The Annals of Probability
1988, Vol. 16, No. 3, 1372–1388

BOOK REVIEW

GALEN R. SHORACK AND JON A. WELLNER, *Empirical Processes with Applications to Statistics,* Wiley, New York, 1986, xxxvii + 938 pages, \$59.95.

REVIEW BY PETER GAENSSLER

University of Munich

My own interest in empirical processes arose after having read Pyke's [8][1] beautiful survey where he underlines his view that "the development of empirical processes provides an excellent illustration of the interplay between statistics and probability and of the increased sophistication of mathematical techniques which have been introduced into these disciplines in recent years." Since then the theory of empirical processes has grown in an enormous way. This growth is extensively covered by the present book, which confirms Pyke's view in a very impressive way.

Based on random samples of size n, one basic idea of the subject is to build up a stochastic process $\mathbb{X}_n$ such that certain statistics T_n of interest can be represented as functionals $T(\mathbb{X}_n)$. If the probabilistic behavior of $\mathbb{X}_n$ (for finite sample sizes n as well as asymptotically as n tends to infinity) is well tractable using intrinsic properties of $\mathbb{X}_n$, then one can expect the same for $T_n = T(\mathbb{X}_n)$. From this observation it is evident that the richer the class of processes $\mathbb{X}_n$ with interesting inherited properties (e.g., martingale properties), the richer the class of statistics T_n that can be successfully treated this way. As to asymptotic results for T_n (e.g., CTL's), they usually (at least for smooth T) can then be inferred from corresponding results for $\mathbb{X}_n$. This is demonstrated in the present book in a very efficient way based in part of the special Skorokhod construction of various processes as opposed to classic weak convergence results (cf. additional remarks that follow).

Another main feature is the construction of special versions of $\mathbb{X}_n$ and versions $\mathbb{X}'_n$ of the limiting process $\mathbb{X}$ such that the sample paths of both versions are a.s. very close together (wrt the supremum metric $\|\cdot\|$) with a rate of convergence that is good enough to infer, e.g., LIL-type results for $\mathbb{X}_n$ from those for $\mathbb{X}'_n$ or $\mathbb{X}$. These are the famous strong approximation results of the Hungarians Komlós, Major and Tusnády (1975), called in the present book the "Hungarian constructions" (and usually referred to as KMT approximations in the literature). In this context a recent major theorem of Mason and van Zwet [(1985), Theorem 12.3.4] gives a very remarkable refinement of the KMT inequality.

A third basic technique stressed by the authors is the reduction to the case of samples with distribution concentrated on the unit interval, in which case the

Received November 1987

[1]References listed by author and date refer directly to pages 901–921 of the book. References listed by bracketed number appear at the end of this article.

processes under consideration can be viewed as random elements (re's) in Skorokhod's $D[0,1]$ space; cf. the classical result 1.1.3 and the fine reductions described on pages 99 and 102 for some basic processes under general alternatives.

Chapter 3 is in some sense the core of the present book. Starting with the simplest processes $\mathbb{X}_n$, this chapter deals with one of the main themes: the convergence and distributions of empirical processes. Here the special construction (Theorem 3.1.1) is basic. It covers the uniform empirical process $\mathbb{U}_n$ [defined by

$$\mathbb{U}_n(t) \equiv \sqrt{n}\,(\mathbb{G}_n(t) - t) \quad \text{for } 0 \le t \le 1,$$

where $\mathbb{G}_n$ denotes the empirical distribution function (df) based on uniform(0, 1) rv's $\xi_1, \dots, \xi_n$]; the uniform quantile process $\mathbb{V}_n \equiv \sqrt{n}\,(\mathbb{G}_n^{-1} - I)$ (with I denoting the identity function); the weighted uniform empirical process $\mathbb{W}_n$ [defined by

$$\mathbb{W}_n(t) \equiv (c'c)^{-1/2} \sum_{i=1}^{n} c_{ni}\left[1_{[\xi_i \le t]} - t\right] \quad \text{for } 0 \le t \le 1,$$

where $\{c_{n1}, \dots, c_{nn};\ n \ge 1\}$ is a triangular array of known constants with $c'c \equiv \sum_{i=1}^{n} c_{ni}^2$ and where the ξ_i's are again uniform(0, 1) rv's. Note that $\mathbb{W}_n = \mathbb{U}_n$ if $c_{ni} \equiv 1$]; and the empirical rank process $\mathbb{R}_n$ [defined by

$$\mathbb{R}_n(t) \equiv (c'c)^{-1/2} \sum_{i=1}^{\langle (n+1)t \rangle} c_{nD_{ni}},$$

where $D_{n1}, \dots, D_{nn}$ are the antiranks defined by $\xi_{D_{ni}} = \xi_{n:i}$, $\xi_{n:1} \le \cdots \le \xi_{n:n}$ being the order statistics of $\xi_1, \dots, \xi_n$, and $\langle \cdot \rangle$ denotes the greatest integer function]. In addition, the general (reduced) weighted empirical process $\mathbb{Z}_n$ defined by

$$\mathbb{Z}_n(t) \equiv (c'c)^{-1/2} \sum_{i=1}^{n} c_{ni}\left[1_{[\xi_{ni} \le G_{ni}(t)]} - G_{ni}(t)\right] \quad \text{for } 0 \le t \le 1$$

is studied, where now $G_{n1}, \dots, G_{nn}$ are arbitrary df's on $[0,1]$ and where the ξ_{ni}'s arise via the reduction described on page 99. The sequential uniform empirical process

$$\mathbb{K}_n(s,t) \equiv n^{-1/2} \sum_{i=1}^{\langle ns \rangle} \left[1_{[0,t]}(\xi_i) - t\right],$$

based on independent uniform(0, 1) rv's $\xi_1, \dots, \xi_n$ [and restricted to $(s,t) \in [0,1]^2$] is shown [following Bickel and Wichura (1971)] to converge weakly to a Kiefer process $\mathbb{K}$, being a normal process with $E\mathbb{K}(s,t) = 0$ and

$$\operatorname{cov}[\mathbb{K}(s_1,t_1), \mathbb{K}(s_2,t_2)] = (s_1 \wedge s_2)(t_1 \wedge t_2 - t_1 t_2)$$

[correcting (5) on page 30]. The special constructions for the processes $\mathbb{X}_n$ and their corresponding limiting processes $\mathbb{X}$ (e.g., with $\mathbb{X} = \mathbb{U} \equiv$ Brownian bridge in case $\mathbb{X}_n = \mathbb{U}_n$, with $\mathbb{X} = \mathbb{V} \equiv -\mathbb{U}$ in case $\mathbb{X}_n = \mathbb{V}_n$, with $\mathbb{X} = \mathbb{W} \equiv \mathbb{U}$ in case

$\mathbb{X}_n = \mathbb{W}_n$ and with $\mathbb{X} = \mathbb{W}$ in case $\mathbb{X}_n = \mathbb{R}_n$ and $\bar{c}_n \equiv n^{-1}\sum_{i=1}^n c_{ni} = 0$] allow for representations of the limiting rv's $T(\mathbb{X})$ of $T(\mathbb{X}_n)$, which is another appealing feature of the presentation chosen by the authors.

The main ingredient consists in the extensive use of martingale theory associated with various processes mentioned before, as pointed out in Section 6 of Chapter 3, and thus making available powerful martingale inequalities which, together with many of the classic inequalities of probability theory presented in Appendix A, "are the key to strong theorems." This is beautifully worked out in the present book!

Finally, another basic technique consists of the treatment of the empirical process via Poisson methods, thereby carefully defining the link between empirical, quantile and Poisson processes (see Chapter 8). Many first proofs of results presented in other chapters make use of these methods (cf., e.g., the Poisson embeddings in Sections 14.5–14.7).

This extremely rich exposition of the theory of empirical processes and its statistical applications may be of limited use for beginners, but surely it will be an invaluable source for all those already acquainted with some parts of that theory and its fundamental role in nonparametric statistics (cf. Chapter 19–23 for the main statistical applications).

The great extent of all kinds of empirical process theory and statistical applications contained in the present book (as indicated by the list of more than 500 references!) forces the reviewer to confine special detailed comments on each chapter to those parts that he thinks are of greatest importance. Otherwise we will only roughly comment on the other parts. In any case, we will follow the clear introductions to each chapter given by the authors (cf. also the survey of results presented in Chapter 1). Thereby we will use the notation as given in the list of special symbols on page xxxv.

After an introduction and survey of some results in Chapter 1, Chapter 2 is concerned with foundations, special spaces and special processes. Weak convergence ($\Rightarrow$) of a sequence of re's in an arbitrary (usually nonseparable) metric space (M, δ) is defined in the sense of Dudley (1966), being further pursued by Wichura [9] and Gaenssler [(1983), Section 3]; cf. also [3] for a presentation in case $(M, \delta) = (D[0,1], \|\cdot\|)$ and [4] for a straightforward generalization to handle also empirical processes based on multivariate observations up to empirical processes based on random data in arbitrary sample spaces and indexed by certain classes of sets or functions, respectively, as considered in Chapter 26. Other more general approaches to the weak convergence of re's in nonseparable spaces are provided by Hoffmann–Jørgensen's theory (cf. [1] for an illustration).

Donsker's (1951) theorem on the weak convergence of the partial sum processes $\mathbb{S}_n$ to Brownian motion $\mathbb{S}$ is established. Special constructions of $\mathbb{S}_n$ that converge a.s. are summarized briefly, Skorokhod's (1965) embedding is presented and the Hungarian construction is given (without proof) after the Wasserstein distance it uses is discussed. Strassen's (1964) theorems on the relative compactness ($\rightsquigarrow$) of scaled Brownian motion $\mathbb{S}(nI)/\sqrt{n}$ and of the partial sum process $\mathbb{S}_n$ are presented. The Darling–Erdös (1956) theorem is stated for the supremum of normalized Brownian motion $\mathbb{S}(t)/\sqrt{t}$.

In addition to the remarks already made on the content of Chapter 3, let us stress here the following main point: Theorem 3.1.1 presents the fundamental "special construction" of various processes $\mathbb{X}_n$ via the construction of a triangular array of row-independent uniform(0, 1) rv's $\{\xi_{ni}: 1 \le i \le n;\ n \ge 1\}$ on which the processes $\mathbb{X}_n$, like $\mathbb{X}_n = \mathbb{U}_n, \mathbb{V}_n, \mathbb{W}_n$ or $\mathbb{R}_n$, are based. These rv's together with the limiting processes like $\mathbb{U}$, $\mathbb{V}$ and $\mathbb{W}$, are all defined on a common probability space, such that, e.g.,

$$\|\mathbb{U}_n - \mathbb{U}\| \underset{\text{a.s.}}{\to} 0 \quad \text{as } n \to \infty. \tag{1}$$

The proof (given in Section 3) is based on Theorem 3.3.1 (showing weak convergence of $\mathbb{Z}_n$) and the famous Skorokhod–Dudley–Wichura Theorem 2.3.4 which allows for replacing processes that converge weakly by equivalent processes whose sample paths converge a.s. The idea is then to establish additional results for the equivalent processes using the a.s. convergence and then claim these results (if possible) for the original processes. Also, compared with classical weak convergence results, the special Skorokhod constructions pursued in the present book whenever possible yield simpler and more intuitive proofs of results. These constructions are especially important for applications to statistics. Note that from (1) it follows at once that for all T being $\|\cdot\|$-continuous a.s. $\mathbb{U}$ one has $T(\text{Skorokhod's } \mathbb{U}_n) \to_{\text{a.s.}} T(\mathbb{U})$, and since $T(\text{Skorokhod's } \mathbb{U}_n) \cong T(\text{any } \mathbb{U}_n)$, one obtains from (1) that $T(\text{any } \mathbb{U}_n) \to_d T(\mathbb{U})$ for all T that are $\|\cdot\|$-continuous a.s. $\mathbb{U}$.

In Section 1 of Chapter 4 on Alternatives and Processes of Residuals it is shown that the key for the asymptotic behavior of the weighted empirical process $\mathbb{E}_n$, defined by

$$\mathbb{E}_n(x) \equiv (c'c)^{1/2} \sum_{i=1}^{n} c_{ni}\left\{1_{[X_{ni} \le x]} - F(x)\right\}, \qquad -\infty < x < \infty,$$

is constituted by the deterministic term $(c'c)^{-1/2}\sum_{i=1}^{n} c_{ni}[F_{ni} - F]$, provided that the df's $F_{n1}, \dots, F_{nn}$ of the underlying independent rv's $X_{n1}, \dots, X_{nn}$ satisfy the "nearly null-type" condition $\max\{F_{ni} - F: 1 \le i \le n\} \to 0$ as $n \to \infty$. A new sort of contiguity condition is specified [which implies the usual one (cf. Exercise 2 on page 157)] on the basis of which the behavior of the deterministic term is particularly simple, and results in a representation of the limiting process pertaining to $\mathbb{E}_n$ on the probability space of the special construction. Analogous results are shown to hold for the empirical rank process $\mathbb{R}_n$.

In Section 2, following Hájek and Sĭdák [(1967), Section VI.4.5], the authors give an interesting expansion of the local asymptotic power of the $\|(\mathbb{G}_n - I)^+\|$ test for local alternatives. Section 3 describes the asymptotic optimality of the empirical df $\mathbb{F}_n$ as an estimator of the underlying df F. Here Beran's (1977) theorem is proved. The asymptotic minimax theorem of Dvoretzky, Kiefer and Wolfowitz (1956), which shows that $\mathbb{F}_n$ is "asymptotic minimax" for a "supremum type" of loss function, is stated without proof. Finally some other small sample optimality properties of $\mathbb{F}_n$, among which the fact that $\mathbb{F}_n$ is the "nonparametric maximum likelihood estimator" of F, are summarized without proof. As to the latter, the reader may consult the recent paper by Gill [5].

The basic problem that motivates Chapter 5 on Integral Tests for Fit and the Estimated Empirical Process is to determine the distribution of the Cramér–von Mises goodness-of-fit statistic

$$W_n^2 \equiv \int_{-\infty}^{\infty} n[\mathbb{F}_n(x) - F(x)]^2 \, dF(x) = \int_0^1 \mathbb{U}_n^2(t) \, dt.$$

Just as an $n \times n$ covariance matrix Σ can be represented as $\Sigma = \sum_{i=1}^n \lambda_j \gamma_j \gamma_j'$, where the λ_j are eigenvalues and the γ_j are orthonormal eigenvectors of Σ, so too the covariance function K of many processes can be represented as

$$K(s,t) = \sum_{j=1}^{\infty} \lambda_j f_j(s) f_j(t) \tag{2}$$

for functions f_j orthonormal wrt the $\mathscr{L}_2$ metric, i.e., λ_j and f_j are eigenvalues and (orthonormal) eigenfunctions of K defined by the relationship $\int_0^1 f(s)K(s,t)\,ds = \lambda f(t)$ for $0 \le t \le 1$. Let $Z_1^*, Z_2^*, \ldots$ be iid $\mathscr{N}(0,1)$ rv's. Just as $\sum_{j=1}^n \sqrt{\lambda_j} Z_j^* \gamma_j$ is a $\mathscr{N}(0, \Sigma)$ random vector, so too the process $\mathbb{X}$, defined by

$$\mathbb{X}(t) = \sum_{j=1}^{\infty} \sqrt{\lambda_j} Z_j^* f_j(t), \tag{3}$$

is a normal process with mean-value function 0 and covariance function $K = K_{\mathbb{X}}$. Integrating (3) we see that

$$\int_0^1 \mathbb{X}^2(t)\,dt = \sum_{j=1}^{\infty} \lambda_j Z_j^{*2} \tag{4}$$

is distributed as a weighted infinite sum of independent chi-square (1) rv's. This heuristic treatment can be made rigorous for fairly general processes $\mathbb{X}$. This general theory is applied to obtain the Kac and Siegert (1947) decomposition (3) of $\mathbb{X} = \mathbb{U}$ and the Durbin and Knott (1972) decomposition of $\mathbb{U}_n$. Then, just as (4) follows from (3), the distributions of W_n^2 and its limiting form $W^2 \equiv \int_0^1 [\mathbb{U}(t)]^2\,dt$ follow. The determination of the distribution of rv's such as W^2 began the application to statistics of the type of methodology that is presented in Section 5. Sections 1–4 provide the theory for using integral tests of fit like W_n^2 to test whether or not a sample comes from a population with completely specified continuous df F. In Section 5 the df F is allowed to depend on a parameter θ whose value is unknown. The natural extension of W_n^2 is then

$$\hat{W}_n^2 \equiv \int_{-\infty}^{\infty} n[F_n(x) - F_{\hat\theta}(x)]^2 \, dF_{\hat\theta}(x) = \int_0^1 \hat{\mathbb{U}}_n^2(t)\,dt$$

for some estimator $\hat\theta$ of θ, $\hat{\mathbb{U}}_n$ being an appropriate estimated empirical process. Following Darling (1955) it is shown that the natural limit of the $\hat{\mathbb{U}}_n$ process is a process of the form (when θ is one dimensional) $\hat{\mathbb{U}} = \mathbb{U} + Zg$, where g is a known function and the rv Z arises as the limit of rv's Z_n of the type $\int_0^1 h\,d\mathbb{U}_n$. Results on convergence of $\hat{\mathbb{U}}_n$ to $\hat{\mathbb{U}}$ are summarized for location, scale and location–scale cases. Section 6 considers the distribution of the natural limit $\hat{W}^2 = \int_0^1 \hat{\mathbb{U}}^2(t)\,dt$ of $\hat{W}_n^2$. In Section 9 it is assumed that $F(\cdot - \theta)$ is the correct df for some value of θ and then θ is estimated by the value $\hat{\theta}_n$ that minimizes

$$W_n^2 = \int_{-\infty}^{\infty} n[\mathbb{F}_n(x) - F(x-\theta)]^2 \, dF(x-\theta).$$

Following Pyke (1970), an improvement of Blackman's (1955) theorem on the asymptotic distribution of $\hat{\theta}_n$ is presented.

Chapter 6 on Martingale Methods starts with a nice heuristic discussion of counting processes and martingale theory associated with the work of Aalen (1976), Aalen and Johansen (1978), Gill (1980, 1983), Rebolledo (1980) and Khmaladze (1981). Conditions under which the heuristics are actually true are given in Appendix B, which itself provides a useful introduction to basic facts on martingale theory in connection with counting processes and stochastic integrals. Most of the results in Appendix B, which also covers important martingale inequalities and Rebolledo's martingale CLT, are stated without proof but the authors took great care to give sufficient references. Given iid rv's $X_{n1}, \ldots, X_{nn}$ with arbitrary df F, Section 1 introduces the cumulative hazard function Λ associated with F and defines the basic martingale $(\mathbb{M}_n(x))$, $-\infty < x < \infty$, by

$$\mathbb{M}_n(x) \equiv \sqrt{n}\left[\mathbb{F}_n(x) - \int_{-\infty}^{\infty} \frac{1-\mathbb{F}_{n-}}{1-F_-}\, dF\right] = \mathbb{U}_n(F(x)) + \int_{-\infty}^{x} \mathbb{U}_n(F_-)\, d\Lambda,$$

where h_- denotes the left-continuous version of h. The natural limiting process to associate with $\mathbb{M}_n$ is

$$\mathbb{M}(x) \equiv \mathbb{U}(F(x)) + \int_{-\infty}^{x} \mathbb{U}(F_-) d\Lambda,$$

where

$$\mathrm{Cov}[\mathbb{M}_n(x), \mathbb{M}_n(y)] = \mathrm{Cov}[\mathbb{M}(x), \mathbb{M}(y)] = V(x \wedge y)$$

with $V(x) \equiv \int_{-\infty}^{x}(1-\Delta\Lambda)\, dF$ $(\Delta\Lambda \equiv \Lambda - \Lambda_-)$. It is then shown that

$$\|\mathbb{M}_n - \mathbb{M}\|_{-\infty}^{\infty} \underset{\text{a.s.}}{\to} 0$$

for the special construction; moreover $\mathbb{M} = \mathbb{S}(V)$ for an appropriate Brownian motion $\mathbb{S}$. Next the predictable variation process $\langle \mathbb{M}_n \rangle$ of $\mathbb{M}_n$ is

$$\langle \mathbb{M}_n \rangle(x) \equiv \int_{-\infty}^{x} (1-\mathbb{F}_{n-})(1-\Delta\Lambda)\, d\Lambda$$

(i.e., $\mathbb{M}_n^2 - \langle \mathbb{M}_n \rangle$ is a 0 mean martingale). Finally extensions to the weighted case of $\mathbb{M}_n$ are considered. In Section 2 convergence in $\|\cdot\psi\|$ metric for U-shaped $\psi \in \mathscr{L}_2(V)$ is treated for the processes $\mathbb{U}_n(F)$ and $\mathbb{W}_n(F)$. Based on the weighted basic martingale $\mathbb{M}_n$, Section 3 deals with the process $\mathbb{K}_n(x) \equiv \int_{-\infty}^{x} h\, d\mathbb{M}_n$ for $-\infty < x < \infty$, where $h \in \mathscr{L}_2(V)$. It is shown that $\mathbb{K}_n$ is a 0 mean martingale and its predictable variation is evaluated. Then the existence of $\mathbb{K}(x) \equiv \int_{-\infty}^{x} h\, d\mathbb{M}$ is proved and convergence of $\mathbb{K}_n$ to $\mathbb{K}$ in $\|\cdot\psi\|$ metrics is studied. Similarly, Section 4 is concerned with processes of the form $\int_{-\infty}^{\cdot} h\, d\mathbb{U}_n(F)$ and $\int_{-\infty}^{\cdot} h\, d\mathbb{W}_n(F)$ for $h \in \mathscr{L}_2(F)$ and its limiting processes $\int_{-\infty}^{\cdot} h\, d\mathbb{U}(F)$ and $\int_{-\infty}^{\cdot} h\, d\mathbb{W}(F)$. Again convergence in $\|\cdot\psi\|$ metrics is proved based on two powerful inequalities (which in turn rely on the evaluated martingale structures, whence, as in many cases before, the Birnbaum–Marshall inequality can be used).

Chapter 7 is concerned with Censored Data and the Product-Limit Estimator. Given iid nonnegative rv's $X_1, \dots, X_n$ ("survival times") with arbitrary df F on $[0, \infty)$ (assumed to be nondegenerate) and iid rv's $Y_1, \dots, Y_n$ ("censoring times") with arbitrary df G and independent of the X's, suppose that the only observable variables are $Z_i \equiv X_i \wedge Y_i$ and $\delta_i \equiv 1_{[X_i \le Y_i]}$ for $i = 1, \dots, n$, on the basis of which one wants to get a reasonable estimator of F. In this so-called random censorship model the Kaplan–Meier (1958) product-limit estimator $\hat{\mathbb{F}}_n$ of F plays the same fundamental role as the empirical df $\mathbb{F}_n$ does in former chapters in case of no censoring (i.e., in case $G \equiv 0$). The main object is to present results for $\hat{\mathbb{F}}_n$ and the corresponding empirical process $\mathbb{X}_n \equiv \sqrt{n}(\hat{\mathbb{F}}_n - F)$ on $[0, \infty)$, where a major novelty in the present approach consists of the efficient usage of martingale theory of counting processes as summarized in Appendix B. Besides $\hat{\mathbb{F}}_n$, the empirical cumulative hazard function $\hat{\Lambda}_n$, serving as an estimator for the cumulative hazard function Λ associated with F, plays a fundamental role through the study of the empirical cumulative hazard process $\mathbb{B}_n \equiv \sqrt{n}(\hat{\Lambda}_n - \Lambda)$ on $[0, \infty)$. The study of this process together with the process $\mathbb{X}_n/(1 - F) \equiv \sqrt{n}(\hat{\mathbb{F}}_n - F)/(1 - F)$ facilitates the study of $\mathbb{X}_n$, since it is shown that both $\mathbb{B}_n$ and $\mathbb{X}_n/(1 - F)$ can be represented as stochastic integrals wrt some basic martingale $\mathbb{M}_n$. By Theorem 7.1.1, there exists a special construction of both the rv's $X_{n1}, Y_{n1}, \dots, X_{nn}, Y_{nn}$ and the Brownian motion $\mathbb{S}$ such that $\|\mathbb{M}_n - \mathbb{S}(V)\|_0^\infty \to_{\text{a.s.}} 0$ as $n \to \infty$, where $V(t) \equiv \int_0^t (1 - G_-)(1 - \Delta\Lambda)\, dF$. Section 3 establishes strong consistency of $\hat{\mathbb{F}}_n$ and $\hat{\Lambda}_n$, i.e., that, e.g.,

$$(5) \qquad \sup_{0 \le t \le \tau} \left| \hat{\mathbb{F}}_n(t) - F(t) \right| \to_{\text{a.s.}} 0 \quad \text{as } n \to \infty,$$

where $\tau \equiv H^{-1}(1)$ and $1 - H \equiv (1 - F)(1 - G)$. At the end of the proof on page 305, after having established that for any $\theta < \tau$,

$$(6) \qquad \sup_{0 \le t \le \theta} \left| \hat{\mathbb{F}}_n(t) - F(t) \right| \to_{\text{a.s.}} 0 \quad \text{as } n \to \infty,$$

the argument that (6) implies (5) seems to be incomplete. Section 4 is concerned with the asymptotic behavior of the processes $\mathbb{B}_n$ and $\mathbb{X}_n$ with limiting processes $\mathbb{B}$ and $\mathbb{X}$ being both representable as stochastic integrals wrt $\mathbb{M} \equiv \mathbb{S}(V)$. It is shown that for the special construction of Theorem 7.1.1, $\|\mathbb{B}_n - \mathbb{B}\|_0^\theta \to_{\text{a.s.}} 0$ and $\|\mathbb{X}_n - \mathbb{X}\|_0^\theta \to_{\text{a.s.}} 0$ as $n \to \infty$ for any $\theta < \tau$.

In Chapter 8 on Poisson and Exponential Representations one finds first the well known representations of uniform order statistics in terms of exponential rv's and as waiting times of a conditional Poisson process. Section 3 establishes representations of uniform quantile processes as exploited by Breiman (1968), and Section 4 is concerned with three different representations (conditional, Chibisov and Kac representations) of the uniform empirical process $\mathbb{U}_n$ showing that many of its properties are inherited from properties of a Poisson process. The representations of $\mathbb{U}_n$ in terms of Poisson processes have the proper distributions for each fixed n, but incorrect distributions when viewed jointly in n. Thus they are unsuitable for proving strong limit theorems involving $\to_{\text{a.s.}}$

type results. Section 5 overcomes this difficulty by use of a two-dimensional Poisson process and Poisson bridge, respectively, to provide a representation of $\mathbb{U}_n$ that has correct distributions when viewed jointly in t and n.

To describe the content of Chapter 9 on Some Exact Distributions, we may again follow closely the author's introduction. In Section 1 formulas are determined for the probability that the uniform empirical df $\mathbb{G}_n$ crosses an arbitrary line in order to obtain the exact distribution of $\|(\mathbb{G}_n - I)^{\pm}\|$ and also the powerful Dvoretzky, Kiefer and Wolfowitz (DKW) inequality. More generally, for general functions g and h recursion formulas for the probabilities $P(g \leq \mathbb{G}_n \leq h$ on $[0,1])$ are evaluated and used to obtain the exact distribution of $\|\mathbb{G}_n - I\|$. There are three types of methods used in this chapter: the analytical, the combinatorial and the Poisson representation methods. Using the analytical method based on exact binomial and uniform calculations, Dempster's (1959) key formula for the probablility that $\mathbb{G}_n$ crosses a line for the first time at height i/n is derived. The analytical method is used throughout Sections 1–3. The combinatorial method is used to rederive Dempster's key formula and additional applications are given. Dwass's (1974) beautiful approach to $\mathbb{G}_n$ based on Poisson processes is also given. Concerning the Poisson representation method, an extension of the Dwass approach using purely probabilistic tools is contained in recent papers by Gaenssler and Gutjahr [2] and Gutjahr and Haeusler [6].

Chapter 10 is concerned with Linear and Nearly Linear Bounds on the Empirical Distribution Function $\mathbb{G}_n$. Section 1 examines from an a.s. point of view how small and how large for fixed $k \geq 1$ the uniform order statistics $\xi_{n:k}$ can be: Two fundamental results of Kiefer (1972) and Robbins and Siegmund (1972) are presented. In Section 2 Lai's (1974) SLLN for $\|(\mathbb{G}_n - I)\psi\|$ is proved. Inequalities for the distributions of $\|\mathbb{G}_n/I\|$ and $\|I/\mathbb{G}_n\|_{\xi_{n:1}}^{1}$ are the content of Section 3. Based on the results of Section 9.1, Section 4 provides "in probability linear bounds" on $\mathbb{G}_n$, which means that for any $\varepsilon > 0$ there exists a constant M_ε such that both $\mathbb{G}_n(t) \leq M_\varepsilon t$ for all t and $\mathbb{G}_n(t) \geq t/M_\varepsilon$ for all $t \geq \xi_{n:1}$ occur with probability exceeding $1 - \varepsilon$. On the other hand, the results of Section 1 show that a.s. linear bounds on $\mathbb{G}_n$ do not exist. So the authors describe three ways around this: The first is to allow the slope of the lines to depend on n; the second approach is to use for the upper (lower) bound a "nearly linear" function that has infinite (zero) slope at $t = 0$, an approach which turns out to be useful in the establishment of a SLLN and a LIL for linear combinations of order statistics. The third approach is to truncate off near zero, thus considering $\|\mathbb{G}_n/I\|_{a_n}^1$, $\|I/\mathbb{G}_n\|_{a_n}^1$, $\|\mathbb{G}_n^{-1}/I\|_{a_n}^1$ and $\|I/\mathbb{G}_n^{-1}\|_{a_n}^1$ as $a_n \downarrow 0$.

Chapter 11 deals with Exponential Inequalities and $\|\cdot/q\|$-Metric Convergence of $\mathbb{U}_n$ and $\mathbb{V}_n$. Treatment of goodness-of-fit statistics often reduces to the consideration of functionals of the form $\int_0^1 \mathbb{U}_n(t)\psi(t)\,dt$, say. Based on the fact that for the special construction one has

$$\|\mathbb{U}_n - \mathbb{U}\| \underset{\text{a.s.}}{\rightarrow} 0, \tag{1}$$

it is tempting to write

$$\left| \int_0^1 \mathbb{U}_n(t)\psi(t)\,dt - \int_0^1 \mathbb{U}(t)\psi(t)\,dt \right| \leq \|(\mathbb{U}_n - \mathbb{U})/q\| \int_0^1 q(t)\psi(t)dt.$$

Then, if $\int_0^1 q\psi\, dt < \infty$ and if (1) can be strengthened to $\mathbb{U}_n \to_{\text{a.s.}} \mathbb{U}$ in the $\|\cdot/q\|$ metric in the sense that $\|(\mathbb{U}_n - \mathbb{U})/q\| \to_{\text{a.s.}} 0$, one can conclude that

$$\int_0^1 \mathbb{U}_n(t)\psi(t)\, dt \to_d \int_0^1 \mathbb{U}(t)\psi(t)\, dt.$$

Proving in probability convergence wrt $\|\cdot/q\|$ metrics leads naturally (as shown in the proof of Theorem 3.7.1) to application of the Pyke–Shorack inequality according to which, for suitable q functions,

$$\lambda^2 P\left(\|\mathbb{U}_n/q\|_0^\theta \ge \lambda\right) \le \int_0^\theta [q(t)]^{-2}\, dt \quad \text{for all } \lambda > 0.$$

Chibisov (1964) proved that

$$\int_0^1 [q(t)]^{-2}\, dt < \infty \tag{7}$$

implies

$$\|(\mathbb{U}_n - \mathbb{U})/q\| \to_p 0 \quad \text{as } n \to \infty. \tag{8}$$

This was reexamined by O'Reilly (1974): For symmetric (around $t = \frac{1}{2}$) functions q, which satisfy both q increasing on $(0, \frac{1}{2}]$ and $t^{-1/2}q(t)$ decreasing on $(0, \frac{1}{2}]$, the Chibisov–O'Reilly theorem refines "(7) implies (8)" to the assertion that (8) holds for the special construction if and only if $\lim_{t\to\infty} q(t)/\sqrt{t \log_2(1/t)} = \infty$. The author's approach in the present chapter for proving $\|\cdot/q\|$ convergence results for $\mathbb{U}_n$ and $\mathbb{V}_n$ is based on good exponential bounds for binomial rv's and for uniform order statistics. These exponential bounds are then extended to neighborhoods of zero for $\mathbb{U}_n$ and $\mathbb{V}_n$, yielding probability inequalities for $\|\mathbb{U}_n^{\#}/q\|_a^b$ and $\|\mathbb{V}_n^{\#}/q\|_a^b$, respectively [where $0 \le a < (1-\delta)b < b \le \delta \le \frac{1}{2}$, and where $f^{\#}$ denotes, simultaneously, anyone of f^+, f^- or $|f|$], on which the proofs in Section 5 of weak convergence of $\mathbb{U}_n$ and $\mathbb{V}_n$ in $\|\cdot/q\|$ metrics are based.

The Hungarian constructions of $\mathbb{K}_n$, $\mathbb{U}_n$ and $\mathbb{V}_n$ are contained in Chapter 12. Let us comment on this by first comparing it with Skorokhod's (1956) construction, which provided the special construction of a triangular array $\{\xi_{ni}, 1 \le i \le n; n \ge 1\}$ of row-independent uniform(0, 1) rv's and a Brownian bridge $\mathbb{U}$, all on a common probability space such that, e.g., for the uniform empirical process $\mathbb{U}_n$ based on $\xi_{n1}, \dots, \xi_{nn}$ ($\equiv$ Skorokhod's $\mathbb{U}_n$) one has

$$\|\mathbb{U}_n - \mathbb{U}\| \underset{\text{a.s.}}{\to} 0 \quad \text{as } n \to \infty. \tag{1}$$

Given that Skorokhod's $\mathbb{U}_n$ is based on a triangular array, one knows nothing about the joint distribution of Skorokhod's $(\mathbb{U}_1, \mathbb{U}_2, \dots)$. Thus this construction can only be used to infer from (1) that

$$T(\text{any } \mathbb{U}_n) \to_d T(\mathbb{U}) \tag{9}$$

[or $T(\text{any } \mathbb{U}_n) \to_p T(\mathbb{U})$] for all T being $\|\cdot\|$ continuous a.s. $\mathbb{U}$, but not to infer $\to_{\text{a.s.}}$ convergence. The Hungarian construction [begun in Csörgő and Révész (1975a) and fundamentally strengthened by Komlós, Major and Tusnády (1975)] improves Skorokhod's construction in that it only uses a single sequence $(\xi_i)_{i\ge1}$

of uniform(0, 1) rv's and a Kiefer process on a common probability space that satisfy

$$\limsup_{n\to\infty} \frac{\sqrt{n}}{(\log n)^2} \|\mathbb{U}_n - \mathbb{B}_n\| \le \text{some } M < \infty \quad \text{a.s.} \tag{10}$$

Here $\mathbb{U}_n$ is the uniform empirical process based on $\xi_1, \dots, \xi_n$ linked with the corresponding sequential uniform process $\mathbb{K}_n$ through $\mathbb{K}_n(k/n, \cdot) = \sqrt{k/n}\,\mathbb{U}_k$ for $1 \le k \le n$; $\mathbb{B}_n \equiv \mathbb{K}(n, \cdot)/\sqrt{n}$ is a Brownian bridge. Since $T(\mathbb{B}_n) \cong T(\mathbb{U})$, this construction also yields (9). But it is also capable of yielding $\to_{\text{a.s.}}$ results though the subscript n on $\mathbb{B}_n$ may make the problem difficult. Its real value is in the rate it establishes. In Section 1 the Hungarian construction of $\mathbb{K}_n$ leading to (10) is summarized without giving a proof establishing the basic construction. Section 2 shows how a sequence of uniform quantile processes $\mathbb{V}_n$ that converge at a rate can be constructed from the partial-sum process of independent exponential rv's. The basic ideas here are due to Breiman (1968) and Brillinger (1969). Although use of Skorokhod embedding gives an a.s. rate of nearly $O(n^{-1/4})$, it is also possible to use the Hungarian construction (as presented in Section 2.7) to obtain an a.s. rate of $O((\log n)/\sqrt{n})$. The constructions and approximation rates discussed in Sections 1 and 2 all concern the supremum metric $\|\cdot\|$. In Section 3 yet another construction of $\mathbb{U}_n$ and $\mathbb{V}_n$ is summarized, which pays close attention to the behavior of these processes near 0 and 1: It is due to Csörgő, Horvath and Mason (1984a): This refined construction is based on the same partial-sum approximation as in Section 2, but particular care is taken in treating the approximation error near 0 and 1. The result is a construction that is suited to the sharper metrics $\|\cdot/q\|$, where $q(t) = [t(1-t)]^{1/2-\nu}$ with $0 \le \nu < \frac{1}{2}$. Finally the aforementioned major theorem of Mason and van Zwet (1985) is presented stating that there exists a sequence of independent uniform(0, 1) rv's $\xi_1, \xi_2, \dots$ and a sequence of Brownian bridge processes $\mathbb{B}_n$ on a common probability space such that for universal positive constants c_i,

$$P\left(\sup_{0\le t\le d/n} |\mathbb{U}_n(t) - \mathbb{B}_n(t)| \ge (c_1 \log d + x)/\sqrt{n}\right) \le c_2 \exp(-c_3 x)$$

$$\text{for all } -\infty < x < \infty \text{ and } 1 \le d \le n,$$

with the same inequality holding for the supremum taken over $1 - d/n \le t \le 1$. Thus

$$\limsup_{n\to\infty} \frac{\sqrt{n}}{\log n} \|\mathbb{U}_n - \mathbb{B}_n\| \le \text{some } M < \infty \quad \text{a.s.}$$

In Chapter 13 on Laws of the Iterated Logarithm Associated with $\mathbb{U}_n$ and $\mathbb{V}_n$, at first Smirnow's (1944) LIL is established, according to which

$$\limsup_{n\to\infty} \frac{\|\mathbb{U}^{\#}\|}{b_n} = \limsup_{n\to\infty} \frac{\|n(\mathbb{G}_n - I)^{\#}\|}{\sqrt{n}\, b_n} = \frac{1}{2} \quad \text{a.s.,}$$

where $b_n \equiv \sqrt{2 \log_2 n}$. Recall that $f^{\#}$ denotes, simultaneously, any one of f^+, f^-

or $|f|$. This result is then strengthened to Chung's (1949) characterization of upper class sequences $\lambda_n \uparrow$:

$$P(\|\mathbb{U}_n\| \geq \lambda_n \text{ i.o.}) = \begin{cases} 0 \\ 1 \end{cases} \text{according as } \sum_{n=1}^{\infty} \frac{\lambda_n^2}{n} \exp(-2\lambda_n^2) \begin{cases} < \infty \\ = \infty \end{cases}.$$

Also an estimate is given for the order of magnitude of the probability that $\|\mathbb{U}_n\|$ ever crosses an $\uparrow$ barrier. The proofs are based on maximal inequalities for $\|\mathbb{U}_n/q\|$ [due to James (1975) and Shorack (1980)] in conjunction with the DKW inequality 9.2.1. Section 3 contains Finkelsteins's (1971) theorem according to which $\mathbb{U}_n/b_n \rightsquigarrow \mathscr{H}$ a.s. wrt $\|\cdot\|$ on $D[0,1]$ as $n \to \infty$, where again $b_n \equiv \sqrt{2\log_2 n}$ and $\mathscr{H} \equiv \{h\colon h$ is absolutely continuous on $[0,1]$ with $h(0) = h(1) = 0$ and $\int_0^1 [h(t)]^2\,dt \leq 1\}$. Also Cassel's (1951) theorem is proved: It shows that, given any $\varepsilon > 0$, for a.e. ω there exists an $N_{\varepsilon,\omega}$ such that for all $n \geq N_{\varepsilon,\omega}$ the increments of $\mathbb{U}_n$ satisfy

$$|\mathbb{U}_n((s,t])/b_n| \leq \sqrt{(t-s)(1-(t-s))} + \varepsilon \quad \text{for all } 0 \leq s < t \leq 1.$$

Combining Finkelstein's theorem with the previous Hungarian construction (10) yields

$$\mathbb{K}(n,\cdot)/\sqrt{n}\, b_n \rightsquigarrow \mathscr{H} \text{ a.s. wrt } \|\cdot\| \quad \text{on } D[0,1] \text{ as } n \to \infty.$$

Section 4 extends these results to $\|\cdot/q\|$ metrics by giving James's (1975) characterization of those q functions for which

$$\mathbb{U}_n/(qb_n) \rightsquigarrow \mathscr{H}_q \equiv \{h/q\colon h \in \mathscr{H}\} \quad \text{a.s. wrt } \|\cdot\| \text{ on } D[0,1].$$

Chapter 14 studies Oscillations of the Empirical Process. For this, let $\mathbb{X}$ denote a stochastic process viewed as a re taking values in $D[0,1]$: If C denotes an interval $(s,t] \subset [0,1]$, let $|C| \equiv t - s$ and $\mathbb{X}(C) \equiv \mathbb{X}(t) - \mathbb{X}(s)$. The modulus $\omega_{\mathbb{X}}$ of continuity of $\mathbb{X}$ is defined by $\omega_{\mathbb{X}}(a) \equiv \sup_{|C| \leq a} |\mathbb{X}(C)|$. Set $\bar{\omega}_{\mathbb{X}}(a) \equiv \sup_{|C| = a} |\mathbb{X}(C)|$ and let the Lipschitz $\frac{1}{2}$ modulus $\tilde{\omega}_{\mathbb{X}}$ be defined by $\tilde{\omega}_{\mathbb{X}}(a) \equiv \sup_{a \leq |C| 1} |\mathbb{X}(C)|/\sqrt{|C|}$. If $\mathbb{X} = \mathbb{U}$, let $\omega \equiv \omega_{\mathbb{U}}$, $\bar{\omega} \equiv \bar{\omega}_{\mathbb{U}}$ and $\tilde{\omega} \equiv \tilde{\omega}_{\mathbb{U}}$. Theorem 14.1.1 establishes Lévy's (1937) classical theorem according to which

$$\lim_{a \downarrow 0} \frac{\omega(a)}{\sqrt{2a \log(1/a)}} = 1 \quad \text{a.s.} \tag{11}$$

Theorems 14.1.2 and 14.1.4, respectively, prove (11) with ω replaced by $\bar{\omega}$ and $\tilde{\omega}$, respectively. Section 2 considers analogous results for $\mathbb{U}_n$ [Stute's (1982) theorem]. The results of Mason et al. (1983) on the a.s. limiting behavior of $\omega_{\mathbb{U}_n}^{\#}(a_n)/[2a_n \log(1/a_n)]^{1/2}$ on the "boundary sequences" of Stute's theorem are also presented together with extensions concerning

$$\omega_n^{\#}(a_n) \equiv \sup_{a_n \leq |C| \leq 1} (\mathbb{U}_n^{\#}(C))/\sqrt{|C|}.$$

There are also two alternative approaches of importance to these same theorems based on the Hungarian construction and Poisson embedding which are carefully developed in Section 4 and in Sections 5–7, respectively.

The main concern of Chapter 15 is the Empirical Difference Process

$$\mathbb{D}_n \equiv \mathbb{U}_n + \mathbb{V}_n = \mathbb{U}_n - \mathbb{U}_n(\mathbb{G}_n^{-1}) + \sqrt{n}\left(\mathbb{G}_n \circ \mathbb{G}_n^{-1} - I\right).$$

The study of $\mathbb{D}_n(t)$ for fixed t was introduced by Bahadur (1966). Kiefer (1967, 1972) obtained the results for the process $\mathbb{D}_n$ presented here, the main result being that (with $b_n \equiv \sqrt{2\log_2 n}$),

$$\limsup_{n\to\infty} \frac{n^{1/4}\|\mathbb{D}_n\|}{\sqrt{b_n \log n}} = \frac{1}{\sqrt{2}} \quad \text{a.s.}$$

This allows functionals of $\mathbb{V}_n$ to be treated as functionals of $\mathbb{U}_n$.

Chapter 16 is concerned with the Normalized Uniform Empirical Process $\mathbb{Z}_n$ and the Normalized Uniform Quantile Process, where $\mathbb{Z}_n$ is defined by

$$\mathbb{Z}_n(t) \equiv \mathbb{U}_n(t)/\sqrt{t(1-t)} \quad \text{for } 0 < t < 1$$

(having mean zero and variance 1). From Chibisov's Theorem 11.5.1 it follows that $\Rightarrow$ fails for $\mathbb{Z}_n$ and from James's Theorem 13.4.1 that $\rightsquigarrow$ fails for $\mathbb{Z}_n/\sqrt{2\log_2 n}$; but in both cases the function $t \to \sqrt{t(1-t)}$ "just missed." This chapter considers the rate at which $\mathbb{Z}_n$ blows up: Let E_v denote the df of the extreme value distribution, i.e.,

$$E_v(x) \equiv \exp(-\exp(-x)) \quad \text{for } -\infty < x < \infty,$$

and let $b_n \equiv \sqrt{2\log_2 n}$ and $c_n \equiv 2\log_2 n + 2^{-1}\log_3 n - 2^{-1}\log 4\pi$. Then the main result established by Jaeschke (1979) and proved in Section 1 states that

$$b_n\|\mathbb{Z}_n^{\pm}\| - c_n \to_d E_v^2 \quad \text{and} \quad b_n\|\mathbb{Z}_n\| - c_n \to_d E_v^4 \quad \text{as } n \to \infty.$$

A reasonable "studentized version" of $\|\mathbb{Z}_n^+\|$ considered by Eicker (1979) is also presented. Section 2 exhibits the a.s. rate of divergence of $\|\mathbb{Z}_n\|_0^{1/2}$ and Section 3 the a.s. behavior of $\|\mathbb{Z}_n\|_{a_n}^{1/2}$ with $a_n \downarrow 0$ presenting a main result of Csáki (1977). Section 4 considers the normalized quantile process showing that [according to Csörgő and Révész (1978a)]

$$\limsup_{n\to\infty} \left\|\mathbb{V}_n/\sqrt{I(1-I)}\right\|_{a_n}^{1-a_n}/b_n \le 2 \quad \text{a.s.,}$$

where $a_n \equiv 9(\log_2 n)/n$.

Chapter 17 on the Uniform Empirical Process Indexed by Intervals and Functions deals with $\mathbb{U}_n$ "indexed" by a collection $\mathscr{F}$ of functions, i.e., with

$$\mathbb{U}_n(f) \equiv \int_0^1 f\, d\mathbb{U}_n = n^{-1/2}\sum_{i=1}^n \left[f(\xi_i) - \int_0^1 f(t)\, dt\right]$$

for $f \in \mathscr{F}$. Instead of $\mathscr{F} = \{1_{[0,t]}\colon 0 \le t \le 1\}$, now $\mathscr{F} = \mathscr{C} \equiv \{1_C\colon C = (s,t], 0 \le s < t \le 1\}$ and $\mathscr{F} = \mathscr{C}(a,b) \equiv \{C \in \mathscr{C}\colon a \le |C| \le b\}$ are considered as index sets for $\mathbb{U}_n$. Section 1 parallels Section 11.2, developing inequalities for $\|\mathbb{U}_n/q\|_{\mathscr{C}(a,b)} \equiv \sup\{|\mathbb{U}_n(C)|/q(|C|)\colon C \in \mathscr{C}(a,b)\}$ and Section 2 is concerned with weak convergence of $\mathbb{U}_n$ in $\|\cdot/q\|_{\mathscr{C}(a_n,1)}$ metrics with $a_n \equiv (\varepsilon\log n)/n$, $\varepsilon > 0$. Indexing by $\mathscr{F} \subset C[0,1]$ is considered (as a special case) in Section 3. The main theorem there implies weak convergence of $\{\mathbb{U}_n(f): f \in \mathscr{F}\}$ when $\mathscr{F} = \{f \in C[0,1]: |f(t) - f(s)| \le |t - s|^{\alpha}\}$, $\alpha > \frac{1}{2}$, representing a result due to Strassen and Dudley (1969) and being closely related to the higher-dimensional results in Chapter 26.

The standardized quantile process $\mathbb{Q}_n$ studied in Chapter 18 is defined by $\mathbb{Q}_n(t) \equiv g(t)\sqrt{n}[\mathbb{F}_n^{-1}(t) - F^{-1}(t)]$ for $0 < t < 1$, where $\mathbb{F}_n$ is the empirical df based on iid rv's X_i, $1 \le i \le n$, with df F. Suppose F has a density f that is positive on (c, d) where $-\infty \le c < d \le \infty$, and zero elsewhere, and let $g \equiv f(F^{-1})$ be the density quantile function. Based on the special construction of Theorem 3.1.1, Section 1 shows weak convergence of $\mathbb{Q}_n$ to Brownian bridge $\mathbb{V} \equiv -\mathbb{U}$ in $\|\cdot/q\|$ metrics. Section 2 establishes that for "smooth" df's F on $(-\infty, \infty)$ the difference $\|\mathbb{Q}_n - \mathbb{V}_n\|$ goes to zero at a rate that is almost $n^{-1/2}$; this is enough to imply that important theorems of Kiefer and Bahadur, Finkelstein and so on extend trivially from $\mathbb{V}_n$ to $\mathbb{Q}_n$ and thus allows for miscellaneous applications presented in the same section.

Chapters 19–23 constitute the main body of statistical applications. Chapter 19 is concerned with L-statistics, i.e., linear combinations of a function of order statistics $X_{n:1} \le \cdots \le X_{n:n}$ of the form

$$T_n \equiv n^{-1} \sum_{i=1}^{n} c_{ni} h(X_{n:i})$$

based on iid rv's X_i, known constants c_{ni} and a known function h of the form $h = h_1 - h_2$ with each h_i increasing and left continuous. Again let $\mathbb{F}_n$ be the empirical df based on $X_1, \dots, X_n$ and define $J_n(t) \equiv c_{ni}$ for $(i-1)/n < t \le i/n$, with $J_n(0) = c_{n1}$ and $\Psi_n(t) \equiv \int_{1/2}^{t} J_n(s)\,ds$ for $0 \le t \le 1$. T_n can then be represented as

$$T_n = \int_0^t h(\mathbb{F}_n^{-1}) J_n\,dt = \int_0^1 h(\mathbb{F}_n^{-1})\,d\Psi_n.$$

A natural centering constant is

$$\mu_n \equiv \int_0^1 h(F^{-1}) J_n\,dt = \int_0^1 h(F^{-1})\,d\Psi_n.$$

Supposing that $X_i = F^{-1}(\xi_i)$ for iid uniform(0,1) rv's ξ_i, let $\mathbb{G}_n$ denote the empirical df based on $\xi_1, \dots, \xi_n$ and let $g \equiv h(F^{-1})$ on $(0,1)$. Then, provided that $J_n \to J$ in some sense [cf. Section 4 for the verification of step $(*)$ under Assumption 1 on page 662],

$$\begin{aligned} T_n - \mu_n &= \int_0^1 g\,d[\Psi_n(\mathbb{G}_n) - \Psi_n] \\ &\underset{(*)}{=} -\int_0^1 [\Psi_n(\mathbb{G}_n) - \Psi_n]\,dg \\ &\underset{(**)}{\approx} -\int_0^1 [\mathbb{G}_n - I] J\,dg \\ &= \frac{1}{n} \sum_{i=1}^{n} Y_i \end{aligned}$$

with $Y_i \equiv \int_0^1 [1_{[\xi_i \le t]} - t] J(t)\, dg(t)$ being iid! Thus the key of this approach is to control the size of the error $\gamma_n \equiv T_n - \mu_n - (1/n)\Sigma_{i=1}^n Y_i$ made in using the approximation of step (* *) in order to obtain a WLN, SLLN, CLT or LIL for T_n. This is achieved under certain "bounded growth" and "smoothness" assumptions on J, assuming also that $J_n \to J$ locally uniformly as $n \to \infty$. Functional versions of the main theorems are also considered. General examples are given and examples build around randomly trimmed and Winsorized means.

The main concern of Chapter 20 is linear rank statistics. For this, let $X_{n1}, \dots, X_{nn}$ be iid having a continuous df. Let $D_{n1}, \dots, D_{nn}$ denote the antiranks (defined by $X_{nD_{ni}} = X_{n:i}$) of $X_{n1}, \dots, X_{nn}$. Let $c_{n1}, \dots, c_{nn}$ denote known weights and $d_{n1}, \dots, d_{nn}$ denote known scores. Then the linear rank statistics are defined as

$$T_n \equiv (c'c)^{-1/2} \sum_{i=1}^n d_{ni} c_{nD_{ni}} = \int_0^1 h_n \, d\mathbb{R}_n = -\int_0^1 \mathbb{R}_n \, dh_n,$$

where h_n is the score(s) function defined by $h_n(t) \equiv d_{ni}$ for $(i-1)/n < t \le i/n$ with h_n right continuous at 0 and where $\mathbb{R}_n$ is the empirical rank process which satisfies (assuming $\bar{c}_n \equiv n^{-1}\Sigma_{i=1}^n c_{ni} = 0$)

$$\|\mathbb{R}_n - \mathbb{W}\| \underset{\text{a.s.}}{\to} 0 \tag{12}$$

for the special construction of Theorem 3.1.1, $\mathbb{W}$ being a Brownian bridge. Section 1 introduces a basic martingale $\mathbb{M}_n = \{\mathbb{M}_n(i/n), 0 \le i \le n\}$ being linked with $\mathbb{R}_n$ through the relation

$$\mathbb{M}_n\left(\frac{1}{n}\right) = \mathbb{R}_n(p_{ni}) + n^{-1} \sum_{j=1}^{i} \frac{\mathbb{R}_n(p_{nj})}{1 - j/n}, \tag{13}$$

with $p_{ni} \equiv i/(n+1)$. Extending $\mathbb{M}_n$ to $[0,1]$ by letting $\mathbb{M}_n(t) = \mathbb{M}_n(i/n)$ for $i/n \le t < (i+1)/n$, the natural limiting process to associate with $\mathbb{M}_n$ is [note (13) and (12)]

$$\mathbb{M}(t) \equiv \mathbb{W}(t) = \int_0^t \frac{\mathbb{W}(s)}{1-s}\, ds.$$

It is shown that

$$\|(\mathbb{M}_n - \mathbb{M})/q\| \underset{\text{a.s.}}{\to} 0 \quad \text{as } n \to \infty \tag{14}$$

for the special construction, provided that $\bar{c}_n = 0$ and $\max\{c_{ni}^2/c'c\colon 1 \le i \le n\} \to 0$ and q is increasing on $[0,1]$ with $\int_0^1 [q(t)]^{-2}\, dt < \infty$. Section 2 studies processes of the form $\mathbb{T}_n = \int_0^{\cdot} h_n \, d\mathbb{R}_n$ so that $\mathbb{T}_n(p_{nk}) = T_{nk} \equiv (c'c)^{-1/2}\Sigma_{i=1}^k d_{ni} c_{nD_{ni}}$ is the contribution made to T_n by the first k order statistics $X_{n:1}, \dots, X_{n:k}$. As to the asymptotic behavior of $\mathbb{T}_n$, again $\|\cdot/q\|$ metrics are considered requiring that $\|[(h_n - h)/q]\| \to 0$, $h/q \in \mathscr{L}_2$, to obtain $\|(\mathbb{T}_n - \mathbb{T})/q\| \to_p 0$ for the special construction under the same assumptions as in (14) and with $\mathbb{T} \equiv \int_0^{\cdot} h \, d\mathbb{W}$. An interesting corollary is obtained by setting $q \equiv 1$ in these hypotheses to conclude that $T_n = \mathbb{T}_n(1) \to_d T = \mathbb{T}(1) \cong \mathscr{N}(0, \sigma_h^2)$ for

$\sigma_h^2 \equiv \|[h]\|^2 - \bar{h}^2$ ($\bar{h} \equiv \int_0^1 h\,dt$) (cf. Hájek and Šidák [(1967), page 163] for this result under the same conditions with $q \equiv 1$). Section 3 considers rank statistics T_n under contiguous alternatives, the asymptotic power of tests based on rank statistics, the Pitman efficiency of tests and asymptotic linearity. Section 4 is devoted to the famous Chernoff–Savage theorem and the final Section 5 presents some exercises for order statistics and spacings.

Chapter 21 on Spacings considers the small-sample distribution theory of uniform spacings $\delta_{ni} \equiv \xi_{n:i} - \xi_{n:i-1}$ where $\xi_{n:1} \leq \cdots \leq \xi_{n:n}$ are the order statistics of uniform(0, 1) rv's $\xi_1, \ldots, \xi_n$. The limiting distribution of $(\delta_{n:1}, \delta_{n:n+1})$ is presented. Uniform spacings can be considered (for fixed n) to be the interarrival times of a renewal process of exponential rv's that has been standardized by dividing by the total sum. Empirical, quantile and weighted empirical processes of such renewal rv's are studied and are specialized to the exponential case. Statistics used to test for a uniform distribution are considered; they are viewed as functionals on the processes of Section 4, which process depends on the type of alternatives considered. Section 6 states some LIL-type results for spacings.

Chapter 22 is concerned with Symmetry. The so-called empirical symmetry process and the empirical rank symmetry process together with identities relating these processes to $\mathbb{U}_n$ are introduced. Testing goodness of fit (supremum tests of fit and integral tests of fit, respectively) are treated for a symmetric df and the processes under contiguity are considered. Some asymptotic results for signed rank statistics under symmetry are presented and estimators for estimating an unknown point of symmetry based on signed rank statistics and on variants of the Cramér–von Mises statistic are established. Finally the problem of estimating the df of a symmetric distribution with unknown point of symmetry is treated.

Further Applications are presented in Chapter 23: Bootstrapping the empirical process; smooth estimators of F with density f based on Kernel estimators of f; the a-shorth estimate of θ when the model is given by iid rv's with df $F_\theta = F(\cdot - \theta)$, where F has a density f that is symmetric about 0; convergence of U-statistic empirical processes; reliability and econometric functions (mean residual life function, Lorenz curve, scaled total time on test function). A unified treatment of the theory of corresponding processes has been developed by Csörgő et al. (1983).

Chapter 24 is concerned with Large Deviations. Section 1 introduces the concept of Bahadur efficiency [based on Bahadur's (1971) fundamental monograph] and presents the key theorem for deriving the exact slope of a test. The fundamental requirement of this theorem is a large deviation result. A large deviation result for binomial rv's is applied to $\mathbb{G}_n(t)$ and extend to $D_{\psi,n}^{\#} \equiv \|(\mathbb{G}_n - I)_\psi^{\#}\|$ for a weight function ψ such that ψ is positive and continuous on $(0, 1)$, symmetric about $t = \frac{1}{2}$, and $\lim_{t\to\infty}\psi(t)$ exists in $[0, \infty]$. Section 3 introduces the Kullback–Leibler information number and gives some of its basic properties. The Sanov problem and theorems giving conditions under which Sanov's conclusion is valid are stated.

Extensions to the case of Independent but Not Identically Distributed Random Variables are considered in Chapter 25. Good theorems for the empirical process in this situation require an extension of the DKW inequality 9.2.1 which is the subject of Section 1, where Bretagnolle's (1980) main inequalities are proved. Just as $n\mathbb{G}_n(t) \cong \text{binomial}(n, t)$, in the present situation of independent but not identically distributed rv's $X_{n1}, \dots, X_{nn}$ with df's $F_{n1}, \dots, F_{nn}$, $n\mathbb{F}_n(x) \equiv \sum_{i=1}^n 1_{(-\infty, x]}(X_{ni})$ has for a fixed x the generalized binomial distribution $[1_{(-\infty,x]}(X_{ni})$ being independent Bernoulli rv's with probabilities of success $p_i \equiv F_{ni}(x)]$. Compared with binomial$(n, \bar{p})$, $\bar{p} \equiv (p_1 + \cdots + p_n)/n$, the generalized binomial distribution is in a very strong sense less dispersed as shown by Hoeffding (1956). This remarkable fact is the subject of Section 2 being used in Section 3 to obtain in probability "linear bounds" on the empirical df $\mathbb{F}_n$. On their basis the weak convergence of the empirical, weighted empirical and quantile process based on independent but not identically distributed rv's in $\|\cdot/q\|$ metrics is established. Section 5 presents a CLT for linear combinations of a function of order statistics of independent but not identically distributed rv's and explores some of its consequences.

The final Chapter 26 on Empirical Measures and Processes for General Spaces gives an outline of the research of empirical processes during the past decade which has aimed to generalize the results to higher-dimensional observations and more abstract sample spaces. Further details and elaboration on these themes are provided by Dudley's (1984) monograph "A course on empirical processes" [*Lecture Notes in Math.* **1097** 1–142. Springer, Berlin], by Pollard's (1984) book on convergence of stochastic processes and by Giné and Zinn (1984).

Altogether the book of Shorack and Wellner will serve as an encyclopedia on empirical processes and its statistical applications. It includes a lot of good exercises, illustrations accompanying the proofs of major results and useful tables. The community of statisticians and probabilists should be grateful to the authors for the immense work they did in writing such a book which can be highly recommended to statisticians as well to the probabilists [when the latter are hopefully willing to follow Kempthorne's (1985) suggestion (*Inst. Math. Statist. Bull.* **14** 321–323) that "the broad needs of society are addressed only by the combination of mathematical probability *and* statistical ideas"].

REFERENCES

[1] Andersen, N. T. and Dobric, V. (1987). The central limit theorem for stochastic processes. *Ann. Probab.* **15** 164–177.

[2] Gaenssler, P. and Gutjahr, M. (1986). On computing exact Kolmogorov and Rényi type distributions. *Colloq. Math. Soc. János Bolyai* **45** 155–174.

[3] Gaenssler, P., Haeusler, E. and Schneemeier, W. (1984). Selected topics on empirical processes. In *Proc. Third Prague Symp. on Asymptotic Statistics* (P. Mandl and M. Husková, eds.) 57–91. North-Holland, Amsterdam.

[4] Gaenssler, P. and Schneemeier, W. (1986). On functional limit theorems for a class of stochastic processes indexed by pseudo-metric parameter spaces (with applications to empirical processes). *Proc. Meeting on Probability in Banach Spaces at Sandbjerg Castle June 1986. Lecture Notes in Math.* Springer, Berlin. To appear.

[5] GILL, R. D. (1987). Non- and semi-parametric maximum likelihood estimators and the von Mises method (Part I). Report MS-R 8709, Dept. Math. Statist., CWI, Amsterdam.

[6] GUTJAHR, M. and HAEUSLER, E. (1986). Elementary probabilistic computations of goodness-of-fit distributions defined by numbers of intersections. *Colloq. Math. Soc. János Bolyai* **45** 211–235.

[7] HOFFMANN-JØRGENSEN, J. (1984). Stochastic processes on Polish spaces. Unpublished.

[8] PYKE, R. (1972). Empirical processes. *Jeffrey-Williams Lectures 1968–1972, Canadian Mathematical Congress, Montreal* 13–43.

[9] WICHURA, M. J. (1968). On the weak convergence of non-Borel probabilities on a metric space. Ph.D. dissertation, Columbia Univ.

DEPARTMENT OF MATHEMATICS
UNIVERSITY OF MUNICH
D-8000 MUNICH 2
WEST GERMANY