CONTRIBUTIONS TO SAMPLE SPACINGS THEORY, I: LIMIT DISTRIBUTIONS OF SUMS OF RATIOS OF SPACINGS

By Saul Blumenthal[1]

Rutgers—The State University

1. Introduction. In this paper, we study the limiting distribution properties and stochastic convergence of certain statistics based on ratios of sample spacings from different populations. The interest in these statistics stems from their connection with the "parametric" two sample hypothesis. This application is discusseb in detail in the companion paper, Blumenthal (1966). The statistics themselves are described below.

Let $X_1, \cdots, X_m$ and $Y_1, \cdots, Y_n$ be a set of $(n + m)$ independent random variables, the first m having common cdf $F(x)$ and the second n having common cdf $G(x)$. Denote the two sets of ordered observations by $X_1' \leqq \cdots \leqq X_m'$ and $Y_1' \leqq \cdots \leqq Y_n'$. The sample spacings, or sample successive differences, from the two sets of random variables are given as

$$(1.1) \qquad \begin{aligned} DX_i &= X'_{i+1} - X_i', & i &= 1, \cdots, m-1, \\ DY_j &= Y'_{j+1} - Y_j', & j &= 1, \cdots, n-1. \end{aligned}$$

If $m = n$, we can define the statistic

$$(1.2) \qquad S_n(r) = \sum_{i=1}^{n-1} (DX_i/DY_i)^r, \qquad 0 < |r| \leqq 1.$$

If $m > n$, a subset $X_{i_1}, \cdots, X_{i_n}$ of $X_1, \cdots, X_m$ can be chosen at random and then $S_n(r)$ can be defined as above. We shall assume $m = n$ whenever we discuss $S_n(r)$.

Under certain assumptions about the behavior of $F(x)$ and $G(x)$ in the tails, limiting distributions are found for $S_n(r)$ in Section 3. It is found that as $|r|$ varies from 0 to 1, the limiting distribution varies over the class of stable distributions with parameter α going from 1 to 2 (see Section 3). Stochastic convergence of $S_n(r)$ to a limit is taken up in Section 4.

In the following section our notation and assumptions are detailed.

2. Preliminary remarks. In this section we introduce the basic tools which are used in Sections 3 and 4. The results of those sections depend on being able to express the set of random variables DX_i/DY_i, $i = 1, \cdots, n - 1$, in terms of a set of independent random variables. To establish this equivalence, the ratio (DX_i/DY_i) will be expanded in a Taylor series as a function of the hazard rate, and certain well known properties of the hazard rate will be used to make the connection with the set of independent random variables. The relation of gen-

Received 21 December 1964; revised 20 December 1965.

[1] Some of this work was supported by the National Science Foundation under Contract No. NSF-G-19126 while the author was at the University of Minnesota.