Statistical Methodology 16 (2014) 1–13

Contents lists available at SciVerse ScienceDirect

Statistical Methodology

journal homepage: www.elsevier.com/locate/stamet

Partial or complete characterization of a bivariate distribution based on one conditional distribution and partial specification of the mode function of the other conditional distribution

Indranil Ghosh [a,*], N. Balakrishnan [b,c]

[a] *Austin Peay State University, Clarksville, TN, United States*

[b] *McMaster University, Hamilton, Ontario, Canada*

[c] *King Abdulaziz University, Jeddah, Saudi Arabia*

ARTICLE INFO

Article history:
Received 19 March 2012
Received in revised form
2 April 2013
Accepted 18 June 2013

Keywords:
Bivariate distribution characterization
Conditional mode
Second-order differential equation

ABSTRACT

There are various ways to characterize a bivariate distribution based on given distributional information. For example, information on both families of conditional densities, i.e., of X given Y and of Y given X, is sufficient to characterize the bivariate distribution. On the other hand, knowledge of both regression functions, i.e., $E(X|Y = y)$ and $E(Y|X = x)$, will be inadequate to determine the joint distribution. In this paper, we discuss to what extent we can characterize (either partially or completely) a bivariate distribution on the basis of complete specification of one family of conditional distributions and partial or complete specification of the mode function of the other family of conditional distributions. This problem is related to an open question mentioned in the paper of Arnold, Castillo and Sarabia (2008) [3].

© 2013 Elsevier B.V. All rights reserved.

1. Introduction

Questions of compatibility of conditional and marginal specifications of distributions are of fundamental importance while modeling data. In an effort to specify bivariate probability models, one

* Corresponding author. Tel.: +1 9512754215.
E-mail addresses: ghoshi@apsu.edu, jamesbond.indranil@gmail.com (I. Ghosh), bala@mcmaster.ca (N. Balakrishnan).

1572-3127/$ – see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.stamet.2013.06.002

is constrained by an inability to visualize the implications of assuming that a given bivariate family of densities will contain a member which will adequately describe the given phenomenon. One of the main difficulties encountered while using probabilistic models to solve real-life problems is the selection of an appropriate model to reflect the reality being observed. One possibility is to use the well-known parametric families of distributions to approximately fit the observed data. However, those so-called "well known" families are too simple in the sense that they depend only on a limited number of parameters, and may not be adequate to model the observed data. In such a situation, one might consider the idea of conditional specification. Specification of joint distributions by means of conditional densities has received considerable attention over the years by authors such as Dawid [9,10], and Gelman and Speed [11,12]. Arnold et al. [2] have discussed this problem for a wide range of families of distributions including exponential families. These models can be useful in situations such as model building in classical statistical settings and in the elicitation and construction of multi-parameter prior distributions in a Bayesian framework.

One of the problems in defining joint densities by specifying their conditionals is the compatibility problem. For example, one possible approach to the specification of the distribution of two-dimensional random variable (X, Y) involves presenting both families of conditional distributions (X given Y and of Y given X). However, the consistency of both the conditional distributions must be checked (see Arnold and Press [6]) to determine if any joint distribution exists with them as its conditional distributions. Several alternative approaches exist in the literature with regard to the problem of determining the possible compatibility of two families of conditional distributions; see, for example, Arnold and Press [6] and Arnold and Gokhale [4]. Also, the problem of determination of most nearly compatible (ε-compatible, as introduced by Arnold and Gokhale [4,5]) has been considered as well.

Here, we consider an alternative approach for identifying to what extent we can characterize either partially or completely the joint probability distribution of any two random variables (X, Y) by specifying one conditional distribution and partial or complete specification of the mode function of the other conditional distribution. Not much work has been done on the characterization of bivariate distributions based on the specification of complete knowledge of one family of conditional distribution and knowledge of the median or mode or percentile function of the other conditional distribution. Recently, Arnold, Castillo and Sarabia [3] provided an interesting discussion in this regard. Motivated by their work, we discuss here the problem of either partial or complete characterization of a bivariate distribution based on the complete specification of $P(X \leq x|Y \leq y)$ and partial or complete specification of the conditional mode function of Y given $X \leq x$. Our discussion here is under the assumption that the distributions of X and Y are both of continuous type. In this paper, we specifically examine the extent to which either a partial or complete specification of the conditional mode function (of the distribution of Y, given $X \leq x$), together with the knowledge of $P(X \leq x|Y \leq y)$, can determine either partially or completely the joint distribution of (X, Y).

The rest of this paper is organized as follows. In Section 2, we discuss the idea of conditional specification of a bivariate distribution under this set up. In Section 3, we present several examples to illustrate the approach of characterizing either partially or completely a bivariate distribution by complete specification of one conditional distribution and either partial or complete specification of the mode function of the other conditional distribution. In Section 4, we make some final comments on this approach and also point out some problems for further study.

2. Characterization of bivariate distribution

Let (X, Y) be a two-dimensional random variable, with joint distribution function

$$F_{X,Y}(x, y) = P(X \leq x, Y \leq y), \quad (x, y) \in \mathbb{R}.$$

Naturally, for $F_{X,Y}(x, y)$ to be a legitimate distribution, it must be monotone in both x and y, must satisfy the conditions $F(-\infty, y) = 0$, $F(x, -\infty) = 0$ and $F(\infty, \infty) = 1$, and must assign nonnegative mass to every rectangle in $\mathbb{R}^2$. At a basic level, the distribution of (X, Y) will be determined by identifying the probability space on which X and Y are defined, say $(\Omega, \mathscr{F}, P)$, and explicitly defining the mappings $X : \Omega \to \mathbb{R}$ and $Y : \Omega \to \mathbb{R}$. But, this may not be feasible often.

Download English Version:

https://daneshyari.com/en/article/7547700

Download Persian Version:

https://daneshyari.com/article/7547700

Daneshyari.com