TESTS FOR MONOTONE FAILURE RATE

FRANK PROSCHAN
and
RONALD PYKE
BOEING SCIENTIFIC RESEARCH LABORATORIES

1. Introduction and summary

In applications such as life-testing and reliability, a useful characterization of distributions is in terms of their failure rates. The failure rate $q(t)$ of a distribution $F(t)$ having a density $f(t)$ is defined by $q(t) = f(t)/\{1 - F(t)\}$ for t such that $F(t) < 1$. The probabilistic interpretation is that corresponding to a failure distribution $F(t)$, $q(t)$ represents the conditional probability density of failure given that failure has not yet occurred by time t.

It is easy to verify that the exponential distribution with density

$$f(t) = \lambda e^{-\lambda t} \tag{1.1}$$

has a constant failure rate λ. Physically this might correspond to a situation in which the object fails if a sufficiently large environmental stress occurs, such stresses being distributed according to a Poisson process. It is assumed that the object develops no greater propensity toward failure as time elapses.

In many physical situations the object does become more vulnerable to failure with increasing age. This is characteristic of objects subject to wear-out—moving parts, human beings past youth, and so on. In such situations one would expect the failure distribution to be characterized by an increasing failure rate. Examples of such distributions are the gamma with density

$$f(t) = \frac{\lambda(\lambda t)^{\alpha-1}e^{-\lambda t}}{\Gamma(\alpha)}, \qquad \lambda > 0, \quad \alpha \geq 1, \quad t \geq 0, \tag{1.2}$$

and the Weibull with density

$$f(t) = \lambda\alpha t^{\alpha-1}e^{-\lambda t^{\alpha}}, \qquad \lambda > 0, \quad \alpha \geq 1, \quad t \geq 0. \tag{1.3}$$

In certain situations, however, it is reasonable to expect that the failure rate will decrease, at least over a certain interval of time. Thus, during the early months of human life, as a result of infancy diseases, the failure rate actually decreases. For certain electronic components, manufacturing defects tend to cause failure early in life, so that the failure rate may be higher during the initial period of usage. Materials which become work-hardened may exhibit a decreasing failure rate during a certain interval of time.

For both theoretical and practical reasons, it is important to distinguish between the case of monotone nondecreasing failure rate (referred to hereafter as increasing failure rate, and abbreviated IFR) and the complementary case (failure rate strictly decreasing over some time interval). Theoretically, knowing that the failure rate is increasing, we may do the following:

(1) obtain bounds on survival probability as functions of the moments, much improved over the usual Chebyshev bounds (see Barlow and Marshall [1]). We may also obtain bounds on the renewal function (see Barlow and Proschan [3]);

(2) obtain inequalities on the moments;

(3) conclude that the survival probability $1 - F(t)$ has the monotone likelihood ratio property in differences of t;

(4) show that certain basic operations, such as convolution and the formation of order statistics, preserve the property of increasing failure rate;

(5) develop certain variation diminishing properties for $1 - F(t)$.

See Barlow, Marshall, and Proschan [2] for a discussion of (2), (3), (4), and (5).

For practical applications, a knowledge that failure rate is increasing is also important, since under this assumption

(1) a policy of planned replacement of aging components may be more economical (see Barlow and Proschan [5]);

(2) the computation of optimum spares kits becomes easier (see Barlow and Proschan [4], chapter 6);

(3) the failure rate of certain types of systems composed of components with increasing failure rate is itself increasing (see Esary and Proschan [9]); and

(4) we may obtain the maximum likelihood estimate of the failure distribution (see Marshall and Proschan [17]).

It is clear that a test to determine whether a sample comes from a population characterized by an increasing failure rate would serve a useful purpose. In this paper, we propose and study the following nonparametric test. Let $X_{n1}, X_{n2}, \cdots, X_{nn}$ be a sample of independent observations from the common distribution F, with density f, where $f(t) = 0$ for $t < 0$, and failure rate $q(t)$. We wish to choose between the following:

(i) *Null Hypothesis*, H_0: $q(t) = \lambda$, λ an unknown positive constant.

(ii) *Alternative Hypothesis*, H_1: $q(t)$ is increasing, but not constant.

The test statistic is computed as follows. Let $T_{n,1} \leq T_{n,2} \leq \cdots \leq T_{n,n}$ be the ordered observations, $D_{n,1} = T_{n,1}$, $D_{n,2} = T_{n,2} - T_{n,1}, \cdots, D_{n,n} = T_{n,n} - T_{n,n-1}$, the spacings, and $\bar{D}_{n,1} = nD_{n,1}$, $\bar{D}_{n,2} = (n-1)\,D_{n,2}, \cdots, \bar{D}_{n,n} = D_{n,n}$, the normalized spacings. Let

$$V_{i,j} = \begin{cases} 1 & \text{if } \bar{D}_{n,i} \geq \bar{D}_{n,j} \quad \text{for} \quad i, j = 1, 2, \cdots, n, \\ 0 & \text{otherwise.} \end{cases} \tag{1.4}$$

The test statistic is

$$V_n = \sum_{i,j=1, i<j}^{n} V_{i,j}. \tag{1.5}$$

We reject the null hypothesis at the α level of significance if $V_n > v_{n,\alpha}$ where $v_{n,\alpha}$ is determined so that $P[V_n > v_{n,\alpha} | H_0] = \alpha$.

Heuristically, we may justify the test as follows. Under the null hypothesis, $\bar{D}_{n,1}, \bar{D}_{n,2}, \cdots, \bar{D}_{n,n}$ are independently distributed, each with density $\lambda e^{-\lambda t}$, as shown in Epstein and Sobel [8], so that $P[V_{i,j} = 1] = \frac{1}{2}$ for $i, j = 1, 2, \cdots, n$, $i \neq j$. However, under the alternative hypothesis, $P[V_{i,j} = 1] > \frac{1}{2}$ for $i < j$, $i, j = 1, 2, \cdots, n$, as will be shown in section 3. Thus, each $V_{i,j}$, and consequently V_n, tends to be larger under the alternative hypothesis, so that rejection of the null hypothesis occurs for large values of V_n. Since under the null hypothesis the distribution of V_n is known (see section 2), we have available $v_{n,\alpha}$.

In section 4 we show that V_n, suitably normalized, is asymptotically normally distributed for a wide class of alternatives. In particular, under mild assumptions when the underlying distribution is IFR, V_n is asymptotically normally distributed. In section 5 we use the criterion of Asymptotic Relative Efficiency (ARE) to compare the test based on V_n with the likelihood ratio test for Weibull alternatives and with the likelihood ratio test for gamma alternatives.

The results of the present paper are discussed in Pyke [18] and in Barlow and Proschan ([4], appendix 2).

2. Distribution under the null hypothesis

As mentioned above, under the null hypothesis, $\bar{D}_{n,1}, \bar{D}_{n,2}, \cdots, \bar{D}_{n,n}$ are independently distributed, each having density $\lambda e^{-\lambda t}$. Thus, all orderings of $\bar{D}_{n,1}, \bar{D}_{n,2}, \cdots, \bar{D}_{n,n}$ are equally likely. Let $P_n(k)$ be the number of orderings of $\bar{D}_{n,1}, \bar{D}_{n,2}, \cdots, \bar{D}_{n,n}$ with exactly k inversions of indices; an inversion of indices $i < j$ occurs when $\bar{D}_{n,i} > \bar{D}_{n,j}$. As shown by Kendall [12] and Mann [16], $P_n(k)$ satisfies the recurrence relation

$$P_n(k) = P_{n-1}(k) + P_{n-1}(k-1) + \cdots + P_{n-1}(k-n+1), \tag{2.1}$$

with $P_n(k) = 0$ for $k < 0$. Since $P[V_n = k] = P_n(k)/n!$, we may use (2.1) to calculate recursively $P[V_n = k]$. Note that the distribution of V_n is independent of λ. Tables are given in both Kendall [12] and Mann [16] for $P[V_n \leq k]$, $n \leq 10$.

To obtain the generating function $H_n(x) = \sum_{k=0}^{\infty} P_n(k)x^k$ of the $P_n(k)$, note that $0 = H_{n+1}(x) - H_n(x)\{1 + x + \cdots + x^n\}$ so that

$$H_n(x) = \prod_{i=1}^{n} \{1 + x + \cdots + x^{i-1}\}. \tag{2.2}$$

Thus, except for the factor $1/n!$, $H_n(x)$ corresponds to the generating function of $\sum_{i=1}^{n} U_i$, where $U_1, U_2, \cdots, U_n$ are n independent variables, with U_i being uniformly distributed on the integers $0, 1, 2, \cdots, i-1$. It follows that μ_n, the mean of V_n, is given by

$$\mu_n = \frac{n(n-1)}{4} \tag{2.3}$$

and σ_n^2, the variance of V_n, by

$$\sigma_n^2 = \frac{(2n+5)(n-1)n}{72}. \tag{2.4}$$

These results are obtained by Mann using a recursion among the moments, and by Feller [10] using a direct probability argument.

Both Kendall and Mann prove that the distribution of V_n is asymptotically normal by the method of moments. Also Dantzig [6] shows that solutions of a general class of recurrence relations, of which (2.1) is a special case, converge to the normal distribution. From (2.2), an immediate proof is obtained using the Lindeberg-Feller normal convergence criterion (Loève, [15], p. 280), by verifying that for given $\epsilon > 0$,

$$g_n(\epsilon) = \frac{1}{\sigma_n^2} \sum_{k=1}^{n} \int_{|x| \geq \epsilon\sigma_n} x^2 \, dF_k \to 0 \qquad \text{as} \quad n \to \infty, \tag{2.5}$$

where F_k is the distribution function of U_k defined just above. Formula (2.5) is readily verified by choosing n_0 large enough to insure that

$$\epsilon(2n_0^3 + 3n_0^2 - 5n_0)/72 > n_0, \tag{2.6}$$

since for $n \geq n_0$, $g_n(\epsilon) = 0$.

3. Unbiasedness of test

In this section we shall show that the test statistic V_n is unbiased, that is, $P[V_n \geq v_{n,\alpha}|H_1] \geq \alpha$ for $0 < \alpha \leq 1$, $n = 2, 3, \cdots$. Moreover, we shall show that under H_1 the $V_{i,j}$ are stochastically ordered, in that

$$P[V_{i,j} = 1|H_1] \geq P[V_{k,\ell} = 1|H_1] \qquad \text{for} \quad i \leq k < \ell \leq j. \tag{3.1}$$

To prove that V_n is an unbiased test statistic, we make the transformation

$$X'_{n,i} = -\ln \bar{F}(X_{n,i}), \tag{3.2}$$

where $\bar{F}(X) = 1 - F(X)$. It follows that

$$P[X'_{n,i} > u] = P[\ln \bar{F}(X_{n,i}) < -u] = P[\bar{F}(X_{n,i}) < e^{-u}] = e^{-u}. \tag{3.3}$$

Thus each $X'_{n,i}$ is distributed according to the exponential distribution with unit mean. Moreover, since the $X_{n,1}, \cdots, X_{n,n}$ are independent, so are the $X'_{n,1}, \cdots, X'_{n,n}$.

Next let $T'_{n,1} < T'_{n,2} < \cdots < T'_{n,n}$ represent the ranked $X'_{n,1}, X'_{n,2}, \cdots, X'_{n,n}$, so that $T'_{n,i} = -\ln \bar{F}(T_{n,i})$, $i = 1, 2, \cdots, n$. Further, let

$$\bar{D}'_{n,i} = \begin{cases} nT'_{n,1}, & i = 1, \\ (n - i + 1)(T'_{n,i} - T'_{n,i-1}), & i = 2, 3, \cdots, n. \end{cases} \tag{3.4}$$

According to Epstein and Sobel [8], the $\bar{D}'_{n,i}$ are independently, identically distributed according to the exponential distribution with unit mean.

Note that $T'_{n,i}$ is an increasing function of $T_{n,i}$. Moreover, as may be verified readily, $T'_{n,i}$ is a convex function of $T_{n,i}$, $i = 1, 2, \cdots, n$. It follows that $\bar{D}'_{n,i} \geq \bar{D}'_{n,j}$ implies $\bar{D}_{n,i} \geq \bar{D}_{n,j}$ for $i < j$, $i, j = 1, 2, \cdots, n$. Thus, $V_{i,j} \geq V'_{i,j}$, $i < j$, $i, j = 1, 2, \cdots, n$, where $V'_{i,j} = 1$ if $\bar{D}'_{n,i} \geq \bar{D}'_{n,j}$. Hence, $V_n \geq V'_n$ where $V'_n = \sum_{i<j} V'_{i,j}$, so that $P[V_n \geq v_{n,\alpha}|H_1] \geq \alpha$, for $0 < \alpha \leq 1$, $n = 2, 3, \cdots$, by Lehmann ([14], p. 73).

To prove (3.1), we note that $[\bar{D}_{n,i} \geq \bar{D}_{n,j}] = [\bar{D}'_{n,i} \geq K_{i,j}\bar{D}'_{n,j}]$ where $K_{i,j}$ is a random variable ≤ 1, for $i < j$, $i, j = 1, 2, \cdots, n$, whereas $[\bar{D}_{n,k} \geq \bar{D}_{n,\ell}] = [\bar{D}'_{n,k} \geq K_{k,\ell}\bar{D}'_{n,\ell}]$, where $K_{i,j} \leq K_{k,\ell}$. Hence, (3.1) follows by Lehmann ([14], p. 73).

The unbiasedness argument goes through for the more general class of statistics given in the first paragraph of section 4 below.

4. Asymptotic distribution under an alternative hypothesis

We now show that V_n when suitably normalized is asymptotically normally distributed for a wide class of alternatives. Since no additional effort is necessary, we shall in fact derive the asymptotic normality of a more general class of functions of spacings, namely those of the form

$$G_n = \sum_{i=2}^{n-1} \sum_{j=i+1}^{n} g(\bar{D}_{n,i}, \bar{D}_{n,j}) \tag{4.1}$$

where $g(x, y)$ is a bounded nonnegative function which is nonincreasing (nondecreasing) in its first (second) coordinate. If $g(x, y)$ is equal to 1 or 0 according as $x \geq y$ or $x < y$, then G_n is essentially equal to V_n. (The reason for the summation beginning at $i = 2$ instead of $i = 1$, is to allow for alternative distributions whose support does not contain the origin.) For convenience, assume $0 \leq g \leq 1$. We remark here that if g satisfies certain differentiability conditions, one may use a "Taylor's expansion" approach to prove the asymptotic normality of G_n. However, such an approach does not work for V_n, the statistic of primary interest in this paper. The more widely applicable approach used below is to find another random variable which is asymptotically equivalent to G_n and whose limiting distribution is more easily derived. Although the main interest of this paper is in the case of alternatives with increasing failure rates, the proof of asymptotic normality given below does not require monotonicity, but rather differentiability (a.e.) of the failure rate function. The limit theorem of this section generalizes the authors' theorem announced in Pyke [18].

The method of proof relies heavily upon a special construction of the random variables under consideration. In this construction, all random variables are functions of a sequence of independent exponential random variables. For the case of uniform random variables, this construction was used by Rényi [19] to obtain the limiting distributions of order statistics and statistics of the Kolmogorov-Smirnov type. In fact, as the last step in the proof of our theorem we shall use, in the form of lemma 4.1, Rényi's approximation of uniform order statistics.

The construction is as follows: let $\{Y_n;\ n \geq 1\}$ be a sequence of independent identically distributed exponential random variables with distribution function $H(y) = 1 - \exp(-y)$ for $y > 0$. For $1 \leq i \leq n$, set

$$Y_{n,i} = \sum_{j=1}^{i} Y_j(n - j + 1)^{-1}, \qquad U_{n,i} = H(Y_{n,i}) = 1 - \exp(-Y_{n,i}). \tag{4.2}$$

Then $(Y_{n,1}, Y_{n,2}, \cdots, Y_{n,n})$ and $(U_{n,1}, U_{n,2}, \cdots, U_{n,n})$ are the order statistics of random samples of size n from, respectively, the exponential distribution H and the uniform distribution function on $(0, 1)$.

Let F be a given alternative distribution function. Throughout most of this section we will use the following weak assumption.

ASSUMPTION 1. *The function F is absolutely continuous and its support is an interval (finite or infinite). Moreover, there exists a density function f of F which is continuous on the interior of this interval.*

The assumption that the support of F is an interval is natural for theorems on spacings, since if the support of F is disconnected, there will always be a fixed number of spacings which do not converge to zero. With assumption 1 in mind, let F^{-1} denote the inverse of F. Define $T_{n,i} = F^{-1}(H(Y_{n,i})) = F^{-1}(U_{n,i})$. Then $(T_{n,1}, T_{n,2}, \cdots, T_{n,n})$ are the order statistics of a random sample on F. Let K denote the composed function $F^{-1}H$, so that $T_{n,i} = K(Y_{n,i})$. Under assumption 1, F is continuously differentiable. Let k and h denote the derivatives of K and H, respectively. Then for all $v > 0$, $k(v) = h(v)/f(K(v))$. Since $h = 1 - H$ on $(0, \infty)$, it follows that if we set $k(H^{-1}(u)) = r(u)$, then

$$r(u) = \frac{(1-u)}{f(F^{-1}(u))} = \frac{1}{q(F^{-1}(u))} \tag{4.3}$$

for all $0 < u < 1$. This is to say that r is the reciprocal of the failure rate after it is transformed, by means of F^{-1}, onto the unit interval. The function r is the key function in the analysis to follow. (For notational convenience, we shall most frequently write r_u for $r(u)$.)

With the above construction of $(T_{n,1}, \cdots, T_{n,n})$, one obtains

$$\begin{aligned} \bar{D}_{n,i} &= (n-i+1)(T_{n,i} - T_{n,i-1}) \\ &= (n-i+1)[K(Y_{n,i}) - K(Y_{n,i-1})] \end{aligned} \tag{4.4}$$

Therefore, by the mean value theorem

$$\bar{D}_{n,i} = (n-i+1)(Y_{n,i} - Y_{n,i-1})k(\theta_{n,i}) = Y_i k(\theta_{n,i}) = Y_i r(A_{n,i}) \tag{4.5}$$

where $A_{n,i} = H(\theta_{n,i}) \in [U_{n,i-1}, U_{n,i}]$ and $\theta_{n,i}$ is the appropriate number in $[Y_{n,i-1}, Y_{n,i}]$. Since $A_{n,i}$ lies between the $(i-1)$-th and i-th order statistics from a uniform $-(0, 1)$ sample, and since $U_{n,i} - U_{n,i-1}$ is $0_p(n^{-1})$, it seems natural to expect that in (4.5) one might hope to replace $A_{n,i}$ with $U_{n,i}$ without affecting the limiting distribution of G_n. This substitution turns out to be possible. On the other hand, it is known that $U_{n,i}$ approximates i/n in the sense that $U_{n,i} - i/n$ is $0_p(n^{-1/2})$. Therefore, one might expect that in (4.5) one may replace $A_{n,i}$ with i/n without affecting the limiting distribution of G_n. This substitution, however, cannot be justified.

For $0 < u < v < 1$, set

$$\mathrm{L}(u, v) = E[g(Y_1 r_u, Y_2 r_v)] \tag{4.6}$$

where Y_1 and Y_2 are independent exponential random variables of mean 1. For $\delta > 0$, define

$$\bar{L}(u, v: \delta) = \sup_{|u-x|<\delta, |v-y|<\delta} L(x, y) \tag{4.7}$$

and define $\underline{L}(u, v: \delta)$ accordingly as the infimum of L over the same region. (It should be understood that only values $0 < x \leq y < 1$ in the domain of L are considered when taking its supremum and infimum.) Write $L^+ = \bar{L} - L$ and $L^- = L - \underline{L}$ to denote the upper and lower oscillations of L. Define

$$\lambda^+(\delta) = \int_0^1 \int_u^1 L^+(u, v: \delta)\, dv\, du \tag{4.8}$$

and define λ^- similarly in terms of L^-. Set $\lambda = \lambda^+ + \lambda^-$.

For most of the remaining results of this section, we require the following assumption which permits the passage of a limit inside an integral and provides an integrable absolute bound on certain integrands which arise.

ASSUMPTION 2. *The derivative r' of r exists a.e. and is continuous a.e. Moreover,*

$$\lim_{\delta \to 0} \delta^{-1}\lambda^+(\delta) = \lim_{\delta \to 0} \delta^{-1}\lambda^-(\delta) = \int_0^1 \int_u^1 K(u, v)\, dv\, du < \infty \tag{4.9}$$

where the integrals are Riemann integrals and where

$$K(u, v) = E\{g(Y_1 r_u, Y_2 r_v)[(Y_1 - 1)|r_v'|/r_v + (1 - Y_2)|r_u'|/r_u]\} \tag{4.10}$$

whenever r_u' and r_v' are defined.

We remark at this point that if one defines $\bar{r}(u: \delta) = \sup_{|u-x|<\delta} r(x)$ and defines $\underline{r}(u: \delta)$ as the analogous infimum, then due to the monotonicity of g one obtains $\bar{L}(u, v: \delta) = E[g(Y_1\bar{r}(u: \delta)), Y_2\underline{r}(v: \delta)]$ where one uses the natural interpretation when $\bar{r} = +\infty$. Moreover, to motivate the definition of K in (4.10), observe that whenever $\bar{r}_u < \infty$ and $\underline{r}_v > 0$,

$$L^+(u, v: \delta) = \int_0^\infty \int_0^\infty e^{-y-z}[g(y\bar{r}_u, z\underline{r}_v) - g(yr_u, zr_v)]\, dy\, dz \tag{4.11}$$

$$= \int_0^\infty \int_0^\infty g(s, t)[(\bar{r}_u\underline{r}_v)^{-1} \exp(-s/\bar{r}_u - t/\underline{r}_v) - (r_u r_v)^{-1} \exp(-s/r_u - t/r_v)]\, ds\, dt.$$

It may then be shown that

$$K(u, v) = |L_1(u, v)| + |L_2(u, v)| \tag{4.12}$$

where $L_1(L_2)$ denotes the partial derivative of L with respect to the first (second) coordinate. One may check that (4.10) and (4.12) are compatible by observing that

$$\begin{aligned} L_1(u, v) &= (r_u'/r_u)E[g(Y_1 r_u, Y_2 r_v)(1 - Y_1)], \\ L_2(u, v) &= (r_v'/r_v)E[g(Y_1 r_u, Y_2 r_v)(1 - Y_2)], \end{aligned} \tag{4.13}$$

whereas the expectations in the above expressions are respectively negative and positive due to the monotonicity properties of g. Observe also, that due to the boundedness of g, it follows from (4.12) and (4.13) that the integrability of $K(u, v)$ is equivalent to $\int_0^1 (d/du)|\log r(u)|\, du < \infty$

For $1 < i < j \leq n$, set

$$(4.14) \qquad G_{n,i,j} = g(\bar{D}_{n,i}, \bar{D}_{n,j}), \qquad S_{n,i,j} = (gY_i r_{i/n}, Y_j r_{j/n}),$$
$$T_{n,i,j} = G_{n,i,j} - S_{n,i,j}$$

and

$$(4.15) \qquad R_{n,i,j} = \left(U_{n,i} - \frac{i}{n}\right) L_1\left(\frac{i}{n}, \frac{j}{n}\right) + \left(U_{n,j} - \frac{j}{n}\right) L_2\left(\frac{i}{n}, \frac{j}{n}\right).$$

Let G_n, S_n, T_n, and R_n denote their respective sums over $1 < i < j \leq n$. We shall show first that G_n is asymptotically equivalent to $S_n + R_n$ in the sense that $n^{-3/2}(G_n - S_n - R_n) \xrightarrow{p} 0$. We then show that $n^{-3/2}R_n - R_n^* \xrightarrow{p} 0$ where R_n^* is defined in (4.55), and then derive the limiting distribution of $n^{-3/2}S_n + R_n^*$. The choice of R_n can be motivated by observing that $L(A_{n,i}, A_{n,j}) - L(i/n, j/n)$ would be equal to the expectation of $G_{n,i,j} - S_{n,i,j}$ with respect to Y_i and Y_j *if it were true that* $(A_{n,i}, A_{n,j})$ *and* (Y_i, Y_j) *were independent.* They are in fact dependent, but a basic step in the proof below is to show that the degree of dependence is negligible. The definition of $R_{n,i,j}$ given above is then a natural Taylor's expansion approximation to $L(A_{n,i}, A_{n,j}) - L(i/n, j/n)$.

THEOREM 4.1. *Under assumptions* 1 *and* 2, $n^{-3/2}(G_n - S_n - R_n) \xrightarrow{p} 0$.

PROOF. Throughout the proof we shall tacitly assume that r' exists and is continuous at the points $\{i/n: 1 \leq i < n, n > 1\}$. This can be done without loss of generality since the proof holds for any other partitions, $t_{n,1} < t_{n,2} < \cdots < t_{n,n}$, for which $|t_{n,i} - i/n| < 1/n$ and the desired properties are satisfied. This is clear since the only role played by these points is in the formation of approximating sums to certain Riemann integrals.

By a well-known theorem of Kolmogorov [13], one may find for each $\epsilon > 0$ a positive number b_ϵ such that for all n, $P(B_n(b_\epsilon n^{-1/2})) > 1 - \epsilon$ where

$$(4.16) \qquad B_n(\delta) = \left[\frac{i}{n} - U_{n,i-1} < \delta,\ U_{n,i} - \frac{i}{n} < \delta;\ 1 \leq i \leq n\right].$$

It therefore suffices to show that $n^{-3/2}(G_n - S_n - R_n)$ converges to zero in probability when restricted to the event $B_n \equiv B_n(bn^{-1/2})$ for all sufficiently large b; that is, to show that $n^{-3/2}(G_n - S_n - R_n)I_{B_n} \xrightarrow{p} 0$ where I_A denotes the indicator function of the event A. We shall assume b to be fixed throughout the proof.

For notational convenience, let us replace summations by integrals. To this end, define

$$(4.17) \qquad \begin{aligned} G_n(u, v) &= n^{1/2}G_{n,i,j}, & S_n(u, v) &= n^{1/2}S_{n,i,j}, \\ T_n(u, v) &= n^{1/2}T_{n,i,j}, & R_n(u, v) &= n^{1/2}R_{n,i,j}, \end{aligned}$$

for $i - 1 < nu \leq i$, $j - 1 < nv \leq j$ and $1 < i < j \leq n$. Define the functions to be zero elsewhere. Write

$$(4.18) \qquad \Delta_n = n^{-3/2}(G_n - S_n - R_n) = \int_0^1 \int_u^1 [T_n(u, v) - R_n(u, v)]\, dv\, du.$$

To prove that $\Delta_n \xrightarrow{p} 0$, we shall show that $E(\Delta_n^2 I_{B_n}) \to 0$. Observe first of all that

$$E(\Delta_n^2 I_{B_n}) \tag{4.19}$$
$$= \int_0^1 \int_u^1 \int_0^1 \int_x^1 E\{[T_n(u, v) - R_n(u, v)][T_n(x, y) - R_n(x, y)]I_{B_n}\} \, dy \, dx \, dv \, du.$$

We wish to show that the limit as $n \to \infty$ may be passed inside these integrals.

On B_n, $|R_n(u, v)| \le bK(u, v)$ by definitions (4.12) and (4.15). To get a bound on $T_n(u, v)$, set $\bar{r}_{n,i} = \bar{r}(i/n: bn^{-1/2})$ and $\underline{r}_{n,i} = \underline{r}(i/n: bn^{-1/2})$. Let $T_n^+(u, v)$ and $T_n^-(u, v)$ denote the positive and negative parts of $T_n(u, v)$, and set

$$\begin{aligned} T_{n,i,j}^{++} &= g(Y_i\bar{r}_{n,i}, Y_j\underline{r}_{n,j}) - S_{n,i,j}, \\ T_{n,i,j}^{--} &= S_{n,i,j} - g(Y_i\underline{r}_{n,i}, Y_j\bar{r}_{n,j}), \\ T_n^{++}(u, v) &= n^{1/2}T_{n,i,j}^{++}, \qquad T_n^{--}(u, v) = n^{1/2}T_{n,i,j}^{--} \end{aligned} \tag{4.20}$$

for all $1 < i < j \le n$, $i - 1 < nu \le i$ and $j - 1 < nv \le j$. On the event B_n, $\underline{r}_{n,i} \le r(A_{n,i}) \le \bar{r}_{n,i}$. Therefore,

$$0 \le T_n^+(u, v)I_{B_n} \le T_n^{++}(u, v), \qquad 0 \le T_n^-(u, v)I_{B_n} \le T_n^{--}(u, v). \tag{4.21}$$

Set

$$T_n^{++} = \int_0^1 \int_0^1 T_n^{++}(u, v) \, dv \, du, \qquad T_n^{--} = \int_0^1 \int_0^1 T_n^{--}(u, v) \, dv \, du. \tag{4.22}$$

We wish to show that the limits of $E(T_n^{++})$ and $E[(T_n^{++})^2]$ exist and may be passed inside the integral signs. By direct computation,

$$\operatorname{var}(T_n^{++}) = n^{-3} \textstyle\sum^* \{E[T_{n,i,j}^{++}T_{n,k,m}^{++}] - E(T_{n,i,j}^{++})E(T_{n,k,m}^{++})\} \tag{4.23}$$

where $\sum^*$ denotes summation over all values of $1 < i < j \le n$ and $1 < k < m \le n$ with the restriction that there are at most 3 distinct values among i, j, k, and m. (The summation over 4 distinct values is zero by independence.) Since the summands are bounded and converge to zero as $n \to \infty$ and as $i/n, j/n, k/n$, and m/n converge to values in $(0, 1)$, it follows from the Lebesgue dominated convergence theorem that the variance of T_n^{++} converges to zero. To show that $E(T_n^{++})$ converges to a finite limit, we compute using (4.7) that

$$\begin{aligned} E(T_n^{++}) &= n^{-3/2} \sum_{i<j} L^+\left(\frac{i}{n}, \frac{j}{n}: bn^{-1/2}\right) \\ &= n^{1/2} \int_0^1 \int_u^1 L_n^+(u, v: bn^{-1/2}) \, dv \, du \end{aligned} \tag{4.24}$$

where $L_n^+(u, v: bn^{-1/2})$ is defined to equal $L^+(u, v: bn^{-1/2})$ when $i - 1 < nu < i$ and $j - 1 < nv \le j$. However, since $L^+ = \bar{L} - L$ and

$$\begin{aligned} |L_n^+(u, v: \delta) - L^+(u, v: \delta)| \le{}& L^+(u, v: \delta + n^{-1}) \\ &- L^+(u, v: \delta) + L^+(u, v: n^{-1}), \end{aligned} \tag{4.25}$$

it follows from assumption 2 that

$$\lim_{n \to \infty} n^{1/2} \int_0^1 \int_u^1 [L_n^+(u, v: bn^{-1/2}) - L^+(u, v: bn^{-1/2})] dv \, du = 0. \tag{4.26}$$

Hence, again by assumption 2, it follows from (4.24) that

$$(4.27)\qquad \lim_{n\to\infty} E(T_n^{++}) = \lim_{n\to\infty} n^{1/2}\lambda^+(bn^{-1/2}) = b\int_0^1 \int_u^1 K(u, v)\, dv\, du < \infty.$$

Similar results may be shown to hold for the first two moments of T_n^{--}. If we now apply these results to the integrand in (4.19), we obtain that

$$(4.28)\qquad E(\Delta_n^2 I_{B_n}) \le \int_0^1 \int_u^1 \int_0^1 \int_x^1 E\{[T_n^{++}(u, v) + T_n^{--}(u, v) + bK(u, v)] \times [T_n^{++}(x, y) + T_n^{--}(x, y) + bK(x, y)]\}\, dy\, dx\, dv\, du$$

and that the limit of the integral on the right-hand side of (4.28) exists and equals the integral of the limit, namely,

$$(4.29)\qquad 9b^2\left\{\int_0^1 \int_u^1 K(u, v)\, dv\, du\right\}^2 < \infty.$$

If $\{f_n\}$ and $\{g_n\}$ are sequences of measurable functions on some measure space and if $g_n \to g$ (a.e.), $f_n \to f$ (a.e.), $|f_n| \le g_n$ and $\int g_n \to \int g < \infty$, then $\int f_n \to \int f$. This is a direct consequence of Fatou's lemma applied to the functions $g_n - f_n$ and $g_n + f_n$. Therefore, if we replace the integrand in (4.19) by f_n and the integrand on the right-hand side of (4.28) by g_n, it follows from this remark and the above results that to prove $E(\Delta_n^2 I_{B_n}) \to 0$, it suffices to show that $f_n \to 0$ (a.e.); that is, to show that

$$(4.30)\qquad \lim_{n\to\infty} E\{[T_n(u, v) - R_n(u, v)][T_n(x, y) - R_n(x, y)]I_{B_n}\} = 0$$

for almost all points (u, v, x, y) with $0 < u < v < 1$ and $0 < x < y < 1$. Equivalently, we will show that each of the four cross-products in (4.30) con-converges to the same finite limit. In what follows, consider (u, v, x, y) to be fixed, and set $i = [nu]$, $j = [nv]$, $k = [nx]$ and $m = [ny]$. The dependence upon n of the indices i, j, k, and m will not be reflected in the notation, but no confusion should result.

To show (4.30), observe first that if (Y_i, Y_j, Y_k, Y_m) could be considered to be independent of all other random variables appearing in (4.30), then $T_n(u, v)$ and $T_n(x, y)$ could be replaced by their expectations with respect to (Y_i, Y_j, Y_k, Y_m), namely $n^{1/2}[L(A_{n,i}, A_{n,j}) - L(i/n, j/n)]$ and $n^{1/2}[L(A_{n,k}, A_{n,m}) - L(k/n, m/n)]$, respectively. It would then be a straightforward matter to show that the limit in (4.30) exists and is zero. Although (Y_i, Y_j, Y_k, Y_m) is in fact not independent of the other random variables in (4.30), the degree of dependence is sufficiently negligible, as we shall now show.

According to the construction of the random variable $U_{n,i}$ given in (4.2), $1 - U_{n,i} = \exp\{-\sum_{m=1}^{i} Y_m/(n - m + 1)\}$. For any set of nonnegative integers J, define $U_{n,t}^J$ by

$$(4.31)\qquad 1 - U_{n,t}^J = (1 - U_{n,i}) \exp\Big\{\sum_{i \ge s \in J} Y_s/(n - s + 1)\Big\}.$$

Clearly, $\{U_{n,t}^J\colon 1 \le i \le n\}$ and $\{Y_s\colon s \in J\}$ are independent sets of random variables. If, moreover, J is finite, then

$$(4.32) \qquad n(U_{n,i} - U^J_{n,i}) = (1 - U_{n,i})[\exp\{\sum_{i \geq s \in J} Y_s/(n - s + 1\} - 1]n$$

is convergent in probability. In the discussion which follows, we set $J = \{i, j, k, m\}$. (Actually, J depends upon n since i, j, k, and m do.)

For all s, $U_{n,s} - A_{n,s} \leq U_{n,s} - U_{n,s-1}$. Moreover, from Chebyshev's inequality, using third moments, one obtains that $nP[U_{n,s} - U_{n,s-1} > \epsilon n^{-1/2}] \to 0$ for all $\epsilon > 0$. Similarly, one can show that $nP[U_{n,s} - U^J_{n,s} > \epsilon n^{-1/2}] \to 0$ since $U^J_{n,s} \geq U_{n,s-4}$. These statements can also be shown to hold for i, j, k, and m simultaneously. Therefore, one obtains that for any $\epsilon > 0$ there is an event $C_{n,J}$ such that on $C_{n,J}$, $|U_{n,s} - U^J_{n,s}| \leq \epsilon$ for all $s \in J$ and such that $nP(C^c_{n,J}) \to 0$. Consequently, each of the four terms in (4.30), when restricted to the complement of $C_{n,J}$, converges to zero. It therefore remains to show that

$$(4.33) \qquad \lim_{n\to\infty} nE\{[T_n(u, v) - R_n(u, v)][T_n(x, y) - R_n(x, y)]I_{B_n}I_{C_{n,J}}\} = 0.$$

For $s - 1 < nu \leq s$, write $Z_n(u) = n^{1/2}(U_{n,s} - s/n)$, $Z^J_n(u) = n^{1/2}(U^J_{n,s} - s/n)$, and let $R_{n,1}(u, v)$ be defined in the same way as $R_n(u, v)$, except that Z^J_n replaces Z_n; that is

$$(4.34) \qquad R_{n,1}(u, v) = Z^J_n(u)L_1(u, v) + Z^J_n(v)L_2(u, v).$$

Set $R_{n,2} = R_n - R_{n,1}$. Then on $C_{n,J}$,

$$(4.35) \qquad |R_{n,2}(u, v)| \leq \epsilon K(u, v).$$

Also, let $T_{n,1}(u, v)$ be the same as $T_n(u, v)$ except that $A_{n,i}$ and $A_{n,j}$ are replaced by $U^J_{n,i}$ and $U^J_{n,j}$; that is,

$$(4.36) \qquad T_{n,1}(u, v) = n^{1/2}[g(Y_i r(U^J_{n,i}), Y_j r(U^J_{n,j})) - S_{n,i,j}].$$

Set $T_{n,2} = T_n - T_{n,1}$. Then on $C_{n,J}$;

$$(4.37) \qquad n^{-1/2}(T_{n,2}(u, v))^+ \leq g(Y_i\bar{r}(U^J_{n,i}\colon \epsilon n^{-1/2}), Y_j\underline{r}(U^J_{n,j}\colon \epsilon n^{-1/2})) - g(Y_i r(U^J_{n,i}), Y_j r(U^J_{n,j}))$$

with an analogous inequality for the negative part of $T_{n,2}$. If one now substitutes $R_n = R_{n,1} + R_{n,2}$ and $T_n = T_{n,1} + T_{n,2}$ into (4.33) and multiplies out the product, one is faced with the problem of evaluating limits for several terms of which two are given by

$$(4.38) \qquad a_n = E[T_{n,2}(u, v)R_{n,1}(x, y)I_{B_n \cap C_{n,J}}]$$

and

$$(4.39) \qquad b_n = E[T_{n,1}(u, v)R_{n,1}(x, y)I_{B_n \cap C_{n,J}}].$$

We shall study only these two terms since the other terms may be treated similarly.

First of all, upon splitting both $T_{n,2}$ and $R_{n,1}$ into their positive and negative parts and using (4.37), one obtains

$$(4.40) \qquad a_n \leq n^{1/2}E\{[L^+(U^J_{n,i}, U^J_{n,j}\colon \epsilon n^{-1/2}) + L^-(U^J_{n,i}, U^J_{n,j}\colon \epsilon n^{-1/2})]2bK(x, y)I_{B_{n,J}}\}$$

where

$$(4.41)\qquad B_n \subset B_{n,J} = \left[n^{1/2}\left| U^J_{n,i} - \frac{i}{n}\right| \leq b + 3n^{-1/2};\ 1 \leq i \leq n \right],$$

since (Y_i, Y_j) is independent of all other random variables including $I_{B_{n,J}}$. Because of (4.11), the random variables $n^{1/2}L^+$ and $n^{1/2}L^-$, which appear in (4.40) are bounded when restricted to $B_{n,J}$ for sufficiently large n. Moreover, each of these random variables converges in law to $\epsilon K(u, v)$, whereas $I_{B_{n,J}}$ converges in law to I_B where $B[|Z(t)| \leq b\colon 0 \leq t \leq 1]$ and where $\{Z(t)\colon 0 \leq t \leq 1\}$ is a tied-down Wiener process. (These remarks follow from the fact that the Z^J_n-process converges weakly to the Z-process and that $\delta^{-1}L^+(u, v\colon \delta) \rightarrow K(u, v)$ uniformly in a neighborhood of (u, v) since r' is assumed to be continuous at u and v.) It therefore follows that the expectation in (4.40) also converges, so that

$$(4.42)\qquad \limsup_{n\rightarrow\infty} a_n \leq 4\epsilon b K(u, v)K(x, y)P(B)$$

which bound converges to zero as $\epsilon \rightarrow 0$. Similar arguments may be applied to show that each of the other terms in (4.33) which involve either $R_{n,2}$ or $T_{n,2}$ converges to zero.

The other type of term which must be considered is like b_n in that neither $T_{n,2}$ nor $R_{n,2}$ appears. Each term of this type can be shown to have the same limit so that the resultant sum of these limits, with the appropriate signs, is zero as required. The proof of (4.33), and hence of the theorem, will then be complete.

To evaluate the limit of b_n, one first uses an argument similar to that used above for a_n to show that

$$(4.43)\qquad b'_n \equiv E\{T_{n,1}(u, v)R_{n,1}(x, y)[I_{B_{n,J}} - I_{B_n \cap C_{n,J}}]\} \rightarrow 0.$$

For this, one uses the fact that $P(B_{n,J}\backslash B_n) \rightarrow 0$. On the other hand, consider

$$\begin{aligned}(4.44)\qquad b''_n \equiv b_n + b'_n &= E\{T_{n,1}(u, v)R_{n,1}(x, y)I_{B_{n,J}}\}\\ &= E\left\{n^{1/2}\left[L(U^J_{n,i}, U^J_{n,j}) - L\left(\frac{i}{n}, \frac{j}{n}\right)\right] R_{n,1}(x, y)I_{B_{n,J}}\right\},\end{aligned}$$

the latter equation resulting from taking expectations with respect to (Y_i, Y_j). It follows from (4.11), (4.15), and (4.41) that the random variable inside the expectation in (4.44) is uniformly bounded for all n and converges in law to $R(u, v)R(x, y)I_B$ where B is as defined above and

$$(4.45)\qquad R(u, v) = Z(u)L_1(u, v) + Z(v)L_2(u, v).$$

Hence, the limit may be passed inside the expectation sign to yield

$$(4.46)\qquad \lim_{n\rightarrow\infty} b''_n = E[R(u, v)R(x, y)I_B].$$

The proof of theorem 4.1 is therefore complete.

In order to complete the derivation of the asymptotic normality of G_n, it only remains to derive the limiting distribution of $n^{-3/2}[S_n + R_n - E(S_n)]$. It should be observed that the asymptotic normality of S_n can be derived straightforwardly by the method of moments, since S_n is a sum of random variables $\{S_{n,i,j}\}$ which are independent unless one or more subscripts are the same. Also, the asymptotic normality of R_n can be derived immediately using the weak convergence of the Z_n-process to the Z-process. However, since we are interested in the limiting distribution of their sum $S_n + R_n$, it is necessary to know something of the joint behavior of S_n and R_n. This joint behavior may be explicitly described by making use of the following lemma: if we write

$$W_n(u) = n^{1/2}W_{n,i} = n^{1/2}(1-u)\sum_{k=1}^{i}(Y_k - 1)(n-k+1)^{-1}, \tag{4.47}$$

for $i - 1 < nu \leq i$, this lemma will enable us to replace $Z_n(u)$ in the definition of R_n given in (4.7) with $W_n(u)$.

LEMMA 4.1. *For all* $i = 1, 2, \cdots, n$,

$$E\left|U_{n,i} - \frac{i}{n} - W_{n,i}\right| < cn^{-1} \tag{4.48}$$

where c is a constant independent of i and n.

PROOF. Since $W_{n,n} = 0$ and $E(1 - U_{n,n}) = (n+1)^{-1}$, we can restrict attention to values of $i < n$. Because of (4.2), one may write

$$\begin{aligned} E[U_{n,i} - i/(n+1) - W_{n,i}]^2 = \operatorname{var}(U_{n,i}) + E(W_{n,i}^2) \\ + 2E[W_{n,i}\exp(-Y_{n,i})]. \end{aligned} \tag{4.49}$$

Clearly, $\operatorname{var}(U_{n,i}) = i(n-i+1)/(n+1)^2(n+2)$, and

$$E(W_{n,i}^2) = \left(1 - \frac{i}{n}\right)^2 \sum_{k=1}^{i}(n-k+1)^{-2}. \tag{4.50}$$

Moreover, straightforward computation yields

$$\begin{aligned} &E[W_{n,i}\exp(-Y_{n,i})] \\ &= \left(1-\frac{i}{n}\right)\sum_{k=1}^{i} E\left\{\frac{Y_k - 1}{n-k+1}e^{-Y_k(n-k+1)^{-1}} \prod_{j=1, j\neq k}^{i} e^{-Y_j(n-j+1)^{-1}}\right\} \\ &= \left(1-\frac{i}{n}\right)\sum_{k=1}^{i}(-1)(n-k+1)^{-1}(n-k+2)^{-1}(n-i+1)(n+1)^{-1} \end{aligned} \tag{4.51}$$

since $E[\exp(-aY)] = (1+a)^{-1}$ and $E[aY\exp(-aY)] = (1+a)^{-2}$ for an exponential random variable Y with mean 1. Therefore, (4.49) becomes

$$\begin{aligned} &\frac{i(n-i+1)}{(n+1)^2(n+2)} + \left(1-\frac{i}{n}\right)^2 \sum_{k=1}^{i}(n-k+1)^{-2} \\ &- 2\left(1-\frac{i}{n}\right)(n-i+1)(n+1)^{-1}\sum_{k=1}^{i}(n-k+1)^{-1}(n-k+2)^{-1}. \end{aligned} \tag{4.52}$$

However, upon approximating the sums in these expressions by integrals, one obtains

$$\frac{i}{(n+1)(n-i+1)} \leq \sum_{k=1}^{i}(n-k+1)^{-2} \leq \frac{i}{n(n-i)} \tag{4.53}$$

and

$$\frac{(i-1)}{(n+2)(n-i+2)} \leq \sum_{k=1}^{i} (n-k+1)^{-1}(n-k+2)^{-1} \leq \frac{i}{n(n-i)}. \tag{4.54}$$

The proof may now be completed by direct computation.

As an application of lemma 4.1, one may prove the following lemma.

LEMMA 4.2. *Under assumption* 1 *and* 2, $n^{-3/2}R_n - R_n^* \xrightarrow{p} 0$ *as* $n \to \infty$, *where*

$$R_n^* = \int_0^1 \int_u^1 [W_n(u)L_1(u, v) + W_n(v)L_2(u, v)]\, dv\, du. \tag{4.55}$$

PROOF. This is immediate, since by definition,

$$n^{-3/2}R_n = \int_0^1 \int_u^1 [Z_n(u)L_1(u, v) + Z_n(v)L_2(u, v)]\, dv\, du, \tag{4.56}$$

and since by lemma 4.1, $E|Z_n(u) - W_n(u)| < cn^{-1/2}$.

The coefficient of $n^{-1/2}(Y_k - 1)$ in (4.55) is straightforwardly checked to be $\beta(k/n - 1/n)$ where β is defined by

$$\begin{aligned}(1-w)\beta(w) = \int_w^1 \int_u^1 (1-u)L_1(u, v)\, dv\, du \\ + \int_w^1 \int_0^v (1-v)L_2(u, v)\, du\, dv.\end{aligned} \tag{4.57}$$

Therefore, theorem 4.1 and lemma 4.2 state that $n^{-3/2}[G_n - E(S_n)]$ has the same limiting distribution function as

$$\begin{aligned}\eta_n &\equiv n^{-3/2}[S_n - E(S_n)] + R_n^* \\ &= n^{-3/2}[S_n - E(S_n)] + n^{-1/2} \sum_{k=2}^{n} (Y_k - 1)\beta\left(\frac{k}{n} - \frac{1}{n}\right)\end{aligned} \tag{4.58}$$

whenever it possesses one. But η_n may be written in the form $\eta_n = n^{-3/2} \sum_{i<j} \eta_{n,i,j}$ where $\eta_{n,i,j}$ and $\eta_{n,k,m}$ are independent random variables (with mean zero) whenever i, j, k, and m are distinct. This leads one to the main theorem of this section, namely theorem 4.2.

THEOREM 4.2. *Under assumptions* 1 *and* 2, $n^{-3/2}[G_n - E(S_n)]$ *converges in law as* $n \to \infty$ *to a normal random variable with mean zero and variance* σ^2 *where* $\sigma^2 = \sigma_S^2 + 2\sigma_{S,R} + \sigma_R^2$, σ_S^2 *is given by* (4.62) *below,*

$$\sigma_{S,R} = -\int_0^1 \int_u^1 \left[\frac{L_1(u, v)r_u}{r_u'\beta(u)} + \frac{L_2(u, v)r_v}{r_v'\beta(v)}\right] dv\, du \tag{4.59}$$

and

$$\sigma_R^2 = \int_0^1 [\beta(u)]^2\, du. \tag{4.60}$$

PROOF. As mentioned above, $n^{-3/2}[G_n - E(S_n)]$ will have the same limiting distribution as η_n where η_n is given by (4.58). Let σ_S^2 and σ_R^2 denote the limiting variances of the first and second terms, respectively, in the right-hand side of (4.57). Then it is immediate that σ_R^2 is given by (4.60). Moreover, $\sigma_R^2 < \infty$, since from (4.57) it follows that for all w, $\beta(w) \leq 2 \int_0^1 \int_u^1 K(u, v)\, dv\, du$, which is finite by assumption 2. To compute σ_S^2, consider

$$(4.61) \qquad \operatorname{var}(n^{-3/2}[S_n - E(S_n)]) = n^{-3} \sum\nolimits^* \{E(S_{n,i,j}S_{n,k,m}) - E(S_{n,i,j})E(S_{n,k,m})\}$$

where $\sum^*$ denotes summation over all values of $i < j$ and $k < m$ with the restriction that there are at most 3 distinct values among i, j, k, and m. [The summation over 4 distinct values is zero by independence.] Since $\sum^* = 2\sum_{i<j}\sum_{j<k} + \sum_{i<j}\sum_{i<k} + \sum_{i<j}\sum_{k<j}$ and since the summands represent bounded Riemann-integrable functions, one obtains that the limiting variance of $n^{-3/2}S_n$ is

$$(4.62a) \qquad \begin{aligned} \sigma_S^2 = {} & \int_0^1 \int_u^1 \int_u^1 [H_1(u, v, w) - L(u, v)L(u, w)]\, dw\, dv\, du \\ & + 2\int_0^1 \int_u^1 \int_v^1 [H_2(u, v, w) - L(u, v)L(v, w)]\, dw\, dv\, du \\ & + \int_0^1 \int_0^w \int_0^w [H_3(u, v, w) - L(u, w)L(v, w)]\, du\, dv\, dw \end{aligned}$$

where

$$(4.62b) \qquad \begin{aligned} H_1(u, v, w) &= E[g(Y_1r_u, Y_2r_v)g(Y_1r_u, Y_3r_w)], \\ H_2(u, v, w) &= E[g(Y_1r_u, Y_2r_v)g(Y_2r_v, Y_3r_w)], \\ H_3(u, v, w) &= E[g(Y_1r_u, Y_3r_w)g(Y_2r_v, Y_3r_w)]. \end{aligned}$$

The limiting covariance term $\sigma_{S,R}$ may also be computed directly from (4.58) and (4.13) to be that given in (4.59).

The proof of the asymptotic normality of η_n may now be completed by the method of moments. By using the independence of $\eta_{n,i,j}$ and $\eta_{n,k,m}$ when the subscripts are distinct, one may show that all terms in the s-th moment of $n^{-3/2}[\eta_n - E(\eta_n)]$, $s > 2$, involve either less than (when s is odd) or less than or equal to (when s is even) $3s/2$ different subscripts in the product of $\eta_{n,i_1,j_1}\eta_{n,i_2,j_2} \cdots \eta_{n,i_s,j_s}$. Since there is a factor of $n^{-3s/2}$ multiplying each such product, the s-th moment converges to zero when s is odd and to $\sigma^s s!2^{-s/2}/(s/2)!$ when s is even, as required.

In this paper we are primarily interested in the case of a nondecreasing failure rate function, or equivalently of a nonincreasing r. For this case, r' is nonnegative, and some simplifications are possible in the statement of the assumptions of theorem 4.2. When r is nonincreasing, $L(u, v)$ is nonincreasing in u and nondecreasing in v. Hence, $\bar{L}(u, v\colon \delta) = L(u - \delta, v + \delta)$. This enables one to check that (4.9) of assumption 2 holds. Moreover, since $|r'_u| = -r'_u$ and $K(u, v) = L_2(u, v) - L_1(u, v)$, one obtains

$$(4.63) \qquad \int_0^1 \int_0^1 K(u, v)\, dv\, du = \int_0^1 [L(u, 1) - L(0, u)]\, du.$$

We therefore have corollary 4.1.

COROLLARY 4.1. *If r is nonincreasing* (IFR) *and r' exists and is continuous a.e. on* $(0, 1)$, *then the conclusion of theorem* 4.2 *holds.*

The conditions of this corollary are satisfied by the examples considered in the following section. We also remark that among IFR distributions, (4.63) equals zero only when r is constant.

It should be observed that if r is piecewise constant, then r' exists and equals 0, a.e. Hence, even if there is a countably infinite number of disjoint open intervals in (0, 1) on each of which r is constant, one still has $r' = 0$, a.e., and hence $n^{-3/2}R_n \xrightarrow{p} 0$. This indicates the difficulty that would arise if one attempted to prove the limiting normality of G_n for general r by first proving it for piecewise constant failure rates and then using these to approximate a general r.

We conclude this section by evaluating the several quantities which appear in the above analysis, for the special case ot the statistic V_n; that is, the case of $g(x, y) = 1$ or 0 according as $x \geq$ or $< y$. For this case,

$$L(u, v) = P[Y_1 r_u \geq Y_2 r_v] = r_u(r_u + r_v)^{-1}, \tag{4.64}$$

$$L_1(u, v) = -L_2(v, u) = r_v r_u'(r_u + r_v)^{-2}, \tag{4.65}$$

$$\begin{aligned} H_1(u, v, w) &= P[Y_1 r_u \geq Y_2 r_v, Y_1 r_u \geq Y_3 r_w] \\ &= r_u(r_u + r_v)^{-1} - r_w(r_w + r_u)^{-1} + r_v r_w(r_v r_w + r_u r_v + r_u r_w)^{-1}, \end{aligned} \tag{4.66}$$

$$\begin{aligned} H_2(u, v, w) &= P[Y_1 r_u \geq Y_2 r_v \geq Y_3 r_w] \\ &= r_u(r_u + r_v)^{-1} - r_u r_w(r_u r_w + r_u r_v + r_w r_v)^{-1}, \end{aligned} \tag{4.67}$$

$$\begin{aligned} H_3(u, v, w) &= P[Y_1 r_u \geq Y_3 r_w, Y_2 r_v \geq Y_3 r_w] \\ &= r_u r_v(r_u r_v + r_u r_w + r_v r_w)^{-1}. \end{aligned} \tag{4.68}$$

All of the results and proofs of this section have centered upon r, the transformed inverse of the failure rate function q. For applications it is desirable to transform the results back into q. For example, the asymptotic mean which will be of central importance in the following section is

$$\mu = \int_0^1 \int_u^1 L(u, v)\, dv\, du = \int_0^\infty \int_x^\infty q(y)[q(x) + q(y)^{-1}f(x)f(y)\, dy\, dx. \tag{4.69}$$

5. Asymptotic relative efficiency

In this section we use the criterion of Asymptotic Relative Efficiency (ARE) to compare the test based on V_n to two other possible tests. For some specified set of alternatives indexed by θ, say, the ARE of one sequence of tests based on a sequence of asymptotically normal test statistics $\{T_n\}$ against a second sequence of tests based on the asymptotically normal test statistics $\{\tau_n\}$ is defined as

$$\left\{\frac{[\mu_T'(\theta_0)]^2}{\sigma_T^2(\theta_0)}\right\}\left\{\frac{[\mu_\tau'(\theta_0)]^2}{\sigma_\tau^2(\theta_0)}\right\}^{-1} \tag{5.1}$$

whenever it exists. In (5.1), $\mu_T(\theta)$ and $\sigma_T^2(\theta)$ denote the limiting mean and variance respectively of $\{T_n\}$, μ_T' denotes the derivative of μ_T with respect to θ, and θ_0 denotes the null hypothesis. A similar interpretation is understood for the notation in the denominator relative to τ.

It should be remarked that before one may actually relate (5.1) to the limiting ratio of the sample sizes required by the two tests to achieve the same power (which is the usual definition of ARE, due to Pitman (see, for example, Fraser ([11], chapter 7)), one must know that the asymptotic normality of the test statistics is uniform in θ as $\theta \to \theta_0$. In order to obtain a uniformity result like this for V_n, one would require something analogous to the Berry-Esseen theorem for double sums of random variables like (4.58), as well as uniform bounds on the several estimates made in the proof of theorem 4.1. With this proviso in mind, it is believed that (5.1) is a useful preliminary measure of the comparative merits of tests. It is computed in what follows for V_n against the likelihood-ratio test for Weibull alternatives and against the likelihood-ratio test for gamma alternatives.

(a) *Likelihood-ratio test for Weibull alternatives.* Suppose that H_0 is as before, but H_1 is specialized to the case in which the underlying distribution is given by the Weibull formula with increasing failure rate:

$$F(x) = 1 - e^{-\lambda x^\theta}, \qquad \lambda > 0, \qquad \theta > 1, \qquad x \geq 0, \tag{5.2}$$

with λ assumed known. The likelihood ratio test is to reject the null hypothesis if

$$\max_{\theta \geq 1} \left[\lambda^n \theta^n \left(\prod_{i=1}^n X_{n,i}\right)^{\theta-1} \exp\left(-\lambda \sum_{i=1}^n X_{n,i}^\theta\right)\right] \Big/ \lambda^n \exp\left(-\lambda \sum_{i=1}^n X_{n,i}\right) > c_\alpha. \tag{5.3}$$

For θ close to 1, after simplification, this is equivalent to rejecting if the test statistic $T_n^W > c'_\alpha$, where

$$T_n^W = \sum_{i=1}^n (1 - \lambda X_{n,i}) \ln X_{n,i}. \tag{5.4}$$

Thus

$$\mu_W(\theta) = \int_0^\infty (1 - \lambda x)(\ln x)\theta\lambda x^{\theta-1} e^{-\lambda x^\theta}\, dx, \tag{5.5}$$

so that

$$\begin{aligned} \mu'_W(1) &= \int_0^\infty \{1 + \ln x - \lambda x \ln x\}\{(1 - \lambda x)(\ln x)\lambda e^{-\lambda x}\}\, dx \\ &= (\ln \lambda + \gamma - 1)^2 + \frac{\pi^2}{6}, \quad \text{where} \quad \gamma = .5772156649 \cdots . \end{aligned} \tag{5.6}$$

Similarly, we may compute $\sigma_W^2(1) = (\ln \lambda + \gamma - 1)^2 + \pi^2/6$. Hence,

$$\frac{\{\mu'_W(1)\}^2}{\sigma_W^2(1)} = (\ln \lambda + \gamma - 1)^2 + \frac{\pi^2}{6}. \tag{5.7}$$

For the test statistic V_n we compute

$$\mu(\theta) = \int_0^\infty \int_x^\infty \frac{q(y)}{q(x) + q(y)} \theta\lambda y^{\theta-1} e^{-\lambda y^\theta} \theta\lambda x^{\theta-1} e^{-\lambda x^\theta}\, dy\, dx \tag{5.8}$$

where the failure rate of the Weibull, $q(x) = \theta\lambda x^{\theta-1}$. After simplification, we find $\mu'(1) = \frac{1}{4} \ln 2$. From section 2 we see that $\sigma^2(1) = 1/36$. Hence,

$$\frac{\{\mu'(1)\}^2}{\sigma^2(1)} = \frac{9}{4}(\ln 2)^2. \tag{5.9}$$

Thus from (5.1), the ARE_W becomes

$$\text{(5.10)} \qquad \text{ARE}_W = \tfrac{9}{4}(\ln 2)^2 \Big/ \left\{(\ln \lambda + \gamma - 1)^2 + \frac{\pi^2}{6}\right\}$$

$$= \frac{1.0809}{(\ln \lambda - .4228)^2 + 1.6449}.$$

Note that as $\lambda \to 0$ or ∞, $\text{ARE}_W \to 0$. For all $\lambda > 0$, $\text{ARE}_W \leq .6571$; equality is attained for $\ln \lambda = .4228$.

(b) *Likelihood ratio test against the gamma distribution.* Next assume that H_1 is specialized to the case in which the underlying distribution is the gamma with increasing failure rate; the corresponding density is

$$\text{(5.11)} \qquad f(x) = \frac{\lambda^\theta x^{\theta-1} e^{-x}}{\Gamma(\theta)}, \qquad \theta \geq 1.$$

The likelihood ratio test is to reject the null hypothesis if

$$\text{(5.12)} \qquad \frac{\max\limits_{\theta \geq 1} \left\{\lambda^{n\theta} \left(\prod\limits_{i=1}^{n} X_{n,i}\right)^{\theta-1} \exp\left(-\lambda \sum\limits_{i=1}^{n} X_{n,i}\right) \Big/ [\Gamma(\theta)]^n\right\}}{\lambda^{n\theta} \exp\left(-\lambda \sum\limits_{i=1}^{n} X_{n,i}\right)} > c_\alpha;$$

or, equivalently for θ close to 1, if $T_n^G = \sum_{i=1}^n \ln X_{n,i} > c'_\alpha$. A similar calculation to that in (a) yields $(\{\mu'_G(1)\}^2/\sigma_G^2(1)) = \pi^2/6$.

For the test statistic V_n, we compute

$$\text{(5.13)} \qquad \mu(\theta) = \int_0^\infty \int_x^\infty \frac{q(y)}{q(x) + q(y)} f(x) f(y)\, dy\, dx$$

$$= \int_0^\infty \int_x^\infty \frac{y^{\theta-1} e^{-\lambda y} \cdot \int_x^\infty t^{\theta-1} e^{-\lambda t}\, dt\, \lambda^{2\theta} (xy)^{\theta-1} e^{-\lambda(x+y)}}{[\Gamma(\theta)]^2 \left\{x^{\theta-1} e^{-\lambda x} \int_y^\infty t^{\theta-1} e^{-\lambda t}\, dt + y^{\theta-1} e^{-\lambda y} \int_x^\infty t^{\theta-1} e^{-\lambda t}\, dt\right\}}\, dx\, dy.$$

After a good deal of calculation, we obtain $\mu'(1) = \frac{1}{2} \ln 2 - \frac{1}{4}$. Thus,

$$\text{(5.14)} \qquad \frac{\{\mu'(1)\}^2}{\sigma^2(1)} = 9\,(\ln 2 - \tfrac{1}{2})^2.$$

Hence,

$$\text{(5.15)} \qquad \text{ARE}_G \frac{9\,(\ln 2 - \frac{1}{2})}{\pi^2/6} = .2040.$$

(c) *Likelihood ratio test against the gamma when the true distribution is the Weibull.* Next let us compute the ARE of the test statistic $T_n^G = \sum_{i=1}^n \ln X_{n,i}$ when the distribution is Weibull, equation (5.2), under H_1. Recall that T_n^G is the asymptotic likelihood ratio test against the gamma distribution.

Since

$$\text{(5.16)} \qquad \mu_{G,W}(\theta) = \int_0^\infty (\ln x) \lambda \theta x^{\theta-1} e^{-\lambda x^\theta}\, dx,$$

after simplification we obtain $\mu'_{G,W}(\theta) = \gamma + \ln \lambda$. Also we obtain $\sigma^2_{G,W}(1) = \pi^2/6$. Hence,

$$\text{ARE}_{G,W} = \frac{6(\gamma + \ln \lambda)^2}{\pi^2 \left\{ (\ln \lambda + \gamma - 1)^2 + \dfrac{\pi^2}{6} \right\}}. \tag{5.17}$$

For $\lambda = 1$, $\text{ARE}_{G,W} = .1111$. Thus for $\lambda = 1$, the V_n test is considerably better since the corresponding value of ARE_W is .5927. However, in general, the ratio

$$\frac{\text{ARE}_{G,W}}{\text{ARE}_W} = \frac{6(\gamma + \ln \lambda)^2}{1.0809\pi^2} \tag{5.18}$$

may be larger than 1; in fact, for λ sufficiently large, the ratio becomes arbitrarily large. This implies that for large λ, the statistic V_n will be considerably inferior to T_n^G.

(d) *Likelihood ratio test against the Weibull when the true distribution is the gamma.* Next let us compute the ARE of the test statistic $T_n^W = \sum_{i=1}^n (1 - \lambda X_{n,i}) \ln X_{n,i}$ when the distribution is gamma, equation (5.11), under H_1. Recall that T_n^W is the asymptotic likelihood ratio test against the Weibull distribution. Since

$$\mu_{W,G}(\theta) = \int_0^\infty (1 - \lambda x)(\ln x) \frac{\lambda^\theta x^{\theta - 1}}{\Gamma(\theta)} e^{-\lambda x}\, dx, \tag{5.19}$$

we obtain after simplification $\mu'_{W,G}(\theta) = \ln \lambda + \gamma$. We also obtain $\sigma^2_{W,G}(\theta) = (\ln \lambda + \gamma - 1)^2 + \pi^2/6$. Thus we have

$$\text{ARE}_{W,G} = \frac{36(\ln \lambda + \gamma)^2}{\pi^4 + 6\pi^2(\ln \lambda + \gamma - 1)^2}. \tag{5.20}$$

At $\lambda = 1$, $\text{ARE}_{W,G} = .1111$. Thus for $\lambda = 1$, the V_n test is better than T_n^W since the corresponding value of $\text{ARE}_W = .2040$. However, in general, the ratio

$$\frac{\text{ARE}_{W,G}}{\text{ARE}_G} = \frac{36(\ln \lambda + \gamma)^2}{.2040\{\pi^4 + 6\pi^2(\ln \lambda + \gamma - 1)^2\}} \tag{5.21}$$

may be made greater than 1 by taking λ sufficiently large. In fact, as $\lambda \to \infty$, the ratio approaches 2.98.

REFERENCES

[1] R. E. Barlow and A. W. Marshall, "Bounds for distributions with monotone hazard rate," Parts I and II, *Ann. Math. Statist.*, Vol. 35 (1964), pp. 1234–1274.

[2] R. E. Barlow, A. W. Marshall, and F. Proschan, "Properties of probability distributions with monotone hazard rate," *Ann. Math. Statist.*, Vol. 34 (1963), pp. 375–389.

[3] R. W. Barlow and F. Proschan, "Comparison of replacement policies, and renewal theory implications," *Ann. Math Statist.*, Vol. 35 (1964), pp. 577–589.

[4] ———, *Mathematical Theory of Reliability*, New York, Wiley, 1965.

[5] ———, "Planned replacement," *Studies in Applied Probability and Management Science*, edited by Arrow, Karlin, and Scarf, Stanford, Stanford University Press, 1962, pp. 63–87.

[6] G. B. DANTZIG, "On a class of distributions that approach the normal distribution function," *Ann. Math. Statist.*, Vol. 10 (1939), pp. 247–253.

[7] B. EPSTEIN, "The exponential distribution and its role in life testing," *Industrial Quality Control*, Vol. 15 (1958), pp. 2–7.

[8] B. EPSTEIN and M. SOBEL, "Life testing," *J. Amer. Statist. Assoc.*, Vol. 48 (1953), pp. 486–502.

[9] J. D. ESARY and F. PROSCHAN, "Relationship between system failure rate and component failure rates," *Technometrics*, Vol. 5 (1963), pp. 183–189.

[10] W. FELLER, "The fundamental limit theorems in probability," *Amer. Math. Soc. Bull.*, Vol. 51 (1945), pp. 800–832.

[11] D. A. S. FRASER, *Nonparametric Methods in Statistics*, New York, Wiley, 1957.

[12] M. G. KENDALL, "A new measure of rank correlation," *Biometrika*, Vol. 30 (1938), pp. 81–93.

[13] A. N. KOLMOGOROV, "Sulla determinazione empirica di une legge di distribuzione," *Giorn. Ist. Ital. Attuari*, Vol. 4 (1933), pp. 83–91.

[14] E. L. LEHMANN, *Testing Statistical Hypotheses*, New York, Wiley, 1959.

[15] M. LOÈVE, *Probability Theory*, Princeton, Van Nostrand, 1960.

[16] H. B. MANN, "Nonparametric tests against trend," *Econometrica*, Vol. 13 (1945), pp. 245–249.

[17] A. W. MARSHALL and F. PROSCHAN, "Maximum likelihood estimation for distributions with monotone failure rate," *Ann. Math. Statist.*, Vol. 36 (1965), pp. 69–77.

[18] R. PYKE, "Spacings," *J. Roy. Statist. Soc. Ser. B*, Vol. 27 (1965), pp. 395–449.

[19] A. RÉNYI, "On the theory of order statistics," *Acta Math. Acad. Sci. Hungar.*, Vol. 4 (1953), pp. 191–231.