Méthodes de calcul et de raisonnement

2021

1. En utilisant la formule des probabilités totales associée au système complet d'événements $\{[X_n = 0], [X_n = 1], [X_n = 2]\}$, on trouve $V_{n+1} = MV_n$ avec
$$M = \begin{pmatrix} 1 & 1/4 & 0 \\ 0 & 1/2 & 0 \\ 0 & 1/4 & 1 \end{pmatrix}.$$

2. En échelonnant la matrice $M - \lambda I_3$, on trouve $\text{Spec}(M) = \{\frac{1}{2}, 1\}$, et en posant
$$P = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 0 & -2 \\ 0 & 1 & 1 \end{pmatrix} \text{ et } D = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \frac{1}{2} \end{pmatrix},$$
on a bien $M = PDP^{-1}$.

3. On montre par une rapide récurrence que pour tout entier n, $M^n = PD^nP^{-1}$. Après calculs, on en déduit que
$$M^n = \begin{pmatrix} 1 & \frac{1}{2} - \frac{1}{2^{n+1}} & 0 \\ 0 & \frac{1}{2^n} & 0 \\ 0 & \frac{1}{2} - \frac{1}{2^{n+1}} & 1 \end{pmatrix}.$$

4. (a) Une rapide récurrence permet de montrer que pour tout entier n, $V_n = M^nV_0$, et donc
$$\mathbb{P}(X_n = 0) = \mathbb{P}(X_0 = 0) + \left(\frac{1}{2} - \frac{1}{2^{n+1}}\right)\mathbb{P}(X_0 = 1)$$
$$\mathbb{P}(X_n = 1) = \frac{1}{2^n}\mathbb{P}(X_0 = 1)$$
$$\mathbb{P}(X_n = 2) = \left(\frac{1}{2} - \frac{1}{2^{n+1}}\right)\mathbb{P}(X_0 = 1) + \mathbb{P}(X_0 = 2)$$
On en déduit l'espérance
$$\mathbb{E}(X_n) = \mathbb{P}(X_n = 1) + 2\mathbb{P}(X_n = 2) = \mathbb{P}(X_0 = 1) + 2\mathbb{P}(X_0 = 2) = \mathbb{E}(X_0).$$
(b) On a
$$\mathbb{P}(X_n \in \{0, 2\}) = \mathbb{P}(X_n = 0) + \mathbb{P}(X_n = 2) \to \mathbb{P}(X_0 = 0) + \mathbb{P}(X_0 = 1) + \mathbb{P}(X_0 = 2) = 1$$
car $\frac{1}{2} \in]-1, 1[$.

5. (a) On reconnait l'espérance d'une variable suivant une loi binomiale $\mathcal{B}(2N, \frac{i}{2N})$. La somme vaut donc i.
(b) En utilisant la formule des probabilités totales associée au système complet d'événements $\{[X_n =$

$i] \mid i \in [\![0, 2N]\!]\}$, on a

$$\begin{aligned}\mathbb{E}(X_{n+1}) &= \sum_{j=0}^{2N} j\mathbb{P}(X_{n+1} = j) \\ &= \sum_{j=0}^{2N} j \sum_{i=0}^{2N} \mathbb{P}_{[X_n=i]}(X_{n+1} = j)\mathbb{P}(X_n = i) \\ &= \sum_{i=0}^{2N} \mathbb{P}(X_n = i)S_i \\ &= \sum_{i=0}^{2N} i\mathbb{P}(X_n = i) \\ &= \mathbb{E}(X_n)\end{aligned}$$

(c) Ainsi, on a pour tout n : $\mathbb{E}(X_n) = \mathbb{E}(X_0)$. En moyenne, le nombre d'allèles de type A reste constant.

6. (a) On a

$$\begin{aligned}\mathbb{P}_{[X_n=k]}(X_{n+1} \in \{0, 2N\}) &= \mathbb{P}_{[X_n=k]}(X_{n+1} = 0) + \mathbb{P}_{[X_n=k]}(X_{n+1} = 2N) \\ &= \left(\frac{2N-k}{2N}\right)^{2N} + \left(\frac{k}{2N}\right)^{2N} \\ &\geqslant 2\frac{1}{(2N)^{2N}}\end{aligned}$$

(b) En utilisant la formule des probabilités totales, on a donc

$$\begin{aligned}u_{n+1} &= \sum_{k=0}^{2N} \mathbb{P}_{[X_n=k]}(X_{n+1} \in \{0, 2N\})\mathbb{P}(X_n = k) \\ &= \mathbb{P}_{[X_n=0]}(X_{n+1} \in \{0, 2N\})\mathbb{P}(X_n = 0) + \mathbb{P}_{[X_n=2N]}(X_{n+1} \in \{0, 2N\})\mathbb{P}(X_n = 2N) \\ &+ \sum_{k=1}^{2N-1} \mathbb{P}_{[X_n=k]}(X_{n+1} \in \{0, 2N\})\mathbb{P}(X_n = k) \\ &\geqslant \mathbb{P}(X_n = 0) + \mathbb{P}(X_n = 2N) + 2\frac{1}{(2N)^{2N}} \sum_{k=1}^{2N-1} \mathbb{P}(X_n = k) \\ &\geqslant u_n + 2\frac{1}{(2N)^{2N}}(1 - u_n)\end{aligned}$$

et on retrouve le résultat demandé.

(c) On reconnaît une suite arithmético-géométrique, et la méthode usuelle donne

$$\forall n \in \mathbb{N},\ (u_0 - 1)(1 - \alpha)^n + 1,$$

et on trouve donc une limite égale à 1 car $|1 - \alpha| < 1$.

(d) On pose $\alpha = \frac{2}{(2N)^{2N}}$, et on considère la suite (w_n) associée. Montrons par récurrence que $u_n \geqslant w_n$.

- C'est vrai pour $n = 0$.
- Soit $n \in \mathbb{N}$, on suppose $u_n \geqslant w_n$. Alors

$$\begin{aligned}u_{n+1} &\geqslant u_n + (1 - u_n)\alpha \\ &\geqslant \alpha + (1 - \alpha)u_n \\ &\geqslant \alpha + (1\alpha)w_n \\ &\geqslant w_{n+1}\end{aligned}$$

On a donc pour tout n : $w_n \leqslant u_n \leqslant 1$, et par théorème d'encadrement des limités, la suite (u_n) converge vers 1.

Ainsi, au bout d'un temps très long, tous les individus auront les mêmes allèles.

7. Chaque individu est du type 1 avec probabilité p_1 et les individus sont indépendants entre eux. Ainsi, N_1 suit une loi binomiale de paramètres N et p_1.

   (On note qu'en général, $N_i \hookrightarrow \mathcal{B}(N, p_i)$.)

8. On a donc $\mathbb{E}(N_1) = Np_1$ et $\mathbb{V}(N_1) = Np_1(1-p_1)$.

9. On peut appliquer directement le théorème de Moivre-Laplace :

$$\lim_{n\to+\infty} \mathbb{P}\left(a \leqslant \frac{N_1 - Np_1}{\sqrt{Np_1(1-p_1)}} \leqslant b\right) = \frac{1}{\sqrt{2\pi}} \int_a^b e^{-t^2/2}\mathrm{d}t.$$

10. (a) Par symétrie de la covariance, la matrice W est symétrique réelle, donc diagonalisable par le théorème spectral.

    (b) On a d'une part

$$\mathbb{V}(aN_1 + bN_2) = a^2\mathbb{V}(N_1) + b^2\mathbb{V}(N_2) + 2ab\operatorname{Cov}(N_1, N_2),$$

    et d'autre part

$$\begin{pmatrix} a & b \end{pmatrix} W \begin{pmatrix} a \\ b \end{pmatrix} = a^2\mathbb{V}(N_1) + b^2\mathbb{V}(N_2) + 2ab\operatorname{Cov}(N_1, N_2).$$

    On a donc bien le résultat voulu.

    (c) Soit $\lambda \in \operatorname{Spec}(W)$ et soit $X = \binom{a}{b}$ un vecteur propre associé. On a alors

$$\begin{pmatrix} a & b \end{pmatrix} W \begin{pmatrix} a \\ b \end{pmatrix} = \|X\|^2\lambda.$$

    On en tire alors, comme la variance est positive et $\|X\| \neq 0$, la positivité de λ.

    (d) Si on avait une valeur propre nulle, alors en considérant $\binom{a}{b}$ un vecteur propre associé, on aurait $\mathbb{V}(aN_1 + bN_2) = 0$.

    La variable $aN_1 + bN_2$ est donc presque sûrement constante. Or cette variable peut prendre (avec probabilité non nulle) les valeurs 0, a et b, et donc n'est pas constante.

    Ainsi, les valeurs propres de W sont strictement positives.

    (e) En notant λ et μ les valeurs propres de W, on a donc par théorème spectral $W = PD^2P^{-1}$ avec $P^{-1} = P^{\mathbf{T}}$ et $D = \begin{pmatrix} \sqrt{\lambda} & 0 \\ 0 & \sqrt{\mu} \end{pmatrix}$ qui est bien inversible, les valeurs propres étant strictement positives.

11. On note que $A^{\mathbf{T}} = PD^{-1}$, et donc

$$AWA^{\mathbf{T}} = D^{-1}P^{-1}PD^2P^{-1}PD^{-1} = I_2.$$

    Ainsi, les variances valent 1 et la covariance 0.

12. (a) On a $N_1 + N_2 = N - N_3$, et donc

$$\mathbb{V}(N_1 + N_2) = \mathbb{V}(N - N_3) = \mathbb{V}(N_3) = Np_3(1-p_3).$$

    (b) On en déduit

$$\operatorname{Cov}(N_1, N_2) = \frac{1}{2}(\mathbb{V}(N_1 + N_2) - \mathbb{V}(N_1) - \mathbb{V}(N_2)) = -Np_1p_2.$$

    (c) On utilise la formule classique d'inversion d'une matrice 2×2 pour trouver

$$W^{-1} = \frac{1}{Np_1p_2p_3} \begin{pmatrix} p_2(1-p_2) & p_1p_2 \\ p_1p_2 & p_1(1-p_1) \end{pmatrix}.$$

(d) On a donc

$$A^{\mathbf{T}}A = PD^{-2}P^{-1} = W^{-1}.$$

On a alors

$$\begin{aligned}
Y_1^2 + Y_2^2 &= (Y_1\ Y_2)\begin{pmatrix} Y_1 \\ Y_2 \end{pmatrix} \\
&= (N_1 - Np_1 \quad N_2 - Np_2)W^{-1}\begin{pmatrix} N_1 - Np_1 \\ N_2 - Np_2 \end{pmatrix} \\
&= \frac{1}{Np_1p_2p_3}(N_1 - Np_1 \quad N_2 - Np_2)\begin{pmatrix} p_2(1-p_2) & p_1p_2 \\ p_1p_2 & p_1(1-p_1) \end{pmatrix}\begin{pmatrix} N_1 - Np_1 \\ N_2 - Np_2 \end{pmatrix} \\
&= \frac{1}{Np_1p_2p_3}(N_1 - Np_1 \quad N_2 - Np_2)\begin{pmatrix} p_2(1-p_2)(N_1 - Np_1) + p_1p_2(N_2 - Np_2) \\ p_1p_2(N_1 - Np_1) + p_1(1-p_1)(N_2 - Np_2) \end{pmatrix} \\
&= \frac{1}{Np_1p_2p_3}\big(p_2(1-p_2)(N_1 - Np_1)^2 + 2p_1p_2(N_1 - Np_1)(N_2 - Np_2) \\
&+ p_1(1-p_1)(N_2 - Np_2)^2\big) \\
&= \frac{1}{Np_1p_2p_3}\big(p_2(p_1+p_3)(N_1 - Np_1)^2 + 2p_1p_2(N_1 - Np_1)(N_2 - Np_2) \\
&+ p_1(p_2+p_3)(N_2 - Np_2)^2\big) \\
&= \frac{(N_1 - Np_1)^2}{Np_1} + \frac{(N_2 - Np_2)^2}{Np_2} \\
&+ \frac{1}{Np_3}((N_1 - Np_1)^2 + 2(N_1 - Np_1)(N_2 - Np_2) + (N_2 - Np_2)^2) \\
&= \frac{(N_1 - Np_1)^2}{Np_1} + \frac{(N_2 - Np_2)^2}{Np_2} + \frac{(N_3 - Np_3)^2}{Np_3}
\end{aligned}$$

13. Il est clair que Z^2 est une variable positive presque sûrement, donc sa fonction de répartition est nulle sur $\mathbb{R}_-$. Soit donc $t > 0$. On a

$$\begin{aligned}
\mathbb{P}(Z^2 \leqslant t) &= \mathbb{P}(-\sqrt{t} \leqslant Z \leqslant \sqrt{t}) \\
&= \frac{1}{\sqrt{2\pi}}\int_{-\sqrt{t}}^{\sqrt{t}} e^{-t^2/2}\mathrm{d}t \\
&= 2\Phi(\sqrt{t}) - 1
\end{aligned}$$

La fonction de répartition est donc continue (car Φ tend vers $\frac{1}{2}$ en 0) et de classe $\mathcal{C}^1$ sauf peut-être en 0. La variable Z^2 est donc à densité, et une densité est donnée par

$$F_{Z^2}\colon t \mapsto \frac{\mathrm{d}}{\mathrm{d}t}\mathbb{P}(Z^2 \leqslant t) = \frac{1}{\sqrt{2\pi t}}e^{-t/2}\mathbb{1}_{\mathbb{R}_+^*}(t).$$

14. On a par formule de Koenig Huygens

$$\mathbb{E}(Z^2) = \mathbb{V}(Z) + \mathbb{E}(Z)^2 = 1.$$

De plus, sous réserve de convergence, on a

$$\mathbb{E}(Z^4) = \int_{-\infty}^{+\infty} \frac{t^4}{\sqrt{2\pi}}e^{-\frac{t^2}{2}}\mathrm{d}t = 2\int_0^{+\infty} \frac{t^4}{\sqrt{2\pi}}e^{-\frac{t^2}{2}}\mathrm{d}t.$$

Soit alors $A > 0$. On définit les fonctions

$$u\colon t \mapsto \frac{t^3}{\sqrt{2\pi}} \text{ et } v\colon t \mapsto -e^{-\frac{t^2}{2}}.$$

u et v sont de classe $\mathcal{C}^2$ sur $[0, A]$, et par intégration par parties,

$$\begin{aligned}
\int_0^A \frac{t^4}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} \mathrm{d}t &= \int_0^A u(t)v'(t)\mathrm{d}t \\
&= [u(t)v(t)]_0^A - \int_0^A u'(t)v(t)\mathrm{d}t \\
&= \int_0^A \frac{3t^2}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} \mathrm{d}t \\
&\to \frac{3}{2}\mathbb{E}(Z^2) = \frac{3}{2}
\end{aligned}$$

Finalement, Z^2 admet un moment d'ordre 2, qui vaut 3.

On a alors par formule de König-Huygens

$$\mathbb{V}(Z^2) = \mathbb{E}(Z^4) - \mathbb{E}(Z^2)^2 = 2.$$

15. La changement de variable $t = xu$ est bien de classe $\mathcal{C}^1$, et on a alors

$$h(x) = \int_0^1 \frac{x\mathrm{d}u}{\sqrt{xu(x - xu)}} = \int_0^1 \frac{\mathrm{d}u}{\sqrt{u(1-u)}}$$

qui ne dépend pas de x.

16. (a) Les variables sont bien indépendantes par lemme de coalition, donc une densité de $Z_1^2 + Z_2^2$ est donnée par le produit de convolution g de f_{Z^2} par elle-même.

Il est clair que g est nulle sur $\mathbb{R}_-$. Soit alors $x > 0$. On a

$$\begin{aligned}
g(x) &= \int_{-\infty}^{+\infty} f_{Z^2}(x-t) f_{Z^2}(t)\mathrm{d}t \\
&= \int_0^x \frac{1}{2\pi} e^{-\frac{x}{2}} \frac{1}{\sqrt{t(x-t)}}\mathrm{d}t \\
&= \frac{e^{-\frac{x}{2}}}{2\pi} C
\end{aligned}$$

17. On sait que l'intégrale de g sur R converge et vaut 1. On a

$$\begin{aligned}
\int_{-\infty}^{+\infty} g(x)\mathrm{d}x &= \frac{C}{2\pi} \int_0^{+\infty} e^{-\frac{x}{2}} \mathrm{d}x \\
&= \frac{C}{\pi}
\end{aligned}$$

On a donc $C = \pi$, et on reconnaît que g est la densité d'une loi exponentielle de paramètre $\frac{1}{2}$.

On a donc

$$\mathbb{E}(Z_1^2 + Z_2^2) = 2 \text{ et } \mathbb{V}(Z_1^2 + Z_2^2) = 4.$$