

Corrigé du Contrôle continu de MB61 n° 1

1. L'évènement $[X_1 \in Y]$ est la réunion disjointe des évènements $[X_1 = c]$ et $[X_1 = t]$. Ainsi,

$$\mathbb{P}(X_1 \in Y) = \mathbb{P}(X_1 = c) + \mathbb{P}(X_1 = t) = 3/5.$$

2. La variable aléatoire $W_n(Y)$ suit une loi de Bernoulli dont le paramètre est donné par

$$\mathbb{P}(W_n(Y) = 1) = \mathbb{P}(X_n \in Y) = 3/5.$$

3. Montrons que les variables aléatoires $W_1(Y)$ et $W_2(Y)$ sont indépendantes. Considérons l'évènement $[W_1(Y) = 0, W_2(Y) = 0]$. Il se réécrit comme $[X_1 \notin Y, X_2 \notin Y]$. Puisque les variables aléatoires X_1 et X_2 sont indépendantes, il vient alors

$$\mathbb{P}(X_1 \notin Y, X_2 \notin Y) = \mathbb{P}(X_1 \notin Y)\mathbb{P}(X_2 \notin Y),$$

ce qui prouve que

$$\mathbb{P}(W_1(Y) = 0, W_2(Y) = 0) = \mathbb{P}(W_1(Y) = 0)\mathbb{P}(W_2(Y) = 0).$$

Le même raisonnement est valable pour les trois autres évènements $[W_1(Y) = 0, W_2(Y) = 1]$, $[W_1(Y) = 1, W_2(Y) = 0]$ et $[W_1(Y) = 1, W_2(Y) = 1]$. Ceci prouve que les variables aléatoires $W_1(Y)$ et $W_2(Y)$ sont indépendantes.

4. Les variables aléatoires $N_n(Y)$ et $R_n(Y)$ représentent respectivement le nombre et la proportion de pyrimidines dans une séquence de longueur n .

5. Comme l'espérance est linéaire, nous avons

$$\mathbb{E}(R_n(Y)) = \frac{1}{n}\mathbb{E}(N_n(Y)).$$

Toujours par linéarité de l'espérance, nous avons

$$\mathbb{E}(N_n(Y)) = \mathbb{E}\left(\sum_{k=1}^n W_k(Y)\right) = \sum_{k=1}^n \mathbb{E}(W_k(Y)) = \frac{3}{5}n.$$

Ainsi,

$$\mathbb{E}(R_n(Y)) = \frac{3}{5}.$$

Nous avons

$$\text{Var}(R_n(Y)) = \frac{1}{n^2}\text{Var}(N_n(Y)),$$

et comme les variables aléatoires $(W_n(Y))_{n \geq 1}$ sont indépendantes, nous avons également

$$\text{Var}\left(\sum_{k=1}^n W_k(Y)\right) = \sum_{k=1}^n \text{Var}(W_k(Y)) = n \frac{3}{5} \left(1 - \frac{3}{5}\right) = 6n/25.$$

Ainsi,

$$\text{Var}(R_n(Y)) = \frac{6}{25n}.$$

6. Soit $\varepsilon > 0$, montrons que

$$\mathbb{P}(|R_n(Y) - 3/5| \geq \varepsilon) \rightarrow 0,$$

quand n tend vers l'infini. D'après l'inégalité de Bienaimé-Tchebychev, nous avons

$$\mathbb{P}(|R_n(Y) - 3/5| \geq \varepsilon) = \mathbb{P}(|R_n(Y) - \mathbb{E}(R_n(Y))| \geq \varepsilon) \leq \frac{\text{Var}(R_n(Y))}{\varepsilon^2} = \frac{6}{25\varepsilon^2 n} \rightarrow 0.$$

7. Si $n = 1$, l'évènement $[T_Y = 1]$ s'écrit $[X_1 \in Y]$, ainsi

$$\mathbb{P}(T_Y = 1) = \mathbb{P}(X_1 \in Y) = \frac{3}{5} = \frac{3}{5} \left(\frac{2}{5}\right)^{1-1}.$$

Si $n > 1$, l'évènement $[T_Y = n]$ s'écrit

$$[X_1 \notin Y, X_2 \notin Y, \dots, X_{n-1} \notin Y, X_n \in Y].$$

Comme les variables aléatoires $X_1, \dots, X_n$ sont indépendantes, il vient

$$\mathbb{P}(T_Y = n) = \mathbb{P}(X_1 \notin Y) \mathbb{P}(X_2 \notin Y) \dots \mathbb{P}(X_{n-1} \notin Y) \mathbb{P}(X_n \in Y) = \underbrace{\left(1 - \frac{3}{5}\right) \dots \left(1 - \frac{3}{5}\right)}_{n-1 \text{ fois}} \frac{3}{5}.$$

On a bien

$$\mathbb{P}(T_Y = n) = \frac{3}{5} \left(\frac{2}{5}\right)^{n-1}.$$

8. En utilisant la remarque de l'énoncé et le fait que les variables aléatoires $X_1, \dots, X_n$ sont indépendantes, il vient

$$\mathbb{P}(T_Y \geq n+1) = \mathbb{P}(X_1 \notin Y) \dots \mathbb{P}(X_n \notin Y) = \left(\frac{2}{5}\right)^n.$$

9. Par définition, nous avons

$$\mathbb{P}(T_Y \geq m+n+1 \mid T_Y \geq n+1) = \frac{\mathbb{P}(T_Y \geq m+n+1, T_Y \geq n+1)}{\mathbb{P}(T_Y \geq n+1)}.$$

Comme

$$[T_Y \geq m+n+1, T_Y \geq n+1] = [T_Y \geq m+n+1],$$

il vient alors

$$\mathbb{P}(T_Y \geq m+n+1 \mid T_Y \geq n+1) = \frac{\mathbb{P}(T_Y \geq m+n+1)}{\mathbb{P}(T_Y \geq n+1)} = \frac{\left(\frac{2}{5}\right)^{n+m}}{\left(\frac{2}{5}\right)^n} = \left(\frac{2}{5}\right)^m.$$

10. La variable aléatoire $W_n(\mathbf{aa})$ suit une loi de Bernoulli dont le paramètre est donné par

$$\mathbb{P}(W_n(\mathbf{aa}) = 1) = \mathbb{P}(X_n = \mathbf{a}, X_{n+1} = \mathbf{a}) = 1/25.$$

11. Les variables aléatoires $W_1(\mathbf{aa})$ et $W_2(\mathbf{aa})$ ne sont pas indépendantes. En effet, nous avons d'une part

$$\mathbb{P}(W_1(\mathbf{aa}) = 1, W_2(\mathbf{aa}) = 1) = \mathbb{P}(X_1 = \mathbf{a}, X_2 = \mathbf{a}) = 1/25,$$

et d'autre part

$$\mathbb{P}(W_1(\mathbf{aa}) = 1)\mathbb{P}(W_2(\mathbf{aa}) = 1) = 1/625.$$

12. Comme l'espérance est linéaire, nous avons

$$\mathbb{E}(R_n(\mathbf{aa})) = \frac{1}{n} \mathbb{E}(N_n(\mathbf{aa})).$$

Toujours par linéarité de l'espérance, nous avons

$$\mathbb{E}(N_n(\mathbf{aa})) = \mathbb{E}\left(\sum_{k=1}^n W_k(\mathbf{aa})\right) = \sum_{k=1}^n \mathbb{E}(W_k(\mathbf{aa})) = (n-1)/25.$$

Ainsi,

$$\mathbb{E}(R_n(Y)) = 1/25.$$

13. Comme $W_1^2(\mathbf{aa}) = W_1(\mathbf{aa})$, il vient

$$\mathbb{E}(W_1^2(\mathbf{aa})) = 1/25.$$

Remarquons que $W_1(\mathbf{aa})W_2(\mathbf{aa})$ ne prend que les valeurs 0 et 1, c'est donc une variable de Bernoulli et son paramètre est donné par

$$\mathbb{P}(W_1(\mathbf{aa})W_2(\mathbf{aa}) = 1) = \mathbb{P}(W_1(\mathbf{aa}) = 1, W_2(\mathbf{aa}) = 1) = 1/25.$$

Ainsi,

$$\mathbb{E}(W_1(\mathbf{aa})W_2(\mathbf{aa})) = 1/25.$$

Remarquons que $W_1(\mathbf{aa})$ et $W_3(\mathbf{aa})$ sont indépendantes. Ainsi,

$$\mathbb{E}(W_1(\mathbf{aa})W_3(\mathbf{aa})) = \mathbb{E}(W_1(\mathbf{aa}))\mathbb{E}(W_3(\mathbf{aa})) = 1/625.$$

14. Nous avons

$$\text{Var}(R_n(\mathbf{aa})) = \frac{1}{n^2} \text{Var}(N_n(\mathbf{aa})).$$

Maintenant, remarquons que

$$[N_n(\mathbf{aa})]^2 = \left(\sum_{i=1}^{n-1} W_i(\mathbf{aa})\right)^2 = \sum_{i=1}^{n-1} [W_i(\mathbf{aa})]^2 + 2 \sum_{1 \leq i < j \leq n-1} W_i(\mathbf{aa})W_j(\mathbf{aa}).$$

Ainsi,

$$\mathbb{E}([N_n(\mathbf{aa})]^2) = \sum_{i=1}^{n-1} \mathbb{E}([W_i(\mathbf{aa})]^2) + 2 \sum_{1 \leq i < j \leq n-1} \mathbb{E}(W_i(\mathbf{aa})W_j(\mathbf{aa})).$$

Comme

$$\begin{aligned} & \sum_{1 \leq i < j \leq n-1} \mathbb{E}(W_i(\mathbf{aa})W_j(\mathbf{aa})) \\ &= \sum_{i=1}^{n-2} \mathbb{E}(W_i(\mathbf{aa})W_{i+1}(\mathbf{aa})) + \sum_{1 \leq i+1 < j \leq n-1} \mathbb{E}(W_i(\mathbf{aa})W_j(\mathbf{aa})) \\ &= (n-2)\mathbb{E}(W_1(\mathbf{aa})W_2(\mathbf{aa})) + \frac{(n-3)(n-2)}{2} \mathbb{E}(W_1(\mathbf{aa})W_3(\mathbf{aa})) \\ &= (n-2)/125 + \frac{(n-3)(n-2)}{2 \times 625}, \end{aligned}$$

il vient

$$\mathbb{E}([N_n(\mathbf{aa})]^2) = (n-1)/25 + 2(n-2)/125 + (n-3)(n-2)/625.$$

Or,

$$\mathbb{E}([N_n(\mathbf{aa})])^2 = (n-1)^2/625,$$

ce qui entraîne

$$\text{Var}(N_n(\mathbf{aa})) = (n-1)/25 + 2(n-2)/125 + (n-3)(n-2)/625 - (n-1)^2/625$$

et au final

$$\text{Var}(R_n(\mathbf{aa})) = \frac{1}{25(n-1)} + \frac{2(n-2)}{125(n-1)^2} + \frac{(n-3)(n-2)}{625(n-1)^2} - \frac{1}{625}.$$

15. En remarquant que $\text{Var}(R_n(\mathbf{aa})) \rightarrow 0$ et en appliquant Bienaimé-Tchebychev, on établit la convergence en probabilité.