



Marches d'ivrognes

Question préliminaire sobre

D'après le cours, si $\alpha \in \mathbb{R}$ alors l'application $u \mapsto (1+u)^\alpha$ est développable en série entière, laquelle série entière possède un rayon de convergence égal à 1 :

$$\forall u \in]-1, 1[, \quad (1+u)^\alpha = \sum_{k=0}^{+\infty} \binom{\alpha}{k} u^k.$$

Pour $t \in]-1, 1[$, on a $-t \in]-1, 1[$, donc on peut appliquer le résultat précédent :

$$\forall t \in]-1, 1[, \quad \frac{1}{\sqrt{1-t}} = (1+(-t))^{-1/2} = \sum_{k=0}^{+\infty} \binom{-1/2}{k} (-1)^k t^k.$$

La relation précédente constitue bien un développement en série entière de rayon de convergence égal à 1 (introduction des $(-1)^k$ ne changeant pas ce rayon). Il reste à simplifier les coefficients binomiaux avec le bidouillage usuel « à la Wallis » où, pour traiter un produit de termes impairs, on introduit les termes pairs qui font apparaître une factorielle. Puisque ces termes pairs peuvent eux-mêmes se combiner en une puissance de deux multipliée par une factorielle, c'est gagné :

$$\begin{aligned} \binom{-1/2}{n} &= \frac{(-\frac{1}{2})(-\frac{3}{2}) \cdots \overbrace{\left(-\frac{1}{2} - n + 1\right)}^{\frac{1-2n}{2} = -\frac{2n-1}{2}}}{n!} = (-1)^n \frac{1 \times 3 \times \cdots \times (2n-1)}{2^n n!} \\ &= (-1)^n \frac{1 \times 3 \times \cdots \times (2n-1) \times 2 \times 4 \times \cdots \times (2n)}{2^n n! (2 \times 1)(2 \times 2) \cdots (2 \times n)} = (-1)^n \frac{(2n)!}{4^n n!^2}. \end{aligned}$$

Finalement :

$$\forall t \in]-1, 1[, \quad \frac{1}{\sqrt{1-t}} = \sum_{n=0}^{+\infty} \frac{1}{4^n} \binom{2n}{n} t^n.$$

OUI, il faut que vous finissiez par savoir faire ce travail, et NON, le regarder 100 fois ne suffit en général pas...

1 Premières propriétés des X_n

1. La loi de U_k est donnée par la modélisation de la marche, qui dit que $U_k = X_k - X_{k-1}$ vaut 1 avec probabilité p et -1 avec la probabilité q .

$$U_k(\Omega) = \{-1, 1\}, \quad \mathbb{P}(U_k = 1) = p \text{ et } \mathbb{P}(U_k = -1) = q.$$

On peut ensuite calculer l'espérance et la variance de U_k via le lemme de transfert. Mais aussi en notant que $V_k = \frac{U_k + 1}{2}$ suit une loi de Bernoulli $\mathcal{B}(p)$, donc $\mathbb{E}(V_k) = p = \frac{\mathbb{E}(U_k) + 1}{2}$ et $\text{Var}(V_k) = pq = \frac{1}{4} \text{Var}(U_k)$.

$$\mathbb{E}(U_k) = 2p - 1 (= p - q) \text{ et } \text{Var}(U_k) = 4pq.$$

2. La somme $\sum_{i=1}^n U_i = \sum_{i=1}^n (X_i - X_{i-1})$ est bien entendu télescopique donc se simplifie en $X_n - X_0$;
et comme $X_0 = 0$:

$$\sum_{i=1}^n U_i = X_n.$$

On va utiliser d'une part la *linéarité* de l'espérance, et d'autre part le fait que pour des variables *mutuellement indépendantes* (ce qui est le cas ici par hypothèse), la variance de la somme est la somme des variances.

$$\mathbb{E}(X_n) = n(p - q) \text{ et } \text{Var}(X_n) = 4npq.$$

3. On a $\mathbb{P}(X_1 = 1 \text{ et } X_2 = -2) = 0$, alors que $\mathbb{P}(X_1 = 1)\mathbb{P}(X_2 = -2) = pq^2 \neq 0$, donc :

$$X_1 \text{ et } X_2 \text{ ne sont pas indépendantes.}$$

4. On a $X_k - X_\ell = \sum_{i=\ell+1}^k U_i$ et $X_\ell = \sum_{i=1}^\ell U_i$. Or $U_1, \dots, U_\ell, \dots, U_k$ sont mutuellement indépendantes. Le *lemme des coalitions* nous dit alors qu'il en va de même pour deux variables construites à l'aide d'une partition de cette famille.

$$X_k - X_\ell \text{ et } X_\ell \text{ sont indépendantes.}$$

Soient $n, m \in \mathbb{Z}$. Par définition des probabilités conditionnelles puis par indépendance de $X_k - X_\ell$ et X_ℓ , on a :

$$\begin{aligned} \mathbb{P}(X_k = n | X_\ell = m) &= \frac{\mathbb{P}(X_k = n \text{ et } X_\ell = m)}{\mathbb{P}(X_\ell = m)} = \frac{\mathbb{P}(X_k - X_\ell = n - m \text{ et } X_\ell = m)}{\mathbb{P}(X_\ell = m)} \\ &= \frac{\mathbb{P}(X_k - X_\ell = n - m) \mathbb{P}(X_\ell = m)}{\mathbb{P}(X_\ell = m)}. \end{aligned}$$

$$\mathbb{P}(X_k = n | X_\ell = m) = \mathbb{P}(X_k - X_\ell = n - m).$$

Maintenant, $X_k - X_\ell$ est la somme de $k - \ell$ variables aléatoires indépendantes suivant une loi $\mathcal{L}$ donnée, et il en va de même pour $X_{k-\ell}$, donc ces deux variables suivent les mêmes lois !

$$\text{Si } m, n \in \mathbb{Z}, \text{ alors } \mathbb{P}(X_k = n | X_\ell = m) = \mathbb{P}(X_{k-\ell} = n - m).$$

2 Probabilité d'être au bar

1. X_1 vaut 1 ou -1 ; X_2 vaut $-2, 0$ ou 2 ... Par récurrence immédiate, on montrerait que X_k est un entier de $[-k, k]$ qui a la même parité que k . Ainsi, X_{2k+1} est un entier impair, donc est non nul.

$$\mathbb{P}(X_{2k+1} = 0) = 0$$

2. On a $X_{2k} = 0$ si et seulement si les $2k$ premiers pas sont constitués de k pas à droite et k pas à gauche. Il y a $\binom{2k}{k}$ façon de choisir l'ensemble E des pas qui seront à droite, et pour chacune de ces configurations, la probabilité d'avoir $X_i = \begin{cases} 1 & \text{si } i \in E \\ -1 & \text{sinon} \end{cases}$ vaut $p^k q^k$. La probabilité de l'union disjointe de ces $\binom{2k}{k}$ événement vaut alors :

$$\mathbb{P}(X_{2k} = 0) = \binom{2k}{k} p^k q^k$$

On pouvait aussi considérer $Y_{2k} = \sum_{i=1}^{2k} \frac{U_i + 1}{2} = \frac{1}{2}X_{2k} + k$ qui est une somme de $2k$ Bernoulli indépendantes donc qui suit une loi binomiale $\mathcal{B}(2k, p)$, donc : $\mathbb{P}(X_{2k}=0) = \mathbb{P}(Y_{2k} = k) = \binom{2k}{k} p^k q^k$

3. Le raisonnement est le même : on a $X_{2k} = 2\ell$ si et seulement si on a fait d pas à droite et $2k - d$ pas à gauche, avec $d - (2k - d) = \ell$, c'est-à-dire $d = k + \ell$. Il y a $\binom{2k}{\ell + k}$ choix des positions pour placer les d pas à droite, et ainsi :

$$\text{Pour } k \in \mathbb{N} \text{ et } \ell \in \llbracket -k, k \rrbracket, \text{ on a } \mathbb{P}(X_{2k} = 2\ell) = \binom{2k}{\ell + k} p^{k+\ell} q^{k-\ell}.$$

Ici encore on pouvait utiliser Y_{2k} .

4. On s'intéresse aux $r > 0$ tels que la suite de terme général $b_n = a_n r^n$ soit bornée. Fixons donc $r > 0$ et intéressons nous à :

$$\frac{b_{n+1}}{b_n} = \dots = \frac{2(2n+1)}{n+1} pqr \xrightarrow{n \rightarrow +\infty} 4pqr.$$

Ainsi :

- Si $r < \frac{1}{4pq}$ alors $\frac{b_{n+1}}{b_n} \xrightarrow{n \rightarrow +\infty} \ell \in]0, 1[$, donc $b_n \xrightarrow{n \rightarrow +\infty} 0$ donc $(a_n r^n)_{n \in \mathbb{N}}$ est bornée.
- Si $r > \frac{1}{4pq}$ alors $\frac{b_{n+1}}{b_n} \xrightarrow{n \rightarrow +\infty} \ell > 1$, donc $|b_n| \xrightarrow{n \rightarrow +\infty} +\infty$ donc $(a_n r^n)_{n \in \mathbb{N}}$ n'est pas bornée.

Par définition, le rayon de convergence de $\sum a_n z^n$ est la borne supérieure de l'ensemble des r tels que $(a_n r^n)_{n \in \mathbb{N}}$ est bornée, donc :

$$\text{Le rayon de convergence de } \sum a_n z^n \text{ est } \frac{1}{4pq}.$$

On se souvient que $q = 1 - p$, et que la fonction $p \mapsto p(1 - p)$ a un comportement maintenant bien connu sur $[0, 1]$: elle est en particulier majorée par $1/4$ (maximum pris en $1/2$).

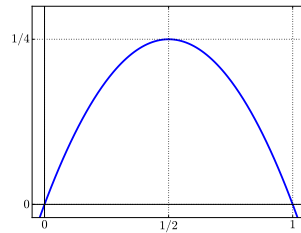


FIGURE 1 – À graver dans votre cerveau

On a donc $4pq \leq 1$, puis :

$$\text{Le rayon de convergence de } \sum a_n z^n \text{ est minoré par } 1.$$

Pour ce point on pouvait aussi se contenter de dire que $(a_n 1^n)$ est bornée (puisque $a_n 1^n$ est une probabilité) !

5. Fixons $z \in]-1, 1[$. D'après ce qui précède, $4pq \leq 1$, donc $4pqz < 1$, donc d'après la question préliminaire :

$$\frac{1}{\sqrt{1-4pqz}} = \sum_{n=0}^{+\infty} \frac{1}{4^n} \binom{2n}{n} (4pqz)^n = \sum_{n=0}^{+\infty} \underbrace{\binom{2n}{n} p^n (1-p)^n}_{\mathbb{P}(X_{2n}=0)=a_n} z^n = A(z),$$

soit comme demandé :

$$\text{Pour tout } z \in]-1, 1[, A(z) = \frac{1}{\sqrt{1-4pqz}}.$$

3 Probabilité de repasser au bar

1. Par exemple via la formule des probabilités totales et en utilisant l'indépendance de U_1 et U_2 :

$$\begin{aligned} t_1 &= \mathbb{P}(X_2 = 0) = \mathbb{P}(U_1 = 1 \text{ et } U_2 = -1) + \mathbb{P}(U_1 = -1 \text{ et } U_2 = 1) \\ &= \underbrace{\mathbb{P}(U_1 = 1)}_p \underbrace{\mathbb{P}(U_2 = -1)}_q + \underbrace{\mathbb{P}(U_1 = -1)}_q \underbrace{\mathbb{P}(U_2 = 1)}_p, \end{aligned}$$

$$\boxed{t_1 = 2pq}$$

2. L'événement (sa fonction indicatrice pour être précis) $(T = 2k)$ est fonction des variables $U_1, \dots, U_k$ alors que l'événement $(X_{2n} - X_{2k} = 0)$ est fonction des variables $U_{2k+1}, \dots, U_{2n}$. Or les variables $U_1, \dots, U_{2k}, \dots, U_{2n}$ sont mutuellement indépendantes; le lemme des coalitions nous dit que les événements (leurs fonctions indicatrices) sont donc également indépendant(e)s.

$$\boxed{\text{L'événement } (T = 2k) \text{ est indépendant de l'événement } (X_{2n} - X_{2k} = 0).}$$

3. Quand $X_{2n} = 0$, la variable aléatoire T prend une valeur dans l'ensemble $\{2, 4, \dots, 2n\}$; on peut donc partitionner l'événement $(X_{2n} = 0)$:

$$\{X_{2n} = 0\} = \bigsqcup_{k=1}^n \{X_{2n} = 0 \text{ et } T = 2k\}$$

puis écrire, sachant que $(X_{2n} = 0 \text{ et } T = 2k)$ est équivalent à $(X_{2n} - X_{2k} = 0 \text{ et } T = 2k)$, et par disjonction/indépendance :

$$a_n = \mathbb{P}(X_{2n} = 0) = \sum_{k=1}^n \mathbb{P}(X_{2n} = 0 \text{ et } T = 2k) = \sum_{k=1}^n \underbrace{\mathbb{P}(T = 2k)}_{=t_k} \underbrace{\mathbb{P}(X_{2n} - X_{2k} = 0 \mid T = 2k)}_{=\mathbb{P}(X_{2n} - X_{2k} = 0)}$$

Or $\mathbb{P}(X_{2n} - X_{2k} = 0) = \mathbb{P}(X_{2n-2k} = 0) = a_{n-k}$ (voir question 1.4), donc :

$$\boxed{a_n = \sum_{k=1}^n t_k a_{n-k}.$$

4. On a la subtile majoration $|\mathbb{P}(T = 2n)| \leq 1$ qui nous assure que $(t_n \cdot 1^n)_{n \in \mathbb{N}}$ est bornée, donc :

$$\boxed{\text{Le rayon de convergence de } \sum t_n z^n \text{ est minoré par } 1.}$$

Maintenant, on sait que le *produit de Cauchy* de deux séries entières de rayon de convergence R_1 et R_2 possède un rayon de convergence R minoré par $\min(R_1, R_2)$, et sur $] -R, R[$ la somme du produit de Cauchy... est le produit des sommes des séries. Ici :

$$\forall z \in]-1, 1[, \quad A(z)G(z) = \underbrace{a_0 t_0}_{=0} + \sum_{n=1}^{+\infty} \underbrace{\left(\sum_{k=1}^n t_k a_{n-k} \right)}_{=a_n} z^n = \sum_{n=0}^{+\infty} a_n t^n = \underbrace{a_0}_{=1},$$

soit comme annoncé :

$$\boxed{\text{Pour tout } z \in]-1, 1[, A(z) - 1 = A(z)G(z).}$$

On se souvient que $A(z) = \frac{1}{\sqrt{1-4pqz}}$, donc :

$$\boxed{\text{Pour tout } z \in]-1, 1[, G(z) = 1 - \sqrt{1-4pqz}.}$$

5. On note que dans la question préliminaire, on a obtenu un développement assez simple d'une expression proche de $G(z)$. Plus proche de $G'(z)$, en fait. C'est parti!
On a d'une part G dérivable sur $[0, \frac{1}{4pq}[$ (donc $[0, 1[$ car $pq \leq 1/4$), avec :

$$\forall z \in [0, 1[, \quad G'(z) = \frac{2pq}{\sqrt{1-4pqz}} = \sum_{k=0}^{+\infty} 2pq \frac{1}{4^k} \binom{2k}{k} (4pqz)^k = \sum_{k=0}^{+\infty} 2 \binom{2k}{k} (pq)^{k+1} z^k.$$

Mais *d'autre part*, la théorie sur les séries entières nous dit que $G'(z)$ est la somme de la série dérivée formellement :

$$\forall z \in [0, 1[, \quad G'(z) = \sum_{n=1}^{+\infty} n t_n z^{n-1} = \sum_{k=0}^{+\infty} (k+1) t_{k+1} z^k.$$

Par *unicité du développement en série entière*¹, les coefficients de ces développements sont égaux, soit encore (après décalage d'indice) :

$$\forall n \in \mathbb{N}^*, \quad t_n = \frac{2(pq)^n \binom{2n-2}{n-1}}{n} = \frac{2(pq)^n (2n-2)!}{n (n-1)!^2} = 2n(pq)^n \frac{(2n-2)!}{n!^2} = \frac{(pq)^n (2n)!}{2n-1 \cdot n!^2},$$

et finalement :

$$\text{Pour tout } n \in \mathbb{N}^*, t_n = \frac{1}{2n-1} \binom{2n}{n} p^n q^n.$$

On aurait aussi pu faire le calcul directement via $G(z) = 1 - (1-u)^{1/2} \dots$

6. G est une fonction génératrice ; donc on peut sortir l'argumentaire générique : la famille d'événements $(\{T = 2n\})_{n \in \mathbb{N}}$ étant totale, la série $\sum \mathbb{P}(T = 2n)$ est convergente (de somme égale à 1 accessoirement), donc si on note $f_n : z \mapsto t_n z^n$, on a $\|f_n\|_{\infty}^{[0,1]} = t_n$, donc la série de fonctions continues $\sum f_n$ converge normalement sur $[0, 1]$. En particulier, la somme est définie et continue sur ce segment.

$$G \text{ est définie et continue sur } [0, 1].$$

On regarde G au voisinage de 1. *D'une part*, $G(z) \xrightarrow{z \rightarrow 1^-} G(1)$ par continuité de G en 1 (question précédente). Et *d'autre part*, l'expression de $G(z)$ étant connue pour $z < 1$, on a

$$G(z) = 1 - \sqrt{1 - 4pqz} \xrightarrow{z \rightarrow 1^-} 1 - \sqrt{1 - 4pq},$$

donc *par unicité de la limite* $G(1) = 1 - \sqrt{1 - 4pq}$ (insistons sur le fait que la relation $G(z) = 1 - \sqrt{1 - 4pqz}$ n'avait PAS été prouvée pour $z = 1$).

Il reste à noter que $G(1) = \sum_{n=1}^{+\infty} t_n$, et par ailleurs :

$$1 - 4pq = 1 - 4p(1-p) = 1 - 4p + 4p^2 = (2p-1)^2.$$

On peut conclure :

$$\sum_{n=1}^{+\infty} t_n = 1 - |2p-1|.$$

7. Si on note E l'événement « l'ivrogne repassera au bar au moins une fois », alors :

$$\mathbb{P}(E) = \mathbb{P}(T > 0) = \sum_{n=1}^{+\infty} \mathbb{P}(T = 2n) = \sum_{n=1}^{+\infty} t_n = 1 - |2p-1|,$$

donc cette probabilité vaut 1 si et seulement si $2p-1 = 0$, c'est-à-dire $p = \frac{1}{2}$.

$$\text{L'ivrogne repasse au bar avec probabilité 1 si et seulement si } p = \frac{1}{2}.$$

8. Le cours nous dit que T est d'espérance finie si et seulement si sa fonction génératrice est dérivable en 1. Or ici, $G(z) = 1 - \sqrt{1-z}$, donc G n'est pas dérivable en 1.

$$T \text{ n'est pas d'espérance finie.}$$

1. Qui se prouve, rappelons le, en récupérant les coefficients comme des dérivées en 0, modulo une factorielle.

On pouvait aussi passer par l'équivalent de Stirling $n! \sim \sqrt{2\pi n} \left(\frac{n}{e}\right)^n$, qui nous dit après agitation des bras :

$$n\mathbb{P}(T = n) \sim \frac{1}{\sqrt{\pi n}},$$

donc $\sum n\mathbb{P}(T = n)$ diverge donc T n'est pas d'espérance finie.

Finalement, l'ivrogne repassera au bar, mais non seulement il ne peut pas borner le temps de premier retour (on le savait déjà), mais même *en moyenne*, le temps de retour n'est pas borné. Penser a contrario au temps moyen de premier PILE dans une suite infinie de lancers de pièces : ce temps n'est pas borné, mais il possède une moyenne (égale à 2) ; les temps d'attente longs sont donc peu probables.

4 Fréquence de passage au bar

1. Le nombre R_n de passages par le bar entre les temps 1 et $2n$ est le nombre de k tels que $Y_k = 1$: c'est donc la somme des Y_k !

$$R_n = \sum_{k=1}^n Y_k.$$

2. Les variables de Bernoulli Y_k ne sont pas du tout indépendantes, mais *l'espérance reste linéaire*. Et comme

$$\mathbb{E}(Y_k) = \mathbb{P}(X_{2k} = 0) = \frac{1}{4^k} \binom{2k}{k},$$

on a bien :

$$\mathbb{E}(R_n) = \sum_{k=1}^n \frac{1}{4^k} \binom{2k}{k}$$

3. On peut établir la formule par récurrence, mais je vous laisse apprécier ce qui suit...

Le produit de Cauchy de deux séries *absolument convergentes* est convergent, et on a donc :

$$\forall x \in]-1, 1[, \quad \frac{1}{1-x} \frac{1}{\sqrt{1-x}} = \sum_{n=0}^{+\infty} \left(\sum_{k=1}^n \frac{1}{4^k} \binom{2k}{k} \right) x^n.$$

Or par dérivation terme à terme d'une série entière dans $] -R, R[$, on a aussi

$$\frac{1}{1-x} \frac{1}{\sqrt{1-x}} = \left(\frac{2}{\sqrt{1-x}} \right)' = \sum_{n=1}^{+\infty} 2n \frac{1}{4^n} \binom{2n}{n} x^{n-1} = \sum_{n=0}^{+\infty} 2(n+1) \frac{1}{4^{n+1}} \binom{2n+2}{n+1} x^n = \sum_{n=0}^{+\infty} \frac{2n+1}{4^n} \binom{2n}{n} x^n.$$

On conclut avec l'unicité du développement en série entière :

$$\text{Pour tout } n \in \mathbb{N}^*, \quad \sum_{k=0}^n \frac{1}{4^k} \binom{2k}{k} = \frac{2n+1}{4^n} \binom{2n}{n}$$

D'après ce qui précède, $\mathbb{E}(R_n) = \frac{2n+1}{4^n} \binom{2n}{n} - 1$. Il reste à dégainer une dernière fois Stirling, pour obtenir (faites silence : je plisse les yeux et calcule de tête...) :

$$\mathbb{E}(R_n) \sim \frac{2}{\sqrt{\pi}} \sqrt{n}$$

Ainsi, $\frac{\mathbb{E}(R_n)}{n} \xrightarrow{n \rightarrow +\infty} 0$.

La fréquence de visite au bar tend vers 0.