Feuille d'exercices 27.
Corrigé de quelques exercices

Exercice 27.7 :

1°) En modélisant la situation à l'aide de variables aléatoires, l'indice I de l'urne choisie vérifie $I \sim \mathcal{B}(\frac{1}{2})$, et la couleur C de la boule tirée (en convenant que $C = 1$ si la boule tirée est noire et que $C = 0$ sinon) vérifie : la loi de "C sachant que $I = 1$" est $\mathcal{B}(\dfrac{n_1}{n_1 + b_1})$ et la loi de "C sachant que $I = 2$" est $\mathcal{B}(\dfrac{n_2}{n_2 + b_2})$.
La formule des probabilités totales donne :

$$\begin{aligned} P(C = 1) &= P(C = 1|I = 1)P(I = 1) + P(C = 1|I = 2)P(I = 2) \\ &= \frac{1}{2}\left(\frac{n_1}{n_1 + b_1} + \frac{n_2}{n_2 + b_2}\right). \end{aligned}$$

2°) Pour répondre à la deuxième question on applique la formule de Bayes :

$$\begin{aligned} P(I = i \mid C = 1) &= \frac{P(I = i)P(C = 1 \mid I = i)}{P(I = 1)P(C = 1 \mid I = 1) + P(I = 2)P(C = 1 \mid I = 2)} \\ &= \frac{\dfrac{n_i}{n_i + b_i}}{\dfrac{n_1}{n_1 + b_1} + \dfrac{n_2}{n_2 + b_2}}. \end{aligned}$$

Exercice 27.8 :

$P(X_1 = k|X_1 + X_2 = n) = \dfrac{P(X_1 = k, X_2 = n - k)}{P(X_1 + X_2 = n)}$, or X_1 et X_2 sont indépendants, et on sait que $X_1 + X_2$ suit une loi binomiale de paramètres $(n_1 + n_2, p)$, donc pour $k \in [\max(0, n - n_2), \min(n, n_1)]$ (sinon la probabilité est nulle),

$$P(X_1 = k|X_1 + X_2 = n) = \frac{\binom{n_1}{k} p^k (1-p)^{n_1-k} \binom{n_2}{n-k} p^{n-k}(1-p)^{n_2-n+k}}{\binom{n_1+n_2}{n} p^n (1-p)^{n_1+n_2-n}},$$

donc $P(X_1 = k|X_1 + X_2 = n) = \dfrac{\binom{n_1}{k}\binom{n_2}{n-k}}{\binom{n_1+n_2}{n}}$.

Il s'agit d'une loi hypergéométrique, mais l'appellation est hors programme. Cela correspond, pour un tirage sans remise de n boules dans une urne de $n_1 + n_2$ boules dont

n_1 sont de couleur 1 et n_2 de couleur 2 (ici il s'agit des n "boules" qui ont réussi le test 1 ou le test 2, chacun de proba p), à la probabilité de tirer k boules de couleur 1.

Exercice 27.9 :

1°) $(Y_3 \leq k) = (X_1 \leq k) \cap (X_2 \leq k) \cap (X_3 \leq k)$, or les variables aléatoires sont indépendantes, et $P(X_i \leq k) = \frac{k}{n}$, donc $P(Y_3 \leq k) = (\frac{k}{n})^3$ (également vrai pour $k = 0$).

Pour $k \in \{1, \dots, n\}$, $P(Y_3 = k) = P(Y_3 \leq k) - P(Y_3 \leq k-1) = (\dfrac{k}{n})^3 - (\dfrac{k-1}{n})^3$.

2°) $(Y_1 \geq k) = (X_1, X_2, X_3 \geq k)$, or les variables aléatoires sont indépendantes, donc $P(Y_1 \geq k) = (\dfrac{n-k+1}{n})^3$ (également vrai pour $k = n+1$).

Pour $k \in \{1, \dots, n\}$, $P(Y_1 = k) = P(Y_1 \geq k) - P(Y_1 \geq k+1) = (\dfrac{n-k+1}{n})^3 - (\dfrac{n-k}{n})^3$.

3°) a) Posons $T_i = \mathbf{1}_{X_i \leq k}$: T_i est un test de Bernoulli, donc $T_i \sim B(\frac{k}{n})$. De plus, T_1, T_2, T_3 sont indépendants, donc d'après le cours, $Z_k = T_1 + T_2 + T_3 \sim B(3, \frac{k}{n})$.

Ainsi, pour tout $h \in \{0, \dots, 3\}$, $P(Z_k = h) = \dbinom{3}{h} (\frac{k}{n})^h (1 - \frac{k}{n})^{3-h}$.

b) $[Z_k \geq 2]$ correspond à l'événement ""deux ou trois indices i sont tels que $X_i \leq k$", donc $[Z_k \geq 2] = [Y_2 \leq k]$. Ainsi, $P(Y_2 \leq k) = 3(\dfrac{k}{n})^2 . \dfrac{n-k}{n} + (\dfrac{k}{n})^3$.

Exercice 27.10 :

1°) a)

◇ Si $X_n = 0$, à l'instant n, les deux boules noires sont dans l'urne U et les deux boules blanches dans V, donc à l'instant $n+1$, de manière certaine, les urnes U et V contiendront toutes deux une boule blanche et une boule noire.

Ainsi, lorsque $P(X_n = 0) > 0$, $P(X_{n+1} = i | X_n = 0)$ vaut 0 si $i \neq 1$ et vaut 1 si $i = 1$.

On peut mener un raisonnement similaire lorsque $X_n = 2$, donc $P(X_{n+1} = i | X_n = 2)$ vaut 0 si $i \neq 1$ et vaut 1 si $i = 1$.

◇ Si $X_n = 1$, les urnes U et V contiennent chacune une boule blanche et une boule noire. Pour obtenir $X_n = 2$, il est nécessaire de tirer la boule noire dans l'urne U (avec probabilité $\frac{1}{2}$) et la boule blanche dans l'urne V (aussi avec probabilité $\frac{1}{2}$). Les deux tirages étant indépendants, $P(X_{n+1} = 2 | X_n = 1) = \frac{1}{4}$.

De même, $P(X_{n+1} = 0 | X_n = 1) = \frac{1}{4}$.

De plus, lorsque $P(X_n = 1) > 0$,

$$1 = P(X_{n+1} \in \{0, 1, 2\} | X_n = 1) = \sum_{j=0}^{2} P(X_{n+1} = j | X_n = 1),$$

donc $P(X_{n+1} = 1 | X_n = 1) = \frac{1}{2}$.

b) Soit $i \in \{1, 2, 3\}$:

$[C_{n+1}]_i = P(X_{n+1} = i-1) = \displaystyle\sum_{j=0}^{2} P(X_{n+1} = i-1, X_n = j)$ d'après la formule des probabilités totales. Lorsque $P(X_n = j) > 0$,

$P(X_{n+1}=i-1, X_n=j) = P(X_{n+1}=i-1|X_n=j)P(X_n=j) = M_{i,j+1}P(X_n=j)$.
Lorsque $P(X_n=j)=0$, on a encore $P(X_{n+1}=i-1, X_n=j) = 0 = M_{i,j+1}P(X_n=j)$, donc $[C_{n+1}]_i = \sum_{j=0}^{2} M_{i,j+1}P(X_n=j) = \sum_{j=1}^{3} M_{i,j}[C_n]_j$, donc $C_{n+1} = MC_n$.
Par récurrence, on montre alors facilement que $C_n = M^n C_0$.

2°) Pour $\lambda \in \mathbb{R}$, $\chi_M(\lambda) = \begin{vmatrix} \lambda & -1/4 & 0 \\ -1 & \lambda - 1/2 & -1 \\ 0 & -1/4 & \lambda \end{vmatrix}$, donc d'après la formule de Sarrus, $\chi_M(\lambda) = \lambda^3 - \frac{\lambda^2}{2} - \frac{\lambda}{4} - \frac{\lambda}{4} = \lambda(\lambda-1)(\lambda+\frac{1}{2})$.
Soit $n \geq 1$: la division euclidienne de X^n par $\chi_M(X)$ s'écrit

$$(1) \ : \ X^n = \chi_M(X)Q(X) + a_nX^2 + b_nX + c_n.$$

D'après le théorème de Cayley-Hamilton, $M^n = a_nM^2 + b_nM + c_nI_3$.
En remplaçant dans (1) X par les valeurs propres de M, on obtient : $0 = c_n$, $1 = a_n + b_n$, $(-\frac{1}{2})^n = \frac{a_n}{4} - \frac{b_n}{2} = \frac{3}{4}a_n - \frac{1}{2}$, donc $c_n = 0$, $a_n = \frac{2}{3} + \frac{4}{3}(-\frac{1}{2})^n$, $b_n = \frac{1}{3} - \frac{4}{3}(-\frac{1}{2})^n$.
Or $C_n = M^n \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$, donc C^n est la troisième colonne de $M^n = a_nM^2 + b_nM$, mais vue la dernière colonne de M, la troisième colonne de M^2 est égale à la seconde colonne de M, donc $C_n = [\frac{2}{3} + \frac{4}{3}(-\frac{1}{2})^n] \begin{pmatrix} 1/4 \\ 1/2 \\ 1/4 \end{pmatrix} + [\frac{1}{3} - \frac{4}{3}(-\frac{1}{2})^n] \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$.
On en déduit que $P(X_n=0) = P(X_n=2) = \frac{1}{6} + \frac{1}{3}(-\frac{1}{2})^n \underset{n\to+\infty}{\longrightarrow} \dfrac{1}{6}$
et $P(X_n=1) = \frac{2}{3} + \frac{4}{3}(-\frac{1}{2})^{n+1} \underset{n\to+\infty}{\longrightarrow} \dfrac{2}{3}$.

Remarque. c'est une variante des urnes d'Ehrenfest. Si l'on remplace 2 boules par N boules, le modèle est plus complexe, mais il permet de modéliser la notion d'irréversibilité.

Exercice 27.11 :

1°) $P(S_n=0, Z_n=0) \leq P(n = S_n + Z_n = 0) = 0$, mais
$P(S_n=0)P(Z_n=0) = P(S_n=0)P(S_n=n) = (1-p)^np^n \neq 0$, donc S_n et Z_n ne sont pas indépendantes.

2°) Pour tout $s \in \mathbb{N}$, $(S=s) = \bigcup_{n\in\mathbb{N}} (S=s, N=n) = \bigcup_{n\in\mathbb{N}} (S_n=s, N=n)$.

On note Ω l'univers de définition de toutes les variables aléatoires de l'énoncé et $\mathcal{F}$ la tribu utilisée sur Ω. Si l'on admet que S_n est une variable aléatoire (comme somme finie de variables aléatoires), alors $(S_n=s)\cap(N=n) \in \mathcal{F}$, donc $(S=s) \in \mathcal{F}$, ce qui prouve que S est bien une variable aléatoire. Alors $Z = N - S$ est aussi une variable aléatoire.

3°) Soit $k, l \in \mathbb{N}$. Soit $\omega \in \Omega$:

$$\begin{aligned} \omega \in (S = k, Z = l) &\iff S(\omega) = k, S(\omega) + Z(\omega) = k + l \\ &\iff N(\omega) = k + l, S_{k+l}(\omega) = k, \text{ donc} \end{aligned}$$

$P(S = k, Z = l) = P(N = k + l, S_{k+l} = k) = P(N = k + l)P(S_{k+l} = k)$,
car N et S_{k+l} sont indépendantes,
donc $P(S = k, Z = l) = e^{-\lambda} \dfrac{\lambda^{k+l}}{(k + l)!} \begin{pmatrix} k + l \\ k \end{pmatrix} p^k (1 - p)^l = e^{-\lambda} \dfrac{(\lambda p)^k}{k!} \dfrac{(\lambda(1 - p))^l}{l!}$.
Ainsi, il existe deux suites (a_k) et (b_l) dans $\mathbb{R}_+$ telles que, pour tout $k, l \in \mathbb{N}$,
$P(S = k, Z = l) = a_k b_l$, avec $a_k = e^{-p\lambda} \dfrac{(\lambda p)^k}{k!}$ et $b_l = e^{(1-p)\lambda} \dfrac{(\lambda(1 - p))^l}{l!}$.
Alors $P(S = k) = \displaystyle\sum_{l=0}^{+\infty} P(S = k, Z = l) = a_k \sum_{l=0}^{+\infty} b_l = a_k$ et de même $P(Z = l) = b_l$.
Alors $P(Z = l)P(S = k) = a_k b_l = P(Z = l, S = k)$, ce qui prouve que S et Z sont indépendantes.
On a $P(S = k) = a_k$, donc $S \sim \mathcal{P}(p\lambda)$ et $Z \sim \mathcal{P}((1 - p)\lambda)$.

Espérance

Exercice 27.12 :

Pour tout $i \in \mathbb{N}_n$, notons B_i l'événement “la boule numéro i est marquée”.
Pour tout $j \in \mathbb{N}_n$, notons $T_{i,j}$ une variable aléatoire de Bernoulli qui teste si, lors du j-ème tirage, c'est la boule i qui est tirée.
Pour tout $i, j \in \mathbb{N}_n$, $T_{i,j} \sim \mathcal{B}(\frac{1}{n})$.
Soit $i \in \mathbb{N}_n$. Les événements $(T_{i,j})_{1 \leq j \leq n}$ sont mutuellement indépendants, donc en posant $Y_i = \displaystyle\sum_{j=1}^{n} T_{i,j}$, Y_i représente le nombre de fois que la boule i est marquée, et d'après le cours, $Y_i \sim \mathcal{B}(n, \frac{1}{n})$.
Alors, $P(B_i) = P(Y_i \geq 1) = 1 - P(Y_i \leq 0) = 1 - (1 - \frac{1}{n})^n$.
Posons $X_i = 1_{B_i}$. Ainsi, $X_i \sim \mathcal{B}(1 - (1 - \frac{1}{n})^n)$.
Le nombre de boules marquées est $X = \displaystyle\sum_{i=1}^{n} X_i$.
Il faut comprendre que l'énoncé s'intéresse à l'espérance de X. Par linéarité de l'espérance,
$E(X) = \displaystyle\sum_{i=1}^{n} E(X_i) = n(1 - (1 - \frac{1}{n})^n)$, or $(1 - \frac{1}{n})^n = e^{n \ln(1 - \frac{1}{n})} = e^{n(-\frac{1}{n} + o(\frac{1}{n}))} \xrightarrow[n \to +\infty]{} \dfrac{1}{e}$,
donc $E(X) \sim n(1 - \frac{1}{e})$.

Exercice 27.13 :

1°) Pour la k-ième personne P_k, on note $X_{n,k}$ la variable aléatoire égale au nombre de lettres qu'elle reçoit.
$X_{n,k}$ peut être considérée comme la somme de $(n-1)$ variables de Bernoulli indépendantes associées à chaque personne de la communauté (valant 1 si cette personne envoie la lettre à P_k, 0 sinon).
Par suite, $X_{n,k}$ suit la loi binomiale $\mathcal{B}(n-1, \frac{1}{n-1})$, donc pour tout $j \in \{0, \ldots, n-1\}$, $p_{j,n} = P(X_{n,k} = j) = \begin{pmatrix} n-1 \\ j \end{pmatrix} (\frac{1}{n-1})^j (1 - \frac{1}{n-1})^{n-1-j}$.
À j fixé, quand $n \to \infty$,
$p_{j,n} = \frac{(n-1)(n-2) \times (n-j)}{(n-1)(n-1) \times (n-1)} \frac{1}{j!} e^{(n-1)\ln(1-\frac{1}{n-1})} (1 - \frac{1}{n-1}))^{-j} \sim \frac{e^{-1}}{j!}$, ce qui correspond à une loi de poisson de paramètre 1.
Nous avons redémontré un résultat classique d'approximation d'une loi binomiale par une loi de Poisson, valable ici car $(n-1)\frac{1}{n-1} = 1 \underset{n \to +\infty}{\longrightarrow} 1$.

2°) On note D_j : "au moins une personne reçoit j lettres". On a $D_j = \bigcup_{k=1}^{n} \{X_{n,k} = j\}$.
Pour $j > n/2$, on sait que cette union est disjointe puisque seule une personne au plus du groupe peut être dans ce cas.
D'où $P(D_j) = nP(X_{n,k} = j) = n \begin{pmatrix} n-1 \\ j \end{pmatrix} (\frac{1}{n-1})^j (1 - \frac{1}{n-1})^{n-1-j}$.

Exercice 27.14 :

1°) Soit $i \in \mathbb{N}_N$. X_i est à valeurs dans $\mathbb{N}^*$ et $P(X_i = k) = p(1-p)^{k-1}$ pour tout $k \in \mathbb{N}^*$, donc lorsque $n \in \mathbb{N}^*$, $P(X_i \leq n) = \sum_{k=1}^{n} p(1-p)^{k-1} = p \sum_{k=0}^{n-1} q^k = p\frac{1-q^n}{1-q} = 1 - q^n$, puis $P(X_i > n) = 1 - P(X_i \leq n) = q^n$.

2°) **a)** Soit $n \in \mathbb{N}^*$. Pour tout $\omega \in \Omega$,
$\omega \in (Y > n) \iff Y(\omega) > n \iff \forall i \in \mathbb{N}_N, \ \ X_i(\omega) > n$,
donc $P(Y > n) = P(X_1 > n, \ldots, X_N > n)$, puis d'après l'indépendance mutuelle de $X_1, \ldots, X_n$, $P(Y > n) = \prod_{i=1}^{N} P(X_i > n) = q^{nN}$ d'après la question précédente.
On en déduit que $P(Y \leq n) = 1 - q^{nN}$, ce qui est également vrai lorsque $n = 0$.
Ainsi, pour tout $n \in \mathbb{N}^*$,
$P(Y = n) = P((Y \leq n) \setminus (Y \leq n-1)) = q^{(n-1)N} - q^{nN} = (q^N)^{n-1}(1 - q^N)$.
b) Ainsi, $Y \sim \mathcal{G}(1 - q^N)$, donc d'après le cours, Y est d'espérance finie et $E(Y) = \frac{1}{1-q^N}$.
On peut expliquer sans calcul pourquoi Y suit une loi géométrique :

Pour tout $i \in \mathbb{N}_N$, sans modifier la loi de X_i, on peut supposer que la variable aléatoire X_i représente l'instant de premier succès d'une série de tests mutuellement indépendants de paramétre p. Considérons que ces N séries de tests sont réalisés en parallèle et que l'on s'arrête dès l'instant où l'un des tests est un succès. Cela revient à considérer que l'on effectue un nouveau test, qui se réalise si et seulement si parmi les N anciens tests, l'un au moins se réalise. Ainsi, le nouveau test est un échec avec une probabilité de $(1-p)^N$, donc est un succés avec une probabilité de $1-q^N$. Y représente l'instant de premier succès de ce nouveau test, il est donc naturel que $Y \sim \mathcal{G}(1-q^N)$.

Exercice 27.15 :

1°) Notons S le nombre de canards survivants. On a $S = \sum_{c \in \mathcal{C}} 1_{c\ survit}$
où $\mathcal{C}$ désigne l'ensemble des canards.
Par linéarité de l'espérance, $E(S) = \sum_{c \in \mathcal{C}} E(1_{c\ survit}) = \sum_{c \in \mathcal{C}} P(c\ survit)$.
Fixons $c \in \mathcal{C}$. Notons X le nombre de tirs réussis sur c. Chaque chasseur tire sur le canard c avec une probabilité de l'atteindre égale à $\frac{p}{20}$ et les chasseurs agissent indépendamment les uns des autres, donc d'après le cours, $X \sim \mathcal{B}(10, \frac{p}{20})$.
Alors $P(c\ survit) = P(X = 0) = (1 - \frac{p}{20})^{10}$, puis $E(S) = 20(1 - \frac{p}{20})^{10}$.

2°) Notons T le nombre de canards touchés. L'énoncé nous demande de calculer $E(T)$.
D'après le cours, $E(T) = \sum_{t=0}^{+\infty} tP(T = t)$, or d'après la formule des probabilités totales,

$P(T = t) = \sum_{n=0}^{+\infty} P(T = t \mid N = n)P(N = n)$, où N est le nombre de canards.
Alors, en utilisant la formule de Fubini pour des familles dénombrables de réels positifs,
$E(T) = \sum_{n=0}^{+\infty} P(N = n)E(T \mid N = n)$,

en convenant que $E(T \mid N = n) = \sum_{t=0}^{+\infty} tP(T = t \mid N = n)$.
On sait que $N \sim \mathcal{P}(\lambda)$ où l'on a posé $\lambda = 15$.
On obtient ainsi que $E(T) = \sum_{n=0}^{+\infty} e^{-\lambda} \frac{\lambda^n}{n!} n(1 - (1 - \frac{p}{n})^{10})$.

Exercice 27.16 :

1°) La probabilité étant uniforme, $P(G) = \dfrac{1}{\text{Card}(\Omega)}$.
Choisir un élément de Ω revient à choisir l'ensemble A de ses arêtes avec les contraintes $\text{Card}(A) = m$ et $A \subset \{(i, j)/1 \leq i < j \leq n\}$, donc cela revient à choisir m éléments parmi $N = \binom{n}{2}$. Ainsi $P(G) = \dfrac{1}{\binom{N}{m}}$.

2°) $i \sim j$ est donc l'ensemble des graphes de Ω possédant l'arête (i,j). Choisir un tel graphe revient à choisir les $m-1$ autres arêtes parmi les $N-1$ arêtes disponibles.

Ainsi $P(i \sim j) = \displaystyle\sum_{\omega \in (i\sim j)} P(\omega) = \frac{\mathrm{Card}(i \sim j)}{\binom{N}{m}} = \frac{\binom{N-1}{m-1}}{\binom{N}{m}}$, donc $P(i \sim j) = \dfrac{m}{N}$.

3°) On suppose que $m, n \geq 3$, sans quoi l'espérance est nulle.
Notons K ce nombre de triangles, qui est donc une variable aléatoire de Ω dans $\mathbb{N}$.
On a $K = \displaystyle\sum_{\substack{\{x,y,z\}\subset S \\ x,y,z \text{distincts}}} 1_{x\sim y, y\sim z, z\sim x}$, donc par linéarité de l'espérance,

$$E(K) = \sum_{\substack{\{x,y,z\}\subset S \\ x,y,z \text{distincts}}} E(1_{x\sim y, y\sim z, z\sim x}) = \sum_{\substack{\{x,y,z\}\subset S \\ x,y,z \text{distincts}}} P(x \sim y, y \sim z, z \sim x).$$

Le même argument utilisé en question 2 prouve que

$$P(x \sim y, y \sim z, z \sim x) = \frac{\binom{N-3}{m-3}}{\binom{N}{m}} = \frac{m(m-1)(m-2)}{N(N-1)(N-2)},$$

donc $E(K) = \dbinom{n}{3} \dfrac{m(m-1)(m-2)}{N(N-1)(N-2)}$.

Variance

Exercice 27.17 :

La théorie des familles sommables de réels positifs justifie les inégalités suivantes :
Pour $h \in \{1, \ldots, k-1\}$, d'après la formule de transfert,

$$\begin{aligned}
E(|X|^h) &= \sum_{x \in X(\Omega)} |x|^h P(X = x) \\
&\leq \sum_{\substack{x \in X(\Omega) \\ |x| \geq 1}} |x|^h P(X = x) + \sum_{\substack{x \in X(\Omega) \\ |x| \leq 1}} |x|^h P(X = x) \\
&\leq \sum_{\substack{x \in X(\Omega) \\ |x| \geq 1}} |x|^k P(X = x) + \sum_{\substack{x \in X(\Omega) \\ |x| \leq 1}} P(X = x) \\
&\leq \sum_{x \in X(\Omega)} |x|^k P(X = x) + \sum_{x \in X(\Omega)} P(X = x) \\
&= E(|X|^k) + 1.
\end{aligned}$$

Exercice 27.18 :

1°) Y_i est à valeurs dans $\{0, 1\}$, donc $Y_i \sim \mathcal{B}(q)$, où
$q = P(Y_i = 1) = P(X_i = X_{i+1} = 1) = p^2$ car X_i et X_{i+1} sont indépendantes.
En conclusion, $Y_i \sim \mathcal{B}(p^2)$.

2°) Par linéarité de l'espérance, $E(Y) = \sum_{i=1}^{n-1} E(Y_i)$, or d'après le cours, $E(Y_i) = p^2$, donc $E(Y) = (n-1)p^2$.

3°) Si $n \leq 1$, $Y = 0$ et $V(Y) = 0$.
Si $n = 2$, $Y = Y_1$ et $V(Y) = V(Y_1) = p^2(1-p^2)$.
Pour la suite, on suppose que $n \geq 3$.
D'après le cours, $V(Y) = \sum_{i=1}^{n-1} V(Y_i) + 2 \sum_{1 \leq i < j \leq n-1} \text{Cov}(Y_i, Y_j)$.
De plus, si $i, j \in \mathbb{N}_{n-1}$ avec $i + 1 < j$, alors $Y_i = X_i X_{i+1}$ et $Y_j = X_j X_{j+1}$ sont indépendantes d'après le lemme de coalition, donc $\text{Cov}(Y_i, Y_j) = 0$.
Ainsi, $V(Y) = (n-1)p^2(1-p^2) + 2\sum_{i=1}^{n-2} \text{Cov}(Y_i, Y_{i+1})$.
Soit $i \in \mathbb{N}_{n-2}$. D'après la formule de Koenig-Huygens,
$\text{Cov}(Y_i, Y_{i+1}) = E(Y_i Y_{i+1}) - E(Y_i)E(Y_{i+1})$, or, en tenant compte du fait que X_{i+1} est à valeurs dans $\{0, 1\}$, on a $X_{i+1}^2 = X_{i+1}$, donc $\text{Cov}(Y_i, Y_{i+1}) = E(X_i X_{i+1} X_{i+2}) - p^4$, mais $X_i X_{i+1} X_{i+2}$ est encore à valeurs dans $\{0, 1\}$, donc
$E(X_i X_{i+1} X_{i+2}) = P(X_i X_{i+1} X_{i+2} = 1) = p^3$. Ainsi, $\text{Cov}(Y_i, Y_{i+1}) = p^3 - p^4$, puis $V(Y) = (n-1)(p^2 - p^4) + 2(n-2)(p^3 - p^4) = (n-1)p^2 + 2(n-2)p^3 - (3n-5)p^4$.

Exercice 27.19 :

$det(M) = \sum_{\sigma \in \mathcal{S}_n} \varepsilon(\sigma) \prod_{j=1}^{n} X_{\sigma(j),j}$,
donc par linéarité de l'espérance puis indépendance des $X_{\sigma(j),j}$,
$E(det(M)) = \sum_{\sigma \in \mathcal{S}_n} \varepsilon(\sigma) \prod_{j=1}^{n} E(X_{\sigma(j),j}) = 0$.
Ainsi d'après la formule de Koenig-Huygens,
$Var(det(M)) = E([det(M)]^2) = E(\sum_{\sigma, \sigma' \in \mathcal{S}_n} \varepsilon(\sigma)\varepsilon(\sigma') \prod_{j=1}^{n} X_{\sigma(j),j} X_{\sigma'(j),j})$.
Soit $\sigma, \sigma' \in \mathcal{S}_n$ telle que $\sigma \neq \sigma'$:
$E(\prod_{j=1}^{n} X_{\sigma(j),j} X_{\sigma'(j),j}) = E(\prod_{\substack{1 \leq j \leq n \\ \sigma(j) = \sigma'(j)}} X_{\sigma(j),j}^2 \prod_{\substack{1 \leq j \leq n \\ \sigma(j) \neq \sigma'(j)}} X_{\sigma(j),j} X_{\sigma'(j),j})$.
Le premier produit vaut 1 et le second est un produit non vide de variables aléatoires mutuellement indépendantes,
donc $E(\prod_{j=1}^{n} X_{\sigma(j),j} X_{\sigma'(j),j}) = \prod_{\substack{1 \leq j \leq n \\ \sigma(j) \neq \sigma'(j)}} E(X_{\sigma(j),j}) E(X_{\sigma'(j),j}) = 0$.
Ainsi, $Var(det(M)) = \sum_{\sigma \in \mathcal{S}_n} \varepsilon(\sigma)^2 E(\prod_{j=1}^{n} X_{\sigma(j),j}^2) = \sum_{\sigma \in \mathcal{S}_n} E(1) = n!$.

Exercice 27.20 :

1°) Si l'on reprend la démonstration de la loi faible des grands nombres, l'inégalité de Bienaymé-Tchebychev donne $P(|S_n - E(S_n)| \geq \varepsilon) \leq \dfrac{Var(S_n)}{\varepsilon^2}$, or $E(S_n) = E(X_1) = 0$ et $Var(S_n) = \frac{1}{n}Var(X_1)$ où $Var(X_1) = E(X_1^2) = 1$, donc $P(|S_n| \geq \varepsilon) \leq \dfrac{1}{n\varepsilon^2}$.
Avec $n = 10\ 000$ et $\varepsilon = 0,1$, on obtient bien $P(|S_{10\ 000}| \geq 0,1) \leq \frac{1}{100}$.

2°) Soit $t \in \mathbb{R}_+$:
$P(X \geq a) \leq P(e^{tX} \geq e^{ta})$, donc d'après l'inégalité de Markov,
$P(X \geq a) \leq \dfrac{E(e^{tX})}{e^{ta}} = e^{-ta+H(t)}$, puis en travaillant dans $\mathbb{R}_+ \cup \{+\infty\}$,
$\ln[P(X \geq a)] \leq -ta + H(t)$, donc par passage à la borne inférieure,
$\ln[P(X \geq a)] \leq \inf\limits_{t\geq 0}(-ta + H(t))$, ce qui permet de conclure.

3°) Soit $t \in \mathbb{R}_+$.
$e^{H(t)} = E(e^{tS_n}) = E(\prod\limits_{i=1}^{n} e^{\frac{t}{n}X_i})$, donc d'après l'indépendance mutuelle des X_i,
$e^{H(t)} = \prod\limits_{i=1}^{n} E(e^{\frac{t}{n}X_i})$, puis d'après la formule de transfert,
$e^{H(t)} = \prod\limits_{i=1}^{n}(e^{-\frac{t}{n}} \times \dfrac{1}{2} + e^{\frac{t}{n}} \times \dfrac{1}{2}) = ch(\dfrac{t}{n})^n$, donc $H(t) = n \ln ch(\frac{t}{n})$.

4°) Les questions 2 et 3 donnent : $P(S_n \geq \varepsilon) \leq e^{\inf_{t\geq 0}(n \ln ch(\frac{t}{n}) - t\varepsilon)}$.
Supposons que $\varepsilon \in]0,1[$. Posons $f_n(t) = n \ln ch(\frac{t}{n}) - t\varepsilon$:
$f_n'(t) = \dfrac{sh\frac{t}{n}}{ch\frac{t}{n}} - \varepsilon = \text{th}\dfrac{t}{n} - \varepsilon$, donc $f_n'(t) = 0 \Longleftrightarrow t = n\ \text{argth}\varepsilon = \dfrac{n}{2}\ln\Big(\dfrac{1+\varepsilon}{1-\varepsilon}\Big)$.
Posons $t_0 = n \ln\sqrt{\dfrac{1+\varepsilon}{1-\varepsilon}}$: $f_n'(t)$ est négatif entre 0 et t_0 puis positif,
donc $\inf\limits_{t\geq 0} f_n(t) = f_n(t_0) = n\ln(\text{ch}(\text{argth}\varepsilon)) - \varepsilon t_0$.
Pour tout $x \in \mathbb{R}$, $\dfrac{1}{\text{ch}^2 x} = 1 - \text{th}^2 x$, donc $\text{ch}x = \dfrac{1}{\sqrt{1-\text{th}^2 x}}$.
On en déduit que $\inf\limits_{t\geq 0} f_n(t) = -\dfrac{n}{2}\Big(\ln(1-\varepsilon^2) + \varepsilon\ln\Big(\dfrac{1+\varepsilon}{1-\varepsilon}\Big)\Big)$.
Avec $n = 10\ 000$ et $\varepsilon = 0,1$, on calcule $f_n(t_0) = -50,0837 \pm 10^{-4}$,
donc $P(S_{10\ 000} \geq 0,1) \leq 1,8 \times 10^{-22}$.
De plus $-S_n = \dfrac{1}{n}\sum\limits_{i=1}^{n}(-X_i)$, mais $-X_i$ suit la même loi que X_i et les $-X_i$ sont mutuellement indépendantes, donc la question 3 est encore valable pour $X = -S_n$ et on en déduit $P(-S_{10\ 000} \geq 0,1) \leq 1,8 \times 10^{-22}$.
Or $(|S_n| \geq 0,1) = (S_n \geq 0,1) \cup (-S_n \geq 0,1)$,
donc $P(|S_n| \geq 0,1) \leq P(S_n \geq 0,1) + P(-S_n \geq 0,1) \leq 3,6 \times 10^{-22}$.

On obtient ainsi une majoration bien meilleure que celle de la première question.

Sommes de Riemann

Exercice 27.21 :

[$\frac{1}{b-a}\int_a^b g(t)dt$ désigne la valeur moyenne de $g(t)$ lorsque t décrit $[a,b]$. Il s'agit donc de montrer que l'image par f de la moyenne des $g(t)$ est inférieure à la moyenne des images par f des $g(t)$. Mais cette propriété est connue pour la moyenne d'un nombre fini de valeurs, d'après la convexité de f. Il reste à passer des sommes finies aux intégrales. C'est possible car une intégrale peut être vue comme limite de sommes finies.]

On sait que la somme de Riemann $\frac{b-a}{n}\sum_{k=1}^{n} g(a+k\frac{b-a}{n})$ tend vers $\int_a^b g(t)dt$ lorsque n tend vers $+\infty$. De même, $\frac{b-a}{n}\sum_{k=1}^{n} f\circ g(a+k\frac{b-a}{n})$ tend vers $\int_a^b f\circ g(t)dt$. De plus, f est convexe, donc pour tout $n\in\mathbb{N}^*$,
$f\left(\frac{1}{n}\sum_{k=1}^{n} g(a+k\frac{b-a}{n})\right)\le\frac{1}{n}\sum_{k=1}^{n} f(g(a+k\frac{b-a}{n}))$. f étant continue, lorsque l'on fait tendre n vers $+\infty$, on obtient l'inégalité de Jensen.

Exercice 27.22 :

[La présence d'une racine $n^{\text{ème}}$ et d'un produit nous incite à passer aux logarithmes.]

Pour tout $n\in\mathbb{N}^*$, posons $x_n=\sqrt[n]{\prod_{k=1}^{n}\Big(1+(\frac{k}{n})^2\Big)}$ et $y_n=\ln(x_n)$.

$y_n=\frac{1}{n}\sum_{k=1}^{n}\ln(1+(\frac{k}{n})^2)=\frac{1}{n}\sum_{k=1}^{n} f(\frac{k}{n})$, où $f:\begin{array}{ccc}[0,1] & \longrightarrow & \mathbb{R}\\ x & \longmapsto & \ln(1+x^2)\end{array}$.

y_n est une somme de Riemann, donc $y_n\underset{n\to+\infty}{\longrightarrow}\int_0^1 f(t)dt$.

[Pour calculer cette intégrale, on gère la présence du logarithme en intégrant par parties.]

Intégrons par parties.

$$\begin{aligned}\int_0^1 f(t)dt &= [x\ln(1+x^2)]_0^1-\int_0^1 x\times\frac{2x}{1+x^2}dx\\ &=\ln(2)+\int_0^1 2\Big(\frac{1}{1+x^2}-1\Big)dx\\ &=\ln(2)-2+2[\arctan(x)]_0^1\\ &=\ln(2)-2+\frac{\pi}{2}.\end{aligned}$$

Ainsi, $y_n\underset{n\to+\infty}{\longrightarrow}\ln(2\sqrt{e^\pi}e^{-2})$ or $x_n=e^{y_n}$, donc $\boxed{x_n\underset{n\to+\infty}{\longrightarrow}\frac{2\sqrt{e^\pi}}{e^2}}$.

Exercice 27.23 :

1°) f étant continue sur le compact $K = [\frac{k\pi}{n}, \frac{(k+1)\pi}{n}]$, elle est bornée et elle atteint ses bornes. On peut donc noter $m = f(\alpha)$ le minimum de f sur K et $M = f(\beta)$ le maximum de f sur K. Alors $f(\alpha)\int_{\frac{k\pi}{n}}^{\frac{(k+1)\pi}{n}} |sin(nx)|dx \leq \int_{\frac{k\pi}{n}}^{\frac{(k+1)\pi}{n}} f(x)|sin(nx)|dx \leq f(\beta)\int_{\frac{k\pi}{n}}^{\frac{(k+1)\pi}{n}} |sin(nx)|dx$. Or $\int_{\frac{k\pi}{n}}^{\frac{(k+1)\pi}{n}} |sin(nx)|dx > 0$, en tant qu'intégrale d'une application continue positive et non identiquement nulle, donc le quotient $\dfrac{\int_{\frac{k\pi}{n}}^{\frac{(k+1)\pi}{n}} f(x)|sin(nx)|dx}{\int_{\frac{k\pi}{n}}^{\frac{(k+1)\pi}{n}} |sin(nx)|dx}$ est compris entre $f(\alpha)$ et $f(\beta)$, mais f est continue, donc on peut appliquer le théorème des valeurs intermédiaires : il existe $\alpha_k \in K$ tel que $\dfrac{\int_{\frac{k\pi}{n}}^{\frac{(k+1)\pi}{n}} f(x)|sin(nx)|dx}{\int_{\frac{k\pi}{n}}^{\frac{(k+1)\pi}{n}} |sin(nx)|dx} = f(\alpha_k)$.

En adaptant cette démonstration, on peut prouver la formule de la moyenne : Soit f et g deux applications de $[a, b]$ dans $\mathbb{R}$ (avec $a, b \in \mathbb{R}$ et $a < b$) telles que f est continue et g positive. Alors il existe $\gamma \in [a, b]$ tel que $\displaystyle\int_a^b f(t)g(t)\ dt = f(\gamma)\int_a^b g(t)\ dt$.

2°) $\displaystyle\int_0^\pi f(x)|\sin nx|dx = \sum_{k=0}^{n-1}\int_{\frac{k\pi}{n}}^{\frac{(k+1)\pi}{n}} f(x)|sin(nx)|dx = \sum_{k=0}^{n-1} f(\alpha_k)\int_{\frac{k\pi}{n}}^{\frac{(k+1)\pi}{n}} |sin(nx)|dx$.
Soit $k \in \{0, \ldots, n-1\}$. L'application $x \longmapsto \sin(nx)$ étant de signe constant sur l'intervalle $[\frac{k\pi}{n}, \frac{(k+1)\pi}{n}]$, $\displaystyle\int_{\frac{k\pi}{n}}^{\frac{(k+1)\pi}{n}} |sin(nx)|dx = \left|\int_{\frac{k\pi}{n}}^{\frac{(k+1)\pi}{n}} sin(nx)dx\right| = \left|\left[\frac{\sin(nx)}{n}\right]_{\frac{k\pi}{n}}^{\frac{(k+1)\pi}{n}}\right| = \frac{2}{n}$, donc $\displaystyle\int_0^\pi f(x)|\sin nx|dx = \sum_{k=0}^{n-1} f(\alpha_k)\frac{2}{n} = \frac{2}{\pi}\sum_{k=0}^{n-1}\frac{\pi}{n}f(\alpha_k) \underset{n\to+\infty}{\longrightarrow} \frac{2}{\pi}\int_0^\pi f(x)dx$, d'après le cours sur les sommes de Riemann.