# 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

$$DX_i = X'_{i+1} - X_i', \qquad i = 1, \cdots, m - 1, \tag{1.1}$$
$$DY_j = Y'_{j+1} - Y_j', \qquad j = 1, \cdots, n - 1.$$

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

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

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 $\alpha$ 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.

eral spacings to exponential variables by means of the hazard rate was used by the author (1963), and by Proschan and Pyke (1962), (1965). For uniform spacings, LeCam (1958) also used the connection with exponential variables to obtain limit theorems.

We now define the hazard rate function and give some of its properties. For a more comprehensive discussion see Barlow, Marshall and Proschan (1963). Let $F(x)$ be a distribution function with density $f(x)$. The hazard rate $h(x)$ is defined

$$h(x) = f(x)/(1 - F(x)). \tag{2.1}$$

The cumulative hazard rate $H(x)$ is

$$H(x) = \int_{-\infty}^{x} h(t)\, dt = -\log (1 - F(x)). \tag{2.2}$$

It is easily seen that the random variable $H(X)$ (where $X$ has cdf $F(x)$) has the standard exponential distribution,

$$P(H(X) \leqq x) = 1 - e^{-x}. \tag{2.3}$$

Since $H(x)$ is an increasing function of $x$, if $X_1' \leqq \cdots \leqq X_m'$ represent the ordered values of $m$ independent random variables with common distribution $F(x)$, then $H(X_1') \leqq \cdots \leqq H(X_m')$ represent the ordered values of $m$ independent exponentially distributed random variables. Let

$$U_i = (m - i)(H(X_{i+1}') - H(X_i')), \qquad i = 0, 1, \cdots, m - 1, \tag{2.4}$$

(with $H(X_0') = 0$). The fundamental result for this paper is that $U_0, \cdots, U_{m-1}$ are a set of $m$ independent exponentially distributed random variables with the distribution (2.3), and that we can write

$$H(X_i') = \textstyle\sum_{j=0}^{i-1} (U_j/(m - j)), \qquad i = 1, \cdots, m. \tag{2.5}$$

The distribution property of $U_0, \cdots, U_{m-1}$ was used by Rényi (1953) in the study of order statistics, and discovered independently by Epstein and Sobel (1954) in the study of life testing plans, but appears to predate both, being found in the work of Sukhatme (1937).

The spacing $DX_i$ is related to the hazard rate through the expansions

$$DX_i = (H(X_{i+1}') - H(X_i'))/h(\bar{X}_i) \tag{2.6}$$

and

$$\begin{aligned} h^r(\bar{X}_i) &= h^r(i) + r[H(\bar{X}_i) - H(i)](h'(i)/h^{2-r}(i)) \\ &\quad + r[H(\bar{X}_i) - H(i)]^2\{[h(\tilde{X}_i)h''(\tilde{X}_i) - (2 - r)(h'(\tilde{X}_i))^2]/h^{4-r}(\tilde{X}_i)\} \end{aligned} \tag{2.7}$$

for any $r$, if the indicated derivatives exist and are continuous, where we have

$$X_i' \leqq \bar{X}_i \leqq X_{i+1}', \qquad i = 1, \cdots, m - 1,$$

and

$$\min (F^{-1}(i/m), \bar{X}_i) \leqq \tilde{X}_i \leqq \max (F^{-1}(i/m), \bar{X}_i), \qquad i = 1, \cdots, m - 1.$$

The notation used is the following: $h(i)$ is an abbreviation for $h(x)$ evaluated at $x = F^{-1}(i/m)$. Similarly for $h'(i)$, $h''(i)$, and $H(i)$. We use the prime notation for derivatives: $h'(x) = (dh(x)/dx)$, etc. $F^{-1}(x)$ is the usual inverse of the distribution function which we assume to be unique whenever used. In fact we assume that $f(x) > 0$ on an open interval $(F^{-1}(u), F^{-1}(v))$, $0 \leqq u < v \leqq 1$ where $u$, $v$ will be apparent from the context of discussion and are in most places taken as 0 and 1.

We shall use corresponding notation for the $Y$'s and $G(x)$. The hazard rate is $q(x)$, the cumulative hazard rate, $Q(x)$, derivatives will be primed, and $q(i)$ will be $q(x)$ evaluated at $x = G^{-1}(i/n)$, etc. The assumptions regarding $f(x)$ apply also to $g(x)$ with the same values of $(u, v)$.

Further, define

$$V_i = (n - i)(Q(Y'_{i+1}) - Q(Y_i')), \qquad i = 0, 1, \cdots, n, \tag{2.8}$$

(with $Q(Y_0') = 0$). As before

$$Q(Y_i') = \sum_{j=0}^{i-1} (V_j/(n - j)), \qquad i = 1, \cdots, n, \tag{2.9}$$

where $V_1, \cdots, V_n$ are independent, exponentially distributed random variables which are also independent of $U_1, \cdots, U_m$. The corresponding versions of (2.6) and (2.7) will not be repeated.

In Theorem 3.0 we shall investigate conditions under which the first order term in (2.7) can be omitted or replaced by a simpler term while in Lemma 3.1 we consider the conditions under which the second order term in (2.7) can be ignored.

In studying convergence and distribution properties under the assumption that $F(x) = G((x - \mu)/\sigma)$, a simplification can be obtained by noting that

$$DX_i = \sigma DX_{1i}, \qquad i = 1, \cdots, m - 1, \tag{2.10}$$

where the random variables $X_{1i}, \cdots, X_{1m}$ are independent with common distribution $G(x)$. Thus any theorem which is true when $F(x) = G(x)$ holds also when $F(x) = G((x - \mu)/\sigma)$ with the obvious substitution of $(DX_i/\sigma)$ for $DX_i$.

For later reference we state two simple facts

$$P((U_i/V_i)^r < x) = x^{(1/r)}/(1 + x^{(1/r)}), \qquad r > 0, \tag{2.11}$$

and

$$E[(U_i/V_i)^r] = \int_0^\infty (x^r/(1 + x)^2)\, dx = \pi r/\sin \pi r, \qquad 0 < r < 1. \tag{2.12}$$

**3. Limiting distributions for ratios.** In this section we shall obtain the limiting distributions for the sums $S_n(r) = \sum (DX_i/DY_i)^r$. (See (1.2.)) We take the index of summation to run from 1 through $n - 1$ throughout this and the following sections unless stated otherwise. We assume $m = n$. The distributions are obtained both for $F(x) = G(x)$ and for $F(x) \neq G(x)$. The results are based on Theorem 3.1 which establishes the connection between these sums and sums of the independent random variables discussed in Section 2. It is then possible to use standard limit theorems found in Gnedenko and Kolmogorov (1954) to establish the forms of the limiting distributions of the sums of independent random variables. Except in the case where the limiting distribution is normal, our results give the value of the characteristic function of the limiting distribution. Of course this identifies the distribution uniquely but without an expression for the distribution function itself, it is not possible to compute percentage points, etc.

Two types of distributions arise below. In Theorems 3.1 and 3.3, we obtain the so-called stable distributions. The characteristic function $\varphi(t)$ of these distributions can be represented by (see Gnedenko and Kolmogorov, p. 164)

$$\log \varphi(t) = i\eta t - c|t|^{\alpha}\{1 + i\beta(t/|t|)\omega(t, \alpha)\}, \tag{3.1}$$

where $\alpha$, $\beta$, $\eta$, $c$ are constants ($\eta$ is any real number, $-1 \leqq \beta \leqq 1$, $0 \leqq \alpha \leqq 2$, $c \geqq 0$) and

$$\begin{aligned} \omega(t, \alpha) &= \tan \pi\alpha/2 \qquad \text{if } \alpha \neq 1 \\ &= (2/\pi) \log |t| \qquad \text{if } \alpha = 1. \end{aligned}$$

In Theorems 3.1 and 3.3, in each case we have $1 \leqq \alpha \leqq 2$. For $\alpha = 2$, (3.1) represents the normal distribution and we shall arrange the result so that the limiting distribution is the standard normal (mean 0, variance unity). We denote this distribution by $\Phi(x)$. For $1 \leqq \alpha < 2$, we denote the distribution (3.1) by $S(\alpha, \beta, \eta, c)$ listing the values of these constants in the order given.

Each of the distributions in Theorems 3.1 and 3.3 has $|\beta| = 1$ and is skewed toward the right. Such distributions have been studied by Mandelbrot (1960) in connection with economic theory and he has labelled them "Pareto-Lévy" distributions (for $1 < \alpha < 2$). He also mentions some tabulation now in progress which when published will make the present results of more immediate usefulness. For discussions of stable distributions in general the reader is referred to Lukacs (1960) and to Fisz (1962).

In Theorem 3.2, we obtain as limiting distributions infinitely divisible (ID) distributions of the class $L$ which arises from limits of sums of independent random variables. These can be represented in general by

$$\begin{aligned} \log \varphi(t) = i\eta t - (\sigma^2 t^2/2) &+ \textstyle\int_{-\infty}^{0} [e^{iut} - 1 - iut/(1 + u^2)]\, dM(u) \\ &+ \textstyle\int_{0}^{\infty} [e^{iut} - 1 - iut/(1 + u^2)]\, dN(u) \end{aligned} \tag{3.2}$$

where $\eta$ is any constant, $\sigma^2 \geqq 0$, and $M(u)$, $N(u)$ satisfy certain regularity conditions (see Gnedenko and Kolmogorov, p. 84). We shall have $\sigma^2 = 0$ in each case in Theorem 3.2 and $M(u) = 0$ (all $u$). Also we shall have $(-N(u)) = u^{-a}$ in each case. Thus we shall indicate the distributions (3.2) simply as $L(\eta, a)$ in the theorem.

Preliminary to Theorem 3.0, we need some restrictions on how far a uniform order statistic can deviate from its stochastic limit.

LEMMA 3.0. *Let $X_1, \cdots, X_n$ represent a random sample from the uniform dis-*

*tribution on* [0, 1]. *Let the ordered values be* $X_1' \leqq \cdots \leqq X_n'$. *Then the following statements will hold*:

$$\text{(3.3a)} \quad \lim_{n\to\infty} P\{(i/n) - X'_{i-1} < n^{-\frac{1}{2}}, X_i' - (i/n) < n^{-\frac{1}{2}}; \quad i = 1, \cdots, n\} = 1,$$

$$\text{(3.3b)} \quad \lim_{n\to\infty} P\{|X'_{i_j} - (i_j/n)| < (\log n)/n; j = 1, \cdots, kn^{\frac{1}{2}}, \quad k \text{ fixed and the indices } i_1, \cdots, i_{kn^{\frac{1}{2}}} \text{ fixed}\} = 1,$$

$$\text{(3.3c)} \quad \lim_{n\to\infty} P\{X_n' < 1 - (n \log n)^{-1}, X_1' > (n \log n)^{-1}\} = 1.$$

PROOF. Equation (3.3a) is a restatement of a theorem of Kolmogorov (1933), while (3.3b) follows easily from the fact that for any collection of events $A_i$, $i = 1, \cdots, m$, (with complements $A_i^c$) $P(\bigcap_{i=1}^m A_i) \geqq 1 - \sum_{i=1}^m P(A_i^c)$. Here, each $A_i^c$ has probability $[1 - ((\log n)/n)]^n$. Finally, (3.3c) is derived from

$$P\{(n \log n)^{-1} < X_1, \cdots, X_n < 1 - (n \log n)^{-1}\} = [1 - 2(n \log n)^{-1}]^n \to e^{-(2/\log n)} \to 1.$$

This concludes the proof.

In the following theorem, we write $X_n \to {}_P Y_n$ to mean that the sequence of random variables $X_n - Y_n$ approaches zero stochastically. In later theorems, it will be convenient to write $\mathfrak{L}\{X_n\} \to F(x)$ to mean that $F(x)$ is the limiting distribution for the sequence $\{X_n\}$ of random variables.

THEOREM 3.0. *Let* $F(x)$ *and* $G(x)$ *be distribution functions whose hazard rates satisfy the conditions stated below. Let the* $DX_i$, $DY_i$, $U_i$, $V_i$, $h(i)$, $q(i)$, etc., *be as defined in Section* 2. *As* $n$ *increases*,

$$\text{(3.4a)} \quad (1/n^r)\sum (DX_i/DY_i)^r \to_P (1/n^r)\sum (U_i/V_i)^r(q(i)/h(i))^r, \quad \tfrac{1}{2} < r \leqq 1,$$

$$\text{(3.4b)} \quad (1/n \log n)^{\frac{1}{2}}\sum (DX_i/DY_i)^{\frac{1}{2}} \to_P (1/n \log n)^{\frac{1}{2}}\sum (U_i/V_i)^{\frac{1}{2}}(q(i)/h(i))^{\frac{1}{2}} \quad r = \tfrac{1}{2},$$

$$\text{(3.4c)} \quad n^{-\frac{1}{2}}\sum (DX_i/DY_i)^r \to_P n^{-\frac{1}{2}}\sum \{(U_i/V_i)^r(q(i)/h(i))^r + (\pi r^2/(n + 1 - i) \sin \pi r)[(V_i - 1)\textstyle\sum_{j=i+1}^{n-1} (h^{-r}(j)q'(j)/q^{2-r}(j)) - (U_i - 1)\sum_{j=i+1}^{n-1} (q^r(j)h'(j)/h^{2+r}(j))]\}, \quad 0 < r < \tfrac{1}{2}.$$

Equations (3.4) are valid also with $DX_i$ and $DY_i$ interchanged, provided that $U_i$ and $V_i$, and $q(x)$ and $h(x)$ are correspondingly interchanged.

Let

$$K(x; r) = h'(F^{-1}(x))q^r(G^{-1}(x))/h^{2+r}(F^{-1}(x)), \qquad 0 \leqq x \leqq 1.$$

Then we require that

$$(1) \quad \textstyle\int_0^1 \{(1 - x)^{-1}\int_x^1 K^2(y; r)dy\}\, dx < \infty$$

(3.5a) and

(2) there exists a function $p(\delta)$ such that

$$\lim_{\delta\to 0}\,[a(\delta, r)\int_{1-p(\delta)}^{1} \{(1-x)^{-1}\int_x^1 K^2(y; r)\,dy\}\,dx] = 0,$$

$$\lim_{\delta\to 0} p(\delta) = 0,$$

$$\lim_{\delta\to 0} (\delta \log \delta/p(\delta)) = 0,$$

where $a(\delta, r) = \max\{1, (-\log \delta)^{2r-1}\}$. Furthermore, let

$$\alpha(x; \delta, r) = \sup_{R(x,y,\delta)} [h(F^{-1}(y))h''(F^{-1}(y)) - (2+r)(h'(F^{-1}(y)))^2/h^{4+r}(F^{-1}(y))]$$

where

$$R(x, y, \delta) = \{(x, y): |y - x| < \delta(x),\, y < 1 - (\delta^2(-\log \delta)^{-1}),\, y > (\delta^2(-\log \delta)^{-1})\}$$

and

$$\begin{aligned} \delta(x) &= \delta, & \delta < x < 1 - \delta,\\ &= \delta^2, & 0 \leqq x \leqq \delta,\, 1 - \delta \leqq x \leqq 1. \end{aligned}$$

Then we require that for some function $v(\delta)$ with $\lim_{\delta\to 0} v(\delta) = 0$ that

(3.5b)

$$\text{(i)}\quad \lim_{\delta\to 0} \sup_{0\leqq x\leqq 1-v(\delta)} [a(\delta, r)\delta^4(-\log \delta)(1-x)^{-2} \int_x^{1-\delta^2} [q^r(y)\alpha(y; \delta, r)]^2\,dy] = 0;$$

$$\text{(ii)}\quad \lim_{\delta\to 0} [a(\delta, r)\int_{1-v(\delta)}^{1-\delta^2} [q^r(y)\alpha(y; \delta, r)]^2\,dy] = 0$$

and that

(3.5c) $$\lim_{\delta\to 0} [b(\delta, r)\int_{\delta^2}^{1-\delta^2} \{q^r(x)\alpha(x; \delta, r)(1-x)^{-1}\}\,dx] = 0,$$

where

$$\begin{aligned} b(\delta, r) &= \delta, & 0 < r < \tfrac{1}{2},\\ &= \delta(-\log \delta)^{-\frac{1}{2}}, & r = \tfrac{1}{2},\\ &= \delta^{2r}, & \tfrac{1}{2} < r < 1,\\ &= \delta^2(-\log \delta)^{\frac{1}{2}}, & r = 1. \end{aligned}$$

All of conditions (3.5) must be satisfied also with $r$ replaced by $(-r)$ and with $q(x)$ and $h(x)$ interchanged, making four sets of conditions in total.

Note: When $0 < r \leqq \frac{1}{2}$, condition (3.5a-1) implies (3.5a-2) as is easily seen. The latter condition is of concern only when $\frac{1}{2} < r \leqq 1$. It can also be shown that when $r = 1$, (3.5b) implies (3.5c). Further remarks on the conditions appear following Lemma 3.1.

PROOF. Using (2.6) and (2.7) and their counterparts for the $DY_i$, we have

$$
\begin{aligned}
\sum (DX_i/DY_i)^r &= \sum (U_i/V_i)^r(q(i)/h(i))^r \\
&- r\sum (U_i/V_i)^r(H(\bar{X}_i) - H(i))(q^r(i)h'(i)/h^{2+r}(i)) \\
&+ r\sum (U_i/V_i)^r(Q(\bar{Y}_i) - Q(i))(h^{-r}(i)q'(i)/q^{2-r}(i)) \\
&-r\sum (U_i/V_i)^r(H(\bar{X}_i) - H(i))^2q^r(i)[h(\tilde{X}_i)h''(\tilde{X}_i) \\
&- (2+r)(h'(\tilde{X}_i))^2/h^{4+r}(\tilde{X}_i)] + r\sum (U_i/V_i)^r(Q(\bar{Y}_i) \\
&- Q(i))^2h^{-r}(i)[q(\tilde{Y}_i)q''(\tilde{Y}_i) - (2-r)(q'(\tilde{Y}_i))^2/q^{4-r}(\tilde{Y}_i)] \\
&- r^2\sum (U_i/V_i)^r(Q(\bar{Y}_i) - Q(i))(H(\bar{X}_i) \\
&- H(i))[q'(i)h'(i)/q^{2-r}(i)h^{2+r}(i)] - r^2\sum (U_i/V_i)^r(Q(\bar{Y}_i) \qquad (3.6) \\
&- Q(i))(H(\bar{X}_i) - H(i))^2[q'(i)(h(\tilde{X}_i)h''(\tilde{X}_i) \\
&- (2+r)(h'(\tilde{X}_i))^2)/q^{2-r}(i)h^{4+r}(\tilde{X}_i)] - r^2\sum (U_i/V_i)^r(Q(\bar{Y}_i) \\
&- Q(i))^2(H(\bar{X}_i) - H(i))[h'(i)(q(\tilde{Y}_i)q''(\tilde{Y}_i) \\
&- (2-r)(q'(\tilde{Y}_i))^2)/h^{2+r}(i)q^{4-r}(\tilde{Y}_i)] - r^2\sum (U_i/V_i)^r(Q(\bar{Y}_i) \\
&- Q(i))^2(H(\bar{X}_i) - H(i))^2[(q(\tilde{Y}_i)q''(\tilde{Y}_i) \\
&- (2-r)(q'(\tilde{Y}_i))^2)(h(\tilde{X}_i)h''(\tilde{X}_i) \\
&- (2+r)(h'(\tilde{X}_i))^2)/q^{4-r}(\tilde{Y}_i)h^{4+r}(\tilde{X}_i)]
\end{aligned}
$$

where

$$X_i' \leqq \bar{X}_i \leqq X_{i+1}', \qquad Y_i' \leqq \bar{Y}_i \leqq Y_{i+1}', \qquad i = 1, \cdots, n-1,$$
$$\tilde{X}_i \,\varepsilon\, [\bar{X}_i, F^{-1}(i/n)], \qquad \tilde{Y}_i \,\varepsilon\, [\bar{Y}_i, G^{-1}(i/n)], \qquad i = 1, \cdots, n-1.$$

The first term on the right in (3.6) is of the form used in (3.4) and we must study the remaining terms. The second and third terms are very similar in nature and we shall show that when properly normalized, they converge stochastically to zero for $\frac{1}{2} \leqq r \leqq 1$, and to the appropriate random variable given on the right side of (3.4c) for $0 < r < \frac{1}{2}$. The remaining terms represent "error" terms and will be shown to converge stochastically to zero, under the conditions (3.5). This latter demonstration will be given in Lemma 3.1. Because of the similarity of the second and third terms, only the second will be studied in detail.

Ignoring the coefficient, $-r$, the second term in (3.6) can be rewritten

$$
\begin{aligned}
(3.7) \quad & \sum (U_i/V_i)^r(H(X_i') - H(i))(q^r(i)h'(i)/h^{2+r}(i)) \\
& + \sum (U_i/V_i)^r(H(\bar{X}_i) - H(X_i'))(q^r(i)h'(i)/h^{2+r}(i)).
\end{aligned}
$$

It is the first term in (3.7) which we shall study further. By the definition of $\bar{X}_i$, we have wp 1, $H(\bar{X}_i) - H(X_i') \leqq H(X_{i+1}') - H(X_i') = U_i/(n-i)$.

It will be seen in the course of the proof that this bound is sufficient to guarantee that whenever the leading term in (3.7) converges to zero or to a random variable having finite variance, the secondary term converges stochastically to zero. A further simplification results from using (2.2) to note that

$$(3.8)\quad H(i) = Q(i) = -\log(1 - i/n) = \textstyle\sum_{j=0}^{i-1} 1/(n-j) + \Delta_n/(n-i)$$

where $|\Delta_n| < B$ (independent of $n$). Using (3.8), and the representation (2.5) for $H(X_i')$, the leading term in (3.7) becomes

$$(3.9)\qquad \textstyle\sum (U_i/V_i)^r(q^r(i)h'(i)/h^{2+r}(i))\sum_{j=0}^{i-1}(U_j - 1)/(n-j).$$

The error in using (3.9) will be stochastically of the same magnitude as the secondary term in (3.7) and will thus converge stochastically to zero. Finally, introducing a change in order of summation, we write (3.9) as

$$(3.10)\quad \textstyle\sum_{j=0}^{n-2}[(U_j - 1)/(n-j)]\sum_{i=j+1}^{n-1}(U_i/V_i)^r(h'(i)q^r(i)/h^{2+r}(i)).$$

The study of the variance of (3.10) is complicated for $\frac{1}{2} \leqq r \leqq 1$ by the non-existence of the $2r$th moment of $(U_i/V_i)$. This difficulty is circumvented by the introduction of the "truncated" variables $U_i'$, $V_i'$ defined by:

$$(3.11)\qquad \begin{aligned} U_i' &= U_i && \text{if } (U_i/V_i) \leqq n\log n\\ &= (n\log n)V_i && \text{if } (U_i/V_i) > n\log n,\\ V_i' &= V_i && \text{if } (U_i/V_i) \geqq 1/n\log n\\ &= (n\log n)U_i && \text{if } (U_i/V_i) < 1/n\log n. \end{aligned}$$

It is easily seen that

$$(3.12)\qquad \begin{aligned} E(U_i' - 1) &= -1/(1 + n\log n),\\ E(U_i' - 1)^2 &= (1 + (n\log n)^2)/(1 + n\log n)^2 < 1 \end{aligned}$$

and that

$$(3.13)\qquad \begin{aligned} E(U_i'/V_i') &\leqq k_1 \log(n\log n),\\ E(U_i'/V_i')^p &\leqq k_2 (n\log n)^{p-1} \qquad (p > 1). \end{aligned}$$

It is also easy to verify that

$$(3.14)\qquad \textstyle\sum_{i=1}^{n} P[(U_i'/V_i') \neq (U_i/V_i)] = 2n/(1 + n\log n)$$

which goes to zero as $n$ increases. Similar expressions obtain for $\sum P(U_i' \neq U_i)$ and $\sum P(V_i' \neq V_i)$. Thus it is possible to substitute $U_i'$ for $U_i$ and $V_i'$ for $V_i$ in (3.10) and study the stochastic convergence of the resultant expression which has a finite second moment. Thus writing

$$(3.15)\quad A_j = (1/(n-j))\textstyle\sum_{i=j+1}^{n-1}(U_i'/V_i')^r(h'(i)q^r(i)/h^{2+r}(i)),$$

we have

$$
\begin{aligned}
&E\{\sum A_j(U_j' - 1)\}^2 \\
(3.16)\qquad &= \sum E(U_j' - 1)^2 A_j^2 + 2\textstyle\sum_{k>j} E(U_j' - 1)(U_k' - 1)A_jA_k \\
&= \sum (1 + (c/n \log n)^2)EA_j^2 + 2\textstyle\sum_{k>j} (d/n \log n)E(U_k' - 1)A_jA_k
\end{aligned}
$$

using (3.12) and the fact that $A_j$ involves $i$ subscripts which are strictly greater than $j$. Now we have

$$
\begin{aligned}
E(A_j^2) = {} & (n-j)^{-2}\textstyle\sum_{i=j+1}^{n-1} E(U_i'/V_i')^{2r}(h'(i)q^r(i)/h^{2+r}(i))^2 \\
(3.17)\qquad & + 2(n-j)^{-2}\textstyle\sum_{k>i\geq j+1} E(U_i'/V_i')^r E(U_k'/V_k')^r \\
& \cdot (h'(i)q^r(i)/h^{2+r}(i))(h'(k)q^r(k)/h^{2+r}(k)),
\end{aligned}
$$

which becomes for $r = 1$ (using (3.13)),

$$
\begin{aligned}
E(A_j^2) = {} & [k_2(n \log n)^2/n^2(1 - (j/n))^2] \textstyle\sum_{i=j+1}^{n-1} (h'(i)q(i)/ h^3(i))^2 \\
(3.17a)\qquad & + [2(k_1 \log (n \log n))^2/n^2(1 - (j/n))^2] \\
& \cdot \textstyle\sum_{k>i>j} (h'(i)q(i)/h^3(i))(h'(k)q(k)/ h^3(k))
\end{aligned}
$$

and for $\frac{1}{2} < r < 1$, using (3.13) and (2.12) (neglecting the slight error of applying (2.12) to $(U'/V')$ instead of $(U/V)$), (3.17) is

$$
\begin{aligned}
E(A_j^2) = {} & [k_2(n \log n)^{2r-1}/n^2(1 - (j/n))^2]\textstyle\sum_{i=j+1}^{n-1} (h'(i)q^r(i)/ h^{2+r}(i))^2 \\
(3.17b)\qquad & + [2(\pi r)^2/(\sin \pi r)^2 n^2(1 - (j/n))^2] \\
& \cdot \textstyle\sum_{k>i>j}(h'(i)q^r(i)/h^{2+r}(i))(h'(k)q^r(k)/h^{2+r}(k)).
\end{aligned}
$$

Now introduce the normalizing constant $(1/n)^{2r}$ into (3.16) from (3.4a), use (3.17) and the leading term of (3.16) becomes

$$
\begin{aligned}
& n^{-1}k_2(\log n)^{2r-1}\textstyle\sum_{j=0}^{n-2} (1/n)(1 - (j/n))^{-2} \\
(3.18)\qquad & \cdot \textstyle\sum_{i=j+1}^{n-1}(h'(i)q^r(i)/h^{2+r}(i))^2(1/n) + [2\gamma(n, r)/n^{2r-1}] \\
& \cdot \textstyle\sum_{j=0}^{n-2} (1/n)(1 - (j/n))^{-2} \\
& \cdot \textstyle\sum_{k\neq i>j} [h'(i)h'(k)q^r(i)q^r(k)/h^{2+r}(i)h^{2+r}(k)](1/n)^2
\end{aligned}
$$

with $\gamma(n, r)$ given by

$$
\begin{aligned}
(3.18a)\qquad \gamma(n, r) &= (k_1 \log (n \log n))^2, & r &= 1, \\
&= (\pi r)^2/(\sin \pi r)^2, & \tfrac{1}{2} &< r < 1.
\end{aligned}
$$

For $r = \frac{1}{2}$, (3.17) becomes (using (3.13) and (2.12)) (neglecting the slight error of using (2.12))

$$
\begin{aligned}
& [k_1(\log (n \log n))/n^2(1 - (j/n))^2]\textstyle\sum_{i=j+1}^{n-1} (h'(i)q^{\frac{1}{2}}(i)/h^{2.5}(i))^2(1/n) \\
& \quad + [\pi^2/4n^2(1 - (j/n))^2]\textstyle\sum_{k\neq i>j} (h'(i)q^{\frac{1}{2}}(i)/h^{2.5}(i))(h'(k)q^{\frac{1}{2}}(k)/h^{2.5}(k)).
\end{aligned}
$$

With the normalization $(1/n \log n)$ from (3.4b) we obtain

$$(3.19)\quad \begin{aligned}&[k_1(\log (n \log n))/n \log n]\textstyle\sum_{j=0}^{n-2} (1/n)(1 - (j/n))^{-2}\\ &\cdot \textstyle\sum_{i=j+1}^{n-1}(h'(i)q^{\frac{1}{2}}(i)/h^{2.5}(i))^2(1/n) + (\pi^2/4)(\log n)^{-1}\\ &\cdot \textstyle\sum_{j=0}^{n-2} (1/n)(1 - (j/n))^{-2}\\ &\cdot \textstyle\sum_{k\neq i>j} (h'(i)h'(k)q^r(i)q^r(k)/h^{2+r}(i)h^{2+r}(k))(1/n)^2.\end{aligned}$$

There is still the double sum in (3.16) to consider and we claim that essentially this term can be obtained by using $E(U_k' - 1)EA_jEA_k$ in place of $E(U_k' - 1)A_jA_k$ since the error of so doing can be shown to be of smaller magnitude than the resulting expression. This approximation leads to

$$(3.20)\quad [C/(n^{2r} \log^2 n)][\textstyle\sum (1/n)\{(1 - (j/n))^{-1}\sum_{i=j+1}^{n-1} (h'(i)q^r(i)/h^{2+r}(i))\}]^2$$

for $\frac{1}{2} < r < 1$, where $C$ depends on $r$ through (2.12), and where the normalizing constant changes to $[C/(n \log^3 n)]$ for $r = \frac{1}{2}$ and to $[C \log (n \log n)/n^2 \log^2 n]$ for $r = 1$. In all of these cases, condition (3.5a-1) assures the convergence to zero of this expression, since the sum in the square brackets is the Riemann sum approximation to

$$(3.20\text{a})\quad \textstyle\int_0^1 [1/(1 - x)]\int_x^1 [h'(F^{-1}(y))q^r(G^{-1}(y))/h^{2+r}(F^{-1}(y))]\, dy\, dx$$

and using the fact that $\int_0^1 g^2(x)\, dx < \infty \Rightarrow \int_0^1 g(x)\, dx < \infty$ we see that if

$$(3.20\text{b})\quad \textstyle\int_0^1 \{(1 - x)^{-1}\int_x^1 K(y; r)\, dy\}^2\, dx < \infty$$

then (3.20a) is also finite. By the Schwartz inequality,

$$(1 - x)^{-2}(\textstyle\int_x^1 K(y; r)\, dy)^2 \leqq (1 - x)^{-1}\int_x^1 K^2(y; r)\, dy$$

so that (3.5a-1) does insure the finiteness of (3.20b) and thus of (3.20a).

For $r < \frac{1}{2}$, there is no need for introducing $U'$ and $V'$ variables and we can consider (3.16) without the primes. Doing so, using (2.14), and normalizing with $n^{-\frac{1}{2}}$ gives (omitting the intermediate steps)

$$(3.21)\quad \begin{aligned}&E\{(n^{\frac{1}{2}})^{-1}\textstyle\sum A_j(U_j - 1)\}^2 = (2\pi r/n \sin 2\pi r)\\ &\cdot\textstyle\sum_{j=0}^{n-2} (1/n)(1 - (j/n))^{-2}\sum_{i=j+1}^{n-1} (h'(i)q^r(i)/h^{2+r}(i))^2(1/n)\\ &\quad + (\pi r/\sin \pi r)^2\textstyle\sum_{j=0}^{n-2} (1/n)(1 - (j/n))^2\\ &\qquad \cdot \textstyle\sum_{k\neq i>j} (h'(i)q^r(i)/h^{2+r}(i))(h'(k)q^r(k)/h^{2+r}(k))(1/n)^2.\end{aligned}$$

Note that the double sum of (3.16) is identically zero when $U'$ is replaced by $U$ so that (3.20) does not arise.

The three expressions (3.18), (3.19) and (3.21) are basically similar and can be summarized as

$$(3.22)\quad \begin{aligned}&(Ca_n(r)/n)\textstyle\sum_{j=0}^{n-2} (1/n)(1 - (j/n))^{-2}\sum_{i=j+1}^{n-1} (h'(i)q^r(i)/h^{2+r}(i))^2(1/n)\\ &+ Db_n(r)\textstyle\sum_{j=0}^{n-2} (1/n)\{(1 - (j/n))^{-1}\sum_{i=j+1}^{n-1} (h'(i)q^r(i)/h^{2+r}(i))(1/n)\}^2\end{aligned}$$

where $C$ and $D$ may depend on $r$, $a_n(r) = 1$ for $0 < r \leqq \frac{1}{2}$ and $(\log n)^{2r-1}$ for $(\frac{1}{2}) < r \leqq 1$, $b_n(r) = 1$ for $0 < r < \frac{1}{2}$, and $b_n(r) \to 0$ as $n$ increases for $(\frac{1}{2}) \leqq r \leqq 1$. Thus, if condition (3.20b) holds, the second sum in (3.22) represents a Riemann sum approximation to the integral and will approach the integral. Thus when multiplied by $b_n(r)$, it converges to zero for $(\frac{1}{2}) \leqq r \leqq 1$, and to the given finite integral of (3.20b) for $0 < r < \frac{1}{2}$. As seen before, (3.20b) does hold whenever (3.5a-1) is satisfied.

The first term in (3.22) converges to zero for all $r$. For some distributions, $\int_0^1 \{(1-x)^{-2} \int_x^1 K^2(y; r)\, dy\}\, dx < \infty$ in which case the preceding argument suffices. It is easily verified that this condition implies (3.5a), and in many cases this integral is divergent, and so is the sum in (3.22) but we shall show that condition ((3.5a)-(2)) is designed to limit the rate of divergence. Write $b_i = (1/ni) \sum_{k=n-i+1}^{n-1} (h'(k)q^r(k)/h^{2+r}(k))^2(1/n)$. Then the first term in (3.22) is (using the Abel sum formula)

$$(3.23) \quad Ca_n(r) \sum_{i=2}^{n} (1/i)b_i = Ca_n(r)\{(1/n)B_n - B_1 - \sum_{i=2}^{n} (1/i(i-1))B_{i-1}\},$$

where

$$(3.24) \quad B_i = \sum_{j=1}^{i} b_j = \sum_{k=n-i}^{n-1} (1/n)(1-(k/n))^{-1} \sum_{j=k+1}^{n-1} (h'(j)q^2(j)/h^{2+r}(j))^2(1/n).$$

Now condition (3.5a-1) asserts that $B_n$ approaches the given finite integral, and $B_1 = 0$ by definition.

Also, for $t_n$ such that $(t_n/n) \to 0$ as $n$ increases, we have $B_{t_n} < \epsilon(\epsilon > 0$, given, arbitrary) for all $n$ sufficiently large since it is seen from (3.24) that

$$B_{t_n} < \int_{1-\delta}^{1} [dx/(1-x)] \int_x^1 [h'(F^{-1}(y))q^2(G^{-1}(y))/h^{2+r}(F^{-1}(y))]^2\, dy < \epsilon$$

provided that $\delta$ is sufficiently small, and $(t_n/n) < \delta$. Since all terms are positive, $B_i < B_{t_n}$, $i = 1, \cdots, t_n$. Also, all $B_i < D < \infty$ for some $D$. Thus, we have

$$\sum_{i=2}^{n} (1/i(i-1))B_n < B_{t_n} \sum_{i=2}^{t_n} (1/i(i-1)) + D\sum_{i=t_n}^{n} (1/i(i-1)) < B_{t_n} + D/t_n .$$

Returning to (3.23), we see that if $t_n$ goes to infinity with $n$, then if $a_n(r)$ is bounded, (3.23) converges to zero. This applies to $0 \leqq r \leqq \frac{1}{2}$. Since we can always choose $t_n$ so that $(\log n/t_n) \to 0$, the difficulty for $\frac{1}{2} < r \leqq 1$ arises with $(\log n)^{2r-1}B_{t_n}$. Condition (3.5a-2) asserts just this with $\delta = n^{-\frac{1}{2}}$. This completes the proof for $\frac{1}{2} \leqq r \leqq 1$. For $0 < r < \frac{1}{2}$, it remains to be shown that

$$n^{-\frac{1}{2}} \sum_{j=0}^{n-2} [(U_j - 1)/(n-j)] \cdot \sum_{i=j+1}^{n-1} [(U_i/V_i)^r - \pi r/\sin \pi r](h'(i)q^r(i)/h^{2+r}(i)) \to_P 0$$

in order to justify writing (3.4c). This is accomplished by straightforward computation of the second moment and will be omitted.

LEMMA 3.1. *Under conditions* (3.5), *the "error terms" of Equation* (3.6) *when appropriately normalized, converge stochastically to zero.*

PROOF. The "leading" error terms in (3.6) are the fourth and fifth. Since these are similar, only the fourth will be studied in detail. The sixth term is a hybrid of the second and third, and will converge to zero whenever those terms converge to proper random variables. The last three terms represent combinations of the fourth or fifth with the second or third and again the methods used herein show that they converge to zero under the conditions of the theorem. Limiting attention to the fourth term, we shall show the stochastic convergence to zero of

$$(3.25)\quad C_n(r)\sum (U_i/V_i)^r(H(\bar{X}_i) - H(i))^2q^r(i) \cdot[(h(\tilde{X}_i)h''(\tilde{X}_i) - (2 + r)(h'(\tilde{X}_i))^2)/h^{4+r}(\tilde{X}_i)]$$

where

$$\begin{aligned} C_n(r) &= n^{-r}, & \tfrac{1}{2} < r \leqq 1,\\ &= (n\log n)^{-\frac{1}{2}}, & r = \tfrac{1}{2},\\ &= n^{-\frac{1}{2}}, & 0 < r < \tfrac{1}{2}. \end{aligned}$$

Writing,

$$\begin{aligned} H(\bar{X}_i) - H(i) &= H(\bar{X}_i) - H(X_i') + H(X_i') - H(i)\\ (3.26)\qquad &< H(X'_{i+1}) - H(X_i') + H(X_i') - H(i)\\ &= [U_i/(n - i)]\\ &\quad + \textstyle\sum_{j=0}^{i-1} (U_j - 1)/(n - j) + B/(n - i) \end{aligned}$$

as was done in the preceding theorem, it is easily seen that (3.25) converges to zero if and only if

$$(3.27)\quad C_n(r)\sum (U_i/V_i)^r(\textstyle\sum_{j=0}^{i-1} (U_j - 1)/(n - j))^2q^r(i) \cdot[(h(\tilde{X}_i)h''(\tilde{X}_i) - (2 + r)(h'(\tilde{X}_i))^2)/h^{4+r}(\tilde{X}_i)]$$

converges to zero.

Because $F(X)$ is uniformly distributed on (0, 1), Lemma 3.0 applies to the ordered values $F(X_1') \leqq \cdots \leqq F(X_n')$. Thus, since

$$|F(\tilde{X}_i) - i/n| < \max(|F(X_i') - i/n|, |F(X'_{i+1}) - i/n|)$$

by the construction of $\tilde{X}_i$, we can use Lemma 3.0 to conclude that

$$(3.28)\quad \lim_{n\to\infty} P\{|F(\tilde{X}_i) - i/n| < \delta_i,\ i = 1, \cdots, n - 1; \cdot F(\tilde{X}_n) \leqq 1 - (1/n\log n),\ F(\tilde{X}_1) \geqq (1/n\log n)\} = 1$$

where

$$\begin{aligned} \delta_i &= n^{-\frac{1}{2}}, & n^{\frac{1}{2}} + 1 \leqq i \leqq n - n^{\frac{1}{2}} - 1\\ &= (n^{-1})\log n, & i \leqq n^{\frac{1}{2}},\ i \geqq n - n^{\frac{1}{2}}. \end{aligned}$$

Write

$$(3.29)\quad \alpha(i) = \sup_{|x-i/n|<\delta_i} [(h(F^{-1}(x))h''(F^{-1}(x)) - (2+r)[h''(F^{-1}(x))]^2)/h^{4+r}(F^{-1}(x))].$$

Following Proschan and Pyke (1965), we shall show the stochastic convergence of (3.27) given the event described in (3.28), since that is sufficient to imply the "unconditional" convergence of (3.27). But given the event of (3.28), we can replace the random variable $([h(\tilde{X}_i)h''(\tilde{X}_i) - (2+r)(h'(\tilde{X}_i))^2]/h^{4+r}(\tilde{X}_i))$ by the nonrandom quantity $\alpha(i)$ given in (3.29). Thus we shall demonstrate the convergence to zero of

$$(3.30)\qquad C_n(r)\sum (U_i/V_i)^r q^r(i)\alpha(i)\left(\sum_{j=0}^{i-1}(U_j-1)/(n-j)\right)^2.$$

The method will be to investigate the second moment of (3.30). Since $E(U_i/V_i)^r$ does not exist for $r \geqq 1$, we again use the equivalent variables $U_i'$, $V_i'$ introduced in (3.11) when necessary. To save space, we shall write out the moments using $U$'s and $V$'s, with the understanding that primes are intended where needed. The expected square of (3.30) consists of two parts

$$(3.31\text{a})\quad C_n^2(r)\sum E(U_i/V_i)^{2r}E\left(\sum_{j=0}^{i-1}[(U_j-1)/(n-j)]\right)^4(q^r(i)\alpha(i))^2;$$

$$(3.31\text{b})\quad 2C_n^2(r)\sum_{k>i}E(U_k/V_k)^r E(U_i/V_i)^r\left(\sum_{L=0}^{k-1}[(U_L-1)/(n-L)]\right)^2 \cdot\left(\sum_{j=0}^{i-1}[(U_j-1)/(n-j)]\right)^2(q^r(i)q^r(k)\alpha(i)\alpha(k)).$$

Writing

$$\sum_{L=0}^{k-1}[(U_L-1)/(n-L)] = \sum_{L=0}^{i-1}[(U_L-1)/(n-L)] + \sum_{L=i+1}^{k-1}[(U_L-1)/(n-L)] + (U_i-1)/(n-i),$$

we have

$$\begin{aligned}
&\left(\sum_{L=0}^{k-1}[(U_L-1)/(n-L)]\right)^2\left(\sum_{j=0}^{i-1}[(U_j-1)/(n-j)]\right)^2 \\
&\quad = \left(\sum_{j=0}^{i-1}[(U_j-1)/(n-j)]\right)^4 \\
&\qquad + \left(\sum_{j=0}^{i-1}[(U_j-1)/(n-j)]\right)^2\left(\sum_{j=i+1}^{k-1}[(U_j-1)/(n-j)]\right)^2 \\
(3.32)\quad &\qquad + ((U_i-1)/(n-i))^2\left(\sum_{j=0}^{i-1}[(U_j-1)/(n-j)]\right)^2 \\
&\qquad + 2((U_i-1)/(n-i))\left(\sum_{j=0}^{i-1}[(U_j-1)/(n-j)]\right)^3 \\
&\qquad + 2((U_i-1)/(n-i))\left(\sum_{j=0}^{i-1}[(U_j-1)/(n-j)]\right)^2 \cdot\left(\sum_{j=i+1}^{k-1}[(U_j-1)/(n-j)]\right) \\
&\qquad + 2\left(\sum_{j=0}^{i-1}[(U_j-1)/(n-j)]\right)^3\left(\sum_{j=i+1}^{k-1}[(U_j-1)/(n-j)]\right).
\end{aligned}$$

The following are easily verified:

$$E\left(\sum_{j=0}^{i-1} [(U_j - 1)/(n - j)]\right)^4 = C/(n - i + 1)^2,$$

$$(3.33) \quad E\left(\sum_{j=0}^{i-1} [(U_j - 1)/(n - j)]\right)^3 = C'/(n - i + 1)^2,$$

$$E\left(\sum_{j=0}^{i-1} [(U_j - 1)/(n - j)]\right)^2 = 1/(n - i + 1),$$

$$E\left(\sum_{j=i+1}^{k-1} [(U_j - 1)/(n - j)]\right)^2 = 1/(n - k + 1) - 1/(n - i + 1),$$

where $C$, $C'$ are of the form $a + o((n - i + 1)^{-2})$. Further, using the fact that the joint density of $T = (U/V)$ and $U$ is

$$f(t, u) = ut^{-2}e^{-u}e^{-(u/t)}; \qquad u \geqq 0, t \geqq 0,$$

we find that

$$(3.34) \quad \begin{aligned} E(UT^r) &\leqq 2/(1 - r), & r < 1, \\ &\leqq 2ET, & r = 1; \\ E(U^2T^r) &\leqq 3/(1 - r), & r < 1, \\ &\leqq 3ET, & r = 1. \end{aligned}$$

Now, use of (2.14) and (3.13) along with (3.33) and the definition of $C_n(r)$ (3.25), gives for (3.31a)

$$(3.35) \qquad a_n(r) \sum (n - i + 1)^{-2}[(q^r(i)\alpha(i))^2/n],$$

where

$$\begin{aligned} a_n(r) &= A, & 0 < r \leqq \tfrac{1}{2} \\ &= A(\log n)^{2r-1}, & \tfrac{1}{2} < r \leqq 1 \end{aligned}$$

and $A$ is of the form $(A' + o(1))$. The use of (3.32), the independence of the $U$'s and $V$'s and (3.33), (3.34) along with (2.14) and (3.13) yield for (3.31b)

$$(3.36) \quad 2b_n(r) \sum_{k>i} [(C/(n - i + 1)^2)q^r(i)q^r(k)\alpha(i)\alpha(k) + (q^r(i)\alpha(i)/(n - i + 1))(q^r(k)\alpha(k)/(n - k + 1))]$$

where

$$\begin{aligned} b_n(r) &= B/n, & 0 < r < \tfrac{1}{2}, \\ &= B/n \log n, & r = \tfrac{1}{2}, \\ &= B/n^{2r}, & \tfrac{1}{2} < r < 1, \\ &= B \log n/n^2, & r = 1, \end{aligned}$$

and again, $B = (B' + o(1))$, and $C = (C' + o(n - i + 1)^2)$. Strictly speaking, (3.35) and (3.36) are valid only when $U$'s and $V$'s appear in (3.30). The correct expressions for $(U')$'s and $(V')$'s in (3.30) are of the nature of $(1 + O(n^{-1}))$

times the expressions (3.35) and (3.36) (which can be verified using (3.12)). Thus it is sufficient to show that both (3.35) and (3.36) converge to zero as $n$ increases. Further simplification in (3.36) is attained by noting that for $k > i$, $(n - k + 1)^{-1} > (n - i + 1)^{-1}$ so that (3.36) converges to zero if and only if

$$(3.37)\quad 2b_n(r) \textstyle\sum_{k>i} (q^r(i)\alpha(i)/(n - i + 1))(q^r(k)\alpha(k)/(n - k + 1))$$
$$= b_n(r)[(\textstyle\sum q^r(i)\alpha(i)/(n - i + 1))^2 - \textstyle\sum (q^r(i)\alpha(i))^2/(n - i + 1)^2]$$

converges to zero. Since $a_n(r) \geqq nb_n(r)$, the second term on the right in (3.37) goes to zero whenever (3.35) does. Thus it is necessary and sufficient to have (3.35) and

$$(3.38)\qquad d_n(r) \textstyle\sum q^r(i)\alpha(i)/(n - i + 1)$$

where

$$\begin{aligned} d_n(r) &= n^{-\frac{1}{2}}, & 0 < r < \tfrac{1}{2}, \\ &= (n \log n)^{-\frac{1}{2}}, & r = \tfrac{1}{2}, \\ &= n^{-r}, & \tfrac{1}{2} < r < 1, \\ &= (\log n)^{\frac{1}{2}} n^{-1}, & r = 1, \end{aligned}$$

converge to zero. In general, both (3.35) and (3.38) must be checked. However, for $r = 1$, (3.35) alone suffices (by the Schwartz inequality).

Using the Abel sum formula as in Theorem 3.0, we find that (3.35) approaches zero as $n$ increases if there is a sequence of constants $t_n$ depending on $n$, such that

$$(3.39)\qquad \begin{aligned} &\lim_{n\to\infty} \{a_n'(r) \log t_n(n - i + 1)^{-2} B_i\} = 0, \quad \text{uniformly in } i = 1, 2, \cdots, t_n - 1, \\ &\text{and} \\ &\lim_{n\to\infty} \{a_n'(r) B_{t_n}\} = 0, \quad \text{where} \\ &a_n'(r) = \max(1, (\log n)^{2r-1}), \\ &B_i = \textstyle\sum_{j=i}^{n-1} (q^r(j)\alpha(j))^2/n. \end{aligned}$$

Conditions (3.39) can be more conveniently expressed by the use of integrals and noting that

$$(3.40)\qquad B_i < \textstyle\int_{i/n}^{1-1/n} q^{2r}(x; n)\alpha^2(x)\, dx$$

wheie $q(x; n)$ is the step function whose value for $(2i - 1)/2n < x \leqq (2i + 1)/2n$ is $q(i/n)$, and

$$(3.41)\qquad \begin{aligned} \alpha(x) &= \sup_{|y-x|<\delta(x)} [(h(F^{-1}(y))h''(F^{-1}(y)) - (2 + r) \\ &\qquad \cdot [h'(F^{-1}(y))]^2)/h^{4+r}(F^{-1}(y))]; \\ \delta(x) &= 2\delta_i \text{ (eq. (3.28))} \quad \text{for} \quad (2i - 1)/2n < x \leqq (2i + 1)/2n. \end{aligned}$$

Since $q'(x)$ exists and is continuous, if $(i/n) \to p$, then the integral in (3.40) will converge or diverge together with $\int_p^{(1-1/n)} q^{2r}(x)\alpha^2(x)\,dx$ so that conditions (3.5b) are sufficient to insure (3.39), making the identification of $\delta$ with $n^{-\frac{1}{2}}$.

The convergence of (3.38) to zero is easily seen to be equivalent to condition (3.5c).

Remarks on Theorem 3.0 and Lemma 3.1. (1) Results (3.4) bear a similarity to those of LeCam (1958) who found that for functions of uniform spacings, the limiting distribution would be the same as the limiting distribution of the same function of exponential variables (as in (3.4a) and (3.4b)) with the possible exception of having an additional variance term as in (3.4c). Neither result is a special case of the other, but both reflect the same phenomenon.

(2) An important special case of Theorem 3.0 occurs for $F(x) = G(x)$, in which case the conditions (3.5) need only be checked for $+r$ and $-r$ and also take on a simpler form since $q(x)$ is then $h(x)$. If $(h'(x)/h^2(x))$ is bounded, then (3.5a) is easily verified. The function $\alpha(x; \delta, r)$ in (3.5) is bounded by $\sup h^{-r}(y)$ with $y$ suitably restricted. Conditions (3.5b) and (3.5c) can then be seen to hold by examination of the limiting case $(h'(x)/h^2(x)) = 1$, which happens for the uniform distribution. This is a limiting case in the sense that if $h'(x) = 0$, or if $h''(x) = 0$, or either tends to zero, the conditions (3.5) are easily verified, but when $(h'(x)/h^2(x))$ is merely bounded, (3.5b) and (3.5c) are difficult to verify, and with the uniform this quantity always equals its bound. The computations involved in finding $\alpha(x; \delta, r)$ are easy for the uniform, $\alpha(x; \delta, r)$ being in fact $\sup |(1 - y)^r|$, and it is somewhat surprising that for the conditions (3.5b and c) to be satisfied, careful attention must be given to the limits on the integration and the definition of $\delta(x)$. (Note that $v(\delta)$ can be taken to be $\delta$.) The integrals in (3.5b, c) diverge if the upper limits are set equal to unity, and if $\delta(x)$ is taken equal to $\delta$ for all $x$, the conditions fail when $r = 1$. In the recent work of Proschan and Pyke (1965), $\delta(x)$ was just $\delta$ and we were quite surprised in trying to check the conditions (3.5) for the uniform to find that this definition of $\delta(x)$ was not sufficient. This makes us regard the uniform as a difficult distribution to handle for sample spacings ratios which is paradoxical since it is the most naturally associated distribution with ordinary sample spacings.

(3) The difficulty of working with uniform spacings is pointed up not only by the complications mentioned in (2), but by the fact that LeCam also had to relate their behavior to that of exponential variables, and then use very delicate manipulations to achieve his results. Exponential variables lend themselves nicely to work with spacings as seen in Section 2. The expansion in terms of the exponential is crucial to Theorem 3.0. This is most easily seen by examining Blumenthal (1962) in which an attempt was made to expand general spacings in terms of uniform spacings. Such an expansion necessitated great restrictions on the densities $f(x)$, most of which have been avoided by the present technique.

(4) We would also conjecture that if $h'(x)/h^2(x)$ is not bounded and $F(x) = G(x)$, (3.5b and c) will not hold. One example where this quantity is not bounded is given by $F(x) = 1 - (\log x)^{-1}$, $x > e$, for which not even (3.5a) holds. It is easily verified that (3.4) is false for this $F(x)$.

Since boundedness of $(h'(x)/h^2(x))$ when $F(x) = G(x)$ is sufficient for (3.4) and "almost necessary" as the preceding example shows, we might ask what sort of distributions satisfy or fail to satisfy this requirement. In terms of the tail behavior of the distribution $F(x)$, $|h'(x)/h^2(x)| < B$ implies $(1 - F(x))^{B+1} < f(x) < (1 - F(x))^{-B+1}$. It also implies that $xf(x)/(1 - F(x))$ (i.e., $xh(x)$) is bounded away from zero.

Roughly speaking, not many distributions will violate these conditions. This is due to the representation $1 - F(x) = e^{-H(x)}$ (see (2.2)), which implies that $H(x) \to \infty$ as $x$ increases. Now suppose that $xh(x)$ approaches zero for large $x$. Writing $g(x)$ for $xh(x)$, it is readily verified that $H(x) \to \infty$ implies that the series $\sum_{n=1}^{\infty} (g(n)/n)$ must diverge, even though $g(n) \to 0$. Thus, the rate of convergence to zero must be very limited. For instance $g(x)$ can behave as $(1/\log x)$ or some slightly more rapidly decreasing function, but not as $(1/\log x)^{1+\beta}$ for any positive $\beta$. The previous example then is typical of such distributions.

In passing, we note that for a distribution having $xh(x)$ approach zero, no moments of positive order exist. Also, even for as badly behaved a distribution as the Cauchy, the quantity $xh(x)$ approaches unity for large $x$.

(5) In examining the proof of Lemma 3.1, it might be noted that it is not necessary to have a second derivative and a second order term in the Taylor series expansion, but great difficulties in handling the error term arise when only a first order expansion is used. Our attempts indicated that when the error term in a first order expansion was suitably restricted, the conditions imposed almost amounted to the assumption of a second derivative, and looked very unappealing. Also deeper troubles with the $\delta(x)$ terms seem to arise. One place in which a first order expansion might not introduce too many complications would be for $\frac{1}{2} \leqq r \leqq 1$ since the first order term vanishes there. This would mean two expansions for the separate cases $r < \frac{1}{2}$, $r \geqq \frac{1}{2}$, and two separate sets of conditions which hardly seem worthwhile in view of the limited additional generality gained.

The limiting distributions of the quantities on the right sides of (3.4) now can be determined by use of theorems in Gnedenko and Kolmogorov.

THEOREM 3.1. *If $F(x) = G(x)$ and $F(x)$ satisfies the conditions of Theorem* 3.0, *then as $n$ increases,*

$$(3.42a)\quad \mathfrak{L}\{\log n[(1/n \log n) \textstyle\sum (DX_i/DY_i) - 1]\} \to S(1, 1, -1, \pi/2), \qquad r = 1;$$

$$(3.42b)\quad \mathfrak{L}\{(1/n^r) \textstyle\sum [(DX_i/DY_i)^r - (\pi r/\sin \pi r)]\} \to S[(1/r), -1, 0, (-(1/r)M(1/r) \cos (\pi/2r))], \qquad \tfrac{1}{2} < r < 1,$$

*where* $M(1/r) = \int_0^\infty [(e^{-y} - 1 + y)/y^{1+(1/r)}]\, dy$;

$$(3.42c)\quad \mathfrak{L}\{(1/n \log n)^{\frac{1}{2}} \textstyle\sum [(DX_i/DY_i)^{\frac{1}{2}} - (\pi/2)]\} \to \Phi(x), \qquad r = \tfrac{1}{2};$$

$$(3.42\text{d})\quad \mathcal{L}\{(1/n\sigma_0^2(n, r))^{\frac{1}{2}} \sum [(DX_i/DY_i)^r - (\pi r/\sin \pi r)]\} \to \Phi(x), \qquad 0 < r < \tfrac{1}{2},$$

*where*

$$\sigma_0^2(n, r) = [2\pi r/\sin 2\pi r - (\pi r)^2/\sin^2 \pi r] - (2\pi r^2/\sin \pi r)^2 \sum [1/n(n - i + 1)] \textstyle\sum_{j=i+1}^n (h'(j)/h^2(j)) + [2(\pi r^2)^2/\sin^2 \pi r] \sum (1/n)\{[1/(n - i + 1)] \sum_{j=i+1}^n (h'(j)/h^2(j))\}^2.$$

PROOF. Lemma 3.1 allows application of the results in Section 35 of Gnedenko and Kolmogorov for sums of independent, identically distributed random variables. Use of Theorem 2 of Section 35 yields both (3.42a) and (3.42b). The constants $\alpha$ and $\beta$ are found by using Equation (15) of Section 34 and Equations (7) and (8) of Section 35. The constant $c$ follows from the definitions used in Section 34. In (3.42a), $\eta$ can be found by direct evaluation of the characteristic function of $[(1/n)\sum (U_i/V_i) - \log n]$. The sums in (3.42b) are in the "domain of normal attraction" of their limit laws.

From Theorem 1, Section 35, (3.42c) follows. Notice that the sum here is not in the "domain of normal attraction" of the normal law since the second moment of $(U_i/V_i)^{\frac{1}{2}}$ does not exist.

In (3.42d), Section 35 does not apply since the quantities on the right side of (3.4c) are not identically distributed. The limiting normality is easily verified by the standard central limit theorem (Theorem 4, Section 21 of Gnedenko and Kolmogorov). A direct computation of the variance of the right side of (3.4c) leads to the given value of $\sigma_0^2(n, r)$.

Next we consider the distributions when $F(x) \neq G(x)$.

THEOREM 3.2. *If $F(x)$ and $G(x)$ satisfy the conditions of Theorem* 3.0, *the sums $A_n(r)$ converge to non zero limits, and $\sigma_1^2(n, r)$ converges to a non zero limit, then the distributions of the quantities below converge to the indicated* ID *laws.*

$$(3.43\text{a})\quad \mathcal{L}\{\log n[(1/n \log nA_n(1)) \sum (DX_i/DY_i) - 1] \to L(-1, 1), \qquad r = 1,$$

$$(3.43\text{b})\quad \mathcal{L}\{(1/n^rA_n(r)) \sum [(DX_i/DY_i)^r - (\pi rq^r(i)/h^r(i) \sin \pi r)] \to L(0, 1/r), \qquad \tfrac{1}{2} < r < 1,$$

$$(3.43\text{c})\quad \mathcal{L}\{(2/n \log nA_n(\tfrac{1}{2}))^{\frac{1}{2}} \sum [(DX_i/DY_i)^{\frac{1}{2}} - (\pi/2)(q(i)/h(i))^{\frac{1}{2}}]\} \to \Phi(x), \qquad r = \tfrac{1}{2},$$

*where* $A_n(r) = (1/n) \sum (q(i)/h(i))^{1/r}$;

$$(3.43\text{d})\quad \mathcal{L}\{(1/n\sigma_1^2(n, r))^{\frac{1}{2}} \sum [(DX_i/DY_i)^r - (\pi rq^r(i)/h^r(i) \sin \pi r)]\} \to \Phi(x), \qquad 0 < r < \tfrac{1}{2},$$

*where*

$$\sigma_1^2(n, r) = (1/n) \sum \{(2\pi r/\sin 2\pi r - (\pi r)^2/(\sin \pi r)^2)(q(i)/h(i))^{2r}$$
$$-2(\pi r^2/\sin \pi r)^2(q(i)/h(i))^r\{[1/(n - i + 1)] \textstyle\sum_{j=i+1}^{n-1} [q^r(j)h'(j)/h^{2+r}(j)]$$
$$+ [1/(n - i + 1)] \textstyle\sum_{j=i+1}^{n-1} [h^{-r}(j)q'(j)/q^{2-r}(j)]\}$$
$$+ (\pi r^2/\sin \pi r)^2\{([1/(n - i + 1)] \textstyle\sum_{j=i+1}^{n-1} [q^r(j)h'(j)/h^{2+r}(j)])^2$$
$$+ ([1/(n - i + 1)] \textstyle\sum_{j=i+1}^{n-1} [h^{-r}(j)q'(j)/q^{2-r}(j)])^2\}\}.$$

PROOF. Using Theorem 3.0, the problem is reduced to one concerning independent random variables. Equations (3.43a) and (3.43b) are verified using Theorem 1, Section 25 of Gnedenko and Kolmogorov (G-K). The fact that $\sigma^2 = 0$ follows from the convergence of $A_n(r)$ and Equation (11), Section 35 of G-K which holds by virtue of our Theorem 3.1.

Equation (3.43c) follows from Theorem 4, Section 26 of G-K. Equation (3.43d) is a result of the standard central limit theorem.

**4. Stochastic convergence.** In this section, we study the stochastic convergence of the statistics $S_n(r)$ whose distributions were considered in the previous section. Results similar to these were obtained by the author (1962) under somewhat more restrictive conditions on the distribution functions. We shall take advantage of the work of the previous section in obtaining the convergence theorems.

THEOREM 4.0. *Let $F(x)$ and $G(x)$ satisfy* (3.5a-1) *and assume there exists a function $v(\delta)$ with*

$$\lim_{\delta\to 0} v(\delta) = 0$$

*such that*

$$\text{(1)} \quad \lim_{\delta\to 0} \sup_{0\le x\le 1-v(\delta)}[a'(\delta, r)\delta^4(-\log \delta)(1 - x)^{-2} \cdot \textstyle\int_x^{1-\delta^2} [q^r(y)\alpha(y; \delta, r)]^2\, dy = 0; \tag{4.1a}$$

$$\text{(2)} \quad \lim_{\delta\to 0} [a'(\delta, r) \textstyle\int_{1-v(\delta)}^{1-\delta^2} [q^r(y)\alpha(y; \delta, r)]^2\, dy] = 0,$$

*where*

$$\begin{aligned} a'(\delta, r) &= \delta^2, & 0 < r < \tfrac{1}{2},\\ &= -\delta^2\log \delta, & r = \tfrac{1}{2},\\ &= \delta^{4(1-r)}(-\log \delta)^{2r-1}, & \tfrac{1}{2} < r < 1,\\ &= (1/(-\log \delta)), & r = 1, \end{aligned}$$

*and $\alpha(y; \delta, r)$ is given in* (3.5).

*Further, we require that*

$$\text{(i)} \quad \lim_{\delta\to 0} \{\delta \textstyle\int_{\delta^2}^{1-\delta^2} q^r(x)\alpha(x; \delta, r)\, dx\} = 0; \tag{4.1b}$$

$$\text{(ii)} \quad \lim_{\delta\to 0} \{\textstyle\int_{1-\delta}^{1-\delta^2} q^r(x)\alpha(x; \delta, r)\, dx\} = 0.$$

*These conditions must hold with r replaced by* $(-r)$ *and/or* $q(x)$ *interchanged with* $h(x)$.

*Then as n increases*

$$(4.2a)\quad (1/n \log n) \sum (DX_i/DY_i) \to_P (1/n \log n) \sum (q(i)U_i/h(i)V_i), \qquad r = 1,$$

$$(4.2b)\quad (1/n) \sum (DX_i/DY_i)^r \to_P (1/n) \sum (q(i)U_i/h(i)V_i)^r, \qquad 0 < r < 1.$$

PROOF. Essentially, Theorem 3.0 and Lemma 3.1 should be repeated with new normalizing constants. The crucial step in Theorem 3.0 is verifying that (3.23) converges to zero. The multiplier $a_n(r)$ now converges to zero for all $r$, so that in following the steps of examining (3.23), it is seen that (3.5a-1) alone suffices.

We must also verify that the second term in (3.22) goes to zero, but this is trivial since the revised $b_n(r) \to 0$ for all $r$.

In going through Lemma 3.1, again it is necessary to check (3.35) and (3.38) with appropriate re-definition of $a_n(r)$ and $d_n(r)$. Condition (4.1a) is a direct counterpart of (3.5b) with the revised $a(\delta, r)$ term. Because $d_n(r)$ for this Theorem is given by

$$\begin{aligned} d_n(r) &= n^{-1}, & 0 < r < 1, \\ &= n^{-1}(\log n)^{-\frac{1}{2}}, & r = 1, \end{aligned}$$

convergence of (3.38) is assured by convergence of

$$(4.3)\qquad \sum (1/n(n - i + 1)q^r(i)\alpha(i).$$

Use of the Abel sum formula gives as an upper bound on (4.3)

$$(4.4)\qquad (B_1/m) + B_m$$

where $B_i = \sum_{j=i}^{n-1} (q^r(j)\alpha(j)/n)$. Taking $m = n^{\frac{1}{2}}$ and identifying $\delta$ with $n^{-\frac{1}{2}}$ lead to condition (4.1b) to guarantee that (4.4) converges to zero. This completes the proof.

Note that all of the conditions of Theorem 4.0 are less restrictive than the corresponding conditions of Theorem 3.0, and the second order conditions are particularly weaker. This is naturally to be expected in view of the smaller normalizing constant and essentially "first order" nature of stochastic convergence. We might add that since $a_n(r)$ does converge to zero, (3.5a-1) is not necessary—the integral can diverge subject to some restraint on the rate of divergence. For $r = 1$, the restraints would be severe—almost equivalent to (3.5a-1), and we can verify that the distribution cited in the fourth remark following Lemma 3.1, provides an example where (4.2a) is false when $F(x) = G(x)$.

THEOREM 4.1. *Let* $F(x)$ *and* $G(x)$ *satisfy the conditions of Theorem* 4.0. *Then as n increases*

$$(4.5a)\quad (1/n \log n) \sum (DX_i/DY_i) \to_P \int_0^1 (g(G^{-1}(x))/f(F^{-1}(x)))\, dx, \qquad r = 1,$$

(4.5b) $1/n \sum (DX_i/DY_i)^r$

$$\to_P (\pi r/\sin \pi r) \int_0^1 (g(G^{-1}(x))/f(F^{-1}(x)))^r \, dx, \qquad 0 < r < 1.$$

PROOF. By Theorem 4.0, we need only consider the sums of independent random variables appearing in relations (4.2).

To establish (4.5a) or (4.5b) for $\frac{1}{2} \leqq r < 1$, the theorems of Section 28 of Gnedenko and Kolmogorov can be used. The strong law of large numbers (Section 27 (ibid)) establishes (4.5b) for $0 < r < \frac{1}{2}$. In this latter case, the functions of $U_i$ and $V_i$ converge in the strong sense of probability one to the given constant, but the convergence in Theorem 4.0 is not with probability one so that this statement cannot be made in (4.5b). The integrals given are the limiting values of the appropriate sums. (These are the quantities $A_n(r)$ which appear in Theorem 3.2.) This completes the proof.

**Acknowledgment.** The author wishes to thank the referee for his careful and painstaking reading and criticism which led eventually to a better paper.

REFERENCES

BARLOW, R. E., MARSHALL, A. W., and PROSCHAN, F. (1963). Properties of probability distributions with monotone hazard rate. *Ann. Math. Statist.* **34** 375–389

BLUMENTHAL, S. (1962). A test of the two-sample problem with nuisance location and scale parameters, and an estimate of the scale parameter. Technical Report No. 58, Department of Statistics, Stanford Univ.

BLUMENTHAL, S. (1963). The asymptotic normality of two test statistics associated with the two-sample problem. *Ann. Math. Statist.* **34** 1513–1523.

BLUMENTHAL, S. (1966). Contributions to sample spacings theory, II: tests of the parametric goodness of fit and two-sample problems. *Ann. Math. Statist.* **37** 925–939.

CRAMÉR, H. (1946). *Mathematical Methods of Statistics*. Princeton Univ. Press.

DARLING, D. A. (1957). The Kolmogorov-Smirnov, Cramér-von Mises tests. *Ann. Math. Statist.* **28** 823–838.

EPSTEIN, B. and SOBEL, M. (1954). Some theorems relevant to life testing from an exponential distribution. *Ann. Math. Statist.* **25** 373–381.

FISZ, M. (1962). Infinitely divisible distributions: recent results and applications. *Ann. Math. Statist.* **33** 68–84.

GNEDENKO, B. V. and KOLMOGOROV, A. N. (1954). *Limit Distributions for Sums of Independent Random Variables*. Addison-Wesley, Cambridge.

KOLMOGOROV, A. N. (1933). Sulla determinazione empirica di une legge di distribuzione. *Giorn. dell'Instit. degli att.* **4** 83–91.

LECAM, L. (1958). Une theoreme sur la division d'une intervalle par des points pris au hasard. *Pub. Inst. Statist. Univ. Paris* **7** 7–16.

LUKACS, E. (1960). *Characteristic Functions*. Hafner, New York

MENDELBROT, B. (1960). The Pareto-Lévy law and the distribution of income. *Int'l. Economic Rev.* **2** 79–106.

PROSCHAN, F. and PYKE, R. (1962). General limit theorems for spacings. *Ann. Math. Statist.* **33** 307.

PROSCHAN, F. and PYKE, R. (1965). Tests for monotone failure rate. Document No. DI-82-0422, Boeing Scientific Research Laboratories.

RÉNYI, A. (1953). On the theory of order statistics. *Acta Math. Acad. Sci. Hungar.* **4** 191–231.

SUKHATME, P. V. (1937). Tests of significance for samples of the $\chi^2$ population with two degrees of freedom. *Ann. Eugen.* **8** 52–56.

WEISS, L. (1961). On the estimation of scale parameters. *Naval Res. Logist. Quart.* **8** 245–256.