LOI MULTINÔMIALE (B3, C7)
(08 / 05 / 2020, © Monfort, Dicostat2005, 2005-2020))

Généralisation directe de la **loi binômiale**, la **loi multinômiale** intervient dans de nombreuses branches de la **Statistique** : **schéma probabiliste** complexe et **théorie des sondages**, **ajustement** d'une **loi de probabilité** et **statistique du chi-deux**, **modèle qualitatif** et **tableau de contingence**.

(i) Soit $n \in \mathbf{N}^*$ un entier naturel, $\Pi_n = \{x \in \mathbf{N}^K : e_K' x = n\}$ un « hyper-plan » dans $\mathbf{N}^K$ et $p = (p_1, ..., p_K) \in S_K$ (**simplexe** de $\mathbf{R}^K$).

Etant donné un **espace probabilisable** $(\Omega, \mathscr{T}, P)$ et un **vecteur aléatoire** $\xi : \Omega \mapsto \mathbf{R}^K$, on dit que la loi P^ξ de ξ est une **loi multinômiale**, ou parfois une **loi polynômiale**, ssi sa **densité** pr à la **mesure de comptage** de Π_n est de la forme :

$$(1) \qquad f(x) = \mathbf{1}(\Pi_n)(x) \,.\, (\Pi_{k=1}^K x_k\,!)^{-1} \,.\, \{(\Sigma_{k=1}^K x_k)\,!\} \,.\, \Pi_{k=1}^K p_k^{\,x(k)}, \qquad \forall\, x \in \mathbf{R}^K.$$

où $\mathbf{1}(A)$ désigne la **fonction indicatrice** d'une partie A et où, par commodité, $x(k)$ désigne x_k .

La loi multinômiale P^ξ est généralement notée $\mathscr{M}_K(n, p)$.

(ii) Cette loi possède les propriétés suivantes :

(a) ses **lois conditionnelles** et ses **lois marginales** sont multinômiales (et, en particulier, binômiales) ;

(b) deux premiers **moments algébriques** :

$$(2)_a \qquad E\,\xi = n\,p \qquad (\text{ie } E\,\xi_k = n\,p_k\,,\ \forall\, k \in N_K^*),$$

$$(2)_b \qquad C(\xi_k, \xi_l) = \begin{cases} -\,n\,p_k\,p_l & \text{si } l \neq k, \\ n\,p_k\,(1 - p_k) & \text{sinon (cf covariance).} \end{cases}$$

Autrement dit, $C(\xi_k, \xi_l) = n\,p_k\,(\delta_{kl} - p_l)$, où δ_{kl} désigne le **symbole de KRONECKER**. La **matrice de covariance** de ξ s'écrit donc :

$$(3) \qquad V\,\xi = n\,.\,(p^\wedge - p\,p'),$$

où $p^\wedge$ est le matricialisé du vecteur p (cf **matricialisation**). De plus, $rg(V\,\xi) = K - 1$;

(c) **fonction génératrice** :

$$(4) \qquad g(u) = (u'\,p)^n\,;$$

(d) **fonction caractéristique** :

(5) $\varphi(t) = \{\Sigma_{k=1}^{K} p_k \exp(i\, t_k)\}^{n}$;

(e) si les **va** $\xi \sim \mathcal{M}_K(n', p)$ et $\eta \sim \mathcal{M}_K(n'', p)$ sont indépendantes (cf **indépendance stochastique**), alors (par **convolution**) :

(6) $\xi + \eta \sim \mathcal{M}_K(n' + n'', p)$;

(f) asymptotiquement, on a la **convergence en loi** :

(7) $\mathcal{L}\{n^{-1/2}(\xi - n\, p)\} \to_{n \to +\infty} \mathcal{N}_K(0, n^{-1} V \xi)$;

(g) si $n \to +\infty$ et si $p_k \to 0+$ ($\forall\, k \in N_{K-1}{}^*$) avec n en sorte que $n\, p_k \to \lambda_k$ ($\forall\, k \in N_{K-1}{}^*$) (cf loi binômiale), alors le vecteur aléatoire $\xi^{\sim} = (\xi_1, \ldots, \xi_{K-1})'$ tend en loi vers une **loi de probabiltié** dont la **densité** pr à la **mesure de comptage** sur $\mathbf{N}^{K-1}$ s'écrit (cf **loi de POISSON multidimensionnelle**) :

(8) $f(x^{\sim}) = \Pi_{k=1}^{K-1} \exp(-\lambda_k)\, \lambda_k^{x(k)} / x_k\,!$,

où $x^{\sim} = (x_1, \ldots, x_{K-1})$ et où $x(k)$ désigne x_k .

(iii) La loi multinômiale trouve son origine dans le **schéma probabiliste** suivant. Soit $(\Omega, \mathcal{T}, P)$ un **espace probabilisé** et $A_1, \ldots, A_K$ des **événements** dont les **probabilités** d'occurrence resp sont $p_k = P(A_k)$, $\forall\, k \in N_K{}^*$, avec $\Sigma_{k=1}^{K} p_k = 1$ (évènements disjoints dont la réunion est Ω). Alors, en répétant n fois, de façon indépendante, l'**épreuve aléatoire** $(\Omega, \mathcal{T}, P)$, la lp de la va ξ dont chaque coordonnée est égale au nombre d'occurrences de A_k ($k \in N_K{}^*$) est la loi multinômiale.