TD8 - Inégalités probabilistes, LGN et TCL

Inégalité de Markov et de Bienaymé-Tchebychev

Exercice 1 *(Comparaison des inégalités de Markov et de Bienaymé-Tchebychev)*

Le nombre de pièces sortant d'une usine en une journée est une variable aléatoire d'espérance 50. On veut estimer la probabilité que la production d'un jour donné dépasse 75 pièces.

En utilisant l'inégalité de Markov, quelle estimation obtient-on sur cette probabilité ? Que peut-on dire de plus sur cette probabilité si on sait que la variance de la production quotidienne est 25 ?

Exercice 2 *(Pièces défectueuses)*

Une usine fabrique des pièces dont une proportion inconnue $p \in]0,1[$ est défectueuse, et on souhaite trouver une valeur approchée de p. On effectue un prélèvement de n pièces. On suppose que le prélèvement se fait sur un échantillon très grand, et donc qu'il peut s'apparenter à une suite de n tirages indépendants avec remise. On note S_n la variable aléatoire égale au nombre de pièces défectueuses et on souhaite quantifier le fait que $\frac{S_n}{n}$ approche p.

a) Quelle est la loi de S_n ? Sa moyenne ? Sa variance ?

b) Démontrer que, pour tout $\varepsilon > 0$

$$\mathbb{P}\left(\left|\frac{S_n}{n} - p\right| \geqslant \varepsilon\right) \leqslant \frac{1}{4n\varepsilon^2}.$$

c) En déduire une condition sur n pour que S_n/n soit une valeur approchée de p à 10^{-2} près avec une probabilité supérieure ou égale à 95%.

Exercice 3 *(Surréservation et inégalité de Bienaymé-Tchebychev)*

Une compagnie aérienne exploite un avion Paris-Montréal d'une capacité de 150 places. Pour ce vol, une analyse statistique a montré qu'un passager ayant réservé son billet se présentait à l'embarquement avec une probabilité de $p = 0{,}75$. La compagnie souhaite optimiser le remplissage de l'avion et souhaite vendre n billets, avec $n > 150$, mais en limitant le risque que plus de 150 personnes se rendent à l'embarquement à moins de 5%. On supposera dans la suite que $np < 150$. On définit la variable aléatoire S_n comme le nombre de personnes, parmi les n ayant réservé un billet, se présentant à l'embarquement.

a) Quelle est la loi de S_n ?

b) En appliquant l'inégalité de Tchebychev à S_n, démontrer que

$$P(S_n \geqslant 150) \leqslant \frac{np(1-p)}{(150-np)^2}.$$

c) Résoudre sur $]0,150[$ l'inéquation $\frac{x(1-p)}{(150-x)^2} \leqslant 0{,}05$.

d) Combien la compagnie peut-elle vendre de billets tout en s'assurant que la probabilité que plus de 150 clients se présentent à l'embarquement est inférieure ou égale à 5% ?

Exercice 4 *(Loi faible des grands nombres)*

Soit $(X_n)_{n\geqslant 1}$ une suite de variables aléatoires indépendantes, de même espérance m et de même variance σ^2. On pose

$$\overline{X}_n = \frac{X_1 + \cdots + X_n}{n}.$$

Démontrer, à l'aide de l'inégalité de Bienaymé-Tchebychev, que pour tout $\varepsilon > 0$,

$$\lim_{n\to\infty} \mathbb{P}\left(\left|\overline{X}_n - m\right| \geqslant \varepsilon\right) = 0.$$

Loi forte des grands nombres et théorème central limite

Exercice 5 *(Autour de la loi binomiale)*

Soient $B_1, \ldots, B_n$ des v.a. de Bernoulli indépendantes de paramètre $p \in [0,1]$ et f une fonction continue sur $[0,1]$.

a) Quelle est la distribution de la variable aléatoire $S_n \coloneqq B_1 + \cdots + B_n$?

b) Exprimer $\mathbb{E}\left(f\left(\frac{S_n}{n}\right)\right)$ en fonction de p, n, et f.

c) Après avoir justifié l'existence de la limite, calculer

$$\lim_{n\to\infty} \sum_{k=0}^{n} \binom{n}{k} p^k (1-p)^{n-k} f\left(\frac{k}{n}\right).$$

Exercice 6 *(Loi forte et loi de Poisson)*

a) Soient X et Y deux variables aléatoires discrètes suivant chacune des lois de Poisson indépendantes de paramètres λ_X et λ_Y. Soit $k \in \mathbb{N}$, calculer $\mathbb{P}(X+Y) = k$. Que peut-on en déduire sur la distribution de la variable aléatoire $Z = X + Y$?

b) Soient $X_1, \ldots X_n$ des variables aléatoires de loi de Poisson, indépendantes, de même paramètre $\lambda > 0$.

(i) Déduire de la question précédente la distribution de la variable aléatoire $Z_n = \sum_{k=1}^n X_k$.

(ii) Soit $f : \mathbb{R}_+ \to \mathbb{R}$ une fonction continue et bornée. Après avoir justifié son existence, calculer

$$\lim_{n\to\infty} \sum_{k=0}^{+\infty} e^{-n\lambda} \frac{(n\lambda)^k}{k!} f\left(\frac{k}{n}\right).$$

Exercice 7 *(TCL et loi de Poisson)*

Soit $(X_i)_{i\geqslant 1}$ une suite de variables aléatoires i.i.d. de Poisson de paramètre $\lambda = 1$.

a) Soit $Z_n = X_1 + \cdots + X_n$ et F_n sa fonction de répartition.

(i) Pour tout n entier, calculer $F_n(n)$.

(ii) Calculer après, avoir justifié son existence,

$$\lim_{n\to\infty} e^{-n} \sum_{k=0}^{n} \frac{n^k}{k!}.$$

Indication : Montrer que $Z_n \sim \mathcal{P}(n)$.

b) Soit f une fonction continue et bornée sur $\mathbb{R}$. Après avoir justifié l'existence de la limite, montrer que

$$\lim_{n\to\infty} e^{-n} \sum_{k\in\mathbb{N}} f\left(\frac{k-n}{\sqrt{n}}\right) \frac{n^k}{k!} = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} f(t) e^{-\frac{t^2}{2}} dt.$$

Exercice 8 *(Moyenne géométrique)*

Soit $(U_n)_{n\geqslant 1}$ une suite de variables aléatoires i.i.d. uniformes sur $[0,1]$, on définit $X_n = \left(\prod_{j=1}^{n} U_j\right)^{1/n}$.

a) Que vaut $\log X_n$?

b) Soit $j \in \mathbb{N}$ fixé, calculer pour tout $x \in \mathbb{R}$, $\mathbb{P}\left(-\log U_j \leqslant x\right)$. En déduire la loi de $-\log U_j$.

c) Montrer que X_n converge p.s. quand $n \to \infty$, et déterminer sa limite.

d) En utilisant le théorème central limite, calculer

$$\lim_{n\to\infty} \mathbb{P}\left(X_n \leqslant e^{-1}\right).$$

Pour approfondir

Exercice 9 *(Une variante de la loi faible des grands nombres)*

Soit $(X_n)_{n\geqslant 1}$ une suite de variables aléatoires deux à deux indépendantes. On suppose que chaque X_n suit une loi de Bernoulli de paramètre p_n. On note $S_n = X_1 + \cdots + X_n$ et on souhaite démontrer que, pour tout $\varepsilon > 0$,

$$\lim_{n\to+\infty} P\left(\left|\frac{S_n}{n} - \frac{1}{n}\sum_{k=1}^{n} p_k\right| \geqslant \varepsilon\right) = 0.$$

Pourquoi ne peut-on pas appliquer directement la loi faible des grands nombres ? Quelle est l'espérance de S_n ? Sa variance ? Démontrer que $V(S_n) \leqslant n$. En déduire le résultat.

Exercice 10 *(Un problème chinois !)*

On suppose qu'à la naissance, la probabilité qu'un nouveau-né soit un garçon est égale à $1/2$. On suppose que tous les couples ont des enfants jusqu'à obtenir un garçon. On souhaite évaluer la proportion de garçons dans une génération de cette population. On note X le nombre d'enfants d'un couple pris au hasard dans la population.

Donner la loi de la variable aléatoire X. On suppose qu'une génération en âge de procréer est constituée de N couples, et on note $X_1, \cdots, X_N$ le nombre d'enfants respectif de chaque couple. On note enfin P la proportion de garçons issus de cette génération. Exprimer P en fonction de $X_1, \ldots, X_N$. Quelle est la limite de P lorsque N tend vers l'infini. Qu'en pensez-vous ?

Exercice 11 *(Théorème de Bernstein-Weierstrass)*

Soit f une fonction continue de $[0,1]$ dans $\mathbb{C}$. Le n-ième polynôme de Bernstein de f , B_n, est défini par

$$B_n(x) = \sum_{k=0}^{n} \binom{n}{k} x^k (1-x)^{n-k} f(k/n).$$

a) Soit $S_n(x)$ la variable aléatoire définie par $S_n(x) = \mathrm{Bin}(n,x)/n$, où $\mathrm{Bin}(n,x)$ est une variable aléatoire de loi binomiale de paramètres n et x.

(i) Vérifier que $B_n(x) = \mathbb{E}\left(f(S_n(x))\right)$.

(ii) Calculer $\mathbb{E}\left(S_n(x)\right)$ et $\mathbb{V}\left(S_n(x)\right)$.

b) Montrer que pour tout $\eta > 0$ et pour tout $x \in [0,1]$

$$\mathbb{P}\left(|S_n(x) - x| \geqslant \eta\right) \leqslant \frac{1}{n\eta^2}.$$

c) On rappelle que f étant continue sur un segment, elle est également uniformément continue.

(i) Rappeler la définition de l'uniforme continuité.

(ii) Montrer que pour tout $\eta > 0$

$$|B_n(x) - f(x)| \leqslant \mathbb{E}\left(|f(x) - f\big(S_n(x)\big)|\mathbb{1}_{|S_n(x)-x|<\eta}\right) + 2\,\|f\|_\infty\, \mathbb{P}\left(|S_n(x) - x| \geqslant \eta\right).$$

(iii) En déduire que pour tout $\varepsilon > 0$, il existe $\eta > 0$ tel que

$$|B_n(x) - f(x)| \leqslant \varepsilon + 2\,\|f\|_\infty\, \mathbb{P}\left(|S_n(x) - x| \geqslant \eta\right).$$

d) Déduire des questions précédentes le Théorème de Bernstein-Weierstrass

$$\sup_{x\in[0,1]} |B_n(x) - f(x)| \underset{n\to\infty}{\to} 0.$$