ÉCHANTILLON SANS REMISE (M3)
(18 / 03 / 2020, © Monfort, Dicostat2005, 2005-2020)

Un **échantillon sans remise** est un échantillon tiré selon un **schéma probabiliste** (**schéma d'urne**) dans lequel toute **unité** tirée n'est pas remise dans l'urne avant le tirage de l'ensemble des unités suivantes : l'**échantillon** est constitué de l'ensemble des unités ainsi tirées. Chaque unité de l'urne peut ainsi figurer 0 ou 1 fois (au plus) dans l'échantillon (cf **sondage exhaustif**, **tirage exhaustif**).

(i) Soit $\Omega = \{\omega_1, ..., \omega_M\}$ un **ensemble** (urne, **population**, etc) fini (Card Ω = M). On note $\mathcal{S}_N$ l'ensemble des **arrangements** sans répétition de N unités parmi les M. On appelle alors **échantillon sans remise**, ou **échantillon sans répétition**, ou **échantillon sans re(m)placement**, de taille N extrait de Ω tout ensemble $A = \{a_1, ..., a_N\}$ qui est élément de $\mathcal{S}_N$. On note alors :

$$(1) \qquad \mathcal{S} = \cup_{N=1}^{\infty} \mathcal{S}_N$$

l'ensemble des échantillons sans remise possibles extraits de Ω.

On a donc :

$$(2) \qquad \text{Card } \mathcal{S}_N = \begin{cases} A_M^N = M (M - 1) \ldots (M - N + 1), & \text{si } N \leq M, \\ 0 \text{ sinon.} & \end{cases}$$

Par suite, $\mathcal{S} \subset \mathcal{A}$ (ensemble des échantillons avec remise).

En particulier, si N = M, le tirage du M-échantillon Ω est appelé **recensement**. On note souvent S (ou s, ou encore I) un élément de $\mathcal{S}$.

(ii) Soit $(\mathcal{Y}, \mathcal{C})$ un **espace d'observation** et $\eta : \Omega \mapsto \mathcal{Y}$ une **variable d'intérêt** étudiée, qui fait l'objet d'un sondage dans Ω. Avec les mêmes notations que pour un **échantillon avec remise**, on peut définir des estimateurs classiques pour des **caractéristiques** théoriques données (ie celles de la population).

Lorsque l'échantillon est sans répétition (ie lorsqu'on note $s \in \mathcal{S}$), on note aussi :

$$\bar{y}_N = N^{-1} \Sigma_{m \in s} Y_m \qquad \text{(moyenne empirique dans s),}$$

$$(3) \qquad T_N = M \, \bar{y}_N \qquad \text{(total empirique),}$$

$$s_N^2 = (N - 1)^{-1} \Sigma_{m \in s} (Y_m - \bar{y}_N)^2 \qquad \text{(variance empirique corrigée).}$$

(iii) L'opération de sommation sur l'échantillon sans remise A est aussi parfois notée selon eg :

$$T_N = \mathbf{S}_{n=1}^{N} y_n \qquad \text{(total empirique)},$$

(4)

$$s_N^2 = (N - 1)^{-1} \mathbf{S}_{m \in s} (Y_m - \bar{y}_N)^2, \quad \text{(variance corrigée)},$$

etc.

L'étude des échantillons sans répétition utilise souvent des **variables de CORNFIELD**.