

# Pas de malchance

## 1 Un peu de géométrie

▷ **Question 1** Une feuille de TD contient 4 coquilles. À chaque relecture, une coquille non corrigée est corrigée avec probabilité  $1/3$ . Les relectures sont indépendantes les unes des autres. Combien de relectures faut-il faire pour que la probabilité qu'il ne subsiste aucune coquille soit supérieure à  $0,9$  ?

## 2 Moi d'abord

▷ **Question 2** Alan et Beth lancent une pièce à tour de rôle, le premier qui fait face a gagné. Alan commence, quelle est la probabilité qu'il gagne ?

▷ **Question 3** Peut-on truquer la pièce pour rendre le jeu équitable ?

▷ **Question 4** Cette fois ils lancent deux dés. Alan gagne s'il obtient 7, Beth gagne si elle obtient 6. Elle s'est déjà fait avoir une fois et exige de commencer. Est-ce un jeu équitable ?

▷ **Question 5** L'une de ces parties risque-t-elle de se prolonger indéfiniment ?

## 3 Estimateurs

Soient  $X_1, \dots, X_n$  des variables aléatoires réelles indépendantes, de même espérance  $m$  et de même variance  $\sigma^2 > 0$ .

▷ **Question 6** On pose  $\overline{X}_n := \frac{1}{n} \sum_{i=1}^n X_i$ . Calculer l'espérance et la variance de  $\overline{X}_n$ . Pourquoi appelle-t-on  $\overline{X}_n$  la moyenne empirique ?

▷ **Question 7** On pose  $S_n^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \overline{X}_n)^2$ . Calculer l'espérance de  $S_n^2$ . Quel est l'intérêt de cette variable ?

▷ **Question 8** Quelle est la différence avec la variance empirique  $V_n := \frac{1}{n} \sum_{i=1}^n (X_i - \overline{X}_n)^2$  ?

▷ **Question 9** Exprimer  $S_n^2$  et  $V_n$  à l'aide de  $\sum_{i=1}^n X_i$ ,  $\sum_{i=1}^n X_i^2$  et  $n$ . Quel est l'intérêt ?

On suppose de plus que pour tout  $i$ ,  $X_i$  est une variable aléatoire uniformément distribuée dans  $[0; \theta]$ .

▷ **Question 10** Estimer  $\theta$  à l'aide de  $\overline{X}_n$ . Quelle est la variance de cet estimateur ?

▷ **Question 11** On pose  $M_n := \max(X_1, \dots, X_n)$ . Estimer  $\theta$  à l'aide de  $M_n$ . Indice : quelle est l'espérance de  $M_n$  ?

▷ **Question 12** Calculer la variance de ce nouvel estimateur et comparer au précédent.

## 4 Réponses aux exercices

### ▷ Question 1, page 1

$$P[\text{La coquille n° } i \text{ est corrigée en au plus } n \text{ relectures}] \\ = 1 - P[\text{Elle n'est pas corrigée après } n \text{ relectures}] = 1 - \left(\frac{2}{3}\right)^n$$

$$\text{Donc } P[\text{Les 4 coquilles sont corrigées en au plus } n \text{ relectures}] = \left(1 - \left(\frac{2}{3}\right)^n\right)^4$$

Enfin, puisque  $n$  est entier,  $\left(1 - \left(\frac{2}{3}\right)^n\right)^4 \geq 0,9 \iff n \geq 10$  (presque 9, mais rigoureusement c'est 10).

### ▷ Question 2, page 1

Alan gagne si la pièce tombe sur face lors d'un lancé de numéro impair (le premier lancé est numéroté 1).

$$\begin{cases} P[\text{Alan gagne}] = \sum_{i=0}^{\infty} \left(\frac{1}{2}\right)^{2i+1} \\ P[\text{Beth gagne}] = \sum_{i=0}^{\infty} \left(\frac{1}{2}\right)^{2i+2} = \frac{1}{2} P[\text{Alan gagne}] \end{cases}$$

$$\text{Donc } P[\text{Alan gagne}] = \frac{2}{3}.$$

### ▷ Question 3, page 1

Posons  $p := P[\text{face}]$ .

$$\begin{cases} P[\text{Alan gagne}] = \sum_{i=0}^{\infty} (1-p)^{2i} \cdot p \\ P[\text{Beth gagne}] = \sum_{i=0}^{\infty} (1-p)^{2i+1} \cdot p = (1-p) \cdot P[\text{Alan gagne}] \end{cases}$$

Donc  $P[\text{Alan gagne}] = P[\text{Beth gagne}] \iff p = 0$ . Le seul moyen de truquer la pièce est de faire en sorte que personne ne gagne.

### ▷ Question 4, page 1

$$\text{Posons } \begin{cases} P_A = P[\text{somme des dés} = 7] = \frac{6}{36} \\ P_B = P[\text{somme des dés} = 6] = \frac{5}{36} \end{cases}.$$

$$P[\text{Beth gagne}] = \sum_{i \geq 0} (1 - P_B)^i (1 - P_A)^i \cdot P_B$$

$$P[\text{Alan gagne}] = (1 - P_B) \sum_{i \geq 0} (1 - P_A)^i (1 - P_B)^i \cdot P_A = (1 - P_B) \frac{P_A}{P_B} P[\text{Beth gagne}] = \frac{31}{30} P[\text{Beth gagne}].$$

Alan est avantagé.

### ▷ Question 5, page 1

Si la pièce est truquée et ne tombe jamais sur face, le jeu est infini.

Sinon, soit  $0 < q \leq \min(P[\text{un lancer d'Alan est gagant}], P[\text{un lancer de Beth est gagant}])$ . La probabilité que le jeu dure au moins  $n$  étapes est majorée par  $(1 - q)^n$ , qui tend vers 0 en  $n \rightarrow \infty$ .

### ▷ Question 6, page 1

$$E[\overline{X}_n] = \frac{1}{n} \sum_{i=1}^n E[X_i] = m.$$

$$\text{Var}[\overline{X}_n] = \text{Var}\left[\frac{1}{n} \sum X_i\right] = \frac{1}{n^2} \text{Var}[\sum X_i] = \frac{1}{n^2} n \sigma^2 = \frac{\sigma^2}{n}.$$

On réalise une expérience avec  $n$  tirages,  $\overline{X}_n$  est une estimation de  $m$ , mais c'est une variable aléatoire dont la valeur dépend de l'expérience. « Moyenne empirique » signifie que c'est la moyenne obtenue selon le résultat de l'expérience, il peut y avoir une erreur par rapport à l'espérance  $m$ .

▷ **Question 7, page 1**

(Toutes les sommes sont pour  $1 \leq i \leq n$ .)

$$\begin{aligned} E(S_n^2) &= E\left[\frac{1}{n-1} \sum (X_i - \bar{X}_n)^2\right] \\ &= \frac{1}{n-1} E\left[\sum X_i^2 - 2 \sum X_i \bar{X}_n + \sum \bar{X}_n^2\right] \\ &= \frac{1}{n-1} E\left[\sum X_i^2 - 2n\bar{X}_n^2 + n\bar{X}_n^2\right] \\ &= \frac{1}{n-1} \sum E[X_i^2] - \frac{n}{n-1} E[\bar{X}_n^2] \\ &= \frac{n}{n-1} E[X_1^2] - \frac{1}{n(n-1)} \left(E[\sum X_i^2] + E[\sum_{i \neq j} X_i X_j]\right) \\ &= \frac{n}{n-1} E[X_1^2] - \frac{1}{n(n-1)} \left(nE[X_1^2] + n(n-1)E[X_1]^2\right) \\ &= E[X_1^2] - E[X_1]^2 = \sigma^2 \end{aligned}$$

$S_n^2$  est donc un moyen d'estimer  $\sigma^2$ .

▷ **Question 8, page 1**

$V_n$  est la variance observée sur les  $n$  tirages. Si  $X_1, \dots, X_n$  n'est qu'un échantillon de la population,  $V_n$  a un *biais* par rapport à la variance de la variable aléatoire  $X_1$  et n'est donc pas une bonne manière d'estimer  $\sigma^2$ .

▷ **Question 9, page 1**

$$S_n^2 = \frac{1}{n-1} \left(\sum X_i^2 - \frac{1}{n} (\sum X_i)^2\right). \quad V_n = \frac{n-1}{n} S_n^2.$$

L'intérêt est de limiter la mémoire nécessaire (par exemple dans une calculatrice). Après avoir effectué  $k$  mesures, on ne retient que  $\sum_{i=1}^k X_i$ ,  $\sum_{i=1}^k X_i^2$  et  $k$ . On peut mettre à jour ces variables au fur et à mesure que l'expérience avance.

▷ **Question 10, page 1**

La densité de  $X$  est  $\frac{1}{\theta}$ .  $E(\bar{X}_n) = E(X) = \int_0^\theta \frac{x}{\theta} dx = \frac{\theta}{2}$ . Un estimateur de  $\theta$  est donc  $2\bar{X}_n$ . Noter que  $E(2\bar{X}_n)$  n'est pas donné par l'expérience, on ne peut donc pas dire que c'est un estimateur.

$$\text{Var}(2\bar{X}_n) = 4 \frac{1}{n} \text{Var}(x) = \frac{4}{n} \left(\int_0^\theta \frac{x^2}{\theta} dx - E(X)^2\right) = \frac{4}{n} \left(\frac{\theta^2}{3} - \frac{\theta^2}{4}\right) = \frac{\theta^2}{3n}.$$

▷ **Question 11, page 1**

$$P[M_n \leq x] = P[\bigcap_{i=1}^n X_i \leq x] = P[X \leq x]^n = \left(\frac{x}{\theta}\right)^n. \quad M_n \text{ a donc pour densité}$$

$$\begin{cases} nx^{n-1}/\theta^n & \text{si } 0 \leq x \leq \theta \\ 0 & \text{sinon} \end{cases}.$$

$$E[M_n] = \int_0^\theta x(nx^{n-1})/\theta^n dx = \frac{n}{n+1} \theta.$$

Un estimateur de  $\theta$  est donc  $\frac{n}{n+1} M_n$ .

▷ **Question 12, page 1**

$$E[M_n^2] = \int_0^\theta x^2(nx^{n-1})/\theta^n dx = \frac{n}{n+2} \theta^2.$$

$$\text{Var}[M_n] = E[M_n^2] - E[M_n]^2 = \frac{n}{n+2} \theta^2 - \frac{n^2}{(n+1)^2} \theta^2 = \frac{n\theta}{(n+2)(n+1)^2}$$

$\text{Var}[M_n] = \Omega\left(\frac{1}{n^2}\right)$  tandis que  $\text{Var}[\bar{X}_n] = \Omega\left(\frac{1}{n}\right)$ . Ce second estimateur est donc plus précis.