules à tirer correspond à une suite de tirage telle qu'il reste à la fin une blanche et une noire. On a alors tiré 4 boules et il faut tirer une cinquième. Toutes les possibilités intermédiaires sont possibles.

Notons, pour tout $k \in [\![1;5]\!]$ B_k (resp N_k) l'événement « obtenir une boule blanche (resp. noire) au k-ième tirage. ».

• $[X = 2]$ sera réalisé ssi on les deux premières boules tirées sont blanches.
$P(X = 2) = P(B_1 \cap B_2) = P(B_1)\mathbb{P}(B_2 \mid B_1) = \frac{2}{6} \times \frac{1}{5} = \frac{1}{15}$

• $[X = 3]$ sera réalisé si on tire la deuxième boule blanche au troisième tirage, après avoir déjà tiré une blanche au premier ou au deuxième tirage. Ces deux issues sont incompatibles et on a :
$P(X = 3) = P(B_1 \cap N_2 \cap B_3) + P(N_1 \cap B_2 \cap B_3)$
$P(X = 3) = P(B_1)\mathbb{P}(N_2 \mid B_1)\,\mathbb{P}(B_3 \mid B_1 \cap B_2) + P(N_1)\mathbb{P}(B_2 \mid N_1)\,\mathbb{P}(B_3 \mid N_1 \cap N_2)$ d'après la formule des probabilités composées.
$P(X = 3) = \frac{2}{6} \times \frac{4}{5} \times \frac{1}{4} + \frac{4}{6} \times \frac{2}{5} \times \frac{1}{4} = \frac{2}{15}$

• $[X = 4]$ peut être réalisé soit en piochant successivement les quatre boules noires, soit en piochant la deuxième boule blanche au quatrième tirage (la première ayant été tirée lors d'un des trois premiers). Il y a donc 4 issues qui réalisent $[X = 4]$ et on a :
$P(X = 4) = P(N_1 \cap N_2 \cap N_3 \cap N_4) + P(B_1 \cap N_2 \cap N_3 \cap B_4) + P(N_1 \cap B_2 \cap N_3 \cap B_4) + P(N_1 \cap N_2 \cap B_3 \cap B_4)$
$P(X = 4) = \frac{4}{6} \times \frac{3}{5} \times \frac{2}{4} \times \frac{1}{3} + \frac{2}{6} \times \frac{4}{5} \times \frac{3}{4} \times \frac{1}{3} + \frac{4}{6} \times \frac{2}{5} \times \frac{3}{4} \times \frac{1}{3} + \frac{4}{6} \times \frac{3}{5} \times \frac{2}{4} \times \frac{1}{3}$
$P(X = 4) = \frac{1}{15} + \frac{1}{15} + \frac{1}{15} + \frac{1}{15}$
$P(X = 4) = \frac{4}{15}$

• $P(X = 5) = 1 - (P(X = 2) + P(X = 3) + P(X = 4)) = \frac{8}{15}$

Finalement, la loi de X est donnée par :

$x \in \Omega$	2	3	4	5
$P(X = x)$	$\frac{1}{15}$	$\frac{2}{15}$	$\frac{4}{15}$	$\frac{8}{15}$

Quelques remarques : Dans ce genre d'exercice un arbre (au brouillon) reste réalisable quoique très grand. Mais si vous êtes perdu, cela peut aider.

Pour prolonger, on peut se demander comment on aurait pu gérer un tel exercice avec 2 boules blanches et n boules noires, voire avec b boules blanches et n boules noires...

Dans un tel exercice, il peut être plus « léger » de noter $B_1N_2B_3$... pour $B_1 \cap N_2 \cap B_3$.... Si on souhaite utiliser une telle convention d'écriture, il suffit de la décrire avant.

```
import random as rd
def X():
    b = 2
    n = 4
    x = 0
    while b*n>0 :
        if rd.randint(1,b+n)<=b:
```

```
            b -= 1
        else:
            n -= 1
        x += 1
    return x

# pour générer un échantillon de taille 100
print([X() for k in range(100)])
```

on peut remplacer le test `b*n>0` *par* `b!=0 and n!=0`

---

exercice 2 ★★☆☆☆

On considère le code ci-dessous.

```
import random as rd

def simul():
    S = []
    for k in range(3):
        S += [rd.randint(1,6)]
    return(S)

def X(L):
    np = 0
    ni = 0
    for i in range(len(L)):
        if L[i]%2 == 0:
            np += 1
        else:
            ni += 1
    return(2*np-ni)
```

```
def Y(L):
    if L[0] == L[1] == L[2]:
        return(10)
    L.sort()
    if L == [1,2,4]:
        return(20)
    return(-1)

# Illustration

for k in range(10):
    L = simul()
    print(L)
    print(X(L))
    print(Y(L))
    print('\n')
```

Décrire en français l'expérience aléatoire simulée, et l'univers. Décrire les variables aléatoires en jeu ici. Pour chaque variable aléatoire, on donnera la loi.

Proposition de corrigé :

• `simul()` simule trois lancers de dés successifs. L'univers Ω est donc l'ensemble des 3-listes de $[\![1;6]\!]$.
Ainsi, $\mathrm{Card}(\Omega) = 6^3$.

• pour une issue, `X` compte le nombre n_p de tirages pairs et le nombre n_i de tirages impairs et renvoie $2n_p - n_i$.
Remarquons que n_p et n_i sont liés : $n_i = 3 - n_p$.
Ainsi, $X = 2n_p - n_i = 3n_p - 3$.
Le nombre de dés pairs peut être égal à 0,1,2 ou 3. Ainsi, $X(\Omega) = \{-3, 0, 3, 6\ \}$
Pour déterminer la loi de X, on peut décrire l'univers, compter les issues... mais le plus simple est de remarquer que n_p est également une variable aléatoire. Si on appelle « succès » l'obtention d'un numéro pair, n_p compte le nombre de succès dans la répétition indépendante de 3 épreuves de Bernoulli de même paramètre $\frac{1}{2}$ (il y a une chance sur deux d'avoir un nombre pair). Donc $n_p \hookrightarrow \mathcal{B}(3, \frac{1}{2})$.
Ainsi,
$P(X = -3) = P(3n_p - 3 = -3) = P(n_p = 0) = \binom{3}{0}\left(\frac{1}{2}\right)^3 = \frac{1}{8}$.
$P(X = 0) = P(3n_p - 3 = 0) = P(n_p = 1) = \binom{3}{1}\left(\frac{1}{2}\right)^3 = \frac{3}{8}$.
$P(X = 3) = P(3n_p - 3 = 3) = P(n_p = 2) = \binom{3}{2}\left(\frac{1}{2}\right)^3 = \frac{3}{8}$.
$P(X = 6) = P(3n_p - 3 = 6) = P(n_p = 3) = \binom{3}{3}\left(\frac{1}{2}\right)^3 = \frac{1}{8}$.

• à une issue, `Y` associe 10 si les trois chiffres sont identiques, 20 si les trois chiffres sont 1, 2, et 4 et -1 dans tous les autres cas.
$[Y = 10]$ est réalisé par les issues dont les trois numéros sont identiques. Donc $P(Y = 10) = \frac{3}{6^3} = \frac{1}{72}$.
Les issues réalisant $[Y = 20]$ sont les $3!$ permutations de $\{1, 2, 4\}$. Ainsi, $P(Y = 20) = \frac{3!}{6^3} = \frac{1}{36}$.
Enfin, $[Y = -1] \cup [Y = 10] \cup [Y = 20] = \Omega$ donc $P(Y = -1) = 1 - P(Y = 10) - P(Y = 20) = \frac{69}{72} = \frac{23}{24}$.

exercice 3 ★☆☆☆☆

Soit X une variable à valeurs dans $\mathbb{Z}$ définie par sa fonction de répartition f_X définie par :

$$f_X(x) = \begin{cases} 0 & \text{si } x < -1 \\ \frac{1}{3} & \text{si } -1 \leqslant x < 2 \\ 1 & \text{si } x \geqslant 2 \end{cases}$$

1. Déterminer la loi de X
2. Soit $T = aX + b$ déterminer la loi de T

Proposition de corrigé :Pour tout $k \in \mathbb{Z}$, on a $P(X = k) = P(X \leqslant k) - P(X \leqslant k-1) = F_X(k) - F_X(k-1)$
Ainsi, si $k < 1$, $P(X = k) = 0$
$P(X = -1) = F_X(-1) - F_X(-2) = \frac{1}{3}$
$P(X = 0) = F_X(0) - F_X(-1) = \frac{1}{3} - \frac{1}{3} = 0$. De même $P(X = 1) = P(X = 2) = 0$
$P(X = 2) = F_X(2) - F_X(1) = 1 - \frac{1}{3} = \frac{2}{3}$
$\forall k > 2, P(X = k) = F_X(k) - F_X(k-1) = 1 - 1 = 0$
Finalement, $\boxed{X(\Omega) = \{-1; 2\}.\ P(X = -1) = \frac{1}{3} \text{ et } P(X = 2) = \frac{2}{3}}$

exercice 4 ★★★☆☆

Soit $n \in \mathbb{N}^*$. Soit X une variable aléatoire réelle définie sur un espace probabilisé $(\Omega, \mathcal{P}(\Omega), \mathbb{P})$.

On suppose que $X(\Omega) = [\![1; n]\!]$ et qu'il existe $\lambda > 0$ tel que :

pour tout $k \in [\![1; n]\!]$, $\quad \mathbb{P}(X = k) = \lambda k$.

1. Déterminer la valeur de λ
2. Calculer l'espérance de X, puis de $Y = X + 1$ et de $Z = X + \dfrac{1}{X}$

Proposition de corrigé :Nous avons déjà traité plusieurs exercices similaires (cf ex précédent.)

Ainsi, $\lambda = \frac{2}{n(n+1)}$ et $E(X) = \frac{2n+1}{3}$.
Par linéarité de l'espérance, on a $E(Y) = E(X+1) = E(X) + 1 = \frac{2n+1}{3} + 1 = \frac{2n+4}{3}$

Z est bien définie var $0 \notin X(\Omega)$ et $x \mapsto x + \frac{1}{x}$ est déinie sur $\mathbb{R}^*$.
$E(Z) = E(X + \frac{1}{X})$
$E(Z) = \displaystyle\sum_{k=1}^{n} (k + \frac{1}{k}) P(X = k)$ avec le théorème de transfert théorème de transfert
$E(Z) = \displaystyle\sum_{k=1}^{n} k P(X = k) + \sum_{k=1}^{n} \frac{1}{k} P(X = k)$ par linéarité de la somme
$E(Z) = E(X) + \displaystyle\sum_{k=1}^{n} \frac{1}{k} \lambda k$
$E(Z) = \frac{2n+1}{3} + n\lambda$
$E(Z) = \frac{2n+1}{3} + \frac{2}{n+1}$
$\boxed{E(Z) = \frac{2n^2+3n+7}{3(n+1)}}$

exercice 5 ★★★★☆

On considère un sac contenant des jetons indiscernables au toucher et numérotés de 1 à 6.

1. On procède à deux tirages consécutifs d'un jeton avec remise et on appelle M le plus grand des deux numéros obtenus et M' le plus petit.

   a. Déterminer la loi de M.

   on pourra considérer les événements $J_{k,n}$: « obtenir le jeton k au n-ième tirage. »

   b. Déterminer la loi de M'.

2. On procède à deux tirages consécutifs d'un jeton sans remise et on appelle N le plus grand des deux numéros obtenus.

   a. Donner $N(\Omega)$.

   b. Déterminer la loi de N.

3. Simuler en Python les variables aléatoires M et N.

Proposition de corrigé :

1. a. Pour $(k,n) \in [\![1;6]\!] \times \{1,2\}$, notons $J_{k,n}$ l'événement « obtenir le jeton k au n-ième tirage. »
$[M=k] = \left(J_{k,1} \cap \bigcup_{i=1}^{k} J_{i,2}\right) \cup \left(J_{k,2} \cap \bigcup_{i=1}^{k-1} J_{i,1}\right) = \bigcup_{i=1}^{k}(J_{k,1} \cap J_{i,2}) \cup \bigcup_{i=1}^{k-1}(J_{k,2} \cap J_{i,2})$ avec la convention qu'une réunion vide est l'ensemble vide pour $k=1$
C'est une réunion d'événements deux à deux incompatibles et, puisque les tirages se font avec remise, on a pour tout $(a,b) \in [\![1;6]\!]^2$ on a : $P(J_{a,1} \cap J_{b,2}) = \frac{1}{36}$.
On a donc, pour tout $k \in [\![1;6]\!]$ avec la convention qu'une somme vide est nulle (pour $k=1$) :

$$P(M=k) = \sum_{i=1}^{k} \frac{1}{36} + \sum_{j=1}^{k-1} \frac{1}{36} = \frac{2k-1}{36}$$

$$\boxed{\text{pour tout } k \in [\![1;6]\!],\ P(M=k) = \frac{2k-1}{36}}$$

comme toujours, vérifier que la loi obtenue vérifie $\sum_{k=1}^{6} P(M=k) = 1$...

Autre méthode. Pour $i \in \{1;2\}$, notons X_i la v.a.r qui, à un tirage associe le numéro du jeton tiré au i-ème tirage. Notons F_{X_i} la fonction de répartition de X_i. Pour tout $k \in [\![1;6]\!]$,
$F_M(k) = P(M \leqslant k) = P([X_1 \leqslant k] \cap [X_2 \leqslant k]) = P(X_1 \leqslant k)P(X_2 \leqslant k)$ car les tirages sont indépendants.
$F_M(k) = \frac{k^2}{36}$

Ainsi, $P(M=k) = F_M(k) - F_M(k-1) = \frac{k^2-(k-1)^2}{36} = \frac{2k-1}{36}$

b. De même, avec les notations précédentes, on a : $M'(\Omega) = [\![1;6]\!]$ et
$[M'=k] = \bigcup_{i=k}^{6}(J_{k,1} \cap J_{i,2}) \cup \bigcup_{i=k+1}^{6} (J_{k,1} \cap J_{i,2})$
$P(M'=k) = \sum_{i=k}^{6} \frac{1}{36} + \sum_{i=k+1}^{6} \frac{1}{36}$ avec la convention de la somme vide égale à 0 pour $k=6$
$P(M'=k) = \frac{6-k+1}{36} + \frac{6-k}{36} = \frac{13-2k}{36}$

$\boxed{\text{pour tout } k \in [\![1;6]\!],\ P(M'=k) = \frac{13-2k}{36}}$

remarque : on peut aussi envisager $P(N \geqslant k) = P([X_1 \geqslant 1] \cap [X_2 \geqslant k]) = P(X_1 \geqslant k)P(X_2 \geqslant k)$... en faisant bien attention aux cas particuliers et que le but est $F_N(k) = P(X \leqslant k) = 1 - P(N \geqslant k+1)$ puisque k est entier.

2. Cas « sans remise »

a. $N(\Omega) = [\![2;6]\!]$. En effet, les deux jetons étant distincts, 1 ne peut être une valeur et le plus petit maximum possible est obetnu avec $(1;2)$. Pour tout $k \in [\![2;6]\!]$, on a :

b. $[N = k] = \left(J_{k,1} \cap \bigcup_{i=1}^{k-1} J_{i,2}\right) \cup \left(J_{k,2} \cap \bigcup_{i=1}^{k-1} J_{i,1}\right) = \bigcup_{i=1}^{k}(J_{k,1} \cap J_{i,2}) \cup \bigcup_{i=1}^{k-1}(J_{k,2} \cap J_{i,2})$

C'est une réunion d'événements deux à deux incompatibles.

$P(M = k) = \sum_{i=1}^{k-1} P(J_{k,1})\mathbb{P}(J_{i,2} \mid J_{k,1}) + \sum_{i=1}^{k-1} P(J_{i,1})\mathbb{P}(J_{k,2} \mid J_{i,1})$

$P(M = k) = \sum_{i=1}^{k-1} \frac{1}{6} \times \frac{1}{5} + \sum_{i=1}^{k-1} \frac{1}{6} \times \frac{1}{5}$

$P(M = k) = \frac{2k-2}{30}$

$\boxed{\forall k \in [\![2;6]\!]\,, P(N = k) = \frac{k-1}{15}}$

Remarque : ici encore, penser à vérifier la somme...
Une alternative aux $J_{i,n}$ aurait été d'expliciter l'univers (ensemble des 2-listes sans répétition d'éléments de $[\![1;6]\!]$, et de dénombrer $[N = k]$.

---

exercice 6 ★★★☆☆

Soit $n \in \mathbb{N}^*$ et X une variable aléatoire qui prend ses valeurs dans $[\![0;n]\!]$. Pour tout $k \in [\![0;n]\!]$, on suppose que $P(X = k) = \ln(a^k)$, où a est un réel strictement positif.

1. Déterminer a pour que X soit une variable aléatoire.
2. Calculer $\mathbf{E}(X)$.
3. Calculer $\mathbf{V}(X)$.

<u>*Proposition de corrigé*</u> :Pour $a \geqslant 1$, on a $\ln(a^k) \geqslant 0$ pour tout $k \in [\![0;n]\!]$. Il suffit maintenant d'avoir $\sum_{k=0}^{n} P(X = k) = 1$ c'est à dire $\sum_{k=0}^{n} \ln(a^k) = 1$ soit $\sum_{k=0}^{n} k\ln(a) = 1$ d'où $\ln(a)\frac{n(n+1)}{2} = 1$ puis $\boxed{a = \exp\left(\frac{2}{n(n+1)}\right)}$

On a alors bien $a > 1$ et, pour cette valeur de a, on a : $\forall k \in [\![0;n;,]\!] \quad P(X = k) = \frac{2}{n(n+1)}k$

$\mathbf{E}(X) = \sum_{k=0}^{n} kP(X = k) = \frac{2}{n(n+1)} \sum_{k=0}^{n} k^2 = \frac{2}{n(n+1)} \times \frac{n(n+1)(2n+1)}{6} = \frac{(2n+1)}{3}$

$\boxed{\mathbf{E}(X) = \frac{2n+1}{3}}$

$\mathbf{V}(X) = \mathbf{E}(X^2) - (\mathbf{E}(X))^2$ d'après la formule de Kœnig Huygens'.

$\mathbf{E}(X^2) = \sum_{k=0}^{n} kP(X = k) = \frac{2}{n(n+1)} \sum_{k=0}^{n} k^3 = \frac{2}{n(n+1)} \times \frac{n^2(n+1)^2}{4} = \frac{n(n+1)}{2}$

Ainsi, $\mathbf{V}(X) = \frac{n(n+1)}{2} - \frac{(2n+1)^2}{9} = \frac{9n^2+9n-2(4n^2+4n+1)}{18} = \frac{n^2+n-2}{18}$

$\boxed{\mathbf{V}(X) = \frac{n^2+n-2}{18}}$

NB : une variance est toujours positive (ou nulle). Dans une formule abstraite, cette petite vérification permet parfois de débusquer des étourderies dans le calcul !

exercice 7 ★★☆☆☆

Soit $n \in \mathbb{N}^*$. On dispose d'une urne contenant n boules numérotées de 1 à n. On commence par tirer une boule, on note X son numéro, puis on remet cette boule dans l'urne, mais on retire toutes les boules dont les numéros sont strictement inférieurs. On effectue alors un nouveau tirage dans l'urne, et on note Y le numéro obtenu.

1. Quelle est la loi de X ? Préciser $\mathbf{E}(X)$ et $\mathbf{V}(X)$
2. Déterminer la loi de Y en fonction des nombres $H_l = \displaystyle\sum_{k=1}^{\ell} \frac{1}{k}$, avec $\ell \in \mathbb{N}^*$.

Proposition de corrigé :

1. Les boules ayant chacune autant de chances d'être tirées, on a :

$$\boxed{X \hookrightarrow \mathcal{U}(\llbracket 1;n \rrbracket).\ E(X) = \tfrac{n+1}{2} \text{ et } V(X) = \tfrac{n^2-1}{12}}$$

2. $Y(\Omega) = \llbracket 1;n \rrbracket$. En effet, en tirant la boule 1 au premier tirage toutes les boules sont disponibles au deuxième. Soit $k \in \llbracket 1;n \rrbracket$. La famille $([X=i])_{1\leqslant i\leqslant n}$ forme un système complet d'événements et la formule des probabilités totales donne alors :

$$P(Y=k) = \sum_{i=1}^{n} P(X=i)\mathbb{P}(Y=k \mid X=i)$$

Or, pour tout $i \in \llbracket 1;n \rrbracket$. Supposons réalisé l'événement $[X=i]$. Avant le deuxième tirage, l'urne contient les boules numérotées de i à n. ainsi, si $k<i$, la boule k n'est pas dans l'urne et $\mathbb{P}(Y=k \mid X=i)=0$ et si $k \geqslant i$, il y a une boule k parmi les $n-i+1$ boules de l'urne donc $\mathbb{P}(Y=k \mid X=i) = \frac{1}{n-i+1}$

Ainsi, $P(Y=k) = \displaystyle\sum_{i=1}^{k} \frac{1}{n}\frac{1}{n-i+1} = \frac{1}{n}\sum_{\ell=n-k+1}^{n} \frac{1}{\ell}$ avec le changement d'indice $j=n-i+1$

Finalement, $\boxed{\forall k \llbracket 1;n \rrbracket, \quad P(Y=k) = \tfrac{1}{n}\left(H_n - H_{n-k}\right)}$

exercice 8 On considère une suite de tirages avec remise dans une urne contenant N boules numérotées de 1 à N. Pour tout $n \geqslant 1$, on note Y_n le nombre de numéros non encore sortis à l'issue du n-ième tirage.

1. Déterminer Y_1.
2. Soit $n \geqslant 2$.
   a. Déterminer $Y_n(\Omega)$.
   b. Montrer que, pour $k \in [\![0; N-1]\!]$, on a :

$$P(Y_n = k) = \frac{N-k}{N} P(Y_{n-1} = k) + \frac{k+1}{N} P(Y_{n-1} = k+1).$$

3. En déduire que la suite $(E(Y_n))_{n \geqslant 1}$ est une suite géométrique et en déduire l'expression explicite de $E(Y_n)$ pour tout $n \geqslant 1$.

Proposition de corrigé :

1. Y_1 est la variable certaine égale à $N-1$. En effet, après un tirage, les $N-1$ boules non tirées ont un numéro non tiré.
2. Soit $n \geqslant 2$.
   a. $Y_n(\Omega) \subset \mathbb{N}$. De plus, on tire au moins un numéro, et il reste donc au maximum $N-1$ numéros non tirés. Si $n < N$, on aura tiré au maximum n numéros distincts et il restera $N-n$ numéros non tirés. Ainsi, $Y_n(\Omega) \subset [\![N-n; N-1]\!]$. Réciproquement, si $k \in [\![N-n; N-1]\!]$, alors $k \geqslant N-n$ donc $n \geqslant N-k$ et on peut tirer $N-k$, numéros distincts aux $N-k$ premiers tirages, et des numéros déjà tirés aux suivants. Il reste alors $N-(N-k) = k$ numéros non tirés donc $k \in Y(\Omega)$. Finalement

$$\boxed{Y_n(\omega) = \begin{cases} [\![N-n; N-1]\!] & \text{si } n < N \\ [\![0; N-1]\!] & \text{sinon} \end{cases}}$$

On peut aussi répondre $Y_n(\Omega) = [\![\max(0, N-n); N-1]\!]$

   b. Soit $k \in [\![0; N-1]\!]$. La famille $([Y_{n-1} = j])_{j \in [\![0; N-2]\!]}$ est un s.c.e et avec la formule des probabilités totales, on a :

$$\begin{aligned} P(Y_n = k) &= \sum_{j=0}^{n} P([Y_{n-1} = j] \cap [Y_n = k]) \\ &= P([Y_{n-1} = k] \cap [Y_n = k]) + P([Y_{n-1} = k+1] \cap [Y_n = k]) \end{aligned}$$

En effet, un tirage supplémentaire ne peut faire varier le nombre de numéros non tirés de plus de 1.

Comme $Y_{n-1}(\Omega) = [\![N-n+1; N-1]\!]$ et $Y_n(\Omega) [\![N-n; N-1]\!]$,
• si $k \notin Y_n(\Omega)$, alors $k \notin Y_{n-1}(\Omega)$ et $k+1 \notin Y_{n-1}(\Omega)$ donc tous les termes de l'égalité sont nuls et celle-ci est vraie.

• Si $k \in Y_n(\Omega)$, alors $k+1 \in Y_{n-1}(\Omega)$ donc $P(Y_{n-1} = k+1) \neq 0$. De plus, si $[Y_{n-1} = k+1]$ est réalisé, $[Y_n = k]$ est réalisé ssi on tire l'un des $k+1$ numéros non encore tirés et $\mathbb{P}(Y_n = k \mid Y_{n-1} = k+1) = \frac{k+1}{N}$.

Si $k = N-n$, $P(Y_{n-1} = k) = 0$ donc

$$P(Y_n = k) = P(Y_{n-1} = k+1)\mathbb{P}(Y_n = k \mid Y_{n-1} = k+1) = \frac{k+1}{N} P(Y_{n-1} = k+1)$$

et l'égalité est vraie.

Enfin, si $k > N-n$, si on suppose réalisé $[Y_{n-1} = k+1]$, $[Y_n = k]$ est réalisé ssi on tire l'un des $N-k$ numéros déjà tirés et $\mathbb{P}(Y_n = k \mid Y_{n-1} = k+1) = \frac{N-k}{N}$. On a alors :

$$P(Y_n = k) = P(Y_{n-1} = k)\mathbb{P}(Y_n = k \mid Y_{n-1} = k) + P(Y_{n-1} = k+1)\mathbb{P}(Y_n = k \mid Y_{n-1} = k+1)$$
$$= \frac{N-k}{N}P(Y_{n-1} = k) + \frac{k+1}{N}P(Y_{n-1} = k+1) \text{ et l'égalité est vraie.}$$

On a prouvé :

$$\boxed{\text{pour } k \in [\![0; N-1]\!], \text{ on a : } P(Y_n = k) = \frac{N-k}{N}P(Y_{n-1} = k) + \frac{k+1}{N}P(Y_{n-1} = k+1).}$$

Attention, l'écriture $\mathbb{P}(A \mid B)$ n'a de sens que si $P(B) \neq 0$ et donc il est nécessaire de distinguer des cas.

3. Comme $P(Y_n = k) = 0$ si $k < N - n$, on peut écrire :

$$E(Y_n) = \sum_{k=N-n} kP(Y_n = k)$$
$$E(Y_n) = \sum_{k=0}^{N-1} kP(Y_n = k)$$
$$E(Y_n) = \sum_{k=0}^{N-1} k\left(\frac{N-k}{N}P(Y_{n-1} = k) + \frac{k+1}{N}P(Y_{n-1} = k+1)\right)$$
$$NE(Y_n) = \sum_{k=0}^{N-1} k(N-k)P(Y_{n-1} = k) + \sum_{k=0}^{N-1} k(k+1)P(Y_{n-1} = k+1)$$
$$NE(Y_n) = \sum_{k=0}^{N-1} k(N-k)P(Y_{n-1} = k) + \sum_{j=1}^{N}(j-1)jP(Y_{n-1} = j) \text{ avec le changement d'indice } j = k+1$$
$$NE(Y_n) = \sum_{k=1}^{N-1} [k(N-k) + k(k-1)]\, P(Y_{n-1} = k) \text{ en enlevant les termes nuls } (P(Y_{n-1} = N) = 0)$$
$$NE(Y_n) = (N-1)\sum_{k=1}^{N-1} kP(Y_{n-1} = k)$$
$$E(Y_n) = \frac{N-1}{N}E(Y_{n-1})$$

$$\boxed{(E(Y_n))_{n\geqslant 1} \text{ est une suite géométrique de raison } \tfrac{N-1}{N}}$$

Or Y_1 est la v.a.r certaine égale à $N-1$ donc $E(Y_1) = N-1$ et,

pour tout entier $n \geqslant 1$, $E(Y_n) = \left(\frac{N-1}{N}\right)^{n-1} E(Y_1) = \left(\frac{N-1}{N}\right)^{n-1}(N-1)$

$$\boxed{\forall n \geqslant 1, E(Y_n) = \frac{(N-1)^n}{N^{n-1}}}$$

---

exercice 9 ★★★★☆

Soit $n \in \mathbb{N}^*$. On lance n fois une pièce de monnaie donnant "pile" avec la probabilité $p \in]0, 1[$ et "face" avec la probabilité $q = 1 - p$.

On appelle k -chaîne une suite de k lancers consécutifs ayant tous donné "pile", cette suite devant être précédée et suivie d'un "face" (à moins qu'elle ne débute ou termine le tirage).

Pour tout k de $\{1, ..., n\}$, on note Y_k la variable aléatoire égale au nombre total de k-chaînes de "pile" obtenues au cours de ces n lancers, et P_k l'événement "on obtient "pile" au k -ième lancer".

Par exemple, avec $n = 11$, si l'on a obtenu les résultats $P_1P_2F_3F_4P_5P_6P_7F_8P_9F_{10}P_{11}$ alors $Y_1 = 2$, $Y_2 = 1$ et $Y_3 = 1$

Le but de cet exercice est de déterminer, pour tout k de $\{1, ..., n\}$, l'espérance de Y_k , notée $E(Y_k)$.

1. Déterminer la loi de Y_n et donner $E(Y_n)$.
2. Soit $n \in \mathbb{N}, n \geqslant 2$. Montrer que $\mathbb{P}(Y_{n-1}=1)=2qp^{n-1}$ et donner $E(Y_{n-1})$.
3. Dans cette question, k désigne un élément de $\{1,\dots,n-2\}$.

   Pour tout i de $\{1,\dots,n\}$, on note $X_{i,k}$ la variable aléatoire qui vaut 1 si une k -chaîne de "pile" commence au i -ème lancer et qui vaut 0 sinon.

   a. Calculer $\mathbb{P}(X_{1,k}=1)$.

   b. Soit $i \in \{2,\dots,n-k\}$. Montrer que $\mathbb{P}(X_{i,k}=1)=q^2p^k$.

   c. Montrer que $\mathbb{P}(X_{n-k+1,k}=1)=qp^k$.

   d. Exprimer Y_k en fonction des variables $X_{i,k}$ puis déterminer $E(Y_k)$.

Proposition de corrigé :

1. Soit $n \in \mathbb{N}^*$. $Y_n(\Omega)=\{0;1\}$ car il ne peut y avoir qu'une n-chaîne parmi n tirages. $[Y_n=1]$ est réalisé si, et seulement si on ne tire que des pile.

   $\mathbb{P}(Y_n=1)=\mathbb{P}(\bigcap_{k=1}^n P_k)=\prod_{k=1}^n \mathbb{P}(P_k)=p^n$

   $\boxed{\text{Ainsi, } X \text{ suit une loi de Bernoulli de paramètre } p^n \text{ et son espérance vaut } E(Y_n)=p^n(1-p^n)}$
2. Une $(n-1)$-chaîne de piles ne peut être obtenue qu'avec $(n-1)$ piles précédés ou suivis d'un « face ». Ainsi, $[Y_{n-1}=1]=(P_1P_2\dots P_{n-1}F_n)\cup(F_1P_2\dots P_n)$
   Ainsi, $P(Y_{n-1}=1)=P(P_1P_2\dots P_{n-1}F_n)+P(F_1P_2\dots P_n)$ car les deux issues sont incompatibles.
   $P(Y_{n-1}=1)=p^{n-1}q+qp^{n-1}$ car les tirages sont indépendants.
   Finalement, $\boxed{P(Y_{n-1}=1)=2qp^{n-1}}$

   $Y_{n-1}(\Omega)=\{0;1\}$. En effet, soit $k \geqslant 2$ dans $X(\omega)$ Alors on aurait $k(n-1)\leqslant n$ soit $k \leqslant \frac{1}{n-2}$ ce qui est contradictoire. Ainsi, $Y_{n-1} \hookrightarrow \mathcal{B}(2qp^{n-1}))$
   $\boxed{E(Y_{n-1})=2qp^{n-1}}$.
3. a. Soit $k \in [\![1;n-2]\!]$. $[X_{1,k}=1]=P_1P_2\dots P_kF_{k+1}$. Ainsi, $\boxed{P(X_{1,k}=1)=p^kq}$.

   b. Soit $i \in [\![2;n-k]\!]$. $[X_{i,k}=1]=F_{i-1}P_iP_{i+1}\dots P_{i+k-1}F_{i+k}$. Ainsi, $\boxed{P(X_{i,k}=1)=qp^kq=q^2p^k}$.
   RQ : avez-vous regardé pourquoi la restriction à $[\![2;n-k]\!]$?

   c. $[X_{n-k+1,k}=1]=F_{n-k}P_{n-k+1}\dots P_n$ donc $\boxed{P(X_{n-k+1,k}=1)=qp^k}$.

   d. Pour tout $k \in [\![1;n]\!]$, $\boxed{Y_k=\sum_{i=1}^{n-k+1} X_{i,k}}$. Remarquons que, pour tout $i \in [\![1;n-k+1]\!]$, la variable $X_{i,k}$ suit une loi de Bernoulli. Par linéarité de l'espérance, on obtient :

   $E(Y_k)=\sum_{i=1}^{n-k+1} E(X_{i,k})$

   $E(Y_k)=E(X_{1,k})+\sum_{i=2}^{n-k} E(X_{i,k})+E(X_{n-k+1,k})$

   $\boxed{E(Y_k)=2qp^k+(n-k-1)q^2p^k}$

---

exercice 10 ★★★★☆

Soit n un entier supérieur ou égal à 2.

On effectue $2n$ tirages au hasard dans une urne contenant n boules numérotées de 1 à n.

Un tirage consiste à extraire une boule de l'urne, la boule tirée étant remise dans l'urne. On note N la variable aléatoire égale au numéro du tirage au cours duquel, pour la première fois, on a obtenu une boule déjà obtenue auparavant.

Rappel : on note $A_n^k=\frac{n!}{(n-k)!}$

1. Montrer que $N(\Omega) = [\![2; n+1]\!]$

2. Montrer que : $\forall k \in [\![1; n]\!]$, $\quad P(N > k) = \dfrac{A_n^k}{n^k}$

3. a. Montrer que : $\forall k \in [\![2; n]\!]$, $\quad P(N = k) = P(N > k-1) - P(N > k)$.

   b. Calculer $P(N = n+1)$ puis en déduire la loi de N.

4. Montrer que l'espérance $E(N)$ de la v.a.r. N est $\displaystyle\sum_{k=0}^{n} \frac{A_n^k}{n^k}$

Proposition de corrigé :

1. Il faut au moins moins deux tirages pour obtenir deux fois la même boule. D'autre part, on peut tirer n boules distinctes lors des n premiers tirages et la $n+1$ -ème est nécepssairement un doublon, et cela peut être le premier doublon. Tous les cas intermédiaires sont possibles et $\boxed{N(\Omega) = [\![2; n+1]\!]}$

2. Soit $k \in [\![1; n]\!]$.
   $\boxed{\text{V1}}$ **Par « reformulation »**
   $[N > k] = \bigcap_{i=1}^{n} [N > i]$
   Ainsi, d'après la formule des probabilités composées (en remarquant que la suite des $([N > j])_j$ est décroissante au sens de l'inclusion,
   $$P(N > k) = P(N > 1) P_{[N>1]}(N > 2) P_{[N>2]}(N > 3) \ldots P_{[N>k-1]}(N > k)$$
   Or, $P(N > 1) = 1$ et, pour tout $i \in [\![1; k-1]\!]$, si on suppose réalisé l'événement $[N > i]$, alors i boules distinctes ont été tirées lors des i premiers tirages, et $[X > i+1]$ est réalisé si on tire une des $(n-i)$ boules distinctes des i premières parmi les n boules de l'urne. On obtient donc :
   $$P(N > k) = 1 \prod_{i=1}^{k-1} \frac{n-i}{n} = \frac{n \prod_{i=1}^{k-1}(n-i)}{n^k} = \frac{A_n^k}{n^k}$$
   $\boxed{\text{V2}}$ **Par dénombrement :**
   L'univers Ω est l'ensemble des $2n$-listes de $[\![1; n]\!]$. Il y a équiprobabilité et
   $$\text{Card}(\Omega) = n^{2n}.$$
   D'autre part, $[N > k]$ est réalisé ssi les k premiers tirages ont des numéros distincts. Or, il y a A_n^k façons de tirer une suite de k numéros distincts dans $[\![1; n]\!]$, et on complète avec une $2n-k$ liste de $[\![1; n]\!]$. Ainsi,
   $$\text{Card}([N > k]) = A_n^k \times n^{2n-k}$$
   Finalement,
   $$P(N > k) = \frac{A_n^k n^{2n-k}}{n^{2n}} = \frac{A_n^k}{n^k}$$

3. a. Soit $k \in [\![2; n]\!]$. $[X > k-1] = [X = k] \cup [X > k]$. La réunion portant sur des événements incompatibles, on a : $P(X > k-1) = P(X = k) + P(X > k)$ et donc $\boxed{P(N = k) = P(N > k-1) - P(N > k)}$.

   b. $P(N = n+1) = P(X > n)$. En effet, si les n premiers tirages sont distincts, le $n+1$ sera nécessairement égal à l'un des n premiers.
   D'où $P(X = n+1) = \dfrac{A_n^n}{n^n} = \dfrac{n!}{n^n}$
   $\forall k \in [\![2; n]\!]$, $P(X = k) = \dfrac{A_n^{k-1}}{n^{k-1}} - \dfrac{A_n^k}{n^k}$ d'après 3.a) et 2.

   > Avec la convention $A_n^k = 0$ si $k > n$, on peut généraliser la formule...

4. Le raisonnement qui suit est très classique, à retenir !
   $$E(N) = \sum_{k=2}^{n+1} k P(N = k)$$
   $$E(N) = (n+1) P(N = n+1) + \sum_{k=2}^{n} k \left(P(N > k-1) - P(N > k) \right)$$
   $$E(N) = (n+1) P(N = n+1) + \sum_{k=2}^{n} k P(N > k-1) - \sum_{k=2}^{n} k P(N > k)$$

$E(N) = (n+1)\frac{A_n^n}{n^n} + \sum_{j=1}^{n-1}(j+1)P(N>j) - \sum_{k=2}^{n} kP(N>k)$ avec le changement d'indice $j = k-1$

$E(N) = (n+1)\frac{A_n^n}{n^n} + \sum_{j=1}^{n-1}P(N>j) + \sum_{j=1}^{n-1} jP(N>j) - \sum_{k=2}^{n} kP(N>k)$

$E(N) = (n+1)\frac{A_n^n}{n^n} + \sum_{j=1}^{n-1}P(N>j) + 1P(N>1) - nP(X>n)$

$E(N) = (n+1)\frac{A_n^n}{n^n} + \sum_{j=1}^{n-1}P(N>j) + 1 - n\frac{A_n^n}{n^n}$

$E(N) = \frac{A_n^n}{n^n} + \sum_{k=1}^{n-1}\frac{A_n^k}{n^k} + \frac{A_n^0}{n^0}$

$\boxed{E(N) = \sum_{k=0}^{n}\frac{A_n^k}{n^k}}$

Bienaymé-Tchebychev facile

exercice 11 Un dé à six faces s'arrête sur six avec une probabilité $p \in \,]0;1[$ à chaque lancer. On le lance indéfiniment, et on note X_n la v.a.r. égale au nombre de fois où le six est sorti au cours des $6n$ premiers lancers ($n \in \mathbb{N}$).

1. Donner la loi de X_n, son espérance, sa variance.
2. Donner une valeur de n pour laquelle on a plus d'une chance sur deux d'obtenir une fréquence d'apparition du six qui s'écarte de moins de 10^{-2} de la valeur $\frac{1}{6}$.
3. Écrire un script Python qui simule $N = 10\,000$ tirages et qui évalue la fréquence d'apparition du six si le dé est honnête $p = \frac{1}{6}$. Comparer la fréquence obtenue à la valeur obtenue précédemment.
4. On lance 6000 fois un dé et on obtient 1105 fois le nombre 6. Est-ce normal ?

Proposition de corrigé :

1. On répète $6n$ fois et indépendemment une même épreuve de Bernoulli de probabilité de succès $\frac{1}{6}$ et X_n compte le nombre de succès obtenus. Ainsi, $X_n \hookrightarrow \mathcal{B}(6n, \frac{1}{6})$.
   D'après le cours, $E(X_n) = 6n\frac{1}{6} = n$ et $V(X_n) = 6n\frac{1}{6}\frac{5}{6} = \frac{5n}{6}$.
2. Soit $n \in \mathbb{N}^*$. La fréquence d'apparition du six est $F_n = \frac{X_n}{6n} = \frac{1}{6n}X_n$.
   Or, $E(F_n) = \frac{1}{6n}E(X_n) = \frac{1}{6}$ par linéarité de l'espérance.
   De plus, $V(F_n) = \left(\frac{1}{6n}\right)^2 V(X_n) = \frac{1}{6^2n^2} \times \frac{5n}{6} = \frac{5}{6^3n}$.
   Ainsi,
   $$P\left(\left|F_n - \frac{1}{6}\right| < 10^{-2}\right) = 1 - P(|F_n - E(F_n)| \geqslant 10^{-2})$$
   Or, avec l'inégalité de Bienaymé-Tchebychev
   $$P(|F_n - E(F_n)| \geqslant 10^{-2}) \leqslant \frac{V(F_n)}{(10^{-2})^2} = \frac{\frac{5}{6^3n}}{10^{-4}} = \frac{5}{6^3.10^{-4}.n}$$
   Ainsi,
   $$P\left(\left|F_n - \frac{1}{6}\right| < 10^{-2}\right) > \frac{1}{2} \iff P(|F_n - E(F_n)| \geqslant 10^{-2}) < \frac{1}{2}$$
   Ainsi, pour $\frac{5}{6^3.10^{-4}.n} < \frac{1}{2}$, c'est à dire si $n > \frac{10^5}{6^3}$, on aura la probabilité voulue.
   Pour $n > 463$ (c'est dire pour plus de 2778 lancers), on a plus d'une chance sur deux d'obtenir une fréquence d'apparition du six qui s'écarte de moins de 10^{-2} de la valeur $\frac{1}{6}$.

> En pratique, on trouverait une valeur plus faible de n par calcul direct avec « BinomialFrep » on un script Python en reproduisant l'effet.

3.

```
import random as rd
N = 10000
C = 0
for k in range(N) :
    if rd.randint(1,6)==6 :
        C += 1
print("la fréquence d'apparition du 6 est : ", C/N)
```

L'écart à $\frac{1}{6}$ est de l'ordre de 4 millièmes, soit bien moindre que $10{*}{*}(-2)$. Même en ramenant $N = 2778$ et en répétant la simulation, on s'aperçoit que l'écart à l'espérance ne dépasse jamais 10^{-2} et on peut conjecturer que la probabilité d'avoir un écart de plus de 10^{-2} est (bien) plus élevée que $\frac{1}{2}$.

Pour se fixer les idées, voici un code qui estime statistiquement la probabilité d'obtenir un écart de moins de 10^{-2} si $n = 463$ (soit 2778 tirages).

```
import random as rd
N = 1000
K = 0
for k in range(N) :
    C = 0
    for j in range(2778) :
        if rd.randint(1,6)==6 :
            C += 1
    if abs(C/N-1/6)>10**(-2) :
        K += 1
print('la probabilité cherchée est environ : ', K/N)
```

On trouve environ $0,83$, soit bien plus que $\frac{1}{2}$.

4. On lance 6000 fois un dé. Alors $n = 1000 > 463$. L'écart à $\frac{1}{6}$ de la fréquence d'apparition du six est $\frac{1105}{6000} - \frac{1}{6} = 0,0175 > 10^{-2}$.
   On peut supposer que dé n'est pas bien équilibré puisque l'estimation précédente est « large » et que l'on dépasse le nombre minimum de tirage requis...

   En deuxième année, la notion d'intervalle de confiance permet de préciser les choses.

exercice 12

1. Soit $N \in \mathbb{N}^*$. Un joueur fait face à une porte fermée à clé. On donne au joueur un trousseau de N clés, et une seule des clés du trousseau ouvre la porte. Le joueur essaie les clés une à une jusqu'à trouver la bonne. Il paye un euro par essai et gagne 10 euros quand il parvient à ouvrir la porte.

   Pour $k \in [\![1; N]\!]$, on note A_k l'événement « à la k-ième tentative, le joueur essaie une clé et ouvre la porte » ; on note T le numéro de l'essai où le joueur trouve la bonne clé.

   a. Exprimer, pour chaque $k \in [\![1; N]\!]$, l'événement $[T = k]$ à l'aide des événements $A_1; A_2; ...; A_k$, puis en déduire la valeur de $P(T = k)$. Identifier la loi de T.

   b. Le joueur n'accepte de jouer que si son espérance de gain est positive ou nulle. Quelles valeurs de N lui conviennent ?

2. Soit $M \in \mathbb{N}^*$. On change le déroulement du jeu. Cette fois l'organisateur fait jouer M joueurs en même temps. Chaque joueur est face à une porte (chacun la sienne), et chaque joueur a droit à seulement un essai de clé pour ouvrir sa porte (on lui donne toujours N clés, dont une seule convient). Le prix pour jouer est le même, ainsi que le gain en cas d'ouverture. On note Y la variable aléatoire qui renvoie le nombre de joueurs qui ouvrent leur porte.

   a. Identifier la loi de Y

   b. On se place du point de vue de l'organisateur. Quelle est son espérance de gain ?

Proposition de corrigé :

1. a. Soit $k \in [\![1; N]\!]$. $[T = k] = \left(\bigcap_{i=1}^{k-1} \overline{A_k}\right) \cap A_k$

   On peut supposer que le joueur prendra une clé différente à chaque essai. Alors, d'après la formule des probabilités composées, on a :

   $P(T = k) = P(\overline{A_1}) \times P_{A_1}(\overline{A_2}) \times ... \times P_{\overline{A_1} \cap ... \cap \overline{A_{k-2}}}(\overline{A_{k-1}}) \times P_{\overline{A_1} \cap ... \cap \overline{A_{k-1}}}(A_k)$.

   Pour $i \in [\![1; k-2]\!]$, si on suppose réalisé l'événement $\overline{A_1} \cap ... \cap \overline{A_i}$, alors i clés ont été essayées, il reste $N - i$ clés parmi lesquelles 1 ouvre la porte et $N - i - 1$ ne l'ouvrent pas. Ainsi,

   $P(T = k) = \frac{N-1}{N} \times \frac{N-2}{N-1} \times ... \times \frac{N-i-1}{N-i} \times ... \times \frac{N-(k-1)}{N-k-2} \frac{1}{N-(k-1)}$

   $P(T = k) = \frac{1}{N}$

   Ici, on a déjà utilise les pointillés dans la formule des probabilités composées, on peut donc continuer. Si on souhaite des notations plus « carrées », on peut noter

   $P(T = k) = \left(\prod_{i=0}^{N} \frac{N - i - 1}{N - i}\right) \times \frac{1}{N-k+1}$, *puis reformuler en quotient de deux produits et simplifier le produit télescopique par un changement d'indice.*

   $$\boxed{\forall k \in [\![1; N]\!], P(T = k) = \frac{1}{N}. \text{ Ainsi, } T \hookrightarrow \mathcal{U}([\![1; N]\!])}$$

   b. On suppose ici que le joueur joue jusqu'à trouver la bonne porte. Notons G la variable aléatoire qui, à une partie associe le gain du joueur. Puisqu'il joue jusqu'à trouver la clé, il gagnera de toutes façons 10€. Il perdra également un euro par essai. Ainsi, $G = 10 - T$.

   Par linéarité de l'espérance, on a $\mathbb{E}(G) = 10 - \mathbb{E}(T) = 10 - \frac{N+1}{2}$

   Le joueur jouera ssi $\mathbb{E}(G) \geqslant 0$ donc ssi $10 - \frac{N+1}{2} \geqslant 0$, càd ssi $N \leqslant 19$.

   $\boxed{\text{Le joueur jouera ssi } N \leqslant 19}$

2. a. Soit $M \in \mathbb{N}^*$. Les M tirages sont des épreuves de Bernoulli indépendantes et de même probabilité de succès $\frac{1}{N}$.
$\boxed{\text{la v.a.r } Y \text{ comptant le nombre de gagnants suit donc une loi binômiale } \mathcal{B}(M, \frac{1}{N})}$

   b. On sait que $\mathbb{E}(Y) = \frac{M}{N}$. L'organisateur gagne 1€ par joueur soit M euros et perd 10€ par gagnant, c'est à dire $10Y$ euros. La v.a.r H décrivant le gain de l'organisateur vérifie donc $H = M - 10Y$ et, par linéarité de l'espérance, $\mathbb{E}(H) = M - 10\mathbb{E}(Y) = M - \frac{10M}{N}$
$\boxed{\text{L'espérance de gain de l'organisateur est } M - \frac{10M}{N}}$

---

exercice 13 En une semaine, un changeur de monnaie a distribué 1000 pièces de monnaie dont 50 sont fausses. Guillaume a reçu 15 pièces de ce changeur. Donner une valeur approchée de la probabilité qu'au moins 3 de ces pièces soient fausses.

Proposition de corrigé :Le nombre de fausses pièces de monnaie reçues par Guillaume est décrit par une variable aléatoire X suivant une loi hypergéométrique $\mathcal{H}(1000, 15, \frac{50}{1000})$.

$$P(X \geqslant 3) = 1 - P(X < 3) = 1 - P(X = 0) - P(X = 1) - P(X = 2)$$

En approchant la loi de X par une loi binômiale $\mathcal{B}(15, \frac{50}{1000})$ (ce qui a un sens car on ne s'intéresse qu'à des évéments $[X = k]$ avec k « petit » par rapport à 50), on obtient la valeur approchée :

$$P(X \geqslant 3) \approx 1 - \binom{15}{0}\left(\frac{5}{1000}\right)^0\left(\frac{995}{1000}\right)^{15} - \binom{15}{1}\left(\frac{5}{1000}\right)^1\left(\frac{995}{1000}\right)^{14} - \binom{15}{2}\left(\frac{5}{1000}\right)^2\left(\frac{995}{1000}\right)^{13}$$

Finalement

$$\boxed{P(X \geqslant 3) \approx 5,44.10^{-5}}$$

INÉGALITÉ BIENAYMÉ TCHEBYCHEV

exercice 14 On choisit au hasard 1000 individus de la population française dont on supposera qu'elle comporte autant d'hommes que de femmes. Soit r la probabilité d'obtenir un nombre d'hommes entre 450 et 550 strictement.

1. Minorer r en utilisant l'inégalité de Bienaymé-Tchebychev .

2. Donner une expression exacte de r, et faire l'application numérique grâce à la machine.

Proposition de corrigé :

1. Notons X la variable aléatoire comptant le nombre d'hommes dans l'échantillon de 1000 personnes. Alors $X \hookrightarrow \mathcal{B}(1000, \frac{1}{2})$ (comme ce n'est pas le but de la question, on peut passer vite sur ce point)

   Ainsi, $E(X) = 1000 \times \frac{1}{2} = 500$ et $V(X) = 1000 \times \frac{1}{2} \times (1 - \frac{1}{2}) = 250$. Or, on a :
$$r = P(450 < X < 550)$$
$$r = P(|\,X - 500\,| < 50)$$
$$r = 1 - P(|\,X - E(X)\,| \geqslant 50)$$
   D'après l'inégalité de Bienaymé-Tchebychev,
$$P(|\,X - E(X)\,| \geqslant 50) \leqslant \frac{V(X)}{50^2} = \frac{250}{2500} = 0,1$$
   Ainsi,
$$\boxed{r \geqslant 0,9}$$

2. L'expression exacte de r est donnée par
$$r = P(450 < X \leqslant 549)$$
$$r = \sum_{k=451}^{5} 49P(X = k)$$
$$r = F(549) - F(450) \text{où } F \text{ est la fonction de répartition de } X$$
   On réalise l'application numérique avec `BinomialFrep` dans la calculatrice ou à l'aide d'un code Python.

   On trouve $0,9983$, soit une probabilité bien plus proche de 1 que la minoration obtenue avec l'inégalité de BT, réputée « très large ».

exercice 15 ★★★★★

[...] Pour ce dernier jeu, le participant lance successivement n boules au hasard dans N cases numérotées de 1 à N avec $N \geq 2$. On suppose que les différents lancers de boules sont indépendants et que la probabilité pour qu'une boule quelconque tombe dans une case donnée est $\frac{1}{N}$. Une case peut contenir plusieurs boules.

Le gain étant fonction du nombre de cases atteintes, on étudie la variable aléatoire T_n égale au nombre de cases non vides à l'issue des n lancers.

1. Déterminer en fonction de n et de N les valeurs prises par T_n.
2. Donner les lois de T_1 et T_2.
3. Déterminer, lorsque $n \geq 2$, la probabilité des événements $[T_n = 1]$, $[T_n = 2]$, $[T_n = n]$. (pour la dernière probabilité on distinguera deux cas $n > N$ et $n \leq N$).
4. A l'aide de la formule des probabilités totales, justifier l'égalité (I) suivante, pour tout entier k tel que $1 \leq k \leq n$,
$$(I) \qquad P(T_{n+1} = k) = \frac{k}{N} P(T_n = k) + \frac{N-k+1}{N} P(T_n = k-1)$$
5. Afin de calculer l'espérance $E(T_n)$ de la variable T_n, on considère la fonction polynômiale G_n définie par :
$$\forall x \in \mathbb{R}, \quad G_n(x) = \sum_{k=1}^{n} P(T_n = k) x^k$$
   a. Quelle est la valeur de $G_n(1)$?
   b. Exprimer $E(T_n)$ en fonction de $G'_n(1)$.
   c. En utilisant la relation (I), montrer que :
$$\forall x \in \mathbb{R}, \quad G_{n+1}(x) = \frac{1}{N}(x - x^2) G'_n(x) + x G_n(x)$$
   d. En dérivant l'expression précédente, en déduire que :
$$E(T_{n+1}) = (1 - \frac{1}{N}) E(T_n) + 1$$
   e. Prouver enfin que l'espérance de la variable T_n est donnée par :
$$E(T_n) = N\left[1 - \left(1 - \frac{1}{N}\right)^n\right]$$

<u>Proposition de corrigé</u> :(Ecricome ECE 2008)

1. Si $n <= N$, $T_n(\Omega) = [\![1; n]\!]$. Si $n > N$, $T_n(\Omega) = [\![1; N]\!]$.
   On peut résumer en $T_n(\Omega) = [\![1; \min(n, N)]\!]$

2. T_1 est la loi certaine égale à 1. $T_2(\Omega) = \{1; 2\}$. $[T_2 = 1]$ est réalisé ssi la deuxième boule arrive dans la même case que la première.
   Le second lancer étant supposé indémendant du premier, on a $P(T_2 = 1) = \frac{1}{N}$.
   On en déduit $P(T_2 = 2) = 1 - \frac{1}{N}$.

3. Soit $n \geq 2$, L'univers Ω comporte autant d'issues que de n-listes d'éléments de $[\![1;N]\!]$, soit N^n énéments.
$[T_n = 1]$ est réalisé par N issues. En effet, toutes les boules sont alors dans une des N cases.
Ainsi, $P(T_n = 1) = \dfrac{N}{N^n} = \dfrac{1}{N^{n-1}}$

$[T_n = 2]$ est réalisé si chacune des n boules arrive dans exactement deux cases. Il y a $\binom{N}{2}$ façons de choisir ces deux cases puis on compte les n-listes de ces deux cases, en enlevant les deux issues pour lesquelles tous les tirages sont dans la même case.
Ainsi, $P(T_n = 2) = \dfrac{\binom{N}{2}(2^n - 2)}{N^n} = \dbinom{N}{2}\dfrac{2^n - 2}{N^n}$
Si $n > N$, $[T_n = n]$ est un événement impossible car il n'y a pas assez de cases. $P(T_n = n) = 0$.
Supposons maintenant $n \leq N$. Alors, $[T_n = n]$ comporte autant d'issues que de n-listes d'éléments distincts de $[\![1;N]\!]$, soit A_N^n
Ainsi, $P(T_n = n) = \dfrac{A_N^n}{N^n}$

4. Soit k tel que $1 \leq k \leq n$.
Les $([T_n = i])_{1 \leqslant n}$ sont de probabilité non nulles et forment un système complet d'événements. La formule des probabilités totales donne donc :

$$P(T_{n+1} = k) = \sum_{i=1}^{n} P(T_n = i) P_{[T_n = i]}(T_{n+1} = k)$$

Or, on ne peut avoir k cases occupées après $n+1$ tirages que si k cases sont déjà occupées après n tirages (et que l'on place la $n+1$-ème boule dans une des k cases déjà occupées) ou si $k-1$ cases sont occupées après n tirages et que la n-ième est placée dans une des $N-(k-1)$ cases vides.

Ainsi, si $i \notin \{k-1, k\}$, $P_{[T_n = i]}(T_{n+1} = k) = 0$ et on obtient :

$$P(T_{n+1} = k) = P(T_n = k) P_{[T_n = k]} P(T_{n+1} = k) + P(T_n = k-1) P_{[T_n = k-1]} P(T_{n+1} = k)$$
$$P(T_{n+1} = k) = \frac{k}{N} P(T_N = k) + \frac{N-k+1}{N} P(T_n = k-1)$$

$\boxed{\text{Ceci prouve la relation } (I)}$.

5.

a. $G_n(1) = \displaystyle\sum_{k=1}^{n} P(T_n = k) 1^k = \sum_{k=1}^{n} P(T_n = k) = 1$

b. $\forall x \in \mathbb{R}, \quad G'_n(x) = \displaystyle\sum_{k=1}^{n} P(T_n = k) k x^{k-1}$.
En particulier, $G'_n(1) = \displaystyle\sum_{k=1}^{n} P(T_n = k) k = E(T_n)$

c. Soit $x \in \mathbb{R}$. $G_{n+1}(x) = \sum_{k=1}^{n+1} P(T_{n+1} = k)x^k$

Montrons que la formule (I) reste valable pour $k = n+1$.
On veut donc prouver :

$$P(T_{n+1} = n+1) = \frac{n+1}{N}P(T_n = n+1) + \frac{N-(n+1)+1}{N}P(T_n = n)$$

Or, on sait que $P(T_n = n+1) = 0$ donc on veut prouver :

$$P(T_{n+1} = n+1) = \frac{N-(n+1)+1}{N}P(T_n = n)$$

• Si $n > N$, $n+1 > N$ donc$P(T_n = n) = P(T_{n+1} = n+1) = 0$
donc l'égalité est vraie.

• Si $n = N$, $n+1 > N$ donc $P(T_{n+1} = n+1) = 0$ et $\frac{N-(n+1)+1}{N} = 0$ donc l'égalité est vraie.

• si $n < N$, alors $n+1 \leqslant N$ donc, d'après 3.,

$$\begin{aligned} P(T_{n+1} = n+1) &= \frac{A_N^{n+1}}{N^{n+1}} \\ &= \frac{(N-n)}{N} \times \frac{A_N^n}{N^n} \\ &= \frac{N-(n+1)+1}{N}P(T_n = n) \end{aligned}$$

Dans tous les cas, (I) est vraie pour $k = n+1$

$$\begin{aligned} G_{n+1}(x) &= \sum_{k=1}^{n+1} P(T_{n+1} = k)x^k \\ &= \sum_{k=1}^{n+1} \left(\frac{k}{N}P(T_N = k) + \frac{N-k+1}{N}P(T_n = k-1)\right)x^k \end{aligned}$$

En remarquant que $P(T_n = n+1 = 0)$, on obtient

$$G_{n+1}(x) = \sum_{k=1}^{n} \frac{k}{N}P(T_N = k)x^k + \sum_{k=1}^{n+1}\left(1 - \frac{(k-1)}{N}\right)P(T_n = k-1)x^k$$

Par linéarité de la somme, on trouve :

$$G_{n+1}(x) = \sum_{k=1}^{n} \frac{k}{N}P(T_N = k)x^k + \sum_{k=1}^{n+1} P(T_n = k-1)x^k - \sum_{k=1}^{n+1} \frac{(k-1)}{N})P(T_n = k-1)x^k$$

Avec le changement d'indice $j = k - 1$, on obtient :

$$G_{n+1}(x) = \frac{x}{N}G_n'(x)+\sum_{j=0}^{n} P(T_n = j)x^{j+1}-\sum_{j=0}^{n} \frac{j}{N}P(T_n = j)x^{j+1}$$

Comme $P(T_n = 0) = 0$, il vient :

$$\begin{aligned} G_{n+1}(x) &= \frac{x}{N}G_n'(x)+x\sum_{j=0}^{n} P(T_n = j)x^j-\frac{x^2}{N}\sum_{j=1}^{n} \frac{j}{N}P(T_n=j)x^{j-1} \\ &= \frac{x}{N}G_n'(x) + xG_n(x) - \frac{x^2}{N}G_n'(x) \\ &= \frac{1}{N}(x - x^2)G_n'(x) + xG_n(x) \end{aligned}$$

d. En dérivant, on trouve, pour tout $x \in \mathbb{R}$:

$$G_{n+1}'(x)=\frac{1}{N}(1-2x)G_n'(x)+\frac{1}{N}(x-x^2)G_n{'}'(x)+G_n(x)+xG_n'(x)$$

En évaluant en 1, on trouve : $E(T_{n+1}) = -\frac{1}{N}E(T_n) + 1 + E(T_n)$

$$\boxed{E(T_{n+1}) = (1 - \frac{1}{N})E(T_n) + 1}$$

e. Le résultat étant donné, on peut le prouver par récurrence sur n.

Toutefois, rédigeons comme s'il n'était pas donné :

On reconnaît une suite arithmético-géométrique d'équation caractéristique $x = \left(1 - \frac{1}{N}\right)x + 1$ qui admet N comme solution. La suite (u_n) de TG : $\forall n \in \mathbb{N}, u_n = E(T_n) - N$ est géométrique de raison $(1 - \frac{1}{N})$ et de premier terme $u_1 = E(T_1) - N = 1 - N$.

Alors, pour tout $n \in \mathbb{N}^*$, on a

$$\begin{aligned} E(T_n) - N &= (1 - N)\left(1 - \frac{1}{N}\right)^{n-1} \\ &= N + (1 - N)\left(1 - \frac{1}{N}\right)^{n-1} \\ &= N\left(1 + \left(\frac{1-N}{N}\right)\left(1 - \frac{1}{N}\right)^{n-1}\right) \\ &= N\left(1 + \left(1 - \frac{1}{N}\right)\left(1 - \frac{1}{N}\right)^{n-1}\right) \end{aligned}$$

$$\boxed{E(T_n) = N\left(1 + \left(1 - \tfrac{1}{N}\right)\left(1 - \tfrac{1}{N}\right)^{n-1}\right)}$$

exercice 16 ★★☆☆☆

Retrouver la formule du binôme de Newton en étudiant la variable aléatoire X suivant une loi $\mathcal{B}\left(n, \frac{a}{a+b}\right)$

Proposition de corrigé : Supposons $X \hookrightarrow \mathcal{B}(n, \frac{a}{a+b})$

Alors, $1 = \sum_{k=0}^{n} P(X=k) = \sum_{k=0}^{n} \binom{n}{k} \left(\frac{a}{a+b}\right)^k \left(\frac{b}{a+b}\right)^{n-k} = \sum_{k=0}^{n} \binom{n}{k} \frac{a^k b^{n-k}}{(a+b)^n} = \frac{1}{(a+b)^n} \sum_{k=0}^{n} \binom{n}{k} a^k b^{n-k}$
d'où le résultat.

exercice 17 ★★★★☆

Une secrétaire effectue n appels pour tenter de joindre n correspondants distincts. Pour chaque appel, elle a une probabilité p d'obtenir son correspondant, et $q = 1-p$ de ne pas le joindre.

1. On note X le nombre de correspondants obtenus. Quelle est la loi de X ? Donner son espérance et sa variance.

2. La secrétaire tente une deuxième fois de joindre les $n - X$ correspondants qu'elle n'a pas pu joindre la première fois. On note Y le nombre de correspondants joints à la deuxième tentative, et $Z = X + Y$. Quelles sont les valeurs que peut prendre Z ?

3. Calculer $\mathbb{P}(Y = h \mid X = k)$ pour les valeurs de k et h pour lesquelles cela a un sens.

4. Pour $\ell \in Z(\Omega)$, calculer $P(Z = \ell)$.

5. Soit n, ℓ, k des entiers naturels tels que $0 \leqslant k \leqslant \ell \leqslant n$. Montrer que
$$\binom{n}{k}\binom{n-k}{\ell-k} = \binom{n}{\ell}\binom{\ell}{k}.$$
En déduire que $P(Z = \ell) = \binom{n}{\ell} p^\ell (1+q)^\ell (q^2)^{n-\ell}$

6. En constatant que $p(1+q) = 1 - q^2$, reconnaitre la loi suivie par Z.

Proposition de corrigé :

1. L'énoncé suggère que le succès d'un appel est indépendant des appels précédents. X compte donc le nombre de succès lors de n répétitions indépendantes d'une épreuve de Bernoulli de même paramètre p. Ainsi,
$$\boxed{X \hookrightarrow \mathcal{B}(n, p)}$$

2. $X(\Omega) = [\![0; n]\!]$. X étant fixé, on a $0 \leqslant X \leqslant n$ et $0 \leqslant Y \leqslant n - X$ donc $0 \leqslant X \leqslant X + Y \leqslant n - X + X$. Ainsi,
$$\boxed{Z(\Omega) = [\![0; n]\!]}$$

3. Pour tout $k \in [\![0; n]\!]$, si on suppose réalisé $[X = k]$, alors pour $P_{[X=k]}$, Y suit une loi binômiale $\mathcal{B}(n-k, p)$
$$\boxed{\text{Pour tous entiers naturels } k, h \text{ tels que } 0 \leqslant k \leqslant k+h \leqslant n, \mathbb{P}(Y = h \mid X = k) = \binom{n-k}{h} p^h q^{n-k-h}}$$

4. La famille $([X = k])_{k \in [\![0;n]\!]}$ forme un système complet d'événements de probabilités non nulles. La formule des probabilités totales donne alors, pour tout $\ell \in [\![0; n]\!]$:

$$\begin{aligned}
P(Z=\ell) &= \sum_{k=0}^{n} P(X=k)\mathbb{P}(Z=\ell \mid X=k) \text{ et, comme } \mathbb{P}(Z=\ell \mid X=k) \text{ si } k>\ell, \\
&= \sum_{k=0}^{\ell} P(X=k)\mathbb{P}(X+Y=\ell \mid X=k) \\
&= \sum_{k=0}^{\ell} P(X=k)\mathbb{P}(Y=\ell-k \mid X=k) \\
&= \sum_{k=0}^{\ell} \binom{n}{k} p^k q^{n-k} \cdot \binom{n-k}{\ell-k} p^{\ell-k} q^{n-k-(\ell-k)} \\
&= \sum_{k=0}^{\ell} \binom{n}{k}\binom{n-k}{\ell-k} p^{\ell} q^{2n-k-\ell}
\end{aligned}$$

$$\boxed{P(Z=\ell) = \sum_{k=0}^{\ell} \binom{n}{k}\binom{n-k}{\ell-k} p^{\ell} q^{2n-k-\ell}.}$$

5. Soit n, ℓ, k des entiers naturels tels que $0 \leqslant k \leqslant \ell \leqslant n$.

$$\begin{aligned}
\binom{n}{k}\binom{n-k}{\ell-k} &= \frac{n!}{k!(n-k)!} \times \frac{(n-k)!}{(\ell-k)!((n-k)-(\ell-k))!} \\
&= \frac{n!}{k!(\ell-k)!(n-\ell)!} \\
&= \frac{n!}{\ell!(n-\ell)!} \times \frac{\ell!}{k!(\ell-k)!} \\
&= \binom{n}{\ell}\binom{\ell}{k}
\end{aligned}$$

$$\begin{aligned}
\text{Ainsi, } P(Z=\ell) &= \sum_{k=0}^{\ell} \binom{n}{k}\binom{n-k}{\ell-k} p^{\ell} q^{2n-k-\ell} \\
&= \sum_{k=0}^{\ell} \binom{n}{\ell}\binom{\ell}{k} p^{\ell} q^{2n-k-\ell} \\
&= \binom{n}{\ell} p^{\ell} q^{2n-2\ell} \sum_{k=0}^{\ell} \binom{\ell}{k} 1^k q^{\ell-k} \\
&= \binom{n}{\ell} p^{\ell} \left(q^2\right)^{n-\ell} (1+q)^{\ell} \text{ avec la formule du binôme}
\end{aligned}$$

$$\boxed{P(Z=\ell) = \binom{n}{\ell} p^{\ell}(1+q)^{\ell}(q^2)^{n-\ell}}$$

6. $p(1+q) = (1-q)(1+q) = 1-q^2$, et on a donc :

$\forall \ell \in [\![0; n]\!], P(Z=\ell) = \binom{n}{\ell}(1-q^2)^{\ell}(q^2)^{n-\ell}$. On reconnaît une loi binômiale et

$$\boxed{Z \hookrightarrow \mathcal{B}(n, 1-q^2)}$$

exercice 18 ★☆☆☆☆

Dans chacune des situations ci-dessous, reconnaître la loi de X.

a. On choisit une carte dans un jeu de 32 cartes numérotées de 1 à 32. X est le numéro de la carte.

b. Un étudiant a dans son placard 20 bouteille d'eau et 10 bouteilles de soda. chaque jour, il emmène une bouteille au hasard. X est le nombre de bouteilles de soda consommées sur 5 jours.

c. On lance un dé à 6 face. X est le numéro obtenu.

d. Un élève répond au hasard au qcm de 10 questions comportant 4 propositions parmi lesquelles exactement une est correcte. X est le nombre de bonnes réponses.

e. On tire sans remise les cartes d'un jeu de 52 cartes. On arrête lors qu'on obtient l'as de cœur. X désigne le nombre de cartes tirées en tout.

f. Une boîte de chocolats contient 20 chocolats au lait et 10 chocolats noirs. On en prélève 3 d'un coup au hasard. X désigne le nombre de chocolats au lait.

g. Pendant chaque cours, Félix regarde sa montre au hasard et note si le nombre de minutes et pair ou impair. X est le nombre de fois où ce nombre de minutes était pair au bout d'une journée de 8 heures.

h. Toto vient d'emménager et ne parvient toujours pas à distinguer les 5 clés de son trousseau. Pour ouvrir la porte, il les essaie une par une jusqu'à trouver la bonne (il ne tests pas deux fois la même clé !). X est son nombre d'essais.

exercice 19 ★★☆☆☆

Une piste rectiligne est divisée en cases numérotées de gauche à droite, chaque case étant représentée par un entier naturel. Une puce se déplace vers la droite de une ou deux cases au hasard avec équiprobabilité à chaque saut. Au départ, elle est sur la case 0.

Soit X_n le numéro de la case occupée par la puce après n sauts et Y_n le nombre de fois où la puce a sauté d'une case au cours des n premiers sauts.

1. Donner la loi, l'espérance et la variance de Y_n.
2. En déduire celles de X_n.
3. Écrire une fonction Python `sauts(n)` simulant les n sauts et renvoyant la position finale de la puce.

Proposition de corrigé :

1. Chaque saut est une épreuve de Bernoulli. En appelant succès le résultat « la longueur du saut est une case », Y_n compte le nombre de succès dans la répétition indépendante de n de ces épreuves de Bernoulli. Ainsi,
$$\boxed{Y_n \hookrightarrow \mathcal{B}(n, \tfrac{1}{2}), \quad \mathbf{E}(Y_n) = \tfrac{n}{2}, \quad \mathbf{V}(Y_n) = \tfrac{n}{4})}$$
2. $X_n = Y_n + 2(n - Y_n) = 2n - Y_n$. Par linéarité de l'espérance, et par propriétés de la variance,
$$\boxed{\mathbf{E}(X_n) = 2n - \tfrac{n}{2} \text{ et } \mathbf{V}(X_n) = \mathbf{V}(Y_n) = \tfrac{n}{4}}$$
3. Écrire une fonction Python `sauts(n)` simulant les n sauts et renvoyant la position finale de la puce.
4. une première proposition proche du déroulement effectif :

```
import random as rd
def saut(n):
    p = 0
    for k in range(1,n+1):
        if rd.random()<0.5 :
            p = p+1
        else:
            p = p+2
    return p
```

une autre plus astucieuse (on avance toujours de 1 (au minimum) et, une fois sur deux, de 1 de plus)

```
import random as rd
def saut(n):
    p = 0
    for k in range(1,n+1):
         p += 1 + int(rd.random()<0.5)
    return p
```

deux autres propositions utilisant l'observation sur les variables aléatoires :

```
import random as rd
def saut(n):
    T = []
    for k in range(n) :
        T.append(int(rd.random()<0.5))
    return 2*n - sum(T)
```

ou

```
import random as rd
def saut(n):
    return 2*n - sum([int(rd.random()<0.5) for k in range(n)])
```

exercice 20 ★★★☆☆

1. ♡ Soit $n \in \mathbb{N}$. Soit X une variable aléatoire telle que $X(\Omega) = [\![0; n]\!]$. Montrer que

$$\mathbf{E}(X) = \sum_{k=1}^{n} P(X \geqslant k)$$

2. **application :** Une urne contient n jetons numérotés de 1 à n. On tire des jetons successivement avec remise jusqu'à obtention d'un numéro inférieur ou égal au précédent. On note X le nombre de jetons tirés.
Par exemple si les tirages donnent $(1, 3, 7, 5)$ alors X prend la valeur 4.

   a. Déterminer $X(\Omega)$.
   b. Calculer $P(X > k)$ pour k compris entre 1 et $n+1$.
   c. En déduire la loi de X.
   d. Calculer $\mathbf{E}(X)$.
   e. Que devient $\mathbf{E}(X)$ lorsque $n \to +\infty$?

Proposition de corrigé :

1. ♡

$$E(X) = \sum_{k=0}^{n} kP(X = k)$$
$$E(X) = \sum_{k=1}^{n} kP(X = k) \text{ car l' terme d'indice 0 est nul}$$
$$E(X) = \sum_{k=1}^{n} k\left(P(X \geqslant k) - P(X \geqslant k+1)\right)$$
$$E(X) = \sum_{k=1}^{n} kP(X \geqslant k) - \sum_{k=1}^{n} kP(X \geqslant k+1) \text{ par linéarité de la somme}$$
$$E(X) = \sum_{k=1}^{n} kP(X \geqslant k) - \sum_{j=2}^{n+1} (j-1)P(X \geqslant j) \text{ avec le changement d'indice } j = k+1$$
$$E(X) = 1P(X \geqslant 1) + \sum_{k=2}^{n} kP(X \geqslant k) - \sum_{k=2}^{n} (k-1)P(X \geqslant k) - (n+1)\underbrace{P(X \geqslant n+1)}_{=0 \text{ car } X(\Omega)=[\![0;n]\!]}$$
$$E(X) = P(X \geqslant 1) + \sum_{k=2}^{n} (k-(k-1))P(X \geqslant k)$$
$$E(X) = \sum_{k=1}^{n} P(X \geqslant k)$$

2. **application :** Une urne contient n jetons numérotés de 1 à n. On tire des jetons successivement avec remise jusqu'à obtention d'un numéro inférieur ou égal au précédent. On note X le nombre de jetons tirés.
Par exemple si les tirages donnent $(1, 3, 7, 5)$ alors X prend la valeur 4.

   a. Il faut tirer au moins deux jetons et après avoir tiré n jetons, le $(n+1)$-ième sera nécessairement égal à un numéro déjà tiré. donc $X(\Omega) \subset [\![2; n+1]\!]$. Réciproquement, soit $k \in [\![2; n+1]\!]$. $[X = k]$ est réalisé par exemple si on tire les numéros 1 à $k-1$ aux $k-1$ premiers tirages puis $k-1$ à nouveau. Donc $[\![2; n+1]\!] \subset X(\Omega)$.
   Par double inclusion, $\boxed{X(\Omega) = [\![2; n+1]\!]}$.
   b. $P(X > n+1) = 0$ car $n+1 \notin X(\Omega)$. Soit $k \in [\![1; n]\!]$.
   $[X > k]$ est réalisé ssi les k premiers numéros tirés le sont par ordre strictement croissant. Chacun de ces tirages est représenté de manière unique[1] par une partie à k éléments de $[\![1; n]\!]$ donc $\mathrm{Card}([X = k]) = \binom{n}{k}$.

[1] on parle de bijection ici

De plus, il y a autant de façons de tirer les k premières boules que de k listes de $[\![1;n]\!]$ soit n^k, et il y a équiprobabilité donc

$$\forall k \in [\![1;n+1]\!], P(X>k) = \frac{\binom{n}{k}}{n^k}$$

On a bien $P(X>n+1)=0$ avec la convention sur les coefficients binomiaux[2]

Ici, on a bien $P(X>k) = P\left(\bigcap_{j=1}^{k})P(X>j\right) = P(X>1)P_{X>1}(X>2)...P_{[X>k-1]}(X>k)$... (méthode très classique à retenir) MAIS il n'est pas simple du tout de déterminer $P_{[X>k-1]}(X>k)$ donc on est obligé de passer par du dénombrement.

c. Soit $k \in X(\Omega) = [\![2;n+1]\!]$.

$$P(X=k) = P(X \geqslant k) - P(X \geqslant k-1)$$
$$P(X=k) = P(X>k-1) - P(X>k)$$
$$P(X=k) = \frac{\binom{n}{k-1}}{n^{k-1}} - \frac{\binom{n}{k}}{n^k}$$

d. X est à valeurs dans $[\![2;n+1]\!]$. Donc $X-2$ est à valeurs dans $[\![0;n-1]\!]$ et avec 1), et

$$\mathbf{E}(X-2) = \sum_{k=1}^{n-1} P(X-2 \geqslant k)$$
$$\mathbf{E}(X)-2 = \sum_{k=1}^{n-1} P(X \geqslant k+2)$$
$$\mathbf{E}(X) = 2 + \sum_{k=1}^{n-1} P(X>k+1)$$
$$\mathbf{E}(X) = 2 + \sum_{k=1}^{n-1} \frac{\binom{n}{k+1}}{n^{k+1}}$$
$$\mathbf{E}(X) = 2 + \sum_{j=2}^{n} \binom{n}{j}\frac{1}{n^j}$$
$$\mathbf{E}(X) = 2 - \binom{n}{0}\frac{1}{n^0} - \binom{n}{1}\frac{1}{n^1} + \sum_{j=0}^{n}\binom{n}{j}\frac{1}{n^j}1^{n-j}$$
$$\mathbf{E}(X) = 0 + \left(1+\frac{1}{n}\right)^n$$
$$\mathbf{E}(X) = 0 + \left(1+\frac{1}{n}\right)^n$$

e. $E(X) = \exp\left(n\ln(1+\frac{1}{n})\right)$.

$\boxed{\text{v1}}$ On sait que $\lim\limits_{x\to+0}\frac{\ln(1+x)}{x} = 1$ donc en posant $x=\frac{1}{u}$, on a $\lim\limits_{u\to+\infty}\frac{\ln(1+\frac{1}{u})}{\frac{1}{u}} = 1$. Ainsi, $\left(1+\frac{1}{n}\right)^n \underset{n\to+\infty}{\longrightarrow} 1$

$\boxed{\text{v2}}$ On sait que $\ln(1+x) \underset{x\to0}{\sim} x$ donc, comme $\lim\limits_{n\to+\infty}\frac{1}{n}=0$, $\left(1+\frac{1}{n}\right)^n \underset{n\to+\infty}{\sim} n\frac{1}{n} \underset{n\to+\infty}{\longrightarrow} 1$

fin commune (on ne passe JAMAIS aux équivalents <u>dans</u> une exponentielle) :

Par continuité de $t \mapsto \mathrm{e}^t$,

$$\boxed{\lim_{n\to+\infty} E(X) = \mathrm{e}.}$$

[2] il n'existe aucune manière de constituer une suite strictement croissantes de $n+1$ nombres de $[\![1;n]\!]$

exercice 21 **un classique** ★★★☆☆

Soit X une variable aléatoire finie. On définit la fonction G_X par :

$$G_X(t) = \mathbf{E}(\mathrm{e}^{tX})$$

1. Que vaut $G_X(0)$? $G'(0)$? $G''(0)$?

2. Exprimer $V(X)$ à l'aide de G.

3. Calculer G_X dans les deux cas suivants :

   a. $X \hookrightarrow \mathcal{U}(\llbracket 1; n \rrbracket)$.

   b. $X \hookrightarrow \mathcal{B}(n, p)$.

4. Retrouver l'expression de $\mathbf{E}(X)$ et $\mathbf{V}(X)$ si $X \hookrightarrow \mathcal{B}(n, p)$.

5. Même question si $X \hookrightarrow \mathcal{U}(\llbracket 1; n \rrbracket)$.

__Proposition de corrigé__ :

1. Notons $X(\Omega) = \{x_1, ..., x_n\}$. Avec le théorème de transfert,
$$G_X(t) = \mathbf{E}(\mathrm{e}^{tX})$$
$$G_X(t) = \sum_{k=0}^{n} \mathrm{e}^{tx_k} P(X = x_k)$$
G_X est dérivable sur $\mathbb{R}$ car les fonctions $t \mapsto \mathrm{e}^{tx_k}$ le sont et
$$G'_X(t) = \sum_{k=0}^{n} x_k \mathrm{e}^{tx_k} P(X = x_k)$$
$$G''_X(t) = \sum_{k=0}^{n} x_k^2 \mathrm{e}^{tx_k} P(X = x_k)$$
Ainsi,
$G_X(0) = \sum_{k=0}^{n} P(X = x_k) = 1$
$G'_X(0) = \sum_{k=0}^{n} x_k P(X = x_k) = E(X)$
$G''_X(0) = \sum_{k=0}^{n} x_k^2 P(X = x_k) = E(X^2)$ (avec le théorème de transfert)[3]

2. Avec la formule de Kœnig Huygens, $V(X) = E(X^2) - E(X)^2 = G''_X(0) - (G'_X(0))^2$.

3. Calculer G_X dans les deux cas suivants :

   En reprenant le raisonnement de la question 1,

   a. si $X \hookrightarrow \mathcal{U}(\llbracket 1; n \rrbracket)$.
$$G_X(t) = \sum_{k=0}^{n} \mathrm{e}^{tk} \frac{1}{n}$$
$$G_X(t) = \frac{1}{n} \sum_{k=0}^{n} (\mathrm{e}^{t})^k$$
$$G_X(t) = \begin{cases} 1 & \text{si } t = 0 \\ \frac{1-e^{(n+1)t}}{n(1-\mathrm{e}^t)} & \text{sinon} \end{cases}$$

   b. si $X \hookrightarrow \mathcal{B}(n, p)$.

[3] utilisé dans «l'autre sens» par rapport à «d'habitude»

$$G_X(t) = \sum_{k=0}^{n} \mathrm{e}^{tk} \binom{n}{k} p^k q^{n-k}$$

$$G_X(t) = \sum_{k=0}^{n} \binom{n}{k} \left(p\mathrm{e}^t\right)^k q^{n-k}$$

$$G_X(t) = \left(p\mathrm{e}^t + (1-p)\right)^n \text{ avec la formule du binôme}$$

4. si $X \hookrightarrow \mathcal{B}(n,p)$.

$$G_X(t) = \left(p\mathrm{e}^t + (1-p)\right)^n$$

$$G'_X(t) = np\mathrm{e}^t \left(p\mathrm{e}^t + (1-p)\right)^{n-1}$$

$$G''_X(t) = (np\mathrm{e}^t)^2 \left(p\mathrm{e}^t + (1-p)\right)^{n-2} + p\mathrm{e}^t \left(p\mathrm{e}^t + (1-p)\right)^{n-1}$$

Il vient alors

$$E(X) = G'_X(0) = np\left(p + (1-p)\right)^{n-1} = np$$

$$E(X^2) = G''_X(0) = np(n-1)p\left(p + (1-p)\right)^{n-2} + np\left(p + (1-p)\right)^{n-1} = n(n-1)p^2 + np$$

$$V(X) = E(X^2) - (E(X))^2 = n(n-1)p^2 + np - (np)^2 = n^2p^2 - np^2 + np - n^2p^2 = np(1-p)$$

On retrouve bien les valeurs de $E(X)$ et $V(X)$.

5. si $X \hookrightarrow \mathcal{U}(\llbracket 1;n \rrbracket)$.

La question ne présente pas d'intérêt particulier. Il vaut mieux utiliser G_X et ses dérivées sous forme de sommes, et on retrouve exactement les mêmes calculs que dans le cours...

profondissement

exercice 22 **oral ESCP 2019 (question courte)** ★★★★★

Soient $p \in]0,1[$ et $n \in \mathbb{N}^*$. On pose $q = 1-p$. Soient trois variables aléatoires X, Y, Z définies sur un même espace probabilisé $(\Omega, \mathcal{A}, P)$ qui sont indépendantes, de lois respectives $\mathcal{B}(n,p)$, $\mathcal{B}(n,p)$ et $\mathcal{B}(n,q)$.

On définit la matrice aléatoire M par :

$$\forall \omega \in \Omega,\ M(\omega) = \begin{pmatrix} X(\omega) & Y(\omega) \\ 0 & Z(\omega) \end{pmatrix}$$

Calculer la probabilité que la matrice M soit diagonalisable.

Indication : La notion de diagonalisabilité ne sera vue qu'en deuxième année. Un étudiant en deuxième année saurait que la matrice triangulaire supérieure $\begin{pmatrix} a & b \\ 0 & c \end{pmatrix}$ *est diagonalisable ssi [*$a \neq c$*] ou [*$a = c$ *et* $b = 0$*]*

<u>*Proposition de corrigé*</u> :Notons E l'événement « la matrice $M(\omega)$ est diagonalisable »

$$\begin{aligned} P(E) &= P([X \neq Z] \cup ([X = Z] \cap [Y = 0])) \\ &= P(X \neq Z) + P([X = Z] \cap [Y = 0]) \text{ car } [X = Z] \text{ et } [X \neq Z] \text{ sont incompatibles} \\ &= 1 - P(X = Z) + P(X = Z)P(Y = 0) \text{ car } X, Y, Z \text{ sont mutuellement indépendantes.} \\ &= 1 + P(X = Z)(P(Y = 0) - 1) \end{aligned}$$

D'une part, $P(Y = 0) = p^n$

D'autre part, la formule des probabilités totales avec le système complet d'événements $\Big([X = k]\Big)_{0 \leqslant k \leqslant n}$ donne :

$$\begin{aligned} P(X = Z) &= \sum_{k=0}^{n} P([X = k] \cap [X = Z]) \\ &= \sum_{k=0}^{n} P([X = k] \cap [Z = k]) \\ &= \sum_{k=0}^{n} P(X = k) \times P(Z = k) \text{ par indépendance de } X \text{ et } Z \\ &= \sum_{k=0}^{n} \binom{n}{k} p^k q^{n-k} \binom{n}{k} q^k p^{n-k} \\ &= (pq)^n \sum_{k=0}^{n} \binom{n}{k}\binom{n}{k} \end{aligned}$$

En développant l'égalité polynômiale

$$(X+1)^n (X+1)^n = (X+1)^{2n}$$

avec la formule du binôme de Newton

$$\sum_{k=0}^{n}\sum_{j=0}^{n} \binom{n}{k}\binom{n}{j} X^{k+j} = \sum_{i=0}^{2n} \binom{2n}{i} X^i$$

En identifiant les coefficients de X^n, on trouve :

$$\sum_{j+k=n} \binom{n}{k}\binom{n}{j} = \binom{2n}{n}$$

Autrement dit,

$$\sum_{k=0}^{n} \binom{n}{k}\binom{n}{n-k} = \binom{2n}{n}$$

Ainsi, comme $\binom{n}{k} = \binom{n}{n-k}$

$$P(X = Y) = \binom{2n}{n}(pq)^n$$

Finalement,

$$\boxed{P(E) = 1 + \binom{2n}{n}(pq)^n\left(p^n - 1\right)}$$

Pour les curieux, un programme qui permet de calculer une estimation de la probabilité cherchée en calculant la fréquence de matrices diagonalisables sur un échantillon de teille 10000 et de comparer la fréquence à la valeur théorique

```
import numpy as np
import numpy.random as rd
n = 10
p = 0.5
q = 1-p
N = 10000
S = 0
for k in range(1,N+1):
    a = rd.binomial(n,p)
    b = rd.binomial(n,p)
    c = rd.binomial(n,q)
    S = S + int((a!=c)or((a==c)and(b==0)))
f = S/N
r = 1+ np.math.factorial(2*n)/(np.math.factorial(n)**2)*((p*q)**n)*(p**n-1)

print(f,r)
```

exercice 23 **oral ESCP 2018 - S -3.11** ★★★★★

Une urne contient exclusivement des boules rouges et noires indiscernables au toucher.

La proportion de boules rouges est $p \in]0,1[$ et celle de boules noires est $q = 1 - p$.

On effectue des tirages successifs avec remise d'une boule de l'urne jusqu'à obtention d'une boule rouge.

Un maximum de n tirages, avec $n \geq 1$, est cependant fixé : on décide de s'arrêter si on n'a pas tiré de boule rouge à l'issue du n-ième tirage.

On note G_n la variable aléatoire égale au rang du tirage d'une boule rouge, si ce rang existe, et qui vaut 0 si aucune boule rouge n'est apparue au cours des n tirages.

1. Dans cette question, n est un entier fixé.
   a. Déterminer la loi de la variable aléatoire G_n.
   b. En utilisant la fonction $x \mapsto \displaystyle\sum_{k=1}^{n} x^k$, montrer que l'on a : $E(G_n) = \dfrac{1 - q^n(1 + np)}{p}$.
2. a. Déterminer la limite de $(E(G_n))_{n\in\mathbb{N}}$ quand n tend vers $+\infty$.
   Interpréter le résultat *(interprétation plus faisable en deuxième année).*
   b. Étudier la convergence en loi de la suite de variables aléatoires (G_n)
   (*cf deuxième année : il s'agit de calculer* $\lim\limits_{n\to+\infty} P(G_n = k)$ *pour tout* $k \in \mathbb{N}$.)
3. Ce processus peut être considéré comme une partie qui est gagnée si une boule rouge est apparue au cours des n tirages.
   On mise de la manière suivante : pour jouer une partie, il faut payer 15 euros. Si la partie est gagnée à la $k$ème boule tirée ($k \leq n$), le joueur gagne $(20 - k)$ euros. On note B_n le gain aléatoire pour une partie.
   a. Exprimer B_n en fonction de G_n, puis déterminer son espérance.
   b. On admet que pour tout réel $x \in]0,1[$ on a $\dfrac{1}{2} < \dfrac{1}{x} + \dfrac{1}{\ln(1-x)} < 1$.
   Quelle valeur de n doit-on choisir afin de maximiser le gain d'une partie ?

Proposition de corrigé :

1. a) On a ici la loi géométrique tronquée :

$$P(G_n = k) = \begin{cases} q^{k-1}p & \text{si } 1 \leq k \leq n \\ q^n & \text{si } k = 0 \end{cases}$$

b) On a alors $E(G_n) = p\sum_{k=1}^{n} kq^{k-1}$. Soit la fonction

$$f(x) = \sum_{k=1}^{n} x^k = \frac{x - x^{n+1}}{1-x}.$$

On dérive

$$f'(x) = \sum_{k=1}^{n} kx^{k-1} = \frac{1-(n+1)x^n + nx^{n+1}}{(1-x)^2}$$

D'où :

$$E(G_n) = pf'(q) = \frac{1-(n+1)q^n + nq^{n+1}}{(1-q)^2} = \frac{1-q^n(1+np)}{p}$$

2. a) On a $n = o(r^n)$ pour $r > 1$, en particulier pour $r = \frac{1}{q} > 1$

d'où : $\lim_{n\to+\infty} nq^n = 0$, et $\lim_{n\to+\infty} E(G_n) = \frac{1}{p}$ qui est l'espérance d'une loi géométrique.

b) Soit un entier $k \geq 1$, on a : $\lim_{n\to+\infty} P[G_n = k] = q^{k-1}p$ (dès que $n > k$)

d'où la convergence en loi vers la loi géométrique.

3. a) On a $B_n = (5 - G_n \cdot \mathbf{1}_{G_n\neq 0}) - 15 \cdot \mathbf{1}_{G_n=0}$, ce qui est une variable aléatoire en tant que fonction de la variable aléatoire discrète G_n. Le calcul donne

$$\begin{aligned} E(B_n) &= \sum_{k=1}^{n}(5-k)P[G_n = k] - 15P[G_n = 0] \\ &= 5(1 - P[G_n = 0]) - E(G_n) - 15P[G_n = 0] \\ &= 5 - E(G_n) - 20P[G_n = 0] \\ &= 5 - \frac{1-q^n(1+np)}{p} - 20q^n \end{aligned}$$

c) Soit g la fonction : $g(x) = 5 - \frac{1-q^x(1+xp)}{p} - 20q^x = 5 - \frac{1}{p} + \frac{q^x}{p} + xq^x - 20q^x$.

Alors $g'(x) = \left[\frac{\ln(q)}{p}) + 1 + x\ln(q) - 20\ln(q)\right] q^x$, et

$$g'(x) \geq 0 \iff \frac{1}{p} + \frac{1}{\ln(q)} + x - 20 \leq 0 \iff x \leq 20 - \frac{1}{p} - \frac{1}{\ln(q)}$$

Ainsi le maximum est obtenu pour $x = 20 - \frac{1}{p} - \frac{1}{\ln(q)}$ or on a vu : $0,5 < \frac{1}{p} + \frac{1}{\ln(q)} < 1 \Rightarrow 19 < x < 19.5$.

Finalement $n = 19$ est la valeur entière la plus proche du maximum.

exercice 24 **oral ENSAE 2013** ★★★★☆

Soit n un entier naturel non nul. Une boite contient $(2n+1)$ jetons bicolores (chaque jeton admet une face rouge et l'autre bleue).

On lance simultanément tous les jetons et on observe leurs faces supérieures.

1. Justifier qu'une, et une seule des deux couleurs apparaît un nombre impair de fois.

   On notera X la variable aléatoire associée à ce nombre.

2. Donner la loi de X.

3. Calculer son espérance et sa variance.

Proposition de corrigé :

1. Notons R le nombre de faces rouges et B le nombre de faces bleues.
   Par l'absurde :
   • Si R et B ont même parité (tous deux pairs ou tous deux impairs), alors $2n+1 = N+B$ est pair, absurde.
   Ainsi, exactement un des nombres R ou B est impair.

2. $X(\Omega) = \{2k+1 \mid k \in [\![0;n]\!]\}$.

   > on peut justifier oralement, droit dans les yeux :
   > En effet, X étant impair l'inclusion $\subset$ est évidente et pour tout $k \in [\![0;n]\!]$, $[X = 2k+1]$ est réalisé si on a par exemple $2k+1$ faces rouges et $2n-2k$ faces bleues.

   Soit $k \in [\![0;n]\!]$. $[X = 2k+1] = [R = 2k+1] \cup [B = 2k+1]$
   Or, si on appelle succès le fait d'obtenir une face rouge, R compte le nombre de succès dans la répétition de $2n+1$ épreuves de Bernoulli inrépendantes et de même paramètre $\frac{1}{2}$. Aussi, $R \hookrightarrow \mathcal{B}\left(2n+1, \frac{1}{2}\right)$
   De même, $B \hookrightarrow \mathcal{B}\left(2n+1, \frac{1}{2}\right)$ (R est la loi binômiale du nombre d'échecs)
   Finalement, par incompatibilité des événements,
   $P(X = 2k+1) = P(R = 2k+1) + P(B = 2k+1)$
   $$P(X = 2k+1) = \binom{2n+1}{2k+1}\left(\frac{1}{2}\right)^k\left(\frac{1}{2}\right)^{2n+1-k} + \binom{2n+1}{2k+1}\left(\frac{1}{2}\right)^k\left(\frac{1}{2}\right)^{2n+1-k} = 2\binom{2n+1}{2k+1}\left(\frac{1}{2}\right)^{2n+1} = \boxed{\frac{1}{2^{2n}}\binom{2n+1}{2k+1}}$$

3. X est une v.a.r finie donc admet une espérance et une variance :
   $$\mathbf{E}(X) = \sum_{k=0}^{n}(2k+1)P(X = 2k+1) = \sum_{k=0}^{n}(2k+1)\frac{1}{2^{2n}}\binom{2n+1}{2k+1} = \frac{1}{2^{2n}}\sum_{k=0}^{n}(2k+1)\frac{2n+1}{2k+1}\binom{2n}{2k} = \frac{2n+1}{2^{2n}}\sum_{k=0}^{n}\binom{2n}{2k}$$
   Or, $2^{2n} = (1+1)^{2n} = \displaystyle\sum_{j=0}^{2n}\binom{2n}{j}$
   et $0 = (1-1)^{2n} = \displaystyle\sum_{j=0}^{2n}\binom{2n}{j}(-1)^j$
   En ajoutant terme à terme, les termes impairs se « compensent » et on trouve : $2^{2n} = 2\displaystyle\sum_{k=0}^{n}\binom{2n}{2k}$ donc $\displaystyle\sum_{k=0}^{n}\binom{2n}{2k} = 2^{2n-1}$
   Ainsi,
   $$\boxed{\mathbf{E}(X) = \frac{2n+1}{2}}$$

   > Le résultat est cohérent (compris entre 0 et $2n+1$) et équilibré entre les valeurs extrêmes, ce qui est typique des situations symétriques. il y a pour chaque jeton autant de chances d'obtenir une face rouge ou bleue.
   > Si vous aviez trouvé autre chose, par exemple $2n+1$, il faut relire les calculs.

$$\begin{aligned}
\mathbf{E}(X^2) &= \sum_{k=0}^{n}(2k+1)^2 P(X=2k+1)\\
&= \frac{2n+1}{2^{2n}}\sum_{k=0}^{n}(2k+1)\binom{2n}{2k} \text{ en procédant comme pour } \mathbf{E}(X)\\
&= \frac{2n+1}{2^{2n}}\sum_{k=1}^{n}2k\binom{2n}{2k} + \frac{2n+1}{2^{2n}}\sum_{k=0}^{n}\binom{2n}{2k}\\
&= \frac{2n+1}{2^{2n}}\sum_{k=1}^{n}2k\frac{2n}{2k}\binom{2n-1}{2k-1} + \frac{2n+1}{2}\\
&= \frac{2n(2n+1)}{2^{2n}}\sum_{k=1}^{n}\binom{2n-1}{2k-1} + \frac{2n+1}{2}
\end{aligned}$$

or, $2^{2n-1} = (1+1)^{2n-1} = \displaystyle\sum_{j=0}^{2n-1}\binom{2n-1}{j}$

et $0 = (1-1)^{2n-1} = \displaystyle\sum_{j=0}^{2n-1}\binom{2n-1}{j}(-1)^j$

Ainsi, en soustrayant, les termes pair se « simplifient » et on obtient :

$2^{2n-1} = 2\displaystyle\sum_{k=1}^{n}\binom{2n-1}{2k-1}$ donc $\displaystyle\sum_{k=1}^{n}\binom{2n-1}{2k-1} = 2^{2n-2}$ et

$\mathbf{E}(X^2) = \dfrac{2n(2n+1)}{4} = \dfrac{n(2n+1)}{2}$

d'où, avec la formule de Kœnig Huygens,

$\mathbf{V}(X) = \mathbf{E}(X^2) - \mathbf{E}(X)^2 = \dfrac{n^2(2n+1)^2}{4} - \dfrac{(2n+1)^2}{4} = \dfrac{(n^2-1)(2n+1)^2}{4}$

$$\boxed{\mathbf{V}(X) = \frac{(n^2-1)(2n+1)^2}{4}}$$