INDÉPENDANCE STOCHASTIQUE (B4, C6, D1)
(25 / 04 / 2020, © Monfort, Dicostat2005, 2005-2020)

L'**indépendance stochastique**, ou **indépendance en probabilité**, est une notion clef du **calcul des probabilités** et de la **Statistique** :

(a) d'une part, elle simplifie de nombreuses **procédures statistiques** : ainsi, un **échantillon indépendant** (eg un **échantillon iid**), une **suite indépendante** (eg une **suite indépendante équidistribuée**) ou encore un **processus purement aléatoire**, sont des concepts relativement simples à étudier. C'est notamment le cas d'un certain nombre de résultats tq la **loi des grands nombres** ou le **théorème de la limite centrale** ;

(b) d'autre part, il est souvent possible d'adopter un cadre statistique, ou une **représentation statistique**, dans lesquels l'**hypothèse** d'indépendance est (quasiment) satisfaite : eg par **transformation des données**, ou lorsque le **phénomène** aléatoire est répétitif et que les conditions qui le génèrent (eg par **expérimentation**, sont invariantes ou stables.

Alors que la notion de **dépendance** comporte des définitions multiples car il en existe diverses « formes », la définition de l'indépendance est, au contraire, unique.

Ces deux notions se définissent (a) pour des **événements aléatoires**, (b) pour des **familles** d'évènements (eg **tribu de parties**), (c) pour des **variables aléatoires** et (d) pour des **statistiques**.

(i) Soit $(\Omega, \mathcal{T}, P)$ un **espace probabilisé**. dans lequel $\mathcal{T}$ est une tribu de parties de Ω.

On dit que :

(a) $A \in \mathcal{T}$ et $B \in \mathcal{T}$ sont deux **événements stochastiquement indépendants**, ou deux **événements indépendants en probabilité**, ou simplement des **événements indépendants**, ssi :

(1) $\quad P(A \cap B) = P(A) \,.\, P(B).$

En termes de **probabilité conditionnelle**, ceci équivaut à :

(2) $\quad P(B / A) = P(B)$ ou encore $P(A / B) = P(A)$,

la réalisation d'un événement n'influant en rien sur celle de l'autre ;

(b) $(A_i)_{i=1,\ldots,n}$ est une **suite** d'événements indépendants (dans leur ensemble) ssi :

(3) $\quad P(A_{i(1)} \cap \ldots \cap A_{i(k)}) = P(A_{i(1)}) \ldots P(A_{i(k)})$, ie $P(\bigcap_{j=1}^{k} A_{i(j)}) = \prod_{j=1}^{k} P(A_{i(j)})$,

pour toute suite (i_1 ,..., i_k), aussi notée (i(1) ,..., i(k)), d'entiers distincts de N_n* ;

(c) plus généralement, les **tribus** $\mathcal{T}_1$,..., $\mathcal{T}_n$ (de parties de Ω) sont des **tribus (stochastiquement) indépendantes** ssi :

(4) $\quad P(\bigcap_{i=1}^n A_i) = \prod_{i=1}^n P(A_i)$,

pour tout **système** (A_1 ,..., A_n) d'**événements** de $\mathcal{T}_1 \times \ldots \times \mathcal{T}_n$, ie tout système tq $A_i \in \mathcal{T}_i$, $\forall i \in N_n^*$.

(d) ainsi, n ≥ 2 événements A_1 ,..., A_n sont définis comme **événements (stochastiquement) indépendants** ssi les k tribus $\mathcal{T}_i$ définies par :

(5) $\quad \mathcal{T}_i = \{A_i , A_i^c , \Omega , \varnothing\}$

sont indépendantes.

Par suite, n ≥ 2 événements $(A_i)_{i=1,\ldots,n}$ sont indépendants ssi, à la fois :

$$P(A_i \cap A_j) = P(A_i) . P(A_j), \qquad \forall (i, j) : i < j,$$

$$(6) \quad P(A_i \cap A_j \cap A_k) = P(A_i) . P(A_j) . P(A_k), \qquad \forall (i, j, k) : i < j < k,$$

$$\ldots$$

$$P(A_1 \cap \ldots \cap A_n) = P(A_1) \ldots P(A_n),$$

ou, de façon équivalente, ssi (en termes absolus) :

(7) $\quad P(\bigcap_{j \in J} A_j) = \prod_{j \in J} P(A_j), \qquad \forall J \subset N_n^*$,

ou encore ssi (en termes conditionnels) :

(8) $\quad P(A_i / A_{j(1)} \cap \ldots \cap A_{j(k)}) = P(A_i), \qquad \forall k \in N_n^*$,

pour tout $i \in \{j_1 ,\ldots, j_k\}^c$ et tout (j_1 ,..., j_k) tq $P(A_{j(1)} \cap \ldots \cap A_{j(k)}) \neq 0$ (en notant aussi (i_1 ,..., i_k) pour (i(1) ,..., i(k))).

(ii) Soit X = (X_1 ,..., X_N) une suite (finie) de **vars** $X_n : \Omega \mapsto$ **R**. On dit que X est une **suite indépendante** ssi la **loi de probabilité** du **vecteur aléatoire** $X : \Omega \mapsto$ **R**^{N} est le produit des lois propres (cf **loi marginale**) des coordonnées X_n, ie ssi (cf **loi conjointe**) :

(9) $P^X = P^{X_1} \otimes ... \otimes P^{X_N}$, ou encore $\mathscr{L}(X) = \mathscr{L}(X_1) \otimes ... \otimes \mathscr{L}(X_N)$.

On montre que :

(a) pour toute suite d'évènements (A_1 ,..., A_N), l'indépendance de la suite (A_1 ,..., A_N) équivaut à celle de la suite de leurs **variables indicatrices** (**1**(A_1) ,..., **1**(A_N)) ;

(b) si (f_1 ,..., f_N) est une suite de fonctions boréliennes, définies sur **R** et à valeurs dans **R**₊ (resp tq les vars f_n (X_n) soient intégrables), on a :

(10) $E \{\Pi_n f_n (X_n)\} = \Pi_n E \{f_n (X_n)\}$.

(iii) Les résultats précédents ont été généralisés.

Ainsi, si $(\mathscr{X}_n , \mathscr{B}_n)_{n=1,...,N}$ est une suite (finie) d'**espaces probabilisables**, on dit que la suite (finie) $X = (X_n)_{n=1,...,N}$ de va $X_n : \Omega \mapsto \mathscr{X}_n$ est une **suite de variables aléatoires (stochastiquement) indépendantes**, ou simplement une **suite indépendante**, ssi les sous-tribus $\mathscr{T}_n$ (où $n \in N_N^*$) de $\mathscr{T}$ qu'elles engendrent resp sont des tribus indépendantes (cf **tribu engendrée**). Pour celà, il faut et il suffit que la **loi conjointe** des va X_n , définie sur la tribu $\otimes_n \mathscr{B}_n$, soit égale au produit des lois $P^{X(n)}$ (**lp propres** aux va X_n), ie :

(11) $P^X ([X \in B]) = \otimes_n P^{X(n)} ([X_n \in B_n])$, $\forall B \in \otimes_n \mathscr{B}_n$,

où $A = (A_1 ,..., A_N)$ et $X = (X_1 ,..., X_N)$, ou encore :

(12) $P^X = \otimes_n P^{X(n)}$.

(en notant X(n) pour désigner X_n).

Si $A = (A_n)_{n \in \mathbf{N}}$ est une suite d'évènements $A_n \in \mathscr{T}$, on dit que A est une **suite (d'événements) stochastiquement indépendante**, ou une **suite indépendante en probabilité**) ssi chacune des suites finies $(A_n)_{n = 0,1,...,N}$ est une suite indépendante. Les notions s'étendent au cas d'une **famille d'événements** (resp d'une famille de va) quelconque en supposant l'indépendance de toute suite finie d'événements (resp de va) (cf aussi **processus stochastique**, **système projectif de probabilités**).

On dit parfois que les variables définissant une suite indépendante $X = (X_n)_{n \in \mathbf{N}}$ sont des **variables orthogonales**, et l'on note alors $X_\alpha \perp X_\beta$, $\forall (\alpha, \beta) \in \mathbf{N}^2_{\neq}$.

(iv) En **Statistique**, l'indépendance intervient souvent en relation avec des questions d'**exhaustivité** et de **liberté**. Ainsi, si $(\Omega, \mathscr{T}, P)$ est un **modèle statistique**, $\mathscr{S}$ une sous-**tribu exhaustive** et quasi-complète de $\mathscr{T}$ (cf **probabilité complète**), et $\mathscr{R}$ une sous-tribu de $\mathscr{T}$ supposée être une tribu libre (cf **liberté**), alors les sous-tribus $\mathscr{R}$ et $\mathscr{S}$ sont indépendantes, $\forall$ P $\in \mathscr{P}$.