Fonction caractéristique et théorème central limite

*On présente ici la notion de **fonction caractéristique** d'une variable aléatoire. Si elle nous permet de caractériser les variables aléatoires usuelles, qu'elles soient discrètes ou continues, nous verrons que cette notion permet aussi d'aller chercher un résultat fondamental : le **théorème central limite** dont les applications peuvent être très pratiques, puisque le calcul d'une probabilité se ramène alors au calcul de l'aire sous une distribution gaussienne.*

Définition Soit $(\Omega, \mathcal{A}, P)$ un espace probabilisé et considérons X une variable aléatoire réelle définie sur Ω. On appelle **fonction caractéristique** de X l'application définie par :

$$\phi_X : t \longmapsto E(e^{iXt})$$

En particulier, si X désigne une variable aléatoire discrète telle que $X(\Omega) = \{x_k,\ k \in I\}$, alors ϕ_X désigne la somme d'une série de fonctions définie par :

$$\phi_X(t) = \sum_{k \in I} p_k e^{ix_k t}$$

Remarque D'ailleurs, à t fixé dans $\mathbb{R}$, on remarque que pour tout $k \in I$, $|p_k e^{ix_k t}| = p_k$ et ainsi, la série converge absolument de sorte que ϕ_X est définie sur $\mathbb{R}$ tout entier.

1. Pour chacune des lois suivantes, donner l'expression de $\phi_X(t)$:

$$X \sim B(p),\ X \sim P(\lambda),\ X \sim G(p)$$

2. (a) Soient $X_1, \ldots, X_n$ des variables aléatoires discrètes mutuellement indépendantes et on note $S_n = X_1 + \ldots + X_n$. Montrer que :

$$\phi_{S_n}(t) = \prod_{k=1}^{n} \phi_{X_k}(t)$$

   (b) On suppose que $X \sim B(n, p)$. Donner alors l'expression de sa fonction caractéristique.

On peut même généraliser cette notion aux **variables aléatoires réelles continues** que vous avez rencontrées en terminale. Ces variables aléatoires sont généralement définies par une fonction de densité f continue par morceaux, positive et intégrable sur $\mathbb{R}$ telle que $\displaystyle\int_{\mathbb{R}} f(x)\, dx = 1$ et pour laquelle :

$$P(Y \in [a, b]) = \int_a^b f(x)\, dx$$

Dans ce cas, on adapte les notations et la fonction caractéristique est donnée sur $\mathbb{R}$ par l'intégrale à paramètre :

$$\phi_X(t) = \int_{\mathbb{R}} e^{ixt} f(x)\, dx$$

Remarque On reconnaît ici la **transformée de Fourier** de f... celle-ci est d'ailleurs bien définie sur $\mathbb{R}$ puisque f est intégrable.

3. (a) On dit que la variable aléatoire réelle Y suit la **loi exponentielle** de paramètre $\lambda > 0$ si elle admet pour densité la fonction :

$$f : x \longmapsto \begin{cases} \lambda e^{-\lambda x} \text{ si } x \geq 0 \\ 0 \text{ sinon} \end{cases}$$

   Calculer $\phi_Y(t)$.

   (b) On dit que la variable aléatoire réelle Y suit la **loi normale** $N(0, 1)$ si elle admet pour densité la fonction :

$$f : x \longmapsto \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$$

   Calculer $\phi_Y(t)$. *On pourra encore montrer que ϕ_Y est solution d'une équation différentielle linéaire...*

4. On se place dans la cas particulier où X est une variable aléatoire discrète vérifiant $X(\Omega) = \mathbb{N}$.
   Etablir que pour tout $n \in \mathbb{N}$,

$$P(X = n) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \phi_X(t) e^{-int}\, dt$$

Remarque On pourra donc retenir que la connaissance de ϕ_X caractérise la loi de X : c'est en cela qu'elle désigne une **fonction caractéristique**. Bien entendu, ce résultat se généralise mais il repose en partie sur l'inversion de la transformée de Fourier, qui n'est pas simple à obtenir.

5. En utilisant la fonction caractérisique, établir que la somme d'un nombre fini de variables aléatoires indépendantes et suivant une loi de Poisson suit encore une loi de Poisson dont on précisera le paramètre.

Pour finir, on admet alors le **théorème de Paul Lévy** :

Théorème 1 (de Lévy).

Soient $(\Omega, \mathcal{A}, P)$ un espace probabilisé et $(X_n), X$ des variables aléatoires réelles définies sur Ω. Alors,

$$X_n \xrightarrow{loi} X \Leftrightarrow \forall\, t \in \mathbb{R},\ \phi_{X_n}(t) \underset{n\to+\infty}{\longrightarrow} \phi_X(t)$$

6. Soit (X_n) une suite de variables aléatoires mutuellement indépendantes et identiquement distribuées. On suppose de plus que les variables aléatoires X_n possèdent un moment d'ordre 2, et on pose pour tout $n \in \mathbb{N}^*$:

$$S_n = \sum_{k=1}^{n} X_k \text{ et } X_n^* = \frac{S_n - E(S_n)}{\sqrt{V(S_n)}}$$

   (a) En notant $m = E(X_i)$ et $\sigma = \sqrt{V(X_i)}$, justifier que $X_n^* = \dfrac{1}{\sqrt{n}}(Y_1 + \ldots + Y_n)$ avec Y_i des variables aléatoires centrées réduites et indépendantes. Montrer dans ce cas que :

$$\phi_{X_n^*}(t) = \left(\phi_{Y_1}(\frac{t}{\sqrt{n}})\right)^n$$

   (b) Justifier rapidement que $Y_1 \in L^2$, puis établir que $\phi'_{Y_1}(0) = 0$ et $\phi''_{Y_1}(0) = -1$ de sorte qu'au voisinage de 0 :

$$\phi_{Y_1}(t) = 1 - \frac{t^2}{2} + o(t^2)$$

   (c) En déduire finalement que :

$$X_n^* = \frac{S_n - E(S_n)}{\sqrt{V(S_n)}} \text{ converge en loi vers la noi normale } N(0,1)$$

Remarque Cette convergence en loi est aussi appelé **théorème central limite**. Avec la **loi faible des grands nombres**, ils constituent deux résultats puissants en probablités, notamment en ce qui concerne leur interprétation :

- la **loi faible des grands nombres** nous permet d'affirmer que la fréquence d'apparition d'un évènement tend vers vers une limite qui n'est rien d'autre que la probabilité que celui-ci se réalise... et ainsi, la répétition d'une expérience nous donnera "une bonne idée" de la probabilité qu'un évènement se réalise.

- le **théorème central limite** nous donne :

$$P(a \le X_n^* \le b) = P(a\sigma(S_n) \le S_n - E(S_n) \le b\sigma(S_n)) \underset{n\to+\infty}{\longrightarrow} \frac{1}{\sqrt{2\pi}} \int_a^b e^{-x^2/2}\, dx$$

  et ainsi, quand la taille de l'échantillon augmente, on peut en déduire "la façon" dont le nombre d'apparitions d'un évènement se disperse autour de sa moyenne. C'est ce qui justifie qu'en statistiques, nos histogrammes tendent vers une distribution gaussienne.
  Concrètement, si on observe les notes obtenues sur les 5 derniers DS, alors on obtient la distribution suivante :