CRRAO Advanced Institute of Mathematics, Statistics and Computer Science (AIMSCS)

Research Report

Author (s): G.G. Hamedani and B.L.S. Prakasa Rao

Title of the Report: **CHARACTERIZATIONS OF PROBABILITY DISTRIBUTIONS THROUGH SUB-INDEPENDENCE, MAX-SUB-INDEPENDENCE AND CONDITIONAL SUB-INDEPENDENCE**

Research Report No.: **RR2015-01**

Date: **January 21, 2015**

Prof. C R Rao Road, University of Hyderabad Campus,
Gachibowli, Hyderabad-500046, INDIA.
www.crraoaimscs.org

Journal of the Indian Statistical Association
Vol.?? No. ?, 201?,

Characterizations of probability distributions through sub-independence, max-sub-independence and conditional sub-independence

G.G. Hamedani

Department of Mathematics, Statistics and Computer Science, Marquette University, Milwaukee, WI 53201-1881, USA

B.L.S. Prakasa Rao

C R Rao Advanced Institute of Mathematics, Statistics and Computer Science, Hyderabad 500046, India

Abstract

Limit theorems as well as other well-known results in probability and statistics are often based on the distribution of the sums of independent random variables. The concept of sub-independence, which is much weaker than that of independence, is shown to be sufficient to yield the conclusions of these theorems and results. It also provides a measure of dissociation between two random variables which is much stronger than uncorrelatedness. We define the concepts of max-sub-independence, conditionally sub-independence and present certain characterizations of distributions based on these concepts as well as that of sub-independence.

Key Words : *Sub-independence, Max-sub-independence, Conditional-sub-independence, Characterization.*

Received : month year

1 Introduction

Limit theorems as well as other well-known results in probability and statistics are often based on the distribution of the sums of independent (and often identically distributed) random variables rather than the joint distribution of the summands. Therefore, the full force of independence of the summands will not be required. In other words, it is the convolution of the marginal distributions which is needed, rather than the joint distribution of the summands which, in the case of independence, is the product of the marginal distributions. This is precisely the reason for the statement: "why assume independence when you can get by with sub-independence".

The concept of sub-independence can help to provide solution for some modeling problems where the variable of interest is the sum of a few components. Examples include a regression model $Y = g(X) + \varepsilon$ where $g(X)$ and ε are uncorrelated and may not be independent. For example, in Bazargan et al. (2007), the return value of significant wave height (Y) is modeled by the sum of a cyclic function of random delay $D, \hat{g}(D)$ and a residual term $\hat{\varepsilon}$. They found that the two components are uncorrelated but not independent and used sub-independence to compute the distribution of the return value. For a detailed study of the concept of sub-independence, we refer the reader to Hamedani (2013). The concepts of conditional independence, conditional association and conditional mixing for sequences of random variables are studied in Prakasa Rao (2009). It is known that if a sequence of random varibles is independent, it need not be conditionally independent and vice versa. We introduce the notion of conditional-sub-independence and obtain some characterizations of distributions for random variables which are conditionally sub-independent. In addtion, we introduce the property of max-sub-independence for a finite set of random variables and obtain a characterization of distributions through max-sub-independence property.

2 Preliminaries

Definition 2.1: Two random variables X and Y are said to be *sub-independent* if

$$\phi_{X,Y}(t,t) = \phi_{X+Y}(t) = \phi_X(t)\phi_Y(t), t \in R$$

where $\phi_X(.)$ denotes the characteristic function of the random variable X and $\phi_{X,Y}(.,.)$ denotes the joint characteristic function of the bivariate random vector (X,Y). A set of random variables $X_1, X_2, \ldots, X_k$ are said to be *sub-independent* if every subset of them is sub-independent, that is

$$\phi_{X_{i_1},\ldots,X_{i_m}}(t,\ldots,t) = \phi_{X_{i_1}+\ldots+X_{i_m}}(t) = \Pi_{r=1}^m \phi_{X_{i_r}}(t), t \in R$$

for all $\{i_1,\ldots,i_m\} \subset \{1,\ldots,k\}$ and for all $m, 2 \leq m \leq k$.

Definition 2.2: Two random variables X and Y are said to be *conditionally sub-independent* given $Z = z$ if

$$\phi_{X,Y}(t,t;z) = \phi_{X+Y}(t;z) = \phi_X(t;z)\phi_Y(t;z), t \in R$$

where $\phi_X(.;z)$ denotes the conditional characteristic function of the random variable X given $Z = z$ and $\phi_{X,Y}(.,.;z)$ denotes the conditional joint characteristic function of the bivariate random vector (X,Y) given $Z = z$. A set of random variables $X_1, X_2, \ldots, X_k$ are said to be *conditionally sub-independent* given $Z = z$ if every subset of them is conditionally sub-independent given $Z = z$, that is

$$\phi_{X_{i_1},\ldots,X_{i_m}}(t,\ldots,t;z) = \phi_{X_{i_1}+\ldots+X_{i_m}}(t;z) = \Pi_{r=1}^m \phi_{X_{i_r}}(t;z), t \in R$$

for all $\{i_1,\ldots,i_m\} \subset \{1,\ldots,k\}$ and for all $m, 2 \leq m \leq k$.

Definition 2.3 : Two random variables X and Y are said to be *max-sub-independent* if

$$P(\max(X,Y) \leq x) = P(X \leq x, Y \leq x) = P(X \leq x)P(Y \leq x), x \in R.$$

A set of random variables $X_1, X_2, \ldots, X_k$ are said to be *max-sub-independent* if every subset of them is max-sub-independent, that is

$$P(X_{i_1} \leq x, \ldots, X_{i_m} \leq x) = \Pi_{r=1}^{m} P(X_{i_r} \leq x), x \in R$$

for all $\{i_1, \ldots, i_m\} \subset \{1, \ldots, k\}$ and for all $m, 2 \leq m \leq k$.

The concept of sub-independence for random variables was introduced by Durairajan (1979). For an expository perspective on sub-independent random variables, see Hamedani (2013). Characterizations of probability distributions based on conditionally independent random variables were discussed in Prakasa Rao (2009, 2013). We introduce now the concept of max-sub-independence following the ideas of max-infinitely divisible and max-stable random variables. Note that the concept of sub-independence is weaker than the concept of independence for random variables and conditional independence does not imply independence and vice versa (cf. Prakasa Rao (2009)).

3 Characterizations

We now state and prove three characterization theorems for probability distributions when sub-independence, conditional sub-independence or max-sub-independence is present.

Theorem 3.1: Let X_1, X_2 and X_3 be sub-independent identically distributed random variables. Let $Y_1 = \alpha X_1 + X_2$ and $Y_2 = (1-\alpha)X_1 + X_3$. Suppose the characteristic function of $Y_1 + Y_2$ does not vanish. Then the distribution of the random variable $Y_1 + Y_2$ determines the characteristic function of X_1 up to a cube root of unity.

Proof : It is obvious that the joint characteristic function of (Y_1, Y_2) is

$$\begin{aligned} \phi_{Y_1,Y_2}(t_1, t_2) &= E[\exp\{it_1(\alpha X_1 + X_2) + it_2((1-\alpha)X_1 + X_3)\}], t_1, t_2 \in R \\ &= E[\exp\{i(\alpha t_1 + (1-\alpha)t_2)X_1 + iX_2 t_1 + iX_3 t_2\}], t_1, t_2 \in R \end{aligned}$$

$$= \phi_{X_1,X_2,X_3}(\alpha t_1 + (1-\alpha)t_2), t_1, t_2), t_1, t_2 \in R.$$

In particular, choosing $t_1 = t_2 = t \in R$, we get that

$$\begin{aligned} \phi_{Y_1,Y_2}(t,t) &= \phi_{Y_1+Y_2}(t) = \phi_{X_1,X_2,X_3}(t,t,t) \\ &= \phi_{X_1}(t)\phi_{X_2}(t)\phi_{X_3}(t) \text{ (by sub-independence of } X_1, X_2 \text{ and } X_3) \\ &= [\phi_{X_1}(t)]^3 \text{ (by the identical distribution of } X_1, X_2 \text{ and } X_3) \end{aligned}$$

Let W_1, W_2 and W_3 be identically distributed random variables which are sub-independent. Let $V_1 = \alpha W_1 + W_2$ and $V_2 = (1-\alpha)W_1 + W_3$. Then

$$\phi_{V_1,V_2}(t,t) = \phi_{V_1+V_2}(t) = [\phi_{W_1}(t)]^3, t \in R. \tag{3.1}$$

Suppose the distribution of $Y_1 + Y_2$ is the same as that of $V_1 + V_2$. Then the characteristic functions of $Y_1 + Y_2$ and $V_1 + V_2$ are the same. The relations obtained above imply that the characteristic functions of the random variables X_1 and W_1 are non zero and have the same third power. Hence they are equal up to a cube root of unity.

The next theorem is analogous to Theorem 3.1 but deals with conditional sub-independent random variables. We omit the proof of Theorem 3.2 as it can be proved along the lines of proof of Theorem 3.1 using conditional characteristic functions.

Theorem 3.2: Let X_1, X_2 and X_3 be conditionally sub-independent identically distributed random variables given a random variable $Z = z$. Let $Y_1 = \alpha X_1 + X_2$ and $Y_2 = (1-\alpha)X_1 + X_3$. Suppose the conditional characteristic function of $Y_1 + Y_2$ does not vanish. Then the distribution of the random variable $Y_1 + Y_2$ given $Z = z$ determines the conditional characteristic function of X_1 given $Z = z$ up to a cube root of unity..

The following result characterizes probability distributions through the max-sub-independence property.

Theorem 3.3: Let X_1, X_2, X_3 be max-sub-independent identically distributed random variables. Let $Y_1 = \max(X_1, X_2)$ and $Y_2 = \max(X_1, X_3)$. Then the joint distribution of $\max(Y_1, Y_2)$ uniquely determines the distributions of X_1.

Proof : It is obvious that

$$\begin{aligned} P(\max(Y_1, Y_2) \leq x) &= P(Y_1 \leq x, Y_2 \leq x) \\ &= P(\max(X_1, X_2) \leq x, \max(X_1, X_3) \leq x) \\ &= P(X_1 \leq x, X_2 \leq x, X_3 \leq x) \\ &= P(X_1 \leq x)P(X_2 \leq x)P(X_3 \leq x) \text{ (by max-sub-independence)} \\ &= [P(X_1 \leq x)]^3 \text{ (since } X_1, X_2, X_3 \text{ are identically distributed)} \end{aligned}$$

and hence

$$P(X_1 \leq x) = [P(\max(Y_1, Y_2) \leq x)]^{1/3}, x \in R.$$

Remark: The result in Theorem 3.3 also holds for two max-sub-independent identically distributed random variables X_1 and X_2 in the sense that the distribution of $\max(X_1, X_2)$ determines the distribution of the component X_1 uniquely.

4 Examples

We now present some examples of sub-independent random variables and remarks on max-sub-independent random variables.

Example 4.1: Let X and Y be identically distributed random variables with support on the integers $\{1, 2, 3\}$ and joint probabilities $p_{ij} = P(X = i, Y = j)$ given by $p_{11} = p_{22} = p_{33} = 1/9, p_{21} = p_{32} = p_{13} = 2/9$ and $p_{12} = p_{23} = p_{31} = 0$. It can be checked that the random variables x and Y are sub-independent where as the random variables X and $-Y$ are not sub-independent.

Example 4.2 : Let $(X, Y$ be a bivariate random vector with joint characteristic function given by

$$\phi_{X,Y}(t_1, t_2) = \exp\{-(t_1^2 + t_2^2)/2\}[1 + \beta t_1 t_2 (t_1 - t_2)^2 \exp\{(t_1^2 + t_2^2)/4\}]$$

for $-\infty < t_1, t_2 < \infty$ for suitable constant β. It can be shown the random variables X and Y are sub-independent with standard normal distribution and the random variable $X + Y$ has the normal distribution with mean zero and variance 2. However the random variables X and $-Y$ are not sub-independent and the random variables $X - Y$ does not have a normal distribution.

For additional examples and a comprehensive survey of properties of sub-independent random variables, see Hamedani (2013).

Recall that two random variables X and Y are said to be max-sub-independent if

$$F_Z(z) = F_X(z)F_Y(z), z \in R$$

where $Z = \max(X, Y)$ and $F_X(.)$ denotes the distribution function of X. If, in addition, the random variables X and Y are identically distributed, then

$$F_Z(z) = [F_X(z)]^2.$$

If the random variable X has a probability density function f_X, the then the random variable Z will also have a probability density function f_Z, and it is easy to see that

$$f_Z(z) = 2f_X(z)F_X(z), z \in R.$$

Suppose that the random variables X and Y arte max-sub-independent with common distribution $N(0, \sigma^2)$, thenthe random variabl;e $Z = \max(X, Y)$ has a probability density and it is given by

$$f_Z(z) = \frac{2}{\sigma}\phi(\frac{z}{\sigma})\Phi(\frac{z}{\sigma}), z \in R$$

where $\phi(.)$ and $\Phi(.)$ denote the standard normal density function and standard normal distribution function respectively. Hence the random variable Z has a skew-normal disatribution with parameter $\lambda = 1$.

Acknowledgments

This work was supported under the scheme "Ramanujan Chair Professor" by grants from the Ministry of Statistics and Programme Implementation, Government of India (M 12012/15(170)/2008-SSD dated 8/9/09), the Government of Andhra Pradesh, India (6292/Plg.XVIII dated 17/1/08) and the Department of Science and Technology, Government of India (SR/S4/MS:516/07 dated 21/4/08) at the CR Rao Advanced Institute of Mathematics, Statistics and Computer Science, Hyderabad, India.

References

[1] **Bazargan, H., Bahai, H., and Aminzadeh-Gohari, A.** (2007). *Calculating the return value using a mathematical model of significant wave height*, *J. Marine Sci. Technol.*, **12**, 34-42.

[2] **Durairajan, T.M.** (1979). *A classroom note on sub-independence*, *Gujarat Statist. Rev.*, **6**, 17-18.

[3] **Hamedani, G.G.** (2013). *Sub-Independence: An expository perspective*, *Commun. Statist. Theory-Methods*, **41**, 3615-3638.

[4] **Prakasa Rao, B.L.S.** (2009). *Conditional independence, conditional mixing and conditional association*, *Ann. Inst. Statist. Math.*, **61**, 441-460.

[5] **Prakasa Rao, B.L.S.** (2013). *Characterization of distributions based on functions of conditionally independent random variables*, *J. Indian Statist. Assoc.*, **51**, 183-200.

G.G. Hamedani
Department of Mathematics, Statistics and Computer Science
Marquette University, Milwaukee, WI 53201-1881, USA
E-mail: g.hamedani@mu.edu

B.L.S. Prakasa Rao
C R Rao Advanced Institute of Mathematics, Statistics and
Computer Science, Hyderabad 500046, India
E-mail: blsprao@crraoaimscs.res.in