TEST D'INDÉPENDANCE (D1, I2, N)
(13 / 11 / 2019, © Monfort, Dicostat2005, 2005-2019)

L'hypothèse, commode mais souvent « cruciale », d'**indépendance** des **variables** considérées ou des **observations** disponibles est souvent faite en **Statistique**. Il est donc important de s'assurer de son bien-fondé, ce qui est l'objet d'un **test d'indépendance**.

Plusieurs **situations statistiques** courantes font intervenir l'**hypothèse d'indépendance**.

(i) Le **problème à un échantillon** concerne l'indépendance à l'intérieur d'un **échantillon aléatoire**. Disposant d'un **modèle d'échantillonnage** dans lequel $(\mathscr{X}, \mathscr{B}) = (\mathscr{X}_0^N, \mathscr{B}_0^{\otimes N})$ et $X = (X_1, ..., X_N) : \Omega \mapsto \mathscr{X}$ représente l'échantillon, le problème de l'**indépendance « interne »**, ou indépendance entre les **variable aléatoire** X_n (n = 1,..., N) revient généralement à vérifier que le **processus stochastique** défini par X est un **processus purement aléatoire**, ie à tester l'**hypothèse de base** :

(1) $H_0 : P^X = P^{X(1)} \otimes ... \otimes P^{X(N)} = \otimes_{n=1}^N P^{X(n)}$,

ou encore, avec une notation alternative :

(1) $H_0 : \mathscr{L}(X) = \mathscr{L}(X_1) \otimes ... \otimes \mathscr{L}(X_N) = \otimes_{n=1}^N \mathscr{L}(X_n)$,

(produit tensoriel des **lois** individuelles $P^{Xn} = \mathscr{L}(X_n)$), où les X(n) désignent, par commodité, les **va** X_n .

Dans le cas où les variables X sont identiquement distribuées (**échantillon iid**), (1) s'écrit :

(2) $H_0 : P^X = (P^\xi)^{\otimes N}$, ou encore $\mathscr{L}(X) = \{\mathscr{L}(\xi)\}^{\otimes N}$

(puissance tensorielle de la loi commune $P^\xi = \mathscr{L}(\xi)$.

Il existe de nombreux tests d'indépendance de ce type. Ces tests correspondent au problème de l'échantillon (purement) aléatoire (cf aussi **processus purement aléatoire**).

(ii) Le **problème de l'indépendance** d'un **couple aléatoire** $\zeta = (\xi, \eta)$, dont la **loi** est $P^{(\xi, \eta)}$, consiste à tester que la **va** η est indépendante de la va ξ au vu d'un **échantillon** $Z = ((X_1 , Y_1) ,..., (X_N , Y_N))$, dans lequel $Z_n = (X_n , Y_n)$ est une **copie** de ζ.

Il s'agit donc de tester l'**hypothèse de base** d'**indépendance entre échantillons**, ou **indépendance « externe »** :

(3) $H_0 : P^\zeta = P^\xi \otimes P^\eta$.

Des tests de ce type sont le **test de FISHER-YATES**, le **test de HOEFFDING**, le **test de KENDALL**, le **test de SPEARMAN** ou encore le **test des quadrants**.

(iii) Dans l'analyse d'un **tableau statistique multidimensionnel** (eg **tableau de contingence**) on peut définir des tests d'indépendance entre les **variables** (ou « **critères** ») définissant le tableau, dont les tests suivants : **test du chi-deux**, test du chi-deux modifié (cf **statistique du chi-deux modifiée**), test de KULLBACK - LEIBLER (fondé sur la **statistique de KULLBACK-LEIBLER**), test de HELLINGER (fondé sur la **statistique de HELLINGER**).

(iv) On assimile souvent à la catégorie des tests d'indépendance, la catégorie des **tests de non corrélation** : **test du coefficient de corrélation**, **test du coefficient de détermination**, etc. Les tests de KENDALL ou de SPEARMAN indiqués ci-dessus sont parfois assimilés à cette catégorie de tests.