Table des matières

Pré-requis

Objectifs

Introduction : répétition d'évènements indépendants

Exemple :
on répète une expérience incertaine sur une grande population de taille N_0, et on aimerait comprendre le phénomène « en moyenne. »

• **Pour le probabiliste**, il s'agit de déterminer un modèle aléatoire $(\Omega, \mathcal{A}, \mathbf{P})$ et une suite $(X_n)_{n\geq 1}$ de variables aléatoires indépendantes, représentant les différents résultats individuels.
Le comportement moyen d'un groupe de N variables est caractérisé par la variable aléatoire moyenne

$$S_N = \frac{X_1 + \dots X_N}{N}$$

Il s'agit de comprendre les lois de S_N pour $N \in [\![1, N_0]\!]$.

• **Pour le statisticien**, on dispose d'un échantillon $(x_1, \dots, x_N)$ correspondant à un sondage d'une partie de la population, et on aimerait tester son comportement moyen, à l'aide d'un modèle probabiliste, pour prédire l'état du système $(x_1, \dots, x_{N_0})$
Pour cela, on s'appuie sur la moyenne observée $\hat{x} = \dfrac{x_1 + \dots x_N}{N}$ sur l'échantillon sondé.

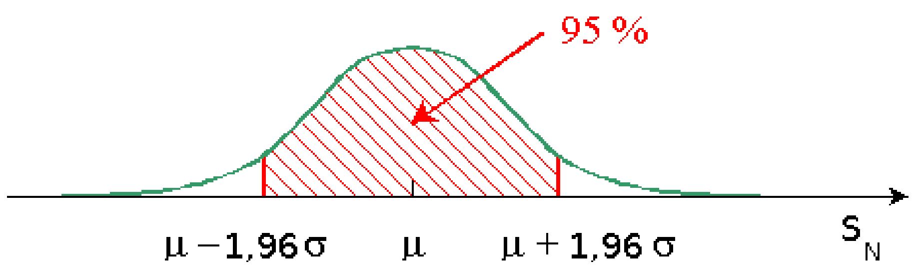
Réponse du programme de lycée :

Si les $(X_n)_{n\geq 1}$ sont indépendantes de même loi usuelle,
admettant une « espérance »$\mu \in \mathbb{R}$ et une « variance » $\sigma^2 > 0$, alors lorsque $N \to +\infty$
$S_N = \dfrac{X_1 + \dots X_N}{N}$ se comporte comme une variable gaussienne de densité $f_X : x \mapsto \dfrac{1}{\sqrt{2\pi\sigma^2}} e^{\frac{-(x-\mu)^2}{2\sigma^2}}$

On a

$$\mathbf{P}\left(\left\{ |S_N - \mu| \leq 1,96 \frac{\sigma}{\sqrt{N}} \right\}\right) \leq 0,95$$

ce qui signifie qu'avec une probabilité supérieur à 95%, $\left| \dfrac{X_1 + \dots X_N}{N} - \mu \right| \leq 1,96 \dfrac{\sigma}{\sqrt{N}}$



Interprétation : la moyenne empirique observée sur l'échantillon $\hat{x} = \dfrac{x_1 + \dots x_N}{N}$ doit être souvent proche de l'espérance μ, et dans l'intervalle de cofiance $I_{0,95} = [\mu - 1,96\sigma, \mu + 1,96\sigma]$ dans 95% des observations.

I. Variables centrées réduites, sommes finies de v.a.i.i.d.

I.1 Variable centrée réduite

Proposition 1.

Soit X une v.a. telle que X^2 admet une espérance et telle que $\sigma(X) > 0$.
Alors $N = \dfrac{X - \mathbb{E}[X]}{\sigma(X)}$ est centrée et réduite, c'est à dire
$\mathbb{E}[N] = 0$ (N centrée)
et
$\sigma(N) = 1$ (N réduite)

I.2 Somme de v.a.i.i.d.

Proposition 2.

Soient $X_1, \dots X_n$ des v.a.i.i.d. telles que X_1 admet une espérance.
En notant $m = \mathbb{E}(X_1)$ et $\sigma = \sigma(X_1)$ et soit $S_n = \sum_{i=1}^{n} X_i$.
Alors $\mathbb{E}[S_n] = nm$ et $\mathbb{V}[S_n] = n\sigma^2$, de sorte que $N = \dfrac{S_n - nm}{\sigma\sqrt{n}}$ est centrée réduite.

II. Comportement asymptotique

II.1 Inégalité de Markov

Proposition 3 (Inégalité de Markov).

Soit X une variable aléatoire discrète, d'espérance finie.
Alors pour tout $t > 0$, on a $\boxed{\mathbf{P}(|X| \geq t) \leq \dfrac{\mathbb{E}(|X|)}{t}}$

démonstration :
On note Y la v.a.r. définie pour tout $\omega \in \Omega$, on a $Y(\omega) = \begin{cases} t & \text{si } |X(\omega)| \geq t \\ 0 & \text{si } |X(\omega)| < t \end{cases}$
Ainsi Y est égale à t si $X \geq t$ et à 0 sinon, et se note $Y(\cdot) = t\ \mathbf{1}_{|X| \geq t}(\cdot)$.
On remarque que pour tout $\omega \in \Omega$, on a $\boxed{|X(\omega)| \geq Y(\omega)}$
(car si ω est tel que $|X(\omega)| \geq t$, alors $|X(\omega)| \geq = Y(\omega)$ et si ω est tel que $|X(\omega)| < t$, alors $|X(\omega)| \geq 0 = Y(\omega)$)
Y est à valeurs dans l'ensemble fini $\{0, 1\}$, donc admet une espérance.
$|X|$ admet une espérance, car la série $\sum |x_k| \mathbf{P}(\{X = x_k\})$ converge, puisque X admet une espérance.
D'après la croissance de l'espérance, on en déduit que :
$\boxed{\mathbb{E}(|X|) \geq \mathbb{E}(Y)}$

or $\mathbb{E}(Y) = 0 \times \mathbf{P}(\{Y=0\}) + t \times \mathbf{P}(\{Y=t\}) = t \times \mathbf{P}(\{Y=t\}) = t \times \mathbf{P}(\{|X| \geq t\})$
d'où $\mathbb{E}(|X|) \geq t\mathbf{P}(|X| \geq t)$, puis on divise par $t > 0$. $\square$

II.2 Inégalité de Bienaymé-Tchebychev

En corollaire de l'inégalité de Markov, on en déduit une seconde inégalité :

Corollaire 4 (inégalité de Markov pour X^2).

Soit X une variable aléatoire discrète, d'espérance et de variance finies.
Alors pour tout $t > 0$, on a $\mathbf{P}(|X| \geq t) \leq \dfrac{\mathbb{E}(X^2)}{t^2}$

démonstration : Markov appliqué à la v.a.r. X^2, en remarquant que $X^2 \geq t^2 \iff |X| \geq t$. $\square$

Proposition 5 (Inégalité de Bienaymé-Tchebychev).

Soit X une variable aléatoire discrète, d'espérance et de variance finies.
Alors pour tout $\varepsilon > 0$, on a : $\boxed{\mathbf{P}(|X - \mathbb{E}(X)| \geq \varepsilon) \leq \dfrac{\mathbb{V}(X)}{\varepsilon^2}}$

démonstration : 2ème inégalité de Markov appliqué à $(X - \mathbb{E}(X))$ $\square$

Remarque 1. La variance (ou l'écart-type) mesure donc la variation par rapport à la moyenne.

II.3 Application : Loi des grands nombres

Théorème 6 (Loi faible des grands nombres).

si $(X_n)_{n\geqslant 1}$ est une suite de variables aléatoires deux à deux indépendantes et de même loi admettant un moment d'ordre 2, alors, si $S_n = \sum_{k=1}^{n} X_k$, $m = \mathbb{E}(X_1)$ et $\sigma = \sigma(X_1)$, on a pour tout $\varepsilon > 0$,

$$\mathbf{P}\left(\left|\frac{1}{n}S_n - m\right| \geqslant \varepsilon\right) \xrightarrow[n\to\infty]{} 0.$$

démonstration : Conséquence de Bienaymé-Tchebychev pour $X = \dfrac{S_n}{n}$, variable d'espérance m et de variance $\dfrac{\sigma^2}{n}$, on a $\boxed{\mathbf{P}\left(\left|\frac{1}{n}S_n - m\right| \geqslant \varepsilon\right) \leq \dfrac{\frac{1}{n} \times \sigma^2}{\varepsilon^2}}$ $\square$

Remarque 2. Intervalle de confiance à 95% pour $n = 1000$ tirages $0,95 = \dfrac{\sigma^2}{n\varepsilon^2}$, donc $\varepsilon = \sqrt{0.95} * \sigma * n$ c'est peu précis.

En pratique celui obtenu par le théorème de la limite centrale est meilleur.

Remarque 3. (HP) Méthode de Monte-Carlo : pour approcher $\int_0^1 f$, on calcule $\frac{\sum_{i=1}^{N} f(X_i)}{N}$ pour des X_i de loi uniforme sur $[0,1]$...

III. Compléments

III.1 Positivité, croissance de l'espérance

En conséquence de la linéarité de l'espérance

Proposition 7.

Si $X_1, \ldots, X_n$ sont des v.a. discrètes admettant une espérance (indépendantes ou non), alors
$\mathbb{E}(X_1 + \cdots + X_n) = \sum_{i=1}^{n} \mathbb{E}(X_i)$

démonstration : Par récurrence sur $n \geq 2$. □

Proposition 8 (Positivité, croissance de l'espérance)**.**

Pour tous X, Y variables aléatoires et tout $\lambda \in \mathbb{R}$, on a :
- [Positivité] Si X est à valeurs positives et admet une espérance, alors $\mathbb{E}(X) \geq 0$
- [Croissance] Si X et Y admettent des espérances et si : $X(\omega) \leq Y(\omega),\ \forall \omega \in \Omega$, alors $\mathbb{E}(X) \leq \mathbb{E}(Y)$

démonstration : Pour la positivité de l'espérance on remarque que $\mathbb{E}(X) = \sum_{k=0}^{+\infty} x_k \mathbf{P}(\{X = x_k\})$ est la somme d'une série convergente positive.
Pour la croissance de l'espérance, il suffit d'appliquer le point précédent à $Y - X \geq 0$ et par linéarité $\mathbb{E}(Y) - \mathbb{E}(X) \geq 0$.

Proposition 9.

Si $|X| \leq Y$ et si Y admet une espérance, alors X estd 'espérance finie et $\mathbb{E}[|X|] \leq \mathbb{E}[Y]$.

III.2 Compléments

Proposition 10.

Si X est positive d'espérance nulle, alors $\{X = 0\}$ est presque-sûr.

Proposition 11.

Si X est à valeurs dans $\mathbb{N}$ admet une espérance, alors $\mathbb{E}(X) = \sum_{n=1}^{+\infty} \mathbf{P}(X \geqslant n)$.

dém : on découpe et on intervertit les sommes convergentes

$$\mathbb{E}(X) = \sum_{k=1}^{+\infty} k\ \mathbf{P}(X = k) = \sum_{k=1}^{+\infty} \left(\sum_{n=1}^{k} \mathbf{P}(X = k) \right)$$

$$\underset{\text{somme série positive}}{=} \sum_{1 \leq n \leq k}^{+\infty} \mathbf{P}(X = k) \underset{interversion}{=} \sum_{n=1}^{+\infty} \left(\sum_{k=n}^{+\infty} \mathbf{P}(X = k) \right) = \sum_{n=1}^{+\infty} \mathbf{P}(X \geqslant n).\ \square$$

Proposition 12 (Inégalité de Cauchy-Schwarz)**.**

Si X et Y sont deux variables aléatoires discrètes telles que X^2 et Y^2 admet des espérance, alors $(\mathbb{E}(XY))^2 \leq \mathbb{E}(X^2)\ \mathbb{E}(Y^2)$.

III.3 Variance d'une somme

Proposition 13.

Si $X_1, \ldots, X_n$ sont des v.a. discrètes admettant une variance, alors

$$\mathbb{V}(X_1 + \cdots + X_n) = \sum_{i=1}^{n} \mathbb{V}(X_i) + \sum_{i \neq j} (E(X_i X_j) - E(X_i)E(X_j))$$

$$= \sum_{i=1}^{n} \mathbb{V}(X_i) + 2 \sum_{1 \leq i < j \leq n} (E(X_i X_j) - E(X_i)E(X_j))$$

démonstration : Calcul direct, par récurrence sur n. $\square$

. $\square$

NOUVEAU Programme PC 2022 :

Contenus	Capacités & commentaires
d) Inégalités probabilistes	
Inégalité de Markov. Inégalité de Bienaymé-Tchebychev. Loi faible des grands nombres : si $(X_n)_{n\geqslant 1}$ est une suite i.i.d. de variables aléatoires de variance finie, alors en notant $S_n = \sum_{k=1}^{n} X_k$ et $m = \mathrm{E}(X_1)$, pour tout $\varepsilon > 0$: $P\left(\left\lvert \frac{S_n}{n} - m \right\rvert \geqslant \varepsilon\right) \xrightarrow[n\to+\infty]{} 0.$	Les étudiants doivent savoir retrouver, avec $\sigma = \sigma(X_1)$: $P\left(\left\lvert \frac{S_n}{n} - m \right\rvert \geqslant \varepsilon\right) \leqslant \frac{\sigma^2}{n\varepsilon^2}.$

Programme PC :

Probabilités

Les chapitres de probabilités permettent de développer les compétences suivantes :
— *modéliser des situations aléatoires par le choix d'un espace probabilisé ou de variables aléatoires adéquats ;*
— *maîtriser un formalisme spécifique aux probabilités.*

3.a) A- Espaces probabilisés

Cette partie a pour objectif la mise en place du cadre général de la théorie des probabilités permettant d'aborder l'étude de processus stochastiques à temps discret. Cette mise en place se veut minimale. En particulier :
— *la notion de tribu ne doit donner lieu à aucun développement théorique autre que sa définition ;*
— *la construction d'espaces probabilisés n'est pas un objectif du programme.*

Contenus	Capacités & commentaires
a) Ensembles dénombrables	
Un ensemble est dit dénombrable s'il est en bijection avec $\mathbb{N}$. Ensembles finis ou dénombrables. Dénombrabilité de $\mathbb{Z}$, d'un produit cartésien de deux ensembles dénombrables.	Un ensemble fini ou dénombrable peut être décrit en extension sous la forme $\{x_n\,;\ n \in \mathbb{N}\}$. Toute autre connaissance sur la dénombrabilité est hors programme.
b) Espace probabilisé	
Si Ω est un ensemble, on appelle *tribu* sur Ω une partie $\mathcal{A}$ de l'ensemble $\mathcal{P}(\Omega)$ des parties de Ω telle que : 1. $\Omega \in \mathcal{A}$, 2. pour tout $A \in \mathcal{A}$, $\overline{A} = \Omega \setminus A \in \mathcal{A}$, 3. pour toute suite $(A_n)_{n\geqslant 0}$ d'éléments de $\mathcal{A}$, la réunion $\bigcup\limits_{n=0}^{+\infty} A_n$ appartient à $\mathcal{A}$. Si Ω est un ensemble et $\mathcal{A}$ une tribu sur Ω, on appelle *probabilité* sur $(\Omega, \mathcal{A})$ une application $\mathbf{P} \colon \mathcal{A} \to [0,1]$ telle que : 1. $\mathbf{P}(\Omega) = 1$, 2. pour toute suite $(A_n)_{n\geqslant 0}$ d'événements incompatibles, $\mathbf{P}\left(\bigcup\limits_{n=0}^{+\infty} A_n\right) = \sum\limits_{n=0}^{+\infty} P(A_n)$. On appelle espace probabilisé un triplet $(\Omega, \mathcal{A}, P)$ où $\mathcal{A}$ est une tribu et $\mathbf{P}$ une probabilité sur $(\Omega, \mathcal{A})$.	L'ensemble Ω est l'univers ; il n'est en général pas précisé. Les éléments de $\mathcal{A}$ sont les événements. Les étudiants doivent savoir expliciter un événement à partir d'autres événements en utilisant la réunion, l'intersection et le complémentaire. On fait le parallèle entre le vocabulaire probabiliste et le vocabulaire ensembliste.

Contenus	Capacités & Commentaires

Propriétés :

— $\bigcap_{n=0}^{+\infty} A_n \in \mathcal{A}$.

— Continuité croissante : si $(A_n)_{n\geqslant 0}$ est une suite d'événements telle que, pour tout n, on ait $A_n \subset A_{n+1}$, alors :

$$\lim_{n\to+\infty} P(A_n) = P\left(\bigcup_{n=0}^{+\infty} A_n\right).$$

— Continuité décroissante : si $(A_n)_{n\geqslant 0}$ est une suite d'événements telle que, pour tout n, on ait $A_{n+1} \subset A_n$, alors :

$$\lim_{n\to+\infty} P(A_n) = P\left(\bigcap_{n=0}^{+\infty} A_n\right).$$

— Sous additivité : si $(A_n)_{n\in\mathbb{N}}$ est une suite d'événements, alors :

$$P\left(\bigcup_{n=0}^{+\infty} A_n\right) \leqslant \sum_{n=0}^{+\infty} P(A_n).$$

c) Conditionnement et indépendance

Si A et B sont deux événements tels que $\mathbf{P}(B) > 0$, on appelle probabilité conditionnelle de A sachant B le réel

$$P_B(A) = \frac{P(A\cap B)}{P(B)}$$

Notation $\mathbf{P}_B(A) = P(A \mid B)$. L'application $\mathbf{P}_B$ est une probabilité sur $(\Omega, \mathcal{A})$.
Ce paragraphe étend rapidement les concepts et résultats vus en première année dans le cadre des univers finis.

Formule des probabilités composées.
Système complet dénombrable d'événements.
Formule des probabilités totales : si $(A_n)_{n\in\mathbb{N}}$ est un système complet d'événements, alors la série $\sum P(B\cap A_n)$ converge et

$$P(B) = \sum_{n=0}^{+\infty} P(B\cap A_n) = \sum_{n=0}^{+\infty} P(B \mid A_n)\ P(A_n)$$

On adopte la convention $\mathbf{P}(B \mid A_n)P(A_n) = 0$ lorsque $\mathbf{P}(A_n) = 0$.
La formule reste valable dans le cas d'une suite $(A_n)_{n\in\mathbb{N}}$ d'événements deux à deux incompatibles tels que $\sum_{n=0}^{+\infty} P(A_n) = 1$.

Formule de Bayes.

Contenus	Capacités & Commentaires
Indépendance de deux événements.	Si $\mathbf{P}(B) > 0$, l'indépendance de A et B équivaut à $\mathbf{P}(A \mid B) = P(A)$.
Indépendance mutuelle d'une famille finie d'événements.	L'indépendance des événements A_i deux à deux n'entraîne pas leur indépendance mutuelle si $n \geqslant 3$.

d') Généralités (sur les variables aléatoires)

Une variable aléatoire discrète X sur $(\Omega, \mathcal{A})$ est une application définie sur Ω dont l'image est finie ou dénombrable et telle que l'image réciproque de tout élément de $X(\Omega)$ appartient à $\mathcal{A}$.	Pour tout $U \subset X(\Omega)$, $X^{-1}(U)$ est un événement.
Notations $(X \in U)$, $\{X \in U\}$.	
Loi d'une variable aléatoire discrète.	
Si X prend ses valeurs dans $\{ x_n \ ; \ n \geqslant 0 \}$, les x_n étant distincts, et si $(p_n)_{n \geqslant 0}$ est une suite de réels positifs vérifiant $\displaystyle\sum_{n=0}^{+\infty} p_n = 1$, alors il existe une probabilité P sur $(\Omega, \mathcal{A})$ telle que $P(X = x_n) = p_n$ pour tout $n \in \mathbb{N}$.	Démonstration hors programme.

d) Lois usuelles

Pour p dans $]0, 1[$, loi géométrique de paramètre p : la variable aléatoire X suit une loi géométrique de paramètre p si et seulement si $$\forall k \in \mathbb{N}^*, \qquad P(X = k) = p(1-p)^{k-1}.$$	Notation $X \hookrightarrow \mathcal{G}(p)$. La loi géométrique peut être interprétée comme rang du premier succès dans une suite illimitée d'épreuves de Bernoulli indépendantes et de même paramètre p.
Série génératrice, espérance et variance.	
Loi de Poisson de paramètre λ.	
$\leftrightarrows$ PC : compteur Geiger.	

b) Espérance et variance

La variable aléatoire réelle discrète X à valeurs dans un ensemble dénombrable $\{x_n \ ; \ n \geqslant 0\}$ est dite d'espérance finie si la série $\sum x_n P(X = x_n)$ est absolument convergente ; si tel est le cas, on appelle espérance de X, noté $\mathbb{E}(X)$, le réel $\displaystyle\sum_{n=0}^{+\infty} x_n P(X = x_n)$.	On admet que la somme $\displaystyle\sum_{n=0}^{+\infty} x_n P(X = x_n)$ ne dépend pas de l'ordre d'énumération. $\leftrightarrows$ PC : énergie moyenne de systèmes à spectre discret.
Si X est à valeurs dans $\mathbb{N}$, alors $\mathbb{E}(X) = \displaystyle\sum_{n=1}^{+\infty} P(X \geqslant n)$.	

Contenus	Capacités & Commentaires
Théorème du transfert : si X est une variable aléatoire et f une application à valeurs réelles définie sur l'image $\{x_n,\ n \in \mathbb{N}\}$ de X, alors $f(X)$ est d'espérance finie si et seulement si la série $\Sigma P(X = x_n)\, f(x_n)$ converge absolument. Dans ce cas, on a : $$\mathbb{E}(f(X)) = \sum_{n=0}^{+\infty} P(X = x_n)\, f(x_n).$$	Démonstration hors programme.