Recueil de questions...

Mélange de démonstrations de cours et de questions classiques

Ce document regroupe quelques démonstrations de résultats qu'il est bon de travailler en vue des concours. Ces questions sont de deux natures :

- démonstration d'un résultat de cours (indiquée par ♥),
- démonstration d'un résultat classique aux écrits ou aux oraux (indiquée par ♠ s'il faut savoir la faire sans indication ou par ♣ si elle serait guidée).

Bien évidemment, il ne faut pas apprendre ces démonstrations... Il faut les travailler, les comprendre : l'objectif étant d'être en mesure de savoir les refaire ou de les adapter à des situations similaires.

Un conseil pour les travailler : pour chacune, résumer les étapes essentielles de la démonstration et les noter. Il faut ensuite se souvenir de ces étapes et être en mesure de gérer en autonomie ce qui se situe entre deux étapes consécutives.

Cette liste est non exhaustive et ne fait que compléter le travail des méthodes et exercices classiques au programme.

Liste des questions classiques

Révisions de première année . . . 3
1 ♥ Formule du binôme de Newton . . . 3
2 ♣ La fameuse somme... . . . 3
3 ♥ Théorème de divergence monotone / Majoration d'une suite croissante convergente . . . 4
4 ♥ Théorème des suites adjacentes . . . 4
5 ♣ Formule de Taylor avec reste intégral . . . 5
6 ♣ Une intégrale classique . . . 6
7 ♣ Sommes de Riemann à gauche (méthode des rectangles à gauche) . . . 7
8 ♣ Série logarithmique . . . 8
9 ♥ Propriétés d'une probabilité . . . 10
10 ♥ Formule des probabilités totales . . . 10
11 ♥ Probabilité conditionnelle . . . 11
12 ♥ Variance d'une VA suivant la loi $\mathscr{G}(p)$ . . . 11
13 ♣ Autour de la trace d'une matrice . . . 12
14 ♣ Matrices stochastiques . . . 13
Chapitre 1 – Compléments sur les suites et les séries . . . 14
15 ♣ Équivalent de la suite des sommes partielles de la série harmonique . . . 14
Chapitre 2 – Couples de variables aléatoires discrètes . . . 15
16 ♥ Stabilité des lois de Poisson . . . 15
17 ♣ Minimum de VA indépendantes suivant une loi géométrique . . . 15
18 ♥ Somme de VA indépendantes de loi de Bernoulli . . . 16
19 ♣ Inégalité de Cauchy-Schwarz pour la covariance . . . 18
20 ♣ Interprétation du coefficient de corrélation linéaire . . . 18
Chapitre 6 – Applications linéaires et matrices . . . 20
21 ♥ Applications linéaires coïncidant sur une base . . . 20
22 ♥ Noyau et image d'une application linéaire . . . 20
23 ♥ Caractérisation de l'injectivité d'une application linéaire . . . 21
24 ♥ Caractérisation des isomorphismes ("triangle de bijectivité") . . . 21
25 ♥ Majoration du rang d'une application linéaire . . . 22
26 ♣ Caractérisation des isomorphismes (bis) . . . 22
27 ♣ Un isomorphisme classique . . . 23
28 ♣ Indice de nilpotence d'un endomorphisme . . . 23
Chapitre 7 – Variables aléatoires à densité . . . 24
29 ♣ Transformée carrée d'une variable aléatoire à densité . . . 24
30 ♣ Expression de $\mathbb{E}(X)$ où X est à densité à valeurs dans $\mathbb{R}^+$ . . . 25
Chapitre 10 – Diagonalisation des matrices carrées . . . 26
31 ♣ Réduction de $U{}^tU$ . . . 26
32 ♣ Valeurs propres et vecteurs propres d'une matrice compagnon . . . 27
Chapitre 11 – Lois à densité usuelles . . . 29
33 ♥ Méthode d'inversion pour la loi exponentielle . . . 29
34 ♣ Moments d'une variable aléatoire suivant la loi $\mathscr{E}(1)$ . . . 29
35 ♣ Parties entière et fractionnaire d'une VA suivant une loi exponentielle . . . 30
36 ♥ Propriétés de la fonction de répartition associée à $\mathscr{N}(0;1)$ . . . 32
37 ♣ Presque une intégrale d'une densité de loi normale... . . . 32
Chapitre 12 – Convergence de suites de variables aléatoires . . . 34
38 ♥ Inégalité de Markov . . . 34
39 ♥ Loi faible des grands nombres . . . 34
40 ♥ Approximation de la loi binomiale par la loi de Poisson . . . 35
41 ♣ Une convergence en loi... . . . 36
Chapitre 15 – Équations différentielles et systèmes différentiels . . . 38
42 ♣ Résolution de $y' + ay = 0$ . . . 38
43 ♥ Structure de l'ensemble des solutions d'une EDL . . . 38

Révisions de première année

1 ♥ Formule du binôme de Newton

Soient $a, b \in \mathbb{R}$.
Pour tout $n \in \mathbb{N}$, on a :

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

☞ Rappel...
Dans sa version matricielle, il est indispensable que les matrices A et B commutent pour l'utiliser.

⋆ Démonstration : Par récurrence...

- **Initialisation.** Pour $n = 0$:
  D'une part :
  $$(a+b)^0 = 1$$
  et d'autre part :
  $$\begin{aligned}\sum_{k=0}^{0} \binom{0}{k} a^k b^{0-k} &= \binom{0}{0} a^0 b^0 \\ &= 1\end{aligned}$$
  L'initialisation est ainsi vérifiée.

- **Hérédité.** Soit $n \in \mathbb{N}$. Supposons que $(a+b)^n = \sum_{k=0}^{n} \binom{n}{k} a^k b^{n-k}$ et montrons que $(a+b)^{n+1} = \sum_{k=0}^{n+1} \binom{n+1}{k} a^k b^{n+1-k}$.

  On a :
  $$\begin{aligned}
  (a+b)^{n+1} &= (a+b)(a+b)^n \\
  &= (a+b)\sum_{k=0}^{n} \binom{n}{k} a^k b^{n-k} && \text{par hypothèse de récurrence} \\
  &= \sum_{k=0}^{n} \binom{n}{k} a^{k+1} b^{n-k} + \sum_{k=0}^{n} \binom{n}{k} a^k b^{n+1-k} \\
  &= \sum_{i=1}^{n+1} \binom{n}{i-1} a^i b^{n-(i-1)} + \sum_{k=0}^{n} \binom{n}{k} a^k b^{n+1-k} && \text{changement d'indice } i = k+1 \text{ dans la première somme} \\
  &= \sum_{i=1}^{n} \binom{n}{i-1} a^i b^{n+1-i} + \binom{n}{n} a^{n+1} b^0 + \binom{n}{0} a^0 b^{n+1} + \sum_{k=1}^{n} \binom{n}{k} a^k b^{n+1-k} && \text{relation de Chasles, licite car } n \geqslant 0 \text{ (si } n = 0\text{, alors les deux sommes sont indexées sur un ensemble vide, donc nulles)} \\
  &= \sum_{k=1}^{n} \left(\binom{n}{k-1} + \binom{n}{k} \right) a^k b^{n+1-k} + \binom{n}{n} a^{n+1} b^0 + \binom{n}{0} a^0 b^{n+1} \\
  &= \sum_{k=1}^{n} \binom{n+1}{k} a^k b^{n+1-k} + \binom{n}{n} a^{n+1} b^0 + \binom{n}{0} a^0 b^{n+1} && \text{relation de Pascal} \\
  &= \sum_{k=1}^{n} \binom{n+1}{k} a^k b^{n+1-k} + \binom{n+1}{n+1} a^{n+1} b^0 + \binom{n+1}{0} a^0 b^{n+1} && \binom{n}{n} = 1 = \binom{n+1}{n+1} \text{ et } \binom{n}{0} = 1 = \binom{n+1}{0} \\
  &= \sum_{k=0}^{n+1} \binom{n+1}{k} a^k b^{n+1-k}
  \end{aligned}$$
  L'hérédité est ainsi établie.

- **Conclusion :** par récurrence, on a ainsi démontré : $\forall n \in \mathbb{N},\ (a+b)^n = \sum_{k=0}^{n} \binom{n}{k} a^k b^{n-k}$.

⋆

2 ♠ La fameuse somme...

Pour tout $n \in \mathbb{N}$, pour tout $x \in \mathbb{R} \setminus \{1\}$:

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

Petite remarque
Bien entendu, puisque le résultat est donné, on peut aussi procéder par récurrence...

⋆ Démonstration : Soit $n \in \mathbb{N}^*$. Considérons la fonction $f : x \longmapsto \sum_{k=1}^{n} x^k$.

- La fonction f est polynomiale, donc dérivable sur $\mathbb{R}$ et, par linéarité de la dérivation :
  $$\forall x \in \mathbb{R},\ f'(x) = \sum_{k=1}^{n} k x^{k-1}$$

- On sait également que pour tout $x \in \mathbb{R} \setminus \{1\}$:

$$f(x) = \sum_{k=0}^{n} x^k - 1 = \frac{1 - x^{n+1}}{1-x} - 1$$

En dérivant sous cette forme, on a, pour tout $x \in \mathbb{R} \setminus \{1\}$:

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

♡ Astuce du chef ! ♡

Pour simplifier les calculs, on n'utilise volontairement pas l'autre formule de somme géométrique donnant directement $\sum_{k=p}^{n} x^k$... On ne met pas non plus sur même dénominateur !

On a ainsi :

$$\forall x \in \mathbb{R} \setminus \{1\}, \ \sum_{k=1}^{n} kx^{k-1} = \frac{1-(n+1)x^n + nx^{n+1}}{(1-x)^2}$$

Conclusion : $\forall x \in \mathbb{R} \setminus \{1\}, \ \sum_{k=1}^{n} kx^k = \dfrac{x-(n+1)x^{n+1} + nx^{n+2}}{(1-x)^2}$.

⋆

3 ♥ Théorème de divergence monotone / Majoration d'une suite croissante convergente

Soit $(u_n)_{n\in\mathbb{N}}$ une suite de réels.

Q1 Si $(u_n)_{n\in\mathbb{N}}$ est **croissante** et **non majorée**, alors $(u_n)_{n\in\mathbb{N}}$ diverge vers $+\infty$.

Q2 Si $(u_n)_{n\in\mathbb{N}}$ est **croissante** et **converge vers un réel ℓ**, alors $(u_n)_{n\in\mathbb{N}}$ est majorée par ℓ.

⋆

Démonstration :

Q1. Supposons que $(u_n)_{n\in\mathbb{N}}$ est croissante et non majorée. On a ainsi :

$$\forall n \in \mathbb{N}, \ u_{n+1} \geqslant u_n \quad ; \quad \forall M \in \mathbb{R}, \ \exists n_0 \in \mathbb{N} \ / \ u_{n_0} > M$$

Montrons :

$$\forall M \in \mathbb{R}, \ \exists n_0 \in \mathbb{N}, \ \forall n \in [\![n_0; +\infty[\![, \ u_n \geqslant M$$

Soit $M \in \mathbb{R}$.
Puisque $(u_n)_{n\in\mathbb{N}}$ n'est pas majorée par M, il existe $n \in \mathbb{N}$ tel que $u_n > M$. Notons n_0 un tel entier.
La suite $(u_n)_{n\in\mathbb{N}}$ étant croissante, on obtient, par récurrence immédiate :

$$\forall n \in [\![n_0; +\infty[\![, \ u_n \geqslant u_{n_0}$$

D'où le résultat.

☞ Rappels...

- $(u_n)_{n\in\mathbb{N}}$ majorée :
$\exists M \in \mathbb{R} \ / \ \forall n \in \mathbb{N}, \ u_n \leqslant M$
- $(u_n)_{n\in\mathbb{N}}$ diverge vers $+\infty$: $\forall M \in \mathbb{R}, \ \exists n_0 \in \mathbb{N}, \ \forall n \in [\![n_0; +\infty[\![, \ u_n \geqslant M$

Q2. Supposons que $(u_n)_{n\in\mathbb{N}}$ est croissante et converge vers un réel ℓ. Montrons que $(u_n)_{n\in\mathbb{N}}$ est majorée par ℓ.
Raisonnons par l'absurde. Supposons alors que $(u_n)_{n\in\mathbb{N}}$ n'est pas majorée par ℓ ; autrement dit, supposons qu'il existe $n \in \mathbb{N}$ tel que $u_n > \ell$. Notons n_0 un tel entier.
La suite $(u_n)_{n\in\mathbb{N}}$ étant croissante, on obtient, par récurrence immédiate :

$$\forall n \in [\![n_0; +\infty[\![, \ u_n \geqslant u_{n_0}$$

Puis, en passant à la limite quand $n \to +\infty$ sur cette inégalité, on obtient :

$$\ell \geqslant u_{n_0}$$

D'où la contradiction.
Par conséquent : la suite $(u_n)_{n\in\mathbb{N}}$ est majorée par ℓ.

☞ Rappel...

$(u_n)_{n\in\mathbb{N}}$ majorée par ℓ :
$\forall n \in \mathbb{N}, \ u_n \leqslant \ell$

⋆

4 ♥ Théorème des suites adjacentes

Deux suites adjacentes convergent et ont la même limite.

Important !

J'insiste sur le "convergent et ont même limite". On ne se contente pas de dire qu'elles ont la même limite, ni qu'elles convergent. Il y a deux informations, les deux sont essentielles.

⋆

Démonstration : Soient $(u_n)_{n\in\mathbb{N}}$ et $(v_n)_{n\in\mathbb{N}}$ deux suites adjacentes telles que :

✓ $(u_n)_{n\in\mathbb{N}}$ est croissante
✓ $(v_n)_{n\in\mathbb{N}}$ est décroissante
✓ $\lim\limits_{n\to+\infty} (v_n - u_n) = 0$

Puisque la suite $(u_n)_{n\in\mathbb{N}}$ est croissante, la suite $(-u_n)_{n\in\mathbb{N}}$ est décroissante. Or $(v_n)_{n\in\mathbb{N}}$ est décroissante, donc la suite $(v_n - u_n)_{n\in\mathbb{N}}$ est décroissante.
Mais, par hypothèse, la suite $(v_n - u_n)_{n\in\mathbb{N}}$ converge vers 0.
Par conséquent : la suite $(v_n - u_n)_{n\in\mathbb{N}}$ est minorée par 0. Ainsi :

$$\forall n \in \mathbb{N}, \ v_n \geqslant u_n$$

Or :

♣ L'idée !

Prouver la convergence de (u_n) et (v_n), pour écrire $\lim\limits_{n\to+\infty} (v_n - u_n) = \lim\limits_{n\to+\infty} v_n - \lim\limits_{n\to+\infty} u_n$. La troisième hypothèse impliquera alors qu'elles convergent vers la même limite... Penser au théorème de convergence monotone, puisque (u_n) et (v_n) sont monotones...

- $(u_n)_{n\in\mathbb{N}}$ est croissante, elle est donc minorée par son premier terme ;
- $(v_n)_{n\in\mathbb{N}}$ est décroissante, elle est donc majorée par son premier terme.

Par transitivité, on obtient ainsi :

$$\forall n \in \mathbb{N},\ v_n \geqslant u_0 \quad ; \quad \forall n \in \mathbb{N},\ u_n \leqslant v_0$$

On a donc :

- $(u_n)_{n\in\mathbb{N}}$ croissante et majorée (par v_0),
- $(v_n)_{n\in\mathbb{N}}$ décroissante et minorée (par u_0).

Par théorème de convergence monotone, ces deux suites convergent respectivement vers des réels ℓ et ℓ'.
On a aussi :

$$\begin{aligned}\lim_{n\to+\infty}(v_n - u_n) &= \lim_{n\to+\infty} v_n - \lim_{n\to+\infty} u_n \qquad \text{(licite, car les deux limites en jeu sont finies)}\\ &= \ell' - \ell\end{aligned}$$

Mais, par hypothèse : $\lim\limits_{n\to+\infty}(v_n - u_n) = 0$. D'où :

$$\ell = \ell'$$

Conclusion : les suites $(u_n)_{n\in\mathbb{N}}$ et $(v_n)_{n\in\mathbb{N}}$ convergent vers la même limite.

⋆

5 ♠ Formule de Taylor avec reste intégral

Soient I un intervalle de $\mathbb{R}$, $a \in I$ et f une fonction définie sur I.
Si f est de classe $\mathscr{C}^\infty$ sur I, alors :

$$\forall x \in I,\ \forall n \in \mathbb{N},\ f(x) = \sum_{k=0}^{n} \frac{(x-a)^k}{k!} f^{(k)}(a) + \int_a^x \frac{(x-t)^n}{n!} f^{(n+1)}(t)dt$$

⋆

Démonstration : Soit $x \in I$. Procédons par récurrence...

- **Initialisation.** Pour $n = 0$.
  On a :

$$\begin{aligned}\sum_{k=0}^{0} \frac{(x-a)^k}{k!} f^{(k)}(a) + \int_a^x \frac{(x-t)^n}{n!} f'(t)dt &= \frac{(x-a)^0}{0!} f(a) + \int_a^x \frac{(x-t)^0}{0!} f'(t)dt\\ &= f(a) + \int_a^x f'(t)dt\\ &= f(a) + \big[f(t)\big]_a^x\\ &= f(a) + f(x) - f(a)\\ &= f(x)\end{aligned}$$

  L'initialisation est ainsi vérifiée.
- **Hérédité.** Soit $n \in \mathbb{N}$.
  Supposons $f(x) = \sum_{k=0}^{n} \frac{(x-a)^k}{k!} f^{(k)}(a) + \int_a^x \frac{(x-t)^n}{n!} f^{(n+1)}(t)dt$ et montrons $f(x) = \sum_{k=0}^{n+1} \frac{(x-a)^k}{k!} f^{(k)}(a) + \int_a^x \frac{(x-t)^{n+1}}{(n+1)!} f^{(n+2)}(t)dt$.
  Par hypothèse de récurrence :

$$f(x) = \sum_{k=0}^{n} \frac{(x-a)^k}{k!} f^{(k)}(a) + \int_a^x \frac{(x-t)^n}{n!} f^{(n+1)}(t)dt$$

  Procédons ensuite à une intégration par parties...
  Posons : $\left|\begin{array}{l} u : t \longmapsto -\dfrac{(x-t)^{n+1}}{(n+1)!} \\ v : t \longmapsto f^{(n+1)}(t) \end{array}\right.$.

  Les fonctions u et v sont $\mathscr{C}^1$ sur le segment $[a, x]$ (ou $[x; a]$) et pour tout $t \in [a; x]$ (ou $[x; a]$) : $\left|\begin{array}{l} u'(t) = \dfrac{(x-t)^n}{n!} \\ v'(t) = f^{(n+2)}(t) \end{array}\right.$.

  On obtient alors :

$$\begin{aligned}f(x) &= \sum_{k=0}^{n} \frac{(x-a)^k}{k!} f^{(k)}(a) + \int_a^x \frac{(x-t)^n}{n!} f^{(n+1)}(t)dt \qquad \text{IPP}\\ &= \sum_{k=0}^{n} \frac{(x-a)^k}{k!} f^{(k)}(a) + \left[-\frac{(x-t)^{n+1}}{(n+1)!} f^{(n+1)}(t)\right]_a^x - \int_a^x -\frac{(x-t)^{n+1}}{(n+1)!} f^{(n+2)}(t)dt\\ &= \sum_{k=0}^{n} \frac{(x-a)^k}{k!} f^{(k)}(a) - \frac{(x-x)^{n+1}}{(n+1)!} f^{(n+1)}(x) + \frac{(x-a)^{n+1}}{(n+1)!} f^{(n+1)}(a) + \int_a^x \frac{(x-t)^{n+1}}{(n+1)!} f^{(n+2)}(t)dt\\ &= \sum_{k=0}^{n+1} \frac{(x-a)^k}{k!} f^{(k)}(a) + \int_a^x \frac{(x-t)^{n+1}}{(n+1)!} f^{(n+2)}(t)dt\end{aligned}$$

  L'hérédité est ainsi établie.

Conclusion : si f est de classe $\mathscr{C}^\infty$ sur I, alors :

$$\forall x \in I,\ \forall n \in \mathbb{N},\ f(x) = \sum_{k=0}^{n} \frac{(x-a)^k}{k!} f^{(k)}(a) + \int_a^x \frac{(x-t)^n}{n!} f^{(n+1)}(t)dt$$

⋆

6 ♣ Une intégrale classique

$$\forall p, q \in \mathbb{N}, \quad \int_0^1 x^p(1-x)^q\,dx = \frac{p!q!}{(p+q+1)!}$$

★ Démonstration : Notons, pour tous $p, q \in \mathbb{N}$, $I(p,q) = \displaystyle\int_0^1 x^p(1-x)^q\,dx$.

- Soit $(p,q) \in \mathbb{N} \times \mathbb{N}^*$. Commençons par une intégration par parties...

  Posons : $\left| \begin{array}{l} u : x \longmapsto \dfrac{x^{p+1}}{p+1} \\ v : x \longmapsto (1-x)^q \end{array} \right.$. Les fonctions u et v sont $\mathscr{C}^1$ sur le segment $[0;1]$ et pour tout $x \in [0;1]$: $\left| \begin{array}{l} u'(x) = x^p \\ v'(x) = -q(1-x)^{q-1} \end{array} \right.$.

  D'où :

$$\begin{aligned} I(p,q) &= \int_0^1 x^p(1-x)^q\,dx \\ &= \left[\frac{x^{p+1}}{p+1}(1-x)^q\right]_0^1 - \int_0^1 -q(1-x)^{q-1}\frac{x^{p+1}}{p+1}\,dx && \text{IPP} \\ &= \frac{q}{p+1}\int_0^1 x^{p+1}(1-x)^{q-1}\,dx && p>0 \text{ et } q+1>0, \text{ donc } 0^p = 0 \text{ et } 0^{q+1} = 0 \\ &= \frac{q}{p+1} I(p+1, q-1) \end{aligned}$$

  Puisque $q \in \mathbb{N}^*$, on obtient :

$$I(p+1, q-1) = \frac{p+1}{q} I(p,q)$$

  On a ainsi établi :

$$\forall (p,q) \in \mathbb{N} \times \mathbb{N}^*, \quad I(p+1,q-1) = \frac{p+1}{q} I(p,q)$$

- Par récurrence, démontrons : $\forall p \in \mathbb{N}, \ \forall q \in \mathbb{N}, \ I(p,q) = \dfrac{p!q!}{(p+q+1)!}$.

  ⋆ **Initialisation.** Pour $p = 0$:
  Soit $q \in \mathbb{N}$. On a :

$$\begin{aligned} I(0,q) &= \int_0^1 (1-x)^q\,dx \\ &= \left[-\frac{(1-x)^{q+1}}{q+1}\right]_0^1 \\ &= \frac{1}{q+1} \end{aligned}$$

  Et :

$$\begin{aligned} \frac{0!q!}{(0+q+1)!} &= \frac{q!}{(q+1)!} \\ &= \frac{1}{q+1} \end{aligned}$$

  D'où :

$$I(0,q) = \frac{1}{q+1}$$

  L'initialisation est ainsi vérifiée.

  ⋆ **Hérédité.** Soit $p \in \mathbb{N}$. Supposons : $\forall q \in \mathbb{N}, \ I(p,q) = \dfrac{p!q!}{(p+q+1)!}$. Montrons : $\forall q \in \mathbb{N}, \ I(p+1,q) = \dfrac{(p+1)!q!}{(p+q+2)!}$.
  Soit $q \in \mathbb{N}$. D'après le point précédent, licite car $p \in \mathbb{N}$ et $q+1 \in \mathbb{N}^*$, on a :

$$\begin{aligned} I(p+1,q) &= \frac{p+1}{q+1} I(p,q+1) && \text{hypothèse de récurrence, licite car } q+1 \in \mathbb{N} \\ &= \frac{p+1}{q+1}\frac{p!(q+1)!}{(p+q+2)!} \\ &= \frac{(p+1)!q!}{(p+q+2)!} \end{aligned}$$

  L'hérédité est ainsi établie.

  > **✘ Attention !**
  > On utilise le point précédent avec $q+1$ au lieu de q (licite, car $q \in \mathbb{N}$, donc $q+1 \in \mathbb{N}^*$) :
  > $$I(p+1,q) = \frac{p+1}{q+1} I(p,q+1)$$

  Conclusion : $\forall p, q \in \mathbb{N}, \ \displaystyle\int_0^1 x^p(1-x)^q\,dx = \frac{p!q!}{(p+q+1)!}$.

Comment retrouver l'expression de $I(p,q)$ si elle n'est pas donnée ?

On a :

$$\begin{aligned} I(p,q) &= \frac{q}{p+1} I(p+1,q-1) \\ &= \frac{q}{p+1}\frac{q-1}{p+2} I(p+2,q-2) \\ &= \frac{q}{p+1}\frac{q-1}{p+2}\frac{q-2}{p+3} I(p+3,q-3) \\ &= \ldots \end{aligned}$$

$$= \frac{q}{p+1}\frac{q-1}{p+2}\frac{q-2}{p+3} \times \dots \times \frac{1}{p+q} I(p+q,0)$$

$$= \frac{q}{p+1}\frac{q-1}{p+2}\frac{q-2}{p+3} \times \dots \times \frac{1}{p+q}\frac{1}{p+q+1}$$

$I(p+q,0) = \int_0^1 x^{p+q}\,dx = \frac{1}{p+q+1}$

$$= \frac{q!}{(p+1)(p+2)\dots(p+q+1)}$$

$$= \frac{q!p!}{(p+q+1)!}$$

⋆

7 ♣ Sommes de Riemann à gauche (méthode des rectangles à gauche)

Soient $a, b \in \mathbb{R}$ tels que $a < b$ et f une fonction définie sur $[a;b]$.
Pour tout $n \in \mathbb{N}^*$, on pose :

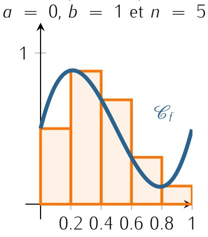
- pour tout $k \in [\![0;n]\!]$, $a_k = a + k\frac{b-a}{n}$
- $S_n = \frac{b-a}{n}\sum_{k=0}^{n-1} f(a_k)$

Si la fonction f **est** $\mathscr{C}^1$ **sur le segment** $[a;b]$, alors la suite $(S_n)_{n\in\mathbb{N}^*}$ converge et :

$$\lim_{n\to+\infty} S_n = \int_a^b f(t)dt$$

Important !

Schéma pour comprendre avec $a = 0$, $b = 1$ et $n = 5$:



La somme S_5 est égale à la somme des aires des 5 rectangles...

Petite remarque

Puisque f est $\mathscr{C}^1$ sur $[a;b]$, elle est en particulier continue sur le segment $[a;b]$, d'où l'existence de $\int_a^b f(t)dt$.

⋆

Démonstration : Objectif : à l'aide du théorème d'encadrement, démontrer que

$$\lim_{n\to+\infty}\left|S_n - \int_a^b f(t)dt\right| = 0$$

Soit $n \in \mathbb{N}^*$.

- Commençons par remarquer que $a_0 = a$, $a_n = b$ et que :

$$\forall k \in [\![0;n-1]\!],\ a_{k+1} - a_k = \frac{b-a}{n}$$

Important !

En fait, $(a_k)_{k\in[\![0;n]\!]}$ fournit une subdivision régulière à pas constant égal à $\frac{b-a}{n}$ du segment $[a;b]$.

- Ensuite :

$$\begin{aligned}
\left|S_n - \int_a^b f(t)dt\right| &= \left|\frac{b-a}{n}\sum_{k=0}^{n-1} f(a_k) - \int_a^b f(t)dt\right| \\
&= \left|\sum_{k=0}^{n-1}\frac{b-a}{n} f(a_k) - \int_a^b f(t)dt\right| \\
&= \left|\sum_{k=0}^{n-1}(a_{k+1}-a_k) f(a_k) - \int_a^b f(t)dt\right| && \text{premier point} \\
&= \left|\sum_{k=0}^{n-1}\int_{a_k}^{a_{k+1}} f(a_k)dt - \int_a^b f(t)dt\right| \\
&= \left|\sum_{k=0}^{n-1}\int_{a_k}^{a_{k+1}} f(a_k)dt - \sum_{k=0}^{n-1}\int_{a_k}^{a_{k+1}} f(t)dt\right| && \text{relation de Chasles, avec } a_0 = a \text{ et } a_n = b \\
&= \left|\sum_{k=0}^{n-1}\int_{a_k}^{a_{k+1}} \big(f(a_k) - f(t)\big)dt\right| && \text{linéarité de la somme et de l'intégrale} \\
&\leqslant \sum_{k=0}^{n-1}\left|\int_{a_k}^{a_{k+1}} \big(f(a_k) - f(t)\big)dt\right| && \text{inégalité triangulaire sur la somme} \\
&\leqslant \sum_{k=0}^{n-1}\int_{a_k}^{a_{k+1}} \big|f(a_k) - f(t)\big|dt && \text{inégalité triangulaire sur l'intégrale, licite car : } \forall k \in [\![0;n-1]\!],\ a_k \leqslant a_{k+1}
\end{aligned}$$

- Majorons l'intégrale, et donc commençons par majorer l'intégrande...
  - ⋆ Puisque f est de classe $\mathscr{C}^1$ sur $[a;b]$, la fonction f' est continue sur $[a;b]$. Par composition, la fonction $|f'|$ est également continue sur le segment $[a;b]$. Ainsi, d'après le théorème des bornes, elle admet un maximum, noté M, sur $[a;b]$.
    On a donc :
    $$\forall x \in [a;b],\ |f'(x)| \leqslant M$$
  - ⋆ Ainsi :
    - ✓ f dérivable sur $[a;b]$ (car de classe $\mathscr{C}^1$),
    - ✓ $\forall x \in [a;b]$, $|f'(x)| \leqslant M$.

D'après l'inégalité des accroissements finis, on a donc :

$$\forall \alpha, \beta \in [a;b],\ \left|f(\alpha) - f(\beta)\right| \leqslant M|\alpha - \beta|$$

Soit $k \in [\![0; n-1]\!]$.
Puisque $a_k \in [a;b]$ et que $[a_k, a_{k+1}] \subset [a;b]$, on obtient d'après le point précédent :

$$\forall t \in [a_k, a_{k+1}],\ \left|f(a_k) - f(t)\right| \leqslant M|a_k - t|$$

Or, pour tout $t \in [a_k; a_{k+1}]$:

$$|a_k - t| \leqslant |a_k - a_{k+1}| = \frac{b-a}{n}$$

D'où :

$$\forall t \in [a_k, a_{k+1}],\ \left|f(a_k) - f(t)\right| \leqslant M\frac{b-a}{n}$$

Ainsi, par croissance de l'intégrale, licite car $a_k \leqslant a_{k+1}$, on obtient :

$$\int_{a_k}^{a_{k+1}} \left|f(a_k) - f(t)\right| dt \leqslant \int_{a_k}^{a_{k+1}} \frac{M(b-a)}{n} dt$$

Or :

$$\begin{aligned}\int_{a_k}^{a_{k+1}} \frac{M(b-a)}{n} dt &= (a_{k+1} - a_k)\frac{M(b-a)}{n} \\ &= \frac{M(b-a)^2}{n^2}\end{aligned}$$

- Par transitivité, on vient d'établir :

$$\forall k \in [\![0; n-1]\!],\ \int_{a_k}^{a_{k+1}} \left|f(a_k) - f(t)\right| dt \leqslant \frac{M(b-a)^2}{n^2}$$

D'où, en sommant de 0 à $n-1$:

$$\sum_{k=0}^{n-1} \int_{a_k}^{a_{k+1}} \left|f(a_k) - f(t)\right| dt \leqslant \frac{M(b-a)^2}{n}$$

✘ Attention !
Le cardinal de $[\![0; n-1]\!]$ est n...

On a ainsi établi :

$$\forall n \in \mathbb{N}^*,\ 0 \leqslant \left|S_n - \int_a^b f(t)dt\right| \leqslant \frac{M(b-a)^2}{n}$$

Or :

$$\lim_{n\to+\infty} \frac{M(b-a)^2}{n} = 0$$

D'où, par théorème d'encadrement :

$$\lim_{n\to+\infty} \left|S_n - \int_a^b f(t)dt\right| = 0$$

Et ainsi :

$$\lim_{n\to+\infty} S_n = \int_a^b f(t)dt$$

✔ Pour s'entraîner...
Écrire un programme **Python** qui, pour une fonction donnée sur un intervalle $[a,b]$ donné, calcule cette somme de Riemann à gauche...

⋆

8 ♣ Série logarithmique

Pour tout $x \in [-1;0]$, la série $\sum_{n\geqslant 1} \frac{x^n}{n}$ est convergente et $\sum_{n=1}^{+\infty} \frac{x^n}{n} = -\ln(1-x)$.

⋆ Démonstration : Soit $x \in [-1;0]$.

- Soit $n \in \mathbb{N}^*$. On a, pour tout $t \in [x;0]$:

$$\begin{aligned}\sum_{k=1}^{n} t^{k-1} &= \sum_{i=0}^{n-1} t^i \\ &= \frac{1-t^n}{1-t}\end{aligned}$$

$t \neq 1$

♣ L'idée !
Pour étudier $\sum kx^{k-1}$, on a dérivé $\sum_{k=0}^{n} x^k$... Pour étudier $\sum_{k\geqslant 1} \frac{x^k}{k}$, on primitive $\sum_{k=1}^{n} x^{k-1}$...

En intégrant de x à 0, licite car la fonction $t \longmapsto \sum_{k=1}^{n} t^{k-1}$ est polynomiale donc continue sur le segment $[x;0]$:

$$\int_x^0 \sum_{k=1}^{n} t^{k-1} dt = \int_x^0 \frac{1-t^n}{1-t} dt$$

Or, par linéarité de l'intégrale, on a :

⋆ d'une part :

$$\begin{aligned}\int_x^0 \sum_{k=1}^{n} t^{k-1} dt &= \sum_{k=1}^{n} \int_x^0 t^{k-1} dt \\ &= \sum_{k=1}^{n} \left[\frac{t^k}{k}\right]_x^0 \\ &= -\sum_{k=1}^{n} \frac{x^k}{k}\end{aligned}$$

$k-1 \neq -1$

✓ Rigueur !
Si on avait $k-1 = -1$, alors il faudrait primitiver $t \longmapsto \frac{1}{t}$, qui se primitive différemment...

⋆ d'autre part :

$$\begin{aligned}\int_x^0 \frac{1-t^n}{1-t}dt &= \int_x^0 \frac{1}{1-t}dt - \int_x^0 \frac{t^n}{1-t}dt \\ &= \big[-\ln(1-t)\big]_x^0 - \int_x^0 \frac{t^n}{1-t}dt \\ &= \ln(1-x) - \int_x^0 \frac{t^n}{1-t}dt\end{aligned}$$

Par conséquent :

$$\sum_{k=1}^{n} \frac{x^k}{k} = -\ln(1-x) + \int_x^0 \frac{t^n}{1-t}dt$$

- Montrons maintenant : $\lim\limits_{n\to+\infty} \int_x^0 \frac{t^n}{1-t}dt = 0$. Soit $n \in \mathbb{N}^*$.

⋆ Par inégalité triangulaire, licite car $x \leqslant 0$, on a :

$$\left|\int_x^0 \frac{t^n}{1-t}dt\right| \leqslant \int_x^0 \left|\frac{t^n}{1-t}\right| dt$$

Or, pour tout $t \in [x;0]$:

$$\begin{aligned}\left|\frac{t^n}{1-t}\right| &= \frac{|t^n|}{|1-t|} \\ &= \frac{|t|^n}{1-t} && 1-t>0 \\ &= \frac{(-t)^n}{1-t} && t \leqslant 0, \text{ donc } |t| = -t\end{aligned}$$

Par conséquent :

$$\left|\int_x^0 \frac{t^n}{1-t}dt\right| \leqslant \int_x^0 \frac{(-t)^n}{1-t}dt$$

⋆ Or :

$$\forall t \in [x;0],\ 1-t \geqslant 1$$

D'où, par décroissance de la fonction inverse sur $\mathbb{R}_*^+$, et comme pour tout $t \in [x;0]$, $(-t)^n \geqslant 0$, on obtient :

$$\forall t \in [x;0],\ \frac{(-t)^n}{1-t} \leqslant (-t)^n$$

Puis, par croissance de l'intégrale, licite car $x \leqslant 0$:

$$\int_x^0 \frac{(-t)^n}{1-t}dt \leqslant \int_x^0 (-t)^n dt$$

⋆ Ensuite, on a :

$$\begin{aligned}\int_x^0 (-t)^n dt &= \left[-\frac{(-t)^{n+1}}{n+1}\right]_x^0 \\ &= \frac{(-x)^{n+1}}{n+1} \\ &\leqslant \frac{1}{n+1} && 0 \leqslant -x \leqslant 1, \text{ puis croissance de } t \longmapsto t^{n+1} \text{ sur } \mathbb{R}^+ \text{ et } n+1>0...\end{aligned}$$

Des trois points précédents, on déduit :

$$\forall n \in \mathbb{N}^*,\ 0 \leqslant \left|\int_{-1}^0 \frac{t^n}{1-t}dt\right| \leqslant \frac{1}{n+1}$$

Mais, $\lim\limits_{n\to+\infty} \frac{1}{n+1} = 0$. D'où, par théorème d'encadrement : $\lim\limits_{n\to+\infty} \left|\int_x^0 \frac{t^n}{1-t}dt\right| = 0$. Et ainsi :

$$\lim_{n\to+\infty} \int_x^0 \frac{t^n}{1-t}dt = 0$$

On obtient finalement :

$$\lim_{n\to+\infty} \left(-\ln(1-x) + \int_x^0 \frac{t^n}{1-t}dt\right) = -\ln(1-x)$$

Autrement dit :

$$\lim_{n\to+\infty} \left(\sum_{k=1}^{n} \frac{x^k}{k}\right) = -\ln(1-x)$$

Conclusion : pour tout $x \in [-1;0]$, la série $\sum\limits_{n\geqslant 1} \frac{x^n}{n}$ est convergente et $\sum\limits_{n=1}^{+\infty} \frac{x^n}{n} = -\ln(1-x)$.

⋆

✘ Attention !
Puisque $t \in [x;0]$, t^n n'est pas de signe constant... On va travailler sur la valeur absolue de l'intégrale.

☞ Rappel...
Pour majorer une fraction à numérateur positif, on minore son dénominateur.

✘ Attention !
$\lim\limits_{n\to+\infty}(-x)^{n+1}$ ne vaut pas toujours 0 quand $x \in [-1;0]$... D'où la manipulation à prévoir ici.
En revanche, on sait bien que :

$$\forall x \in [-1;0],\ \lim_{n\to+\infty} \frac{(-x)^{n+1}}{n+1} = 0$$

que l'on démontre grâce à l'encadrement proposé ici et au théorème d'encadrement...

9 ♥ Propriétés d'une probabilité

Soient $(\Omega, \mathcal{A}, \mathbb{P})$ un espace probabilisé et $A, B \in \mathcal{A}$.

Q1 Si $A \subset B$, alors $\mathbb{P}(A) \leqslant \mathbb{P}(B)$ *(croissance de $\mathbb{P}$)*

Q2 $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B)$ *(formule de Poincaré)*

Démonstration :

Q1. Supposons que $A \subset B$. On a :

$$\begin{aligned} B &= \Omega \cap B \\ &= (A \cup \overline{A}) \cap B && A \cup \overline{A} = \Omega \\ &= (A \cap B) \cup (\overline{A} \cap B) && \text{distributivité de } \cap \text{ sur } \cup \\ &= A \cup (\overline{A} \cap B) && A \subset B, \text{ donc } A \cap B = A \end{aligned}$$

Or, les évènements A et $\overline{A} \cap B$ sont incompatibles, d'où :

$$\mathbb{P}(B) = \mathbb{P}(A) + \mathbb{P}(\overline{A} \cap B)$$

Et comme $\mathbb{P}(\overline{A} \cap B) \geqslant 0$, on a bien :

$$\mathbb{P}(A) \leqslant \mathbb{P}(B)$$

Q2. On a :

$$\begin{aligned} A \cup B &= \Omega \cap (A \cup B) \\ &= (A \cup \overline{A}) \cap (A \cup B) && A \cup \overline{A} = \Omega \\ &= A \cup (\overline{A} \cap B) && \text{"factorisation" de } \cup \text{ sur } \cap \end{aligned}$$

Or, les évènements A et $\overline{A} \cap B$ sont incompatibles. D'où :

$$\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(\overline{A} \cap B)$$

Ensuite, d'après la formule des probabilités totales avec $(A, \overline{A})$ comme système complet d'évènements, on a :

$$\mathbb{P}(B) = \mathbb{P}(A \cap B) + \mathbb{P}(\overline{A} \cap B)$$

Par conséquent :

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

On obtient finalement :

$$\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B)$$

♣ L'idée ! Écrire B comme l'union disjointe de A et d'un autre évènement.

♣ L'idée ! Écrire $A \cup B$ comme l'union disjointe de A et d'un autre évènement.

À retenir... $A \cup B = A \cup (\overline{A} \cap B)$... et à démontrer en partant du membre de droite si besoin.

10 ♥ Formule des probabilités totales

Soient $(\Omega, \mathcal{A}, \mathbb{P})$ un espace probabilisé et $(A_i)_{i\in\mathbb{N}}$ une famille d'évènements.
Si $(A_i)_{i\in\mathbb{N}}$ **est système complet d'évènements**, alors, pour tout $B \in \mathcal{A}$, la série $\sum_{i\in\mathbb{N}} \mathbb{P}(A_i \cap B)$ est convergente et :

$$\mathbb{P}(B) = \sum_{i=0}^{+\infty} \mathbb{P}(A_i \cap B)$$

Démonstration : Supposons que $(A_i)_{i\in\mathbb{N}}$ est un système complet d'évènements.
Soit $B \in \mathcal{A}$. On a :

$$\begin{aligned} B &= \Omega \cap B \\ &= \left(\bigcup_{i=0}^{+\infty} A_i\right) \cap B && \text{car } (A_i)_{i\in\mathbb{N}} \text{ est un sce, donc } \bigcup_{i=0}^{+\infty} A_i = \Omega \\ &= \bigcup_{i=0}^{+\infty} (A_i \cap B) && \text{distributivité de } \cap \text{ sur } \cup \end{aligned}$$

Les évènements de cette union sont-ils deux à deux incompatibles ? Soient $i, j \in \mathbb{N}$. Supposons que $i \neq j$. On a :

$$\begin{aligned} (A_i \cap B) \cap (A_j \cap B) &= (A_i \cap A_j) \cap B && \text{par associativité et commutativité de } \cap \\ &= \varnothing \cap B && \text{car } (A_i)_{i\in\mathbb{N}} \text{ est un sce} \\ &= \varnothing \end{aligned}$$

L'union $\bigcup_{i=0}^{+\infty} (A_i \cap B)$ est donc une union d'évènements deux à deux incompatibles... D'où :

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

Conclusion :

$$\mathbb{P}(B) = \sum_{i=0}^{+\infty} \mathbb{P}(A_i \cap B)$$

♣ L'idée ! Écrire B comme union d'évènements deux à deux incompatibles...

11 ♥ Probabilité conditionnelle

Soient $(\Omega, \mathcal{A}, \mathbb{P})$ un espace probabilisé et $A \in \mathcal{A}$.
Si $\mathbb{P}(A) \neq 0$, alors l'application $\mathbb{P}_A : B \longmapsto \dfrac{\mathbb{P}(A \cap B)}{\mathbb{P}(A)}$ est une probabilité sur $(\Omega, \mathcal{A})$.

> ☞ **Rappel...**
> Il faut que :
> - $\mathbb{P}_A$ soit définie sur $\mathcal{A}$, à valeurs dans $[0;1]$,
> - $\mathbb{P}_A(\Omega) = 1$,
> - si $(B_n)_{n\in\mathbb{N}}$ est une suite d'évènements deux à deux incompatibles, alors $\mathbb{P}_A\left(\bigcup_{n=0}^{+\infty} B_n\right) = \sum_{n=0}^{+\infty} \mathbb{P}_A(B_n)$ (cette dernière propriété est appelée **σ-additivité**).

⋆ Démonstration : Supposons que $\mathbb{P}(A) \neq 0$.

- Soit $B \in \mathcal{A}$. Puisque $\mathcal{A}$ est stable par intersection, on a $A \cap B \in \mathcal{A}$, ainsi $\mathbb{P}(A \cap B)$ existe. Or $\mathbb{P}(A) \neq 0$, donc $\dfrac{\mathbb{P}(A \cap B)}{\mathbb{P}(A)}$ existe.
  Par conséquent : $\mathbb{P}_A$ est une application définie sur $\mathcal{A}$.
- Soit $B \in \mathcal{A}$.
  - ⋆ $\mathbb{P}_A(B)$ est le quotient de deux réels positifs, donc : $\mathbb{P}_A(B) \geqslant 0$.
  - ⋆ Ensuite, on sait que $A \cap B \subset A$. D'où, par croissance de $\mathbb{P}$ (qui est une probabilité) :
  $$\mathbb{P}(A \cap B) \leqslant \mathbb{P}(A)$$
  Puis, en divisant par $\mathbb{P}(A) \neq 0$, on obtient :
  $$\mathbb{P}_A(B) \leqslant 1$$
  Par conséquent : $\mathbb{P}_A$ est à valeurs dans $[0;1]$.
- Puis :
$$\begin{aligned}\mathbb{P}_A(\Omega) &= \frac{\mathbb{P}(A \cap \Omega)}{\mathbb{P}(A)} \\ &= \frac{\mathbb{P}(A)}{\mathbb{P}(A)} \\ &= 1\end{aligned}$$
- Enfin, considérons $(B_n)_{n\in\mathbb{N}}$ une suite d'évènements deux à deux incompatibles. On a :
$$\begin{aligned}\mathbb{P}_A\left(\bigcup_{n=0}^{+\infty} B_n\right) &= \frac{\mathbb{P}\left(A \cap \bigcup_{n=0}^{+\infty} B_n\right)}{\mathbb{P}(A)} \\ &\overset{\text{distributivité de } \cap \text{ sur } \cup}{=} \frac{\mathbb{P}\left(\bigcup_{n=0}^{+\infty} (A \cap B_n)\right)}{\mathbb{P}(A)}\end{aligned}$$

  Montrons que la famille $(A \cap B_n)_{n\in\mathbb{N}}$ est une suite d'évènements deux à deux incompatibles. Soient $i, j \in \mathbb{N}$. Supposons $i \neq j$. On a :
$$\begin{aligned}(A \cap B_i) \cap (A \cap B_j) &= A \cap (B_i \cap B_j) && \text{par associativité et commutativité de } \cap \\ &= A \cap \varnothing && \text{car } (B_n)_{n\in\mathbb{N}} \text{ est une famille d'évènements deux à deux incompatibles} \\ &= \varnothing\end{aligned}$$
  La famille $(A \cap B_n)_{n\in\mathbb{N}}$ est donc une famille d'évènements deux à deux incompatibles. D'où, en reprenant, puisque $\mathbb{P}$ est une probabilité :
$$\begin{aligned}\mathbb{P}_A\left(\bigcup_{n=0}^{+\infty} B_n\right) &= \frac{\sum_{n=0}^{+\infty} \mathbb{P}(A \cap B_n)}{\mathbb{P}(A)} \\ &= \sum_{n=0}^{+\infty} \frac{\mathbb{P}(A \cap B_n)}{\mathbb{P}(A)} \\ &= \sum_{n=0}^{+\infty} \mathbb{P}_A(B_n)\end{aligned}$$

Conclusion : si $\mathbb{P}(A) \neq 0$, alors l'application $\mathbb{P}_A$ est une probabilité sur $(\Omega, \mathcal{A})$. ⋆

12 ♥ Variance d'une VA suivant la loi $\mathscr{G}(p)$

Soient $(\Omega, \mathcal{A}, \mathbb{P})$ un espace probabilisé et X une variable aléatoire sur $(\Omega, \mathcal{A}, \mathbb{P})$. Soit $p \in]0;1[$.
Si $X \hookrightarrow \mathscr{G}(p)$, alors X admet une variance et $\mathbb{V}(X) = \dfrac{1-p}{p^2}$.

> ☞ **Rappel...**
> Si $X \hookrightarrow \mathscr{G}(p)$, alors :
> - $X(\Omega) = \mathbb{N}^*$
> - $\forall n \in \mathbb{N}^*,\ \mathbb{P}([X = n]) = p(1-p)^{n-1}$.

⋆ Démonstration : Supposons $X \hookrightarrow \mathscr{G}(p)$.
On admet que X possède une espérance égale à $\dfrac{1}{p}$. Ensuite :

- Par théorème de transfert :

X admet un moment d'ordre 2 si, et seulement si, la série $\sum_{n\in X(\Omega)} n^2 \mathbb{P}([X = n])$ est absolument convergente
si, et seulement si, la série $\sum_{n\geqslant 1} n^2 \mathbb{P}([X = n])$ est convergente, car il s'agit d'une série à terme général positif

- Soit $N \in \mathbb{N}$, suffisamment proche de $+\infty$. On a :

$$\begin{aligned}\sum_{n=1}^{N} n^2 \mathbb{P}\big([X=n]\big) &= \sum_{n=1}^{N} n^2 p(1-p)^{n-1} \\ &= p\sum_{n=1}^{N} \big(n(n-1)+n\big)(1-p)^{n-1} \\ &= p(1-p)\sum_{n=1}^{N} n(n-1)(1-p)^{n-2} + p\sum_{n=1}^{N} n(1-p)^{n-1}\end{aligned}$$

linéarité de la somme

Or, $p \in]0;1[$, donc $1-p \in]-1;1[$. Ainsi les séries $\sum_{n\geqslant 1} n(n-1)(1-p)^{n-2}$ et $\sum_{n\geqslant 1} n(1-p)^{n-1}$ sont des séries géométriques convergentes.
Par conséquent, la série $\sum_{n\geqslant 1} n^2\mathbb{P}\big([X=n]\big)$ est une combinaison linéaire de séries convergentes, elle est donc également convergente.

- On en déduit que X admet un moment d'ordre 2 et :

$$\begin{aligned}\mathbb{E}(X^2) &= \sum_{n=1}^{+\infty} n^2\mathbb{P}\big([X=n]\big) \\ &= p(1-p)\sum_{n=1}^{+\infty} n(n-1)(1-p)^{n-2} + p\sum_{n=1}^{+\infty} n(1-p)^{n-1} \\ &= p(1-p)\frac{2}{(1-(1-p))^3} + \mathbb{E}(X) \\ &= \frac{2(1-p)}{p^2} + \frac{1}{p} \\ &= \frac{2-p}{p^2}\end{aligned}$$

- Ainsi, d'après la formule de Koenig-Huygens, X admet une variance et :

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

♡ Astuce du chef ! ♡
A l'oral, on dit "variance égale moment d'ordre 2 moins carré de l'espérance".

Conclusion : X admet une variance et $\mathbb{V}(X) = \dfrac{1-p}{p^2}$.

⋆

13 ♣ Autour de la trace d'une matrice

Définition. Soient $n \in [\![2;+\infty[\![$ et $A = (a_{i,j})_{i,j\in[\![1;n]\!]} \in \mathcal{M}_n(\mathbb{R})$. On appelle **trace de A**, notée $\operatorname{tr}(A)$, le réel défini par : $\operatorname{tr}(A) = \sum_{i=1}^{n} a_{i,i}$.

Autremement dit :
La trace d'une matrice carrée est la somme de ses coefficients diagonaux.

Résultats.

Q1. Pour tous $\lambda, \mu \in \mathbb{R}$ et toutes matrices $A, B \in \mathcal{M}_n(\mathbb{R})$, on a $\operatorname{tr}(\lambda A+\mu B) = \lambda\operatorname{tr}(A)+\mu\operatorname{tr}(B)$.

Q2. Pour toutes matrices $A, B \in \mathcal{M}_n(\mathbb{R})$, on a $\operatorname{tr}(AB) = \operatorname{tr}(BA)$.

⋆ Démonstration :

Q1. Soient $\lambda, \mu \in \mathbb{R}$ et $A, B \in \mathcal{M}_n(\mathbb{R})$. Notons $C = \lambda A + \mu B$ ainsi que $A = (a_{i,j})_{i,j\in[\![1;n]\!]}$, $B = (b_{i,j})_{i,j\in[\![1;n]\!]}$ et $C = (c_{i,j})_{i,j\in[\![1;n]\!]}$.
Par définition de la multiplication scalaire et de l'addition matricielle, on a :

$$\forall i,j \in [\![1;n]\!],\ c_{i,j} = \lambda a_{i,j} + \mu b_{i,j}$$

Par conséquent :

$$\begin{aligned}\operatorname{tr}(\lambda A+\mu B) &= \operatorname{tr}(C) \\ &= \sum_{k=1}^{n} c_{k,k} \\ &= \sum_{k=1}^{n} (\lambda a_{k,k} + \mu b_{k,k}) \\ &= \lambda\sum_{k=1}^{n} a_{k,k} + \mu\sum_{k=1}^{n} b_{k,k} \\ &= \lambda\operatorname{tr}(A)+\mu\operatorname{tr}(B)\end{aligned}$$

linéarité de la somme

Q2. Soient $A, B \in \mathcal{M}_n(\mathbb{R})$. Notons $C = AB$ et $D = BA$ ainsi que $A = (a_{i,j})_{i,j\in[\![1;n]\!]}$, $B = (b_{i,j})_{i,j\in[\![1;n]\!]}$, $C = (c_{i,j})_{i,j\in[\![1;n]\!]}$ et $D = (d_{i,j})_{i,j\in[\![1;n]\!]}$.
Par définition de la multiplication matricielle, on a :

$$\forall i,j \in [\![1;n]\!],\ c_{i,j} = \sum_{k=1}^{n} a_{i,k}b_{k,j}\ \ ;\ \ d_{i,j} = \sum_{k=1}^{n} b_{i,k}a_{k,j}$$

Par conséquent :

$$\begin{aligned}\mathrm{tr}(AB) &= \mathrm{tr}(C)\\ &= \sum_{i=1}^{n} c_{i,i}\\ &= \sum_{i=1}^{n}\sum_{k=1}^{n} a_{i,k}b_{k,i}\end{aligned}$$

Également :

$$\begin{aligned}\mathrm{tr}(BA) &= \mathrm{tr}(D)\\ &= \sum_{i=1}^{n} d_{i,i}\\ &= \sum_{i=1}^{n}\sum_{k=1}^{n} b_{i,k}a_{k,i}\end{aligned}$$

Enfin :

$$\begin{aligned}\mathrm{tr}(AB) &= \sum_{i=1}^{n}\sum_{k=1}^{n} a_{i,k}b_{k,i}\\ &= \sum_{1\leqslant i,k\leqslant n} a_{i,k}b_{k,i}\\ &= \sum_{k=1}^{n}\sum_{i=1}^{n} b_{k,i}a_{i,k} && \text{les indices des sommes sont muets}\\ &= \sum_{i=1}^{n}\sum_{k=1}^{n} b_{i,k}a_{k,i}\\ &= \mathrm{tr}(BA)\end{aligned}$$

Réflexe !

On peut toujours permuter deux sommes finies. Il faut en revanche parfois veiller aux indices qui peuvent dépendre l'un de l'autre. Dans ce cas, ne pas hésiter à réécrire les deux sommes sous la forme d'une somme double avant. En gros, le schéma est le suivant :

$$\begin{aligned}\sum_i\sum_j &= \sum_{i,j}\\ &= \sum_j\sum_i\end{aligned}$$

Pour info...

On en déduit que si M et N sont semblables, alors $\mathrm{tr}(M) = \mathrm{tr}(N)$. En effet, si M et N sont semblables, il existe P inversible telle que $M = PNP^{-1}$.
Ainsi, en utilisant $\mathrm{tr}(AB) = \mathrm{tr}(BA)$, avec $A = P^{-1}$ et $B = NP$, on a :

$$\begin{aligned}\mathrm{tr}(M) &= \mathrm{tr}(P^{-1}NP)\\ &= \mathrm{tr}(NPP^{-1})\\ &= \mathrm{tr}(N)\end{aligned}$$

.

14 ♣ Matrices stochastiques

Définition. On dit qu'une matrice carrée est **stochastique** lorsque ses coefficients sont positifs ou nuls et que la somme des coefficients de chaque ligne est égale à 1.
Résultat. Le produit de deux matrices stochastiques est une matrice stochastique.

Démonstration : Soient $n \in \mathbb{N}^*$ et $A, B \in \mathcal{M}_n(\mathbb{R})$ deux matrices stochastiques. Notons $C = AB$ et $C = (c_{i,j})_{i,j\in[\![1;n]\!]}$.

- C est bien une matrice carrée...
- On a, pour tous $i, j \in [\![1;n]\!]$: $c_{i,j} = \sum_{k=1}^{n} a_{i,k}b_{k,j}$. Or A et B sont stochastiques, donc leurs coefficients sont positifs.

  Par conséquent :

  $$\forall i, j \in [\![1;n]\!],\ c_{i,j} \geqslant 0$$

- Également, pour tout $i \in [\![1;n]\!]$:

$$\begin{aligned}\sum_{j=1}^{n} c_{i,j} &= \sum_{j=1}^{n}\sum_{k=1}^{n} a_{i,k}b_{k,j} && \text{permutation des sommes}\\ &= \sum_{k=1}^{n}\sum_{j=1}^{n} a_{i,k}b_{k,j}\\ &= \sum_{k=1}^{n} a_{i,k}\sum_{j=1}^{n} b_{k,j} && B \text{ est stochastique}\\ &= \sum_{k=1}^{n} a_{i,k} && A \text{ est stochastique}\\ &= 1\end{aligned}$$

Par conséquent , C est stochastique.

15 ♣ ÉQUIVALENT DE LA SUITE DES SOMMES PARTIELLES DE LA SÉRIE HARMONIQUE

On a :

$$\sum_{k=1}^{n} \frac{1}{k} \underset{n\to+\infty}{\sim} \ln(n)$$

★ Classique ! ★
On utilise une comparaison série/intégrale pour déterminer un équivalent de la suite des sommes partielles d'une série divergente, ou un équivalent du reste d'une série convergente.

DÉMONSTRATION : Pour tout $n \in [\![2;+\infty[\![$, on pose $S_n = \sum_{k=1}^{n} \frac{1}{k}$.

Soit $n \in [\![2;+\infty[\![$.

- Soit $k \in [\![1;n]\!]$. Par décroissance de la fonction inverse sur $\mathbb{R}_*^+$, donc sur $[k;k+1]$, on a :

$$\frac{1}{k+1} \leqslant \frac{1}{x} \leqslant \frac{1}{k}$$

Puis, par croissance de l'intégrale, licite car $k \leqslant k+1$:

$$\int_k^{k+1} \frac{1}{k+1}dx \leqslant \int_k^{k+1} \frac{1}{x}dx \leqslant \int_k^{k+1} \frac{1}{k}dx$$

Autrement dit :

$$\frac{1}{k+1} \leqslant \ln(k+1) - \ln(k) \leqslant \frac{1}{k}$$

On aime...
Il est toujours agréable de voir des schémas sur la copie, cette étape en est l'occasion...

À retenir...
On retient la méthode mise en place pour établir cet encadrement classique !

- On a donc :

$$\forall k \in [\![1;n]\!],\ \frac{1}{k+1} \leqslant \ln(k+1) - \ln(k) \leqslant \frac{1}{k}$$

D'où, en sommant de 1 à $n-1$, licite car $n \geqslant 2$, et par télescopage :

$$\sum_{k=1}^{n-1} \frac{1}{k+1} \leqslant \ln(n) \leqslant \sum_{k=1}^{n-1} \frac{1}{k}$$

Or :

  - ⋆ Avec le changement d'indice $i = k+1$, on a :

$$\sum_{k=1}^{n-1} \frac{1}{k+1} = \sum_{i=2}^{n} \frac{1}{i} = S_n - 1$$

  - ⋆ et :

$$\sum_{k=1}^{n-1} \frac{1}{k} = S_n - \frac{1}{n}$$

D'où :

$$S_n - 1 \leqslant \ln(n) \leqslant S_n - \frac{1}{n}$$

-
  - ⋆ De l'inégalité de gauche, on déduit :

$$S_n \leqslant \ln(n) + 1$$

  - ⋆ De l'inégalité de droite, on déduit :

$$\ln(n) + \frac{1}{n} \leqslant S_n$$

- On obtient ainsi :

$$\ln(n) + \frac{1}{n} \leqslant S_n \leqslant \ln(n) + 1$$

Et, comme $n \geqslant 2$, on a $\ln(n) > 0$. D'où :

$$1 + \frac{1}{n\ln(n)} \leqslant \frac{S_n}{\ln(n)} \leqslant 1 + \frac{1}{\ln(n)}$$

On a finalement établi :

$$\forall n \in [\![2;+\infty[\![,\ 1 + \frac{1}{n\ln(n)} \leqslant \frac{S_n}{\ln(n)} \leqslant 1 + \frac{1}{\ln(n)}$$

Mais :

$$\lim_{n\to+\infty} \left(1 + \frac{1}{n\ln(n)}\right) = 1 = \lim_{n\to+\infty} \left(1 + \frac{1}{\ln(n)}\right)$$

Par théorème d'encadrement, on conclut $\lim_{n\to+\infty} \frac{S_n}{\ln(n)} = 1$ et ainsi :

$$S_n \underset{n\to+\infty}{\sim} \ln(n)$$

Conclusion : $\sum_{k=1}^{n} \frac{1}{k} \underset{n\to+\infty}{\sim} \ln(n)$.

Chapitre 2 – Couples de variables aléatoires discrètes

16 ♥ Stabilité des lois de Poisson

Soient $(\Omega, \mathcal{A}, \mathbb{P})$ un espace probabilisé, $\lambda, \mu \in \mathbb{R}_*^+$ ainsi que X, Y deux variables aléatoires sur $(\Omega, \mathcal{A}, \mathbb{P})$.
Si X et Y sont **indépendantes** telles que $\boldsymbol{X \hookrightarrow \mathscr{P}(\lambda)}$ et $\boldsymbol{Y \hookrightarrow \mathscr{P}(\mu)}$; alors : $X + Y \hookrightarrow \mathscr{P}(\lambda + \mu)$.

Démonstration :

- Puisque $X \hookrightarrow \mathscr{P}(\lambda)$ et $Y \hookrightarrow \mathscr{P}(\mu)$, on a : $X(\Omega) = \mathbb{N}$ et $Y(\Omega) = \mathbb{N}$.
  Par conséquent : $(X+Y)(\Omega) \subset \mathbb{N}$.
- Soit $n \in \mathbb{N}$. D'après la formule des probabilités totales, avec $\big([X = k]\big)_{k\in\mathbb{N}}$ comme système complet d'évènements, on a :

$$\begin{aligned}
\mathbb{P}\big([X+Y=n]\big) &= \sum_{k=0}^{+\infty} \mathbb{P}\big([X+Y=n] \cap [X=k]\big) \\
&= \sum_{k=0}^{+\infty} \mathbb{P}\big([Y=n-k] \cap [X=k]\big) \\
&= \sum_{k=0}^{+\infty} \mathbb{P}\big([Y=n-k]\big)\mathbb{P}\big([X=k]\big) && \text{indépendance de } X \text{ et } Y \\
&= \sum_{k=0}^{n} \mathrm{e}^{-\mu}\frac{\mu^{n-k}}{(n-k)!}\mathrm{e}^{-\lambda}\frac{\lambda^k}{k!} && \begin{array}{l}\forall k \in \mathbb{N} : n-k \in Y(\Omega) \iff k \leqslant n \\ \text{Donc : } \forall k \in [\![n+1; +\infty[\![,\ \mathbb{P}\big([Y=n-k]\big) = 0. \\ \forall k \in \mathbb{N},\ k \in X(\Omega)\end{array} \\
&= \mathrm{e}^{-(\mu+\lambda)} \sum_{k=0}^{n} \frac{1}{k!(n-k)!}\lambda^k \mu^{n-k} \\
&= \frac{1}{n!}\mathrm{e}^{-(\lambda+\mu)} \sum_{k=0}^{n} \binom{n}{k}\lambda^k\mu^{n-k} \\
&= \frac{1}{n!}\mathrm{e}^{-(\lambda+\mu)}(\lambda+\mu)^n && \text{formule du binôme de Newton}
\end{aligned}$$

- Pour finir, montrons que $(X+Y)(\Omega) = \mathbb{N}$.
  $\boxed{\subset}$ Déjà établie.
  $\boxed{\supset}$ Soit $n \in \mathbb{N}$. D'après ce qui précède : $\mathbb{P}\big([X+Y=n]\big) = \dfrac{1}{n!}\mathrm{e}^{-(\lambda+\mu)}(\lambda+\mu)^n$. Ainsi :
  $$\mathbb{P}\big([X+Y=n]\big) \neq 0$$
  Par conséquent :
  $$[X+Y=n] \neq \varnothing$$
  Et donc :
  $$n \in (X+Y)(\Omega)$$
  D'où l'inclusion recherchée.

Conclusion : $(X+Y)(\Omega) = \mathbb{N}$ et pour tout $n \in \mathbb{N}$, $\mathbb{P}\big([X+Y=n]\big) = \mathrm{e}^{-(\lambda+\mu)}\dfrac{(\lambda+\mu)^n}{n!}$.
Ainsi : $X + Y \hookrightarrow \mathscr{P}(\lambda+\mu)$.

★Subtil...★
On ne peut pas conclure sur l'égalité pour l'instant... En effet, on sait qu'il existe au moins une issue réalisant $[X = 0]$, et au moins une réalisant $[Y = 0]$... Mais rien ne dit qu'il y a bien au moins une issue commune à ces deux évènements.

✘ Attention !
On veut remplacer $\mathbb{P}\big([X=k]\big)$ et $\mathbb{P}\big([Y=n-k]\big)$...
Et si $Z \hookrightarrow \mathscr{P}(\lambda)$, alors, pour tout $z \in \mathbb{R}$:
$$\mathbb{P}([Z=z]) = \begin{cases} \mathrm{e}^{-\lambda}\dfrac{\lambda^z}{z!} & \text{si } z \in \mathbb{N} \\ 0 & \text{si } z \notin \mathbb{N} \end{cases}$$
De façon générale, quand on veut remplacer une probabilité $\mathbb{P}([Z=z])$, on regarde toujours si $z \in Z(\Omega)$ ou non !

✘ Attention !
On sait que
$$A = \varnothing \Longrightarrow \mathbb{P}(A) = 0$$
et donc sa contraposée :
$$\mathbb{P}(A) \neq 0 \Longrightarrow A \neq \varnothing$$
Les réciproques sont fausses (penser aux évènements quasi-impossibles).

♣ Méthode !
Voici comment on justifie rapidement l'autre inclusion une fois les probabilités connues. Son caractère simple et indépendant du contexte justifie que cette étape n'est, en pratique, pas indispensable...

17 ♠ Minimum de VA indépendantes suivant une loi géométrique

Soit $(\Omega, \mathcal{A}, \mathbb{P})$ un espace probabilisé. Soient $p \in]0;1[$, $q = 1 - p$ ainsi que X et Y deux variables aléatoires sur $(\Omega, \mathcal{A}, \mathbb{P})$. On pose $M = \min(X, Y)$ et on admet que M est une variable aléatoire sur $(\Omega, \mathcal{A}, \mathbb{P})$.
Si X et Y sont **indépendantes** et que : $\boldsymbol{X \hookrightarrow \mathscr{G}(p)}$ et $\boldsymbol{Y \hookrightarrow \mathscr{G}(p)}$; alors la variable aléatoire M suit la loi géométrique de paramètre $1 - q^2$.

Démonstration :

- Puisque X et Y suivent des lois géométriques, on a : $X(\Omega) = \mathbb{N}^*$ et $Y(\Omega) = \mathbb{N}^*$. Par conséquent : $M(\Omega) \subset \mathbb{N}^*$.
- Soit $n \in \mathbb{N}$. On a :
  $$[M > n] = [X > n] \cap [Y > n]$$
  Ainsi :
  $$\begin{aligned}
  \mathbb{P}\big([M>n]\big) &= \mathbb{P}\big([X>n] \cap [Y>n]\big) \\
  &= \mathbb{P}\big([X>n]\big)\mathbb{P}\big([Y>n]\big) && \text{indépendance de } X \text{ et } Y \\
  &= \Big(\mathbb{P}\big([X>n]\big)\Big)^2 && X \text{ et } Y \text{ suivent la même loi}
  \end{aligned}$$
  Or $X(\Omega) = \mathbb{N}^*$, donc :
  $$[X > n] = \bigcup_{k=n+1}^{+\infty} [X = k]$$

★Subtil...★
On ne peut pas encore conclure sur l'égalité...

♣ L'idée !
On travaille sur $[M > n]$, facile à interpréter... L'objectif étant d'obtenir $\mathbb{P}\big([M=n]\big)$, il faudra un lien entre $[M = n]$ et $[M > n]$.

Ainsi, par incompatibilité des évènements de la famille $\big([X=k]\big)_{k\in\mathbb{N}^*}$, la série $\sum\limits_{k\geqslant n+1} \mathbb{P}\big([X=k]\big)$ est convergente et :

$$\begin{aligned}
\mathbb{P}\big([X>n]\big) &= \sum_{k=n+1}^{+\infty} \mathbb{P}\big([X=k]\big) \\
&= \sum_{k=n+1}^{n} q^{k-1}p && \text{changement d'indice } i=k-(n+1) \\
&= p\sum_{i=0}^{+\infty} q^{i+n} \\
&= pq^n \frac{1}{1-q} && p=1-q \\
&= q^n
\end{aligned}$$

On obtient ainsi :

$$\begin{aligned}
\mathbb{P}\big([M>n]\big) &= \Big(1-\big(1-q^n\big)\Big)^2 \\
&= q^{2n}
\end{aligned}$$

- Soit ensuite $n \in \mathbb{N}^*$.
  - ⋆ On a :
  $$[M \geqslant n] = [M=n] \cup [M>n]$$
  Or, les évènements $[M=n]$ et $[M>n]$ sont incompatibles, d'où :
  $$\mathbb{P}\big([M\geqslant n]\big) = \mathbb{P}\big([M=n]\big) + \mathbb{P}\big([M>n]\big)$$
  - ⋆ Et ainsi :
  $$\begin{aligned}
  \mathbb{P}\big([M=n]\big) &= \mathbb{P}\big([M\geqslant n]\big) - \mathbb{P}\big([M>n]\big) && M \text{ est à valeurs entières} \\
  &= \mathbb{P}\big([M>n-1]\big) - \mathbb{P}\big([M>n]\big) && \text{résultat encadré, licite car } n-1, n \in \mathbb{N}\ (n\in\mathbb{N}^*) \\
  &= q^{2n-2} - q^{2n} \\
  &= q^{2n-2}(1-q^2) \\
  &= \big(1-(1-q^2)\big)^{n-1}(1-q^2)
  \end{aligned}$$

> **★ Classique ! ★**
> L'obtention de $\mathbb{P}\big([M=n]\big) = \mathbb{P}\big([M\geqslant n-1]\big) - \mathbb{P}\big([M\geqslant n]\big)$ est un grand classique, à maitriser parfaitement sans oublier d'argument !

- Pour finir, montrons que $M(\Omega) = \mathbb{N}^*$.
  - $\boxed{\subset}$ Déjà établie.
  - $\boxed{\supset}$ Soit $n\in\mathbb{N}^*$. D'après ce qui précède : $\mathbb{P}\big([M=n]\big) = \big(1-(1-q^2)\big)^{n-1}(1-q^2)$. Ainsi :
  $$\mathbb{P}\big([M=n]\big) \neq 0$$
  Par conséquent :
  $$[M=n] \neq \varnothing$$
  Et donc :
  $$n \in M(\Omega)$$
  D'où l'inclusion recherchée.

> **✘ Attention !**
> On sait que
> $$A=\varnothing \Longrightarrow \mathbb{P}(A)=0$$
> et donc sa contraposée :
> $$\mathbb{P}(A)\neq 0 \Longrightarrow A\neq\varnothing$$
> Les réciproques sont fausses (penser aux évènements quasi-impossibles).

Conclusion : $M(\Omega)=\mathbb{N}^*$ et pour tout $n\in\mathbb{N}^*$, $\mathbb{P}\big([M=n]\big) = \big(1-(1-q^2)\big)^{n-1}(1-q^2)$.
Ainsi : $M \hookrightarrow \mathscr{G}(1-q^2)$.

⋆

18 ♥ Somme de VA indépendantes de loi de Bernoulli

Soient $(\Omega, \mathcal{A}, \mathbb{P})$ un espace probabilisé, $p\in]0;1[$ et $(X_k)_{k\in\mathbb{N}^*}$ une suite de variables aléatoires sur $(\Omega,\mathcal{A},\mathbb{P})$.
Si $(X_k)_{k\in\mathbb{N}^*}$ est une suite de variables aléatoires **indépendantes** telles que : $\forall k \in \mathbb{N}^*,\ X_k \hookrightarrow \mathscr{B}(p)$; alors :

$$\forall n\in\mathbb{N}^*,\ \sum_{k=1}^{n} X_k \hookrightarrow \mathscr{B}(n;p)$$

⋆ Démonstration : Donnons deux démonstrations de ce résultat. Pour tout $n\in\mathbb{N}^*$, notons $S_n = \sum\limits_{k=1}^{n} X_k$.

1. **Démonstration 1.**
   - Soit $n\in\mathbb{N}^*$. On sait que pour tout $k\in[\![1;n]\!]$, $X_k(\Omega)=\{0;1\}$, donc $S_n(\Omega)\subset[\![0;n]\!]$.
   - Démontrons par récurrence :
   $$\forall n\in\mathbb{N}^*,\ \forall k\in[\![0;n]\!],\ \mathbb{P}\big([S_n=k]\big) = \binom{n}{k}p^k(1-p)^{n-k}$$
     - ⋆ **Initialisation.** Pour $n=1$:
     On a $S_1 = X_1$. D'où :
     $$\begin{aligned}
     \mathbb{P}\big([S_1=0]\big) &= 1-p \\
     &= \binom{1}{0}p^0(1-p)^{1-0}
     \end{aligned}$$
     et :
     $$\begin{aligned}
     \mathbb{P}\big([S_1=1]\big) &= p \\
     &= \binom{1}{1}p^1(1-p)^{1-1}
     \end{aligned}$$
     L'initialisation est ainsi vérifiée.

> **Petite remarque**
> On pense à une récurrence puisque $S_{n+1} = S_n + X_{n+1}$...

- ⋆ **Hérédité.** Soit $n \in \mathbb{N}^*$.
Supposons "$\forall k \in [\![0;n]\!],\ \mathbb{P}\big([S_n = k]\big) = \binom{n}{k} p^k (1-p)^{n-k}$" ;
et montrons "$\forall k \in [\![0;n+1]\!],\ \mathbb{P}\big([S_{n+1} = k]\big) = \binom{n+1}{k} p^k (1-p)^{n+1-k}$".
Soit $k \in [\![0;n+1]\!]$. Remarquons que $S_{n+1} = S_n + X_{n+1}$. On a alors :

$$\begin{aligned}
\mathbb{P}\big([S_{n+1} = k]\big) &= \mathbb{P}\big([S_n + X_{n+1} = k]\big)\\
&= \mathbb{P}\big([S_n + X_{n+1} = k] \cap [X_{n+1} = 0]\big) + \mathbb{P}\big([S_n + X_{n+1} = k] \cap [X_{n+1} = 1]\big)\\
&= \mathbb{P}\big([S_n = k] \cap [X_{n+1} = 0]\big) + \mathbb{P}\big([S_n = k-1] \cap [X_{n+1} = 1]\big)\\
&= \mathbb{P}\big([S_n = k]\big)\mathbb{P}\big([X_{n+1} = 0]\big) + \mathbb{P}\big([S_n = k-1]\big)\mathbb{P}\big([X_{n+1} = 1]\big)\\
&= (1-p)\mathbb{P}\big([S_n = k]\big) + p\mathbb{P}\big([S_n = k-1]\big)
\end{aligned}$$

formule des probabilités totales avec $\big([X_{n+1} = 0], [X_{n+1} = 1]\big)$ comme sce

$X_1, \dots, X_{n+1}$ sont indépendantes donc, par lemme des coalitions, S_n et X_{n+1} sont indépendantes

✘ Attention !
$\mathbb{P}\big([S_n = i]\big) = 0$ si $i \notin [\![0;n]\!]$...

Distinguons maintenant trois cas :
  - ◇ Si $k = 0$:
  Dans ce cas, $\mathbb{P}\big([S_n = k-1]\big) = \mathbb{P}\big([S_n = -1]\big) = 0$, car $S_n(\Omega) \subset [\![0;n]\!]$.
  Et alors :

$$\begin{aligned}
\mathbb{P}\big([S_{n+1} = 0]\big) &= (1-p)\mathbb{P}\big([S_n = 0]\big)\\
&= (1-p)\binom{n}{0} p^0 (1-p)^n\\
&= \binom{n+1}{0} p^0 (1-p)^{n+1}
\end{aligned}$$

par hypothèse de récurrence

  - ◇ Si $k = n+1$:
  Dans ce cas, $\mathbb{P}\big([S_n = k]\big) = \mathbb{P}\big([S_n = n+1]\big) = 0$, car $S_n(\Omega) \subset [\![0;n]\!]$.
  Et alors :

$$\begin{aligned}
\mathbb{P}\big([S_{n+1} = n+1]\big) &= p\mathbb{P}\big([S_n = n]\big)\\
&= p\binom{n}{n} p^n (1-p)^0\\
&= \binom{n+1}{n+1} p^{n+1} (1-p)^{n+1-(n+1)}
\end{aligned}$$

hypothèse de récurrence

  - ◇ Si $k \in [\![1;n]\!]$.
  On obtient :

$$\begin{aligned}
\mathbb{P}\big([S_{n+1} = k]\big) &= (1-p)\mathbb{P}\big([S_n = k]\big) + p\mathbb{P}\big([S_n = k-1]\big)\\
&= (1-p)\binom{n}{k} p^k (1-p)^{n-k} + p\binom{n}{k-1} p^{k-1} (1-p)^{n-(k-1)}\\
&= \left(\binom{n}{k} + \binom{n}{k-1}\right) p^k (1-p)^{n+1-k}\\
&= \binom{n+1}{k} p^k (1-p)^{n+1-k}
\end{aligned}$$

hypothèse de récurrence, avec $k, k-1 \in [\![0;n]\!]$

triangle de Pascal

Dans les trois cas, on a bien établi :

$$\mathbb{P}\big([S_{n+1} = k]\big) = \binom{n+1}{k} p^k (1-p)^{n+1-k}$$

L'hérédité est ainsi établie.
- ⋆ **Conclusion :** $\forall n \in \mathbb{N},\ \forall k \in [\![0;n]\!],\ \mathbb{P}\big([S_n = k]\big) = \binom{n}{k} p^k (1-p)^{n-k}$.

✘ Attention !
On sait que
$$A = \varnothing \Longrightarrow \mathbb{P}(A) = 0$$
et donc sa contraposée :
$$\mathbb{P}(A) \neq 0 \Longrightarrow A \neq \varnothing$$
Les réciproques sont fausses (penser aux évènements quasi-impossibles).

- D'après le point précédent, on a également : $\forall n \in \mathbb{N}^*,\ \forall k \in [\![0;n]\!],\ \mathbb{P}\big([S_n = k]\big) \neq 0$.
D'où :
$$\forall n \in \mathbb{N}^*,\ \forall k \in [\![0;n]\!],\ [S_n = k] \neq \varnothing$$
Et ainsi :
$$\forall n \in \mathbb{N}^*,\ [\![0;n]\!] \subset S_n(\Omega)$$

Conclusion : pour tout $n \in \mathbb{N}^*$, $S_n \hookrightarrow \mathscr{B}(n;p)$.

2. **Démonstration 2.**
- Soit $n \in \mathbb{N}^*$.
  - ⋆ On sait que pour tout $k \in [\![1;n]\!]$, $X_k(\Omega) = \{0;1\}$, donc $S_n(\Omega) \subset [\![0;n]\!]$.
  - ⋆ Soit $k \in [\![0;n]\!]$.

$[S_n = k]$ est réalisé si, et seulement si, la somme des réalisations des variables aléatoires $X_1, \dots, X_n$ est égale à k
si, et seulement si, k des n variables aléatoires $X_1, \dots, X_n$ prennent 1 comme valeur, les autres prennent 0

$\forall i \in [\![1;n]\!],\ X_i(\Omega) = \{0;1\}$

L'évènement $[S_n = k]$ est alors constitué de $\binom{n}{k}$ issues différentes ayant toutes la même probabilité d'apparition, égale à la probabilité $\mathbb{P}\left(\left(\bigcap_{i=1}^{k}[X_i = 1]\right) \cap \left(\bigcap_{i=k+1}^{n}[X_i = 0]\right)\right)$.
Or, par indépendance des variables aléatoires $X_1, X_2, \dots, X_n$, on a :

$$\begin{aligned}
\mathbb{P}\left(\left(\bigcap_{i=1}^{k}[X_i = 1]\right) \cap \left(\bigcap_{i=k+1}^{n}[X_i = 0]\right)\right) &= \left(\prod_{i=1}^{k}\mathbb{P}\big([X_i = 1]\big)\right) \times \left(\prod_{i=k+1}^{n}\mathbb{P}\big([X_i = 0]\big)\right)\\
&= p^k (1-p)^{n-k}
\end{aligned}$$

$\forall i \in [\![1;n]\!],\ X_i \hookrightarrow \mathscr{B}(p)$

☞ Rappel...
Si a et b sont des entiers tels que $a \leqslant b$, alors :
$\mathrm{Card}\big([\![a;b]\!]\big) = b - a + 1$

Par conséquent :

$$\mathbb{P}([S_n = k]) = \binom{n}{k} p^k (1-p)^{n-k}$$

- D'après le point précédent, on a également : $\forall n \in \mathbb{N}^*,\ \forall k \in [\![0;n]\!],\ \mathbb{P}([S_n = k]) \neq 0$.
  D'où :
  $$\forall n \in \mathbb{N}^*,\ \forall k \in [\![0;n]\!],\ [S_n = k] \neq \varnothing$$
  Et ainsi :
  $$\forall n \in \mathbb{N}^*,\ [\![0;n]\!] \subset S_n(\Omega)$$

Conclusion : pour tout $n \in \mathbb{N}^*$, $S_n \hookrightarrow \mathscr{B}(n;p)$.

✘ Attention !
On sait que
$$A = \varnothing \Longrightarrow \mathbb{P}(A) = 0$$
et donc sa contraposée :
$$\mathbb{P}(A) \neq 0 \Longrightarrow A \neq \varnothing$$
Les réciproques sont fausses (penser aux évènements quasi-impossibles).

19 ♣ Inégalité de Cauchy-Schwarz pour la covariance

Soient $(\Omega, \mathcal{A}, \mathbb{P})$ un espace probabilisé et X, Y deux variables aléatoires discrètes sur $(\Omega, \mathcal{A}, \mathbb{P})$.
Si X et Y **admettent une variance**, alors :

$$\big(\mathrm{Cov}(X,Y)\big)^2 \leqslant \mathbb{V}(X)\mathbb{V}(Y)$$

Démonstration : Supposons que X et Y admettent une variance.
Soit $t \in \mathbb{R}$. La variable aléatoire $tX + Y$ est une combinaison linéaire de deux variables aléatoires admettant une variance, par conséquent, elle admet également une variance et :

$$\begin{aligned}\mathbb{V}(tX+Y) &= \mathbb{V}(tX) + 2\mathrm{Cov}(tX,Y) + \mathbb{V}(Y) \\ &= t^2\mathbb{V}(X) + 2t\mathrm{Cov}(X,Y) + \mathbb{V}(Y)\end{aligned}$$

linéarité à gauche de la covariance

Distinguons deux cas :

- Si $\mathbb{V}(X) = 0$.
  Dans ce cas, d'après ce qui précède :
  $$\forall t \in \mathbb{R},\ \mathbb{V}(tX+Y) = 2t\mathrm{Cov}(X,Y) + \mathbb{V}(Y)$$
  Or, on sait qu'une variance est toujours positive, d'où :
  $$\forall t \in \mathbb{R},\ 2t\mathrm{Cov}(X,Y) + \mathbb{V}(Y) \geqslant 0$$
  Ce qui n'est possible que si $\mathrm{Cov}(X,Y) = 0$. L'inégalité de Cauchy-Schwarz est alors vérifiée.
- Si $\mathbb{V}(X) \neq 0$.
  Dans ce cas, la fonction $t \longmapsto t^2\mathbb{V}(X) + 2t\mathrm{Cov}(X,Y) + \mathbb{V}(Y)$ est une fonction polynomiale de degré 2 positive sur $\mathbb{R}$, car égale à la fonction $t \longmapsto \mathbb{V}(tX+Y)$ et qu'une variance est toujours positive.
  Par conséquent, son discriminant est négatif ou nul. Or, ce discriminant est égal à $4\big(\mathrm{Cov}(X,Y)\big)^2 - 4\mathbb{V}(X)\mathbb{V}(Y)$...
  On en déduit :
  $$\big(\mathrm{Cov}(X,Y)\big)^2 \leqslant \mathbb{V}(X)\mathbb{V}(Y)$$

Pourquoi ?
Une fonction affine de signe constant est une fonction constante... On peut le justifier rapidement en raisonnant par l'absurde en supposant que $\mathrm{Cov}(X,Y) \neq 0$...

20 ♣ Interprétation du coefficient de corrélation linéaire

Soient $(\Omega, \mathcal{A}, \mathbb{P})$ ainsi que X et Y deux variables aléatoires discrètes sur $(\Omega, \mathcal{A}, \mathbb{P})$.
Si X et Y **admettent une variance non nulle**, alors :

- $\rho(X,Y) = 1$ si, et seulement si, l'une des variables aléatoires est presque-sûrement fonction affine strictement croissante de l'autre ;
- $\rho(X,Y) = -1$ si, et seulement si, l'une des variables aléatoires est presque-sûrement fonction affine strictement décroissante de l'autre ;

Petite remarque
Il est indispensable d'avoir travaillé la démonstration précédente pour comprendre celle-ci. En effet, la présente démonstration utilise une partie du raisonnement mis en place dans QC19.

Démonstration : Supposons que X et Y admettent une variance non nulle.

- Remarquons déjà que :

  Y est presque-sûrement fonction affine de X si, et seulement si, X est presque-sûrement fonction affine de Y.

  En effet :
  - ⋆ si Y est presque-sûrement fonction affine de X, alors il existe deux réels a, b tels que $\mathbb{P}([Y = aX + b]) = 1$.
    Montrons que $a \neq 0$. Raisonnons par l'absurde et supposons $a = 0$.
    Dans ce cas :
    $$\mathbb{P}([Y = b]) = 1$$
    Autrement dit, la variable aléatoire Y est presque-sûrement constante ; et donc de variance nulle : ce qui contredit l'hypothèse initiale. Par conséquent : $a \neq 0$.
    On obtient alors :
    $$\mathbb{P}\left(\left[X = \frac{1}{a}Y - \frac{b}{a}\right]\right)$$
    La variable aléatoire X est donc presque-sûrement fonction affine de Y.
  - ⋆ De manière analogue (ou par symétrie des rôles de X et Y).

  Dans la suite, établissons donc les résultats dans le cas de Y fonction affine de X.

- Ensuite :

$$\begin{aligned}
|\rho(X,Y)| = 1 &\iff \rho(X,Y)^2 = 1 \\
&\iff \left(\frac{\mathrm{Cov}(X,Y)}{\sigma(X)\sigma(Y)}\right)^2 = 1 \\
&\iff \frac{\big(\mathrm{Cov}(X,Y)\big)^2}{\mathbb{V}(X)\mathbb{V}(Y)} = 1 && \text{$\mathbb{V}(X)$ et $\mathbb{V}(Y)$ sont non nulles} \\
&\iff \big(\mathrm{Cov}(X,Y)\big)^2 = \mathbb{V}(X)\mathbb{V}(Y)
\end{aligned}$$

Par conséquent :

$\lvert\rho(X,Y)\rvert = 1$	si, et seulement si,	il y a égalité dans l'inégalité de Cauchy-Schwarz
	si, et seulement si,	le discriminant de $t \longmapsto \mathbb{V}(tX+Y)$ est nul (voir démo précédente)
	si, et seulement si,	$t \longmapsto \mathbb{V}(tX+Y)$ possède une unique racine
	si, et seulement si,	il existe un unique réel α tel que $\mathbb{V}(\alpha X+Y) = 0$
	si, et seulement si,	il existe un unique réel α tel que la variable aléatoire $\alpha X+Y$ est presque-sûrement constante
	si, et seulement si :	$\exists ! \alpha, b \in \mathbb{R}$ / $\mathbb{P}\big([\alpha X+Y = b]\big) = 1$
	si, et seulement si :	$\exists ! a, b \in \mathbb{R}$ / $\mathbb{P}\big([Y = aX+b]\big) = 1$

Considérons ensuite deux tels réels a et b. On a ainsi, d'après la formule de Koenig-Huygens :

$$\begin{aligned}
\mathrm{Cov}(X,Y) &= \mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y) && \mathbb{P}\big([Y = aX+b]\big) = 1 \\
&= \mathbb{E}\big(X(aX+b)\big) - \mathbb{E}(X)\mathbb{E}(aX+b) && \text{linéarité de l'espérance} \\
&= a\mathbb{E}(X^2) + b\mathbb{E}(X) - a\mathbb{E}(X)^2 - b\mathbb{E}(X) \\
&= a\big(\mathbb{E}(X^2) - \mathbb{E}(X)^2\big) && \text{formule de Koenig-Huygens} \\
&= a\mathbb{V}(X)
\end{aligned}$$

Peut être utile...

$\mathbb{P}\big([X = Y]\big) = 1 \Rightarrow \mathbb{E}(X) = \mathbb{E}(Y)$

Et ainsi :

$$\rho(X,Y) = \frac{a\mathbb{V}(X)}{\sigma(X)\sigma(Y)}$$

Enfin, puisque $\mathbb{V}(X) \geqslant 0$, $\sigma(X) > 0$ et $\sigma(Y) > 0$, on obtient que $\rho(X,Y)$ est du signe de a.

Conclusion : $|\rho(X,Y)| = 1$ si, et seulement si, l'une des variables aléatoires est presque-sûrement fonction affine de l'autre ; et, dans ce cas, le coefficient dominant de la fonction affine est du signe de $\rho(X,Y)$.

⋆

21 ♥ Applications linéaires coïncidant sur une base

Soient E, F deux espaces vectoriels de dimension finie ainsi que $f, g \in \mathscr{L}(E, F)$.
Si **f et g coïncident sur une base de E**, alors f et g sont égales.

Autrement dit : Une application linéaire définie sur E est entièrement définie par l'image qu'elle renvoie aux vecteurs d'une base de E.

DÉMONSTRATION : Notons $n = \dim(E)$ et considérons une base $(e_1, e_2, \ldots, e_n)$ de E. Supposons que f et g coïncident sur cette base ; autrement dit, supposons : $\forall i \in [\![1; n]\!],\ f(e_i) = g(e_i)$. Montrons que $f = g$; autrement dit, montrons : $\forall x \in E,\ f(x) = g(x)$.

Soit $x \in E$. Puisque $(e_1, e_2, \ldots, e_n)$ est une base de E, il existe des uniques réels $\lambda_1, \lambda_2, \ldots, \lambda_n$ tels que : $x = \sum_{i=1}^{n} \lambda_i e_i$. On a alors :

$$\begin{aligned}
f(x) &= f\left(\sum_{i=1}^{n} \lambda_i e_i\right) && \text{linéarité de } f \\
&= \sum_{i=1}^{n} \lambda_i f(e_i) && \text{hypothèse} \\
&= \sum_{i=1}^{n} \lambda_i g(e_i) && \text{linéarité de } g \\
&= g\left(\sum_{i=1}^{n} \lambda_i e_i\right) \\
&= g(x)
\end{aligned}$$

On a ainsi établi : $\forall x \in E,\ f(x) = g(x)$.

Petite remarque Puisque l'unicité des λ_i n'a pas été utilisée, seul le caractère générateur de la famille $(e_1, e_2, \ldots, e_n)$ suffit.

22 ♥ Noyau et image d'une application linéaire

Soient E, F deux espaces vectoriels et $f \in \mathscr{L}(E, F)$.

Q1 $\ker(f)$ est un sous-espace vectoriel de E.

Q2 $\mathrm{Im}(f)$ est un sous-espace vectoriel de F.

☞ Rappels...

$$\ker(f) = \{x \in E \ /\ f(x) = 0_F\}$$
$$\mathrm{Im}(f) = \{y \in F / \exists x \in E / y = f(x)\}$$
ou bien :
$$\mathrm{Im}(f) = \{f(x),\ x \in E\}$$

DÉMONSTRATION :

Q1.
- Par définition : $\ker(f) \subset E$.
- Puisque f est linéaire, $f(0_E) = 0_F$. Donc $0_E \in \ker(f)$. Ainsi, $\ker(f)$ est non vide.
- Montrons que $\ker(f)$ est stable par combinaison linéaire.
  Soient $u, v \in \ker(f)$ et $\lambda, \mu \in \mathbb{R}$. Montrons que $\lambda u + \mu v \in \ker(f)$.
  - ⋆ On a déjà $u, v \in E$ et E étant un espace vectoriel, on obtient : $\lambda u + \mu v \in E$.
  - ⋆ Ensuite :

$$\begin{aligned}
f(\lambda u + \mu v) &= \lambda f(u) + \mu f(v) && \text{par linéarité de } f \\
&= \lambda \times 0_F + \mu \times 0_F && \text{car } u, v \in \ker(f) \\
&= 0_F
\end{aligned}$$

  Ainsi :
$$\lambda u + \mu v \in \ker(f)$$

Conclusion : $\ker(f)$ est un sous-espace vectoriel de E.

Q2.
- Par définition : $\mathrm{Im}(f) \subset F$.
- Puisque f est linéaire, on a $f(0_E) = 0_F$. Donc $0_F \in \mathrm{Im}(f)$. Ainsi, $\mathrm{Im}(f)$ est non vide.
- Montrons que $\mathrm{Im}(f)$ est stable par combinaison linéaire.
  Soient $y, z \in \mathrm{Im}(f)$ et $\lambda, \mu \in \mathbb{R}$. Montrons que $\lambda y + \mu z \in \mathrm{Im}(f)$.
  - ⋆ On a déjà $y, z \in F$ et F étant un espace vectoriel, on obtient : $\lambda y + \mu z \in F$.
  - ⋆ Ensuite :
    Puisque $y \in \mathrm{Im}(f)$, il existe $u \in E$ tel que $y = f(u)$. Considérons un tel u.
    Puisque $z \in \mathrm{Im}(f)$, il existe $v \in E$ tel que $z = f(v)$. Considérons un tel v.
    Par conséquent :

$$\begin{aligned}
\lambda y + \mu z &= \lambda f(u) + \mu f(v) && \text{linéarité de } f \\
&= f(\lambda u + \mu v)
\end{aligned}$$

    Or E est un espace vectoriel, donc $\lambda u + \mu v \in E$. Et ainsi, $\lambda y + \mu z \in \mathrm{Im}(f)$.

Conclusion : $\mathrm{Im}(f)$ est un sous-espace vectoriel de F.

23 ♥ Caractérisation de l'injectivité d'une application linéaire

Soient E, F deux espaces vectoriels et $f \in \mathscr{L}(E, F)$.
L'application linéaire f est injective, si, et seulement si, $\ker(f) = \{0_E\}$.

☞ Rappel...
$\ker(f) = \{x \in E \ / \ f(x) = 0_F\}$

⋆ Démonstration : Raisonnons par double implication...

$\boxed{\Longrightarrow}$ Supposons que f est injective. Montrons que $\ker(f) = \{0_E\}$. Raisonnons par double-inclusion.

- $\boxed{\supset}$ Immédiat, car f est linéaire, donc $f(0_E) = 0_F$.
- $\boxed{\subset}$ Soit $x \in \ker(f)$. Ainsi :

$$\begin{aligned} f(x) &= 0_F \\ &= f(0_E) && \text{linéarité de } f \end{aligned}$$

Par injectivité de f, on obtient : $x = 0_E$. D'où : $\ker(f) \subset \{0_E\}$.

Par conséquent : $\ker(f) = \{0_E\}$.

$\boxed{\Longleftarrow}$ Supposons que $\ker(f) = \{0_E\}$. Montrons que f est injective. Soient $x, y \in E$. On a :

☞ Rappel...
f injective sur E :
$\forall x, y \in E, \big(f(x) = f(y) \Rightarrow x = y\big)$

$$\begin{aligned} f(x) = f(y) &\Longrightarrow f(x) - f(y) = 0_F \\ &\Longrightarrow f(x - y) = 0_F && \text{linéarité de } f \\ &\Longrightarrow x - y \in \ker(f) \\ &\Longrightarrow x - y = 0_E && \ker(f) = \{0_E\} \\ &\Longrightarrow x = y \end{aligned}$$

Par conséquent : f est injective.

⋆

24 ♥ Caractérisation des isomorphismes ("triangle de bijectivité")

Soient E et F deux espaces vectoriels de dimension finie et $f \in \mathscr{L}(E, F)$.
Si $\dim(E) = \dim(F)$, alors :

$$\begin{tikzcd} f \text{ injective} \arrow[rr, leftrightarrow] \arrow[dr, leftrightarrow] & & f \text{ surjective} \arrow[dl, leftrightarrow] \\ & f \text{ bijective} & \end{tikzcd}$$

Important !
En pratique, si $\dim(E) = \dim(F)$ (et en particulier si $E = F$), alors on démontrera la bijectivité de f en démontrant son injectivité (par le noyau).

⋆ Démonstration : Supposons que $\dim(E) = \dim(F)$.

- Montrons que l'injectivité de f équivaut à sa surjectivité.
  Puisque E est de dimension finie, d'après le théorème du rang :

$$\dim(E) = \dim\big(\ker(f)\big) + \mathrm{rg}(f)$$

  Ensuite :

$$\begin{aligned} \big(f \text{ est injective}\big) &\iff \ker(f) = \{0_E\} \\ &\iff \dim\big(\ker(f)\big) = 0 \\ &\iff \mathrm{rg}(f) = \dim(E) && \text{théorème du rang} \\ &\iff \mathrm{rg}(f) = \dim(F) && \dim(E) = \dim(F) \\ &\iff \mathrm{Im}(f) = F \\ &\iff \big(f \text{ est surjective}\big) \end{aligned}$$

☞ Rappels...
• Le singleton $\{0_E\}$ est le seul sous-espace vectoriel de E de dimension 0.
• $\mathrm{Im}(f)$ est un ssev de F
• le seul ssev de F de dimension égale à $\dim(F)$ est F lui-même

- Montrons que f est injective si, et seulement si, f est bijective.

  $\boxed{\Longleftarrow}$ On sait déjà que la bijectivité de f implique son injectivité (par définition).

  $\boxed{\Longrightarrow}$ D'après ce qui précède, l'injectivité de f implique sa surjectivité. Par conséquent, si f est injective, elle est également surjective et donc bijective. L'injectivité de f implique donc sa bijectivité.

  Conclusion : f est injective si, et seulement si, f est bijective.
- De la même façon, on démontre que f est surjective si, et seulement si, f est bijective.

⋆

25 ♥ Majoration du rang d'une application linéaire

Soient E et F deux espaces vectoriels de dimension finie ainsi que $f \in \mathscr{L}(E,F)$.
Le rang de f est fini et : $\mathrm{rg}(f) \leqslant \min\big(\dim(E); \dim(F)\big)$.

Démonstration :

- Par définition, $\mathrm{rg}(f) = \dim\big(\mathrm{Im}(f)\big)$. Or $\mathrm{Im}(f)$ est un sous-espace vectoriel de F, qui lui est de dimension finie. Ainsi, $\mathrm{Im}(f)$ est de dimension finie et donc le rang de f est fini.
- Montrons que $\mathrm{rg}(f) \leqslant \dim(E)$ ET $\mathrm{rg}(f) \leqslant \dim(F)$.
  - ⋆ D'après ce qui précède, on a déjà :
$$\mathrm{rg}(f) \leqslant \dim(F)$$
  - ⋆ Notons $n = \dim(E)$ et considérons $(e_1, e_2, \dots, e_n)$ une base de E. On sait que :
$$\mathrm{Im}(f) = \mathrm{Vect}\big(f(e_1), f(e_2), \dots, f(e_n)\big)$$
    Autrement dit, la famille $\big(f(e_1), f(e_2), \dots, f(e_n)\big)$ est génératrice de $\mathrm{Im}(f)$. Et par conséquent :
$$\mathrm{Card}\big(f(e_1), f(e_2), \dots, f(e_n)\big) \geqslant \dim\big(\mathrm{Im}(f)\big)$$
    On a ainsi :
$$\dim(E) \geqslant \mathrm{rg}(f)$$

  Par conséquent :
$$\mathrm{rg}(f) \leqslant \big(\dim(E); \dim(F)\big)$$

À retenir...

$x \leqslant \min(a,b) \iff \begin{cases} x \leqslant a \\ \text{ET} \\ x \leqslant b \end{cases}$

☞ Rappels...

- Si $\mathscr{F}$ est une famille libre de E, alors $\mathrm{Card}(\mathscr{F}) \leqslant \dim(E)$.
- Si $\mathscr{F}$ est une famille génératrice de E, alors $\mathrm{Card}(\mathscr{F}) \geqslant \dim(E)$.

26 Caractérisation des isomorphismes (bis)

Soient E et F deux espaces vectoriels de dimensions finies, ainsi que $f \in \mathscr{L}(E,F)$. Notons $n = \dim(E)$ et considérons $(e_1, e_2, \dots, e_n)$ une base de E.

Q1 f est surjective si, et seulement si, $\big(f(e_1), f(e_2), \dots, f(e_n)\big)$ est génératrice de F.

Q2 f est injective si, et seulement si, $\big(f(e_1), f(e_2), \dots, f(e_n)\big)$ est libre dans F.

Q3 f est un isomorphisme si, et seulement si, $\big(f(e_1), f(e_2), \dots, f(e_n)\big)$ est une base de F.

À retenir...

Une application linéaire est un isomorphisme si, et seulement si, elle transforme une base en une base.
Ce résultat peut servir dans les deux sens : pour justifier rapidement qu'une famille est une base (car elle est l'image d'une base par un isomorphisme) ou pour démontrer qu'une application linéaire est bijective (car elle transforme une base canonique en une autre base par exemple).

Démonstration :

Q1. Puisque $(e_1 \dots, e_n)$ est génératrice de E : $\mathrm{Im}(f) = \mathrm{Vect}\big(f(e_1), \dots, f(e_n)\big)$.
Or : f est surjective si, et seulement si, $\mathrm{Im}(f) = F$.
D'où : f est surjective si, et seulement si, $F = \mathrm{Vect}\big(f(e_1), \dots, f(e_n)\big)$.
Conclusion : f est surjective si, et seulement si, $\big(f(e_1), \dots, f(e_n)\big)$ est génératrice de F.

Petite remarque

Seul le caractère générateur de $(e_1, \dots, e_n)$ suffit ici.

Q2. Raisonnons par double-implication...

$\boxed{\Longrightarrow}$ Supposons f injective. Montrons que $\big(f(e_1), \dots, f(e_n)\big)$ est libre dans F.

Soient $a_1, a_2, \dots, a_n \in \mathbb{R}$. Supposons $\displaystyle\sum_{i=1}^{n} a_i f(e_i) = 0_F$.

Par linéarité de f, on a ainsi :
$$f\left(\sum_{i=1}^{n} a_i e_i\right) = 0_F$$
Par conséquent :
$$\sum_{i=1}^{n} a_i e_i \in \ker(f)$$
Mais f est injective, donc $\ker(f) = \{0_E\}$. D'où :
$$\sum_{i=1}^{n} a_i e_i = 0_E$$
Or la famille $(e_1, \dots, e_n)$ est une base de E, elle est donc en particulier libre. D'où :
$$\forall i \in [\![1;m]\!],\ a_i = 0$$
Conclusion : la famille $\big(f(e_1), \dots, f(e_n)\big)$ est libre.

Petite remarque

Seule la liberté de $(e_1, \dots, e_n)$ suffit ici.

$\boxed{\Longleftarrow}$ Supposons que la famille $\big(f(e_1), \dots, f(e_n)\big)$ est libre. Montrons que f est injective en montrant que $\ker(f) = \{0_E\}$.

Soit $x \in E$. Puisque $(e_1, e_2, \dots, e_n)$ est une base de E, il existe des uniques réels $x_1, x_2, \dots, x_n$ tels que $x = \displaystyle\sum_{i=1}^{n} x_i e_i$.

Ainsi :
$$\begin{aligned} x \in \ker(f) &\iff f(x) = 0_F \\ &\iff f\left(\sum_{i=1}^{n} x_i e_i\right) = 0_F && \text{linéarité de } f \\ &\iff \sum_{i=1}^{n} x_i f(e_i) = 0_F && \text{liberté de la famille } \big(f(e_1), f(e_2), \dots, f(e_n)\big) \end{aligned}$$

$$\iff \forall i \in [\![1;n]\!],\ x_i = 0$$
$$\iff x = 0_E$$

D'où : $\ker(f) = \{0_E\}$.
Conclusion : f est injective.

Q3. Immédiat d'après les deux résultats précédents...

⋆

27 ♠ Un isomorphisme classique

Soient $n \in [\![2;+\infty[\![$ et $a_1, \ldots, a_n$ des réels.

Si $a_1, \ldots, a_n$ sont deux à deux distincts, alors l'application $T : \begin{vmatrix} \mathbb{R}_{n-1}[x] & \longrightarrow & \mathbb{R}^n \\ P & \longmapsto & \big(P(a_1), \ldots, P(a_n)\big) \end{vmatrix}$ est un isomorphisme.

☞ Rappel...
Un isomorphisme est une application linéaire bijective.

⋆

Démonstration : Supposons que les réels $a_1, \ldots, a_n$ sont deux à deux distincts.

- Montrons déjà que T est linéaire. Soient $P, Q \in \mathbb{R}_{n-1}[x]$ et $\lambda, \mu \in \mathbb{R}$.
  On a :
  $$\begin{aligned} T(\lambda P + \mu Q) &= \big((\lambda P + \mu Q)(a_1), \ldots, (\lambda P + \mu Q)(a_n)\big) \\ &= \big(\lambda P(a_1) + \mu Q(a_1), \ldots, \lambda P(a_n) + \mu Q(a_n)\big) \\ &= \lambda\big(Q(a_1), \ldots, Q(a_n)\big) + \mu\big(Q(a_1), \ldots, Q(a_n)\big) \\ &= \lambda T(P) + \mu T(Q) \end{aligned}$$
  linéarité de l'évaluation

  Conclusion : T est une application linéaire.
- Ensuite, on a $\dim\big(R_{n-1}[x]\big) = n = \dim(\mathbb{R}^n)$. Donc pour démontrer que T est bijective, montrons que T est injective...
  Soit $P \in E$. On a :
  $$\begin{aligned} P \in \ker(T) &\iff T(P) = 0 \\ &\iff \forall i \in [\![1;n]\!],\ P(a_i) = 0 \end{aligned}$$
  Ainsi, les réels $a_1, a_2, \ldots, a_n$ sont des racines distinctes de P. La fonction polynomiale P possède donc au moins n racines distinctes ; mais elle est de degré inférieur ou égal à $n-1$. Par conséquent : P est nulle.
  On obtient :
  $$\ker(T) = \{0_E\}$$
  L'application T est ainsi injective. D'après la remarque du départ, T est ainsi bijective.

☞ Rappel...
Le triangle de bijectivité est valable dès lors que espaces de départ et d'arrivée ont même dimension ; il n'est pas nécessaire qu'ils soient identiques !

⋆

28 ♠ Indice de nilpotence d'un endomorphisme

Soient E un espace vectoriel de dimension finie notée n et f un endomorphisme de E.
On suppose qu'il existe $k \in [\![2;+\infty[\![$ tel que : $f^k = 0_{\mathscr{L}(E)}$ et $f^{k-1} \neq 0_{\mathscr{L}(E)}$. Dans ce cas : $k \leqslant n$.

Confusion d'objets !
On rappelle que, lorsqu'il s'agit d'endomorphismes, f^k désigne $\underbrace{f \circ f \circ \ldots \circ f}_{k \text{ fois}}$.

⋆

Démonstration :

- Puisque f^{k-1} n'est pas l'endomorphisme nul, il existe un vecteur $x \in E$ tel que $f^{k-1}(x) \neq 0_E$.
  Considérons donc un tel x et montrons que la famille $\big(x, f(x), \ldots, f^{k-1}(x)\big)$ est libre.
  Soient $a_0, a_1, \ldots, a_{k-1} \in \mathbb{R}$. Supposons que $a_0 x + a_1 f(x) + \ldots + a_{k-1} f^{k-1}(x) = 0_E$ et notons $(\star)$ cette égalité.
  - ⋆ En appliquant f^{k-1} à cette égalité, on obtient, par linéarité de f^{k-1} :
    $$a_0 f^{k-1}(x) + a_1 f^k(x) + \ldots + a_{k-1} f^{2k-2}(x) = 0_E$$
    Puisque $f^k = 0_{\mathscr{L}(E)}$, on a aussi, par récurrence immédiate :
    $$\forall i \in [\![k;+\infty[\![,\ f^i = 0_{\mathscr{L}(E)}$$
    D'où :
    $$a_0 f^{k-1}(x) = 0_E$$
    Et puisque $f^{k-1}(x) \neq 0_E$, on obtient :
    $$a_0 = 0$$
  - ⋆ Par conséquent, l'égalité $(\star)$ devient :
    $$a_1 f(x) + \ldots + a_{k-1} f^{k-1}(x) = 0_E \qquad (\star)$$
    Puis, en appliquant f^{k-2} à cette égalité, on va obtenir, de façon analogue :
    $$a_1 = 0$$
  - ⋆ Et en réitérant, on obtiendra successivement $a_2 = 0$, puis $a_3 = 0$,..., jusqu'à $a_{k-2} = 0$. Restera alors :
    $$a_{k-1} f^{k-1}(x) = 0_E$$
    Et comme $f^{k-1}(x) \neq 0_E$, on aura $a_{k-1} = 0$.

  Conclusion : $\forall i \in [\![0;k-1]\!],\ a_i = 0$. La famille $\big(x, f(x), \ldots, f^{k-1}(x)\big)$ est donc libre.
- Puisque $\big(x, f(x), \ldots, f^{k-1}(x)\big)$ est libre, on a :
  $$\mathrm{Card}\big(x, f(x), \ldots, f^{k-1}(x)\big) \leqslant \dim(E)$$
  Autrement dit :
  $$k \leqslant n$$

☞ Rappels...
- Si $\mathscr{F}$ est une famille libre de E, alors $\mathrm{Card}(\mathscr{F}) \leqslant \dim(E)$.
- Si $\mathscr{F}$ est une famille génératrice de E, alors $\mathrm{Card}(\mathscr{F}) \geqslant \dim(E)$.

Petite remarque
On peut bien évidemment adapter la démonstration dans le cas d'une matrice...

⋆

29 Transformée carrée d'une variable aléatoire à densité

Soient $(\Omega, \mathcal{A}, \mathbb{P})$ un espace probabilisé et X une variable aléatoire à densité de densité f_X, continue sur $\mathbb{R}$.
Dans ce cas, la variable aléatoire X^2 est une variable aléatoire à densité et admet pour densité la fonction
$$f : x \longmapsto \begin{cases} \frac{1}{2\sqrt{x}}\big(f_X(\sqrt{x}) + f_X(-\sqrt{x})\big) & \text{si } x > 0 \\ 0 & \text{si } x \leqslant 0 \end{cases}.$$

★ Démonstration : Notons $Y = X^2$ ainsi que F_X et F_Y les fonctions de répartition des variables aléatoires X et Y respectivement.

- Puisque $Y = X^2$, on a $Y(\Omega) \subset \mathbb{R}^+$.
- Soit $x \in \mathbb{R}$.
  - ⋆ Si $x < 0$:
$$\begin{aligned} F_Y(x) &= \mathbb{P}\big([Y \leqslant x]\big) \\ &= \mathbb{P}(\varnothing) \quad \text{↲ } Y(\Omega) \subset \mathbb{R}^+ \text{ et } x < 0 \\ &= 0 \end{aligned}$$
  - ⋆ Si $x \geqslant 0$:
$$\begin{aligned} F_Y(x) &= \mathbb{P}\big([Y \leqslant x]\big) \\ &= \mathbb{P}\big([X^2 \leqslant x]\big) \\ &= \mathbb{P}\big([-\sqrt{x} \leqslant X \leqslant \sqrt{x}]\big) \quad \text{↲ } x \geqslant 0 \\ &= F_X(\sqrt{x}) - F_X(-\sqrt{x}) \quad \text{↲ } X \text{ est à densité} \end{aligned}$$

  Conclusion : $\forall x \in \mathbb{R},\ F_Y(x) = \begin{cases} F_X(\sqrt{x}) - F_X(-\sqrt{x}) & \text{si } x \geqslant 0 \\ 0 & \text{si } x < 0 \end{cases}$
- Ensuite :
  - ⋆ La fonction F_Y est :
    - ◇ continue sur $]-\infty; 0[$ car constante sur cet intervalle,
    - ◇ continue sur $]0; +\infty[$ comme composée de fonctions continues sur les intervalles adéquats, F_X l'étant puisque X est à densité,
    - ◇ continue en 0 car :
      d'une part
$$\lim_{\substack{x \to 0 \\ x<0}} F_Y(x) = 0$$
      et d'autre part
$$\begin{aligned} \lim_{\substack{x \to 0 \\ x>0}} F_Y(x) &= \lim_{\substack{x \to 0 \\ x>0}} F_X(\sqrt{x}) - F_X(-\sqrt{x}) \quad \text{↲ } F_X \text{ est continue en } 0 \\ &= F_X(0) - F_X(0) \\ &= 0 \end{aligned}$$
      de sorte que
$$\lim_{\substack{x \to 0 \\ x<0}} F_Y(x) = F_Y(0) = \lim_{\substack{x \to 0 \\ x>0}} F_Y(x)$$

    Par conséquent, la fonction F_Y est continue sur $\mathbb{R}$.
  - ⋆ Par des arguments similaires à la continuité, la fonction F_Y est $\mathscr{C}^1$ sur $\mathbb{R}$ sauf éventuellement en 0 (la fonction F_X est $\mathscr{C}^1$ sur $\mathbb{R}$ car f_X est continue sur $\mathbb{R}$).

  On en déduit que la variable aléatoire Y est à densité et on obtient une de ses densités, notée f_Y, en :
  - ⋆ dérivant F_Y sur les intervalles ouverts :
    - ◇ pour tout $x \in]-\infty; 0[$, $f_Y(x) = 0$;
    - ◇ pour tout $x \in]0; +\infty[$,
$$\begin{aligned} f_Y(x) &= F_Y'(x) \\ &= \frac{1}{2\sqrt{x}} F_X'(\sqrt{x}) - \frac{-1}{2\sqrt{x}} F_X'(-\sqrt{x}) \\ &= \frac{1}{2\sqrt{x}} f_X(\sqrt{x}) + \frac{1}{2\sqrt{x}} f_X(-\sqrt{x}) \\ &= \frac{1}{2\sqrt{x}}\big(f_X(\sqrt{x}) + f_X(-\sqrt{x})\big) \end{aligned}$$
  - ⋆ posant $f_Y(0) = 0$

Conclusion : X^2 est à densité et admet pour densité la fonction $f : x \longmapsto \begin{cases} \frac{1}{2\sqrt{x}}\big(f_X(\sqrt{x}) + f_X(-\sqrt{x})\big) & \text{si } x > 0 \\ 0 & \text{si } x \leqslant 0 \end{cases}$. ★

Petite remarque
Si $X(\Omega) = \mathbb{R}$, alors on a l'égalité $Y(\Omega) = \mathbb{R}^+$.

☞ Rappel...
$\forall t \in \mathbb{R},\ \forall a \in \mathbb{R}^+$:
$t^2 \leqslant a \iff -\sqrt{a} \leqslant t \leqslant \sqrt{a}$

Important !
Si on connaît F_X, on continue bien évidemment le calcul de $F_Y(x)$. Sinon, on peut poursuivre comme présenté ici.

☞ Rappel...
La continuité d'une fonction, c'est ce qui permet de "faire entrer la limite à l'intérieur".

♡ Astuce du chef ! ♡
Pour la continuité, on pourrait dire que F_Y est continue sur $[0; +\infty[$ et se dispenser alors du travail sur la limite à gauche en 0...
Mais l'avantage de travailler sur les intervalles ouverts est que l'on procède exactement de la même façon pour le caractère $\mathscr{C}^1$. Cela donne donc un vrai sens à la phrase "par des arguments similaires".

Petite remarque
On remplacerait alors f_X, qui serait connue, pour obtenir que f_Y est la densité d'une loi usuelle ou déjà étudiée dans l'exercice...

30 ♣ Expression de $\mathbb{E}(X)$ où X est à densité à valeurs dans $\mathbb{R}^+$

Soient $(\Omega, \mathcal{A}, \mathbb{P})$ un espace probabilisé et X une variable aléatoire à densité sur $(\Omega, \mathcal{A}, \mathbb{P})$.
Si : $X(\Omega) \subset \mathbb{R}^+$, X admet une densité f_X continue sur $\mathbb{R}^+$ et X admet une espérance ; alors : l'intégrale $\displaystyle\int_0^{+\infty} \mathbb{P}\big([X > t]\big)\,dt$ est convergente et $\mathbb{E}(X) = \displaystyle\int_0^{+\infty} \mathbb{P}\big([X > t]\big)\,dt$.

⋆

Démonstration : Supposons que $X(\Omega) \subset \mathbb{R}^+$, que X admet une densité f_X continue sur $\mathbb{R}^+$ et que X admet une espérance.
Notons F_X la fonction de répartition de X.
On a, pour tout $t \in \mathbb{R}^+$, $\mathbb{P}\big([X > t]\big) = 1 - F_X(t)$. Or F_X est la fonction de répartition d'une variable aléatoire à densité, elle est donc continue sur $\mathbb{R}^+$. Ainsi, l'intégrale $\displaystyle\int_0^{+\infty} \big(1 - F_X(t)\big)\,dt$ est impropre en $+\infty$ seulement.

- Soit $B \in \mathbb{R}^+$.
  Posons : $\left|\begin{array}{l} u : t \longmapsto 1 - F_X(t) \\ v : t \longmapsto t \end{array}\right.$. Les fonctions u et v sont $\mathscr{C}^1$ sur le segment $[0; B]$ (F_X est $\mathscr{C}^1$ sur $\mathbb{R}^+$, comme primitive de f_X, qui est continue sur $\mathbb{R}^+$) et pour tout $t \in [0; B]$: $\left|\begin{array}{l} u'(t) = -f_X(t) \\ v'(t) = 1 \end{array}\right.$.
  Par intégration par parties, on obtient :
  $$\begin{aligned}\int_0^B \big(1 - F_X(t)\big)\,dt &= \Big[t\big(1 - F_X(t)\big)\Big]_0^B - \int_0^B -t f_X(t)dt \\ &= B\big(1 - F_X(B)\big) + \int_0^B t f_X(t)dt\end{aligned}$$
- Ensuite :
  ⋆ on sait déjà que X admet une espérance, donc $\displaystyle\int_0^{+\infty} t f_X(t)dt$ converge et, puisque $X(\Omega) \subset \mathbb{R}^+$, on a :
  $$\mathbb{E}(X) = \int_0^{+\infty} t f_X(t)dt$$
  ⋆ Démontrons donc : $\displaystyle\lim_{B\to+\infty} B\big(1 - F_X(B)\big) = 0$.
  Soit $B \in \mathbb{R}^+$. On a :
  $$\begin{aligned}B\big(1 - F_X(B)\big) &= B\mathbb{P}\big([X > B]\big) \\ &= B\int_B^{+\infty} f_X(t)dt \\ &= \int_B^{+\infty} B f_X(t)dt\end{aligned}$$
  Pour commencer :
  $$\forall t \in [B; +\infty[, \;\; B \leqslant t$$
  D'où, puisque f_X est positive sur $\mathbb{R}$:
  $$\forall t \in [B; +\infty[, \;\; B f_X(t) \leqslant t f_X(t)$$
  Puis, par croissance de l'intégrale, licite car $B < +\infty$:
  $$\int_B^{+\infty} B f_X(t)dt \leqslant \int_B^{+\infty} t f_X(t)dt$$
  On a donc établi :
  $$0 \leqslant B\big(1 - F_X(B)\big) \leqslant \int_B^{+\infty} t f_X(t)dt$$
  Autrement dit, par relation de Chasles, licite car $\displaystyle\int_0^{+\infty} t f_X(t)dt$ est convergente et que $B \geqslant 0$:
  $$0 \leqslant B\big(1 - F_X(B)\big) \leqslant \int_0^{+\infty} t f_X(t)dt - \int_0^B t f_X(t)dt$$
  Par passage à la limite quand $B \to +\infty$ on obtient, par théorème d'encadrement :
  $$\lim_{B\to+\infty} B\big(1 - F_X(B)\big) = 0$$

Réflexe !
On veut démontrer qu'une quantité positive tend vers 0... on pense au théorème d'encadrement !
Et pour majorer une intégrale, on commence par majorer l'intégrande.

On obtient finalement :
$$\lim_{B\to+\infty} \int_0^B \big(1 - F_X(t)\big)\,dt = \int_0^{+\infty} t f_X(t)dt = \mathbb{E}(X)$$

Petite remarque
On met en parallèle cet exercice avec l'exercice 25 du travail estival...

Conclusion : l'intégrale $\displaystyle\int_0^{+\infty} \mathbb{P}\big([X > t]\big)\,dt$ est convergente et $\mathbb{E}(X) = \displaystyle\int_0^{+\infty} \mathbb{P}\big([X > t]\big)\,dt$.

⋆

Chapitre 10 – Diagonalisation des matrices carrées

31 Réduction de $U\,{}^tU$

Soit $U = \begin{pmatrix} a \\ b \\ c \end{pmatrix}$ une matrice non nulle de $\mathcal{M}_{3,1}(\mathbb{R})$.

La matrice $U\,{}^tU$ est diagonalisable, et semblable à la matrice $\begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & a^2+b^2+c^2 \end{pmatrix}$.

Important !
- La matrice $U\,{}^tU$ est une matrice de $\mathcal{M}_3(\mathbb{R})$.
- La matrice ${}^tUU$ est une matrice de $\mathcal{M}_1(\mathbb{R})$, que l'on assimile à un réel.

★ Démonstration : Notons $A = U\,{}^tU$.
- Remarquons que A est symétrique. En effet :
$$\begin{aligned} {}^tA &= {}^t\big(U\,{}^tU\big) \\ &= {}^t\big({}^tU\big)\,{}^tU \\ &= U\,{}^tU \\ &= A \end{aligned}$$
Conclusion : la matrice A est diagonalisable.
- On a :
$$A = \begin{pmatrix} a^2 & ab & ac \\ ab & b^2 & bc \\ ac & bc & c^2 \end{pmatrix}$$

Petite remarque On retrouve que A est symétrique...

D'où :
$$\begin{aligned} \mathrm{rg}(A) &= \mathrm{rg}\left(\begin{pmatrix} a^2 & ab & ac \\ ab & b^2 & bc \\ ac & bc & c^2 \end{pmatrix}\right) \\ &= \mathrm{rg}\left(\left(\begin{pmatrix} a^2 \\ ab \\ ac \end{pmatrix}, \begin{pmatrix} ab \\ b^2 \\ bc \end{pmatrix}, \begin{pmatrix} ac \\ bc \\ c^2 \end{pmatrix}\right)\right) \\ &= \mathrm{rg}\big((aU, bU, cU)\big) && \text{au moins un des } a, b, c \text{ est non nul} \\ &= \mathrm{rg}\big((U)\big) && U \text{ est non nul} \\ &= 1 \end{aligned}$$

✘ Attention ! Si $x = y = z = 0$, alors $\mathrm{rg}\big(xU, yU, zU\big) = 0 \neq \mathrm{rg}(U)$...

Par conséquent, 0 est valeur propre de A et, par théorème du rang :
$$\dim\big(\ker(A)\big) = 2$$
- Ensuite, remarquons que :
$$\begin{aligned} AU &= U\,{}^tUU \\ &= U(a^2+b^2+c^2) \end{aligned}$$
Puisque U est non nul, on en déduit que $a^2+b^2+c^2$ est valeur propre de A, et U en est un vecteur propre associé. Notons au passage que, puisque $U \neq 0_{3,1}$, on a $a^2+b^2+c^2 \neq 0$.

Petite remarque Je sais que vous n'auriez pas fait ainsi et que l'on peut calculer en utilisant l'expression explicite de A... Mais c'est plus élégant, et si je ne suis pas là pour vous montrer des choses élégantes parfois, à quoi suis-je utile ?! **On retient au passage :**
$${}^tUU = a^2+b^2+c^2$$

- D'après ce qui précède :
  - ⋆ 0 est valeur propre de A et l'espace propre associé à est dimension 2 ; notons une base (V_1, V_2) de cet espace propre (qui est $\ker(A)$) ;
  - ⋆ $a^2+b^2+c^2$ est valeur propre de A et l'espace propre associé est de dimension 1, engendré par U.

  La famille (V_1, V_2, U) est ainsi :
  - ⋆ libre car elle est la concaténation de familles libres de vecteurs propres associées à des valeurs propres différentes,
  - ⋆ de cardinal 3, égal à la dimension de $\mathcal{M}_{3,1}(\mathbb{R})$.

  La famille (V_1, V_2, U) est donc une base de vecteur propre de A.

  Conclusion : on retrouve ainsi que la matrice A est diagonalisable et semblable à la matrice $\begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & a^2+b^2+c^2 \end{pmatrix}$.

★

Et si on demande une matrice de passage ?

Pour cela, il faudrait déterminer des vecteurs V_1 et V_2 convenant...

Le travail sur le rang de A nous permet de constater que les vecteurs $\begin{pmatrix} b \\ -a \\ 0 \end{pmatrix}$, $\begin{pmatrix} c \\ 0 \\ -a \end{pmatrix}$ et $\begin{pmatrix} 0 \\ c \\ -b \end{pmatrix}$ appartiennent à $\ker(A)$.

Puis on en choisit deux de ces trois en guise de V_1 et V_2...

- Si $a \neq 0$: on prend $V_1 = \begin{pmatrix} c \\ 0 \\ -a \end{pmatrix}$ et $V_2 = \begin{pmatrix} b \\ -a \\ 0 \end{pmatrix}$
- Si $a = 0$ et $b \neq 0$: on prend $V_1 = \begin{pmatrix} 0 \\ c \\ -b \end{pmatrix}$ et $V_2 = \begin{pmatrix} b \\ -a \\ 0 \end{pmatrix}$

- Si $a = b = 0$ et $c \neq 0$: on prend $V_1 = \begin{pmatrix} c \\ 0 \\ -a \end{pmatrix}$ et $V_2 = \begin{pmatrix} 0 \\ c \\ -b \end{pmatrix}$

Dans tous les cas, la famille (V_1, V_2) est une famille de $\ker(A)$ qui est :

- ⋆ libre car constituée de deux vecteurs non colinéaires...
- ⋆ de cardinal 2, égal à la dimension de $\ker(A)$.

La famille (V_1, V_2) est donc une base de $\ker(A)$.

Conclusion : on choisit alors comme matrice P la matrice de passage de la base canonique vers la base (V_1, V_2, U) ainsi construite, de sorte que $A = PDP^{-1}$, où $D = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & a^2+b^2+c^2 \end{pmatrix}$.

Petite remarque

On prend les deux que l'on veut si les trois réels a, b, c sont non nuls !

32 ♣ Valeurs propres et vecteurs propres d'une matrice compagnon

Soient $a_0, a_1, a_2 \in \mathbb{R}$ ainsi que $C = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ -a_0 & -a_1 & -a_2 \end{pmatrix} \in \mathcal{M}_3(\mathbb{R})$.

Q1 Les valeurs propres de C sont les racines du polynôme $X^3 + a_2X^2 + a_1X + a_0$.

Q2 Pour tout $\lambda \in \mathrm{Sp}(C)$, l'espace propre $E_\lambda(C)$ est de dimension 1 et engendré par le vecteur $\begin{pmatrix} 1 \\ \lambda \\ \lambda^2 \end{pmatrix}$.

☞ Pour info...

On peut démontrer (même avec les outils d'ECG) que ce polynôme est le plus petit (au sens du degré) polynôme unitaire (coefficient dominant égal à 1) qui soit annulateur de C.

⋆ Démonstration :

Q1. Soit $\lambda \in \mathbb{R}$. On sait que :

$$\big(\lambda \text{ est valeur propre de } C\big) \iff \mathrm{rg}(C - \lambda I_3) < 3$$

Or :

$$\begin{aligned}
\mathrm{rg}(C-\lambda I_3) &= \mathrm{rg}\left(\begin{pmatrix} -\lambda & 1 & 0 \\ 0 & -\lambda & 1 \\ -a_0 & -a_1 & -a_2-\lambda \end{pmatrix}\right) \\
&\underset{C_2 \leftrightarrow C_1}{=} \mathrm{rg}\left(\begin{pmatrix} 1 & -\lambda & 0 \\ -\lambda & 0 & 1 \\ -a_1 & -a_0 & -a_2-\lambda \end{pmatrix}\right) \\
&\underset{\substack{L_2 \leftarrow L_2 + \lambda L_1 \\ L_3 \leftarrow L_3 + a_1 L_1}}{=} \mathrm{rg}\left(\begin{pmatrix} 1 & -\lambda & 0 \\ 0 & -\lambda^2 & 1 \\ 0 & -a_0-a_1\lambda & -a_2-\lambda \end{pmatrix}\right) \\
&\underset{C_3 \leftrightarrow C_2}{=} \mathrm{rg}\left(\begin{pmatrix} 1 & 0 & -\lambda \\ 0 & 1 & -\lambda^2 \\ 0 & -a_2-\lambda & -a_0-a_1\lambda \end{pmatrix}\right) \\
&\underset{L_3 \leftarrow L_3 + (a_2+\lambda)L_2}{=} \mathrm{rg}\left(\begin{pmatrix} 1 & 0 & -\lambda \\ 0 & 1 & -\lambda^2 \\ 0 & 0 & -a_0-a_1\lambda-a_2\lambda^2-\lambda^3 \end{pmatrix}\right)
\end{aligned}$$

♣ Méthode !

Les opérations élémentaires sur les lignes et colonnes conservent le rang.
On cherche donc à échelonner la matrice grâce à ces opérations pour déterminer une CNS pour que son rang ne soit pas maximal.
Attention : on veille toujours à ce que le pivot choisi soit non nul...

Mais, la matrice $\begin{pmatrix} 1 & 0 & -\lambda \\ 0 & 1 & -\lambda^2 \\ 0 & 0 & -a_0-a_1\lambda-a_2\lambda^2-\lambda^3 \end{pmatrix}$ est triangulaire supérieure ; elle est donc de rang maximal si, et seulement si, tous ses coefficients diagonaux sont non nuls.
Par conséquent :

$$\begin{aligned}
\mathrm{rg}(C-\lambda I_3) < 3 &\iff -a_0 - a_1\lambda - a_2\lambda^2 - \lambda^3 = 0 \\
&\iff \lambda^3 + a_2\lambda^2 + a_1\lambda + a_0 = 0
\end{aligned}$$

Conclusion : les valeurs propres de C sont les racines du polynôme $X^3 + a_2X^2 + a_1X + a_0$.

Q2. Soit $\lambda \in \mathrm{Sp}(C)$.

- Remarquons que, puisque les matrices $\begin{pmatrix} 1 \\ -\lambda \\ -a_1 \end{pmatrix}$ et $\begin{pmatrix} 0 \\ 1 \\ -a_2\lambda \end{pmatrix}$ ne sont pas colinéaires, la matrice $\begin{pmatrix} -\lambda & 1 & 0 \\ 0 & -\lambda & 1 \\ -a_0 & -a_1 & -a_2-\lambda \end{pmatrix}$ est au moins de rang 2.
- Mais λ est valeur propre de C, donc la matrice $\begin{pmatrix} -\lambda & 1 & 0 \\ 0 & -\lambda & 1 \\ -a_0 & -a_1 & -a_2\lambda \end{pmatrix}$ n'est pas de rang 3.

Petite remarque

Cette remarque pouvait également être faite à partir de sa réduction sous forme de matrice triangulaire à l'étape précédente...

On en dédit que $\mathrm{rg}(C - \lambda I_3) = 2$ et ainsi, par théorème du rang :

$$3 = rg(C - \lambda I_3) + \dim\big(\ker(C - \lambda I_3)\big)$$

Autrement dit :

$$\dim\big(\ker(C - \lambda I_3)\big) = 1$$

On remarque ensuite que :

$$\begin{pmatrix} -\lambda & 1 & 0 \\ 0 & -\lambda & 1 \\ -a_0 & -a_1 & -a_2-\lambda \end{pmatrix} \begin{pmatrix} 1 \\ \lambda \\ \lambda^2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ -a_0 - a_1\lambda - a_2\lambda^2 - \lambda^3 \end{pmatrix}$$

question précédente

$$= \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

Par conséquent, la famille $\left(\begin{pmatrix} 1 \\ \lambda \\ \lambda^2 \end{pmatrix}\right)$ est une famille de $\ker(C-\lambda I_3)$ qui est :

- ⋆ libre car constituée d'un unique vecteur non nul,
- ⋆ de cardinal 1, égal à la dimension de $\ker(C-\lambda I_3)$.

La famille $\left(\begin{pmatrix} 1 \\ \lambda \\ \lambda^2 \end{pmatrix}\right)$ est donc une base de $\ker(C-\lambda I_3)$.

Conclusion : pour tout $\lambda \in \mathrm{Sp}(C)$, l'espace propre $E_\lambda(C)$ est de dimension 1 et engendré par le vecteur $\begin{pmatrix} 1 \\ \lambda \\ \lambda^2 \end{pmatrix}$.

⋆

♣ Indication...
On cherche une matrice dans $\ker(C-\lambda I_3)$... On regarde les deux premières lignes de la matrice $C-\lambda I_3$, qui suggèrent la matrice $\begin{pmatrix} 1 \\ \lambda \\ \lambda^2 \end{pmatrix}$...

☞ Pour info...
Des deux points, on déduit que la matrice C est diagonalisable si, et seulement si, le polynôme $X^3 + a_2X^2 + a_1X + a_0$ admet trois racines distinctes...

33 ♥ MÉTHODE D'INVERSION POUR LA LOI EXPONENTIELLE

Soient $(\Omega, \mathcal{A}, \mathbb{P})$ un espace probabilisé, $\lambda \in \mathbb{R}_*^+$ et X une variable aléatoire sur $(\Omega, \mathcal{A}, \mathbb{P})$.
Si $X \hookrightarrow \mathscr{U}([0;1[)$, alors $\dfrac{-1}{\lambda}\ln(1-X) \hookrightarrow \mathscr{E}(\lambda)$.

DÉMONSTRATION : Supposons que $X \hookrightarrow \mathscr{U}([0;1[)$ et notons $Y = \dfrac{-1}{\lambda}\ln(1-X)$. Notons également $h : x \longmapsto \dfrac{-1}{\lambda}\ln(1-x)$, définie sur $[0;1[$; ainsi que F_Y la fonction de répartition de Y.

- On a :

$$\begin{aligned} Y(\Omega) &= \big(h(X)\big)(\Omega) \\ &= h\big(X(\Omega)\big) && X \hookrightarrow \mathscr{U}([0;1[) \\ &= h\big([0;1[\big) && h \text{ est continue et strictement croissante sur } [0;1[\\ &= [h(0); \lim_1 h[&& \text{car } \lambda > 0 \\ &= [0;+\infty[\end{aligned}$$

★Subtil...★
La continuité permet d'affirmer que $h([0;1[)$ est un intervalle.
En effet : "l'image d'un intervalle par une fonction continue est un intervalle" (version bis du TVI).

- Soit $x \in \mathbb{R}$. Distinguons deux cas :
  - ⋆ Si $x < 0$:
    On a :

$$\begin{aligned} F_Y(x) &= \mathbb{P}\big([Y \leqslant x]\big) \\ &= \mathbb{P}(\varnothing) && x<0 \text{ et } Y(\Omega) = [0;+\infty[\\ &= 0 \end{aligned}$$

  - ⋆ Si $x \geqslant 0$:
    On a :

$$\begin{aligned} F_Y(x) &= \mathbb{P}\big([Y \leqslant x]\big) \\ &= \mathbb{P}\left(\left[\frac{-1}{\lambda}\ln(1-X) \leqslant x\right]\right) && \lambda > 0 \\ &= \mathbb{P}\big([\ln(1-X) \geqslant -\lambda x]\big) && \text{stricte croissance de exp sur } \mathbb{R} \\ &= \mathbb{P}\big([1-X \geqslant \mathrm{e}^{-\lambda x}]\big) \\ &= \mathbb{P}\big([X \leqslant 1-\mathrm{e}^{-\lambda x}]\big) \\ &= F_X(1-\mathrm{e}^{-\lambda x}) \end{aligned}$$

Important !
L'argument de **stricte** croissance est indispensable ! En effet, cachée derrière cette égalité de probabilités, il y a une égalité d'ensembles :
$[\ln(1-X) \geqslant -\lambda x] = [X \leqslant 1-\mathrm{e}^{-\lambda x}]$.
Et on a, par stricte croissance de l'exponentielle (pour conserver l'équivalence), pour tout $\omega \in \Omega$:
$\ln\big(1-X(\omega)\big) \geqslant -\lambda x \Longleftrightarrow 1-X(\omega) \geqslant \mathrm{e}^{-\lambda x}$

Or : $x \geqslant 0$ et $\lambda > 0$, donc
$$-\lambda x \leqslant 0$$
D'où par croissance de exp sur $\mathbb{R}$:
$$\mathrm{e}^{-\lambda x} \leqslant 1$$
Ainsi :
$$1-\mathrm{e}^{-\lambda x} \geqslant 0$$
Et comme $\mathrm{e}^{-\lambda x} > 0$, on a $1-\mathrm{e}^{-\lambda x} < 1$. D'où :
$$0 \leqslant 1-\mathrm{e}^{-\lambda x} < 1$$
Or $X \hookrightarrow \mathscr{U}\big([0;1[\big)$, d'où :
$$F_X(1-\mathrm{e}^{-\lambda x}) = 1-\mathrm{e}^{-\lambda x}$$
Par conséquent :
$$F_Y(x) = 1-\mathrm{e}^{-\lambda x}$$

Petite remarque
On peut ne pas nécessairement détailler cette partie à l'écrit...

☞ Rappel...
La fonction de répartition d'une VA suivant la loi uniforme sur $[0;1]$ (ou $[0;1[$, ou $]0;1]$, ou $]0;1[$) est :
$$F : x \longmapsto \begin{cases} 0 & \text{si } x < 0 \\ x & \text{si } x \in [0;1] \\ 1 & \text{si } x > 1 \end{cases}$$

On a finalement obtenu :
$$\forall x \in \mathbb{R},\ F_Y(x) = \begin{cases} 0 & \text{si } x < 0 \\ 1-\mathrm{e}^{-\lambda x} & \text{si } x \geqslant 0 \end{cases}$$
On reconnaît ici la fonction de répartition d'une variable aléatoire suivant la loi exponentielle de paramètre λ. Or, la fonction de répartition caractérise la loi...
Conclusion : $Y \hookrightarrow \mathscr{E}(\lambda)$.

34 ♣ MOMENTS D'UNE VARIABLE ALÉATOIRE SUIVANT LA LOI $\mathscr{E}(1)$

Soient $(\Omega, \mathcal{A}, \mathbb{P})$ un espace probabilisé et X une variable aléatoire définie sur cet espace.
Si $X \hookrightarrow \mathscr{E}(1)$, alors pour tout $r \in \mathbb{N}$, X admet un moment d'ordre r et $\mathbb{E}(X^r) = r!$.

DÉMONSTRATION : Supposons que $X \hookrightarrow \mathscr{E}(1)$. Considérons $X(\Omega) = \mathbb{R}^+$ et notons $f_X : x \longmapsto \begin{cases} \mathrm{e}^{-x} & \text{si } x \geqslant 0 \\ 0 & \text{si } x < 0 \end{cases}$ une densité de X.

- **Existence.**
  Soit $r \in \mathbb{N}$.
  - ⋆ Par théorème de transfert, licite car la fonction $x \longmapsto x^r$ est continue sur $X(\Omega)$: :

X admet un moment d'ordre r si, et seulement si, l'intégrale $\displaystyle\int_{-\infty}^{+\infty} x^r f_X(x)dx$ est absolument convergente

si, et seulement si, l'intégrale $\displaystyle\int_{0}^{+\infty} x^r f_X(x)dx$ est convergente, car f_X est nulle sur $]-\infty;0[$ et l'intégrande est positive sur $[0;+\infty[$

si, et seulement si, l'intégrale $\displaystyle\int_{0}^{+\infty} x^r \mathrm{e}^{-x}dx$ est convergente.

⋆ Or, $x \longmapsto x^r\mathrm{e}^{-x}$ est continue sur $\mathbb{R}^+$, donc $\displaystyle\int_{0}^{+\infty} x^r \mathrm{e}^{-x}dx$ est impropre en $+\infty$ seulement. De surcroît :

◇ $x^r\mathrm{e}^{-x} \underset{x\to+\infty}{=} \mathrm{o}\left(\dfrac{1}{x^2}\right)$ (immédiat car $\displaystyle\lim_{x\to+\infty} x^{r+2}\mathrm{e}^{-x} = 0$ par croissance comparée) ;

◇ $\forall x \in \mathbb{R}^+,\ x^r\mathrm{e}^{-x} \geqslant 0,\ \dfrac{1}{x^2} \geqslant 0$;

◇ l'intégrale $\displaystyle\int_{1}^{+\infty} \frac{1}{x^2}dx$ est une intégrale de Riemann impropre en $+\infty$ d'exposant $2>1$, elle est donc convergente.

Ainsi, par critère de négligeabilité sur les intégrales à intégrandes positives, l'intégrale $\displaystyle\int_{1}^{+\infty} x^r \mathrm{e}^{-x}dx$ est convergente.

Puisque $\displaystyle\int_{0}^{1} x^r \mathrm{e}^{-x}dx$ n'est pas impropre (intégrande continue sur le segment d'intégration), on en déduit que l'intégrale $\displaystyle\int_{0}^{+\infty} x^r \mathrm{e}^{-x}dx$ est convergente.

Conclusion : X admet un moment d'ordre r.

- **Calcul.**

⋆ Démontrons : $\forall r \in \mathbb{N},\ \mathbb{E}(X^{r+1}) = (r+1)\mathbb{E}(X^r)$.

Soit $r \in \mathbb{N}^*$. Soit $B \in [0;+\infty[$.

Posons : $\left|\begin{array}{l} u : x \longmapsto x^{r+1} \\ v : x \longmapsto -\mathrm{e}^{-x}\end{array}\right.$. Les fonctions u et v sont $\mathscr{C}^1$ sur le segment $[0;B]$ et pour tout $x\in[0;B]$: $\left|\begin{array}{l} u'(x) = (r+1)x^r \\ v'(x) = \mathrm{e}^{-x}\end{array}\right.$.

Par intégration par parties, on obtient :

$$\begin{aligned}\int_0^B x^{r+1}\mathrm{e}^{-x}dx &= \left[-x^{r+1}\mathrm{e}^{-x}\right]_0^B - \int_0^B -(r+1)x^r\mathrm{e}^{-x}dx \\ &= -B^{r+1}\mathrm{e}^{-B} + (r+1)\int_0^B x^r\mathrm{e}^{-x}dx\end{aligned}$$

Or :

◇ puisque $r+1>0$, par croissance comparée, on a :

$$\lim_{B\to+\infty} B^{r+1}\mathrm{e}^{-B} = 0$$

◇ les intégrales $\displaystyle\int_0^{+\infty} x^{r+1}\mathrm{e}^{-x}dx$ et $\displaystyle\int_0^{+\infty} x^{r}\mathrm{e}^{-x}dx$ sont convergentes d'après le point précédent.

Donc, en passant à la imite quand $B\to+\infty$, on obtient :

$$\int_0^{+\infty} x^{r+1}\mathrm{e}^{-x}dx = (r+1)\int_0^{+\infty} x^{r}\mathrm{e}^{-x}dx$$

Autrement dit :

$$\mathbb{E}(X^{r+1}) = (r+1)\mathbb{E}(X^r)$$

⋆ On a ainsi :

$$\mathbb{E}(X^0) = \mathbb{E}(1) = 1 \quad ; \quad \forall r \in \mathbb{N},\ \mathbb{E}(X^{r+1}) = (r+1)\mathbb{E}(X^r)$$

D'où, par récurrence immédiate :

$$\forall r \in \mathbb{N},\ \mathbb{E}(X^r) = r!$$

Conclusion : pour tout $r\in\mathbb{N}$, X admet un moment d'ordre r et $\mathbb{E}(X^r) = r!$.

⋆

✓ Rigueur !

Si on avait $r+1 \leqslant 0$, il n'y aurait pas de FI... Pas de FI : pas de CC !

Petite remarque

Seule la convergence de $\displaystyle\int_0^{+\infty} x^r\mathrm{e}^{-x}dx$ est nécessaire... Cette convergence et le fait que $\displaystyle\lim_{B\to+\infty} B^{r+1}\mathrm{e}^{-B} = 0$ garantirait (si on ne le savait pas) la convergence de $\displaystyle\int_0^{+\infty} x^{r+1}\mathrm{e}^{-x}dx$ en obtenant une limite finie quand $B\to+\infty$.

Important !

Il est indispensable de mentionner que $\mathbb{E}(X^0) = 1$... Si on avait $\mathbb{E}(X^0) = 17$, on aurait $\forall r \in \mathbb{N},\ \mathbb{E}(X^r) = \dfrac{(r+17)!}{17!}$, qui est différent de $r!$.

35 Parties entière et fractionnaire d'une VA suivant une loi exponentielle

Soient $\lambda \in \mathbb{R}_*^+$, $(\Omega, \mathcal{A}, \mathbb{P})$ un espace probabilisé et X une variable aléatoire sur $(\Omega, \mathcal{A}, \mathbb{P})$.

Q1 Si $X \hookrightarrow \mathscr{E}(\lambda)$, alors $\lfloor X\rfloor + 1 \hookrightarrow \mathscr{G}\left(1-\mathrm{e}^{-\lambda}\right)$.

Q2 Si $X \hookrightarrow \mathscr{E}(\lambda)$, alors $X - \lfloor X\rfloor$ est à densité, de densité la fonction $f : x \longmapsto \begin{cases} \dfrac{\lambda\mathrm{e}^{-\lambda x}}{1-\mathrm{e}^{-\lambda}} & \text{si } x \in]0;1[\\ 0 & \text{sinon}\end{cases}$.

⋆

Démonstration :

Q1. Supposons que X suit la loi exponentielle de paramètre λ. Notons $Y = \lfloor X\rfloor$.

- Déterminons la loi de Y.

⋆ Puisque $X \hookrightarrow \mathscr{E}(\lambda)$, on considère $X(\Omega) = \mathbb{R}^+$. Ainsi : $Y(\Omega) = \mathbb{N}$.

⋆ Soit $n\in\mathbb{N}$. On a :

$$\begin{aligned}\mathbb{P}\left([Y=n]\right) &= \mathbb{P}\left([\lfloor X\rfloor = n]\right) \\ &= \mathbb{P}\left([n \leqslant X < n+1]\right) \\ &= F_X(n+1) - F_X(n) && X \text{ est à densité et on note } F_X \text{ sa fonction de répartition} \\ &= \left(1-\mathrm{e}^{-\lambda(n+1)}\right) - \left(1-\mathrm{e}^{-\lambda n}\right) && n, n+1 \geqslant 0 \\ &= \mathrm{e}^{-\lambda n} - \mathrm{e}^{-\lambda(n+1)} \\ &= \left(\mathrm{e}^{-\lambda}\right)^n\left(1-\mathrm{e}^{-\lambda}\right)\end{aligned}$$

☞ Rappels...

Deux rappels sur la partie entière :
- $\forall x\in\mathbb{R},\ \forall k\in\mathbb{Z}$, $\lfloor x\rfloor = k \iff k\leqslant x < k+1$
- $\forall x\in\mathbb{R},\ \lfloor x\rfloor \leqslant x < \lfloor x\rfloor + 1$

- Déduisons alors la loi de $Y+1$.
  - $\star$ Puisque $Y(\Omega) = \mathbb{N}$, on en déduit $(Y+1)(\Omega) = \mathbb{N}^*$.
  - $\star$ Soit $n \in \mathbb{N}^*$. On a :

$$\begin{aligned} \mathbb{P}\big([Y+1=n]\big) &= \mathbb{P}\big([Y=n-1]\big) \\ &= \big(\mathrm{e}^{-\lambda}\big)^{n-1}\big(1-\mathrm{e}^{-\lambda}\big) \end{aligned}$$

point précédent, licite car $n-1 \in \mathbb{N}$ (car $n-1 \in \mathbb{N}^*$)

  Par conséquent : $Y+1 \hookrightarrow \mathscr{G}\big(1-\mathrm{e}^{-\lambda}\big)$.

Conclusion : $\lfloor X \rfloor + 1 \hookrightarrow \mathscr{G}\big(1-\mathrm{e}^{-\lambda}\big)$.

Q2. Supposons que X suit la loi exponentielle de paramètre λ. Notons $Z = X - \lfloor X \rfloor$.

- On sait que :

$$\forall x \in \mathbb{R},\ \lfloor x \rfloor \leqslant x < \lfloor x \rfloor + 1$$

  Ainsi :

$$\forall \omega \in \Omega,\ 0 \leqslant Z(\omega) < 1$$

  D'où :

$$Z(\Omega) \subset [0;1[$$

- Notons F_Z la fonction de répartition de Z. Soit $x \in \mathbb{R}$.
  - $\star$ Si $x < 0$:

$$\begin{aligned} F_Z(x) &= \mathbb{P}\big([Z \leqslant x]\big) \\ &= \mathbb{P}(\varnothing) \\ &= 0 \end{aligned}$$

$Z(\Omega) \subset [0;1[$ et $x < 0$

  - $\star$ Si $x \geqslant 1$:

$$\begin{aligned} F_Z(x) &= \mathbb{P}\big([Z \leqslant x]\big) \\ &= \mathbb{P}(\Omega) \\ &= 1 \end{aligned}$$

$Z(\Omega) \subset [0;1[$ et $x \geqslant 1$

  - $\star$ Si $x \in [0;1[$:

$$\begin{aligned} F_Z(x) &= \mathbb{P}\big([Z \leqslant x]\big) \\ &= \mathbb{P}\big([X - Y \leqslant x]\big) \\ &= \sum_{n=0}^{+\infty} \mathbb{P}\big([Y=n] \cap [X-Y \leqslant x]\big) \\ &= \sum_{n=0}^{+\infty} \mathbb{P}\big([\lfloor X \rfloor = n] \cap [X \leqslant n+x]\big) \\ &= \sum_{n=0}^{+\infty} \mathbb{P}\big([n \leqslant X < n+1] \cap [X \leqslant n+x]\big) \\ &= \sum_{n=0}^{+\infty} \mathbb{P}\big([n \leqslant X \leqslant n+x]\big) \\ &= \sum_{n=0}^{+\infty} \big(F_X(n+x) - F_X(n)\big) \\ &= \sum_{n=0}^{+\infty} \Big(\big(1-\mathrm{e}^{-\lambda(n+x)}\big) - \big(1-\mathrm{e}^{-\lambda n}\big)\Big) \\ &= \sum_{n=0}^{+\infty} \big(\mathrm{e}^{-\lambda n} - \mathrm{e}^{-\lambda(n+x)}\big) \\ &= \big(1-\mathrm{e}^{-\lambda x}\big) \sum_{n=0}^{+\infty} \big(\mathrm{e}^{-\lambda}\big)^n \\ &= \big(1-\mathrm{e}^{-\lambda x}\big) \frac{1}{1-\mathrm{e}^{-\lambda}} \\ &= \frac{1-\mathrm{e}^{-\lambda x}}{1-\mathrm{e}^{-\lambda}} \end{aligned}$$

formule des probabilités totales, avec $\big([Y=n]\big)_{n\in\mathbb{N}}$ comme système complet d'évènements

$x < 1$, donc $[X \leqslant n+x] \subset [X < n+1]$

X est à densité

$n, n+x \geqslant 0$

Petite remarque

Il n'est pas nécessaire de mentionner l'argument $\mathrm{e}^{-\lambda} \in]-1;1[$ puisque l'on sait déjà, par la FPT, que la série en jeu est convergente...

  Par conséquent :

$$\forall x \in \mathbb{R},\ F_Z(x) = \begin{cases} 0 & \text{si } x < 0 \\ \dfrac{1-\mathrm{e}^{-\lambda x}}{1-\mathrm{e}^{-\lambda}} & \text{si } x \in [0;1[\\ 1 & \text{si } x \geqslant 1 \end{cases}$$

- Ensuite :
  - $\star$ La fonction F_Z est :
    - $\diamond$ continue sur $]-\infty;0[$ car constante sur cet intervalle,
    - $\diamond$ continue sur $]1;+\infty[$ car constante sur cet intervalle,
    - $\diamond$ continue en 0 car :

$$\lim_{\substack{x\to 0\\ x<0}} F_Z(x) = 0 = F_Z(0) = \lim_{\substack{x\to 0\\ x>0}} \frac{1-\mathrm{e}^{-\lambda x}}{1-\mathrm{e}^{-\lambda}}$$

      et continue en 1 car :

$$\lim_{\substack{x\to 1\\ x<1}} \frac{1-\mathrm{e}^{-\lambda x}}{1-\mathrm{e}^{-\lambda}} = 1 = \lim_{\substack{x\to 1\\ x>1}} F_Z(x)$$

Par conséquent, la fonction F_Z est continue sur $\mathbb{R}$.

- ⋆ Par des arguments similaires à la continuité, la fonction F_Z est de classe $\mathscr{C}^1$ sur $\mathbb{R}$ sauf éventuellement en 0 et en 1.

On en déduit que la variable aléatoire Z est à densité et on obtient une de ses densités, notée f_Z en :

- ⋆ dérivant F_Z sur les intervalles ouverts :
  - ◇ pour tout $x \in]-\infty; 0[$, $f_Z(x) = F_Z'(x) = 0$,
  - ◇ pour tout $x \in]0; 1[$, $f_Z(x) = F_Z'(x) = \dfrac{\lambda e^{-\lambda x}}{1 - e^{-\lambda}}$,
  - ◇ pour tout $x \in]1; +\infty[$, $f_Z(x) = F_Z'(x) = 0$.
- ⋆ posant $f_Z(0) = f_Z(1) = 0$.

Petite remarque

On peut également poser $f_Z(0) = \dfrac{\lambda}{1 - e^{-\lambda}}$ et $f_Z(1) = \dfrac{\lambda e^{-\lambda}}{1 - e^{-\lambda}}$...

Conclusion : Z est à densité et admet pour densité la fonction $f_Z : x \longmapsto \begin{cases} \dfrac{\lambda e^{-\lambda x}}{1 - e^{-\lambda}} & \text{si } x \in]0; 1[\\ 0 & \text{sinon} \end{cases}$.

⋆

36 ♥ Propriétés de la fonction de répartition associée à $\mathscr{N}(0;1)$

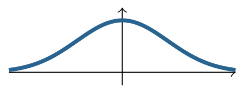
Soient $(\Omega, \mathcal{A}, \mathbb{P})$ un espace probabilisé et X une variable aléatoire sur $(\Omega, \mathcal{A}, \mathbb{P})$.
Supposons que $X \hookrightarrow \mathscr{N}(0;1)$ et notons Φ la fonction de répartition de X.

Q1 $\Phi(0) = \dfrac{1}{2}$

Q2 $\forall x \in \mathbb{R},\ \Phi(-x) = 1 - \Phi(x)$

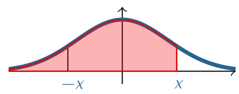
Preuve par le dessin !

Q1#



L'aire totale vaut 1, donc par symétrie (parité), l'aire sur $]-\infty; 0]$ vaut $\dfrac{1}{2}$.

Q2#



$\Phi(x)$ = aire en rouge. Et par symétrie, on remarque que $\Phi(x) + \Phi(-x) = 1$.

⋆ Démonstration : On rappelle que la fonction $\varphi : t \longmapsto \dfrac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}}$ est une densité de X. De façon immédiate, φ est paire.

Q1. On a :

$$\begin{aligned} \Phi(0) &= \mathbb{P}([X \leqslant 0]) && \text{φ est une densité de X} \\ &= \int_{-\infty}^{0} \varphi(t)dt && \text{φ est paire} \\ &= \frac{1}{2}\int_{-\infty}^{+\infty} \varphi(t)dt && \text{φ est une densité de probabilité} \\ &= \frac{1}{2} \end{aligned}$$

Q2. Soit $x \in \mathbb{R}$. On a : $\Phi(-x) = \displaystyle\int_{-\infty}^{-x} \varphi(t)dt$. Effectuons le changement de variable $u = -t$:

$$\left| \begin{array}{ll} u & = -t \\ t & = -u \end{array} \right. \quad ; \quad \left| \begin{array}{ll} du & = -dt \\ dt & = -du \end{array} \right. \quad ;$$

$t =$	$-\infty$	$-x$
$u =$	$+\infty$	x

Ce changement de variable est bien licite, puisque la fonction $u \longmapsto -u$ est $\mathscr{C}^1$ sur $[x; +\infty[$. On a ainsi :

$$\begin{aligned} \int_{-\infty}^{-x} \varphi(t)dt &= \int_{+\infty}^{x} \varphi(-u)(-du) \\ &= \int_{x}^{+\infty} \varphi(-u)du && \text{φ est paire} \\ &= \int_{x}^{+\infty} \varphi(u)du && \text{relation de Chasles} \\ &= \int_{-\infty}^{+\infty} \varphi(u)du - \int_{-\infty}^{x} \varphi(u)du && \text{φ est une densité de probabilité} \\ &= 1 - \Phi(x) \end{aligned}$$

★Subtil...★

Si on pose $u = g(t)$, g doit être bijective et c'est la fonction g^{-1} qui doit être $\mathscr{C}^1$...
Pour ne pas se tromper, on devrait toujours poser $t = \ldots$ même si c'est parfois moins naturel !

Important !

Les deux résultats ne reposent que sur le fait que φ est une densité paire !

⋆

37 ♣ Presque une intégrale d'une densité de loi normale...

Pour tout $(\mu, \sigma) \in \mathbb{R} \times \mathbb{R}_*^+$, l'intégrale $\displaystyle\int_0^{+\infty} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} dx$ est convergente et

$$\int_0^{+\infty} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} dx = \sqrt{2\pi\sigma^2}\,\Phi\left(\frac{\mu}{\sigma}\right)$$

où Φ désigne la fonction de répartition d'une variable aléatoire suivant la loi normale centrée réduite.

⋆ Démonstration : Soit $(\mu, \sigma) \in \mathbb{R} \times \mathbb{R}_*^+$.

♣ L'idée !

Se ramener à une densité d'une loi normale centrée réduite...

- On sait que la fonction $x \longmapsto \dfrac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$ est une densité d'une variable aléatoire suivant la loi $\mathcal{N}(\mu;\sigma^2)$.

  Par conséquent, l'intégrale $\displaystyle\int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} dx$ est convergente, et donc l'intégrale $\displaystyle\int_0^{+\infty} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} dx$ est également convergente.
- Proposons ensuite deux méthodes pour exprimer cette intégrale en fonction de Φ.
  1. **A l'aide d'un changement de variable.**
     Effectuons le changement de variable $t = \dfrac{x-\mu}{\sigma}$:

     $$\left|\begin{array}{ll} t & = \dfrac{x-\mu}{\sigma} \\ x & = \sigma t + \mu \end{array}\right. \quad ; \quad \left|\begin{array}{ll} dt & = \dfrac{1}{\sigma} dx \\ dx & = \sigma dt \end{array}\right. \quad ;$$

$x =$	0	$+\infty$
$t =$	$\dfrac{-\mu}{\sigma}$	$+\infty$ (car $\sigma > 0$)

     ☞ Rappel...
     Les changements de variable affines sont autorisés sur les intégrales impropres (pas les IPP), après avoir établi la convergence de l'intégrale initiale.

     Ce changement de variable est affine, donc licite sur cette intégrale impropre et :

     $$\begin{aligned}\int_0^{+\infty} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} dx &= \int_{-\frac{\mu}{\sigma}}^{+\infty} e^{-\frac{t^2}{2}} \sigma dt \\ &= \sigma\sqrt{2\pi} \int_{-\frac{\mu}{\sigma}}^{+\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt \\ &= \sqrt{2\pi\sigma^2} \left(\int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt - \int_{-\infty}^{-\frac{\mu}{\sigma}} \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt \right) \\ &= \sqrt{2\pi\sigma^2} \left(1 - \Phi\left(-\frac{\mu}{\sigma}\right)\right) \\ &= \sqrt{2\pi\sigma^2}\, \Phi\left(\frac{\mu}{\sigma}\right)\end{aligned}$$

     relation de Chasles, les intégrales en jeu étant bien convergentes

     on reconnaît l'intégrale d'une densité d'une variable aléatoire de loi $\mathcal{N}(0;1)$

     $\forall x \in \mathbb{R},\ 1 - \Phi(-x) = \Phi(x)$
  2. **A l'aide d'un résultat de cours.**
     Notons X une variable aléatoire, définie sur un certain espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$, suivant la loi $\mathcal{N}(\mu;\sigma^2)$ dont une densité est la fonction $x \longmapsto \dfrac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$.
     On a ainsi :

     $$\begin{aligned}\int_0^{+\infty} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} dx &= \sqrt{2\pi\sigma^2} \int_0^{+\infty} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} dx \\ &= \sqrt{2\pi\sigma^2}\, \mathbb{P}\big([X \geqslant 0]\big) \\ &= \sqrt{2\pi\sigma^2}\, \mathbb{P}\left(\left[\frac{X-\mu}{\sigma} \geqslant \frac{-\mu}{\sigma}\right]\right) \\ &= \sqrt{2\pi\sigma^2} \left(1 - \Phi\left(\frac{-\mu}{\sigma}\right)\right) \\ &= \sqrt{2\pi\sigma^2}\, \Phi\left(\frac{\mu}{\sigma}\right)\end{aligned}$$

     $\sigma > 0$

     $X \hookrightarrow \mathcal{N}(\mu;\sigma^2)$, donc $\dfrac{X-\mu}{\sigma} \hookrightarrow \mathcal{N}(0;1)$ et $\dfrac{X-\mu}{\sigma}$ est à densité

     $\forall x \in \mathbb{R},\ 1 - \Phi(-x) = \Phi(x)$

⋆

38 ♥ INÉGALITÉ DE MARKOV

Soient $(\Omega, \mathcal{A}, \mathbb{P})$ un espace probabilisé et X une variable aléatoire sur $(\Omega, \mathcal{A}, \mathbb{P})$.
Si X est **à valeurs positives** et **admet une espérance**, alors :

$$\forall a > 0,\ \mathbb{P}\big([X \geqslant a]\big) \leqslant \frac{\mathbb{E}(X)}{a}$$

Important ! Il faut savoir faire les trois !

DÉMONSTRATION : Proposons trois démonstrations...

1. **Si X est discrète.**
Soit $a \in \mathbb{R}_*^+$. On a :

$$\begin{aligned} a\mathbb{P}([X \geqslant a]) &= a \sum_{x \in X(\Omega),\ x \geqslant a} \mathbb{P}([X = x]) \\ &= \sum_{x \in X(\Omega),\ x \geqslant a} a\mathbb{P}([X = x]) \\ &\leqslant \sum_{x \in X(\Omega),\ x \geqslant a} x\mathbb{P}([X = x]) \\ &\leqslant \underbrace{\sum_{x \in X(\Omega)} x\mathbb{P}([X = x])}_{=\mathbb{E}(X)} \end{aligned}$$

$\forall x \geqslant a$, $a\mathbb{P}([X = x]) \leqslant x\mathbb{P}([X = x])$ (car une probabilité est positive)

X est à valeurs positives

D'où le résultat, puisque $a > 0$.

2. **Si X est à densité.**
Notons f une densité de X. Soit $a \in \mathbb{R}_*^+$. On a :

$$\begin{aligned} a\mathbb{P}([X \geqslant a]) &= a \int_a^{+\infty} f(x)dx \\ &= \int_a^{+\infty} af(x)dx \\ &\leqslant \int_a^{+\infty} xf(x)dx \\ &\leqslant \underbrace{\int_0^{+\infty} xf(x)dx}_{=\mathbb{E}(X),\ \text{car } X(\Omega) \subset \mathbb{R}^+} \end{aligned}$$

$\forall x \geqslant a,\ af(x) \leqslant xf(x)$ (car f positive) et par croissance de l'intégrale

$a > 0$ et pour tout $x \in \mathbb{R}^+,\ xf(x) \geqslant 0$

D'où le résultat, puisque $a > 0$.

3. **Cas général.**
Soit $a \in \mathbb{R}_*^+$. Introduisons la variable aléatoire Y définie sur $(\Omega, \mathcal{A}, \mathbb{P})$ par :

$$\forall \omega \in \Omega,\ Y(\omega) = \begin{cases} a & \text{si } X(\omega) \geqslant a \\ 0 & \text{sinon} \end{cases}$$

- $Y(\Omega) = \{0; a\}$ est un ensemble fini, donc Y possède une espérance et :

$$\begin{aligned} \mathbb{E}(Y) &= 0\mathbb{P}\big([Y = 0]\big) + a\mathbb{P}\big([Y = a]\big) \\ &= a\mathbb{P}\big([X \geqslant a]\big) \end{aligned}$$

$[Y = a] = [X \geqslant a]$

Important ! Il s'agit de démontrer :

$$\forall \omega \in \Omega,\ Y(\omega) \leqslant X(\omega)$$

- Montrons que $Y \leqslant X$. Soit $\omega \in \Omega$. Deux cas se présentent :
  - ⋆ si $\omega \in [X \geqslant a]$:
  Dans ce cas, $Y(\omega) = a$ et $X(\omega) \geqslant a$. D'où : $Y(\omega) \leqslant X(\omega)$.
  - ⋆ si $\omega \notin [X \geqslant a]$:
  Dans ce cas, $Y(\omega) = 0$. Mais X est à valeurs positives, donc $X(\omega) \geqslant 0$. D'où : $Y(\omega) \leqslant X(\omega)$.

  Conclusion : $\forall \omega \in \Omega,\ Y(\omega) \leqslant X(\omega)$.
- Puisque $Y \leqslant X$ et que Y et X possèdent une espérance, par croissance de l'espérance, on a :

$$\mathbb{E}(Y) \leqslant \mathbb{E}(X)$$

Autrement dit :

$$a\mathbb{P}\big([X \geqslant a]\big) \leqslant \mathbb{E}(X)$$

Et comme $a > 0$, on obtient :

$$\mathbb{P}\big([X \geqslant a]\big) \leqslant \frac{\mathbb{E}(X)}{a}$$

☞ **Rappel...**
Croissance de l'espérance : Si X et Y sont deux VA telles que :
- X et Y ont des espérances
- Y est presque-sûrement inférieure ou égale à X, autrement dit $\mathbb{P}\big([Y \leqslant X]\big) = 1$ (pas nécessaire que ce soit partout le cas)

Alors : $\mathbb{E}(Y) \leqslant \mathbb{E}(X)$.

⋆

39 ♥ LOI FAIBLE DES GRANDS NOMBRES

Soient $(\Omega, \mathcal{A}, \mathbb{P})$ un espace probabilisé et $(X_k)_{k \in \mathbb{N}^*}$ une suite de variables aléatoires sur $(\Omega, \mathcal{A}, \mathbb{P})$.
Pour tout $n \in \mathbb{N}^*$, on note $\overline{X_n} = \dfrac{1}{n} \displaystyle\sum_{k=1}^{n} X_k$.

Vocabulaire
$\overline{X_n}$ est la **moyenne empirique** de $X_1, X_2, \ldots, X_n$.

Si $(X_k)_{k\in\mathbb{N}^*}$ est une suite de variables aléatoires **indépendantes, admettant toutes la même espérance** m et **la même variance** σ^2, alors :

$$\forall \varepsilon > 0,\ \lim_{n\to+\infty} \mathbb{P}\left([|\overline{X_n} - m| \geqslant \varepsilon]\right) = 0$$

Vocabulaire

On dit que la suite $(\overline{X_n})_{n\in\mathbb{N}^*}$ converge en probabilité vers la variable aléatoire constante égale à m.

⋆

DÉMONSTRATION : Soit $\varepsilon > 0$.

- Soit $n \in \mathbb{N}^*$. La variable aléatoire $\overline{X_n}$ est une somme de variables aléatoires admettant une variance ; elle admet donc une variance et en particulier une espérance. Puis :

$$\begin{aligned}\mathbb{E}(\overline{X_n}) &= \mathbb{E}\left(\frac{1}{n}\sum_{k=1}^{n} X_k\right) \\ &= \frac{1}{n}\sum_{k=1}^{n}\mathbb{E}(X_k) \\ &= m\end{aligned}$$

linéarité de l'espérance

$\forall k \in \mathbb{N}^*,\ \mathbb{E}(X_k) = m$

et

$$\begin{aligned}\mathbb{V}(\overline{X_n}) &= \mathbb{V}\left(\frac{1}{n}\sum_{k=1}^{n} X_k\right) \\ &= \frac{1}{n^2}\sum_{k=1}^{n}\mathbb{V}(X_k) \\ &= \frac{1}{n}\sigma^2\end{aligned}$$

indépendance de $X_1, X_2, \ldots, X_n$

$\forall k \in \mathbb{N}^*,\ \mathbb{V}(X_k) = \sigma^2$

- Soit $n \in \mathbb{N}^*$. D'après l'inégalité de Bienaymé-Tchebychev, licite car $\overline{X_n}$ admet une variance, on a :

$$\mathbb{P}\left([|\overline{X_n} - \mathbb{E}(\overline{X_n})| \geqslant \varepsilon]\right) \leqslant \frac{\mathbb{V}(\overline{X_n})}{\varepsilon^2}$$

Autrement dit, d'après le point précédent :

$$\mathbb{P}\left([|\overline{X_n} - m| \geqslant \varepsilon]\right) \leqslant \frac{\sigma^2}{n\varepsilon^2}$$

- On a ainsi :

$$\forall n \in \mathbb{N}^*,\ 0 \leqslant \mathbb{P}\left([|\overline{X_n} - m| \geqslant \varepsilon]\right) \leqslant \frac{\sigma^2}{n\varepsilon^2}$$

Or :

$$\lim_{n\to+\infty} \frac{\sigma^2}{n\varepsilon^2} = 0$$

D'où le résultat par théorème d'encadrement.

Conclusion : $\forall \varepsilon > 0,\ \lim_{n\to+\infty} \mathbb{P}\left([|\overline{X_n} - m| \geqslant \varepsilon]\right) = 0.$

⋆

40 ♥ APPROXIMATION DE LA LOI BINOMIALE PAR LA LOI DE POISSON

Soient $(\Omega, \mathcal{A}, \mathbb{P})$ un espace probabilisé, $(X_n)_{n\in\mathbb{N}^*}$ une suite de variables aléatoires définies sur cet espace, $(p_n)_{n\in\mathbb{N}^*}$ une suite de réels de l'intervalle $]0;1[$ et $\lambda \in \mathbb{R}_*^+$.

Si : $\boldsymbol{\forall n \in \mathbb{N}^*,\ X_n \hookrightarrow \mathscr{B}(n; p_n)}$ et $\boldsymbol{\lim_{n\to+\infty} np_n = \lambda}$; alors : la suite $(X_n)_{n\in\mathbb{N}^*}$ converge en loi vers une variable aléatoire suivant la loi de Poisson de paramètre λ. Autrement dit :

$$\forall k \in \mathbb{N},\ \lim_{n\to+\infty}\left(\binom{n}{k} p_n^k (1-p_n)^{n-k}\right) = \mathrm{e}^{-\lambda}\frac{\lambda^k}{k!}$$

⋆

DÉMONSTRATION : Soit $k \in \mathbb{N}$. On a, pour $n \in [\![k; +\infty[\![$:

$$\begin{aligned}\binom{n}{k} p_n^k (1-p_n)^{n-k} &= \frac{n!}{k!(n-k)!} p_n^k (1-p_n)^{n-k} \\ &= \frac{n \times (n-1) \times \ldots \times (n-k+1)(n-k)!}{k!(n-k)!} p_n^k (1-p_n)^{n-k} \\ &= \frac{n \times (n-1) \times \ldots \times (n-k+1)}{k!} p_n^k (1-p_n)^{n-k} \\ &= \frac{n^k\left(\left(1-\frac{1}{n}\right)\left(1-\frac{2}{n}\right)\ldots\left(1-\frac{k-1}{n}\right)\right)}{k!} p_n^k (1-p_n)^{n-k} \\ &= \frac{(np_n)^k}{k!}\left(\left(1-\frac{1}{n}\right)\left(1-\frac{2}{n}\right)\ldots\left(1-\frac{k-1}{n}\right)\right)\frac{(1-p_n)^n}{(1-p_n)^k}\end{aligned}$$

$p_n \neq 1$

Petite remarque

n est destiné à tendre vers $+\infty$, donc il n'est pas dérangeant de le prendre supérieur ou égal à k.

Pourquoi ?

On factorise chaque facteur par n. Et de $n-k+1$ à n, il y a k facteurs (car $\mathrm{Card}\left([\![n-k+1, n]\!]\right) = n-(n-k+1)+1 = k$). On peut aussi dire que chaque facteur est équivalent à n... Mais ça ne dispense pas de les compter !

- Par opération :

$$\lim_{n\to+\infty} \frac{(np_n)^k}{k!} = \frac{\lambda^k}{k!}$$

- Par opération :

$$\lim_{n\to+\infty}\left(\left(1-\frac{1}{n}\right)\left(1-\frac{2}{n}\right)\ldots\left(1-\frac{k-1}{n}\right)\right) = 1$$

- De plus, puisque $\lim_{n\to+\infty} np_n = \lambda$, on a : $p_n \underset{n\to+\infty}{\sim} \dfrac{\lambda}{n}$. Ainsi $\lim_{n\to+\infty} p_n = 0$ et donc :
  - par opération :
$$\lim_{n\to+\infty} (1-p_n)^k = 1$$
  - et :
$$(1-p_n)^n = \exp\left(n\ln(1-p_n)\right)$$
RÉFLEXE !

    Puisque $\lim_{n\to+\infty} -p_n = 0$ on a (équivalent usuel) :
$$\ln(1-p_n) \underset{n\to+\infty}{\sim} -p_n$$
    D'où :
$$n\ln(1-p_n) \underset{n\to+\infty}{\sim} -np_n$$
    Par conséquent :
$$\lim_{n\to+\infty} n\ln(1-p_n) = -\lambda$$
    D'où, par composition de limites :
$$\lim_{n\to+\infty} \exp(n\ln(1-p_n)) = e^{-\lambda}$$

Par produit, on obtient finalement :
$$\lim_{n\to+\infty}\left(\binom{n}{k} p_n^k (1-p_n)^{n-k}\right) = e^{-\lambda}\frac{\lambda^k}{k!}$$

⋆

✘ Attention !
$\lim_{n\to+\infty}(1-p_n)^k$ s'obtient par opération (car k ne bouge pas...) alors que $(1-p_n)^n$ donne *a priori* une FI "1^∞".

Petite remarque
On ne peut pas composer les équivalents par une fonction. Donc on passe à la limite, puis on conclut pas composition de limites.

41 ♠ Une convergence en loi...

Soient $(\Omega, \mathcal{A}, \mathbb{P})$ un espace probabilisé et $(X_k)_{k\in\mathbb{N}^*}$ une suite de variables aléatoires indépendantes suivant toutes la loi uniforme sur $[0;1]$. Dans ce cas, la suite $\left(n\min(X_1,\dots,X_n)\right)_{n\in\mathbb{N}^*}$ converge en loi vers une variable aléatoire suivant la loi exponentielle de paramètre 1.

⋆

Démonstration :

- Soit $n\in\mathbb{N}^*$. Notons $Y_n = n\min(X_1,\dots,X_n)$ et F_{Y_n} la fonction de répartition de Y_n. On a, pour tout $x\in\mathbb{R}$:
$$\begin{aligned}
F_{Y_n}(x) &= \mathbb{P}([Y_n\leqslant x])\\
&= \mathbb{P}([n\min(X_1,\dots,X_n)\leqslant x])\\
&= \mathbb{P}\left(\left[\min(X_1,\dots,X_n)\leqslant \frac{x}{n}\right]\right) && n>0\\
&= 1-\mathbb{P}\left(\left[\min(X_1,\dots,X_n) > \frac{x}{n}\right]\right)\\
&= 1-\mathbb{P}\left(\bigcap_{k=1}^{n}\left[X_k>\frac{x}{n}\right]\right)\\
&= 1-\prod_{k=1}^{n}\mathbb{P}\left(\left[X_k>\frac{x}{n}\right]\right) && X_1,\dots,X_n \text{ sont indépendantes}\\
&= 1-\prod_{k=1}^{n}\left(1-\mathbb{P}\left(\left[X_k\leqslant\frac{x}{n}\right]\right)\right)\\
&= 1-\left(1-F\left(\frac{x}{n}\right)\right)^n && X_1,\dots,X_n \text{ ont même loi et on note } F \text{ leur fonction de répartition commune}\\
&= \begin{cases} 1-(1-0)^n & \text{si } \frac{x}{n}<0\\ 1-\left(1-\frac{x}{n}\right)^n & \text{si } \frac{x}{n}\in[0;1]\\ 1-(1-1)^n & \text{si } \frac{x}{n}>1\end{cases}\\
&= \begin{cases} 0 & \text{si } x<0\\ 1-\left(1-\frac{x}{n}\right)^n & \text{si } x\in[0;n]\\ 1 & \text{si } x>n\end{cases}
\end{aligned}$$

- Soit maintenant $x\in\mathbb{R}$.
  - Si $x<0$:
    On a : $\forall n\in\mathbb{N}^*,\ F_{Y_n}(x)=0$. D'où :
$$\lim_{n\to+\infty} F_{Y_n}(x) = 0$$
  - Si $x\geqslant 0$:
    Puisque $\lim_{n\to+\infty} n = +\infty$, il existe un rang n_0 (on peut prendre $n_0 = \lfloor x\rfloor + 1$ par exemple...), que nous considérons ensuite, tel que :
$$\forall n\in [\![n_0;+\infty[\![,\ n\geqslant x$$
    Ainsi :
$$\forall n\in [\![n_0;+\infty[\![,\ F_{Y_n}(x) = 1-\left(1-\frac{x}{n}\right)^n$$
    Or, pour tout $n\geqslant n_0$, $1-\dfrac{x}{n}>0$ (car $n_0>x$). Et on a donc :
$$\forall n\in [\![n_0;+\infty[\![,\ \left(1-\frac{x}{n}\right)^n = \exp\left(n\ln\left(1-\frac{x}{n}\right)\right)$$
RÉFLEXE !

    Mais :

Petite remarque
On peut choisir de commencer par réfléchir à l'ensemble image de Y_n... Sans mal, on trouve $Y_n(\Omega)\subset[0;n]$, ce qui nous guide ensuite pour la disjonction de cas. Je choisis ici une autre présentation, pour vous montrer comment on peut également procéder.

☞ Rappel...
La fonction de répartition d'une VA suivant la loi uniforme sur $[0;1]$ (ou $[0;1[$, ou $]0;1]$, ou $]0;1[$) est :
$$F : x\longmapsto\begin{cases}0 & \text{si } x<0\\ x & \text{si } x\in[0;1]\\ 1 & \text{si } x>1\end{cases}$$

♡ Astuce du chef ! ♡
Il faut être attentif quand on rédige sans procéder au préalable à la disjonction de cas sur x... C'est au moment de "remplacer" qu'on doit veiller à **remplacer bêtement** !!

Petite remarque
De façon générale, on quantifie x sur l'ensemble sur lequel la fonction de répartition de la loi limite est continue. Ici, comme il s'agit d'une loi exponentielle, on sait que la fonction de répartition associée est continue sur $\mathbb{R}$.

♣ Méthode !
- Les cas sont les mêmes que dans l'expression de la fonction de répartition de la loi limite...
- Si on ne connaît pas la loi limite, on peut "passer à la limite quand $n\to+\infty$ dans les intervalles des cas de F_{Y_n} pour voir"...

◇ $\lim\limits_{n\to+\infty} \dfrac{-x}{n} = 0$,

◇ $\ln(1+u) = u + \underset{u\to 0}{\mathrm{o}}(u)$

D'où :

$$\ln\left(1-\frac{x}{n}\right) = \frac{-x}{n} + \underset{n\to+\infty}{\mathrm{o}}\left(\frac{1}{n}\right)$$

Et ainsi :

$$n\ln\left(1-\frac{x}{n}\right) = -x + \underset{n\to+\infty}{\mathrm{o}}(1)$$

Par conséquent :

$$\lim_{n\to+\infty} n\ln\left(1-\frac{x}{n}\right) = -x$$

Par continuité de l'exponentielle en $-x$, on obtient :

$$\lim_{n\to+\infty} \left(1-\frac{x}{n}\right)^n = \mathrm{e}^{-x}$$

Et ainsi :

$$\lim_{n\to+\infty} F_{Y_n}(x) = 1-\mathrm{e}^{-x}$$

On a donc établi :

$$\forall x\in\mathbb{R},\ \lim_{n\to+\infty} F_{Y_n}(x) = \left\{\begin{array}{ll} 0 & \text{si } x<0 \\ 1-\mathrm{e}^{-x} & \text{si } x\geqslant 0 \end{array}\right.$$

Conclusion : la suite $(Y_n)_{n\in\mathbb{N}^*}$ converge en loi vers une variable aléatoire suivant la loi exponentielle de paramètre 1.

⋆

✓ Rigueur !

On utilise un DL d'ordre 1 plutôt qu'un équivalent, qui poserait souci dans le cas $x=0$.
Deux possibilités donc :
• contourner le souci en utilisant un DL, comme fait ici ;
• inclure le cas "$x=0$" dans le cas "$x<0$", puisque $F_{Y_n}(0)=0$...

☞ Pour info...

Dans le cas où les VA $X_1, \ldots, X_n$ suivent une loi uniforme sur $[0;1]$, les variables aléatoires $1-\max(X_1,\ldots,X_n)$ et $\min(X_1,\ldots,X_n)$ ont même loi. Par conséquent, on trouve la même convergence en loi pour la suite $\big(n\max(X_1,\ldots,X_n)\big)_{n\in\mathbb{N}^*}$, c'est également un classique.

Chapitre 15 – Équations différentielles et systèmes différentiels

42 ♠ Résolution de $y' + ay = 0$

Soient I un intervalle de $\mathbb{R}$, a une fonction continue sur I et f une fonction définie sur I.
Notons A une primitive de a sur I. On a :

$$\left(f \text{ est solution de l'EDL1 } y' + ay = 0\right) \iff \left(\exists \lambda \in \mathbb{R},\ \forall x \in I,\ f(x) = \lambda e^{-A(x)}\right)$$

Autrement dit : l'ensemble des solutions de $y' + ay = 0$, d'inconnue $y \in \mathscr{C}^1(I, \mathbb{R})$, est :

$$\left\{x \in I \mapsto \lambda e^{-A(x)} \ / \ \lambda \in \mathbb{R}\right\}$$

Petite remarque — Puisque a est continue sur l'intervalle I, elle admet des primitives sur I, qui sont de classe $\mathscr{C}^1$ sur I.

☞ Rappel... — Si F et G sont deux sous-ensembles d'un ensemble E, alors :
$E = F \iff \forall x \in E,\ (x \in F \iff x \in G)$

⋆ Démonstration : Il s'agit de montrer une équivalence. Raisonnons par double-implication :

$\boxed{\Longleftarrow}$ Supposons qu'il existe $\lambda \in \mathbb{R}$ tel que : $\forall x \in I,\ f(x) = \lambda e^{-A(x)}$.
Considérons un tel λ.

- ⋆ Puisque A est une primitive d'une fonction continue sur I, elle est de classe $\mathscr{C}^1$ sur I. Par conséquent, la fonction f est $\mathscr{C}^1$ sur I.
- ⋆ Pour tout $x \in I$:

$$\begin{aligned} f'(x) + af(x) &= -A'(x)\lambda e^{-A(x)} + a(x)\lambda e^{-A(x)} && \text{A est une primitive de } a \\ &= -a(x)\lambda e^{-A(x)} + a(x)\lambda e^{-A(x)} \\ &= 0 \end{aligned}$$

Ainsi f est solution de $y' + ay = 0$.

$\boxed{\Longrightarrow}$ Soit f une solution de $y' + ay = 0$. Montrons :

$$\exists \lambda \in \mathbb{R} \ / \ \forall x \in I,\ f(x) = \lambda e^{-A(x)}$$

Travaillons déjà par équivalences... On a :

$$\exists \lambda \in \mathbb{R} \ / \ \forall x \in I,\ f(x) = \lambda e^{-A(x)} \iff \exists \lambda \in \mathbb{R} \ / \ \forall x \in I,\ f(x)e^{A(x)} = \lambda$$

Posons alors la fonction $g : x \longmapsto f(x)e^{A(x)}$ et montrons qu'elle est constante sur I.
La fonction f est $\mathscr{C}^1$ sur I, car elle est solution de $y' + ay = 0$, donc la fonction g est un produit de deux fonctions dérivables (A est $\mathscr{C}^1$ sur I donc $\exp \circ A$ l'est également) sur I. Ainsi, g est dérivable sur I et, pour tout $x \in I$:

$$\begin{aligned} g'(x) &= f'(x)e^{A(x)} + f(x)A'(x)e^{A(x)} \\ &= e^{A(x)}\left(f'(x) + a(x)f(x)\right) \\ &= 0 && f \text{ est solution de } y' + ay = 0 \end{aligned}$$

Par conséquent, la fonction g, étant dérivable et de dérivée nulle sur un intervalle, est constante sur I.
Ainsi :

$$\exists \lambda \in \mathbb{R} \ / \ \forall x \in I,\ g(x) = \lambda$$

Autrement dit :

$$\exists \lambda \in \mathbb{R} \ / \ \forall x \in I,\ f(x) = \lambda e^{-A(x)}$$

Conclusion : $\left(f \text{ est solution de l'EDL1 } y' + ay = 0\right) \iff \left(\exists \lambda \in \mathbb{R},\ \forall x \in I,\ f(x) = \lambda e^{-A(x)}\right)$ ⋆

✓ Rigueur ! — Une solution de $y' + ay = 0$ est une fonction f qui vérifie :
- f est $\mathscr{C}^1$ sur I,
- $f' + af = 0$.

Question : — Une fonction dérivable et de dérivée nulle sur $\mathbb{R}^*$ est-elle constante sur $\mathbb{R}^*$?

Cas particulier — Dans le cas où a est constante, la fonction $x \longmapsto ax$ est une primitive de a, donc l'ensemble des solutions de $y' + ay = 0$ est $\left\{x \in I \mapsto \lambda e^{-ax} \ / \ \lambda \in \mathbb{R}\right\}$ (c'est le cas au programme).

43 ♥ Structure de l'ensemble des solutions d'une EDL

Soient (E) une équation différentielle linéaire et (E_H) son équation différentielle linéaire homogène associée.
Notons S_E l'ensemble des solutions de E, S_H l'ensemble des solutions de (E_H).

Q1 S_H est un espace vectoriel.

Q2 Toute solution de (E) est obtenue en ajoutant à une solution particulière de (E) une solution quelconque de (E_H). Autrement dit :

solution générale de l'EDL	=	solution particulière de l'EDL	+	solution générale de l'équation homogène associée

Ou encore, en notant f_p une solution particulière de E :

$$S_E = \left\{f_p + f_H \ / \ f_H \in S_H\right\}$$

⋆ Démonstration : Démontrons ce résultat dans le cas particulier d'une EDL1 normalisée à coefficient constant.
Soit $a \in \mathbb{R}$, I un intervalle de $\mathbb{R}$, $b \in \mathscr{C}(I, R)$ et :

$$(E) : y' + ay = b$$

d'inconnue $y \in \mathscr{C}^1(I, \mathbb{R})$. On a : $S_H = \{y \in \mathscr{C}^1(I, \mathbb{R}) \ / \ y' + ay = 0\}$.

Petite remarque — La démonstration est analogue dans le cas général...

Q1.

- $S_H \subset \mathscr{C}^1(I, \mathbb{R})$ et $\mathscr{C}^1(I, \mathbb{R})$ est un espace vectoriel.
- S_H est non vide car la fonction nulle vérifie l'équation $y' + ay = 0$.
- Montrons que S_H est stable par combinaison linéaire.
  Soient $f, g \in S_H$ et $\lambda, \mu \in \mathbb{R}$.
  - ⋆ Puisque $f, g \in S_H$, on a $f, g \in \mathscr{C}^1(I, \mathbb{R})$. Donc $\lambda f + \mu g \in \mathscr{C}^1(I, \mathbb{R})$, car $\mathscr{C}^1(I, \mathbb{R})$ est stable par combinaison linéaire.
  - ⋆ Ensuite, par linéarité de la dérivation :

$$\begin{aligned}(\lambda f + \mu g)' + a(\lambda f + \mu g) &= \lambda f' + \mu g' + \lambda a f + \mu a g \\ &= \lambda(f' + af) + \mu(g' + ag) \\ &= 0 \qquad f, g \in S_H\end{aligned}$$

  Par conséquent : $\lambda f + \mu g \in S_H$. Ainsi, S_H est stable par combinaison linéaire.

Important !

Être dans S_H c'est deux choses : être de classe $\mathscr{C}^1$ sur I et vérifier l'équation $y' + ay = 0$.

Conclusion : S_H est un espace vectoriel.

Q2. Soit $y \in \mathscr{C}^1(I, \mathbb{R})$ et supposons connue une solution particulière de (E), notée f_p. On a :

$$\begin{aligned}y \in S_E &\iff y' + ay = b \\ &\iff y' + ay = f_p' + af_p \qquad f_p \text{ solution de } (E)\text{, donc } f_p' + af_p = b \\ &\iff (y - f_p)' + a(y - f_p) = 0 \qquad \text{linéarité de la dérivation} \\ &\iff y - f_p \in S_H \\ &\iff \exists f_H \in S_H \ / \ y - f_p = f_H \\ &\iff \exists f_H \in S_H \ / \ y = f_p + f_H\end{aligned}$$

Conclusion : $S_E = \left\{f_p + f_H \ / \ f_H \in S_H\right\}$.

⋆