第 11 卷第 2 期　　空军工程大学学报(自然科学版)　　Vol. 11　No. 2
2010 年 4 月　　JOURNAL OF AIR FORCE ENGINEERING UNIVERSITY(NATURAL SCIENCE EDITION)　　Apr. 2010

二维多态联合概率分布及信息传输理论研究*

刘进忙[1]，　张娅珣[1]，　袁修久[2]

(1. 空军工程大学　导弹学院，陕西　三原　713800；2. 空军工程大学　理学院，陕西　西安　710051)

摘　要：根据不相关性与独立的等价概念，给出了有限离散随机变量的二维联合概率分布的一种描述方法。将两点值的情况推广到多点离散值的情况，当给出两个随机变量的边缘分布及合适的有限阶幂互协方差函数(或高阶相关系数)时，可得两维联合概率分布列，使两个随机变量的不相关性与独立性有一系统地描述和一个完整的表达形式。经过仿真验证了此方法的有效性。将得到的表示形式应用在数据信道通信的分析中，用信道两端的相关系数来描述输入输出的信息耦合，对信息理论是一个很好的补充。该方法与信息论中的平均互信息、数据处理定理相对应，得到非常美观的公式及非常理想的特性。

关键词：相关系数；联合概率分布；信道矩阵

DOI：10.3969/j.issn.1009-3516.2010.02.016

中图分类号：TP274；O21　　**文献标识码**：A　　**文章编号**：1009-3516(2010)02-0073-04

两个独立随机变量的联合概率密度函数等于各自的概率密度函数的乘积[1]。当2个边缘概率密度函数的乘积不等于联合概率密度函数时，称为2个随机变量不独立。因此，由联合概率密度函数可唯一确定2个随机变量的边缘概率密度函数，反之不能成立[1]。为此，文中将进一步讨论两者等价的条件。

在随机变量的数字特征分析中，若2个随机变量的(线性)相关系数定义为 r_{XY}，或用2个随机变量的协方差 $\mathrm{cov}(X,Y)$ 来描述。当 $r_{XY}=0$ 或 $\mathrm{cov}(X,Y)=0$ 为不相关。一般情况下，(线性)不相关是独立的必要条件，并非充分条件。即独立必定是(线性)不相关的 $r_{X_y}=0$，而线性不相关不能推出独立。

由不相关性与独立性等价的方法，将协方差函数定义推广为高阶的形式：

$$\mathrm{cov}(X^m,Y^n)\triangleq E[X^m-E(X^m)][Y^n-E(Y^n)]=E[X^mY^n]-E[X^m]E[Y^n]=r_{X^mY^n}\sigma_{X^m}\sigma_{Y^n},\ m,n=1,2,\cdots \tag{1}$$

式中：r 若随机变量 X、Y 满足 $\mathrm{cov}(X^m,Y^n)=0(m,n=1,2,\cdots)$ 与随机变量 X、Y 独立等价。

显然，(线性)不相关($m=n=1$)仅仅是(高阶)不相关($m,n=1,2,\cdots$)的特例。

因此给出了高阶不相关与独立的逻辑图，见图1。

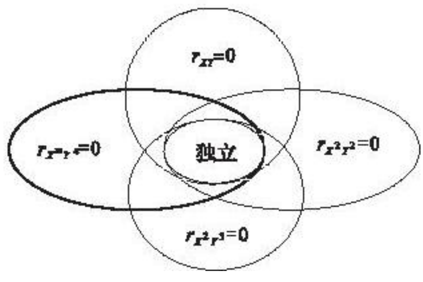

图 1　不相关与独立逻辑关系

Fig. 1　The logic relation with uncorrelation and independence

2个随机变量只取2个值的情况下，不相关与独立等价[1]。通过细致分析，得到多点离散值的情况，给出了一种二维离散随机变量的联合分布表达形式。并讨论了各参数隔离的问题，给出了一个非常清晰的描述方法。

* 收稿日期：2009-09-23

作者简介：刘进忙(1958-)，男，陕西渭南人，教授，博士生导师，主要从事信号与信息处理研究．
E-mail：cathyqueencn@yahoo.com.cn

1 2 个有限离散随机变量的联合概率分布统一描述

根据文献[4],设随机变量 $\boldsymbol{X}$ 的概率空间 $\begin{pmatrix} x_1 & x_2 & \cdots & x_m \\ p_1 & p_2 & \cdots & p_m \end{pmatrix}$,$p_i > 0$,$\sum_{i=1}^{m} p_i = 1$

随机变量 $\boldsymbol{Y}$ 的概率空间为 $\begin{pmatrix} y_1 & y_2 & \cdots & y_n \\ q_1 & q_2 & \cdots & q_n \end{pmatrix}$,$q_j > 0$,$\sum_{j=1}^{n} q_i = 1$

定理 1 所有互不相同分量 x_i 从小到大排列表示行,所有互不相同分量 y_j 从小到大排列表示列,其 $\boldsymbol{X}$、$\boldsymbol{Y}$ 的联合概率分布列矩阵为:

$$\begin{pmatrix} f_{11} & f_{11} & \cdots & f_{1n} \\ f_{21} & f_{22} & \cdots & f_{2n} \\ \cdots & \cdots & & \cdots \\ f_{m1} & f_{m2} & \cdots & f_{mn} \end{pmatrix} = \begin{pmatrix} 1 & 1 & \cdots & 1 \\ x_1 & x_2 & \cdots & x_m \\ \cdots & \cdots & & \cdots \\ x_1^{m-1} & x_2^{m-1} & \cdots & x_m^{m-1} \end{pmatrix}^{-1} E(\boldsymbol{X}_m \boldsymbol{Y}_n^{\mathrm{T}}) \begin{pmatrix} 1 & y_1 & \cdots & y_1^{n-1} \\ 1 & y_2 & \cdots & y_2^{n-1} \\ \cdots & \cdots & & \cdots \\ 1 & y_n & \cdots & y_n^{n-1} \end{pmatrix} = \begin{pmatrix} p_1 \\ p_2 \\ \cdots \\ p_m \end{pmatrix} (q_1 \quad q_2 \quad \cdots \quad q_n) +$$

$$\begin{pmatrix} 1 & 1 & \cdots & 1 \\ x_1 & x_2 & \cdots & x_m \\ \cdots & \cdots & & \cdots \\ x_1^{m-1} & x_2^{m-1} & \cdots & x_m^{m-1} \end{pmatrix}^{-1} \operatorname{cov}(\boldsymbol{X}_m, \boldsymbol{Y}_n) \begin{pmatrix} 1 & y_1 & \cdots & y_1^{n-1} \\ 1 & y_2 & \cdots & y_2^{n-1} \\ \cdots & \cdots & & \cdots \\ 1 & y_n & \cdots & y_n^{n-1} \end{pmatrix}^{-1} \tag{2}$$

2 基正交函数的幂协方差阵

根据文献[6],由于各幂协方差函数是线性无关的,为排除他们之间耦合关系,采用正交多项式矩阵代替矩阵 $\boldsymbol{X}$ 的 vandermonde 矩阵。可设 $\varphi_i(x)$ 为第 i 个加权正交多项式函数,满足如下关系:

故离散矩阵表示形式:

$$\begin{pmatrix} \varphi_0(x_1) & \varphi_0(x_2) & \cdots & \varphi_0(x_m) \\ \varphi_1(x_1) & \varphi_1(x_2) & \cdots & \varphi_1(x_m) \\ \cdots & \cdots & & \cdots \\ \varphi_{m-1}(x_1) & \varphi_{m-1}(x_2) & \cdots & \varphi_{m-1}(x_m) \end{pmatrix} = \begin{pmatrix} 1 & & & & \\ a_{11} & 1 & & & \\ a_{21} & a_{22} & 1 & & \\ \cdots & \cdots & & \ddots & \\ a_{(m-1)1} & a_{(m-1)2} & \cdots & \cdots & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 & \cdots & 1 \\ x_1 & x_2 & \cdots & x_m \\ x_1^2 & x_2^2 & \cdots & x_m^2 \\ \cdots & \cdots & & \cdots \\ x_1^{m-1} & x_2^{m-1} & \cdots & x_m^{m-1} \end{pmatrix} \tag{3}$$

式中:若采用幂级数函数,有 $\varphi_0(\boldsymbol{X}) = 1$; $\varphi_1(\boldsymbol{X}) = \boldsymbol{X} - E(\boldsymbol{X})$;

$$\varphi_k(\boldsymbol{X}) = \boldsymbol{X}^k - [1 \quad \boldsymbol{X} \quad \cdots \quad \boldsymbol{X}^{k-1}] \begin{bmatrix} 1 & E(\boldsymbol{X}) & \cdots & E(\boldsymbol{X}^{k-1}) \\ E(\boldsymbol{X}) & E(\boldsymbol{X}^2) & \cdots & E(\boldsymbol{X}^k) \\ \cdots & \cdots & & \cdots \\ E(\boldsymbol{X}^{k-1}) & E(\boldsymbol{X}^k) & \cdots & E(\boldsymbol{X}^{2k-1}) \end{bmatrix}^{-1} \begin{bmatrix} E(\boldsymbol{X}^k) \\ E(\boldsymbol{X}^{k+1}) \\ \cdots \\ E(\boldsymbol{X}^{2k-1}) \end{bmatrix}, k = 1, 2, \cdots, m-1 \tag{4}$$

设式(3)左端矩阵为 Φ_X,则有:

$$\Phi_X \boldsymbol{P}_m = [E(\varphi_0(\boldsymbol{X})) \quad E(\varphi_1(\boldsymbol{X})) \quad \cdots \quad E(\varphi_{m-1}(\boldsymbol{X}))]^{\mathrm{T}} = [1 \quad 0 \quad \cdots \quad 0] \tag{5}$$

由于式(3)左端矩阵为 Φ_X 是行加权正交的,可对 Φ_X 矩阵进行归一化处理,处理后的矩阵为 $\overline{\Phi}_X = \mathrm{diag}(1, \sigma_{\varphi 1}^{-1} \cdots, \sigma_{\varphi m}^{-1}) \Phi_X$,故有 $\overline{\Phi}_X \boldsymbol{P} \overline{\Phi}_X^{\mathrm{T}} = \boldsymbol{I}$,$\overline{\Phi}_X^{-1} = P \overline{\Phi}_X^{\mathrm{T}}$。$\boldsymbol{P} \triangleq \mathrm{diag}(p_1, p_2, \cdots, p_m)$。同理有 $\overline{\varphi}_n(y_j)$ 及 $\overline{\Phi}_Y$。

定理 2 给定互不相同分量 x_i 及 y_j,$i = 1, 2, \cdots, m$; $j = 1, 2, \cdots, n$,及 2 个边缘分布列,采用加权(分布列)正交多项式矩阵和各阶相关系数,$\boldsymbol{X}$、$\boldsymbol{Y}$ 的联合概率分布列矩阵具有以下关系:

$$[f_{ij}]_{m \times n} = \overline{\Phi}_X^{\mathrm{T}} [r_{(i-1)(j-1)}]_{m \times n} \overline{\Phi}_Y = \begin{pmatrix} p_1 \\ \cdots \\ p_m \end{pmatrix} (q_1 \quad \cdots \quad q_n) + \boldsymbol{A}_{m \times (m-1)} [r_{ij}]_{(m-1) \times (n-1)} \boldsymbol{B}_{(n-1) \times n} \tag{6}$$

式中:$\boldsymbol{A}_{m\times(m-1)}=\begin{bmatrix} p_1\bar{\varphi}_1(x_1) & \cdots & p_1\bar{\varphi}_{m-1}(x_1) \\ p_2\bar{\varphi}_1(x_2) & \cdots & p_2\bar{\varphi}_{m-1}(x_2) \\ \cdots & & \cdots \\ p_m\bar{\varphi}_1(x_m) & \cdots & p_m\bar{\varphi}_{m-1}(x_m) \end{bmatrix}$为 $\boldsymbol{P}\bar{\boldsymbol{\Phi}}_X^{\mathrm{T}}$ 矩阵去掉第 1 列的矩阵;$\boldsymbol{B}_{(n-1)\times n}$ 为 $\bar{\boldsymbol{\Phi}}_Y\boldsymbol{Q}$ 矩阵去掉第 1 行的矩阵; $\boldsymbol{Q}\triangleq\mathrm{diag}(q_1,q_2,\cdots,q_n)$。

通过仿真,验证了该公式的正确性和有效性。

3　信道传输问题分析

1)级联信道相关性分析

根据文献[2]可知,若级联离散二元无记忆信道构成马尔可夫链,如有级联关系:$X\to[B_{XY}]\to Y\to[B_{YZ}]\to Z\to[B_{ZW}]\to W$,则:$B_{XZ}=B_{XY}B_{YZ}$,$B_{XW}=B_{XY}B_{YZ}B_{ZW}$。

由 $|B_{XZ}|=|r_{\varphi XZ}|\sqrt{\frac{s_1s_2\cdots s_m}{p_1p_2\cdots p_m}}=|B_{XY}B_{YZ}|=|B_{XY}||B_{YZ}|=|r_{\varphi XY}|\sqrt{\frac{q_1q_2\cdots q_m}{p_1p_2\cdots p_m}}|r_{\varphi YZ}|\sqrt{\frac{s_1s_2\cdots s_m}{q_1q_2\cdots q_m}}$,则有:$|r_{\varphi XZ}|=|r_{\varphi XY}||r_{\varphi YZ}|$,同理有: $|r_{\varphi XW}|=|r_{\varphi XY}||r_{\varphi YZ}||r_{\varphi ZW}|$。

它很好地反映了数据处理定理的重要性质:级联信道只会丢失更多的信息。线性相关系数越级联越小,反映了级联信息的丢失。

2)信道输入输出幂相关系数阵极值分析

若给定信道,当$|B_{XY}|=0$ 时,$|r_{\varphi XY}|=0$;信道没有信息耦合,信道前后变量联合无关;

$$\frac{\partial}{\partial p_i}|r_{\varphi XY}|^2_{(m-1)\times(m-1)}=|B_{XY}|^2_{m\times m}\frac{p_1p_2\cdots p_m}{q_1q_2\cdots q_m}\left[\frac{1}{p_i}-\sum_{j=1}^{m}\frac{b_{ij}}{q_j}\right]=0,\ q_j=\sum_{i=1}^{m}p_ib_{ij}\qquad j=1,2,\cdots,m$$

改变 $p_1,p_2,\cdots,p_m$ 可使$|r_{\varphi XY}|^2$ 达到最大值。可以证明:当 $p_1=\frac{\sqrt{\bar{b}_2b_2}}{\sqrt{\bar{b}_1b_1}+\sqrt{\bar{b}_2b_2}}$,$p_2=\frac{\sqrt{\bar{b}_1b_1}}{\sqrt{\bar{b}_1b_1}+\sqrt{\bar{b}_2b_2}}$时,有:$r_{XY\max}=\sqrt{\bar{b}_1\bar{b}_2}-\sqrt{b_1b_2}$,即:$|r_{XY}|\leqslant|\sqrt{\bar{b}_1\bar{b}_2}-\sqrt{b_1b_2}|$。

而 $q_1=\sqrt{\bar{b}_1b_2}\frac{\sqrt{\bar{b}_1\bar{b}_2}+\sqrt{b_1b_2}}{\sqrt{\bar{b}_1b_1}+\sqrt{\bar{b}_2b_2}}$, $q_2=\sqrt{b_1\bar{b}_2}\frac{\sqrt{\bar{b}_1\bar{b}_2}+\sqrt{b_1b_2}}{\sqrt{\bar{b}_1b_1}+\sqrt{\bar{b}_2b_2}}$,$\frac{p_1}{q_1}=\frac{\sqrt{\bar{b}_2/\bar{b}_1}}{\sqrt{\bar{b}_1\bar{b}_2}+\sqrt{b_1b_2}}$,$\frac{p_2}{q_2}=\frac{\sqrt{\bar{b}_1/\bar{b}_2}}{\sqrt{\bar{b}_1\bar{b}_2}+\sqrt{b_1b_2}}$,而此时 $\sqrt{\frac{q_1q_2}{p_1p_2}}=\sqrt{\bar{b}_1\bar{b}_2}+\sqrt{b_1b_2}$;$r_{XY\max}$相当于信道能够传输的最大的相关系数值,这时输入分布达到与信道的匹配。可与信息论课程中的信道容量所对应。r_{XY}具有与输入分布的极大值和最差信道的极小值($|B_{XY}|=0$)。

若信道对称 $b_1=b_2=b$ 时,最大的输入输出耦合系数 $r_{XY\max}=\bar{b}-b$。当 $b=0$,$r_{XY}=1$ 时,$r_{XY\max}=\sqrt{\bar{b}_1\bar{b}_2}-\sqrt{b_1b_2}$即:$|r_{XY}|\leqslant|\sqrt{\bar{b}_1\bar{b}_2}-\sqrt{b_1b_2}|$。

它的特点类似于平均互信息的极值性,但比平均互信息的(非线性)计算简单得多,易于理解,且概念非常清楚,二者结论非常相似。由于二元数据通信的普遍性,其结论具有重要意义。

4　总结

根据相关性与独立的等价概念,给出了有限离散随机变量的二维联合概率分布的一种描述方法。当给出 2 个随机变量的边缘分布及合适的有限阶互协方差函数(或高阶相关系数、正交多项式系数),可得两维联合概率分布列,使 2 个随机变量的相关性与独立性有一系统地描述,对这一问题有一个清晰认识,在教学和科研中都具有重要作用。

参考文献:

[1]　谢兴武,李宏伟. 概率统计释难解疑[M]. 北京:科学出版社,2007.

XIE Xingwu, LI Hongwei. The Solution to Possibility and Statistics[M]. Beijing: Science Press, 2007. (in Chinese)
[2] 傅祖芸. 信息论——基础理论与应用[M]. 北京:电子工业出版社,2001.
FU Zuyun. The Information: Basic Theory and Application[M]. Beijing: Electronics Industry Press, 2001. (in Chinese)
[3] 郑阿奇. MATLAB 实用教程[M]. 北京:电子工业出版社,2004.
ZHENG Aqi. MATLAB Tutorial[M]. Beijing: Electronics Industry Press, 2004. (in Chinese)
[4] 陈俊雅,王秀英. 概率论与数理统计中的反例[M]. 天津:科学技术出版社,1993.
CHEN Junya, WANG Xiuying. Counterexamples of Possibility and Mathematical Statistics[M]. Tianjin: Science and Technology Publishing Company, 1993. (in Chinese)
[5] 朱华,黄辉宁. 随机信号分析[M]. 北京:北京理工大学出版社,2002.
ZHU Hua, HUANG Huining. Random Signal Analysis[M]. Beijing: Beijing Science University Publishing Company, 2002. (in Chinese)
[6] 柳重堪. 正交函数及其应用[M]. 北京:国防工业出版社,1982.
LIU Zhongkan. Orthonormal Function's Application[M]. Beijing: National Defence Industry Press, 1982. (in Chinese)
[7] IEEE 1109.01 – 2007. The $\alpha-\mu$ Distribution: A Physical Fading Model for the Stacy Distribution[S].
[8] IEEE 1109.03 – 2007. Performance Analysis of MIMO – MRC in Double – Correlated Rayleigh Environments[S].
[9] IEEE 1109.07 – 2008. On the Mutual Information Distribution of OFDM – Based Spatial Multiplexing: Exact Variance and Outage Approximation[S].
[10] IEEE 1109.2007. On the Multivariate Weibull Fading Model with Arbitrary Correlation Matrix[S].

(编辑:徐楠楠)

Two Dimensional Joint Probability Distribution of Multi – states and the Research on the Information Transfers

LIU Jin – mang[1], ZHANG Ya – xun[1], YUAN Xiu – jiu[2]
(1. Missile Institute, Air Force Engineering University, Sanyuan 713800, Shaanxi, China; 2. Sciences Institute, Air Force Engineering University, Xi'an 710051, China)

Abstract: According to the theory that the independence is equivalent to the uncorrelation, a description of joint probability distribution of finite discrete random variable is found. Through extending the values of two – state to that of multi – state, when two random variables' marginal distributions and right high steps correlative coefficients are given, two dimensional joint probability row can be got to make independence and uncorrelation have an integrated description and expressive form.. This method is verified through simulations. By the way, the expression obtained is applied to the analysis of data channel communications. The in – out information coupling is described by channel's port correlative coefficients, which is a fine supplement to the information theory. This method is corresponding with information theory's average mutual information and data processing theorem, and a perfect formula and ideal characteristics are obtained.

Key words: correlative coefficient; joint probability distribution; channel matrix