TD 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 vérifie : la loi de "C sachant que $I = 1$" est $\mathcal{B}(\dfrac{n_1}{n_1 + b_1})$, en convenant que la couleur noire est représentée par l'entier 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 27.13 :

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 $\mathcal{B}(n - 1, p)$ où $p_n = \frac{1}{n-1}$
et pour $j \in \{0, \ldots, n - 1\}$: $p_{j,n} = P(X_{n,k} = j) = \begin{pmatrix} n-1 \\ j \end{pmatrix} (\dfrac{1}{n-1})^j (1 - \dfrac{1}{n-1})^{n-1-j}$.
À j fixé, quand $n \to \infty$,
$p_{j,n} = \dfrac{(n-1)(n-2) \times (n-j)}{(n-1)(n-1) \times (n-1)} \dfrac{1}{j!} e^{(n-1)\ln(1-\frac{1}{n-1})} (1 - \dfrac{1}{n-1}))^{-j} \sim \dfrac{e^{-1}}{j!}$, 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 $(n - 1)p_n = 1 \underset{n \to +\infty}{\longrightarrow} 1$.

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 27.16 :

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

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 $P(i \sim j) = \sum_{\omega \in (i \sim j)} P(\omega) = \frac{\mathrm{Card}(i \sim j)}{\begin{pmatrix} N \\ m \end{pmatrix}} = \frac{\begin{pmatrix} N-1 \\ m-1 \end{pmatrix}}{\begin{pmatrix} N \\ m \end{pmatrix}}$, donc $P(i \sim j) = \frac{m}{N}$.

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 $K = \sum_{\substack{\{x,y,z\} \subset S \\ x,y,z \text{distincts}}} 1_{x \sim y, y \sim z, z \sim x}$, 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
$P(x \sim y, y \sim z, z \sim x) = \frac{\begin{pmatrix} N-3 \\ m-3 \end{pmatrix}}{\begin{pmatrix} N \\ m \end{pmatrix}} = \frac{m(m-1)(m-2)}{N(N-1)(N-2)}$,
donc $E(K) = \begin{pmatrix} n \\ 3 \end{pmatrix} \frac{m(m-1)(m-2)}{N(N-1)(N-2)}$.

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 $Y_i = X_iX_{i+1}$ et $Y_j = X_jX_{j+1}$ 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,
$\text{Cov}(Y_i, Y_{i+1}) = E(Y_iY_{i+1}) - E(Y_i)E(Y_{i+1})$, donc, en tenant compte du fait que X_{i+1} est à valeurs dans $\{0,1\}$, donc $X_{i+1}^2 = X_{i+1}$, on a $\text{Cov}(Y_i, Y_{i+1}) = E(X_iX_{i+1}X_{i+2}) - p^4$, mais $X_iX_{i+1}X_{i+2}$ est encore à valeurs dans $\{0,1\}$, donc
$E(X_iX_{i+1}X_{i+2}) = P(X_iX_{i+1}X_{i+2} = 1) = p^3$. 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 $\sigma\neq\sigma'$:
$E(\prod_{j=1}^n X_{\sigma(j),j}X_{\sigma'(j),j}) = E(\prod_{\substack{1\leq j\leq n\\ \sigma(j)=\sigma'(j)}} X^2_{\sigma(j),j}\prod_{\substack{1\leq j\leq n\\ \sigma(j)\neq\sigma'(j)}} X_{\sigma(j),j}X_{\sigma'(j),j})$.
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, $Var(det(M)) = \sum_{\sigma\in\mathcal{S}_n}\varepsilon(\sigma)^2E(\prod_{j=1}^n X^2_{\sigma(j),j}) = \sum_{\sigma\in\mathcal{S}_n}E(1) = n!$.

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,
$\ln[P(X\geq a)]\leq \inf_{t\geq 0}(-ta+H(t))$, ce qui permet de conclure.

3°) Soit $t\in\mathbb{R}_+$.
$e^{H(t)} = E(e^{tS_n}) = E(\prod_{i=1}^n e^{\frac{t}{n}X_i})$, donc d'après l'indépendance mutuelle des X_i,
$e^{H(t)} = \prod_{i=1}^n E(e^{\frac{t}{n}X_i})$, puis d'après la formule de transfert,
$e^{H(t)} = \prod_{i=1}^n (e^{-\frac{t}{n}}\times\dfrac{1}{2}+e^{\frac{t}{n}}\times\dfrac{1}{2}) = ch(\dfrac{t}{n})^n$, 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 $f_n(t) = n\ln ch(\frac{t}{n})-t\varepsilon$:
$f_n'(t) = \dfrac{sh\frac{t}{n}}{ch\frac{t}{n}}-\varepsilon = \mathrm{th}\dfrac{t}{n}-\varepsilon$, donc $f_n'(t)=0\Longleftrightarrow t = n\ \mathrm{argth}\varepsilon = \dfrac{n}{2}\ln\Big(\dfrac{1+\varepsilon}{1-\varepsilon}\Big)$.

Posons $t_0 = n\ln\sqrt{\dfrac{1+\varepsilon}{1-\varepsilon}}$: $f_n'(t)$ est négatif entre 0 et t_0 puis positif,
donc $\inf\limits_{t\geq 0} f_n(t) = f_n(t_0) = n\ln(\mathrm{ch}(\mathrm{argth}\varepsilon)) - \varepsilon t_0$.
Pour tout $x\in\mathbb{R}$, $\dfrac{1}{\mathrm{ch}^2 x} = 1 - \mathrm{th}^2 x$, donc $\mathrm{ch}x = \dfrac{1}{\sqrt{1-\mathrm{th}^2 x}}$.
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 $P(S_{10\ 000} \geq 0,1) \leq 1,77\times 10^{-22}$.
De plus $-S_n = \dfrac{1}{n}\displaystyle\sum_{i=1}^{n}(-X_i)$, 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 $P(-S_{10\ 000} \geq 0,1) \leq 1,77\times 10^{-22}$.
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.

Exercice 27.21 :

[$\dfrac{1}{b-a}\displaystyle\int_a^b g(t)dt$ 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 $\dfrac{b-a}{n}\displaystyle\sum_{k=1}^{n} g(a+k\frac{b-a}{n})$ tend vers $\displaystyle\int_a^b g(t)dt$ lorsque n tend vers $+\infty$. De même, $\dfrac{b-a}{n}\displaystyle\sum_{k=1}^{n} f\circ g(a+k\frac{b-a}{n})$ tend vers $\displaystyle\int_a^b f\circ g(t)dt$. De plus, f est convexe, donc pour tout $n\in\mathbb{N}^*$, $f\left(\dfrac{1}{n}\displaystyle\sum_{k=1}^{n} g(a+k\frac{b-a}{n})\right) \leq \dfrac{1}{n}\displaystyle\sum_{k=1}^{n} f(g(a+k\frac{b-a}{n}))$. 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 $x_n = \sqrt[n]{\displaystyle\prod_{k=1}^{n}\Big(1+(\frac{k}{n})^2\Big)}$ et $y_n = \ln(x_n)$.

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

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

[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 $f(\alpha)\int_{\frac{k\pi}{n}}^{\frac{(k+1)\pi}{n}} |sin(nx)|dx \le \int_{\frac{k\pi}{n}}^{\frac{(k+1)\pi}{n}} f(x)|sin(nx)|dx \le f(\beta)\int_{\frac{k\pi}{n}}^{\frac{(k+1)\pi}{n}} |sin(nx)|dx$. 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 $\int_a^b f(t)g(t)\ dt = f(\gamma)\int_a^b g(t)\ dt$.

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