

CHAPITRE 25 : LOI FAIBLE DES GRANDS NOMBRES

Proba 5

Chapitre «en mode TD». On se place dans un espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$

Rappels

Propriété 1 (Échauffement - somme de v.a. indépendantes). Si X, Y sont deux v.a.r.d indépendantes qui admettent une variance, $X + Y$ admet aussi une variance $V(X + Y) = V(X) + V(Y)$.

Exercice (Démonstration). On rappelle que pour X et Y des v.a. indépendantes on a $E(XY) = E(X)E(Y)$. Montrer :

$$V(X + Y) = V(X) + V(Y)$$

On admet ici que les espérances et variances écrites sont bien définies. *Revenir à la définition de la variance !*

Exercice (Rappel sur une autre approximation vue au chapitre 20 pour les variables aléatoires binômiales). On considère une assemblée de n personnes et on note $p_{n,k}$ la probabilité qu'exactly k de ces personnes soient nées le premier janvier. En faisant les approximations qui vous semblent raisonnables, approximer $p_{600;1}$. On pourra utiliser la calculatrice.

I. Inégalités de Markov et de Bienaymé-Chebyshev

♡

Propriété 2 (Inégalité de Markov). Soit X une variable aléatoire réelle discrète admettant une espérance. On a alors pour tout réel $\lambda > 0$:

$$\mathbb{P}(X \geq \lambda) \leq \frac{E(X)}{\lambda}$$

Exercice (Démonstration). On se propose de démontrer ici l'inégalité de Markov. On prend donc X une variable aléatoire avec $X(\Omega) = \{x_k | k \in \mathbb{N}\}$ admettant une espérance et λ un réel strictement positif.

1. Rappeler la définition de « X admet une espérance», et écrire cette espérance en fonction des x_k .
2. On note $\mathbb{1}_\lambda$ la fonction qui vaut 1 sur $[\lambda; +\infty[$ et 0 sur $]-\infty; \lambda[$. Montrer :

$$E(X) \geq \sum_{k=0}^{+\infty} x_k \mathbb{1}_\lambda(x_k) \mathbb{P}(X = x_k) \geq \sum_{k=0}^{+\infty} \lambda \mathbb{1}_\lambda(x_k) \mathbb{P}(X = x_k)$$

3. En déduire enfin : $E(X) \geq \lambda \mathbb{P}(X \geq \lambda)$ et conclure.

Remarque. Cette inégalité ne demande que très peu de connaissances sur X et sa répartition : elle permet une majoration de $\mathbb{P}(X \geq \lambda)$ très universelle mais donc nécessairement imprécise.

Exercice. Une chaîne de production produit en moyenne 5 pièces par jour.

1. Quelle est la probabilité qu'elle en produise un certain jour plus de 8 ? (on donnera seulement une majoration)
2. Supposons qu'on **sait** que le nombre de pièces produites suit une loi $\mathcal{B}(10, \frac{1}{2})$. Dans ce cas, quelle est cette probabilité ? En donner une valeur approchée.

♡

Propriété 3. Soit X une variable aléatoire réelle discrète qui admet **une espérance et une variance**. Soit $\varepsilon > 0$. Alors :

$$\mathbb{P}(|X - E(X)| \geq \varepsilon) \leq \frac{V(X)}{\varepsilon^2}$$

Exercice (Démonstration). Soit X une v.a.r.d. admettant une variance et $\varepsilon > 0$.

1. Rappeler la définition de la variance de X
2. Montrer que $\mathbb{P}(|X - E(X)| \geq \varepsilon) = \mathbb{P}((X - E(X))^2 \geq \varepsilon^2)$
3. En utilisant l'inégalité de Markov, montrer enfin l'inégalité de Bienaymé-Chebyshev.

Exercice. On reprend la chaîne de production précédente. Supposons qu'en plus d'avoir une moyenne de 5 pièces par jour, la variance du nombre de pièces produites par jour est de 1. Que peut-on dire maintenant de la probabilité de produire 8 pièces au moins ? *Indication : on pourra commencer par écrire : $\mathbb{P}(X \geq 8) \leq \mathbb{P}((X \geq 8) \cup (X \leq 2))$*

Exercice (Vers la loi des grands nombres). On possède un dé dont on suppose qu'il est truqué : la face 6 «semble» tomber plus souvent que les autres. En effet, en ayant lancé 1020 fois le dé (les lancers sont supposés équilibrés), on a obtenu 270 fois le chiffre 6. On cherche à montrer que **si le dé était équilibré, ceci serait peu probable**.

Remarque : $1020 = 6 \times 170$

1. On note pour chaque entier $n = 1, \dots, 1020$ $X_n = \begin{cases} 1 & \text{si le } n\text{-ième dé donne 6} \\ 0 & \text{sinon} \end{cases}$.
En supposant que le dé soit équilibré, quelle loi suit X_n ? Préciser alors l'espérance et la variance de X_n
2. On note $S_n = \sum_{k=1}^n X_k$. Que représente S_n ? Donner la valeur de S_{1020} dans notre expérience, préciser sa loi, son espérance et sa variance.
3. En utilisant l'inégalité de Bienaymé-Chebyshev, majorer $\mathbb{P}(S_{1020} \geq 270)$. Le fait d'obtenir 270 fois le chiffre 6 est-il fréquent si le dé est équilibré ?
4. Si on suppose que X_1 suit une loi de Bernoulli avec un autre paramètre, quel paramètre proposeriez-vous ?

II. Loi faible des grands nombres

♥

Théorème 4. Soit (X_n) une suite de variables aléatoires **indépendantes et de même loi** qui admettent une espérance $E(X_1) = m$ et une **variance** $V(X_1) = \sigma^2$. Soit $\overline{X}_n = \frac{1}{n} \sum_{i=1}^n X_i$ et $\varepsilon > 0$.

$$\mathbb{P}(|\overline{X}_n - m| \geq \varepsilon) \xrightarrow{n \rightarrow +\infty} 0$$

♠

Remarque. On dira (l'année prochaine) que $\overline{X}_n$ **converge en probabilité** vers m .

Exercice (Démonstration).

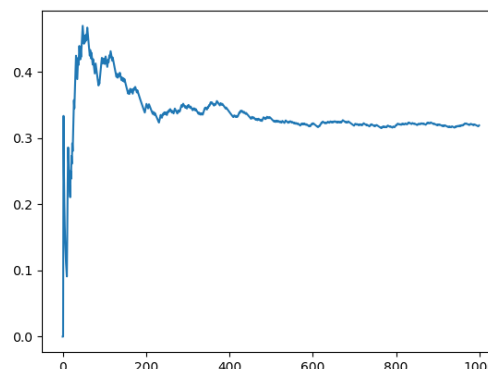
1. Justifier que pour tout entier n , $\overline{X}_n$ admet une variance, donner l'espérance et la variance de $\overline{X}_n$
2. En appliquant l'inégalité de Bienaymé-Chebyshev, majorer $\mathbb{P}(|\overline{X}_n - m| \geq \varepsilon)$ puis démontrer la propriété précédente.

Exercice. On lance 1000 fois une pièce truquée qui tombe avec une probabilité p sur Pile. On obtient k fois la face Pile. Justifier à l'aide du théorème précédent pourquoi on affirme « p est proche de $\frac{k}{1000}$ ».

- Exercice.*
1. Implémenter une fonction `lance_pieces()` qui renvoie 1 avec une probabilité p et 0 avec une probabilité $1 - p$
 2. Compléter le code suivant, pour qu'il demande un entier n à l'utilisateur et permette de calculer $\overline{X}_k$ pour $k = 1, \dots, n$ et d'afficher les valeurs successives

```
import matplotlib.pyplot as plt
n = .....
..... = []
s = .....
for k in range(1,n+1):
    s = s + .....
    liste_de_moyennes.append(s/k)
.....(liste_de_moyennes)
.....()
```

3. Justifier qu'à la ligne 5 on choisisse les arguments 1 et $n+1$ dans le `range`
4. On lance le programme avec $n = 1000$ et on obtient le graphique ci-contre.
Donner une valeur approchée de p , et montrer graphiquement pour $\varepsilon = 0,01$, ce que représente l'information « $|\overline{X}_n - m| < \varepsilon$ avec grande probabilité»





Remarque. On observe sur le graphique précédent que la suite X_n semble même converger vers m . La **loi faible des grands nombres** porte en fait son nom en référence à la :

Théorème 5 (Loi forte des grands nombres). Soit (X_n) une suite de variables aléatoires indépendantes et de même loi qui admettent une espérance $E(X_1) = m$ et $\overline{X_n}$ défini comme précédemment. Alors

$$\mathbb{P}(\overline{X_n} \xrightarrow[n \rightarrow +\infty]{} m) = 1$$

Ce résultat n'est pas au programme.

III. Exercices d'application

Exercice. On considère X une variable aléatoire possédant une espérance de 6 et une variance de 6. Appliquer, lorsque c'est possible, l'inégalité de Bienaymé-Chebyshev, pour majorer ou minorer les probabilités des événements suivants :

- | | | |
|----------------------|---------------------|----------------|
| 1. $2 \leq X \leq 6$ | 3. $X \leq 7$ | 5. $X \geq 11$ |
| 2. $5 < X < 7$ | 4. $ X - 6 \geq 1$ | 6. $X \geq 4$ |

Exercice. Soit X une variable aléatoire suivant une loi géométrique de paramètre $\frac{1}{n}$ ($n \in \mathbb{N}^*$)

- Montrer que $\mathbb{P}(X \geq n^2) \leq \frac{1}{n}$
- Montrer que $\mathbb{P}(|X - n| \geq n) \leq 1 - \frac{1}{n}$

Exercice (Théorème de Bernoulli). Soit $p \in]0; 1[$ et $(X_n)_{n \in \mathbb{N}^*}$ une suite de variables aléatoires définies sur un même $(\Omega, \mathcal{A}, \mathbb{P})$ telles que pour tout $n \in \mathbb{N}^*$, $X_n \hookrightarrow \mathcal{B}(n; p)$. Notons pour tout entier $n \in \mathbb{N}^*$: $Z_n = \frac{X_n}{n}$.

Montrer que pour tout $\varepsilon > 0$,

$$\mathbb{P}(|Z_n - p| \geq \varepsilon) \xrightarrow[n \rightarrow +\infty]{} 0$$

Exercice («Surréservation», issu de Bibmath). Une compagnie aérienne exploite un avion Paris-Montréal d'une capacité de 150 places. Chaque passager ayant un billet se présente avec une probabilité de $p = 0,75$. La compagnie souhaite vendre n billets ($n > 150$) en limitant à 5% le risque d'avoir plus de 150 personnes se rendant à l'embarquement. On supposera que $np > 150$ et on définit S_n la variable aléatoire représentant le nombre de personnes se présentant à l'embarquement.

- Quelle est la loi de S_n ?
- En remarquant $S_n \geq 150 \subset (|S_n - np| \geq 150 - np)$, montrer :

$$\mathbb{P}(S_n \geq 150) \leq \frac{np(1-p)}{(150 - np)^2}$$

- Résoudre sur $]0; 150[$ l'inéquation $\frac{x(1-p)}{(150-x)^2} \leq 0,05$
- Combien la compagnie peut-elle vendre de billets ?

Exercice (Inégalité de Chernoff). On lance une pièce de monnaie équilibrée une infinité de fois, on note $X_k = -1$ si le k -ième lancer donne pile et $X_k = 1$ sinon. Pour $n \in \mathbb{N}^*$, on note $S_n = \frac{1}{n} \sum_{k=1}^n X_k$

- En utilisant la loi faible des grands nombres, montrer :

$$\mathbb{P}(|S_{10000}| \geq 0,1) \leq 0,01$$

- Soit X une v.a.r.d. telle que pour tout $t \geq 0$, $\exp(tX)$ admet une espérance. On note pour tout $t \in \mathbb{R}_+$: $H(t) = \ln(E(e^{tX}))$.

Montrer que pour tout $a \in \mathbb{R}$, pour tout $t \geq 0$, $\mathbb{P}(X \geq a) \leq \exp(H(t) - ta)$

- En appliquant le résultat précédent à S_n , montrer qu'avec la même notation $H(t) = n \ln \left(\frac{e^{\frac{t}{n}} + e^{-\frac{t}{n}}}{2} \right)$

- En déduire (calculatrice ...) $\mathbb{P}(|S_{10000}| \geq 0,1) \leq 3,6 \times 10^{-22}$

On a trouvé une approximation beaucoup plus précise, sans outils supplémentaires (mais avec une hypothèse plus forte sur X) !