Théorèmes limites des probabilités

Soit X une variable aléatoire réelle (v.a.r.) admettant une moyenne (ou espérance mathématique) μ et un écart-type σ. Soit $\left(X_i\right)_{i\in\mathbb{N}_*}$ une suite de v.a.r., indépendantes et identiquement distribuées (i.i.d.) comme X ("identiquement distribuée comme X" signifie "de même distribution (ou loi) de probabilité que X") ; $\left(X_i\right)_{i=1,\dots,n}$ est appelé échantillon statistique de taille n.

On pose $\bar{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i$, cette v.a.r. est appelée moyenne d'échantillonnage.

Les propriétés de l'espérance mathématique et de la variance permettent d'établir les égalités suivantes : $E\left(\bar{X}_n\right) = \mu$ et $\mathrm{var}\left(\bar{X}_n\right) = \frac{\sigma^2}{n}$; l'espérance de la suite $\left(\bar{X}_n\right)$ est une suite constante égale à μ et la variance de la suite $\left(\bar{X}_n\right)$ converge vers 0 lorsque n tend vers l'infini.

Le premier théorème limite (*loi faible des grands nombres*) consiste à dire que la suite de v.a.r. $\left(\bar{X}_n\right)$ converge en probabilité vers une v.a.r. d'espérance μ et de variance nulle, c'est-à-dire, vers la v.a.r. constante (dégénérée) μ.

(*Définition* : une suite de v.a.r. $\left(Y_n\right)$ *converge en probabilité* vers une v.a.r. Y (noté $\left(Y_n\right)\xrightarrow{\mathrm{Pr}} Y$) si $\forall\varepsilon > 0,\ \lim_{n\to\infty} P\left(\left|Y_n - Y\right| > \varepsilon\right) = 0.$)

On souhaite préciser la vitesse de convergence de la suite $\left(\bar{X}_n\right)$ vers μ.

On pose : $\bar{X}_n^* = \left(\frac{\bar{X}_n - \mu}{\sigma/\sqrt{n}}\right)$; la v.a.r. $\bar{X}_n^*$ est obtenue par centrage et réduction de la v.a.r. $\bar{X}_n$ ($\bar{X}_n^*$ est appelée parfois v.a.r. $\bar{X}_n$ standardisée).

Il s'agit donc d'une v.a.r. centrée réduite : $E\left(\bar{X}_n^*\right) = 0$ et $\mathrm{var}\left(\bar{X}_n^*\right) = 1$

Le deuxième théorème limite (*théorème central limite*) consiste à dire que la suite $\left(\bar{X}_n^*\right)$ converge en loi vers la loi normale centrée réduite, notée $\mathcal{N}(0,1)$.

(*Définition* : une suite de v.a.r. $\left(Y_n\right)$ *converge en loi* vers une v.a.r. Y (noté $\left(Y_n\right)\xrightarrow{Loi} Y$) si on a, en tout point de continuité y de la fonction de répartition F_Y de Y, $\lim_{n\to\infty} F_{Y_n}(y) = F_Y(y).$)

En résumé, les deux théorèmes limites des probabilités sont :

Loi faible des grands nombres :	$\left(\bar{X}_n\right)\xrightarrow{\mathrm{Pr}}\mu$
Théorème central limite :	$\left(\frac{\bar{X}_n - \mu}{\sigma/\sqrt{n}}\right)\xrightarrow{Loi}\mathcal{N}(0,1)$

Cas particulier

Si X est une v.a.r. de loi de Bernoulli de probabilité p (probabilité de succès lors d'une épreuve aléatoire). On a alors $\mu = p$ et $\sigma = \sqrt{p(1-p)}$, $\sum_{i=1}^{n} X_i$ (nombre de succès lors de n épreuves aléatoires indépendantes) suit une loi binomiale de paramètres n et p et $\bar{X}_n$ (notée F_n) est la fréquence de succès lors des n épreuves.

Les deux théorèmes limites deviennent dans ce cas particulier :

Loi faible des grands nombres : $\left(F_n\right) \xrightarrow{\text{Pr}} p$

Théorème central limite : $\left(\dfrac{F_n - p}{\sqrt{p(1-p)/n}}\right) \xrightarrow{Loi} \mathcal{N}(0,1)$

Extension du théorème central limite

On pose : $V_n = \dfrac{1}{n}\sum_{i=1}^{n}\left(X_i - \bar{X}_n\right)^2$ (variance d'échantillonnage).

Les propriétés sur les espérances et variances permettent d'établir l'égalité : $E\left(V_n\right) = \dfrac{n-1}{n}\sigma^2$.

Pour $n > 1$, on pose : $S_n^2 = \dfrac{n}{n-1}V_n = \dfrac{1}{n-1}\sum_{i=1}^{n}\left(X_i - \bar{X}_n\right)^2$, on a alors $E\left(S_n^2\right) = \sigma^2$.

La formule définissant V_n est la "formule de variance pour une population finie", alors que celle définissant S_n^2 est la "formule de variance pour un échantillon" (appelée parfois variance corrigée) car, lorsque l'on travaille sur un échantillon, c'est pour estimer au mieux les paramètres inconnus de la loi de probabilité sous-jacente ; l'espérance de la v.a.r. S_n^2 est égale à σ^2 ; on dit que la v.a.r. S_n^2 est un estimateur sans biais de σ^2 ce qui n'est pas le cas de V_n.

Le théorème central limite peut être généralisé au cas où l'on remplace le paramètre inconnu σ par la v.a.r. S_n (racine carrée de S_n^2) ou par la v.a.r. $S_n^{'}$ (racine carrée de V_n) :

$$\left(\frac{\bar{X}_n - \mu}{S_n / \sqrt{n}}\right) \xrightarrow{Loi} \mathcal{N}(0,1) \quad \text{et} \quad \left(\frac{F_n - \mu}{\sqrt{F_n(1-F_n)/n}}\right) \xrightarrow{Loi} \mathcal{N}(0,1)$$