

Probabilités géométriques

Charles SUQUET

U.F.R. de Mathématiques Pures et Appliquées
Université des Sciences et Technologies de Lille

<http://math.univ-lille1.fr/~suquet/>

1 Introduction

Que peut bien raconter un probabiliste qui a imprudemment promis une conférence dans le cadre de l'exposition *Au delà du compas...? Géométrie et Théorie des Probabilités* peuvent sembler deux domaines assez éloignés. Pourtant, à y regarder de plus près, et malgré leur différence d'âge, il y a bien des rencontres entre ces deux branches des mathématiques. Ce n'est pas pour rien que l'on désigne parfois la théorie des probabilités comme la « géométrie du hasard ».

À travers ce texte, issu d'une conférence d'une heure, je souhaite montrer comment à partir de quelques problèmes classiques de probabilités portant sur des objets géométriques élémentaires, on peut se poser des questions touchant à la nature même de la modélisation probabiliste. Nous verrons successivement le problème de la *tige brisée*, de *l'aiguille de Buffon* et le *paradoxe de Bertrand*. Les énoncés des problèmes étudiés devraient être compréhensibles par des collégiens et j'espère qu'une partie des solutions proposées le seront aussi. Certains développements ont été écrits en pensant à leurs enseignants et aux étudiants de Licence de Mathématiques. J'ai pris le parti d'illustrer ce texte par un maximum de dessins, quitte à ce que certains paraissent superflus aux lecteurs avancés. Voici la liste des ingrédients utilisés.

- droites, segments ;
- triangles, carrés ;
- cercles ;
- hyperboles, sinusoides (juste une pincée, facultative) ;
- calculs d'aires ;
- formules (le moins possible) ;
- hasard à volonté, loi des grands nombres ;
- simulations informatiques.

2 La tige brisée

Regardons pour commencer un simple énoncé.

On découpe « au hasard » un segment de longueur 1 en trois morceaux. Quelle probabilité a-t-on de pouvoir former un triangle avec ces trois morceaux ?

Il y a un moyen simple d'expérimenter sur cette question. On prend un paquet de spaghetti, on met de l'eau à bouillir¹ et pour peu que l'on accepte de manger ses spaghetti coupés², on les brise l'un après l'autre en 3 morceaux et on regarde à chaque fois s'il est possible de former un triangle avec les trois morceaux. La moyenne du nombre de succès devrait être « proche » de la réponse théorique.

2.1 Un peu de géométrie

Avant de nous demander comment comprendre le « découpage au hasard » d'un segment en trois morceaux, voyons la question de géométrie élémentaire associée à cet énoncé : quelles conditions doivent vérifier les longueurs de trois segments pour pouvoir être les côtés d'un même triangle ? Plutôt que de donner une résolution formelle du problème, je vous propose une approche visuelle que j'espère compréhensible par des collégiens.



FIG. 1 – Découpage

Convenons d'appeler O et I les extrémités du segment et A , B les deux points de rupture, A étant plus proche de O que B , voir la figure 1.

Voici d'abord à la figure 2 un cas où la construction du triangle est possible. En faisant tourner les segments AO et BI respectivement autour de A et B , les extrémités O et I décrivent deux cercles qui se coupent en deux points. Chacun de ces points d'intersection peut servir de troisième sommet C au triangle que l'on cherche à construire.

-
1. Pour éviter le gaspillage.
 2. Il est des sacrifices à la science plus douloureux.

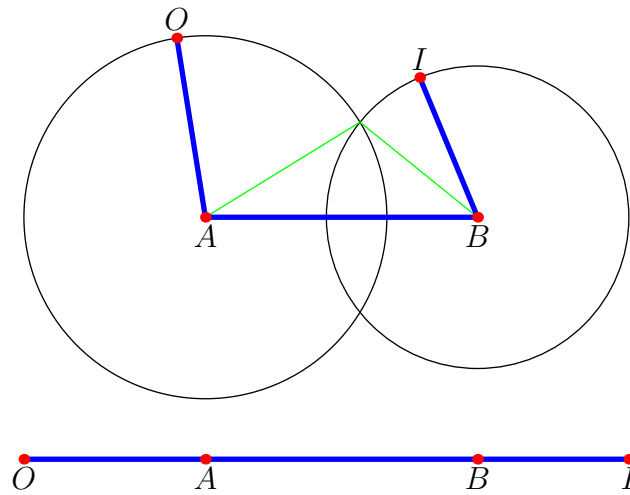


FIG. 2 – Construction du triangle : possible

La figure 3 montre un cas où cela ne marche pas : si la longueur AB est supérieure à la somme des « rayons » $OA + BI$, les deux cercles ne se coupent pas et il n'existe aucun point C vérifiant à la fois $AC = AO$ et $BC = BI$. La construction du triangle est alors impossible.

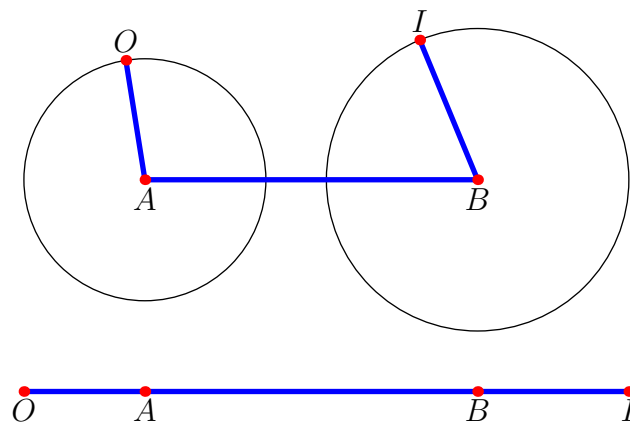


FIG. 3 – Construction du triangle : impossible

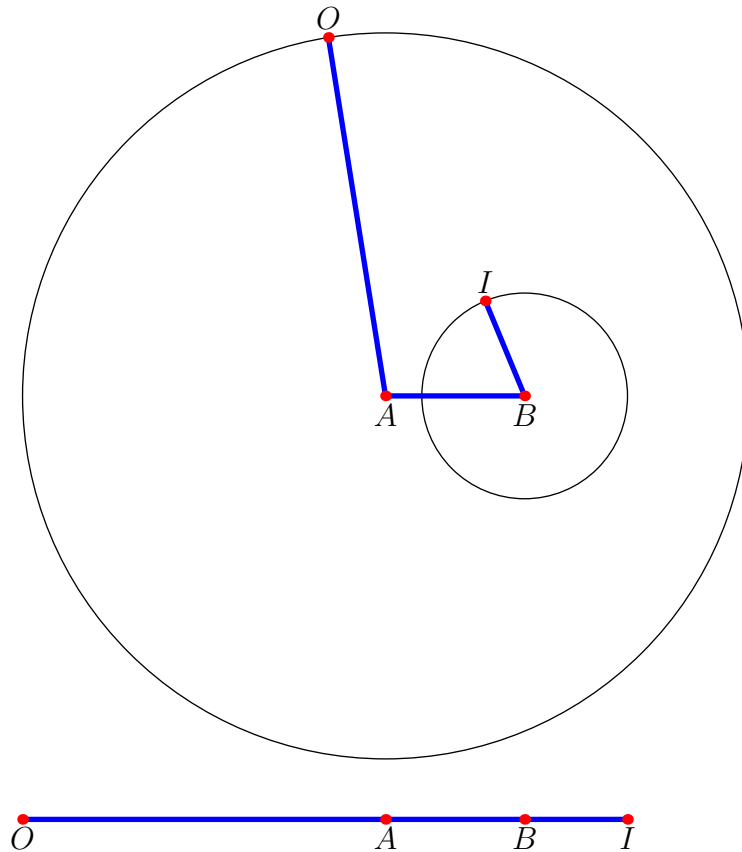


FIG. 4 – Construction du triangle : impossible

Dans le cas de la figure 3, la construction échoue parce que AB est « trop grand ». Elle peut aussi échouer si AB est « trop petit », comme on peut le voir sur la figure 4. Remarquons qu'ici, $AB + BI$ est inférieur à AO (faire tourner I dans le sens des aiguilles d'une montre sur le cercle de centre B jusqu'à ce qu'il soit aligné avec A et B).

Arrivé là, si j'affirme que l'on a vu les trois cas de figure possibles, on pourra m'objecter que j'ai oublié de regarder ce qui se passe en permutant les rôles des trois morceaux. Voici un argument qui devrait nous éviter de faire toutes les figures possibles. Appelons d la plus grande des 3 longueurs de segment OA , AB et BI notons r et r' les deux autres. La construction d'un triangle de côtés d , r et r' est possible si et seulement si les deux cercles de rayons r et r' et de centres distants de d ont au moins un point d'intersection³. Ceci équivaut à l'inégalité $d \leq r + r'$. Si cette inégalité est vérifiée, comme d est la plus grande des trois longueurs, on a automatiquement $r \leq d + r'$ et $r' \leq d + r$.

On pouvait d'ailleurs voir dès le début que ces trois conditions étaient nécessaires en raison de *l'inégalité triangulaire* : dans un triangle, la longueur de chaque côté est inférieure ou égale⁴ à la somme des deux autres.

3. Si vous n'êtes pas convaincu, revoyez les figures 2 et 3 en changeant si nécessaire les noms des points de façon à remplacer le segment AB par le plus long des trois segments.

4. Le cas d'égalité est celui du triangle « aplati », où les trois sommets sont alignés. Dans ce cas les cercles de rayon r et r' sont tangents.

Pour conclure cette discussion, la construction d'un triangle avec les trois morceaux de la tige est possible si et seulement si aucun des trois n'a une longueur strictement supérieure à la somme des deux autres. On peut simplifier cette condition en se rappelant que la somme des trois longueurs vaut 1. La condition équivaut alors à : *chacune des trois longueurs est inférieure ou égale à 1/2*. Si cette affirmation ne vous paraît pas évidente, on peut la vérifier comme suit. Notons ℓ , ℓ' et ℓ'' les trois longueurs. Si $\ell > \ell' + \ell''$, alors comme $\ell' + \ell'' = 1 - \ell$, on en déduit que $\ell > 1 - \ell$, d'où $2\ell > 1$ soit $\ell > 1/2$. D'autre part si $\ell \leq 1/2$, alors $\ell' + \ell'' = 1 - \ell \geq 1 - 1/2 = 1/2 \geq \ell$, d'où $\ell \leq \ell' + \ell''$. On a donc équivalence entre les inégalités $\ell \leq \ell' + \ell''$ et $\ell \leq 1/2$. Il est clair que l'on obtient deux équivalences du même type en permutant les rôles de ℓ , ℓ' , ℓ'' .

Nous pouvons maintenant reformuler le problème de la tige brisée comme suit.

On découpe « au hasard » un segment de longueur 1 en trois morceaux. Quelle est la probabilité que chacun des trois morceaux ait une longueur d'au plus 1/2 ?

2.2 Modélisation

Il est temps de préciser ce que l'on entend par « découpage au hasard ». En réalité il y a une multitude d'interprétations possibles. Nous allons en proposer une⁵ qui permet une résolution relativement élémentaire du problème et qui répond aux deux exigences suivantes.

1. La méthode de choix d'un point de rupture ne doit privilégier *a priori* aucun point du segment.
2. Le choix d'un point de rupture ne doit avoir aucune influence sur le choix de l'autre point de rupture.

Pour réaliser « l'indépendance des choix de points de rupture » décrite par la condition 2, on a recours à deux expérimentateurs disposant chacun d'un feuille de papier sur laquelle est dessiné un segment OI de même longueur que la tige⁶ à briser. On demande à chacun d'eux de marquer avec un crayon bien taillé un point du segment, sans regarder ce que fait l'autre. On rassemble ensuite les résultats et on découpe la tige aux deux points indiqués.

On peut maintenant remarquer que le choix de deux points de rupture au hasard sur le segment OI équivaut au *choix de deux nombres au hasard* X et Y entre 0 et 1. En notant respectivement $\min(X, Y)$ et $\max(X, Y)$ le plus petit et le plus grand de ces deux nombres, les longueurs des trois morceaux de tige sont $OA = \min(X, Y)$, $AB = \max(X, Y) - \min(X, Y) = |X - Y|$ et $BI = 1 - \max(X, Y)$, voir la figure 5.

5. D'autres procédures seront discutées dans la sous-section 2.6.

6. On l'aura compris, l'énoncé parle d'un « segment de longueur 1 », uniquement pour des raisons de confort. D'ailleurs, comme on ne précise pas l'unité de longueur, il suffit de décréter que l'unité de longueur est précisément la longueur de la tige à briser, par exemple celle d'un spaghetti sec.

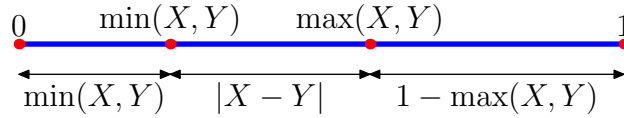


FIG. 5 – Modélisation des points de rupture par deux nombres au hasard X et Y

Ce choix équivaut à celui d'un point au hasard de coordonnées (X, Y) dans le carré unité, voir la figure 6.

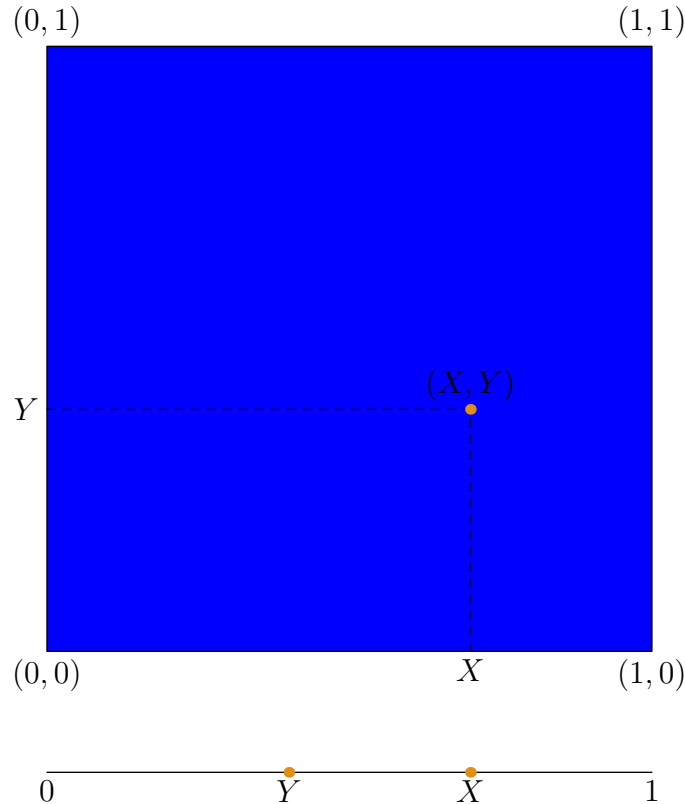


FIG. 6 – Le choix de X et Y dans $[0, 1]$ équivaut au choix du point (X, Y) du carré unité

Voyons maintenant la signification de la condition 1 pour le choix des nombres au hasard X et Y . S'il y avait en tout n nombres dans le segment $[0, 1]$, il nous suffirait de dire : « pour ne privilégier aucun de ces nombres, attribuons à chacun la probabilité $1/n$ d'être choisi ». On pourrait même faire ce choix avec une urne contenant n boules numérotées de 1 à n et en tirant une boule dont le numéro nous indiquerait le nombre à choisir. Hélas, dans un segment il y a une infinité de points et cette première approche naïve ne fonctionne pas. Sans se décourager, essayons en une autre. Découpons par une graduation régulière le segment $[0, 1]$ en 10 parts égales et prenons une urne contenant 10 boules numérotées de 0 à 9. On extrait une boule et selon son numéro, disons i , on choisit le segment $[\frac{i}{10}, \frac{i+1}{10}[$ si $i < 9$ et $[\frac{9}{10}, 1]$ si $i = 9$. Avec ce procédé, chacun des

segments $[0, \frac{1}{10}[, \dots, [\frac{9}{10}, 1]$ a une chance sur 10 d'être choisi, en raison de la composition de l'urne. Remarquons au passage que cette probabilité de choix est égale pour chacun de ces segments à sa longueur⁷. Quel rapport avec le choix de X ? Et bien décidons que X devra être dans le segment ainsi choisi. Par exemple si j'ai tiré la boule numéro 3, je peux dire à ce stade que X est dans $[0,3; 0,4[$, ou encore que $0,3 \leq X < 0,4$. Évidemment, je ne vais pas m'en satisfaire et je vais recommencer avec une graduation dix fois plus fine, découpant le segment initial en 100 parties égales. Plutôt que de prendre une urne avec 100 boules numérotées de 0 à 99, je vais faire deux tirages avec remise, le premier tirage indiquant le chiffre des dizaines et le deuxième celui des unités⁸. Par exemple si le premier tirage amène la boule n° 3 et le deuxième la boule n° 7, je choisis le segment $[0,37; 0,38[$. À l'issue de ces deux tirages, je peux dire que X vérifie $0,37 \leq X < 0,38$. Si j'avais parié *avant* la réalisation des tirages que X vérifie cet encadrement, ma probabilité de gagner aurait été de $1/100$. On peut continuer ainsi en améliorant la précision d'un facteur 10 à chaque étape. En effectuant n tirages avec remise, on arrive à encadrer X entre des bornes distantes de $10^{-n} = 0,00\dots 01$ (n chiffres après la virgule). Par exemple si l'on veut découper un spaghetti de 25 cm, avec 6 tirages, on arrive à une précision de 0,25 millième de millimètre. Pour arriver à effectuer le découpage avec une telle précision, il faudrait découper le spaghetti au laser (quel gâchis!).

Me voilà donc muni d'une méthode *pratique* de choix de X . Après tout, si je choisissais un point au hasard dans le segment OI en le marquant au crayon, la précision dans la détermination de ce point ne pourrait être meilleure que l'épaisseur de la pointe du crayon. On m'objectera avec raison que les points de la géométrie n'ont pas d'épaisseur et que même s'il est impossible physiquement de tracer un point dénué d'épaisseur, cela n'a pas empêché les géomètres de raisonner sur des objets idéaux : points, droites, cercles, ... dont les tracés physiques ne sont que des approximations. On devrait donc pouvoir raisonner sur le nombre au hasard X indépendamment de la précision choisie pour le représenter physiquement, de même qu'on peut raisonner sur le milieu d'un segment que l'on est incapable de tracer avec une précision parfaite. Avec le procédé adopté ci-dessus, il suffit de faire — par la pensée! — une infinité de tirages avec remise. Regardons dans ces conditions quelle peut être la probabilité que X vaille une valeur donnée à l'avance, par exemple $1/3$. Pour obtenir $X = 1/3$ au bout d'une infinité de tirages, il faut avoir les encadrements successifs : $0,3 \leq X < 0,4$, puis $0,33 \leq X < 0,34$, ... et au n^e tirage :

$$0, \underbrace{333\dots 3333}_{n \text{ chiffres}} \leq X < 0, \underbrace{333\dots 333}_{n-1 \text{ chiffres}} 4.$$

La probabilité que cet encadrement se réalise est celle de sortir la boule n° 3 lors de chacun des n premiers tirages dans l'urne, c'est donc 10^{-n} . L'évènement $\{X = 1/3\}$ a clairement une probabilité inférieure ou égale à celle de l'encadrement ci-dessus. Donc

7. Parce que le segment initial a pour longueur 1. S'il avait pour longueur ℓ , cette probabilité $1/10$ serait égale au quotient de la longueur du segment à choisir par celle du segment initial soit $(\ell/10)/\ell = 1/10$.

8. Les deux procédés, tirage dans une urne contenant 100 boules ou deux tirages avec remise dans une urne contenant 10 boules, donnent la même probabilité $1/100$ de choix d'un nombre entier entre 0 et 99. La vérification de cette affirmation est laissée en exercice.

on a pour tout n entier, $P(X = 1/3) \leq 10^{-n}$. Comme on est libre de choisir n aussi grand que l'on veut, donc 10^{-n} aussi petit que l'on veut, cela ne laisse pas d'autre choix à $P(X = 1/3)$ que d'être nul ! On pourrait voir de même que pour n'importe quel nombre a de $[0, 1]$ fixé à l'avance, $P(X = a) = 0$. Ce résultat peut sembler décevant. En fait il montre simplement que les quantités $P(X = a)$ ne sont pas pertinentes pour traiter le problème qui nous intéresse. Au lieu de la probabilité que X prenne une valeur donnée, intéressons nous à celle qu'il appartienne à un *intervalle* donné.

La probabilité que X appartienne à un intervalle $[a_n, b_n[$ dont les bornes sont des nombres décimaux de $[0, 1]$ ayant n chiffres après la virgule, ne fait intervenir que les n premiers tirages et il est facile de voir comme ci-dessus que c'est exactement la longueur $b_n - a_n$ de l'intervalle. Par exemple $P(0,347 \leq X < 0,521)$ est la probabilité que dans une urne contenant 1000 boules⁹, on sorte une boule de numéro supérieur ou égal à 347 et inférieur strictement à 521. Il y a pour cela $521 - 347 = 174$ possibilités sur 1000, d'où une probabilité de $0,174 = 0,521 - 0,347$.

Que peut-on dire si les bornes a et b de l'intervalle ne sont pas des nombres décimaux ? Un exemple simple de cette situation est l'intervalle $[\frac{1}{3}, \frac{2}{3}[$. En effet ces deux bornes n'ont que des développements décimaux illimités avec une infinité de chiffres non nuls : $\frac{1}{3} = 0,333\ 333\ 333\ \dots$, $\frac{2}{3} = 0,666\ 666\ 666\ \dots$. On ne peut donc pas régler la question en un nombre fini de tirages. Notons alors pour tout entier $n \geq 1$, a_n la meilleure approximation par défaut de $1/3$ par un décimal de la forme $k/10^n$, k entier. De même notons a'_n la meilleure approximation par excès, b_n la meilleure approximation par excès de $2/3$ et b'_n la meilleure approximation par défaut. La table 1 donne les 4 premiers termes de ces suites.

n	a_n	a'_n	b'_n	b_n
1	0,3	0,4	0,6	0,7
2	0,33	0,34	0,66	0,67
3	0,333	0,334	0,666	0,667
4	0,333 3	0,333 4	0,666 6	0,666 7

TAB. 1 – Suites des valeurs approchées décimales de $1/3$ et $2/3$

Pour chaque n , l'intervalle $[a_n, b_n[$ contient l'intervalle $[1/3, 2/3[$ qui lui même contient $[a'_n, b'_n[$. On en déduit que

$$P(a'_n \leq X < b'_n) \leq P\left(\frac{1}{3} \leq X < \frac{2}{3}\right) \leq P(a_n \leq X < b_n),$$

d'où

$$b'_n - a'_n = 0, \underbrace{333 \dots 333}_n 2 \leq P\left(\frac{1}{3} \leq X < \frac{2}{3}\right) \leq b_n - a_n = 0, \underbrace{333 \dots 333}_n 4.$$

9. En utilisant encore l'équivalence entre 3 tirages avec remise parmi $0, \dots, 9$ ou un seul tirage parmi $0, \dots, 999$.

En faisant tendre n vers l'infini, on voit alors que $b'_n - a'_n$ converge en croissant vers $1/3$ et $b_n - a_n$ converge en décroissant vers $1/3$. La probabilité $P\left(\frac{1}{3} \leq X < \frac{2}{3}\right)$ qui elle *ne dépend pas de n* , vaut donc obligatoirement $1/3$, c'est bien la longueur de l'intervalle $\left[\frac{1}{3}, \frac{2}{3}\right[$. On peut généraliser ce résultat avec n'importe quel intervalle $[a, b[$ de $[0, 1]$.

Finalement nous conviendrons que le choix du nombre au hasard X dans $[0, 1]$ vérifie la condition

$$P(X \in [a, b]) = P(a \leq X < b) = b - a,$$

pour tout intervalle $[a, b[$ inclus dans $[0, 1]$. On dit que la variable aléatoire X suit la *loi uniforme* sur $[0, 1]$. Notons d'ailleurs que pour tout sous-intervalle de $[0, 1]$, qu'il contienne ou non ses bornes, la probabilité que X tombe dans cet intervalle est égale à la longueur de cet intervalle. Vérifions le par exemple pour un intervalle fermé $[a, b]$. Comme $P(X = b) = 0$, on a $P(X \in [a, b]) = P(X \in [a, b]) + P(X = b) = P(X \in [a, b]) = b - a$. Bien entendu, le choix au hasard du nombre Y vérifiera les mêmes propriétés.

Comme nous l'avons vu plus haut, le choix au hasard « indépendamment l'un de l'autre » de deux nombres X et Y dans $[0, 1]$ équivaut, cf. figure 6, au choix au hasard d'un point de coordonnées (X, Y) dans un carré de côté 1. Nous *admettrons* que le choix de ce point (X, Y) dans le carré vérifie la propriété suivante : pour toute région R convenable du carré unité¹⁰

$$P((X, Y) \in R) = \text{aire}(R).$$

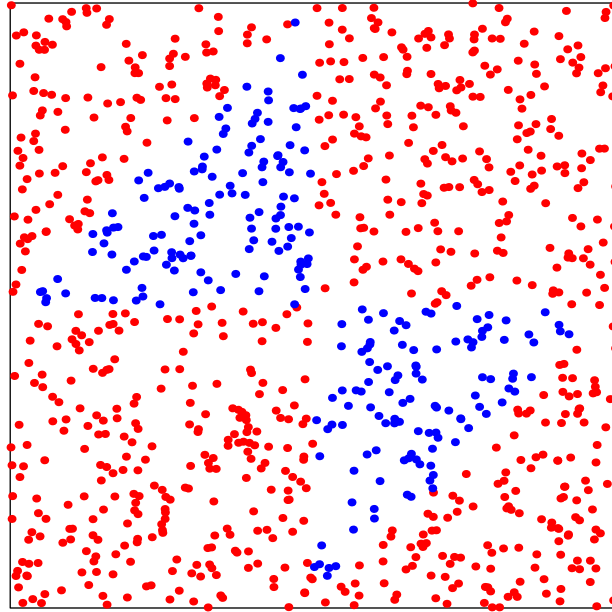
On dit que le vecteur aléatoire (X, Y) suit la loi uniforme sur le carré unité. Nous ne détaillerons pas trop ce que signifie « région convenable ». Pour notre usage, il suffit que cela marche pour des régions qui s'obtiennent en assemblant un nombre fini de triangles. On pourrait d'ailleurs reprendre l'étude faite ci-dessus avec les segments, en remplaçant les graduations par des quadrillages de plus en plus fins pour justifier cette formule dans le cas où R est un triangle.

2.3 Solution graphique

La figure 7 représente le résultat de la simulation du découpage en trois de 1 000 tiges. On a exploité pour cela le générateur de nombres aléatoires du logiciel METAPOST utilisé pour produire la plupart des figures qui illustrent ce texte. Plus précisément, on a écrit un petit programme qui génère 1 000 couples de nombres au hasard (X, Y) et qui teste pour chacun de ces couples si l'une des quantités $\min(X, Y)$, $|X - Y|$ et $1 - \max(X, Y)$ dépasse $1/2$. Si la réponse est oui, la construction du triangle avec les trois morceaux de tige est impossible et on trace le point (X, Y) en rouge. Sinon, la construction du triangle est possible et on trace le point (X, Y) en bleu. Le nombre total de points bleus est affiché à droite du triangle¹¹.

10. Pour la loi uniforme sur un carré quelconque C , la probabilité d'appartenance de (X, Y) au sous-ensemble R de C serait, $\text{aire}(R)/\text{aire}(C)$. Rappelons qu'une probabilité n'a pas d'unité et que lorsque j'écris $P((X, Y) \in R) = \text{aire}(R)$, c'est un abus de langage. Par exemple si le carré unité a pour aire 1 m^2 et si $\text{aire}(R) = 0,37 \text{ m}^2$, il faut comprendre que $P((X, Y) \in R) = 0,37$.

11. Comme j'avais envie de m'amuser un peu, j'ai inclus le code METAPOST du programme de génération de la figure 7 *directement* dans le source L^AT_EX 2_ε de ce programme. Concrètement cela implique



243

FIG. 7 – Simulation 1 000 tiges brisées, en bleu les points (X, Y) de succès

Cette simulation nous donne une idée de ce que l'on va trouver, mais nous avons maintenant un outillage de géométrie élémentaire pour résoudre notre problème. Appelons R l'ensemble des points de coordonnées (x, y) du carré C qui représentent un découpage « réussi » de la tige, c'est-à-dire permettant la construction d'un triangle avec les trois morceaux. D'après l'étude de la sous-section 2.1,

$$R = \left\{ (x, y) \in C; \min(x, y) \leq \frac{1}{2}, |x - y| \leq \frac{1}{2}, 1 - \max(x, y) \leq \frac{1}{2} \right\}.$$

Tout ce qu'il nous reste à faire est donc de représenter R en éliminant tous les points du carré qui ne vérifient pas l'un de ces 3 conditions et de calculer l'aire de R . L'élimination des points ne vérifiant pas l'une des 3 conditions est illustrée par les figures 8 à 10. Le résultat de ces trois éliminations nous montre que R est le papillon bleu de la figure 11. Son aire est $\frac{1}{4}$. En conclusion :

La probabilité de pouvoir former un triangle avec les trois morceaux de la tige brisée est $1/4$.

qu'à chaque recompilation du source $\text{\LaTeX} 2\epsilon$, la figure 7 *change* et en particulier que je ne peux prévoir au moment où je tape ceci le nombre et l'emplacement précis des points bleus que vous observerez en lisant ce document. Il en va de même pour toutes les figures de ce document utilisant le générateur de nombres aléatoires de \METAPOST . Outre l'aspect ludique de la chose, l'intérêt est que j'échappe ainsi au soupçon d'avoir choisi les simulations qui marchent le mieux.

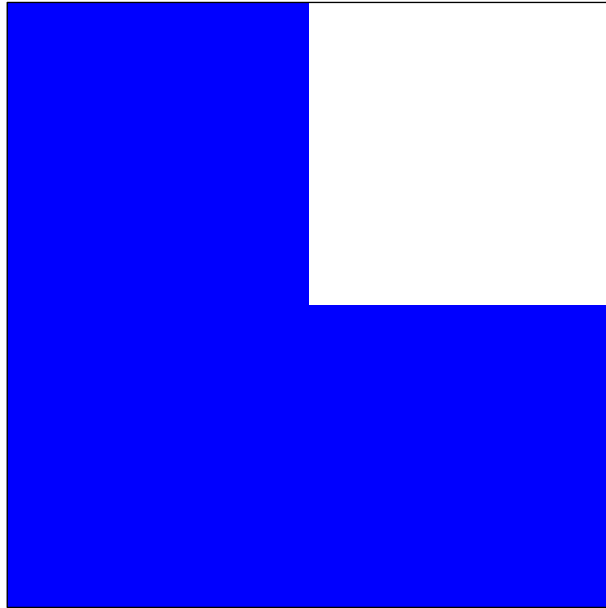


FIG. 8 – Suppression des (x, y) tels que $\min(x, y) > 1/2$

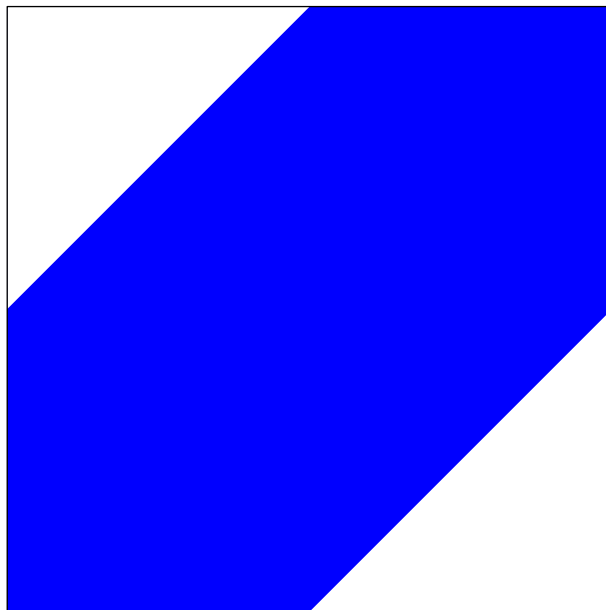


FIG. 9 – Suppression des (x, y) tels que $|x - y| > 1/2$

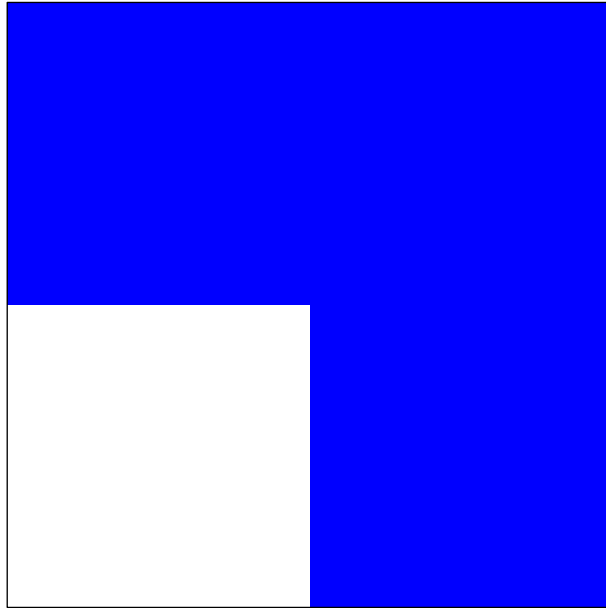


FIG. 10 – Suppression des (x, y) tels que $1 - \max(x, y) > 1/2$, soit $\max(x, y) < 1/2$

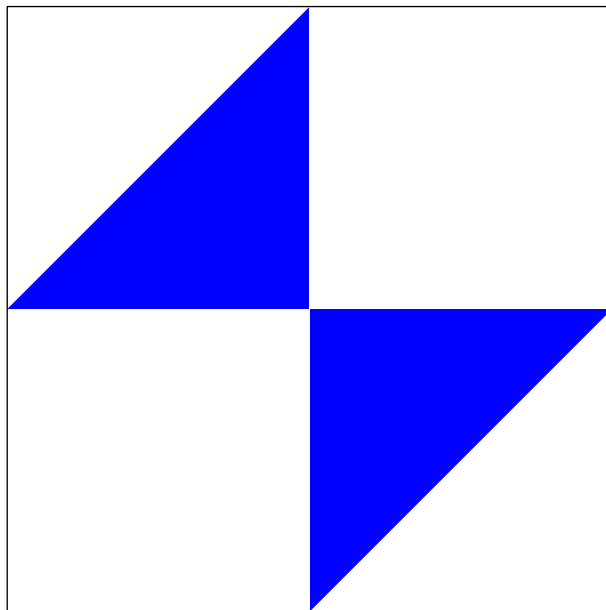


FIG. 11 – Zone R de succès, d'aire $\frac{1}{4}$

2.4 Espérance et loi des grands nombres

La figure 7 ressemble à une vue impressionniste du papillon de la figure 11. Elle nous permet de voir qu'en répétant un grand nombre de fois l'expérience du bris de la tige, on peut se faire une idée des probabilités associées à cette expérience. C'est là une

manifestation de ce que l'on appelle les *lois des grands nombres*. La plus simple des lois des grands nombres est la loi des grands nombres pour les fréquences.

Théorème (loi forte des grands nombres pour les fréquences). *Soit S_n le nombre de succès obtenus au cours de n épreuves répétées indépendantes, avec pour chaque épreuve probabilité de succès p . Quand n tend vers l'infini, la fréquence de succès S_n/n converge avec probabilité 1 vers p .*

On dit aussi que S_n/n converge *presque-sûrement* vers p . La signification pratique de cet énoncé est qu'avec un grand nombre d'épreuves, la fréquence de succès observée est « proche » de la probabilité de succès p . Ainsi sur la figure 7 où le nombre d'épreuves est $n = 1\,000$, la proportion de points bleus est proche de $1/4$.

Avant de voir une forme plus générale de la loi des grands nombres, il nous faut introduire la notion d'*espérance mathématique* d'une variable aléatoire. Il s'agit en quelque sorte de la valeur moyenne d'un nombre au hasard. Lorsque la variable aléatoire X ne peut prendre qu'un nombre fini de valeurs (par exemple les points d'un dé), l'espérance $\mathbf{E}X$ de X est effectivement la moyenne de ces valeurs, chacune ayant pour coefficient sa probabilité d'apparition. Dans le cas du dé, les valeurs possibles de X sont $1, 2, \dots, 6$ et si le dé est équilibré, ces valeurs ont toutes même probabilité d'apparition $\frac{1}{6}$, d'où $\mathbf{E}X = \frac{1}{6}(1 + 2 + \dots + 6) = \frac{21}{6} = 3,5$. On peut donner une interprétation géométrique à cette formule : $\mathbf{E}X$ est le barycentre (ou centre de gravité) du système de masses ponctuelles $(1, \frac{1}{6}), \dots, (6, \frac{1}{6})$ localisées sur le même segment, figure 12. Plus concrètement, on peut imaginer du fil de pêche en nylon tendu et de masse négligeable, sur lequel on a inséré des plombs de même masse $\frac{1}{6}$, aux graduations $1, 2, \dots, 6$. Alors $\mathbf{E}X$ est la distance à l'origine 0 du centre de gravité de la ligne lestée.

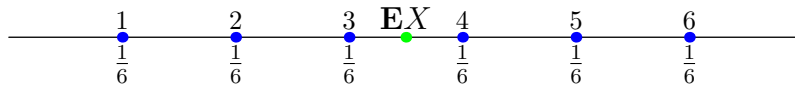


FIG. 12 – L'espérance du nombre de points d'un dé vue comme centre de gravité

Les choses se corsent lorsque X peut prendre une infinité de valeurs distinctes, par exemple si X suit la loi uniforme sur $[0, 1]$. En gardant à l'esprit l'interprétation comme centre de gravité, on s'attend dans le cas de la loi uniforme sur $[0, 1]$ à ce que $\mathbf{E}X$ vaille $\frac{1}{2}$ (le centre de gravité d'une tige homogène est son milieu). Une introduction rigoureuse de la notion d'espérance mathématique sortirait largement du cadre de cet exposé. Pour les lecteurs connaissant l'intégrale de Riemann, un bon compromis est la définition suivante qui bien que non classique, a l'avantage d'utiliser seulement des mathématiques du niveau au plus Bac+2. Pour une introduction à cette définition, voir [3]. Pour les autres lecteurs, ils peuvent se contenter de l'interprétation intuitive de l'espérance comme centre de gravité, qui devrait leur permettre de comprendre l'essentiel de ce document.

Définition (espérance mathématique d'une variable aléatoire positive). *Soit X une va-*

riable aléatoire positive. Son espérance notée $\mathbf{E}X$ est la quantité (éventuellement infinie)

$$\mathbf{E}X = \int_0^{+\infty} P(X > t) dt.$$

Laissant en exercice au lecteur le soin de vérifier que cette formule redonne bien la valeur 3,5 dans le cas d'un dé équilibré¹², regardons le cas d'une variable aléatoire X de loi uniforme sur $[0, 1]$. Dans ce cas, X la probabilité $P(X > 1)$ est nulle, donc pour tout $t \geq 1$, $P(X > t) = 0$, ce qui réduit l'intégrale entre 0 et $+\infty$ à une intégrale sur $[0, 1]$. De plus pour $0 \leq t < 1$, $P(X > t)$ est la probabilité que X tombe dans l'intervalle $]t, 1]$, c'est donc la longueur de cet intervalle soit $1 - t$. Ainsi

$$\mathbf{E}X = \int_0^1 P(X > t) dt = \int_0^1 (1 - t) dt = \dots = \frac{1}{2}.$$

On peut d'ailleurs retrouver géométriquement cette valeur en notant que l'intégrale $\int_0^1 (1 - t) dt$ est l'aire de la région délimitée par les axes du repère et la droite d'équation $y = 1 - t$. Cette région est simplement le triangle rectangle de sommets $(0, 0)$, $(0, 1)$ et $(1, 0)$, soit la moitié du carré unité. La valeur $1/2$ ainsi trouvée correspond bien à l'interprétation intuitive de l'espérance comme un barycentre. La loi uniforme sur $[0, 1]$ correspond à la répartition de masse sur une tige homogène de longueur 1 et dans ce cas, le barycentre ou centre de gravité de la tige est son milieu, d'où la valeur $1/2$.

Dans ce document, toutes les variables aléatoires utilisées sont positives. Nous nous contenterons donc de l'espérance des variables aléatoires positives et nous n'énoncerons la loi forte des grands nombres que pour ce type de variables.

Théorème (loi forte des gands nombres pour des variables aléatoires positives). *Soit $(X_k)_{k \geq 1}$ une suite de variables aléatoires positives, indépendantes et de même loi et $S_n = X_1 + \dots + X_n$ la somme des n premiers termes de cette suite. Quand n tend vers l'infini, S_n/n tend avec probabilité 1 vers $\mathbf{E}X_1$.*

Notons que si les X_k ont même loi, elles ont toutes même espérance. La loi forte des grands nombres pour les fréquences est un cas particulier de la loi forte des grands nombres ci-dessus. En effet dans une suite d'épreuves répétées indépendantes avec pour chacune probabilité de succès p , définissons les variables aléatoires X_k par

$$X_k = \begin{cases} 1 & \text{si la } k^{\text{e}} \text{ épreuve donne un succès,} \\ 0 & \text{si la } k^{\text{e}} \text{ épreuve donne un échec.} \end{cases}$$

Alors la somme S_n des X_k pour $k = 1$ jusqu'à n est exactement le nombre de succès lors des n premières épreuves : dans cette somme, les termes valent 0 ou 1 et il y a autant de 1 que de succès. Les X_k sont indépendantes et de même loi, de par la nature des épreuves répétées. Donc par la loi forte des grands nombres pour des variables aléatoires positives, S_n/n , qui est aussi la fréquence des succès en n épreuves, converge avec probabilité 1 vers $\mathbf{E}X_1$. Il ne nous reste plus qu'à calculer $\mathbf{E}X_1$ en notant que les seules valeurs prises par X_1 sont 0 avec probabilité $1 - p$ et 1 avec probabilité p , d'où

$$\mathbf{E}X_1 = 0 \times (1 - p) + 1 \times p = p.$$

12. Un conseil : représenter la fonction $t \mapsto P(X > t)$ dans ce cas et se ramener à un calcul d'aire.

2.5 Encore un peu de géométrie à l'appui des probabilités

La tige brisée m'ayant servi de prétexte pour parler d'espérance et de loi des grands nombres, je ne saurais résister au désir d'une digression supplémentaire initiée par l'énoncé que voici.

On choisit au hasard deux nombres X et Y entre 0 et 1, par exemple avec le générateur de nombres aléatoires d'un ordinateur. On forme alors les nombres au hasard $\min(X, Y)$ et $\max(X, Y)$. Trouver leur espérance par une méthode géométrique.

Outre son intérêt comme illustration des notions exposées ci-dessus, la résolution de ce problème nous fournira l'espérance des longueurs des trois morceaux de la tige brisée.

2.5.1 Simulations

Pour avoir une idée du résultat, on répète un grand nombre n de fois la génération de X et Y et on regarde la moyenne des résultats obtenus pour $\min(X, Y)$ et pour $\max(X, Y)$. Nous utilisons cette fois le logiciel Scilab¹³ mieux adapté à ce type de simulations. Voici une première sortie écran de l'appel à la fonction `minmax` qui calcule ces deux moyennes pour $n = 1\ 000$ couples simulés.

```
-->minmax(1000)
```

```
Simulation de 1000 couples de variables aléatoires uniformes sur [0,1]
```

```
Moyenne des minima : 0.337781
```

```
Moyenne des maxima : 0.661799
```

et affiche le graphique des fluctuations de ces moyennes entre 1 et n , cf. figure 13.

13. Scilab est un logiciel libre développé par l'INRIA. Il existe en versions Linux et Windows. On peut le télécharger depuis le site de l'INRIA à l'URL :

<http://scilabsoft.inria.fr/>

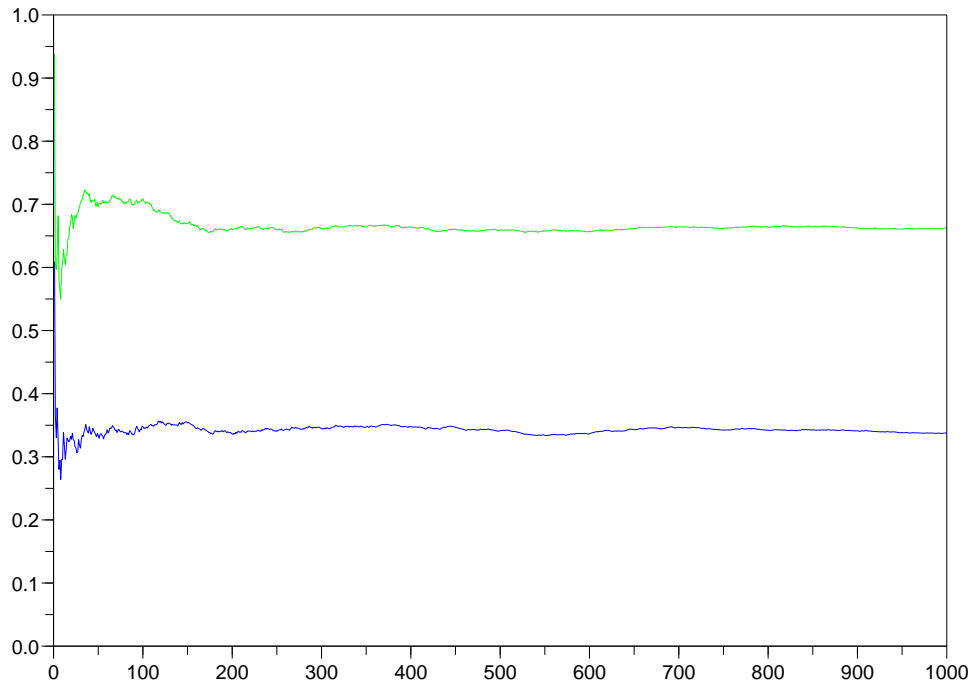


FIG. 13 – Fluctuations des moyennes du min et du max de deux nombres au hasard

Sans trop de mauvaise foi, il semble que les deux suites de moyennes convergent vers $1/3$ pour le min et $2/3$ pour le max. Pour conforter cette première impression, faisons un nouvel essai avec $n = 100\ 000$ couples simulés au lieu de 1 000.

```
-->minmax(100000)
```

```
Simulation de 100000 couples de variables aléatoires uniformes sur [0,1]
```

```
Moyenne des minima : 0.334307
```

```
Moyenne des maxima : 0.666646
```

2.5.2 Résolution géométrique

Nous avons admis plus haut que le point aléatoire de coordonnées (X, Y) suit la loi uniforme sur le carré unité $OABC$. Il en résulte que le point aléatoire de coordonnées $(\min(X, Y), \max(X, Y))$ suit la loi uniforme sur le triangle OBC de sommets $(0, 0)$, $(1, 1)$, $(0, 1)$, cf. figure 14.

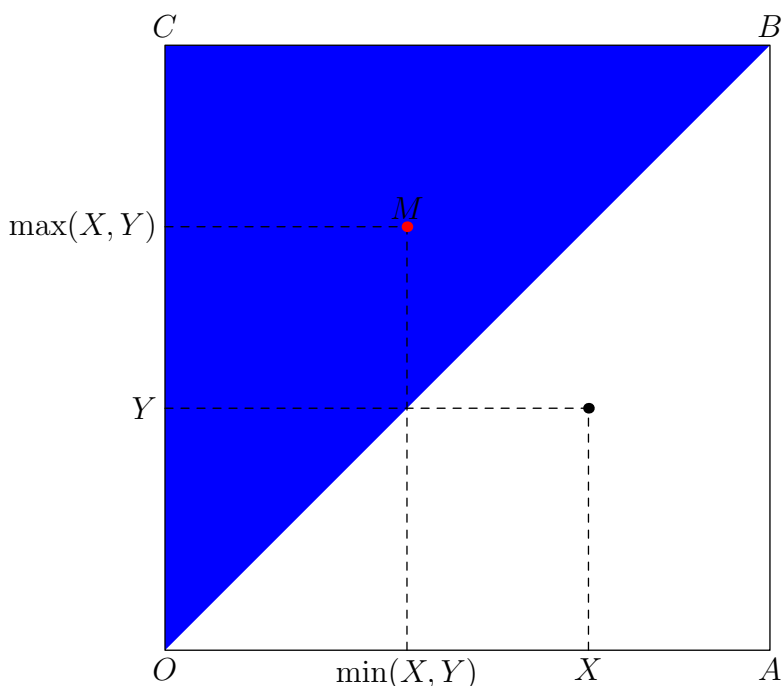


FIG. 14 – M de coordonnées $\min(X, Y)$ et $\max(X, Y)$ suit la loi uniforme sur un triangle

Sans prétendre démontrer cette affirmation, en voici une justification intuitive. On peut voir la loi d'un vecteur aléatoire (un couple de deux variables aléatoires) comme une répartition de masse dans le plan : si R est un ensemble « convenable » du plan $P((X, Y) \in R)$ peut alors s'interpréter comme la masse de l'ensemble R . De ce point de vue, la loi uniforme sur le carré $OABC$ peut être vue comme la répartition de masses d'une plaque de carton homogène de masse totale 1 recouvrant le carré. Comme on passe de (X, Y) à $(\min(X, Y), \max(X, Y))$ par repliement du carton le long de la diagonale OB , on peut facilement se convaincre que la répartition de masse associée au vecteur $(\min(X, Y), \max(X, Y))$ est celle d'une plaque de carton homogène recouvrant le triangle OBC et de masse totale 1 (la masse du triangle OBC n'était que de $1/2$ pour la loi de (X, Y)). Autrement dit, le vecteur aléatoire $(\min(X, Y), \max(X, Y))$ suit la loi uniforme sur le triangle OBC .

Avec l'interprétation de la loi d'un vecteur aléatoire comme une répartition de masse, on peut définir l'espérance d'un vecteur aléatoire comme le *centre de gravité* de cette répartition de masse. Nous admettons que l'espérance du vecteur est égale au vecteur des espérances des coordonnées, autrement dit que $\mathbf{E}(X, Y) = (\mathbf{E}X, \mathbf{E}Y)$. Géométriquement, ceci correspond à la conservation du barycentre par projection sur les axes.

Enfin nous admettons que le centre de gravité d'une plaque triangulaire homogène est le même que celui de trois masses égales localisées aux sommets du triangle, l'isobarycentre des sommets. Ceci nous permet de nous brancher sur une propriété bien connue en géométrie élémentaire : l'isobarycentre des sommets d'un triangle est le point d'intersection G des trois médianes. On sait de plus que sur chacune des médianes joignant un sommet au milieu du côté opposé, ce point G se trouve aux deux tiers de la longueur

de la médiane en partant du sommet.

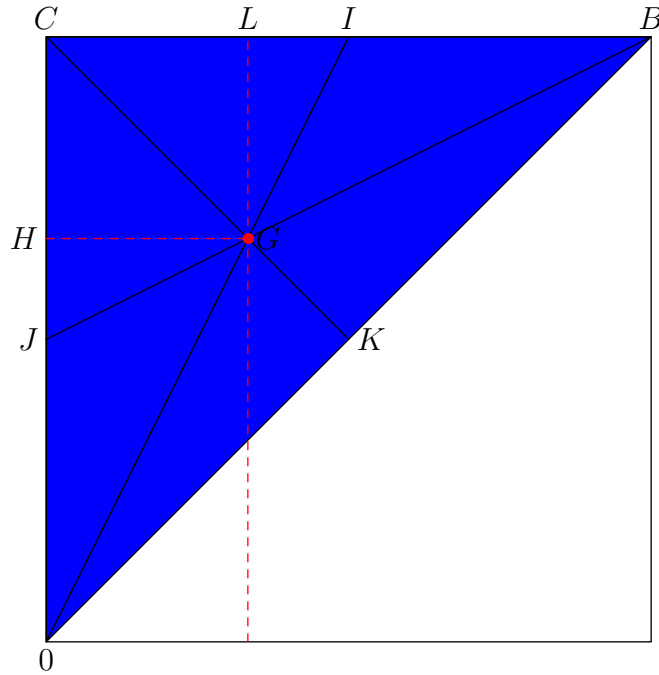


FIG. 15 – $G = \mathbf{E}(\min(X, Y), \max(X, Y)) = \left(\frac{1}{3}, \frac{2}{3}\right)$

Examinons maintenant la figure 15. Le centre de gravité G du triangle OBC a pour coordonnées $\mathbf{E} \min(X, Y)$ et $\mathbf{E} \max(X, Y)$. Il nous suffit donc de calculer les distances CL et OH , ce qui peut se faire par deux invocations du vénérable Thalès. Les droites sécantes OI et OC sont coupées par les parallèles HG et CI , donc $\frac{OH}{OC} = \frac{OG}{OI} = \frac{2}{3}$ et comme $OC = 1$, on en déduit que

$$\mathbf{E} \max(X, Y) = OH = \frac{2}{3}.$$

De même les droites sécantes BC et BJ sont coupées par les parallèles CJ et LG , donc $\frac{BL}{BC} = \frac{BG}{BJ} = \frac{2}{3}$ et comme $BC = 1$, $BL = \frac{2}{3}$, d'où $CL = \frac{1}{3}$ et

$$\mathbf{E} \min(X, Y) = CL = \frac{1}{3}.$$

2.5.3 Vérification calculatoire

La résolution géométrique présentée ci-dessus nous ayant amené à admettre des propriétés relatives à la loi du vecteur aléatoire (X, Y) et au centre de gravité, il n'est pas superflu de procéder à une vérification calculatoire du résultat obtenu.

Pour calculer $\mathbf{E} \min(X, Y)$, commençons par expliciter $P(\min(X, Y) > t)$ pour t réel positif. Les variables aléatoires X et Y suivant la loi uniforme sur $[0, 1]$, il est clair que

$P(X > 1) = P(Y > 1) = 0$, d'où $P(\min(X, Y) > 1) = 0$ et $P(\min(X, Y) > t) = 0$ pour tout $t \geq 1$. Pour $t \in [0, 1]$, on a

$$P(\min(X, Y) > t) = P(X > t \text{ et } Y > t) = P(X > t)P(Y > t),$$

la deuxième égalité résultant de l'indépendance de X et Y . Nous avons vu plus haut que pour $t \in [0, 1]$, $P(X > t) = 1 - t$ et comme Y a même loi que X , ceci vaut aussi pour Y . Finalement

$$P(\min(X, Y) > t) = \begin{cases} (1 - t)^2 & \text{si } 0 \leq t \leq 1, \\ 0 & \text{si } t > 1. \end{cases}$$

On en déduit que

$$\mathbf{E} \min(X, Y) = \int_0^{+\infty} P(\min(X, Y) > t) dt = \int_0^1 (1 - t)^2 dt = \left[\frac{-(1 - t)^3}{3} \right]_0^1 = \frac{1}{3}.$$

Pour calculer $\mathbf{E} \max(X, Y)$, on peut se passer d'un calcul d'intégrale en exploitant l'additivité de l'espérance et l'identité suivante :

$$X + Y = \min(X, Y) + \max(X, Y).$$

En prenant l'espérance des deux membres, il vient

$$\frac{1}{2} + \frac{1}{2} = \frac{1}{3} + \mathbf{E} \max(X, Y), \quad \text{d'où} \quad \mathbf{E} \max(X, Y) = \frac{2}{3}.$$

Remarquons pour finir que $|X - Y| = \max(X, Y) - \min(X, Y)$, d'où par linéarité de l'espérance $\mathbf{E}|X - Y| = \mathbf{E} \max(X, Y) - \mathbf{E} \min(X, Y) = \frac{2}{3} - \frac{1}{3} = \frac{1}{3}$. D'autre part $\mathbf{E}(1 - \max(X, Y)) = \mathbf{E}1 - \mathbf{E} \max(X, Y) = 1 - \frac{2}{3} = \frac{1}{3}$. Les longueurs aléatoires des trois morceaux de la tige brisée ont donc même espérance $\frac{1}{3}$. En fait on pourrait montrer que ces trois longueurs aléatoires ont *même loi*, ce qui rendrait superflu le calcul ci-dessus : si elles ont même loi, elles ont même espérance c et comme la somme des trois longueurs est la constante 1, par additivité de l'espérance on a $3c = 1$, d'où $c = \frac{1}{3}$.

2.6 D'autres modèles ?

Le sujet de la tige brisée est loin d'être épuisé. Dans le cadre de ce texte, il me reste à envisager quelques modèles alternatifs. La procédure de découpage adoptée ci-dessus (choix des points de rupture indépendants et de même loi uniforme sur $[0, 1]$) avait l'avantage de nous conduire à des calculs simples. On peut lui opposer d'autres procédures peut-être plus naturelles, comme par exemple :

Procédure 2 : On choisit d'abord un premier point de fracture X , suivant la loi uniforme sur $[0, 1]$, on découpe et ensuite on tire à pile ou face entre le morceau de gauche et celui de droite. On choisit alors un deuxième point de rupture suivant une loi uniforme sur le morceau ainsi sélectionné.

Procédure 3 : On choisit d'abord un premier point de fracture X , suivant la loi uniforme sur $[0, 1]$, on découpe et ensuite on sélectionne le plus long des deux morceaux obtenus¹⁴ sur lequel on choisit le point de rupture suivant la loi uniforme sur ce morceau.

Commençons par quelques simulations qui nous font subodorer que la probabilité de constructibilité du triangle n'est plus $1/4$, ni avec la procédure 2, ni avec la 3.

```
-->tige_brisee1(100000)
```

Bris de 100000 tiges, procédure 1

Fréquence de constructibilité du triangle : 0.250580

```
-->tige_brisee1(100000)
```

Bris de 100000 tiges, procédure 1

Fréquence de constructibilité du triangle : 0.247780

```
-->tige_brisee2(100000)
```

Bris de 100000 tiges, procédure 2

Fréquence de constructibilité du triangle : 0.192200

```
-->tige_brisee2(100000)
```

Bris de 100000 tiges, procédure 2

Fréquence de constructibilité du triangle : 0.191800

```
-->tige_brisee2(100000)
```

Bris de 100000 tiges, procédure 2

Fréquence de constructibilité du triangle : 0.193270

```
-->tige_brisee2(100000)
```

Bris de 100000 tiges, procédure 2

Fréquence de constructibilité du triangle : 0.193210

14. La procédure qui consisterait à choisir le plus court des deux donnerait une probabilité nulle de construire un triangle. Si cela ne vous paraît pas évident, relisez la conclusion de la sous-section 2.1.

```
-->tige_brisee3(100000)
```

Bris de 100000 tiges, procédure 3

Fréquence de constructibilité du triangle : 0.386640

```
-->tige_brisee3(100000)
```

Bris de 100000 tiges, procédure 3

Fréquence de constructibilité du triangle : 0.385710

```
-->tige_brisee3(100000)
```

Bris de 100000 tiges, procédure 3

Fréquence de constructibilité du triangle : 0.385900

```
-->tige_brisee3(100000)
```

Bris de 100000 tiges, procédure 3

Fréquence de constructibilité du triangle : 0.386890

Essayons maintenant de trouver quelles probabilités exactes se cachent derrière ces valeurs approchées 0,19... pour la procédure 2 et 0,38... pour la 3. Pour la simplification des calculs, il est commode de faire un petit détour par la procédure 2 bis qui consiste à choisir systématiquement le morceau gauche, donc $[0, X]$, après la première rupture. Pour découper le morceau gauche, on génère une variable aléatoire Z indépendante de X et de loi uniforme sur $[0, 1]$ et on découpe au point XZ , obtenant finalement les 3 morceaux de longueurs respectives XZ , $X(1 - Z)$ et $1 - X$. Puisque X et Z sont indépendantes et de même loi uniforme sur $[0, 1]$, le vecteur aléatoire (X, Z) suit la loi uniforme sur le carré unité C . La probabilité de constructibilité du triangle est alors égale à l'aire de la région R définie par

$$R = \left\{ (x, z) \in C; xz \leq \frac{1}{2}, x(1 - z) \leq \frac{1}{2}, 1 - x \leq \frac{1}{2} \right\}.$$

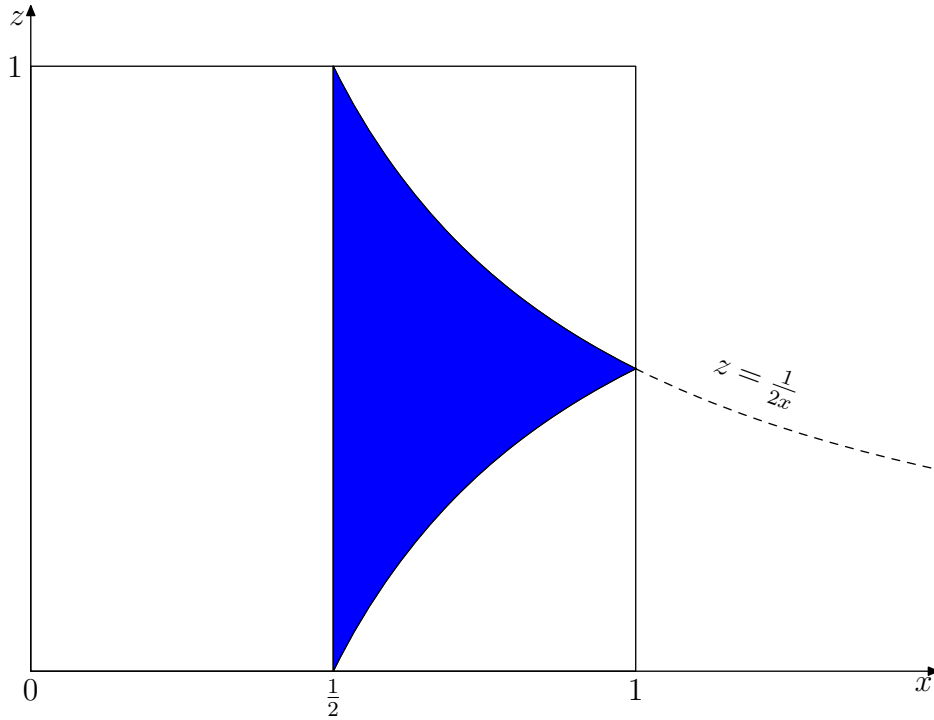


FIG. 16 – Région de succès, procédure 2 bis : choix systématique du morceau gauche

Cette région représentée à la figure 16 est le triangle curviligne délimité par les hyperboles d'équations $xz = \frac{1}{2}$, $x(1-z) = \frac{1}{2}$ et la droite d'équation $x = \frac{1}{2}$. Notons que le deuxième arc d'hyperbole est symétrique du premier par rapport à la droite d'équation $z = \frac{1}{2}$. On en déduit facilement que

$$\text{aire}(R) = 2 \left(\int_{\frac{1}{2}}^1 \frac{1}{2x} dx - \frac{1}{4} \right) = \int_{\frac{1}{2}}^1 \frac{1}{x} dx - \frac{1}{2} = \ln 2 - \frac{1}{2}.$$

Ainsi la probabilité de constructibilité du triangle pour la procédure 2 bis est

$$p_{2b} = \ln 2 - \frac{1}{2} \simeq 0,193\ 147.$$

Par symétrie, il est clair que pour la procédure 2 ter qui consiste à choisir systématiquement le morceau de droite (donc $[1 - X, 1]$) après la première rupture, on obtient la même probabilité de constructibilité : $p_{2t} = p_{2b}$. La région de succès R' est d'ailleurs la symétrique de R par rapport à la droite d'équation $x = 1/2$.

Revenons à la procédure 2 et convenons que si la pièce de monnaie donne face on choisit le morceau gauche et si elle donne pile on choisit le morceau droit. La probabilité de constructibilité du triangle sachant que l'on a choisi le morceau gauche est précisément p_{2b} . De même probabilité de constructibilité du triangle sachant que l'on a choisi le morceau droit est p_{2t} . On obtient alors la probabilité de constructibilité (ou « succès ») p_2 en conditionnant par les deux cas possibles (formule des probabilités totales) :

$$p_2 = P(\text{succès} \mid \text{face})P(\text{face}) + P(\text{succès} \mid \text{pile})P(\text{pile}) = \frac{1}{2}p_{2b} + \frac{1}{2}p_{2t} = p_{2b} = \ln 2 - \frac{1}{2}.$$

Notons au passage que le résultat serait le même avec une pièce de monnaie non équilibrée. Les simulations présentées ci-dessus semblent bien s'accorder avec la valeur trouvée pour p_2 .

Le calcul pour la procédure 3 utilise la même méthode que ci-dessus : exprimer l'évènement « succès » à l'aide de (X, Z) et se ramener à un calcul d'aire. Si $X \leq \frac{1}{2}$, les trois morceaux de la tige ont pour longueurs respectives X , $Z(1 - X)$ et $(1 - Z)(1 - X)$, tandis que si $X > \frac{1}{2}$, ces longueurs sont $1 - X$, $(1 - Z)X$ et ZX . On voit alors facilement que la région de succès dans le plan des x, z est le quadrilatère curviligne représenté à la figure 17 qui peut s'obtenir comme $R \cup R'$, où R est le triangle curviligne de succès pour la procédure 2 bis représenté à la figure 16 et R' est son symétrique par rapport à la droite d'équation $x = \frac{1}{2}$. On en déduit que

$$p_3 = 2p_2 = 2 \ln 2 - 1 \simeq 0,386\ 294.$$

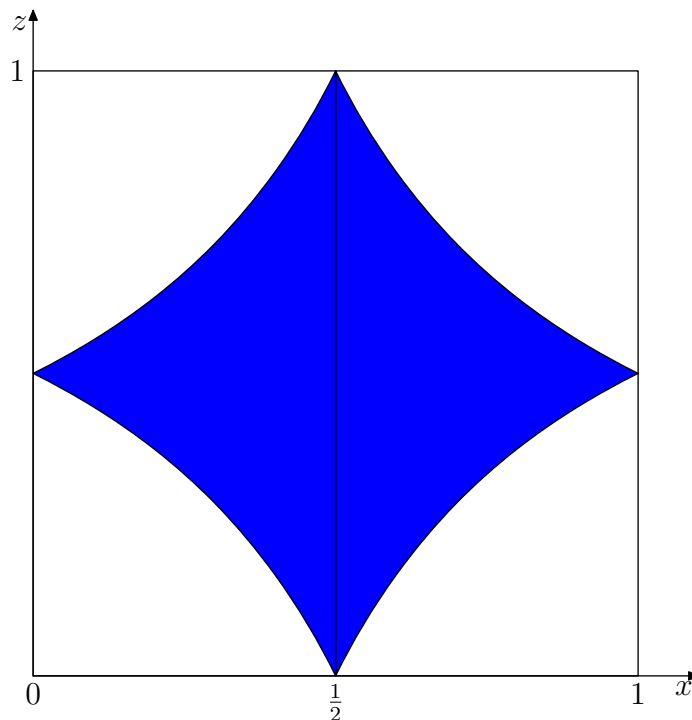


FIG. 17 – Région de succès, procédure 3 : choix du morceau le plus long

3 L'aiguille de Buffon

Georges Louis LECLERC, comte de BUFFON (1707–1788) est d'abord connu comme l'auteur de l'*Histoire naturelle* et l'organisateur du Jardin des Plantes de Paris. Son oeuvre imposante comprend aussi des écrits mathématiques. Dans son *Mémoire sur le jeu de franc carreau*, présenté à l'Académie des Sciences les 14 et 17 mars 1736, on peut trouver le célèbre problème de l'aiguille.



FIG. 18 – Georges Louis LECLERC, comte de BUFFON (1707–1788)

Sur un parquet dont les rainures sont distantes de a , on laisse tomber une aiguille de longueur $\ell < a$. On parie avant le lancer qu'elle va couper l'une des rainures.

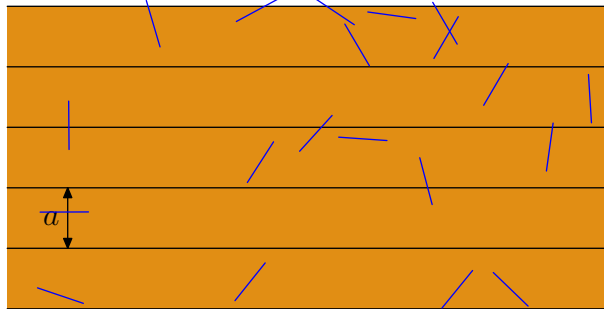


FIG. 19 – Aiguilles lancées sur un parquet

Soit N le nombre aléatoire d'intersections de l'aiguille avec les rainures. Comme $\ell < a$, on a $N = 0$ ou 1 . L'espérance de N est donc

$$\mathbf{E}N = 0 \times P(N = 0) + 1 \times P(N = 1) = P(N = 1).$$

Buffon a montré que

$$\mathbf{E}N = \frac{2\ell}{\pi a}$$

D'après la loi des grands nombres, si on répète un grand nombre de fois ce lancer d'aiguille et que l'on note M la moyenne des valeurs observées pour N , on a

$$M \simeq \mathbf{E}N = \frac{2\ell}{\pi a}$$

d'où

$$\pi \simeq \frac{2\ell}{Ma}$$

Ceci fournit une méthode expérimentale pour obtenir une valeur approchée du nombre π .

3.1 Faisons travailler Scilab

On peut facilement réaliser l'expérience soi-même pour tester la véracité des affirmations ci-dessus. Plus paresseusement, voici quelques résultats de simulations confiées à Scilab¹⁵.

```
-->buffon(20000,10,6.75)
```

Distance inter-rainures : 10.0 cm, lancer de 20000 aiguilles de 6.8 cm

Nombre d'intersections : 8613

Fréquence des intersections : 0.4306

Probabilité d'intersection d'après Buffon : 0.429718

Estimation de pi : 3.1348

Erreur d'estimation : -0.0068 Erreur relative : -0.22%

```
-->buffon(20000,10,6.75)
```

Distance inter-rainures : 10.0 cm, lancer de 20000 aiguilles de 6.8 cm

Nombre d'intersections : 8547

Fréquence des intersections : 0.4274

Probabilité d'intersection d'après Buffon : 0.429718

Estimation de pi : 3.1590

Erreur d'estimation : 0.0174 Erreur relative : 0.55%

```
-->buffon(20000,10,9)
```

Distance inter-rainures : 10.0 cm, lancer de 20000 aiguilles de 9.0 cm

Nombre d'intersections : 11439

Fréquence des intersections : 0.5719

15. C'est moins probant que de réaliser l'expérience, car le programme sous-jacent utilise des hypothèses de modélisation qui sont celles exposées ci-dessous, sous-section 3.3.

Probabilité d'intersection d'après Buffon : 0.572958

Estimation de pi : 3.1471

Erreur d'estimation : 0.0055 Erreur relative : 0.18%

-->buffon(20000,10,9)

Distance inter-rainures : 10.0 cm, lancer de 20000 aiguilles de 9.0 cm

Nombre d'intersections : 11436

Fréquence des intersections : 0.5718

Probabilité d'intersection d'après Buffon : 0.572958

Estimation de pi : 3.1480

Erreur d'estimation : 0.0064 Erreur relative : 0.20%

-->buffon(80000,10,9)

Distance inter-rainures : 10.0 cm, lancer de 80000 aiguilles de 9.0 cm

Nombre d'intersections : 45833

Fréquence des intersections : 0.5729

Probabilité d'intersection d'après Buffon : 0.572958

Estimation de pi : 3.1418

Erreur d'estimation : 0.0002 Erreur relative : 0.01%

-->buffon(80000,10,9)

Distance inter-rainures : 10.0 cm, lancer de 80000 aiguilles de 9.0 cm

Nombre d'intersections : 45756

Fréquence des intersections : 0.5719

Probabilité d'intersection d'après Buffon : 0.572958

Estimation de pi : 3.1471

Erreur d'estimation : 0.0055 Erreur relative : 0.18%

3.2 Tentative d'explication

Voici une méthode astucieuse due à Émile Borel. On admet que si on remplace l'aiguille par une ligne brisée, l'espérance du nombre d'intersections avec les rainures est la somme des espérances de chaque segment considéré comme une aiguille individuelle. On en déduit alors que pour une ligne brisée, l'espérance de N est proportionnelle à la longueur totale ℓ de la ligne :

$$\mathbf{E}N = c\ell,$$

où c est une constante ne dépendant que de a .

On admet ensuite que cette propriété reste vraie pour des « aiguilles » courbes qui peuvent s'obtenir comme limites de lignes brisées. C'est le cas en particulier pour un anneau circulaire.

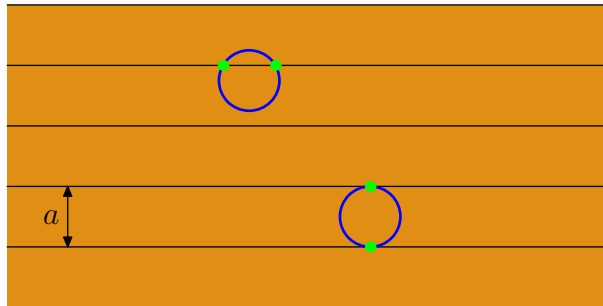


FIG. 20 – Pour un anneau de diamètre a , $N = 2$ donc $\mathbf{E}N = 2$

Si cet anneau a pour diamètre a , N vaut toujours 2, donc $\mathbf{E}N = 2$. Comme ℓ vaut alors πa , on en déduit

$$c\pi a = 2 \quad \text{d'où} \quad c = \frac{2}{\pi a}.$$

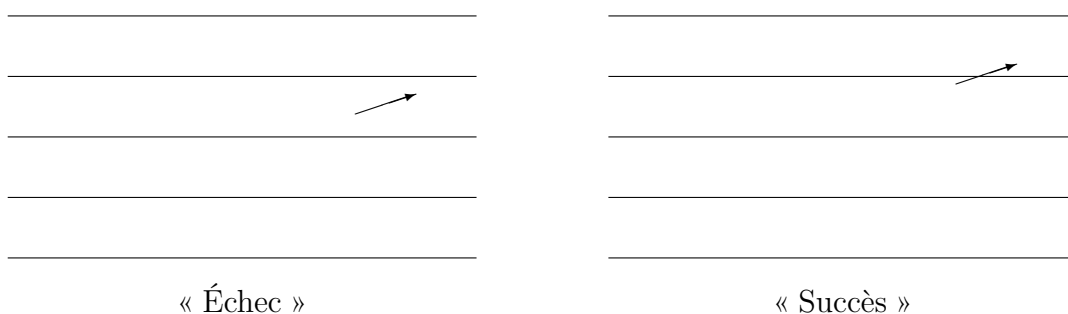
En reportant cette valeur de c dans la formule générale, on obtient bien

$$\mathbf{E}N = \frac{2\ell}{\pi a}.$$

3.3 Le cours auquel vous avez échappé

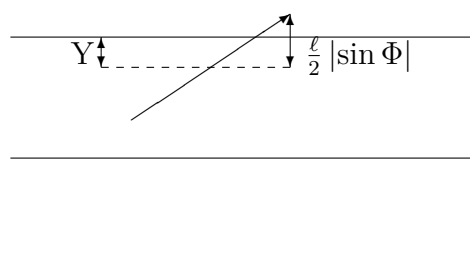
Pour les lecteurs qui souhaiteraient une solution plus formelle, j'ai collé ici un extrait de mon cours de Licence de Mathématiques (3^e année) : *Intégration et Probabilités élémentaires* [3]. Dans ce texte, λ_1 et λ_2 désignent la mesure de Lebesgue en dimension 1 et 2 respectivement.

À titre d'illustration historique de la loi forte des grands nombres, nous présentons maintenant une méthode expérimentale pour obtenir une approximation numérique du nombre π dont l'intérêt est essentiellement d'ordre culturel ¹⁶. Cette méthode a été proposée en 1777 par le célèbre naturaliste Buffon. On trace sur une surface plane horizontale des droites parallèles équidistantes, séparées par une distance a (on peut par exemple utiliser les rainures d'un parquet). On laisse tomber sur cette surface une aiguille de longueur $\ell \leq a$ et une fois l'aiguille immobilisée, on observe si elle coupe l'une des droites du réseau. On répète l'expérience en notant la fréquence des intersections. Lorsque le nombre d'expériences augmente indéfiniment, cette fréquence converge selon Buffon vers $p = \frac{2\ell}{\pi a}$ permettant ainsi d'obtenir une estimation expérimentale du nombre π .



Cherchons une modélisation de cette expérience. On note Y la distance du milieu de l'aiguille à la droite du réseau la plus proche. Y prend ses valeurs dans $[0, \frac{a}{2}]$. On note Φ une mesure de l'angle entre les droites du réseau (toutes orientées dans le même sens) et l'aiguille orientée du chas vers la pointe. Φ prend ses valeurs dans $[0, 2\pi]$ (par exemple) ¹⁷.

Y et Φ sont des variables aléatoires. La connaissance du couple $(Y(\omega), \Phi(\omega))$ suffit pour savoir s'il y a ou non intersection.



Nous ferons les hypothèses suivantes sur les variables aléatoires Y et Φ :

- (H_1) Y suit la loi uniforme sur $[0, \frac{a}{2}]$.
- (H_2) Φ suit la loi uniforme sur $[0, 2\pi]$.
- (H_3) Y et Φ sont indépendantes.

On note E l'événement « l'aiguille coupe l'une des droites du réseau ». La longueur de la projection de la demi-aiguille sur une droite orthogonale au réseau est $Z = \frac{\ell}{2} |\sin \Phi|$. Il

16. C'est aussi un excellent problème de révision.

17. On pourrait aussi utiliser les angles de droites, Φ serait alors à valeurs dans un intervalle de longueur π .

Il y a donc intersection si et seulement si la distance Y du centre de l'aiguille à la droite du réseau la plus proche est inférieure ou égale à Z . Ceci nous permet d'écrire l'évènement E sous la forme :

$$E = \left\{ Y \leq \frac{\ell}{2} |\sin \Phi| \right\}.$$

Comme Y et Φ sont indépendantes, la loi du couple est le produit des lois marginales : $P_{(Y,\Phi)} = P_Y \otimes P_\Phi$. Comme ces lois marginales sont à densités par rapport à λ_1 , on en déduit que $P_{(Y,\Phi)}$ est à densité $f_Y \otimes f_\Phi$ par rapport à λ_2 . D'où

$$f_{(Y,\Phi)}(y, t) = f_Y(y)f_\Phi(t) = \frac{2}{a} \mathbf{1}_{[0, a/2]}(y) \frac{1}{2\pi} \mathbf{1}_{[0, 2\pi]}(t) = \frac{1}{a\pi} \mathbf{1}_{[0, a/2] \times [0, 2\pi]}(y, t).$$

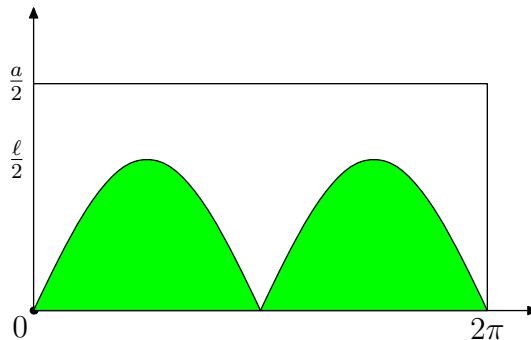
On voit ainsi que le couple (Y, Φ) suit la loi uniforme sur le rectangle $[0, a/2] \times [0, 2\pi]$. Notons D le borélien de \mathbb{R}^2 défini par

$$D := \left\{ (t, y) \in [0, 2\pi] \times [0, a/2]; y \leq \frac{\ell}{2} |\sin t| \right\}.$$

Comme l'évènement E s'écrit aussi $\{(\Phi, Y) \in D\}$, on peut calculer $P(E)$ en utilisant la loi du couple (Φ, Y) qui est la loi uniforme sur $[0, 2\pi] \times [0, a/2]$:

$$\begin{aligned} P(E) &= P((\Phi, Y) \in D) = P_{(\Phi, Y)}(D) \\ &= \int_D f_{(\Phi, Y)}(y, t) \, dy \, dt \\ &= \frac{1}{a\pi} \lambda_2(D \cap [0, 2\pi] \times [0, a/2]) \\ &= \frac{1}{a\pi} \lambda_2(D), \end{aligned}$$

en remarquant que $D \subset [0, 2\pi] \times [0, a/2]$. Le calcul de $P(E)$ se réduit ainsi à celui de l'aire de l'hypographe de la fonction $g : t \mapsto \frac{\ell}{2} |\sin t| \mathbf{1}_{[0, 2\pi]}(t)$.



Par conséquent

$$\lambda_2(D) = \int_0^{2\pi} \frac{\ell}{2} |\sin t| \, dt = \ell \int_0^\pi \sin t \, dt = 2\ell.$$

Finalement,

$$P(E) = \frac{2\ell}{a\pi}.$$

On effectue une suite de lancers de l'aiguille et on note E_i l'évènement « lors du i^e lancer, l'aiguille intersecte une des droites du réseau ». On pose $X_i = \mathbf{1}_{E_i}$ et

$$F_n := \frac{1}{n} \sum_{i=1}^n X_i.$$

Les X_i sont des variables aléatoires de Bernoulli, de paramètre $p := P(E_i) = P(E)$. Elles sont clairement intégrables, puisque bornées. Les E_i forment une suite d'évènements mutuellement indépendants et de même probabilité p . Il en résulte que les X_i forment une suite de variables aléatoires indépendantes et de même loi Bern(p). Par la loi forte des grands nombres pour des variables i.i.d. et intégrables,

$$F_n := \frac{1}{n} \sum_{i=1}^n X_i \xrightarrow[n \rightarrow +\infty]{\text{p.s.}} \mathbf{E}X_1 = P(E).$$

Compte-tenu du calcul de $P(E)$, on peut réécrire ce résultat sous la forme :

$$\frac{2\ell}{aF_n} \xrightarrow[n \rightarrow +\infty]{\text{p.s.}} \pi.$$

L'interprétation physique est la suivante. Si on réalise *une* série de lancers avec n grand, la valeur $F_n(\omega)$ observée nous fournira l'approximation

$$\frac{2\ell}{aF_n(\omega)} \simeq \pi.$$

On considère ainsi qu'il est physiquement impossible d'observer un ω n'appartenant pas à l'évènement de probabilité 1 $\{F_n \text{ converge vers } p\}$.

4 Le paradoxe de Bertrand

Dans son livre *Calcul des Probabilités* [1] publié en 1889, Joseph BERTRAND utilise le problème suivant pour discuter les fondements du calcul des probabilités lorsque l'ensemble des réalisations possibles est un « continu » (un ensemble infini dont on ne peut numéroter les éléments, comme un segment, un carré, un disque...).

On trace une corde « au hasard » sur un cercle. Quelle est la probabilité qu'elle soit plus longue que le côté d'un triangle équilatéral inscrit dans le cercle ?

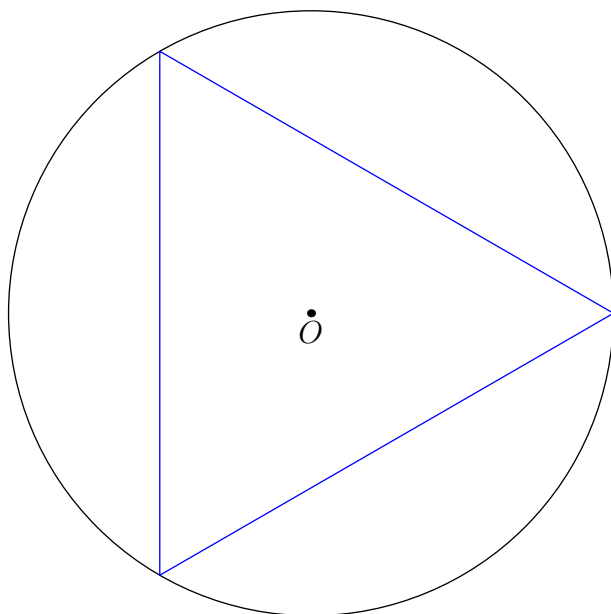


FIG. 21 – Triangle équilatéral inscrit dans le cercle

Ce problème qui a fait couler beaucoup d'encre chez les mathématiciens est resté célèbre sous le nom de paradoxe de Bertrand. Le paradoxe est dans l'existence de plusieurs solutions donnant des résultats différents pour la probabilité cherchée et qui peuvent toutes paraître assez naturelles. Voyons cela de plus près. Les simulations présentées ci-dessous utilisent le générateur de nombres au hasard de `METAPOST`. Les cordes plus longues que le côté du triangle équilatéral sont tracées en vert, les autres en rouge. Le nombre de cordes vertes est compté automatiquement et affiché en haut à droite du cercle. Pour la lisibilité des figures, on n'a généré et tracé qu'un petit nombre de cordes (40 en tout). Il est donc illusoire d'attendre une approximation précise de la probabilité cherchée par la fréquence de succès observée. On peut quand même noter des différences sensibles entre les nombres de cordes vertes pour les trois solutions.

4.1 Première solution

Pour tracer la corde AB de milieu M , il suffit de choisir au hasard la distance OM entre 0 et R rayon du cercle, ainsi que l'angle \widehat{IOM} entre 0 et 360° , figure 22. On considère ici que la distance aléatoire OM et la mesure de l'angle aléatoire \widehat{IOM} sont des variables aléatoires indépendantes, de loi uniforme¹⁸, respectivement sur $[0, R]$ et sur $[0, 360]$. La probabilité cherchée s'écrit dans ce cas $P(OM < \frac{1}{2})$ et vaut $\frac{1}{2}$.

18. Je considère ici que la loi de la mesure de l'angle est uniforme sur $[0, 360]$ par simplicité, mais en réalité le point important est l'indépendance des deux variables aléatoires choisies et on trouverait le même résultat final avec *n'importe quelle autre loi* pour la mesure de l'angle \widehat{IOM} .

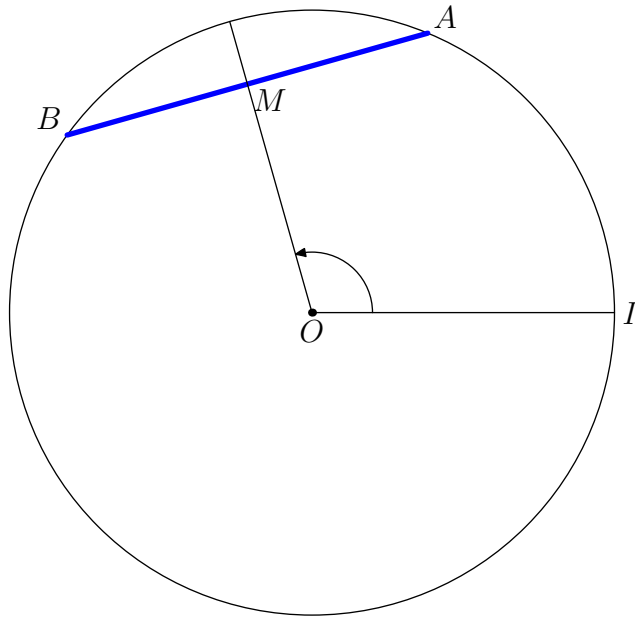


FIG. 22 – Choix de la corde par la distance OM et l'angle \widehat{IOM}

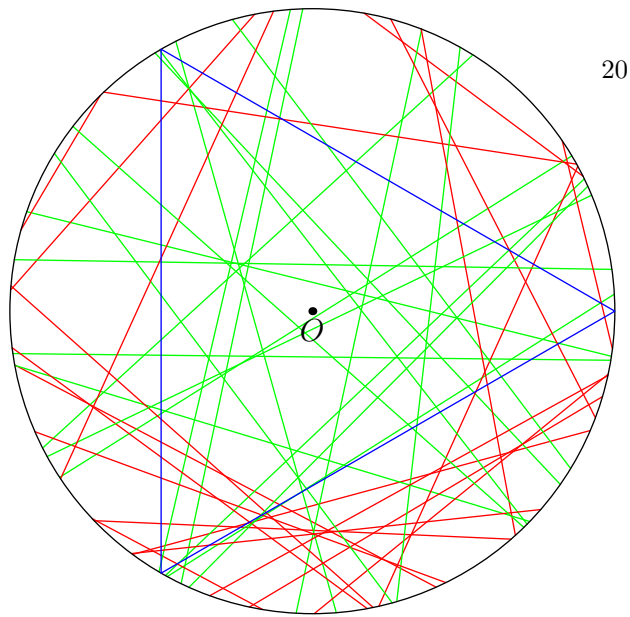


FIG. 23 – 40 cordes tracées au hasard (loi uniforme sur le rayon)

4.2 Deuxième solution

Pour tracer une corde au hasard, il suffit, à une rotation près, de choisir au hasard l'angle au centre.

En pratique, on effectue une première rotation de centre O et d'angle aléatoire de loi uniforme sur $[0, 360]$ qui transforme OI en OA , figure 24.

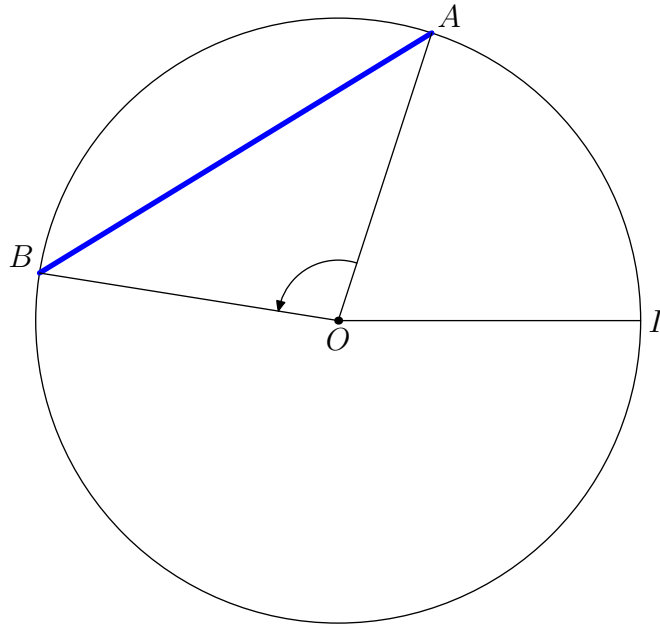


FIG. 24 – Choix de la corde par l'angle au centre

Ensuite on effectue indépendamment une rotation aléatoire de même loi (cela signifie en pratique que l'on va choisir la valeur d'une deuxième variable aléatoire uniforme sur $[0, 360]$ pour l'angle) qui transforme OA en OB . On regarde enfin la longueur de la corde AB ainsi obtenue. Ayant été trop bavard jusqu'ici, je me contenterai de la preuve muette suggérée par la figure 26 pour vous convaincre que la probabilité cherchée est maintenant $\frac{1}{3}$.

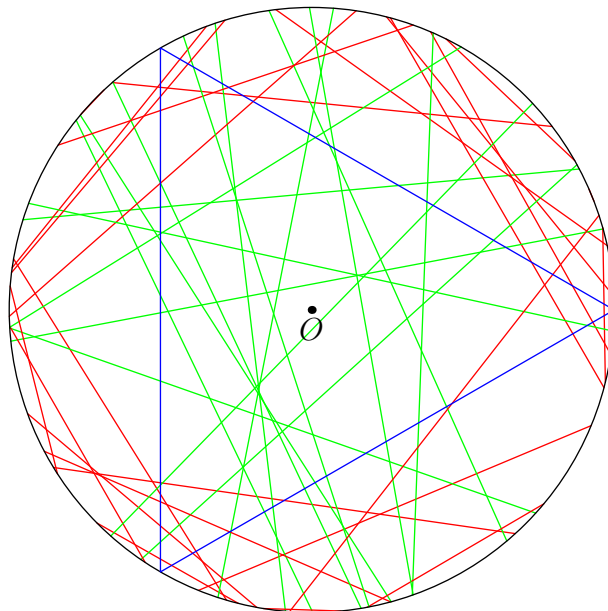


FIG. 25 – 40 cordes tracées au hasard (loi uniforme sur l'angle au centre)

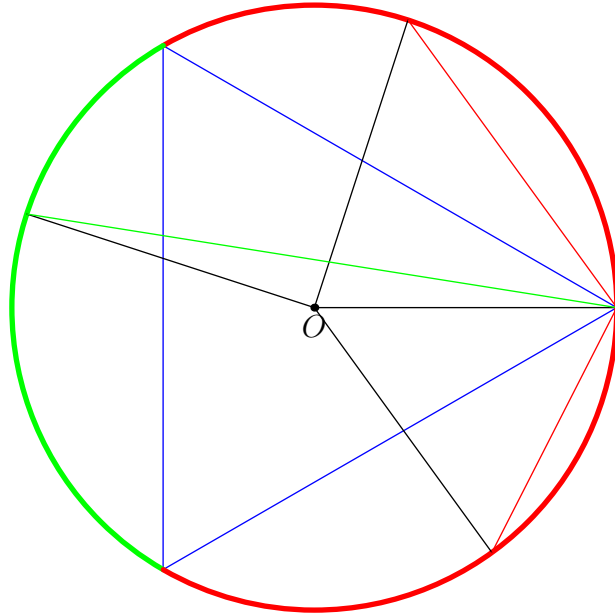


FIG. 26 – Probabilité de succès $1/3$ pour la méthode de l'angle au centre

4.3 Troisième solution

La corde AB du cercle est connue dès que l'on connaît son milieu M . On considère maintenant que ce point aléatoire M suit la loi uniforme sur le disque.

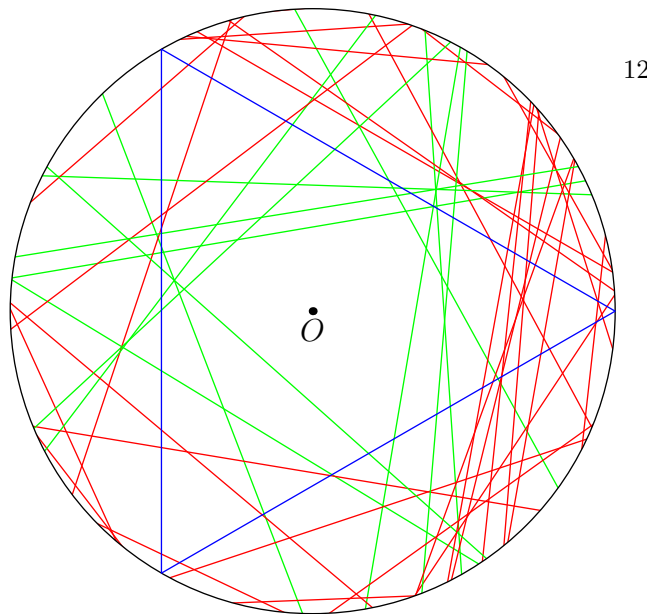


FIG. 27 – 40 cordes tracées au hasard (loi uniforme sur le disque pour le milieu de la corde)

La figure 28 représente la zone de succès pour M en vert et nous servira de preuve muette pour voir que la probabilité cherchée est maintenant $\frac{1}{4}$.

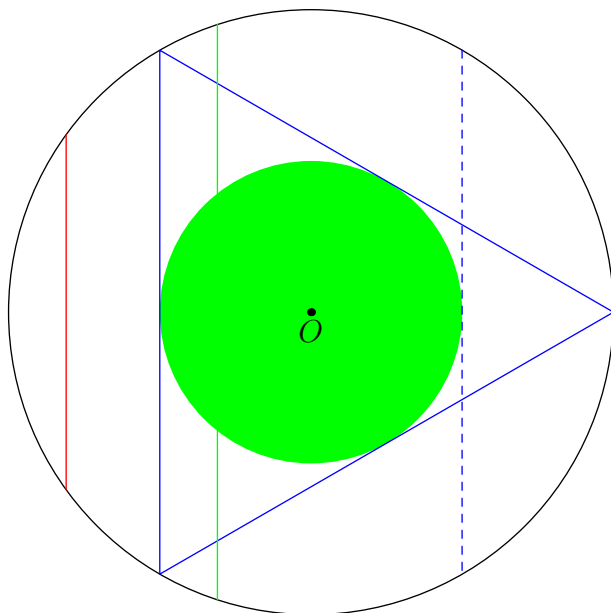


FIG. 28 – Probabilité de succès $1/4$ pour le choix du milieu de la corde par loi uniforme sur le disque

4.4 Commentaire

Mathématiquement, chacune des trois solutions esquissées ci-dessus est correcte¹⁹. La controverse entre mathématiciens suscitée par le problème de Bertrand ne portait pas sur les raisonnements amenant ces résultats différents, mais sur la question de savoir quelle était l'*hypothèse naturelle*. Le fond de l'affaire est que l'expression « tracer une corde au hasard » ne peut avoir de sens mathématique précis sans hypothèses de modélisation supplémentaires. De ce point de vue, la résolution du paradoxe de Bertrand consiste à inverser la question : quelles procédures expérimentales de tracé d'une corde, chacun de ces modèles représente-t-il correctement ?

Pour la solution 1, on peut par exemple imaginer un insecte qui se déplace à vitesse constante sur la grande aiguille d'une horloge en rebroussant chemin à chaque extrémité. À un instant quelconque, on photographie sa position, construisant ainsi le point M de la figure 22, et on trace la corde de milieu M . Pour la solution 2, on peut utiliser une roulette que l'on lance 2 fois pour construire les points A et B de la figure 24 (ou une horloge à deux aiguilles photographiée à un instant quelconque). Pour la solution 3, on peut prendre un couvercle de boîte de camembert et le placer au grand air un jour où le temps est à la pluie. On prend alors pour M le point d'impact de la première goutte de pluie tombant sur le couvercle. On pourrait multiplier les scénarios à l'infini... Pour

19. Bien sûr, il y aurait encore un peu de travail à faire pour traiter proprement l'invariance par rotation utilisée implicitement dans les solutions 1 et 2.

finir j'en suggère un nouveau qui à lui seul pourrait donner un beau sujet de travail pour étudiants de licence : on choisit le point M comme le point d'impact d'une flèche sur une cible circulaire. En général, on peut modéliser ceci par un vecteur aléatoire gaussien et cette situation est déjà suffisamment riche pour une étude qui sortirait du cadre de ce texte. Disons simplement que dans ce cas, au bout de n tentatives, on obtient un nuage de points (les centres des cordes) ayant approximativement une forme elliptique, cette forme dépendant entre autres, de l'habileté du tireur, du vent, de sa vision, etc. En jouant sur tous ces paramètres, on peut trouver n'importe quelle valeur entre 0 et 1 pour la probabilité cherchée.

Références

- [1] BERTRAND J., Calcul des Probabilités, Gauthier-Villars, Paris 1889, réédition J. Gabay 2007, ISBN 2-87647-128-0.
- [2] BARBAZO E., BOUSCASSE J.-M., POMÈS R., PURJOU J., PURJOU M., TERRACHER P.-H., L'esprit des lois continues ou quelques aspects du calcul des probabilités au lycée, IREM d'Aquitaine Talence, 2003.
- [3] SUQUET Ch., Intégration et Probabilités Élémentaires, IPÉ-Math306, 2006–2007, photocopié de cours de Licence de Mathématiques (L3) à Lille 1, téléchargeable à <http://math.uni-lille1.fr/~ipeis/>