

Promote the Use of Two-dimensional Continuous Random Variables Conditional Distribution

Feixue Huang

Department of Economics, Dalian University of Technology, Dalian 116024, China

Tel: 86-411-8470-7210 E-mail: software666@163.com

Cheng Li

Department of Economics, Dalian University of Technology, Dalian 116024, China

E-mail: lchengzi2008@yahoo.com.cn

Abstract

Conditional distribution reflects the dependency link among random variables, but two-dimensional random variables Conditional Distribution has some limitations. In order to rich the content of conditional distribution this paper gives the extension of conditional distribution and examples in the case of continuous random variables. For the given definition of conditional distribution of three-dimensional continuous random variables, it also gives the proof. This article obtains the extension strictly in accordance with the definition of two-dimensional random variables and it uses the theory of conditional probability to get the proof. So it can get conditional distributions after changing the condition to enrich the contents of the conditional distribution.

Keywords: Promote the use, Continuous random variables, Conditional distribution, Multi-dimensional random variables

Introduction

The relationships of two-dimensional random variables (X, Y) are mainly divided into two types: independence and dependence. The more evaluations X and Y have, the more conditional distributions will be there. Each conditional distribution describes one state-specific distribution from a side. So the contents of conditional distribution are richer and its applications are broader. For in many issues values of concerned random variables tend to influence each other, this makes conditional distribution as a powerful tool for studying dependencies among variables(Mao Shi-song, Cheng Yi-ming, Pu Xiao-long. 2004). Conditional distribution for random variables derives from conditional probability for random events, so there is a close relationship between the two and the approaches to handle them are the same, but conditional distribution is more complex to deal with(Feng Tai, Liu De-yin. 1985)(Liang Yi. 1998).

In recent years, the main research direction in China is the research of conditional eigenvalues and the extension and application of conditional distribution. Conditional eigenvalues are mainly pointed to conditional expectation, which is expectation under conditional distribution. In this area the main contributions are Wang Cheng, Zou Hailei gave the definition and researched the characteristics of random variables conditional particular function based on measurement and integral theory(Wang Cheng, Xou Hai-lei. 2008); Zhang Mei used minimum mean-squared error to solve one kind problems of best prediction and took examples to analyze the application of conditional expectation in practical prediction problems(Zhang Mei. 2006); Xu hui and Wu Guogeng educed total probability formula of discrete and continuous random variables based on conditional expectation and indicative function I_A of random events A(Xu Hui, Wu Guo-gen. 2004). For extension of conditional distribution, primacy researches are Cheng Feiyue generalized the condition distribution of the poisson process arrival time(Chen Fei-yue. 2006); Yang Jingping, etc. investigated the marginal recursive equations on excess-of-loss reinsurance treaty under the assumption that the number of claims belongs to the family consisting of Poisson, binomial and negative binomial, and that the severity distribution hand bounded continuous density function(Yang Jing-ping, Wang Xiao-qian, Cheng Shi-hong. 2006); Hu Duanping gave a express formula of distribution for elliptically contoured matrix distributions and showed that condition distribution of elliptically contoured matrix distribution is elliptically contoured distribution yet(Hu Duan-ping. 2001). In the above study, they focused on the application of the conditional distribution showing the research significance of the conditional distribution, but there is few studies researching the nature of the conditional distribution.

In comparison, the studies of conditional distribution abroad are more in-depth. For example, Rodney C.L. Wolff, etc studied the methods of evaluating conditional distribution function; Jushan Bai investigated the dynamic model of testing parametric conditional distributions; Peter Hall and Qiwei Yao discussed approximating

conditional distribution function using dimension reduction; Bruce E. Hansen studied nonparametric estimation of smooth conditional distributions; Persi Diaconis, Bernd Sturmfels analyzed conditional distributions using algebraic algorithms for sampling. They have shown that the researches of conditional distribution are multi-faceted and more complex while make against undergraduate teaching.

In this respect the paper begins to discuss and analyze from the basic content of conditional distribution and educates general formulas with certain conditions on the basis of the definition of conditional distribution. It first starts from the two-dimensional random variable conditional distribution in the case of continuous random variables and changes the given conditions to obtain the extensions of conditional distribution and then gives extensions of conditional distribution when there are three-dimensional random variables. And it proves the definition of the given three-dimensional continuous random variables conditional distribution. This paper is to solve the conditional distribution of multidimensional random variables under the given conditions and its results can be used for teaching, expending the knowledge of the conditional distribution and facilitating people's calculations.

1. Extension of continuous random variables conditional distribution

Extension1.1 Set $p_{X,Y}(x,y)$ the joint density function of two-dimensional random variables (X,Y) and let $Z = X + Y$, of which the density function is $p_Z(z)$. $p(x,y)$, $p_Z(z)$ are continuous everywhere. When $X + Y = z$, the conditional density function of X is $p_{X|Z}(x|z) = \frac{p_{X,Y}(x, z-x)}{p_Z(z)}$ and the conditional density

function of Y is $p_{Y|Z}(y|z) = \frac{p_{X,Y}(z-y, y)}{p_Z(z)}$.

To prove: From the Variable Transformation, note $\begin{cases} u = x, \\ v = x + y \end{cases}$, its inverse function is $\begin{cases} x = u, \\ y = v - u \end{cases}$, then

$$J = \begin{vmatrix} \frac{\partial x}{\partial u} & \frac{\partial y}{\partial u} \\ \frac{\partial x}{\partial v} & \frac{\partial y}{\partial v} \end{vmatrix} = \begin{vmatrix} 1 & -1 \\ 0 & 1 \end{vmatrix} = 1.$$

So the joint density function of (U,V) is

$$p_{U,V}(u,v) = p_{X,Y}(x(u,v), y(u,v)) |J|. \quad (1)$$

That the joint density function of (X,Z) is $p_{X,Z}(x,z) = p_{X,Y}(x, z-x)$.

$$\begin{aligned} P(X \leq x | X + Y = z) &= \lim_{h \rightarrow 0} P(X \leq x | z \leq Z \leq z + h) \\ &= \lim_{h \rightarrow 0} \frac{P(X \leq x, z \leq Z \leq z + h)}{P(z \leq Z \leq z + h)} \\ &= \lim_{h \rightarrow 0} \frac{\int_{-\infty}^x \int_z^{z+h} p_{X,Z}(x,z) dz dx}{\int_z^{z+h} p_Z(z) dz} \\ &= \lim_{h \rightarrow 0} \frac{\int_{-\infty}^x \left\{ \frac{1}{h} \int_z^{z+h} p_{X,Z}(x,z) dz \right\} dx}{\frac{1}{h} \int_z^{z+h} p_Z(z) dz}. \end{aligned} \quad (2)$$

From mean value theorem of integrals we get

$$\lim_{h \rightarrow 0} \frac{1}{h} \int_z^{z+h} p_Z(z) dz = p_Z(z), \quad (3)$$

$$\lim_{h \rightarrow 0} \frac{1}{h} \int_z^{z+h} p_{X,Z}(x,z) dz = p_{X,Z}(x,z). \quad (4)$$

So $P(X \leq x | X + Y = z) = \int_{-\infty}^x \frac{p_{X,Z}(x,z)}{p_Z(z)} dx = \int_{-\infty}^x \frac{p_{X,Y}(x, z-x)}{p_Z(z)} dx$, where the conditional density

function is $p_{X|Z}(x|z) = \frac{p_{X,Y}(x, z-x)}{p_Z(z)}$.

Example1.2 Set two-d random variable $(X, Y) \sim N(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, 0)$ and $X+Y=U$. Given the condition of $X+Y=u$ to solve the condition density of X and Y .

To solve: For the marginal distribution of two-dimensional normal distribution is one-dimensional normal distribution, viz. $X \sim N(\mu_1, \sigma_1^2)$, $Y \sim N(\mu_2, \sigma_2^2)$.

Because X and Y are independent, so $U = X+Y \sim N(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2)$, where the density function of U is

$$p_U(u) = \frac{1}{\sqrt{2\pi(\sigma_1^2 + \sigma_2^2)}} \exp\left\{-\frac{1}{2} \frac{(u - \mu_1 - \mu_2)^2}{\sigma_1^2 + \sigma_2^2}\right\}. \quad (5)$$

For X and Y are independent, we get

$$\begin{aligned} p_{X,Y}(x, u-x) &= p_X(x) p_Y(u-x) \\ &= \frac{1}{\sqrt{2\pi}\sigma_1} e^{-\frac{(x-\mu_1)^2}{2\sigma_1^2}} \cdot \frac{1}{\sqrt{2\pi}\sigma_2} e^{-\frac{(u-x-\mu_2)^2}{2\sigma_2^2}} \\ &= \frac{1}{2\pi\sigma_1\sigma_2} e^{-\frac{(x-\mu_1)^2}{2\sigma_1^2} - \frac{(u-x-\mu_2)^2}{2\sigma_2^2}}, \end{aligned} \quad (6)$$

$$\begin{aligned} p_{X,Y}(u-y, y) &= p_X(u-y) p_Y(y) \\ &= \frac{1}{\sqrt{2\pi}\sigma_1} e^{-\frac{(u-y-\mu_1)^2}{2\sigma_1^2}} \cdot \frac{1}{\sqrt{2\pi}\sigma_2} e^{-\frac{(y-\mu_2)^2}{2\sigma_2^2}} \\ &= \frac{1}{2\pi\sigma_1\sigma_2} e^{-\frac{(u-y-\mu_1)^2}{2\sigma_1^2} - \frac{(y-\mu_2)^2}{2\sigma_2^2}}. \end{aligned} \quad (7)$$

$$\begin{aligned} \text{So } p_{X|U}(x|u) &= \frac{p_{X,Y}(x, u-x)}{p_U(u)} \\ &= \frac{\frac{1}{2\pi\sigma_1\sigma_2} e^{-\frac{(x-\mu_1)^2}{2\sigma_1^2} - \frac{(u-x-\mu_2)^2}{2\sigma_2^2}}}{\frac{1}{\sqrt{2\pi(\sigma_1^2 + \sigma_2^2)}} \exp\left\{-\frac{1}{2} \frac{(u - \mu_1 - \mu_2)^2}{\sigma_1^2 + \sigma_2^2}\right\}} \\ &= \frac{\sqrt{2\pi(\sigma_1^2 + \sigma_2^2)}}{2\pi\sigma_1\sigma_2} \exp\left\{-\frac{(x-\mu_1)^2}{2\sigma_1^2} - \frac{(u-x-\mu_2)^2}{2\sigma_2^2} + \frac{1}{2} \frac{(u - \mu_1 - \mu_2)^2}{\sigma_1^2 + \sigma_2^2}\right\}; \end{aligned} \quad (8)$$

$$\begin{aligned} p_{Y|U}(y|u) &= \frac{p_{X,Y}(u-y, y)}{p_U(u)} = \frac{\frac{1}{2\pi\sigma_1\sigma_2} e^{-\frac{(u-y-\mu_1)^2}{2\sigma_1^2} - \frac{(y-\mu_2)^2}{2\sigma_2^2}}}{\frac{1}{\sqrt{2\pi(\sigma_1^2 + \sigma_2^2)}} \exp\left\{-\frac{1}{2} \frac{(u - \mu_1 - \mu_2)^2}{\sigma_1^2 + \sigma_2^2}\right\}} \\ &= \frac{\sqrt{2\pi(\sigma_1^2 + \sigma_2^2)}}{2\pi\sigma_1\sigma_2} \exp\left\{-\frac{(u-y-\mu_1)^2}{2\sigma_1^2} - \frac{(y-\mu_2)^2}{2\sigma_2^2} + \frac{1}{2} \frac{(u - \mu_1 - \mu_2)^2}{\sigma_1^2 + \sigma_2^2}\right\}. \end{aligned} \quad (9)$$

According to the formula, it is easy to be drawn.

2. Proof of three-dimensional continuous random variables conditional distribution

Definition: Set three-dimensional continuous random variables (X_1, X_2, X_3) . Its density function $f(x_1, x_2, x_3)$ and each marginal density function are continuous everywhere. Given the condition of $X_1 = x_1$ and $X_2 = x_2$, the conditional distribution of X_3 is

$$F(x_3 | x_1, x_2) = \frac{\int_{-\infty}^{x_3} f(x_1, x_2, x_3) dx_3}{\int_{-\infty}^{+\infty} f(x_1, x_2, x_3) dx_3}, \quad (10)$$

and conditional density function is

$$f(x_3 | x_1, x_2) = \frac{f(x_1, x_2, x_3)}{\int_{-\infty}^{+\infty} f(x_1, x_2, x_3) dx_3} \quad (\text{M·Fes. 1962}). \quad (11)$$

To prove: According to the proof of two-dimensional continuous random variables conditional distribution, we can give the conditional distribution of three-dimensional continuous random variables.

$$\begin{aligned} P(X_3 \leq x_3 | X_1 = x_1, X_2 = x_2) &= \lim_{h \rightarrow 0} \lim_{k \rightarrow 0} P(X_3 \leq x_3 | x_1 \leq X_1 \leq x_1 + h, x_2 \leq X_2 \leq x_2 + k) \\ &= \lim_{h \rightarrow 0} \lim_{k \rightarrow 0} \frac{P(X_3 \leq x_3, x_1 \leq X_1 \leq x_1 + h, x_2 \leq X_2 \leq x_2 + k)}{P(x_1 \leq X_1 \leq x_1 + h, x_2 \leq X_2 \leq x_2 + k)} \\ &= \lim_{h \rightarrow 0} \lim_{k \rightarrow 0} \frac{\int_{-\infty}^{x_3} \int_{x_1}^{x_1+h} \int_{x_2}^{x_2+k} f(u, v, t) dv du dt}{\int_{x_1}^{x_1+h} \int_{x_2}^{x_2+k} f_{X_1, X_2}(u, v) dv du} \\ &= \lim_{h \rightarrow 0} \lim_{k \rightarrow 0} \frac{\int_{-\infty}^{x_3} \left\{ \frac{1}{h} \int_{x_1}^{x_1+h} \left\{ \frac{1}{k} \int_{x_2}^{x_2+k} f(u, v, t) dv \right\} du \right\} dt}{\frac{1}{h} \int_{x_1}^{x_1+h} \left\{ \frac{1}{k} \int_{x_2}^{x_2+k} f_{X_1, X_2}(u, v) dv \right\} du}. \end{aligned} \quad (12)$$

From mean value theorem of integrals we get

$$\lim_{k \rightarrow 0} \frac{1}{k} \int_{x_2}^{x_2+k} f(u, v, t) dv = f(u, x_2, t), \quad (13)$$

$$\lim_{h \rightarrow 0} \frac{1}{h} \int_{x_1}^{x_1+h} f(u, x_2, t) du = f(x_1, x_2, t), \quad (14)$$

$$\lim_{k \rightarrow 0} \frac{1}{k} \int_{x_2}^{x_2+k} f_{X_1, X_2}(u, v) dv = f_{X_1, X_2}(u, x_2), \quad (15)$$

$$\lim_{h \rightarrow 0} \frac{1}{h} \int_{x_1}^{x_1+h} f_{X_1, X_2}(u, x_2) du = f_{X_1, X_2}(x_1, x_2). \quad (16)$$

Then we can get conditional distribution function

$$P(X_3 \leq x_3 | X_1 = x_1, X_2 = x_2) = \int_{-\infty}^{x_3} \frac{f(x_1, x_2, t)}{f_{X_1, X_2}(x_1, x_2)} dt, \quad (17)$$

$$F(x_3 | x_1, x_2) = \int_{-\infty}^{x_3} \frac{f(x_1, x_2, t)}{f_{X_1, X_2}(x_1, x_2)} dt = \frac{\int_{-\infty}^{x_3} f(x_1, x_2, x_3) dx_3}{\int_{-\infty}^{+\infty} f(x_1, x_2, x_3) dx_3}. \quad (18)$$

There is following conditional density function

$$f(x_3 | x_1, x_2) = \frac{f(x_1, x_2, x_3)}{\int_{-\infty}^{+\infty} f(x_1, x_2, x_3) dx_3}. \quad (19)$$

In a similar way, under the condition of $X_1 = x_1$, conditional distribution and conditional density of X_2 and X_3 are

$$F(x_3, x_2 | x_1) = \frac{\int_{-\infty}^{x_3} \int_{-\infty}^{x_2} f(x_1, x_2, x_3) dx_2 dx_3}{\int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f(x_1, x_2, x_3) dx_2 dx_3}, \quad (20)$$

$$f(x_3, x_2 | x_1) = \frac{f(x_1, x_2, x_3)}{\int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f(x_1, x_2, x_3) dx_2 dx_3}. \quad (21)$$

Example Three-dimensional continuous random variables (X, Y, Z) subject to Uniform distribution on D, of which the density function is

$$p(x, y, z) = \begin{cases} \frac{3}{4\pi r^3}, & x^2 + y^2 + z^2 \leq r^2; \\ 0, & x^2 + y^2 + z^2 > r^2. \end{cases}$$

I) Given the condition of (X, Y) to solve conditional distribution of Z .

To solve: We can get the density function of (X, Y) from the subject:

$$p_{X,Y}(x, y) = \int_{-\infty}^{+\infty} p(x, y, z) dz = \begin{cases} \int_{-\sqrt{r^2-x^2-y^2}}^{\sqrt{r^2-x^2-y^2}} \frac{3}{4\pi r^3} dy = \frac{3}{2\pi r^3} \sqrt{r^2-x^2-y^2} & x^2 + y^2 \leq r^2; \\ 0 & \text{others.} \end{cases} \quad (22)$$

So when $x^2 + y^2 \leq r^2$, there is

$$p_{Z|X,Y}(z|x, y) = \frac{p(x, y, z)}{p_{X,Y}(x, y)} = \begin{cases} \frac{1}{2\sqrt{r^2-x^2-y^2}} & -\sqrt{r^2-x^2-y^2} \leq z \leq \sqrt{r^2-x^2-y^2}; \\ 0 & \text{others.} \end{cases}, \quad (23)$$

II) Given the condition of Z to solve conditional distribution of (X, Y) .

To solve: We can get the density function of Z from the subject:

$$p_Z(z) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} p(x, y, z) dx dy = \begin{cases} \int_0^{2\pi} \int_0^{\sqrt{r^2-z^2}} \frac{3}{4\pi r^3} \rho d\rho = \frac{3(r^2-z^2)}{4r^3} & -r \leq z \leq r; \\ 0 & \text{others.} \end{cases} \quad (24)$$

So when $-r \leq z \leq r$, there is

$$p_{X,Y|Z}(x, y|z) = \frac{p(x, y, z)}{p_Z(z)} = \begin{cases} \frac{1}{\pi(r^2-z^2)} & x^2 + y^2 \leq r^2 - z^2; \\ 0 & \text{others.} \end{cases} \quad (25)$$

3. Conclusions

This paper mainly discusses conditional distributions of multidimensional random variables and its related examples given certain conditions in the case of continuous situation. It changes the original condition of one fixed variable into more complex conditions, for example the condition that the sum of two variables is fixed in two-dimensional situation, which is easy to solve with discrete random variables and can be extended in continuous situation. This paper gives the proof.

While in the condition of three-dimensional continuous random variables, this article proves one variable's conditional distribution when the other two are fixed to make the conditional distribution theory precise.

This article gets the above results strictly according to two-dimensional random variables conditional distribution and applies the theory of conditional probability in the proof.

Conditional distribution can be applied in the life and work to resolve practical problems. The application of using conditional distribution theory to carry out scientific analysis and calculations with real data is an

important reflection of the usefulness of conditional distribution. In addition, for there are more and more discussions about multidimensional random variables in reality this paper extends the conditional distribution to provide new train of thought for the research in some extent and it aims at enriching the content of conditional distribution, deepening the understanding of it and applying it well in practical.

Acknowledgements

This work is supported by the funds project under the Ministry of Education of the PRC for young people who are devoted to the researches of humanities and social sciences under Grant No. 09YJC790025; software + X research of Dalian University of Technology under Grant No. 842301.

References

- Bruce E. Hansen. Nonparametric Estimation of Smooth Conditional Distributions [EB/OL]. (2004-05)[2009-7-30]. <http://www.ssc.wisc.edu/~bhansen/papers/cdf.pdf>.
- Chen Fei-yue. (2006). Generalization of the Condition Distribution of the Poisson Process Arrival Time. *Journal of Changsha University of Electric Power (Natural Science)*, 2006, 21(2):75-79.
- Feng Tai, Liu De-yin. (1985). Conditional Probability and Conditional Distribution. *Distance Education in China*, 1985, (10):8-11.
- Hu Duan-ping. (2001). The Condition Distribution of Elliptically Contoured Matrix Distributions. *Communication on Applied Mathematics and Computation*, 2001, 15(2):41-44.
- Jushan Bai. (2003). Testing Parametric Conditional Distributions of Dynamic Models. *The Review of Economics and Statistics*, 2003, 85(3):531-549.
- Liang Yi. (1998). Conditional Probability and Conditional Distribution. *Journal of Higher Correspondence Education (Natural Sciences)*, 1998,(5):29-31,52.
- M-Fes. (1962). *Probability and Mathematical Statistics*. Wang fu-bao, translate. Shanghai: Science & Technology Press of Shanghai, 1962.
- Mao Shi-song, Cheng Yi-ming, Pu Xiao-long. (2004). *Probability Theory and Mathematical Statistics Tutorial*. Beijing: Advanced Education Press, 2004.
- Persi Diaconis, Bernd Sturmfels. (1998). Algebraic algorithms for sampling from conditional distributions. *The Annals of statistics*, 1998,26(1):363-397.
- Peter Hall, Qiwei Yao. (2005). Approximating Conditional Distribution Functions Using Dimension Reduction. *The Annals of statistics*, 2005, 33(3):1404-1421.
- Rodney C.L. Wolff, Qiwei Yao, Peter Hall. (1999). Methods for Estimating a Conditional Distribution Function. *Journal of the American Statistical Association*, 1999, 94(445):154-163.
- Wang Cheng, Xou Hai-lei. (2008). Some Properties of Conditional Characteristic function. *Journal of China Jiliang University*, 2008, 19(4):360-362.
- Xu Hui, Wu Guo-gen. (2004). Full Probability Formula Based on the Conditional Expectation and its Applications]. *Journal of East China Institute of Technology*, 2004, 27(2):193-195.
- Yang Jing-ping, Wang Xiao-qian, Cheng Shi-hong. (2006). Conditional Recursive Equations on Excess-of-Loss Reinsurance. *Applied Mathematics and Mechanics*, 2006, 27(8):931-939.
- Zhang Mei. (2006). Using Conditional Expectation to solve Problems of Best Prediction. *Journal of Shanxi Institute of Education*, 2006, 22(2):83-84.