**VARIABLE NOMINALE (C1, C4, C6, G, J, K, L)**
(27 / 04 / 2020, © Monfort, Dicostat2005, 2005-2020)

(i) Soit $(\Omega, \mathcal{T}, P)$ un **espace probabilisé** fondamental et $(\mathcal{K}, \mathcal{D})$ un **espace d'observation** mesurable (parfois appelé **espace nominal**). Soit :

(a) $\mathcal{K} = \{k_1, ..., k_M\}$ un **ensemble** fini, dont les éléments $k_m$ sont appelés **modalités** ou **états** ;

(b) $\kappa : \Omega \mapsto \mathcal{K}$ une **variable qualitative** « simple » (ie non représentable par un couple ni par un n-uple de variables).

On dit que $\kappa$ est une **variable (qualitative) nominale**, ou un **attribut nominal**, ou encore un **descripteur nominal**, ssi il n'existe pas de **structure** de **préordre** ou d'**ordre** $\geq$ définie sur $\mathcal{K}$. Autrement dit, les modalités, ou états, $k_m \in \mathcal{K}$ (m = 1 ,..., M) de $\kappa$ ne sont pas « spontanément » comparables à l'aide d'un préordre ou d'un ordre.

(ii) La **loi de probabilité**, notée $P^\kappa$ ou $\mathcal{L}(\kappa)$, de $\kappa$ peut alors être définie sur $\mathcal{D}$ à partir de P et de $\kappa$ (cf **loi qualitative**). On peut comparer ce type de lois avec la loi d'une variable numérique. Ainsi :

(a) sous des conditions larges, on peut associer à $P^\kappa$ un **mode**, ie une (ou plusieurs) modalités $k_m$ de k correspondant à la plus grande valeur des composantes $p_m$ de $P^\kappa$ ;

(b) dans certaines **situations**, le **statisticien** peut exprimer des préférences entre modalités de $\kappa$ à travers une **fonction d'utilité** $u : \mathcal{K} \mapsto \mathbf{R}$ associant à chaque modalité $k_m \in \mathcal{K}$ un nombre réel $u_m = u(k_m)$, avec $u_{m-1} \leq_u u_m$, $\forall m \in N_M^* \setminus \{1\}$. Par suite, $\kappa$ est transformée en une **variable réelle ordonnée**, ie dont les valeurs sont ordonnées par $\leq_u$ (cf **variable ordinale**), et la suite numérique $u = (u_m)_{m=1,...,M}$ associe alors indirectement (ou implicitement) à $P^\kappa$ d'autres **caractéristiques légales** : eg **fonction de répartition**, **quantile** (dont la **médiane**).

(iii) Soit $N_H^* = \{1, ..., H\}$ un ensemble (fini) d'indices donné, $(\mathcal{K}_h, \mathcal{D}_h)_{h=1,...,H}$ une **suite** d'espaces analogues à $(\mathcal{K}, \mathcal{D})$ et $(\kappa_h)_{h=1,...,H}$ une suite de variables nominales simples du type précédent, avec $\kappa_h : \Omega \mapsto \mathcal{K}_h$ et $\mathcal{K}_h = \{k_{h,1}, ..., k_{h,M(h)}\}$ ($M_h$ modalités), $\forall h \in N_H^*$.

La suite $(\mathscr{K}_h , \mathscr{D}_h)_{h=1,\ldots,H}$ précédente est appelée **suite d'espaces nominaux**. De même, la suite $(\kappa_h)_{h=1,\ldots,H}$ précédente est appelée **suite de variables nominales**, ou simplement « **suite nominale** » : on parle aussi de **variable qualitative nominale multiple** ou de **variable nominale multiple**.

On note $(\mathscr{K} , \mathscr{D})$ la **structure produit** (mesurable), avec $\mathscr{K} = \prod_{h=1,\ldots,H} \mathscr{K}_h$ et $\mathscr{D} = \bigotimes_{h=1,\ldots,H} \mathscr{D}_h$ , ainsi que $\kappa = (\kappa_h)_{h=1,\ldots,H}$ la suite des variables nominales considérées.