PROPORTION EMPIRIQUE (F3)
(21 / 11 / 2020, © Monfort, Dicostat2005, 2005-2020)

(i) Soit $(\Omega, \mathcal{T}, P)$ un **espace probabilisé**, $(\mathcal{X}, \mathcal{B})$ un **espace d'observation**, $X = (X_1, \ldots, X_N)$ un **échantillon iid** comme la **variable parente** $\xi : \Omega \mapsto \mathcal{X}$, dont la **loi** est notée P^ξ, et $B \in \mathcal{B}$.

On note $p = P^\xi(B)$ la **proportion** **(théorique)** des éléments ω de Ω tq $\xi(\omega) \in B$, ie la **probabilité** que ξ « envoie » dans B :

(1) $\quad p = P([\xi \in B]), \qquad \text{avec } [\xi \in B] = \{\omega \in \Omega : \xi(\omega) \in B\}.$

La **proportion empirique** est alors définie selon :

(2) $\quad F_N = N^{-1} . \Sigma_{n=1}^{N} \mathbf{1}([X_n \in B]),$

où $\mathbf{1}(C)$ désigne la **fonction indicatrice** d'une **partie** C de $\mathcal{X}$.

(ii) Comme $E\, \mathbf{1}([X_n \in B]) = p$, on montre que :

(3) $$\begin{cases} E\, F_N = p & \text{(estimateur sans biais)}, \\ V\, F_N = N^{-1} . p . (1 - p). \end{cases}$$

Par suite (cf **théorème de la limite centrale**), on établit la **normalité asymptotique** de F_N suivante (**convergence en loi**) :

(4) $\quad \mathcal{L}\{(F_N - E\, F_N) / (V\, F_N)^{1/2}\} \to_{N \to +\infty} \mathcal{N}_1(0, 1)$ (**loi normale réduite**).

Dans ce qui précède, les moments sont calculés avec P, eg $E\, F_N = \int_\Omega F_N(\omega)\, dP(\omega)$.