STATISTICAL INFERENCE FOR A BIVARIATE EXPONENTIAL DISTRIBUTION BASED ON GROUPED DATA*

YE CINAN

Abstract. Consider the bivariate exponential distribution due to Marshall and Olkin [2], whose survival function is $F(x, y) = \exp[-\lambda_1 x - \lambda_2 y - \lambda_{12} \max(x, y)]$ ($x \geq 0, y \geq 0$) with unknown parameters $\lambda_1 > 0$, $\lambda_2 > 0$ and $\lambda_{12} \geq 0$. Based on grouped data, a new estimator for $\lambda_1$, $\lambda_2$ and $\lambda_{12}$ is derived and its asymptotic properties are discussed. Besides, some test procedures of equal marginals and independence are given. A simulation result is given, too.

1. Introduction

Many bivariate extensions of the exponential distribution have been proposed by several authors. Marshall and Olkin [2] proposed a bivariate exponential model $(X, Y)$ whose survival function is

$$F(x, y) = P\{X \geq x, Y \geq y\} = \exp[-\lambda_1 x - \lambda_2 y - \lambda_{12} \max(x, y)], \ (x \geq 0, y \geq 0) \tag{1.1}$$

with unknown parameters $\lambda_1 > 0, \lambda_2 > 0$ and $\lambda_{12} \geq 0$. We refer to it as MOBVE $(\lambda_1, \lambda_2, \lambda_{12})$. This distribution can be derived from “fatal shock” model or “non-fatal shock” model [2]. It is clear that $X$ and $Y$ are independent if and only if $\lambda_{12} = 0$. It can be proved that this bivariate distribution is the unique one with exponential marginal distribution and memoryless property. Unfortunately, the joint survival function (1.1) has a singular part which poses a major hurdle to deriving any inference procedure for the parameters $\lambda_1, \lambda_2$ and $\lambda_{12}$. The purpose of this paper is to present a new estimator for $\lambda_1, \lambda_2$ and $\lambda_{12}$ based on grouped data. The advantage of the new estimator is to have simple expression to calculate easily and nice asymptotic properties.

In Section 2, the new estimator is derived. In Section 3, asymptotic properties of the estimator are discussed, and some test procedures of equal marginals and independence are given. In Section 4, a simulation result is given.

*Supported by the Jiangsu Provincial Natural Science Foundation.
2. New Estimator for \((\lambda_1, \lambda_2, \lambda_{12})\)

Let \((x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)\) be a sample of size \(n\) drawn from two-dimensional population \((X, Y)\) with MOBVE \((\lambda_1, \lambda_2, \lambda_{12})\) distribution. The support \([0, \infty) \times [0, \infty)\) of this distribution is partitioned into \((k + 1)^2\) \((k \geq 2)\) rectangles \(\{(x, y) : a_{i-1} < x < a_i, b_{j-1} < y < b_j\}\) \((i, j = 1, \ldots, k + 1, a_0 = b_0 = 0, a_{k+1} = b_{k+1} = \infty)\) which are denoted by \([a_{i-1}, a_i; b_{j-1}, b_j]\). For convenience assume that \(a_i = b_i = ic\) for every \(i = 0, 1, \ldots, k\), where \(c\) is a fixed positive constant. Then we can count the number of observations, \(n_{ij}\), falling in the rectangle \([a_{i-1}, a_i; b_{j-1}, b_j]\), \(i, j = 1, \ldots, k + 1\), and \(n = \sum_{i=1}^{k+1} \sum_{j=1}^{k+1} n_{ij}\).

Defining that \(p_{ij} \triangleq P\{a_{i-1} < X < a_i, b_{j-1} < Y < b_j\}\), for \(i, j = 1, \ldots, k + 1\), we get

\[
p_{ij} = F(a_{i-1}, b_{j-1}) - F(a_i, b_{j-1}) - F(a_{i-1}, b_j) + F(a_i, b_j)
= \exp[-\lambda_1 a_{i-1} - \lambda_2 b_{j-1} - \lambda_{12} \max(a_{i-1}, b_{j-1})] \\
- \exp[-\lambda_1 a_i - \lambda_2 b_{j-1} - \lambda_{12} \max(a_i, b_{j-1})] \\
- \exp[-\lambda_1 a_{i-1} - \lambda_2 b_j - \lambda_{12} \max(a_{i-1}, b_j)] \\
+ \exp[-\lambda_1 a_i - \lambda_2 b_j - \lambda_{12} \max(a_i, b_j)]
\]

\((i, j = 1, \ldots, k)\). Furthermore, if \(i > j\),

\[
p_{ij} = e^{-(\lambda_1 i + \lambda_2 j + \lambda_{12} i)c} (e^{\lambda_2 c} - 1)[e^{(\lambda_1 + \lambda_{12})c} - 1].
\]

It follows that, for \(i > j\),

\[
\ln p_{ij} = -(\lambda_1 i + \lambda_2 j + \lambda_{12} i)c + \ln(e^{\lambda_2 c} - 1) + \ln[e^{(\lambda_1 + \lambda_{12})c} - 1]
\]

and

\[
\ln p_{i,j-1} = -[\lambda_1 i + \lambda_2 (j - 1) + \lambda_{12} i]c + \ln((e^{\lambda_2 c} - 1)(e^{(\lambda_1 + \lambda_{12})c} - 1)),
\]

which implies that for \(i = 3, \ldots, k\) and \(j = 2, \ldots, i - 1\),

\[
\ln p_{i,j-1} - \ln p_{ij} = \lambda_2 c. \tag{2.1}
\]

For \(i = k + 1\) and \(j = 1, 2, \ldots, k\)

\[
p_{k+1,j} = P\{a_k < X, b_{j-1} \leq Y\} - P\{a_k < X, b_j < Y\}
= e^{-(\lambda_1 k + \lambda_2 j + \lambda_{12} k)c} (e^{\lambda_2 c} - 1),
\]

from which it follows that

\[
\ln p_{k+1,j-1} - \ln p_{k+1,j} = \lambda_2 c. \tag{2.2}
\]