Probabilités

Hugo SACCU.

TD n° 4

Exercice 1. Running Time.

Q1. Soit $X_n \sim \mathcal{U}(\{0,1\}^n)$. On a : $\mathbb{E}[T(X_n)] \leq k \cdot n^2$ avec k fixé.

D'après l'inégalité de Markov :

$$P(T(X_n) \geq n^2 f(n)) \leq \frac{\mathbb{E}[T(X_n)]}{n^2 f(n)} = \frac{k}{f(n)} \xrightarrow[n \to +\infty]{} 0.$$

Q2. Notons $T_{max} = \max_{\omega \in \Omega} T(\omega)$.

$$k n^2 \geq \mathbb{E}[T] = \sum_{\omega \in \Omega} T(\omega) \, 1/2^n \geq T_{max} \frac{1}{2^n}$$

D'où $T_{max} \leq k \cdot n^2 2^n$. On conclut que $T_{max} = O(n^2 2^n)$.

Exercice 2. Coquilles dans un TD

Q1. Notons N le temps de relecture pour enlever les 4 coquilles. et N_i pour la i-ème coquille.

$N_i \sim \mathcal{G}(1/3)$ et $N = \max(N_1, N_2, N_3, N_4)$.

$$\mathbb{P}(N \leq n) \underset{\text{indépendance}}{=} \prod_{i=1}^{4} \mathbb{P}(N_i \leq n) = \prod_{i=1}^{4} \sum_{k=1}^{n} \mathbb{P}(N_i = k)$$

$$= \prod_{i=1}^{4} \sum \left(\frac{2}{3}\right)^k \frac{1}{3}$$

$$= \prod_{i=1}^{4} \cancel{\frac{1}{3}} \times \frac{1-\left(\frac{2}{3}\right)^n}{1/3}$$

$$= \left(1 - \left(\frac{2}{3}\right)^n\right)^4$$

Avec $n = 10$, on a $\mathbb{P}(N \leq 10) \approx 0{,}9$.

Q2. Le problème du collectionneur de vignettes, mais, à chaque étape, on peut corriger plusieurs coquilles, alors qu'on ne peut avoir qu'un magnet.

Q3.. D'après Tchebychev,

$$\mathbb{P}(10-n < X < 10+n) = 1 - \mathbb{P}(|X-\mu| \geq n) \leq 1-\frac{\sigma^2}{n^2}$$

On a bien le résultat demandé pour $n \geq 50$.

Exercice 3. Tester la pièce.

On lance n fois la pièce. Le nombre de "Pile"s est $X \sim \mathcal{B}(n,p)$.

$$\mathbb{E}[X] = n \cdot p \qquad \mathrm{Var}[X] = np(1-p)$$

$$\mathbb{P}(|X - \mathbb{E}[X]| \geq n/10) \leq \frac{\mathrm{Var}[X]}{(n/10)^2} = \frac{p(1-p)}{n/100} \leq \frac{25}{n}$$

Pour avoir une probabilité d'au moins 0,9, il faut

$$0,9 \leq 1 - 25/n \iff 25/n \leq 0,1 \iff n \geq 250$$

Exercice 4. Comparer Markov, Tchebychev et Chernoff.

$$X \sim \mathcal{B}(n, 1/6) \qquad \mathbb{E}[X] = n/6 \qquad \mathrm{Var}[X] = 5n/36.$$

Markov : $\mathbb{P}(X \geq n/4) \leq \dfrac{n/6}{n/4} = \dfrac{4}{6}$

Tchebychev : $\mathbb{P}(|X - \mathbb{E}[X]| \leq n/12) \leq \dfrac{5n/36}{(n/12)^2} = \dfrac{144 \times 5}{36} \times \dfrac{1}{n} = 20/n$

Chernoff : $\mathbb{P}(X \leq (1+\frac{1}{2})\mathbb{E}[X]) \leq \exp\left(-\dfrac{1/4}{5/2}\dfrac{n}{6}\right) = \exp\left(-\dfrac{n}{60}\right)$.

Exercice 5. Chernoff Bound Interval.

Q1. On a : $\mathbb{E}[e^{\lambda Y}] = (1-p) + p e^{\lambda}$

$= 1 - \mathbb{E}[X] + \mathbb{E}[X] e^{\lambda}$

Convexité de $\exp(\lambda -)$ et croissance de $\mathbb{E}[-]$ $\curvearrowright$ $= \mathbb{E}[(1-X)e^{0} + e^{\lambda} X]$

$\geq \mathbb{E}[e^{\lambda X + (1-X)\cdot 0}] \geq \mathbb{E}[e^{\lambda X}]$

Q2. Pour $\lambda > 0$,

$$\mathbb{P}(X \geq (1+\varepsilon)\mu) = \mathbb{P}(e^{\lambda X} \geq e^{\lambda\mu(1+\varepsilon)})$$

$$\leq \frac{\mathbb{E}[e^{\lambda X}]}{e^{\lambda\mu(1+\varepsilon)}}$$

Par inégalité de Markov

De plus, $\mathbb{E}[e^{\lambda X}] = \mathbb{E}[e^{\lambda(X_1 + \cdots + X_n)}]$

$= \prod \mathbb{E}[e^{\lambda X_i}]$

$= e^{\sum p_i(e^{\lambda}-1)}$

En effet,

$$\mathbb{E}[e^{\lambda X_i}] \leq (1-p_i) + p_i e^{\lambda}$$
$$= 1 + p_i(e^{\lambda}-1)$$
$$\leq e^{p_i(e^{\lambda}-1)}$$

D'où, $\mathbb{P}(X \geq (1+\varepsilon)\mu) \leq \dfrac{e^{\mu(e^{\lambda}-1)}}{e^{\lambda\mu(1+\varepsilon)}}$ où $\lambda = \ln(1+\varepsilon)$.

Exercice 6. Fonction génératrice

1) $G_X(z) = \mathbb{E}[z^X] = \sum_{n=0}^{+\infty} z^n \, \mathbb{P}(X=n)$

$G'_X(z) = \sum_{n=1}^{+\infty} n z^{n-1} \, \mathbb{P}(X=n) \Rightarrow \mathbb{E}[X] = G'_X(1)$

$G''_X(z) = \sum_{n=2}^{+\infty} n(n-1) z^{n-2} \, \mathbb{P}(X=n)$

$$\mathrm{Var}[X] = G''_X(1) + G_X'(1) - (G_X'(1))^2$$

Q2. $G_X(z) = \mathbb{E}[z^X] = \sum_{k=0}^{+\infty} z^k \mathcal{C}(\lambda) \frac{\lambda^k}{k!}$

$$= \mathcal{C}(\lambda) \exp(\lambda z)$$

Q3. $G_X(1) = \sum_{n=0}^{+\infty} \mathbb{P}(X=n) = 1$

d'où $\mathcal{C}(\lambda) \exp(\lambda) = 1$ pour tout $\lambda \in \mathbb{R}$

On en déduit $\mathcal{C}(\lambda) = \exp(-\lambda)$.

Q4. $G_X(z) = \exp(\lambda(z-1))$

$\mathbb{E}[X] = G'_X(1) = \lambda e^{-\lambda} \cdot e^{\lambda} = \lambda$

$$\mathrm{Var}[X] = G''_X(1) + G'_X(1) - (G'_X(1))^2$$
$$= \lambda^2 \cancel{e^{-\lambda} e^{\lambda}} \quad \lambda \quad - \lambda^2$$
$$= \lambda$$

Q5. $X \sim \mathcal{B}(n, p)$.

$G_X(z) = \sum_{k=0}^{+\infty} z^k \, \mathbb{P}(X = k)$

$= \sum_{k=0}^{n} z^k \binom{n}{k} p^k (1-p)^{n-k}$

$= (pz + 1 - p)^n$

$\mathbb{E}[X] = G'_X(1) = np(p+1-p)^{n-1}$
$= np$

$\mathrm{Var}[X] = G''_X(1) + G'_X(1) - G'_X(1)^2$

$= n(n-1)p^2 + np - n^2p^2$
$= p^2(\cancel{n^2} - n - \cancel{n^2}) + np$
$= np(1-p)$.

Q6. $G_S(z) = \mathbb{E}[z^S] = \sum_{n=0}^{+\infty} \mathbb{E}[z^S \mid N=n] \cdot \mathbb{P}(N=n)$

$$= \sum_{n=1}^{+\infty} \mathbb{E}[z^{X_1+\cdots+X_n}] \cdot \mathbb{P}(N=n) + \mathbb{P}(N=0)$$

$$= \sum_{n=1}^{+\infty} \mathbb{E}[z^{X_1} \times \cdots \times z^{X_n}] \cdot \mathbb{P}(N=n) + \mathbb{P}(N=0)$$

car les $(X_i)_{i \geq 0}$ sont i.i.d

$$= \sum_{n=1}^{+\infty} \left(\mathbb{E}[z^{X_1}]\right)^n \cdot \mathbb{P}(N=n)$$

$$= G_N(G_{X_1}(z))$$

Si $X_i \sim \mathcal{G}(q)$ et $N \sim \mathcal{G}(p)$,

$$G_N(z) = \sum_{n=1}^{+\infty} z^n q (1-q)^{n-1} = zq \sum_{n=0} (1-q)^n z^n$$

$$= zq / (1-(1-q)z)$$

D'où

$$G_S(z) = \frac{pzq}{(1-(1-q)z)(1-(1-p)(qz/1-(1-q)z))}$$

$$= \frac{pzq}{1 - z + \cancel{qz} - \cancel{qz} + pqz}$$

$$= \frac{(pq)z}{1-(1-pq)z}$$

On peut en déduire que $S \sim \mathcal{G}(pq)$.

Exercice 7. Probabilités conditionnelles.

Q1. On pose $\Omega' = \Omega \setminus Y^{-1}(\{0\})$, puis

$$X : \mathbb{N}^* \longrightarrow \Omega' \qquad\qquad \mathbb{P}' : \Omega' \longrightarrow [0,1]$$
$$n \longmapsto Y(n). \qquad\qquad \omega \longmapsto \mathbb{P}(\omega) / \mathbb{P}(Y \neq 0).$$

Q2. $\mathbb{E}[X^2] - \mathbb{E}[X]^2 = \mathrm{Var}[X] \geq 0$

D'où $\mathbb{E}[X^2] \geq \mathbb{E}[X]^2$.

Q3. $\mathbb{E}[Y] = \sum_{y=0}^{+\infty} y\, \mathbb{P}(Y=y) = \sum_{y=1}^{+\infty} \underbrace{y}_{\geq 1} \mathbb{P}(Y=y) \geq \mathbb{P}(Y \geq 1) = \mathbb{P}(Y \neq 0)$

$$\frac{\mathbb{E}[Y]^2}{\mathbb{E}[Y^2]} = \frac{\mathbb{E}[X]^2 \cdot \mathbb{P}(Y \neq 0)^{\cancel{2}}}{\mathbb{E}[X^2] \cdot \cancel{\mathbb{P}(Y \neq 0)}} \underset{\text{par Q2.}}{\leq} \mathbb{P}(Y \neq 0) \cdot \times 1$$

Exercice 8. Bucket Sort.

Q1. Si on reçoit $x = \overline{b_k \dots b_0}^2$, on le met dans le bucket numéroté $\overline{b_k \dots b_{k-m}}^2$ (c'est un bitshift de $k-m$).

Q2. Soit $X \sim \mathcal{U}(⟦0, 2^k - 1⟧)$.
$\mathbb{P}(\text{On place } X \text{ dans le seau } i) = 1/n$.

D'où,

$$\mathbb{P}(X_i = \ell) = \binom{n}{\ell} \left(\frac{1}{n}\right)^{\ell} \cdot \left(1 - \frac{1}{n}\right)^{n-\ell}$$

C'est une loi binomiale $\mathcal{B}(n, 1/n)$.

Q3. Notons T le temps de calcul du bucketsort.

$$\mathbb{E}[T] = \sum_{i=0}^{2^{k-1}} \mathbb{E}[T_i] \quad + \mathcal{O}(n)$$

$$\mathbb{E}[T_i] = \mathbb{E}[M X_i^2] = M\left(\mathrm{Var}[X_i] + \mathbb{E}[X_i]^2\right)$$
$$\leq M\left(n \cdot \frac{1}{n} \cdot \frac{n-1}{n} + \frac{n^2}{n^2}\right) = M n\left[2 - 1/n\right] \leq 2Mn$$

D'où, $\mathbb{E}[T] = \mathcal{O}(n)$.

Fim du TD

TD n° 5

Exercice 1. Algorithme probabiliste pour calculer la médiane.

Q1. La partie (a) est en $O(n)$,
puis (b) en $O(n^{3/4} \underbrace{\log n}_{O(n^{1/4})}) = O(n)$,
puis (c) en $O(n)$
puis (d) en $O(n)$
et enfin (e) en $O(n^{3/4} \log n) = O(\log n)$.

D'où l'algorithme est en $O(n)$.

Q2 Erreur 1 : On a $\mathbb{E}[\text{card } F] = \sum \mathbb{E}[Y_i] = n \cdot n^{-1/4} = n^{3/4}$

Par inégalité de Bienaymé-Tchebychev appliquée à

$$\mathbb{P}(|X - \mathbb{E}[X]| \geq a) \leq \frac{Var[X]}{a^2}.$$

TD n° 6

Exercice 1. Black jack.

Bof ...

Loi faible des grands nombres :

Soit $\bar{X}$ le nombre de parties de Blackjack gagnées.

$$\mathbb{P}(|\bar{X} - 1/2| \geq 5\%) \leq \frac{\mathrm{Var}[X]}{n(5\%)^2} \leq \frac{1}{n \cdot (5\%)^2}$$

Pour avoir $\mathbb{P}(\text{il ne triche pas}) \geq 10\%$ il suffit d'avoir $n \geq \frac{1}{(10\%)(5\%)^2} = 4000$

D'après Chernoff,

$$\mathbb{P}(X \geq \tfrac{1}{2} + (5\%)) \leq \exp(-2(5\%)^2/n)$$

$$\exp(-2(5\%)^2/n) \leq 10\% \iff 2 \cdot (5\%)^2/n \leq \ln(10)$$
$$\iff n \geq \ln(10) / (2 \cdot (5\%)^2)$$

D'où $n \geq 461$. Beaucoup plus précis...

Exercice 2. Random Algorithm.

Algorithme $\mathcal{B}$:

On lance $k(x)$ fois l'algorithme $\mathcal{A}$ sur l'entrée x.

On note (A_i) les résultats des executions de $\mathcal{A}$.

On renvoie la réponse majoritaire

Si $x \notin L$, alors on a : $\mathbb{E}[X] \leq k/3$.
Si $x \in L$, alors on a : $\mathbb{E}[X] \geq 3k/4$.

Par inégalité de Chernoff, on a:

- si $x \notin L$, alors

$$\mathbb{P}(\mathcal{B}(x) = 1) = \mathbb{P}(X \geq k/2) \leq \exp\left(-\frac{(1/2)^2}{2+\frac{1}{2}} \cdot \frac{k}{3}\right) = \exp\left(-\frac{k}{30}\right).$$

- si $x \in L$, alors

$$\mathbb{P}(\mathcal{B}(x) = 0) = \mathbb{P}(X \leq k/2) \leq \exp\left(-\frac{1}{2 \cdot 9} \times \frac{3k}{4}\right) = \exp(-k/24)$$

$$\Rightarrow k = 30\,|x| \,/\, \log_2 e$$

Exercice 3. Intercrupteurs

Partie I.

Q1 On pose $Y := \mathbb{1}_{X \geq 1/4}$. On a:

$$1 = \mathbb{E}[X] = \mathbb{E}[XY] + \underbrace{\mathbb{E}[X \mathbb{1}_{X \leq 1/4}]}_{\leq 1/4}$$

d'où

$$\sqrt{3\,\underbrace{\mathbb{E}[Y^2]}_{=\mathbb{E}[Y] = \mathbb{P}(X \geq 1/4)}} \geq \mathbb{E}[XY] \geq \frac{3}{4}$$

D'où $\gamma = 3/16$.

Q2. $\mathbb{E}[Y^2] = \mathrm{Var}[Y] + \mathbb{E}[Y]^2$

$= \frac{1}{n}\sum \mathrm{Var}[X_i]$ et $\mathbb{E}[Y] = 0$ car $\mathbb{E}[X_i] = 0$

$= 1$

$$\mathbb{E}[Y^4] = \frac{1}{n^2}\,\mathbb{E}\left[\sum X_i X_j X_k X_\ell\right]$$

$$= \frac{1}{n^2}\sum \mathbb{E}[X_i X_j X_k X_\ell]$$

Cas 1: $i \neq j, k, \ell$ $\Rightarrow$ $\mathbb{E}[X_i X_j X_k X_\ell] = 0$

Cas 2 : $i = k \neq j = \ell \implies \mathbb{E}[X_i X_j X_k X_\ell] = 1$

Cas 3 : $i = j = k = \ell \implies \mathbb{E}[X_i X_j X_k X_\ell] = 1$

D'où $\mathbb{E}[Y^4] = \frac{1}{n^2}\Big(n + \underbrace{3}_{\substack{i=j\neq k=\ell \\ i=k\neq j=\ell \\ i=\ell\neq j=k}} \cdot \underbrace{n(n-1)}_{\text{choix de } i \text{ et choix du } 2^{nd} \text{ indice}}\Big) = 3 - \frac{2}{n} \leq 3.$

$$\overset{Q1}{\gamma \geq} \mathbb{P}(Y^2 \geq 1/4) \overset{\text{passage à la racine}}{=} \mathbb{P}\left(\frac{1}{\sqrt{n}}|X_1 + \cdots + X_n| \geq 1/2\right)$$

$$= \mathbb{P}(|X_1 + \cdots + X_n| \geq \sqrt{n}/2) \underset{\text{Markov}}{\leq} \frac{\mathbb{E}[|X_1 + \cdots + X_n|]}{\sqrt{n}/2}$$

D'où, $\mathbb{E}[|X_1 + \cdots + X_n|] \geq \frac{\gamma}{2} \cdot \sqrt{n}$.