The Annals of Mathematical Statistics
1972, Vol. 43, No. 5, 1400-1411

CONVERGENCE OF QUANTILE AND SPACINGS PROCESSES WITH APPLICATIONS[1]

BY GALEN R. SHORACK

University of Washington and Mathematisch Centrum

The quantile process was shown by Bickel to converge in the uniform metric on intervals $[a, b]$ with $0 < a < b < 1$. By introducing appropriate new supremum metrics, this result is extended to all of (0, 1). Hence a natural process of ordered spacings from the uniform distribution converges in certain supremum metrics. This is used to establish the limiting normality of a large family of statistics based on ordered spacings, which can be used in testing for exponentiality. The non-null case is also considered.

I. QUANTILES

1. Introduction. Let $X_1, \cdots, X_n$ be a random sample from a df F and let $\mathbb{F}_n$ denote the empirical df. We wish to study the quantile process on (0, 1) defined as

$$n^{\frac{1}{2}}(\mathbb{F}_n^{-1} - F^{-1}) .$$

The appendix of Shorack (1972) should be regarded as a preliminary part of this paper. (It contains a number of results from Pyke and Shorack (1968) in a form we will find convenient.) Theorems and equations from that appendix will be referred to routinely as Theorem A1 and (A1) respectively, etc. In particular, special independent Uniform (0, 1) rv's $\xi_1, \cdots, \xi_n$ are defined in Section A1. These rv's have empirical df Γ_n, and quantile process V_n defined in (A2). Also V_n converges to a special Brownian bridge V in the sense of (A4).

The quantile process has the same finite dimensional distributions as does the process on (0, 1) defined as

$$n^{\frac{1}{2}}[F^{-1}(\Gamma_n^{-1}) - F^{-1}] .$$

We will now study the convergence of the process

$$Q_n = n^{\frac{1}{2}}[g(\Gamma_n^{-1}) - g] \tag{1}$$

on (0, 1). The functions g considered below are quite general continuous functions. The most interesting case of course is when $g = F^{-1}$ for some df F, but we do not require this. However, we will refer to Q_n as the *quantile process*.

Theorem 1 below is the main theorem of Part I. However when the more difficult Condition 1 can be verified, Corollary 1 gives a stronger conclusion. Corollary 2 allows g' to have discontinuities (it should be regarded as a useful

---

Received February 5, 1971; revised February, 1972.

[1] Research for this paper was carried out under NSF Contract, Number GP-13739, while the author was visiting the Mathematisch Centrum in Amsterdam.

AMS 1970 *subject classifications.* Primary 62E20, 60K10; Secondary 62G30, 60B10.

Key words and phrases. Quantile process convergence, testing exponentiality, renewal process.

technique, rather than as a useful result). Corollary 3 allows g to depend on n. We first obtain asymptotic normality of the quantiles in Proposition 1 (which has been proved many times). See Bickel (1967) for Proposition 2.

We comment briefly now on Part II below. The ordered spacings of a renewal process are closely related to the obvious set of order statistics. In Section 6 we will introduce a problem for this kind of spacings, and use Theorem 1 in our solution of it. (Some results of Shorack (1972) could also have been made to depend on Theorem 1.)

Now for differentiable functions g on (0, 1)

$$Q_n = A_n V_n \qquad \text{where} \quad A_n = [g(\Gamma_n^{-1}) - g]/(\Gamma_n^{-1} - I) \tag{2}$$

is a difference quotient defined (for a.e. fixed ω) by left continuity at the most n values of t where $\Gamma_n^{-1} = I$. We define the Q process at all points in (0, 1) where g' exists by

$$Q = g'V . \tag{3}$$

Note that the covariance function of Q is

$$K_Q(s, t) = (s \wedge t - st)g'(s)g'(t) .$$

2. The quantiles. For $0 < p < 1$ let X_{np} denote the $([np] + 1)$th order statistic, where [ ] denotes the greatest integer function. Note that $X_{np} = g(\xi_{np}) = g(\Gamma_n^{-1}(p))$, where in this section we set $g = F^{-1}$.

PROPOSITION 1. (*Asymptotic normality of the sample quantiles*). *Suppose* $0 < p_1 < \cdots < p_\kappa < 1$ *and* $g'(p_k)$ *exists for* $1 \leqq k \leqq \kappa$. *Then the random vector*

$$(n^{\frac{1}{2}}(X_{np_1} - g(p_1)), \cdots, n^{\frac{1}{2}}(X_{np_\kappa} - g(p_\kappa)))$$

is asymptotically multivariate normal with mean vector 0 *and covariance matrix* $||\sigma_{jk}||$ *given by*

$$\sigma_{jk} = p_j(1 - p_k)g'(p_j)g'(p_k) \qquad \text{for} \ \ 1 \leqq j \leqq k \leqq \kappa .$$

PROOF. Suppose $\kappa = 1$. Then $n^{\frac{1}{2}}(X_{np} - g(p)) =_{\text{a.s.}} A_n(p)V_n(p)$. Now $V_n(p) \rightarrow_e V(p)$ and $A_n(p) \rightarrow_e g'(p)$ since $\Gamma_n^{-1}(p) \rightarrow_e p$ and $g'(p)$ exists. Thus $n^{\frac{1}{2}}(X_{np} - g(p)) \rightarrow_{\text{a.s.}} g'(p)V(p)$. For $\kappa > 1$ we simply observe that a random vector converges a.s. if and only if each of its coordinates does. The vector $(g'(p_1)V(p_1), \cdots, g'(p_\kappa)V(p_\kappa))$ clearly has the stated normal distribution. □

3. The main theorem on convergence of the quantile process.

LEMMA 1. *If* g *has a nonzero continuous derivative* g' *on* (0, 1), *then for all* $\varepsilon > 0$

$$\rho^{\varepsilon}_{|g'|}(A_n, g') \equiv \sup_{\varepsilon \leqq t \leqq 1-\varepsilon} |A_n(t) - g'(t)|/|g'(t)| \rightarrow_e 0 \qquad \text{as} \ \ n \rightarrow \infty .$$

PROOF. By the mean value theorem $|A_n(t) - g'(t)|$ equals $|g'(s) - g'(t)|$ for some s between t and $\Gamma_n^{-1}(t)$. For n exceeding some $n_{\varepsilon,\omega}$ we have $\varepsilon/2 \leqq \Gamma_n^{-1}(t) \leqq 1 - \varepsilon/2$ for all $\varepsilon \leqq t \leqq 1 - \varepsilon$. Also $\rho(\Gamma_n^{-1}, I) \rightarrow_e 0$ and g' is uniformly continuous on $[\varepsilon/2, 1 - \varepsilon/2]$. Hence $\sup_{\varepsilon \leqq t \leqq 1-\varepsilon} |A_n(t) - g'(t)| \rightarrow_e 0$. Finally, $|g'|$ is bounded away from 0 on $[\varepsilon, 1 - \varepsilon]$. □

PROPOSITION 2. *If g has a continuous derivative on $[\alpha, \beta]$ for $0 < \alpha < \beta < 1$, then for any $\alpha < a < b < \beta$ we have*

$$\sup_{a \leqq t \leqq b} |Q_n(t) - Q(t)| \to_e 0 \qquad \text{as } n \to \infty .$$

PROOF. Write $Q_n - Q = (A_n - g')V_n + g'(V_n - V)$; and apply Lemma 1 above and Remark A5 and (A4). □

CONDITION 0. g has a nonzero continuous derivative g' on (0, 1). Also $|g'| \leqq R$ on (0, 1) where R is a reproducing u-shaped (increasing) function of Definition A3 for which

$$\zeta(t)R(t)/g'(t) \to 0 \qquad \text{as } t \to 0 \text{ or } 1 \ (\text{as } t \to 1) .$$

(See Lemma A4 for the definition of ζ in terms of q. For $q = [I(1 - I)]^{\frac{1}{2}-\delta}$ for some $\delta > 0$, we could take $\zeta = [I(1 - I)]^{\delta/2}$.)

THEOREM 1 (*Convergence of the quantile process in supremum metrics*). *Suppose Condition* 0 *holds for a particular q in the class $\mathscr{Q}$ of Definition* A1. *Then $\rho_{q|g'|}(Q_n^*, Q) \to_p 0$ as $n \to \infty$; where $*$ restricts functions on* (0, 1) *to* $[1/n, 1 - 1/n]$ *in the sense of Definition* A2.

PROOF. Let Condition 0 hold with Ru-shaped. Using the triangle inequality

$$\rho_{q|g'|}(Q_n^*, Q) \leqq (\alpha_n + M)\rho_{q\zeta}(V_n^*, V) + \alpha_n \rho_{q\zeta}(V, 0) = o_p(1) + \alpha_n O_p(1)$$

where $\alpha_n = \rho_{|g'|/\zeta}(A_n^*, g')$ and $M = \rho(\zeta, 0)$. From the mean value theorem $|A_n(t) - g'(t)| = |g'(s) - g'(t)| \leqq R(s) + R(t)$ for some s between t and $\Gamma_n^{-1}(t)$. Thus on the set $S_{n,\varepsilon}$ of Lemma A3 we have $|A_n^* - g'| \leqq 2R_\beta \leqq 2M_\beta R$ with $\beta = \beta_\varepsilon$. Thus for $\varepsilon > 0$ and some $\delta = \delta_\varepsilon > 0$ sufficiently small we have from the limit condition of Condition 0 that

$$\chi(S_{n,\varepsilon})\alpha_n \leqq \varepsilon + \rho^{\delta}_{|g'|}(A_n, g')\rho(\zeta, 0) ;$$

this can for large n be made (by Lemma 1) to exceed 2ε with probability not exceeding $1 - \varepsilon$. Thus $\alpha_n \to_p 0$.

The proof for R increasing is analogous. □

EXAMPLE 1. Let $F(x) = 1 - e^{-x}$ for $x \geqq 0$. Then $g = F^{-1} = -\log(1 - I)$ and $g' = (1 - I)^{-1}$. Condition 0 holds for every q in $\mathscr{Q}$ with $R = (1 - I)^{-1}$. Thus for all q in $\mathscr{Q}$

$$\rho_{q/(1-I)}(Q_n^*, Q) \to_p 0 \qquad \text{as } n \to \infty .$$

Note also that $*$ need only restrict functions on (0, 1) to $[1/n, 1)$ in this example.

4. Some variations on the main theorem. Condition 1 below is in the spirit of Chernoff, *et al.* (1967).

CONDITION 1. g has nonzero continuous derivative on (0, 1). There exists a $\delta > 0$ such that for all positive β in some neighborhood of 0 there exists $0 < M_\beta < \infty$ such that $|g'(s)/g'(t)| < M_\beta$ whenever $\beta t \leqq s \leqq t + \beta$ and $t \leqq \delta$ and whenever $\beta(1 - t) \leqq 1 - s \leqq (1 - t) + \beta$ and $t \leqq \delta$.

COROLLARY 1. *If Condition* 1 *holds, then for all* q *in* $\mathscr{Q}$ *we have* $\rho_{q|g'|}(Q_n^*, Q) \to_p 0$ *as* $n \to \infty$.

PROOF. Just work Remark A4 into the proof of Theorem 1. □

CONDITION 2. g is absolutely continuous on $(0, 1)$ and g' exists a.e. $|\nu|$. Also $|g'| \leqq R$ a.e. with respect to Lebesgue measure on $(0, 1)$ where R is a reproducing u-shaped (or increasing) function for which $\int_0^1 qRd|\nu| < \infty$ for some q in $\mathscr{Q}$.

COROLLARY 2 (*Convergence of the quantile process in integral metrics*). *If Condition* 2 *holds, then* $\|Q_n^* - Q\|_\nu \to_p 0$ *as* $n \to \infty$ *for* $\| \quad \|_\nu$ *as defined in Section* A2.

PROOF. Basically similar to the proof of Theorem 1. The triangle inequality reduces the problem to one of showing $\int_0^1 q|A_n^* - g'|d|\nu| \to_p 0$. Pointwise convergence of the integrand for every fixed ω is trivial. Note that for the set $S_{n,\varepsilon}$ of Lemma A3

$$\chi(S_{n,\varepsilon})|A_n^* - g'| \leqq 2\chi(S_{n,\varepsilon})(R \vee R(\Gamma_n^{-1})) \leqq 2R_\beta \leqq 2M_\beta R$$

by writing $A_n = \int_I^{\Gamma_n^{-1}} g'(s)\, ds/(\Gamma_n^{-1} - I)$. □

5. Uniformity. We next take up a theorem aimed particularly at letting $g_n = F_n^{-1}$ where F_n is a sequence of df's that converges to a fixed F_0. We now redefine so that

$$Q_n = n^{\frac{1}{2}}[g_n(\Gamma_n^{-1}) - g_n] \qquad \text{and} \qquad Q = g_0'V\,.$$

CONDITION 3. g_0 has a nonzero continuous derivative g_0' on $(0, 1)$. Each g_n has a continuous derivative g_n' on $(0, 1)$. Also $|g_n'| \leqq R$ on $(0, 1)$ for all n where R is a reproducing u-shaped (increasing) function for which

$$\zeta(t)R(t)/|g_0'(t)| \to 0 \qquad \text{as } t \to 0 \text{ or } 1 \text{ (as } t \to 1)\,.$$

For any $\varepsilon > 0$ the functions g_n restricted to $[\varepsilon, 1 - \varepsilon]$ form a uniformly equicontinuous family for which $\sup_{\varepsilon \leqq t \leqq 1-\varepsilon} |g_n(t) - g_0(t)| \to 0$ as $n \to \infty$.

COROLLARY 3. *If Condition* 3 *holds for particular* q *in* $\mathscr{Q}$, *then* $\rho_{q|g_0'|}(Q_n^*, Q) \to_p 0$ *as* $n \to \infty$.

PROOF. This is but a minor variation on the proof of Theorem 1. □

EXAMPLE 2. Let $F_n(x) = 1 - \exp\{-x^{(1+\theta_n)}\}$ for $x \geqq 0$ where $\theta_n \to 0$ as $n \to \infty$. Let $F_0(x) = 1 - e^{-x}$ for $x \geqq 0$. This is the case of Weibull df's converging to the exponential. Now $g_n = [-\log(1 - I)]^{1/(1+\theta_n)}$ and $g_0 = -\log(1 - I)$. Let $R = I^{-\delta}(1 - I)^{-(1+\delta)}$ for any $\delta > 0$. ($I^{-\delta}(1 - I)^{-1}$ suffices for R when $\theta_n > 0$ and $(1 - I)^{-(1+\delta)}$ suffices when $\theta_n < 0$.) We conclude from Corollary 3 that $\rho_\phi(Q_n^*, Q) \to_p 0$ as $n \to \infty$ for any $\phi = I^{\frac{1}{2}-\delta}(1 - I)^{-\frac{1}{2}-\delta}$ with $\delta > 0$. We may let * restrict functions on $(0, 1)$ to $[1/n, 1)$ in this conclusion.

II. SPACINGS

6. A problem. Let $X_1, \cdots, X_n$ be independent rv's having df F with $F(0) = 0$ and having empirical df $\mathbb{F}_n$. Let $g = F^{-1}$. Let $\eta \equiv E(X) > 0$ and let $\operatorname{Var}[X] < \infty$.

Let $D_i = X_i/\bar{X}$ for $1 \leqq i \leqq n$ where $\bar{X} = n^{-1}\sum_{i=1}^n X_i$; and let $0 \leqq D_{n1} \leqq \cdots \leqq D_{nn}$ denote the ordered values. Note that $D_{ni} = X_{ni}/\bar{X}$ where $0 \leqq X_{n1} \leqq \cdots \leqq X_{nn}$ are the order statistics.

If F is an exponential df, then $D_1/n, \cdots, D_n/n$ are distributed as the n spacings from a sample of size $n-1$ from the Uniform (0, 1) distribution.

A standard statistical problem is to test the hypothesis that a given stochastic process is a Poisson process against the alternative that it is some other renewal process. Many statistics proposed for this problem are of the form

$$T_n = n^{-1}\sum_1^n c_{ni} h(D_{ni})\,;$$

where h is a specified function and where the c_{ni}'s form a triangular array of known constants. The null hypothesis of course specifies an exponential df F.

We will establish the asymptotic normality of T_n (under regularity) in Theorem 6 as a consequence of the convergence in Theorem 3 of a certain process H_n based on the D_{ni}'s. The null hypothesis will receive special attention in Theorems 2 and 5. Theorem 4 (first proved by Pyke (1967)) and Corollary 4 are digressions.

Related results are contained in Darling (1953), Pyke (1967), Bickel and Doksum (1969) and Bickel (1969). See Example 5 below in this regard.

7. The spacings processes. The natural *ordered spacings process* or *inverse spacings process* is defined on [0, 1] by

$$H_n(t) = n^{\frac{1}{2}}[D_{ni} - g(t)/\eta] \tag{4}$$

for $(i-1)/n < t \leqq i/n$ and $1 \leqq i \leqq n$ with $H_n(0) = 0$. Note that

$$H_n = n^{\frac{1}{2}}[g(\Gamma_n^{-1})/\bar{X} - g/\eta] \qquad \text{on } (0, 1)\,.$$

For the Brownian bridge $U = -V$ of Section A1 we define

$$H = -g'U/\eta - Zg/\eta^2 \qquad \text{on } (0, 1)\,; \tag{5}$$

where

$$Z = -\int_0^\infty U(F)\,dI$$

is a $N(0, \mathrm{Var}[X])$ rv having

$$\mathrm{Cov}[Z, U(t)] = -\int_0^\infty [t \wedge F(x) - tF(x)]\,dx\,.$$

In case $F(x) = 1 - e^{-x}$ for $x \geqq 0$, we give the processes H_n, H the special labels D_n, D and call D_n the *ordered uniform spacings process.* Thus

$$D = -U/(1-I) + Z\log(1-I)\,. \tag{6}$$

(Note that for exponential F, the processes H_n and H do not depend on η.) Note that D is a normal process on [0, 1) having continuous sample paths, mean value function 0 and covariance function

$$K_D(s, t) = s/(1-s) - \log(1-s)\log(1-t)$$

for $0 \leqq s \leqq t < 1$. (There will be no need to confuse this process D_n with the spacing D_n.)

Less important for our purposes is the *empirical spacings process*

$$G_n(y) = n^{\frac{1}{2}}[\mathbb{F}_n(\bar{X}y) - F(\eta y)] \qquad \text{for } 0 \leqq y < \infty . \tag{7}$$

Note that $n\mathbb{F}_n(\bar{X}y)$ equals the number of D_i's not exceeding y. Define

$$G(y) = U(F(\eta y)) + Zyf(\eta y) \qquad \text{for } 0 \leqq y < \infty \tag{8}$$

where f is the density function of F.

THEOREM 2 (*Convergence of the ordered uniform spacings process in supremum metrics*). *Let F be an exponential* df. *Then for all q in $\mathcal{Q}$ we have*

$$\rho_{q/(1-I)}(D_n^*, D) \to_p 0 \qquad as \quad n \to \infty$$

where $*$ restricts functions on $[0, 1)$ *to* $[1/n, 1)$.

THEOREM 3 (*Convergence of the ordered spacings process in supremum metrics*). *Suppose $g = F^{-1}$ is such that the conclusion of Theorem* 1 *holds. Suppose g/qg' is a bounded function that approaches* 0 *as t approaches* 0 *or* 1. *Suppose $\int_0^\infty q(F)\,dI < \infty$. Then for this particular q we have*

$$\rho_{qg'}(H_n^*, H) \to_p 0 \qquad as \quad n \to \infty ,$$

where $*$ restricts functions on $(0, 1)$ *to* $[1/n, 1 - 1/n]$.

PROOFS. Consider first Theorem 3. Let $Z_n = n^{\frac{1}{2}}(\bar{X} - \eta)$. Then

$$H_n = (Q_n - Z_n g/\eta)/\bar{X} \qquad \text{and} \qquad H = (Q - Zg/\eta)/\eta . \tag{9}$$

Thus

$$\begin{aligned}\rho_{qg'}(H_n^*, H) \leqq\ & \rho_{qg'}(Q_n^*, Q)/\bar{X} + \rho_{qg'}(Q, 0)|1/\bar{X} - 1/\eta| \\ & + \rho_{qg'}(g^*, 0)|Z_n - Z|/\bar{X}\eta + \rho_{qg'}(g^*, 0)|Z|\,|1/\bar{X} - 1/\eta|/\eta \\ & + \rho_{qg'}(g, g^*)|Z|/\eta^2 .\end{aligned}$$

Now $\rho_{qg'}(Q_n^*, Q) = o_p(1)$, $\bar{X} \to_{\text{a.s.}} \eta > 0$, $\rho_{qg'}(Q, 0) = O_p(1)$, $1/\bar{X} \to_{\text{a.s.}} 1/\eta$, $|Z| = O_p(1)$, $\rho_{qg'}(g, 0) < \infty$ and $\rho_{qg'}(g, g^*) \to 0$. To establish that $\rho_{qg'}(H_n^*, H) \to_p 0$, it thus suffices to show that $Z_n \to_p Z$. It was communicated to me by R. Pyke that $Z_n = -n^{\frac{1}{2}} \int_0^\infty (\mathbb{F}_n - F)\,dI = -\int_0^\infty U_n(F)\,dI$. Thus

$$\begin{aligned}|Z_n - Z| &\leqq \rho_{q(F)}(U_n(F), U(F)) \textstyle\int_0^\infty q(F)\,dI \\ &\leqq \rho_q(U_n, U) \textstyle\int_0^\infty q(F)\,dI \to_p 0 .\end{aligned}$$

Consider Theorem 2 next. The hypotheses of Theorem 3 hold; and it is a trivial matter to change the definition of $*$. Note that $\int_0^\infty q(F)\,dI = \int_0^1 [q/(1 - I)]\,dI$ for exponential F. Thus $\rho_{q/(1-I)}(D_n^*, D) \to_p 0$ for all q in $\mathcal{Q}$ for which $\int_0^1 [q/(1 - I)]\,dI < \infty$. But any q in $\mathcal{Q}$ is clearly bounded below by a member of $\mathcal{Q}$ for which this integral is finite. □

COROLLARY 4 (*Convergence of the ordered spacings process in integral metrics*).

Suppose $g = F^{-1}$ is such that the conclusion of Corollary 2 holds. Suppose $\|g\|_\nu < \infty$ and $\|qg'\|_\nu < \infty$ for some q in $\mathscr{Q}$. Then

$$\|H_n{}^* - H\|_\nu \to_p 0 \qquad \textit{as } n \to \infty$$

*where * restricts functions on* $(0, 1)$ *to* $[1/n, 1 - 1/n]$.

PROOF. Simply replace $\rho_{qg'}$ by $\| \ \|_\nu$ in the first inequality in the proof of Theorem 3. □

REMARK 1. It is well known that $a_{ni} \equiv E(D_{ni}) = \sum_1^i (n - j + 1)^{-1}$ for $1 \leqq i \leqq n$ when F is any exponential df. Let

$$(10) \qquad \bar{D}_n = n^{\frac{1}{2}}[-\log(1 - \Gamma_n{}^{-1})/\bar{X} - a_n] \qquad \text{on } [0, 1)$$

where $a_n(t) = a_{ni}$ for $(i - 1)/n < t \leqq i/n$ and $1 \leqq i \leqq n$ and $a_n(0) = 0$. Then $\bar{D}_n$ may replace D_n in Theorem 2.

PROOF. It suffices to show that $\rho_{qg'}(a_n, g) \to 0$ for any q in $\mathscr{Q}$ where $g = -\log(1 - I)$. Now

$$g(i/(n + 1)) = \int_{n-i+1}^{n+1} x^{-1}\, dx \leqq a_{ni} \leqq \int_{n-i+\frac{1}{2}}^{n+\frac{1}{2}} x^{-1}\, dx = g(i/(n + \tfrac{1}{2}))$$

and $g((i - 1)/n) \leqq g(t) \leqq g(i/n)$ for $(i - 1)/n < t \leqq i/n$. Since $(i - 1)/n < i/(n + 1) < i/(n + \frac{1}{2}) < i/n$ we have for $(i - 1)/n < s,\ t \leqq i/n$ that

$$|a_n(t) - g(t)| \leqq |g((i - 1)/n) - g(i/n)| = g'(s)/n \leqq (n - i)^{-1};$$

and thus $(1 - t)|a_n(t) - g(t)| \leqq [1 - (i - 1)/n]/(n - i) \leqq 2/n$ for $1 \leqq i \leqq n - 1$. This yields the claim. □

EXAMPLE 3. In testing for exponentiality, the standard exponential probability plot suggests a statistic whose null distribution may be represented as

$$\sum_1^n (D_{ni}/a_{ni} - 1)^2 = \int_0^1 \bar{D}_n/a_n)^2\, dI .$$

An easy application of Theorem 2 and Remark 1 shows that this statistic is asymptotically distributed as $\int_0^1 (D/g)^2\, dI$.

EXAMPLE 4 (Linear combinations of ordered uniform spacings). Let $F(x) = 1 - e^{-x}$ for $x \geqq 0$ so that $g = -\log(1 - I)$ and consider

$$(11) \qquad T_n = n^{-1} \sum_1^n c_{ni} D_{ni} .$$

We suppose that there exist functions C_n on $(0, 1)$ and a signed measure ν on $(0, 1)$ such that

$$c_{ni}/n = \int_{(i-1)/n}^{i/n} C_n\, d\nu \qquad \text{for } 1 \leqq i \leqq n .$$

Suppose also that

(i) For all n sufficiently large we have $|C_n| \leqq \psi$ a.e. $|\nu|$ where $\int_0^1 [q/(1 - I)]\psi\, d|\nu| < \infty$ for some q in $\mathscr{Q}$,
(ii) $c_{n1} = o(n^{\frac{3}{2}})$,
(iii) $C_n \to C$ a.e. $|\nu|$ as $n \to \infty$ for some function C on $(0, 1)$ and

(iv) $n^{\frac{1}{2}} \int_0^1 (C_n^* - C) g \, d\nu \to 0$ as $n \to \infty$ where $*$ restricts functions on $[0, 1)$ to $[1/n, 1)$.

Then

$$n^{\frac{1}{2}}(T_n - \mu) \to_d N(0, \sigma^2)$$

with $\mu = \int_0^1 Cg \, d\nu$ and $\sigma^2 = \int_0^1 \int_0^1 K_D(s, t) C(s) C(t) \, d\nu(s) \, d\nu(t)$ finite.

PROOF. This is similar to Theorem 2 of [9]. It is a corollary to Theorem 5 below; or an easy consequence of Theorem 2 by writing $n^{\frac{1}{2}}(T_n - \mu) = T_n^* + \gamma_n + \theta_n$ where $T_n^* = \int_0^1 C_n D_n^* \, d\nu$, $\gamma_n = n^{-\frac{1}{2}} c_{n1} D_{n1}$ and $\theta_n = n^{\frac{1}{2}} \int_0^1 (C_n^* - C) g \, d\nu$ and considering $|T_n^* - \int_0^1 CD \, d\nu|$. (As pointed out to me by P. Bickel, asymptotic normality of T_n can also be established by writing T_n as a linear combination of the i.i.d. normalized spacings $|(n - i + 1)(X_{ni} - X_{n,i-1})|$ divided by $\bar{X}$.)

EXAMPLE 4a. Replace c_{ni} in (11) by the a_{ni} of Remark 1. Jackson (1967) proposed this statistic to test for exponentiality. Its null distribution follows easily from Example 4 with $C_n = a_n$, $C = g$ and $d\nu = dI$.

THEOREM 4 (*Convergence of the empirical spacings process*). *Let F be an exponential* df. *Then*

$$\rho(G_n, G) \to_p 0 \qquad \textit{as } n \to \infty .$$

PROOF. Weak convergence of G_n to G for exponential F was proved in Pyke (1965). We write

$$G_n = U_n(F(\bar{X}I)) + A_n I(Z_n - Z) + A_n IZ$$

with $A_n = [F(\bar{X}I) - F(\eta I)]/(\bar{X} - \eta)I$. Thus

$$\rho(G_n, G) \leqq \rho(U_n(F(\bar{X}I)), U(F(\eta I))) + |Z_n - Z| \rho(A_n I, 0) + |Z| \rho(A_n I, f(\eta I) I) .$$

Now $\rho(U_n(F(\bar{X}I)), U(F(\eta I))) \to_{\text{a.s.}} 0$ for any continuous F having positive mean since then $\rho(F(\bar{X}I), F(\eta I)) \to_{\text{a.s.}} 0$. Now $|Z_n - Z| \to_p 0$ whenever $\int_0^\infty q(F) \, dI < \infty$ by the proof of Theorem 3. The mean value theorem easily handles the two terms involving A_n when F is exponential. □

REMARK 2. The conclusion of Theorem 4 holds for many nonexponential F also. The additional conditions needed are $\int_0^\infty q(F) \, dI < \infty$, $\rho(A_n I, 0) = O_p(1)$ and $\rho(A_n I, f(\eta I) I) = o_p(1)$. It seems better to check these last two in particular cases using the mean value theorem than to give unnatural sufficient conditions.

8. Functions of ordered uniform spacings. In this section we prove asymptotic normality of

$$T_n = n^{-1} \textstyle\sum_{i=1}^n c_{ni} h(D_{ni}) \; ; \qquad (12)$$

where h is a fixed known function, the c_{ni}'s form a triangular array of known constants and $D_{ni} = X_{ni}/\bar{X}$ where $0 < X_{ni} < \cdots < X_{nn}$ are the order statistics of a sample $X_1, \cdots, X_n$ of size n from the df $F(x) = 1 - e^{-x}$ for $x \geqq 0$. Let $\mathbb{F}_n$ denote the empirical df of the sample. Let $g = -\log(1 - I) = F^{-1}$.

As in Example 4 and [9], we suppose throughout this section that there exist functions C_n on (0, 1) and a signed measure ν on (0, 1) such that

$$c_{ni}/n = \int_{(i-1)/n}^{i/n} C_n \, d\nu \qquad \text{for } 1 \leqq i \leqq n . \tag{13}$$

Thus

$$T_n = \int_0^1 h(\mathbb{F}_n^{-1}/\bar{X}) C_n \, d\nu . \tag{14}$$

Let C denote a fixed measurable function on (0, 1) and let

$$\mu = \int_0^1 h(g) C \, d\nu \tag{15}$$

and

$$\sigma^2 = \int_0^1 \int_0^1 K_D(s, t) h'(g(s)) h'(g(t)) C(s) C(t) \, d\nu(s) \, d\nu(t) \tag{16}$$

provided these exist. Let * restrict functions on [0, 1) to [1/n, 1). If $C_n{}^*$ replaces C_n in (15) and (16) the resulting quantities will be called μ_n and $\sigma_n{}^2$.

Now

$$n^{\frac{1}{2}}(T_n - \mu) = T_n{}^* + \gamma_n + \theta_n \tag{17}$$

where

$$\begin{aligned} T_n{}^* &= \int D_n{}^* A_n C_n \, d\nu , \\ A_n &= [h(\mathbb{F}_n^{-1}/\bar{X}) - h(g)]/[\mathbb{F}_n^{-1}/\bar{X} - g] , \\ \gamma_n &= n^{-\frac{1}{2}} c_{n1} h(D_{n1}) \end{aligned}$$

and

$$\theta_n = n^{\frac{1}{2}} \int_0^1 (C_n{}^* - C) h(g) \, d\nu.$$

(Left continuity is used to define A_n at the most finite number of points, for each fixed ω, at which it might otherwise be undefined.)

(F1) (i) For all large n we have $|C_n| \leqq \psi$ a.e. $|\nu|$ where $\int_0^1 q|h'(g)g'|\psi \, d|\nu| < \infty$ for some q in $\mathcal{Q}$.

(ii) $\int_0^1 q|(A_n{}^* - h'(g))g'|\psi \, d|\nu| \to_p 0$ as $n \to \infty$, for this same q.

(F2) $\gamma_n \to_p 0$ as $n \to \infty$.

(F3) $C_n \to C$ a.e. $|\nu|$ as $n \to \infty$.

(F4) $n^{\frac{1}{2}} \int_0^1 (C_n{}^* - C) h(g) \, d\nu \to 0$ as $n \to \infty$.

(F1, 2) h has a continuous derivative h' on $(0, \infty)$; and

$$|h^{(i)}| \leqq M_1 I^{r_1 - i} + M_2 I^{-r_2 - i} + M_3 \qquad \text{on} \quad (0, \infty) \ \text{ for } \ i = 0, 1$$

for some $r_1 > 0$, $0 < r_2 < 1$ and $M_1, M_2, M_3 \geqq 0$. Also for all large n we have $|C_n| \leqq \psi$ a.e. $|\nu|$ where

$$\int_0^1 q[M_1|g|^{r_1 - 1} + M_2|g|^{-r_2 - 1} + M_3]|g'|\psi \, d|\nu| < \infty$$

for some q in $\mathcal{Q}$. Finally $c_{n1} = o(n^{\frac{1}{2}+\alpha})$ where

$$\begin{aligned} \alpha &\equiv -r_2 && \text{if } \ M_2 > 0 \\ &\equiv 0 && \text{if } \ M_2 = 0 \ \text{ but } \ M_3 > 0 \\ &\equiv r_1 && \text{if } \ M_2 = M_3 = 0 \ \text{ but } \ M_1 > 0 . \end{aligned}$$

THEOREM 5 (*Asymptotic normality of linear combinations of functions of ordered uniform spacings*). *Condition* (F1, 2) *implies* (F1) *and* (F2). *If* (F1), (F2), (F3) *and* (F4) *hold, then*

$$n^{\frac{1}{2}}(T_n - \mu) \to_d N(0, \sigma^2)$$

with μ *of* (15) *and* σ^2 *of* (16) *finite. If only* (F1) *and* (F2) *hold, then*

$$n^{\frac{1}{2}}(T_n - \mu_n)/\sigma_n \to_d N(0, 1)$$

provided $\liminf_{n\to\infty} \sigma_n^2 > 0$.

PROOF. We will prove that $n^{\frac{1}{2}}(T_n - \mu) \to_d N(0, \sigma^2)$. Let $T = \int_0^1 Dh'(g)C\,d\nu$, which is a $N(0, \sigma^2)$ rv by (F1)(i). Also

$$\begin{aligned}|T_n^* - T| &= |\textstyle\int_0^1 [(D_n^* - D)(A_n^* - h'(g) + h'(g))C_n + D(A_n^* - h'(g))C_n \\ &\quad + Dh'(g)(C_n - C)]\,d\nu| \\ &\leqq \rho_{qg'}(D_n^*, D)[\textstyle\int_0^1 q|(A_n^* - h'(g))g'|\psi\,d|\nu| \\ &\quad + \textstyle\int_0^1 q|h'(g)g'|\psi\,d|\nu|] + \rho_{qg'}(D, 0)\textstyle\int_0^1 q|(A_n^* - h'(g))g'|\psi\,d|\nu| \\ &\quad + \rho_{qg'}(D, 0)\textstyle\int_0^1 q|h'(g)g'|\,|C_n - C|\,d|\nu|\,.\end{aligned}$$

Now $\rho_{qg'}(D_n^*, D) = o_p(1)$, $\rho_{qg'}(D, 0) = O_p(1)$ and $\int_0^1 q|h'(g)g'|\,|C_n - C|\,d|\nu| \to_p 0$ by (F1)(i), (F3) and the dominated convergence theorem. Thus $T_n^* \to_p T$ under (F1) and (F3). Referring to (17), (F2) and (F4) show that $n^{\frac{1}{2}}(T_n - \mu) \to_p T$.

The proof for $n^{\frac{1}{2}}(T_n - \mu_n)/\sigma_n$ is even easier. Note that we can divide by σ_n without destroying $\to_p$ since σ_n is bounded away from 0.

We now show that (F1, 2) implies (F2). As in the basis for Remark 1 we have $E(D_{n1}^r) = E(X_{n1}^r)/E(\bar{X}^r)$ and $E(X_{n1}^r) = E(X^r)/n^r$. Now $2\sum_1^n X_i$ is a Chi-square $(2n)$ rv; so it is easy to check that $E(\bar{X}^r) \to 1$. Thus $E(D_{n1}^r)n^r \to E(X^r)$ for any r such that $E(X^r) < \infty$ (i.e. such that $r > -1$). Thus for some finite M_ε

$$\begin{aligned}P(|\gamma_n| \geqq \varepsilon) &= P(|h(D_{n1})| \geqq \varepsilon n^{\frac{1}{2}}/c_{n1}) \leqq (c_{n1}/\varepsilon n^{\frac{1}{2}})E(|h(D_{n1})|) \\ &\leqq (c_{n1}/\varepsilon n^{\frac{1}{2}})[M_1 E(D_{n1}^{r_1}) + M_2 E(D_{n1}^{-r_2}) + M_3] \\ &\leqq M_\varepsilon c_{n1} n^{-\frac{1}{2}}[M_1 n^{-r_1} + M_2 n^{r_2} + M_3]\,.\end{aligned}$$

We now show that (F1, 2) implies (F1). Now (F1)(i) is trivially true. Since h' exists a.e. $|\nu|$, we have for every fixed ω that the difference quotient A_n converges to $h'(g)$ a.e. $|\nu|$. We would thus like, for every fixed ω, to apply the dominated convergence theorem to claim (F1)(ii). For fixed ω we will in fact be able to bound $\chi(S_{n,\varepsilon})|A_n^*|$ (see Lemma A3 for $S_{n,\varepsilon}$) by $M_\omega R(g)$ for some constant M_ω where $R \equiv M_1 I^{r_1 - 1} + M_2 I^{-r_2 - 1} + M_3$. Thus

$$\chi(S_{n,\varepsilon})\int_0^1 q|A_n^* - h'(g)|\,|g'|\psi\,d|\nu| \to_{\text{a.s.}} 0 \qquad \text{as } n \to \infty$$

follows from the dominated convergence theorem; from which (F1)(ii) easily follows.

It thus remains only to bound $\chi(S_{n,\varepsilon})|A_n^*|$ by $M_\omega R(g)$ for each fixed ω. Now

$$\begin{aligned}|A_n| &= |\textstyle\int_g^{g(\Gamma_n^{-1})/\bar{X}} h'(y)\,dy|/|g(\Gamma_n^{-1})/\bar{X} - g| \\ &\leqq R(g) \vee R(g(\Gamma_n^{-1})/\bar{X}) \leqq M_\omega[R(g) \vee R(g(\Gamma_n^{-1}))]\,.\end{aligned}$$

It thus suffices if we bound $\chi(S_{n,\varepsilon})R(g(\Gamma_n^{-1}))^*$ by $M_\varepsilon R(g)$ for $0 < t \leqq \delta$ and $1 - \delta \leqq t < 1$ for some fixed $\delta > 0$ (bounds for $\delta \leqq t \leqq 1 - \delta$ are trivial). In fact we need only consider $0 < t \leqq \delta$ when Case 1: $M_2 > 0$ or $M_1 > 0$ with $r_1 < 1$; and we need only consider $1 - \delta \leqq t < 1$ when Case 2: $M_1 > 0$ and $r_1 > 1$.

CASE 1. Since $\Gamma_n^{-1} \geqq \beta_\varepsilon I$ on $S_{n\varepsilon}$ by Lemma A3, we have for $0 < t \leqq \delta$ that

$$\chi(S_{n,\varepsilon})R(g(\Gamma_n^{-1}))^* \leqq R(g(\beta_\varepsilon I)) \leqq M_\varepsilon R(g) ;$$

the second step holds since for any given $0 < \beta,\ \delta < 1$ we can find a small constant $M \equiv M_{\beta,\delta} > 0$ for which

$$g(\beta t) \geqq Mg(t) \qquad \text{for } 0 < t \leqq \delta . \tag{18}$$

(For exponential F statement (18) is equivalent to finding a small $M > 0$ such that $1 - \beta t \leqq (1 - t)^M$ for $0 < t \leqq \delta$.)

CASE 2. Since $\Gamma_n^{-1} \leqq 1 - \beta_\varepsilon(1 - I)$ on $S_{n,\varepsilon}$ by Lemma A3, we have for $1 - \delta \leqq t < 1$ that

$$\chi(S_{n,\varepsilon})R(g(\Gamma_n^{-1}))^* \leqq R(g(1 - \beta_\varepsilon(1 - I)) \leqq M_\varepsilon R(g) ;$$

the second step holds since for any given $0 < \beta,\ \delta < 1$ we can find a large constant $M \equiv M_{\beta,\delta} > 0$ for which

$$g(1 - \beta(1 - t)) \leqq Mg(t) \qquad \text{for } 1 - \delta \leqq t < 1 . \tag{19}$$

(For exponential F statement (19) is equivalent to finding a large $M > 0$ such that $\beta(1 - t) \geqq (1 - t)^M$ for $1 - \delta \leqq t < 1$.)

(In Theorem 6 when we consider nonexponential df's, we will assume (18) and (19) in our hypotheses.) □

REMARK 3. If $c_{ni} = 1$ for all i and n, then (12) reduces to

$$T_n = n^{-1} \textstyle\sum_1^n h(D_{ni}) . \tag{20}$$

Such statistics were considered by Darling (1953) and Pyke (1965).

EXAMPLE 5. If $h(x)$ equals x^r for $r > -1$, $|x - 1|^r$ for $r \geqq 1$, $\log x$ or the indicator function of and interval, then T_n of (20) is asymptotically normal by Theorem 5. These include the specific h's yielding asymptotic normality that were considered by Darling (1953). Also note that Theorem 5 allows more general c_{ni}'s in these examples. However, Theorem 5 as applied to (20) itself is not as strong as the theorem contained in the discussion between Kingman and Pyke in [7].

EXAMPLE 5a. If $h(x) = x^{-1}$, then the hypotheses of Theorem 5 fail for (20). (This must necessarily happen, since this rv was shown by Darling (1953) to have a nonnormal limit.)

9. Functions of ordered uniform spacings. In this section we remark on the statistic T_n of (12) for more general F. Let $\gamma_n = n^{-\frac{1}{2}}[c_{n1}h(D_{n1}) + c_{nn}h(D_{nn})]$ and let * restrict functions on (0, 1) to $[1/n, 1 - 1/n]$. Then (17) still holds.

THEOREM 6. *Suppose* (F1, 2), (F2), (F3) and (F4) *hold with the new definitions of* γ_n *and* $*$. *Suppose* g *has a continuous derivative* g' *and* (18) *and* (19) *hold. Suppose* $\rho_{qg'}(H_n^*, H) \to_p 0$ *as* $n \to \infty$ *for some* q *in* $\mathscr{Q}$; *and suppose* g/qg' *is a bounded function. Then*

$$n^{\frac{1}{2}}(T_n - \mu) \to_d N(0, \sigma^2)$$

where $\mu = \int_0^1 h(g)C\, d\nu$ *and*

$$\sigma^2 = \int_0^1 \int_0^1 K_H(s, t)h'(g(s))h'(g(t))C(s)C(t)\, d\nu(s)\, d\nu(t)$$

are finite.

PROOF. Under these conditions the proof of Theorem 5 goes through without change for general g. □

REMARK 4. It is clear that minor variations in this approach allow us to let F depend on n. The key part of a proof is Corollary 3. More general theorems (at a considerable increase in complexity of notation and some increase in complexity of thought) can be based on a version of Corollary 4 using $\|\ \|_\nu$.

REFERENCES

[1] BICKEL, P. (1967). Some contributions to the theory of order statistics. *Proc. Fifth Berkeley Symp. Math. Statist. Prob.* 575–591. Univ. of California Press.

[2] BICKEL, P. (1969). Tests for monotone failure rate, II. *Ann. Math. Statist.* **40** 1250–1260.

[3] BICKEL, P. and DOKSUM, K. (1969). Tests for monotone failure rate based on normalized spacings. *Ann. Math. Statist.* **40** 1216–1235.

[4] CHERNOFF, H., GASTWIRTH, J. and JOHNS, M. (1967). Asymptotic distribution of linear combinations of functions of order statistics with applications to estimation. *Ann. Math. Statist.* **38** 52–72.

[5] DARLING, D. (1953). On a class of problems relating to the random division of an interval. *Ann. Math. Statist.* **24** 239–253.

[6] JACKSON, O. (1967). An analysis of departures from the exponential distribution. *J. Roy. Statist. Soc., Ser. B.* **29** 540–549.

[7] PYKE, R. (1967). Spacings. *J. Roy. Statist. Soc., Ser. B.* **27** 395–449.

[8] PYKE, R. and SHORACK, G. (1968). Weak convergence of a two-sample empirical process and a new approach to Chernoff-Savage theorems. *Ann. Math. Statist.* **39** 755–771.

[9] SHORACK, G. (1972). Functions of order statistics. *Ann. Math. Statist.* **43** 412–427.

DEPARTMENT OF MATHEMATICS
UNIVERSITY OF WASHINGTON
SEATTLE, WASHINGTON 98105