ROCZNIK NAUKOWO–DYDAKTYCZNY WSP W KRAKOWIE

Zeszyt 51 Prace matematyczne VII Rok 1974

Stanisław Wołodźko

SUR L'ASSOCIATIVITÉ DE LA DÉCOMPOSITION BIDIMENTIELLE DE LA VARIABLE ALÉATOIRE

Soit (X,Y) une variable aléatoire bidimentielle de type discret possédant une associativite de décomposition $((x_i,x_j),p_{ij})(i,j = 1,\ldots)$. Nons savons, à partir de la théorie élémentaire du calcul de probabilité /ex.[1], p.48/, que la décomposition frontale de la variable aléatoire X est un ensemble de paires $(x_i,p_{i.})$ où $p_{i.}=\sum_j p_{ij}$, et la décomposition frontale de la variable Y un ensemble de paires $(y_j, p_{.j})$ où $p_{.j}=\sum_i p_{ij}$. Cette note a pour but un essai détermination de l'associativité de la décomposition de la variable aléatoire (X,Y), quand les décompositions frontales des variables correspondantes sont données. D'abord, nous étudierons le problème dans le cas où les variables aléatoires X et Y prennent avec la probabilité 1, un nombre fini de valeurs. Sans réduire les généralites des solutions, nous pouvons supposer que les ensembles des valeurs des deux variables aléatoires considérées, pris avec la probabilité 1, sont équipotents /puisque nous pouvons toujors ajounter, à l'ensemble de la valeur relative de la variable aléatoire, quelques valeurs prises avec la probabilité zéro/. Ensuite, determinons l'associativité de la décomposition de la variable aléatoire (X, Y) recevant avec la probabilite 1 une quantité dénombrable de valeurs.

1. Considérons comme données les décompositions suivantes:

/1/ (x_i,p_i) $(i=1,\ldots, n+1)$ - décompositions de la variable aléatoire X,

/2/ (y_j,q_j) $(j=1,\ldots, n+1)$ - décompositions de la variable aleatoire Y,

où n est un nombre naturel donné. Cherchons l'associativité des décompositions de la variable aléatoire X,Y telles que de décompositions données des variables aléatoires X et Y sont leurs décompositions frontales.

Soit la suite

$$p_{ij} = P(X = x_i,\ Y = y_j) \quad \text{pour} \quad i,j = 1,\ldots, n+1$$

définissant la décomposition recherchée. Si nous indiquons

$$\alpha_{ij} = p_{ij} - p_i q_j \qquad (i,j = 1,\dots, n+1),$$

nous obtenons

/3/ $$\sum_{i=1}^{n+1} \alpha_{ij} = \sum_{i=1}^{n+1} p_{ij} - \sum_{i=1}^{n+1} p_i q_j = \sum_{i=1}^{n+1} p_{ij} - q_j \sum_{i=1}^{n+1} p_i \quad (j=1,\dots, n+1).$$

Du fait que /1/est une décomposition de la variable aléatoire il resulte que $\sum_{i=1}^{n+1} p_i = 1$, de là,alors que la décomposition /1/ est une décomposition frontale de la variable aléatoire (X, Y) dans la décomposition associative de la variable aléatoire (X,Y) nous obtenons $\sum_{i=1}^{n+1} p_{ij} = q_j$. Par conséquent de /3/ nous concluons que

/4/ $$\sum_{i=1}^{n+1} \alpha_{ij} = 0 \quad \text{pour} \quad j=1,\dots, n+1.$$

Analogiquement nous obtenons

/5/ $$\sum_{j=1}^{n+1} \alpha_{ij} = 0 \quad \text{pour} \quad i=1,\dots, n+1.$$

Nous obtenons de /5/ et /4/

/6/ $$\sum_{i,j=1}^{n+1} \alpha_{ij} = 0.$$

Il est facile de vérifier que de la formule

/7/ $$\alpha_{ij} = \varepsilon_{ij} - \varepsilon_{i-1j} - \varepsilon_{ij-1} + \varepsilon_{i-1j-1} \qquad (i,j = 1,\dots, n+1)$$

nous obtenons chaque solution du systeme /4/ et /5/, où les paramétres ε_{ks} sont des nombres naturels arbitraires, tels que

/8/ $$\varepsilon_{ko} = \varepsilon_{os} = \varepsilon_{n+1,s} = \varepsilon_{k,n+1} = 0 \quad (k,s = 1,\dots, n+1).$$

Remarquons, en plus, que pour des réels donnés α_{ij} les parametres ε_{ks} sont définis univoquement.

Prouvons ce qui suit

<u>Théorème 1</u>. Supposons que la décomposition associative de la variable aléatoire (X,Y) est sous forme $((x_i, y_j), p_{ij})$ $(i,j = 1,\dots,n+1)$, où les décompositions frontales des variables X et Y sont homologues à /1/ et /2/. Soit en plus $\alpha_{ij} = p_{ij} - p_i q_j$ $(i,j = 1,\dots, n+1)$. Il existe alors univoquement une famille definie de parametres $\{\varepsilon_{ks}\}$ $(k,s = 1,\dots,n)$ qui remplissent les conditions d'apres /7/ et /8/ et pour $i,j = 1,\dots,n$ les inegalités

/9/ $\max\{-p_i^* q_j^*, -p_i q_j + \varepsilon_{ij-1} + \varepsilon_{i-1j} - \varepsilon_{i-1j-1}\} \leq \varepsilon_{ij} \leq \min\{p_i^* q_j + \varepsilon_{ij-1}, p_i q_j^* + \varepsilon_{i-1j}\}$,
où $p_i^* = (1-p_1-\ldots-p_i)$, $q_j^* = (1-q_1-\ldots-q_j)$.

Démonstration. L'existence et l'unicité de la famille des paramètres $\{\varepsilon_{ks}\}$ remplissant des formules /7/ et /8/, résultent de là que les nombres α_{ij} remplissent le système /4/, /5/. Il reste à demontrer /9/. Remarquons, dans ce but, que les nombres p_{ij} sont des probabilités, alors $p_{ij} = p_i q_j + \alpha_{ij} \geqslant 0$. De là et de /7/ il résulte que

/10/ $$\varepsilon_{ij} - \varepsilon_{i-1j} - \varepsilon_{ij-1} + \varepsilon_{i-1j-1} \geqslant -p_i q_j \quad \text{pour} \quad i,j=1, \quad n+1.$$

Désignons ensuite par le symbole E_{ij} la phrase "le paramètre ε_{ij} remplit l'inégalité /9/". Dans les étapes suivantes nous entreprenons la démonstration de cette inégalité

1°. E_{nn},

2°. $E_{in} \Longrightarrow E_{i-1n}$ $(i=2,\ldots, n)$,

3°. $E_{nj} \Longrightarrow E_{nj-1}$ $(j=2,\ldots, n$,

4°. $E_{i+1j} \wedge E_{ij+1} \Longrightarrow E_{ij}$ $(i,j = 1,\ldots, n-1)$.

Ad 1°. De l'inégalité /10/ nous déduisons toutes les conditions que remplit le paramètre ε_{nn}. En posant successivement $i = j = n+1$; $i = n+1$, $j = n$; $i = n$, $j = n+1$ et $i = j = n$, nous obtenons

$$\varepsilon_{nn} \geqslant -p_{n+1} q_{n+1},$$

$$\varepsilon_{nn} \leqslant p_{n+1} q_n + \varepsilon_{n\,n-1},$$

$$\varepsilon_{nn} \leqslant p_{n+1} + \varepsilon_{n-1\,n},$$

$$\varepsilon_{nn} \geqslant -p_n q_n + \varepsilon_{n-1\,n} + \varepsilon_{nn-1} - \varepsilon_{n\,n-1}.$$

De là, nous obtenons

$$\max\{-p_{n+1}q_{n+1}, -p_n q_n + \varepsilon_{n-1n} + \varepsilon_{nn-1} - \varepsilon_{n-1\,n-1}\} \leqslant \varepsilon_{nn} \leqslant \min\{p_{n+1}q_n + \varepsilon_{nn-1}, p_n q_{n+1} + \varepsilon_{n-1n}\}.$$

Nous avons donc obtenu E_{nn}, parce que $p_{n+1} = p_n^*$ et $q_{n+1} = q_n^*$.

Ad 2°. Choisissons un nombre naturel arbitraire i tel que $2 \leqslant i \leqslant n$ et supposons E_{in}, c'est-à-dire:

$$\max\{-p_i^* q_n^*, -p_i q_n + \varepsilon_{in-1} + \varepsilon_{i-1n} - \varepsilon_{i-1n-1}\} \leqslant \varepsilon_{in} \leqslant \min\{p_i^* q_n + \varepsilon_{in-1}, p_i q_n^* + \varepsilon_{i-1n}\}.$$

De là, nous obtenons les conditions suivantes pour le paramètre ε_{i-1n}:

$$-p_i^* q_n^* \leqslant p_i q_n^* + \varepsilon_{i-1n}$$

et

$$-p_i q_n + \varepsilon_{i-1n} + \varepsilon_{i-1n-1} \leqslant p_i^* q_n^*.$$

Ajoutons-leurs les conditions obtenues de /10/

$$-p_i\, q_n \leqslant \varepsilon_{i-1n} - \varepsilon_{i-2n} - \varepsilon_{i-1\;n-1} + \varepsilon_{i-2\;n-1},$$

$$-p_{i-1}\, q_{n+1} \leqslant -\varepsilon_{i-1\;n} + \varepsilon_{i-2\;n}.$$

Écrivons sous une forme équivalente les quatre inégalités déjà citées:

$$\varepsilon_{i-1n} \geqslant -p_i^*\, q_n^* - p_i\, q_n^* = -p_{i-1}^*\, q_n^*,$$

$$\varepsilon_{i-1n} \leqslant p_i\, q_n^* + p_i\, q_n + \varepsilon_{i-1\;n-1} = p_i\, q_{n-1}^* + \varepsilon_{i-1\;n-1},$$

$$\varepsilon_{i-1n} \leqslant p_{i-1}\, q_n^* + \varepsilon_{i-2\;n},$$

$$\varepsilon_{i-1n} \geqslant -p_{i-1}\, q_n + \varepsilon_{i-2n} + \varepsilon_{i-1n} - \varepsilon_{i-2\;n-1}.$$

De là, E_{i-1n} en résulte directement.

Ad 3^o. Démonstration analogue à la deuxième.

Ad 4^o. Soit $i,j<n$ des nombres naturels arbitraires donnés et supposons que E_{i+1j} et E_{ij+1}. Alors de E_{i+1j} nous avons

$$-p_{i+1}^*\, q_j^* \quad p_i q_j^* + \varepsilon_{ij}$$

et

$$-p_{i+1}\, q_j + \varepsilon_{ij} + \varepsilon_{i+1\;j-1} - \varepsilon_{ij-1} \leqslant p_{i+1}^*\, q_1 + \varepsilon_{i+1\;j-1}.$$

Par contre de E_{ij+1} il résulte que

$$-p_i\, q_{j+1} + \varepsilon_{ij} + \varepsilon_{i-1\;j+1} - \varepsilon_{i-1j} \leqslant p_i\, q_{j+1}^* + \varepsilon_{i-1\;j+1}.$$

Ecrivons les trois conditions ci-dessus sous forme suivantes

$$\varepsilon_{ij} \geqslant -p_i^*\, q_j,$$

$$\varepsilon_{ij} \leqslant p_i^*\, q_j + \varepsilon_{ij-1},$$

$$\varepsilon_{ij} \leqslant p_i\, q_j^* + \varepsilon_{i-1j}.$$

D'autre part de /10/ nous obtenons

$$\varepsilon_{ij} \geqslant -p\;\; q_j + \varepsilon_{i-1j} + \varepsilon_{ij-1} - \varepsilon_{i-1\;j-1}.$$

Des quatre dernières inégalités nous obtenons E_{ij}.

2. Dans le théorème 1 nous avons donné une condition necessaire afin que l'ensemble de paramètres $\{\varepsilon_{ks}\}$ definisse la décomposition associative de la variable aléatoire (X,Y), quand les decompositions frontales des variables aléatoires X et Y sont donnés. Démontrons que cette condition est également suffisante.

<u>Théorème 2</u>. Supposons que les décompositions des variables aléatoires X et Y sont successivement /1/ et /2/ et en plus que la famille des paramètres $\{\varepsilon_{ks}\}$ $(k,s = 1,\ldots, n)$ remplisse la condition /9/. Alors la suite p_{ij} definie par la formule.

$$p_{ij} = p_i\, q_j + \alpha_{ij} \qquad (i,j = 1,\ldots, n+1),$$

où

/11/ $$\alpha_{ij} = \varepsilon_{ij} - \varepsilon_{ij-1} - \varepsilon_{i-1j} + \varepsilon_{i-1\; j-1}$$

et

/12/ $$\varepsilon_{io} = \varepsilon_{oj} = \varepsilon_{in+1} = \varepsilon_{n+1j} = 0$$

définit la décomposition de la variable aléatoire (X,Y) telle que les décompositions des variables aléatoires X et Y sont leurs décompositions frontales.

<u>Démonstration</u>. De /11/ et /12/ il résulte que les nombres α_{ij} remplissent le système /4/, /5/, alors il en résulte /6/. De là, nous obtenons tout de suite que

$$\sum_{i,j=1}^{n+1} p_{ij} = \sum_{i,j=1}^{n+1} p_i\, q_j + \sum_{i,j=1}^{n+1} \alpha_{ij} = \sum_{i=1}^{n+1} p_i \sum_{j=1}^{n+1} q_j = 1,$$

$$p_{i.} = \sum_{j=1}^{n+1} p_{ij} = \sum_{j=1}^{n+1} p_i\, q_j + \sum_{j=1}^{n+1} \alpha_{ij} = p_i \quad (i=1,\ldots, n+1)$$

et $$p_{.j} = \sum_{i=1}^{n+1} p_{ij} = \sum_{i=1}^{n+1} p_i q_j + \sum_{i=1}^{n+1} \alpha_{ij} = q_j \quad (j=1,\ldots, n+1).$$

Remarquons ensuite que de /9/ nous avons

$$- p_i\, q_j + \varepsilon_{ij-1} + \varepsilon_{i-1j} - \varepsilon_{i-1\; j-1} \leqslant \varepsilon_{ij}.$$

Ce qui veut dire que $p_{ij} = p_i\, q_j + \alpha_{ij} \geqslant 0$. Ce qui termine la démonstration.

Démonstrons encore de quelle manière en pratique nous déterminons en s'appuyant le théorème 2, la décomposition associative de la variable aléatoire (X,Y) avec des décompositions frontales données. Soient X et Y des variables aléatoires possédant les décompositions /1/ et /2/.

La décomposition associative recherchée est obtenue de la manière suivante. Prenons un nombre quelconque ε_{11} de l'intervalle

/13/ $$\left[\max\left\{-(1-p_1)(1-q_1),\ -p_1 q_1\right\},\ \min\left\{(1-p_1)q_1,(1-q_1)p_1\right\}\right].$$

Nous déterminons le parametre ε_{12} de l'intervalle

$$\left[\max\left\{-(1-p_1)(1-q_1-q_2),-p_1q_2+\varepsilon_{11}\right\},\min\left\{p_1(1-q_1-q_2),(1-p_1)q_2+\varepsilon_{11}\right\}\right].$$

Prenons ensuite ε_{21} et ε_{22} des intervalles relatives etc.Du théorème 1 il résulte que, de cette manière nous obtenons toutes les décompositions associatives de la variable aléatoire (X,Y).

3. Démontrons maintenant que les théorèmes 1 et 2 sont également valables pour les variables aléatoires prenants avec la probabilité 1 une quantité dénombrable de valeurs.

Pour la démonstration du théorème 1 dans le cas où $n = +\infty$, nous construirons la famille $\{\varepsilon_{ks}\}$ de la manière suivante. Soit $m \geqslant 2$ un nombre naturel arbitraire. Remplaçons par des décompositions relatives, les décompositions des variables aléatoires X, Y et (X,Y) données

$$(x_i,\tilde{p}_i),(y_j,\tilde{q}_j) \quad \text{et} \quad ((x_i,y_j),\ \tilde{p}_{ij}) \quad (i,j = 1,\dots,\ n+1),$$

où
$$\tilde{p}_i = p_i \ (i = 1,\dots,\ m), \quad \tilde{p}_{m+1} = \sum_{i=m+1}^{\infty} p_i,$$

$$\tilde{q}_j = q_j \ (j=1,\dots,\ m), \quad \tilde{q}_{m+1} = \sum_{j=m+1}^{\infty} q_j,$$

$$\tilde{p}_{ij} = p_{ij}\,(i,j = 1,\dots,\ m), \quad \tilde{p}_{im+1} = \sum_{j=m+1}^{\infty} p_{ij},$$

$$\tilde{p}_{m+1j} = \sum_{i=m+1}^{\infty} p_{ij}$$

et
$$p_{m+1\ m+1} = \sum_{i,j=\ m+1}^{\infty} p_{ij}.$$

Soit en plus

$$\tilde{\alpha}_{ij} = \tilde{p}_{ij} - \tilde{p}_i\ \tilde{q}_j \qquad (i,j = 1,\dots,\ n+1).$$

De là et du théorème 1 il résulte qu'il existe univoquement une famille de paramètres définie $\{\varepsilon_{ks}\}$ $(k,s = 1,\dots,\ m)$ remplissant /9/ et telle que

/14/ $$\tilde{\alpha}_{ij} = \tilde{p}_i\ \tilde{q}_j + \tilde{\alpha}_{ij},$$

où

/15/ $$\tilde{\alpha}_{ij} = \varepsilon_{ij} - \varepsilon_{ij-1} - \varepsilon_{i-1j} + \varepsilon_{i-1\,j-1}$$

et

/16/ $$\varepsilon_{io} = \varepsilon_{oj} = \varepsilon_{im+1} = \varepsilon_{m+1\,j} = 0$$

pour $ij = 1,\dots, m+1$. En concevant que $m \to +\infty$ nous obtenons alors une famille de paramètres $\{\varepsilon_{ks}\}$ $(k,s = 1,\dots)$ remplissant /9/,/14,/15/ et /16/ pour $i,j = 1,\dots$.

Pour la démonstration du théorème 2 dans le cas où $n = +\infty$ il suffit de démontrer que

/17/ $$\sum_{i=1}^{\infty} \alpha_{ij} = 0 \quad \text{pour} \quad i = 1,\dots,$$

/18/ $$\sum_{j=1}^{\infty} \alpha_{ij} = 0 \quad \text{pour} \quad j = 1,\dots,$$

et

/19/ $$\sum_{i,j=1}^{\infty} \alpha_{ij} = 0.$$

Dans ce but remarquons que de /9/ nous obtenons pour les nombres naturels arbitraires i,j

$$p_i^* q_j^* \leqslant \varepsilon_{ij} \quad \text{et} \quad \varepsilon_{ij} \leqslant p_i^* q_j + \varepsilon_{ij-1}.$$

De la deuxième inégalité, nous concluons, que $\varepsilon_{ij} \leqslant p_i^*(q_1 + \dots + q_j)$. Puisque $\lim\limits_{i\to\infty} p_i = 0$, donc, nous avons $\lim \varepsilon_{ij} = 0$ pour $j = 1,\dots$. En plus, compte tenu de $\lim\limits_{j\to\infty}(q_1 + \dots + q_j) = 1$, des inégalités ci-dessus il résulte que $\lim\limits_{ij\to\infty} \varepsilon_{ij} = 0$. Analogiquement nous démontrons que $\lim\limits_{i,j\to\infty} \varepsilon_{ij} = 0$ pour $i = 1,\dots$.

D'autre part nous avons

$$\sum_{i=1}^{n} \alpha_{ij} = \varepsilon_{in} - \varepsilon_{i-1n},$$

$$\sum_{j=1}^{n} \alpha_{ij} = \varepsilon_{nj} - \varepsilon_{nj-1}$$

et $$\sum_{i,j=1}^{n} \alpha_{ij} = \varepsilon_{nn}.$$

Des trois dernières inégalités résulte successivement /17/, /18/ et /19/.

R e m a r q u e s   f i n a l e s. $\varepsilon_{ij} = 0$ pour $i,j = 1,\dots$ est un des choix possibles des paramètres $\{\varepsilon_{ks}\}$. Si et seulement si tous les nombres se neutralisent, ce qui veut dire que $p_{ij} = p_i q_j$ pour $i,j = 1,\dots$. La neutralisation de tous les paramètres ε_{ij} est équivalents à l'indépendance des variables aléatoires X et Y. Les paramètres ε_{ij} ont la propriété suivante. Soit F une distribuante de la variable aléatoire (X,Y), et F_1, F_2 des distribuantes correspondantes des variables aléatoires X et Y. Si nous désignons

$$A_{ij} = \{(x,y):\ x_i < x \leqslant x_{i+1},\ y_j < y \leqslant y_{j+1}\} \quad \text{pour} \quad i,j = 1,\dots,$$

il est possible vérifier que

$$F(x,y) = F_1(x)\, F_2(y) + \sum_{i,j=1}^{\infty} \varepsilon_{ij}\, \chi_{A_{ij}},$$

où $\chi_{A_{ij}}$ est la fonction caractéristique du rectangle A_{ij}.

Par contre, par le nombre α_{ij} la covariance μ_{11} de la variable aléatoire (X,Y) s'exprime directement. Nous obtenons à savoir, la formule $\mu_{11} = \sum_{i,j=1}^{\infty} \alpha_{ij}\, x_i y_j$.

Pour indiquer les possibilités d'application des résultats obtenus dans les notes, considérons l'exemple suivant: Considérons l'expérience aléatoire D, dans laquelle la probabilité de succes est égale à $p_1 = \frac{7}{10}$, tandis que la probabilité d'échec est égale à $p_2 = \frac{3}{10}$. Répétons à deux réprisses l'expérience D, donc, de la formule /13/

$$p_{11} = \frac{49}{100} + \varepsilon_{11}, \quad \text{où} \quad -\frac{9}{100} \leqslant \varepsilon_{11} \leqslant \frac{21}{100},$$

nous obtiendrons la probabilité de deux succés. Par ailleurs la probabilité p_{11} est un nombre de l'intervalle $\left[\frac{4}{10}, \frac{7}{10}\right]$.

T r a v a i l   c i t é

[1] M. Fisz, Rachunek prawdopodobieństwa i statystyka matematyczna, Warszawa 1958.