Metrika (1999) 49: 53–69

Metrika

Goodness-of-fit tests for discrete models based on the integrated distribution function

Bernhard Klar

Institut für Mathematische Stochastik, Universität Karlsruhe, Englerstr. 2, D-76128 Karlsruhe, Germany (e-mail: Bernhard.Klar@math.uni-karlsruhe.de)

Received: July 1998

Abstract. This paper presents a new widely applicable omnibus test for discrete distributions which is based on the difference between the integrated distribution function $\Psi(t) = \int_t^\infty (1 - F(x))\,dx$ and its empirical counterpart. A bootstrap version of the test for common lattice models has accurate error rates even for small samples and exhibits high power with respect to competitive procedures over a large range of alternatives.

Key words: Goodness-of-fit test, integrated distribution function, discrete distribution, empirical process, sequence space, parametric bootstrap

1 Introduction

Model validation is an important problem of data analysis. Whereas a variety of omnibus tests are available for continuous probability models, many goodness-of-fit tests for discrete distributions such as Pearson's chi-squared test or the tests proposed in [8], [16] and [19], lack the property of consistency.

For testing the hypothesis that the underlying distribution function (df) is some completely specified df F, classical test statistics, such as those of Cramér-von Mises or Kolmogorov-Smirnov, are functionals of the empirical process $\mathscr{Z}_n = \sqrt{n}(F_n - F)$, where F_n is the empirical df of a random sample of size n.

Apart from consistency, a salient feature of these test statistics is that their distributions do not depend on the hypothetical df F provided F is continuous. In the case of composite hypotheses the test statistics are still distribution-free (i.e. the distribution does not depend on the true parameter value) if equivariant estimators for location and scale are used; this characteristic, however, is lost if shape parameters are present.

If the underlying distribution is discrete, both test statistics are not asymptotically distribution-free, not even in the case of a simple null hypothesis. For this reason, it is impossible to compute the asymptotic distributions, and extensive tables with simulated quantiles are necessary to perform the test, at least when this is carried out conventionally. Moreover, if unknown parameters are present, these tables can only be evaluated at the estimated parameter values, which entails a further approximation error.

In view of the omnipresence of high-speed computers, goodness-of-fit tests for discrete distributions based on the empirical process regained interest because the asymptotic distribution can be approximated by a parametric bootstrap procedure (see [11] for the case of the Kolmogorov-Smirnov and the Cramér-von Mises statistic).

This paper presents a new broadly-applicable omnibus test for discrete distributions which is based on the difference between the *integrated* distribution function $\Psi(t) = \int_t^\infty (1 - F(x))\,dx$ and its empirical counterpart.

The test statistic can also be represented by means of the empirical process (see (3) below), but it is not possible to use the distribution theory given in [11], since the test statistic is not a continuous functional on the space of sequences $x = (x_k)_{k\geq 0}$ converging to zero. However, a suitable setting is the Banach space ℓ_1 of all sequences with $\sum_{k\geq 0} |x_k| < \infty$.

The setting is described more precisely in the next section. Section 3 deals with the weak convergence of the empirical process in ℓ_1 and with the asymptotic null distribution of the test statistic. A parametric bootstrap procedure used to perform the test is shown to maintain asymptotically a given size α. Furthermore, the consistency of the test is established.

Simulations in Section 4 indicate that the test has accurate error rates for sample sizes of at least 20 and exhibits a favourable power over a range of alternatives. At the end of this section, an implementation of the test in S-Plus is given.

2 The new test statistic

On a common probability space $(\Omega, \mathscr{A}, P)$, let $X, X_1, X_2, \ldots, X_n, \ldots$ be a sequence of independent and identically distributed (iid) $\mathbb{N}_0$-valued random variables with common df F. Let

$$\mathscr{F}_\Theta = \{F(\cdot, \vartheta) : \vartheta \in \Theta\}$$

be a parametric family of discrete distribution functions such that the set of points of discontinuity of $F(\cdot, \vartheta)$ is a subset of $\mathbb{N}_0$. We assume that $E_\vartheta(X)$ is finite for each $\vartheta \in \Theta$, where the parameter space Θ is an open subset of $\mathbb{R}^s$. The problem is to test the hypothesis $H_0 : F \in \mathscr{F}_\Theta$ against the general alternative $H_1 : F \notin \mathscr{F}_\Theta$.

For this purpose, we propose a test statistic based on the *integrated distribution function* (idf). The latter, for a positive random variable X with $EX < \infty$, is defined by

$$\Psi_X(t) := E(X - t)^+ = \int_t^\infty \bar{F}(x)\,dx,$$

where $\bar{F} = 1 - F$ denotes the survival function. Specifically, if X is a discrete random variable with probability mass function f and finite expectation,

$$\Psi_X(t) = \sum_{k=[t]+1}^{\infty} (k-t) f(k) = ([t]+1-t)\bar{F}([t]) + \sum_{k=[t]+1}^{\infty} \bar{F}(k),$$

where $[t]$ denotes the integer part of t. The idf is convex and decreasing on $[0, \infty)$ with $\Psi_X(0) = EX$. Moreover, a distribution is uniquely determined by the idf (see [17]).

The idf plays a certain role in sequential optimization and in the comparison of stochastic orders (see [13] and [17] as well as the literature cited therein). In risk theory, it is known under the name "stop-loss transform". The definitions differ to some extent; in [13], the term *integrated distribution function* is used for $\int_0^t F(x)\,dx$. The definition given above has the advantage that Ψ_X is bounded for positive random variables with finite expectation.

To perform the test, the empirical integrated distribution function (eidf), defined by

$$\Psi_n(t) = \int_t^{\infty} \bar{F}_n(x)\,dx = \frac{1}{n}\sum_{i=1}^{n}(X_i - t)\mathbf{1}\{X_i > t\},$$

where $\mathbf{1}$ denotes the indicator function and $F_n(x) = n^{-1}\sum_{j=1}^{n}\mathbf{1}\{X_j \le x\}$ is the empirical df of $X_1, \ldots, X_n$, is compared with the estimated idf

$$\hat{\Psi}(t) = \int_t^{\infty} \bar{F}(x, \hat{\vartheta}_n)\,dx.$$

Here, $\hat{\vartheta}_n = \hat{\vartheta}_n(X_1, \ldots, X_n) \in \Theta$ is a suitable estimator of the parameter vector ϑ.

Clearly, $\Psi_n(0) = \bar{X}_n$, where $\bar{X}_n = n^{-1}\sum_{i=1}^{n} X_i$. If $0 = X_{(0)} \le X_{(1)} \le \cdots \le X_{(n)}$ denotes the ordered sample of $X_1, \ldots, X_n$, the eidf, for $X_{(i-1)} < t \le X_{(i)}$, can be written as

$$\Psi_n(t) = \frac{1}{n}\sum_{j=i+1}^{n}(X_{(j)} - X_{(j-1)})(n-j+1) + \frac{n-i+1}{n}(X_{(i)} - t).$$

In particular,

$$\Psi_n(X_{(i)}) = \frac{1}{n}\sum_{j=i+1}^{n}(X_{(j)} - X_{(j-1)})(n-j+1), \quad i = 1, \ldots, n.$$

Note that $\Psi_n(X_{(i)}) = \bar{X}_n - t_n(i/n)$, where

$$t_n\left(\frac{i}{n}\right) = \frac{1}{n}\sum_{j=1}^{i}(X_{(j)} - X_{(j-1)})(n-j+1), \quad i = 1, \ldots, n,$$

is the *total time on test*, which is well-known in reliability theory. The theoretical *total time on test transform* $t(p) = \int_0^{F^{-1}(p)} \bar{F}(s)\,ds$ and the idf are related by $t(p) + \Psi(F^{-1}(p)) = EX$ $(0 < p < 1)$. An overview over the connections between these and other related quantities like the *Lorenz curve* can be found in [5] and [10]. Barlow and Campo [1] sum up properties of the total time on test transform and study goodness-of-fit tests against special classes of alternatives which are based on this quantity.

The new test statistic is

$$T_n = \sup_{t \geq 0} \sqrt{n}\,|\Psi_n(t) - \hat{\Psi}(t)|, \tag{1}$$

formed by analogy with the Kolmogorov-Smirnov-statistic for the empirical process. Since Ψ_n and $\hat{\Psi}$ are linear on the interval $(k, k+1]$, $k \in \mathbb{N}_0$, we obtain

$$T_n = \sup_{k \in \mathbb{N}_0} \sqrt{n}\,|\Psi_n(k) - \hat{\Psi}(k)| = \max_{0 \leq k \leq M} \sqrt{n}\,|\Psi_n(k) - \hat{\Psi}(k)|, \tag{2}$$

where $M = \max_{1 \leq i \leq n} X_i$. T_n can be written as a functional of the estimated (discrete) empirical process $\mathscr{Z}_n = (Z_{n,k})_{k \geq 0}$, where $Z_{n,k} = \sqrt{n}(F_n(k) - F(k, \hat{\vartheta}_n))$: from

$$\Psi_n(k) - \hat{\Psi}(k) = \sum_{j=k}^{\infty} (F(j, \hat{\vartheta}_n) - F_n(j))$$

it follows that

$$T_n = h(\mathscr{Z}_n) = \sup_{k \geq 0} \left| \sum_{j \geq k} Z_{n,j} \right|, \tag{3}$$

where, for a sequence $x = (x_k)_{k \geq 0}$, $h(\cdot)$ is given by $h(x) = \sup_{k \geq 0} |\sum_{j \geq k} x_j|$. However, since $h(\cdot)$ is not defined on the space of all sequences converging to zero, the results in [11] are not applicable. A suitable setting is the Banach space ℓ_1 of all sequences $x = (x_k)_{k \geq 0}$ with $\sum_{k \geq 0} |x_k| < \infty$, equipped with the norm $\|x\| = \sum_{k \geq 0} |x_k|$. $\mathscr{Z}_n$ is then regarded as a random element in ℓ_1. It is readily seen that $h(\cdot)$ is a continuous functional on ℓ_1. Consequently, the weak convergence of the test statistic follows from the weak convergence of the empirical process (see Corollary 3.2).

3 Results

In this section, we consider the situation of a triangular array $X_{n1}, X_{n2}, \ldots, X_{nn}$ of row-wise iid random variables having the df $F(\cdot, \vartheta_n)$, where $\vartheta_n \in \Theta$, $n \geq 1$, $\lim_{n\to\infty} \vartheta_n = \vartheta$, $\vartheta \in \Theta$. This generalization is necessary to prove the convergence of the bootstrap procedure described later. Throughout the rest of the paper, $\xrightarrow{\mathscr{D}}$ denotes weak convergence of random variables or stochastic processes, $\xrightarrow{P}$ stands for convergence in probability, and v' is the transpose of a row vector v.

We need the following regularity conditions, which hold for standard families of discrete distributions (see Section 4).

(R1) There exists a measurable function $l : \mathbb{N}_0 \times \Theta \to \mathbb{R}^s$ such that

$$\sqrt{n}(\hat{\vartheta}_n - \vartheta_n) = \frac{1}{\sqrt{n}} \sum_{j=1}^{n} l(X_{nj}, \vartheta_n) + \varepsilon_n,$$

where $\varepsilon_n \xrightarrow{P_{\vartheta_n}} 0$ as $n \to \infty$. l has the following properties:
 (i) $E_\vartheta\, l(X, \vartheta) = 0,\ \vartheta \in \Theta$.
 (ii) For fixed $k \in \mathbb{N}_0$, $l(k, \cdot)$ is a continuous function of ϑ.
 (iii) $D(\vartheta) = E[l(X, \vartheta)' l(X, \vartheta)]$ is a finite nonnegative definite matrix that depends continuously on ϑ.

(R2) There is an integer n_0 such that

$$\lim_{l \to \infty} \sup_{n \geq n_0} \sum_{k \geq l} \sqrt{1 - F(k, \vartheta_n)} = 0.$$

(R3) For each $k \geq 0$, the gradient vector $\nabla_\vartheta F(k, \vartheta)$ of $F(k, \vartheta)$ exists and depends continuously on ϑ. Moreover,

$$\lim_{l \to \infty} \sup_{\vartheta^* \in \mathscr{U}(\vartheta)} \sum_{k \geq l} \left| \frac{\partial F(k, \vartheta^*)}{\partial \theta_j} \right| = 0, \quad j = 1, \ldots, s,$$

in some neighbourhood $\mathscr{U}(\vartheta)$ of $\vartheta = (\theta_1, \ldots, \theta_s)'$.

3.1 Theorem. *Under assumptions (R1)–(R3),*

$$\mathscr{Z}_n \xrightarrow{\mathscr{D}} \mathscr{W}.$$

Here, $\mathscr{W} = (W_k)_{k \geq 0}$ is a Gaussian process in ℓ_1 such that, for $k, m \geq 0$ and $\vartheta \in \Theta$, $E_\vartheta[W_k] = 0$, and the covariance function is given by

$$\begin{aligned} C_\vartheta(k, m) = {} & F(min(k, m), \vartheta) - F(k, \vartheta) F(m, \vartheta) \\ & - J(k, \vartheta) \nabla_\vartheta F(m, \vartheta)' - J(m, \vartheta) \nabla_\vartheta F(k, \vartheta)' \\ & + \nabla_\vartheta F(k, \vartheta) D(\vartheta) \nabla_\vartheta F(m, \vartheta)', \end{aligned} \tag{4}$$

where

$$J(k, \vartheta) = E_\vartheta[l(X, \vartheta) \mathbf{1}\{X \leq k\}] = \sum_{j=0}^{k} l(j, \vartheta) f(j, \vartheta).$$

Proof: Since the space ℓ_p, $1 \leq p < \infty$, is separable, the σ-algebra $\mathscr{B}$ of Borel-sets of ℓ_p is generated by the spheres $K(x, \varepsilon) = \{y \in \ell_p : \|x - y\|_p < \varepsilon\}$ $(x \in \ell_p, \varepsilon > 0)$, and it is easily seen that $\mathscr{B}$ coincides with the projection σ-

algebra $\mathscr{M}$, i.e., the smallest σ-algebra on ℓ_p such that the projections $x \to \pi_k x := x_k$ $(x \in \ell_p, k \geq 0)$ are measurable.

Obviously, $\mathscr{Z}_n$ is a $(\mathscr{A}, \mathscr{M})$-measurable mapping from Ω into ℓ_1. Hence, $\mathscr{Z}_n$ is $(\mathscr{A}, \mathscr{B})$-measurable, i.e. the distribution $Q_n = P \circ \mathscr{Z}_n^{-1}$ is a Borel probability measure on ℓ_1, and we can utilize the theory of weak convergence on metric spaces (see, e.g., [3]).

Now, a subset M of $\ell_p, 1 \leq p < \infty$ is relatively compact if and only if M is bounded and $\lim_{n\to\infty} \sup_{x\in M} \sum_{j\geq n} |x_j|^p = 0$. Using this condition and Prokhorov's Theorem, the following tightness conditions can be proved similarly as Lemma 2.1 in [11], putting $A_j = \{x \in \ell_1 : \sum_{k \geq l(j)} |x_k| \leq 1/j\}$ in the proof of Lemma 2.1 in [11].

A sequence $\{Q_n : n \geq 1\}$ of probability measures on $(\ell_p, \mathscr{B})$ is tight if and only if these conditions hold:

(i) For each positive δ and $l \in \mathbb{N}_0$ there exists a finite constant K such that

$$Q_n(\{x \in l_p : |x_l| \leq K\}) \geq 1 - \delta, \quad n \in \mathbb{N}.$$

(ii) For each positive δ and η, there is an integer l such that

$$Q_n\left(\left\{x \in l_p : \sum_{k \geq l} |x_k|^p \leq \eta\right\}\right) \geq 1 - \delta, \quad n \in \mathbb{N}.$$

To establish the weak convergence of $\mathscr{Z}_n$ to $\mathscr{W}$, one first has to show the convergence of finite-dimensional distributions to centered normal distributions with covariances given in (4); this can be done as in [11]. Consequently, the sequence $(Z_{n,l})_{n\geq 1}$ is tight for fixed $l \in \mathbb{N}_0$, i.e. condition (i) above is satisfied. To verify condition (ii) in the case $p = 1$, we have to find integers l and n_0 such that $P_{\vartheta_n}(\sum_{k\geq l} |Z_{n,k}| > \eta) \leq \delta$ for each $n \geq n_0$. Since $Z_{n,k} = T_{n,k} + V_{n,k}$, where $T_{n,k} = \sqrt{n}(F_n(k) - F(k, \vartheta_n))$ and $V_{n,k} = \sqrt{n}(F(k, \vartheta_n) - F(k, \hat{\vartheta}_n))$, this is equivalent to the existence of integers l and n_0 such that

$$P_{\vartheta_n}\left(\sum_{k\geq l} |T_{n,k}| > \frac{\eta}{2}\right) \leq \frac{\delta}{2}, \quad n \geq n_0, \tag{5}$$

and

$$P_{\vartheta_n}\left(\sum_{k\geq l} |V_{n,k}| > \frac{\eta}{2}\right) \leq \frac{\delta}{2}, \quad n \geq n_0. \tag{6}$$

By Markov's and Jensen's inequality, (5) follows from

$$\begin{aligned} P_{\vartheta_n}\left(\sum_{k\geq l} |T_{n,k}| > \frac{\eta}{2}\right) &\leq \frac{2}{\eta} E_{\vartheta_n}\left[\sum_{k\geq l} |T_{n,k}|\right] \\ &= \frac{2}{\eta} \sum_{k\geq l} E_{\vartheta_n}[|T_{n,k}|] \end{aligned}$$

$$\leq \frac{2}{\eta}\sum_{k\geq l}\sqrt{E_{\vartheta_n}[|T_{n,k}^2|]}$$

$$\leq \frac{2}{\eta}\sum_{k\geq l}\sqrt{1-F(k,\vartheta_n)} \to 0 \quad (l\to\infty),$$

where the convergence is uniform in $n \geq n_0$ in view of (R2).

To prove (6), we need the following well-known fact: given a vector $z \in \mathbb{R}^s$ and a $(r\times s)$-matrix $A = (a_{kj})$, the matrix norm $\|A\|_1 = \max_{1\leq j\leq s}\sum_{1\leq k\leq r}\cdot |a_{kj}|$ satisfies $\|Az\|_1 \leq \|A\|_1\|z\|_1$, where $\|\cdot\|_1$ denotes the ℓ_1-norm in $\mathbb{R}^s$. By the mean value theorem for vector valued functions,

$$\sum_{k=l}^{l+m}|V_{n,k}| = \sum_{k=l}^{l+m}|\sqrt{n}(F(k,\vartheta_n)-F(k,\hat\vartheta_n))|$$

$$\leq \|\sqrt{n}(\hat\vartheta_n-\vartheta_n)\|_1 \max_{\vartheta^*\in S}\max_{1\leq j\leq s}\sum_{k=l}^{l+m}\left|\frac{\partial F(k,\vartheta^*)}{\partial\theta_j}\right|$$

$$\leq \|\sqrt{n}(\hat\vartheta_n-\vartheta_n)\|_1\, a_l$$

for $m \in \mathbb{N}$ and $\hat\vartheta_n, \vartheta_n \in \mathscr{U}(\vartheta)$, where S is the line between ϑ_n and $\hat\vartheta_n$, and

$$a_l = \sup_{\vartheta^*\in\mathscr{U}(\vartheta)}\max_{1\leq j\leq s}\sum_{k=l}^{\infty}\left|\frac{\partial F(k,\vartheta^*)}{\partial\theta_j}\right|.$$

By (R3), $\lim_{l\to\infty} a_l = 0$ for a sufficiently small neighbourhood $\mathscr{U}(\vartheta)$ of ϑ. Now, (6) can be established like the corresponding condition in the proof of Theorem 3.1 in [11]. ■

Remark: The proof of Theorem 3.1 indicates that conditions (R2) and (R3) are by no means necessary. However, they are quite easy to verify in applications; a condition corresponding to (R3) is quite common in similar cases (see, e.g., condition (iv) on page 793 of Burke et al. [4]).

In the following, we make the general assumption that the hypothetical distribution $F(\cdot,\vartheta)$ is non-degenerate and satisfies assumptions (R1)–(R3). Since $h(\cdot)$ is continuous, the next result follows from Theorem 3.1 and the continuous mapping theorem.

3.2 Corollary. *Under* $P_{\vartheta_n}, \vartheta_n \in \Theta$, *where* $\lim_{n\to\infty}\vartheta_n = \vartheta \in \Theta$, *we have*

$$T_n \xrightarrow{\mathscr{D}} h(\mathscr{W}) = \sup_{k\geq 0}\left|\sum_{j\geq k} W_j\right|.$$

The distribution $H_\vartheta(t) = P_\vartheta(h(\mathscr{W})\leq t)$ is difficult to calculate; furthermore, it depends on the unknown true value of ϑ. Hence, the hypothetical distribution $H_{n,\vartheta}(t) = P_\vartheta(T_n \leq t)$ of T_n is approximated by the following bootstrap procedure.

Given $X_1, \ldots, X_n$, first compute $\hat{\vartheta}_n$ and $T_n(X_1, \ldots, X_n)$. Then compute $T^*_{j,n} = T_n(X^*_{j1}, \ldots, X^*_{jn}), 1 \le j \le b$, where, conditionally on $X_1, \ldots, X_n$, the random variables $X^*_{j1}, \ldots, X^*_{jn}, 1 \le j \le b$, are iid with df $F(\cdot, \hat{\vartheta}_n)$. Denoting by $c^*_{n,b}(\alpha)$ the $(1-\alpha)$-quantile of $H^*_{n,b}$, where

$$H^*_{n,b}(t) = \frac{1}{b}\sum_{j=1}^{b} \mathbf{1}\{T^*_{j,n} \le t\}$$

is the empirical df of $T^*_{1,n}, \ldots, T^*_{b,n}$, the hypothesis H_0 is rejected at level α if T_n exceeds $c^*_{n,b}(\alpha)$.

To prove that this *bootstrap test* based on T_n has asymptotic level α and is consistent against alternatives with finite expectation, we need the following results.

3.3 Lemma. *Let $Q = P \circ \mathscr{W}^{-1}$ denote the distribution of $\mathscr{W}$. The distribution Qh^{-1}* of $h(\mathscr{W})$ *is concentrated on* $(0, \infty)$ *and absolutely continuous with respect to Lebesgue measure λ; the density $dQh^{-1}/d\lambda$ is strictly positive on* $(0, \infty)$. *In particular, the df H_ϑ is continuous and strictly increasing on* $[0, \infty)$.

Proof: The functional h is a norm on ℓ_1 and hence convex. Let N_Q denote the *linear* support of Q (see [7], p. 2797), which, for Gaussian measures, coincides with the *topological* support ([7], Theorem 5.1.3). Since Q is centered, $0 \in N_Q$ (see [7], p. 2814) and consequently $\inf_{x \in N_Q} h(x) = 0$. By Theorem 7.1 of [7], the distribution Qh^{-1} of $h(\mathscr{W})$ is concentrated on $[0, \infty)$ and is absolutely continuous on $(0, \infty)$; the Lebesgue density $dQh^{-1}/d\lambda$ is strictly positive on $(0, \infty)$. By assumption, the distribution $F(\cdot, \vartheta)$ is non-degenerate and hence Q is not the Dirac measure in 0. Since $h(\mathscr{W}) = 0$ if and only if $\mathscr{W} = 0$, we have $P(\mathscr{W} = 0) = P(h(\mathscr{W}) = 0) = 0$. Therefore, Qh^{-1} is concentrated on $(0, \infty)$. ■

3.4 Lemma. *Suppose that $X_1, X_2, \ldots$ have finite expectation and $\hat{\vartheta}_n = \hat{\vartheta}_n(X_1, \ldots, X_n)$ converges to $\vartheta \in \Theta$ in probability. Let $\|f\|_\infty = \sup_{-\infty < t < \infty} |f(t)|$ for a function $f : \mathbb{R} \to \mathbb{R}$. Then*

a) $\|H^*_{n,b} - H_\vartheta\|_\infty \xrightarrow{P} 0$ *as $n, b \to \infty$.*
b) $c^*_{n,b}(\alpha) \xrightarrow{P} H_\vartheta^{-1}(1-\alpha)$ *as $n, b \to \infty$.*

Proof: Since H_ϑ is continuous by Lemma 3.3, Pólya's theorem states that the assertion of Corollary 3.2 is equivalent to $\lim_{n\to\infty} \|H_{n,\vartheta_n} - H_\vartheta\|_\infty = 0$ for each sequence $\vartheta_n \in \Theta$ converging to ϑ. Combining this with the convergence in probability of $\hat{\vartheta}_n$ to ϑ we obtain $\|H_{n,\hat{\vartheta}_n} - H_\vartheta\|_\infty \xrightarrow{P} 0$ as $n \to \infty$. On the other hand, $\|H^*_{n,b} - H_{n,\hat{\vartheta}_n}\|_\infty$ converges to zero almost surely (see [11], p.89). Hence, a) follows from the triangle inequality. Since, by Lemma 3.3, H_ϑ is continuous and strictly increasing on $[0, \infty)$, assertion b) follows from a). ■

3.5 Theorem. *a) For each $\vartheta \in \Theta$, $\lim P_\vartheta(T_n > c^*_{n,b}(\alpha)) = \alpha$ as $n, b \to \infty$.*

b) Let $X_1, X_2, \ldots$ be an iid sequence of discrete random variables having df F with

$$\rho := \inf_{\vartheta \in \Theta} \sup_{t \in \mathbb{N}_0} |F(t) - F(t, \vartheta)| > 0.$$

If the expectation of X_1 *is finite and* $\hat{\vartheta}_n \in \Theta$ *converges in probability to* $\vartheta \in \Theta$, *the test is consistent, i.e.* $\lim P(T_n > c^*_{n,b}(\alpha)) = 1$ *as* $n, b \to \infty$.

Proof: a) follows by Lemma 3.4 b), since the assumptions of 3.4 are satisfied by the hypothetical distribution. For proving b), note that

$$T_n = \sup_{k\geq 0}\left|\sum_{j\geq k}\sqrt{n}(F_n(j) - F(j,\hat{\vartheta}_n))\right|$$

$$\geq \sqrt{n}\sup_{k\geq 0}\left|\left|\sum_{j\geq k}(F_n(j) - F(j))\right| - \left|\sum_{j\geq k}(F(j) - F(j,\hat{\vartheta}_n))\right|\right|. \quad (7)$$

Now, for each j, $1 - F_n(j)$ converges almost surely to $1 - F(j)$, and

$$\lim_{n\to\infty}\sum_{j=0}^{\infty}(1 - F_n(j)) = \lim_{n\to\infty}\overline{X}_n = EX = \sum_{j=0}^{\infty}(1 - F(j)) \quad a.s.$$

By a version of Scheffé's theorem ([20], p. 862),

$$\sup_{k\geq 0}\left|\sum_{j\geq k}(F_n(j) - F(j))\right| = \sup_{k\geq 0}\left|\sum_{j\geq k}(1 - F(j)) - (1 - F_n(j))\right| \to 0 \quad a.s.$$

On the other hand, as $\rho > 0$, we get $\liminf_{n\to\infty}\sup_{k\geq 0}\cdot |\sum_{j\geq k}(F(j) - F(j,\vartheta_n))| > 0$ for any sequence ϑ_n from Θ. In view of (7), $\lim_{n\to\infty} T_n = \infty$ *a.s.* Combining this with Lemma 3.4 (note that the assumptions of 3.4 are satisfied by the alternative distribution), assertion b) follows. ■

Remarks:

1. The assumptions concerning the alternative distribution can be weakened. However, besides establishing the convergence of T_n to infinity, the behaviour of the critical value $c^*_{n,b}(\alpha)$ has to be controlled. For example, it suffices that $c^*_{n,b}(\alpha)$ stays bounded in probability (cp. [11], p. 91).
2. Lemma 3.4 and Theorem 3.5a) remain valid for each continuous functional on ℓ_1 satisfying the assumptions of Lemma 3.3. As an example, take the ℓ_1-norm itself with pertaining test statistic

$$\tilde{T}_n = \sum_{k\geq 0}|Z_{n,k}|.$$

Note that, apart from the term $\sqrt{n}$, $\tilde{T}_n$ corresponds to Mallow's metric for $r = 1$ between F_n and $F(\cdot,\hat{\vartheta}_n)$ (see, e.g., Bickel and Freedman [2]). A further example is the modified Cramér-von Mises statistic

$$W^2_{mod} = \sum_{k\geq 0} Z^2_{n,k},$$

introduced by Choulakian et al. [6] and Spinelli und Stephens [21] as a test statistic for discrete distributions with a finite number of classes and for Poissonity, respectively; for distribution theory in the second case, a finite number of classes was assumed. Clearly, W^2_{mod} gives greater weight to deviations in the tails than the usual Cramér-von Mises statistic.
The proof of Theorem 3.5b) shows that this part remains valid for both statistics.

3. The setting of the last section may be used to study the asymptotic power of the eidf test with respect to contiguous alternatives. To this end, let

$$\mathscr{F}_\Theta = \{F(\cdot,\vartheta) : \vartheta \in \Theta\} \subset \{\tilde{F}(\cdot,\vartheta,\beta) : \vartheta \in \Theta, \beta \in B \subset \mathbb{R}^q\},$$

where $\tilde{F}(\cdot,\vartheta,0) = F(\cdot,\vartheta)$. Contiguous alternatives to $H_{n0} : X_{n1},\ldots,X_{nn}$ iid according to $F(\cdot,\vartheta)$ are then given by

$$H_{n1}\colon X_{n1} \sim \tilde{F}\left(\cdot,\vartheta,\frac{\gamma}{\sqrt{n}}\right) \quad \left(\frac{\gamma}{\sqrt{n}} \in B, n = 1,2,\ldots\right).$$

Under suitable regularity conditions, the asymptotic distribution of the estimated discrete empirical process under H_{n1} will be that of $(W_k+\delta_k)_{k\geq 0}$, where $(W_k)_{k\geq 0}$ is the Gaussian sequence of Theorem 3.1 and $(\delta_k)_{k\geq 0}$ is given by

$$\delta_k = \gamma\nabla_\beta\tilde{F}(k,\vartheta,\beta)|'_{\beta=0} - \nabla_\beta E_{\vartheta,\beta}[l(X,\vartheta)]|_{\beta=0}\gamma'\nabla_\vartheta F(k,\vartheta)'.$$

Thus, the asymptotic power of the test at level α for testing H_{n0} against H_{n1} is

$$P(h((W_k+\delta_k)_{k\geq 0}) \geq c_\alpha) = P\left(\sup_{k\geq 0}\left|\sum_{j\geq k} W_k + \delta_k\right| \geq c_\alpha\right),$$

where $c_\alpha = H_\vartheta^{-1}(1-\alpha)$. For details, see [15].

4 Examples and simulation

In this section, the finite sample properties of the bootstrap tests based on T_n, $\tilde{T}_n$ and W^2_{mod} are assessed by means of a simulation study. To allow for a comparison with the results of previous Monte Carlo experiments ([8], [18]), the hypothetical distributions are the Poisson, the positive Poisson, the geometric and the logseries distribution.

To calculate the statistics in practice, we obtain from (2) and $\Psi_{\hat{\vartheta}_n}(0) = E_{\hat{\vartheta}_n}(X)$

$$T_n = \sqrt{n}\sup_{1\leq k\leq M}\left|\overline{X}_n - E_{\hat{\vartheta}_n}(X) + \sum_{j=0}^{k-1}(F_n(j) - F(j,\hat{\vartheta}_n))\right|, \tag{8}$$

Table 1. 10%-level powers of the three tests for the Poisson model

Distribution	T_n $n = 50$	T_n $n = 200$	$\tilde{T}_n$ $n = 50$	$\tilde{T}_n$ $n = 200$	W^2_{mod} $n = 50$	W^2_{mod} $n = 200$
$P(3)$	10.0	9.5	9.7	9.8	10.0	9.6
$P(7)$	10.1	10.1	9.7	10.1	9.5	10.2
$Bin(10, .2)$	25	66	20	57	20	54
$Bin(10, .5)$	94	100	89	100	87	100
$NB(5, .71)$	47	93	44	90	40	88
$NB(10, .5)$	93	100	92	100	88	100
$NB(1, .3)$	44	89	43	88	41	87
$NB(1, .9)$	100	100	100	100	100	100
$NA(5, .2)$	60	98	58	98	56	98
$NA(2, .5)$	23	53	22	49	20	47
$P \circ L(.5, .5)$	27	60	26	57	23	51
$P \circ L(.8, .2)$	14	21	14	20	13	17
$P0(.1, 3)$	42	90	44	90	46	94
$P0(.3, 1)$	41	91	42	91	41	91
$PB(.9, 2, .9)$	18	40	18	40	18	43
$PB(.9, 20, .9)$	17	36	16	36	17	38
$PNB(.1, 2, .5)$	86	100	84	100	80	100

where $M = \max_{1 \le j \le n} X_j$. In the following examples, $\hat{\vartheta}_n$ always coincides with the moment estimator of EX, and therefore $E_{\hat{\vartheta}_n}(X) = \overline{X}_n$. In a similar way,

$$\tilde{T}_n = \sum_{k=0}^{M} |Z_{n,k}| + \sqrt{n} E_{\hat{\vartheta}_n}(X) - \sqrt{n} \sum_{k=0}^{M} (1 - F(k, \hat{\vartheta}_n)).$$

In the case of W^2_{mod}, the infinite sum has to be truncated.

4.1 Example. As a first example, consider the hypothesis of Poissonity. Here, $\Theta = \{\vartheta \in \mathbb{R} : \vartheta > 0\}$ and $f(j, \vartheta) = e^{-\vartheta} \vartheta^j / j!$, $j \ge 0$. The estimator $\hat{\vartheta}_n = \overline{X}_n$-satisfies (R1); since $\partial F(k, \vartheta)/\partial \vartheta = -e^{-\vartheta} \vartheta^k / k!$, it is easy to see that assumption (R3) holds. For $k + 1 \ge \vartheta$, Proposition 1 (ii) in [9] with $m = 1$ yields

$$1 - F(k, \vartheta) < \frac{k+2}{k+2-\vartheta} f(k+1, \vartheta)$$

and thus, for sufficiently large $l \in \mathbb{N}$ with $\vartheta_* = \inf_{n \ge n_0} \{\vartheta_n\}$,

$$\sup_{n \ge n_0} \sum_{k \ge l} \sqrt{1 - F(k, \vartheta_n)} \le 2 \sum_{k \ge l} \sqrt{f(k+1, \vartheta_*)}.$$

Hence, (R2) is satisfied as well.

Table 1 shows estimates of 10%-powers of the tests for the Poisson model based on T_n, $\tilde{T}_n$ and W^2_{mod} for samples sizes $n = 50$ and $n = 200$. The entries in this and the following table are the percentages of 1000 Monte Carlo samples that resulted in rejection of the hypothesis of Poissonity. The bootstrap sample size b was taken to be 200.

Besides standard distributions, the alternatives in Table 1 are Poisson-logseries mixtures $P \circ L$, Poisson-binomial mixtures PB, a Poisson-negative binomial mixture PNB and the Poisson with added zeros $P0$. These and the other alternative models used in this section are defined as follows.

- $Bin(n,p)$: $f(k) = \binom{n}{k} p^k (1-p)^{n-k}$, $k = 0, \ldots, n$;
- $NB(r,p)$: $f(k) = \binom{r+k-1}{k} p^r (1-p)^k$, $k = 0, 1, \ldots$;
- $U(0..n)$: $f(k) = 1/(n+1)$, $k = 0, 1, \ldots, n$;
- $L(\vartheta)$: $f(k) = -[\ln(1-\vartheta)]^{-1} \vartheta^k / k$, $k = 1, 2, \ldots$;
- $NA(\varphi_1, \varphi_2)$: $f(k) = \dfrac{e^{-\varphi_1} \varphi_2^k}{k!} \sum_{j=0}^{\infty} \dfrac{(\varphi_1 e^{-\varphi_2})^j j^k}{j!}$, $k = 0, 1, \ldots$;
- $Z(\varphi)$: $f(k) = (\sum_{j=1}^{\infty} j^{-(\varphi+1)})^{-1} k^{-(\varphi+1)}$, $k = 1, 2, \ldots$;
- $P \circ L(\varphi_1, \varphi_2)$: $f(k) = -[\ln(1-\varphi_2)]^{-1} \varphi_1^k / k! \sum_{j=1}^{\infty} j^{k-1} (\varphi_2 e^{-\varphi_1})^j$, $k = 0, 1, \ldots$;
- $P0(\varphi_1, \varphi_2)$: $P(X=0) = \varphi_1 = 1 - P(X=Y)$, $Y \sim P(\varphi_2)$;
- $PB(\varphi_1, m, \varphi_2)$: $P(X = Y_1) = \varphi_1 = 1 - P(X = Y_2)$, $Y_1 \sim P(m\varphi_2)$, $Y_2 \sim Bin(m, \varphi_2)$;
- $PNB(\varphi_1, m, \varphi_2)$: $P(X = Y_1) = \varphi_1 = 1 - P(X = Y_2)$, $Y_1 \sim P(m(1-\varphi_2)/\varphi_2)$, $Y_2 \sim NB(m, \varphi_2)$;
- $\varphi P(\vartheta_1) + (1-\varphi) P(\vartheta_2)$: $P(X = Y_1) = \varphi = 1 - P(X = Y_2)$, $Y_1 \sim P(\vartheta_1)$, $Y_2 \sim P(\vartheta_2)$.

Bin^+ denotes the translated binomial distribution: If $Y \sim Bin(n,p)$, then $X = Y + 1 \sim Bin^+(n,p)$; and similarly for other distributions. If $Y \sim L(\vartheta)(Z(\varphi))$, then $X = Y - 1 \sim L^-(\vartheta)(Z^-(\varphi))$.

In Table VI of [8], the same alternatives and sample sizes as in Table 1 were used. Therein, some parameter values seem to be incorrect; hence, we interchanged the parameters in the Neyman type-A distribution $NA(\varphi_1, \varphi_2)$; for the Poisson-binomial mixture $PB(\varphi_1, m, \varphi_2)$ we used $\varphi_1 = 0.9$ instead of $\varphi_1 = 0.1$, where φ_1 is the weight of the Poisson distribution.

Note that all tests used in [8] (Pearson's χ^2 test, Fisher's dispersion test, the smooth test of Rayner and Best [19] as well as a new test based on the empirical generating function) do not share the property of consistency. However, they are asymptotically distribution-free under H_0 such that asymptotical critical values can easily be obtained.

As a first result, it can be seen that the tests based on T_n, $\tilde{T}_n$ resp. W^2_{mod} behave quite similar, with some advantages for T_n. Since other simulations give the same impression, we do not report the results for $\tilde{T}_n$ resp. W^2_{mod} in the following tables. In most cases, the test based on T_n is not worse than the best test used in [8]; particularly for the sample size $n = 200$, the eidf test frequently has the highest power.

Table 2 shows the power of the test for Poissonity for small samples ($n = 20$ resp. $n = 31$). The same sample sizes and alternatives were used in [18], Table 3, as well as in [8], Table VI. Besides standard distributions, the alternatives include the Zeta distribution and a Poisson-Poisson mixture.

Nakamura and Pérez-Abreu [18] compared the dispersion test, the smooth test and a probability generating function test proposed in [16], as well as their own test which is tailored to the Poisson distribution. In contrast to the other

Table 2. Powers of the eidf test for Poissonity

Distribution	Level:	$n = 20$ 0.10	0.05	$n = 31$ 0.10	0.05
$P(3)$		10.7	4.0	10.7	3.1
$P(7)$		8.6	6.1	10.1	4.9
$Bin(6, 0.5)$		57	37	76	56
$Bin(3, 0.4)$		39	23	52	32
$NB(10, 0.7)$		27	17	36	28
$NB(1, 0.4)$		74	62	90	82
$U(0..4)$		14	8	17	8
$U(0..8)$		65	53	83	73
$L^-(0.6)$		56	44	70	57
$L^-(0.8)$		89	84	98	95
$Z^-(0.8)$		99	98	100	100
$\frac{1}{2}(P(2) + P(6))$		73	61	90	80

three procedures, the last test is an omnibus test, but it is not distribution-free.

It is difficult to compare the small sample results of Table 2 with those of other simulation studies. In most cases, the asymptotically distribution-free tests are conservative for small sample size; thus, their critical values have to be corrected empirically to maintain the nominal level. Clearly, these corrections enhance the power of the test, but they require tables of critical values that depend on n and ϑ; therefore, their usage will be avoided in practical applications.

On the other hand, the empirical level of the test of Nakamura and Pérez-Abreu [18] is far above the nominal level for small sample sizes (about 0.08 for $n = 20$ and $\alpha = 0.05$, see Table 2 in [18]). Performing this test by means of a parametric bootstrap, the nominal level is maintained very closely, but powers decrease strongly; for some alternatives, power is only about two-thirds of the values given in Table 3 in [18].

As a conclusion, the eidf test does particularly well against heavy-tailed distributions like $L^-(.8)$ or $Z^-(.8)$, whereas it does not detect the uniform distribution $U(0, \ldots, 4)$ which, like the Poisson distribution, satisfies $E(X) = Var(X)$. The result for the Zeta distribution $Z^-(.8)$ indicates that the test is able to detect distributions with infinite expectation.

Since for samples this small none of the tests for Poissonity listed above detects alternatives where expectation equals variance, we prefer Fisher's dispersion test for samples of size 20 or 30 due to its simplicity. However, typical diagnostic pitfalls associated with Fisher's dispersion index should be avoided (see [12]).

4.2 Example. As further examples, we consider tests for the geometric, the logseries and the positive Poisson distribution.

In case of the geometric distribution, $\Theta = \{\vartheta \in \mathbb{R} : 0 < \vartheta < 1\}$ and $f(j, \vartheta) = (1 - \vartheta)^j \vartheta$, $j \geq 0$. The maximum likelihood estimator (mle) $\hat{\vartheta}_n = (1 + \overline{X}_n)^{-1}$ satisfies (R1). Since $1 - F(k, \vartheta) = (1 - \vartheta)^{k+1}$ and $\partial F(k, \vartheta)/\partial \vartheta = -(k + 1)(1 - \vartheta)^k$, it is readily seen that assumptions (R2) and (R3) hold.

For the logseries distribution, $\Theta = \{\vartheta \in \mathbb{R} : 0 < \vartheta < 1\}$ and $f(j, \vartheta) = a(\vartheta)\vartheta^j/j, j \geq 1$, where $a(\vartheta) = -(\ln(1-\vartheta))^{-1}$.

The mle $\hat{\vartheta}_n$ is the (unique) solution of the equation $\overline{X}_n = -\vartheta((1-\vartheta)\ln(1-\vartheta))^{-1}$ and satisfies (R1). (R2) follows from

$$1 - F(k, \vartheta) < \frac{a(\vartheta)\vartheta^{k+1}}{(k+1)(1-\vartheta)}, \tag{9}$$

(see [14], p. 292). Combining this with the inequality

$$\begin{aligned} \left|\frac{\partial F(k,\vartheta)}{\partial \vartheta}\right| &= \left|\frac{\partial(1-F(k,\vartheta))}{\partial \vartheta}\right| \\ &= \left|\frac{a'(\vartheta)}{a(\vartheta)}\sum_{j>k} a(\vartheta)\frac{\vartheta^j}{j} + a(\vartheta)\sum_{j>k}\vartheta^{j-1}\right| \\ &< \frac{-a'(\vartheta)}{a(\vartheta)}(1-F(k,\vartheta)) + \frac{a(\vartheta)\vartheta^k}{1-\vartheta}, \end{aligned}$$

reveals that (R3) holds as well.

The positive Poisson distribution arises from the Poisson distribution by setting $f(0, \vartheta) = 0$ and rescaling appropriately. One obtains the pmf $f(j, \vartheta) = e^{-\vartheta}\vartheta^j/(j! \cdot (1 - e^{-\vartheta})), j \geq 1$. The mle $\hat{\vartheta}_n$ is the solution of $\overline{X}_n = \vartheta(1 - e^{-\vartheta})^{-1}$ and satisfies (R1). Since (R2) and (R3) are satisfied by the Poisson distribution, the same holds for the positive Poisson.

Table 3 shows powers of the eidf test for sample sizes $n = 50$ and $n = 200$ for the three null distributions described above. Again, the entries are the percentages of 1000 Monte Carlo samples that resulted in rejection of H_0, and the bootstrap sample size was 200. The table can be compared with Table VII of [8]. There, the chi-squared test and the smooth test for the geometric distribution are used besides the new test of Epps.

For each of the three hypotheses, the eidf test behaves similarly as the probability generating function test of Epps [8], which was the best one in his study. Particularly in the case $n = 50$, the test of Epps has an egde on the eidf test for the alternatives under consideration.

4.3 Example. The final example illustrates the application of the eidf test for Poissonity. On $n = 243$ days, the number of trades in American Home Products Corporation during 1 : 00-1 : 30 p.m. were observed (see [8], p. 1477). On 33 days, there was no trade at all, on 55 days there was just one trade, etc. The vector of absolute frequencies can be found in the last but three line of the S-Plus implementation of the test below. Function `eidfstat` computes the eidf statistic according to (8) (with $\overline{X}_n - E_{\hat{\vartheta}_n}(X) = 0$). In function `eidftest`, an approximate p-value is computed by means of the parametric bootstrap. Program execution with $b = 1000$ bootstrap replications resulted in the p-value 0. Thus, the observed data are not compatible with the Poisson model.

Table 3. 5%-level powers of the eidf test for geometric, logseries and positive Poisson distribution

H_0: Geometric			H_0: Logseries		
Distribution	$n = 50$	$n = 200$	Distribution	$n = 50$	$n = 200$
$G(0.5)$	3.9	5.0	$L(.5)$	4.3	5.1
$NA(2, .5)$	13	56	$P^+(1.05)$	45	99
$NA(4, .5)$	69	100	$P^+(1.2)$	55	100
$NA(5, .2)$	41	98	$P^+(1.3)$	62	100
$NA(10, .2)$	92	100	$G^+(.4)$	23	90
$P(.3)$	13	52	$G^+(.33)$	35	97
$P(.5)$	26	90	$G^+(.25)$	43	100
$P(.7)$	43	100	$Z(0.7)$	72	100
$P \circ L(.5, .7)$	11	36	$Z(1.0)$	40	100
$P \circ L(.5, .5)$	19	68	$Z(1.3)$	66	100
$P \circ L(.5, .3)$	22	83	$Z(2.0)$	40	89
$P \circ L(.3, .5)$	10	39			
$P0(.2, 1.)$	24	89			
$P0(.2, 2.)$	59	100			
$P0(.5, 1.)$	4	9			

H_0: Positive Poisson		
Distribution	$n = 50$	$n = 200$
$P^+(3)$	4.4	4.2
$Bin^+(10, .2)$	46	99
$L(.3)$	16	47
$L(.5)$	42	92
$L(.7)$	87	100
$G^+(.4)$	84	100
$G^+(.33)$	95	100
$NA^+(.5, 2.0)$	100	100
$Z(1.0)$	100	100
$Z(2.0)$	74	100

```
eidfstat <- function(n, freq, cdf)   #computes eidf
statistic
{
  edf <- cumsum(freq)/n   #empirical df
  zsum <- cumsum(edf-cdf)
  tt <- sqrt(n)*max(abs(zsum))   #test statistic
  return(tt)
}

eidftest <- function(freq, b)   #parametric bootstrap
{
  n <- sum(freq)   #number of observations
  xmax <- length(freq) - 1   #largest observation
  xbar <- 0:xmax %*% freq / n   #sample mean
  cdf <- ppois(0:xmax, xbar)   #cumulative df
  tn <- eidfstat(n, freq, cdf)
  count <- 0
  for (i in 1:b) {
```

```
    yp <- rpois( n, xbar)   #Poisson sample of size n
    ybar <- mean(yp)
    freq <- tabulate(yp+1)   #absolute frequencies
    cdf <- ppois(0:max(yp), max(ybar,10^-9))
    tn2 <- eidfstat(n, freq, cdf)
    if (tn2 >= tn) count <- count + 1 }
  return(count/b)
}

# data:  number x of trades in American Home Products,
# x=0,1,...,12; n=243; freq:  absolute frequencies
freq <- c(33,55,68,38,20,11,8,7,2,0,0,0,1)
b <- 1000   #number of Bootstrap replications
pval <- eidftest(freq, b)
print(pval)   #p-value
```

5 Concluding remarks

This paper presented a new test statistic based on the empirical integrated distribution function. The test is an omnibus test, i.e. it is consistent, which seems to be rather exceptional for goodness-of-fit tests for discrete distributions. The limited simulation study in Section 4 indicates that the eidf test has high power and that the nominal level is maintained closely for samples of size 20 or larger.

The eidf test is not tailored to a special null distribution, but can be utilized for each hypothesis satisfying the regularity conditions of Section 3. Example 4.3 shows that test implementation is not very hard. Thereby, it is quite easy to modify the program for testing other hypotheses.

Acknowledgements. This work is part of the author's doctoral thesis at the University of Karlsruhe, written under the supervision of Professor Norbert Henze, whose guidance is gratefully appreciated.

References

[1] Barlow RE, Campo R (1975) Total time on test processes and applications to failure data analysis. In: Barlow RE, Fussell J, Singpurwalla ND (eds.) Reliability and fault tree analysis, SIAM: 451–481
[2] Bickel PJ, Freedman DA (1981) Some asymptotic theory for the bootstrap. Ann. Statist. 9:1196–1217
[3] Billingsley P (1968) Convergence of probability measures. J. Wiley & Sons, New York
[4] Burke, MD, Csörgő, M, Csörgő S, Révész P (1978) Approximations of the empirical process when parameters are estimated. Ann. Prob. 7:790–810
[5] Chandra M, Singpurwalla ND (1981) Relationships between some notions which are common to reliability theory and economics. Mathematics of Operations Research 6:113–121
[6] Choulakian V, Lockhart RA, Stephens MA (1994) Cramér-von Mises statistics for discrete distributions. Canad. J. Statist. 22:125–137
[7] Davydov YA, Lifshits MA (1985) Fibering method in some probabilistic problems. J. Sov. Math. 31:2796–2858
[8] Epps TW (1995) A test of fit for lattice distributions. Commun. Statist.-Theor. Meth. 24:1455–1479

[9] Glynn PW (1987) Upper bounds on Poisson tail distributions. Oper. Res. Lett. 6:9–14
[10] Heilmann W (1985) Transformations of claim distributions. Mitteilungen der Vereinigung schweiz. Versicherungsmathematiker 1:57–69
[11] Henze N (1996) Empirical-distribution-function goodness-of-fit tests for discrete models. Canad. J. Statist. 24:81–93
[12] Henze N, Klar B (1996) Properly rescaled components of smooth tests of fit are diagnostic. Austral. Journ. Statist. 38:61–74
[13] Hinderer K (1980) The integrated distribution function and its applications for some inventory problems with linear demand pattern, arbitrary demand distribution and without fixed ordering cost. Math. Operationsforsch. Statist., Ser. Optimization 11:299–310
[14] Johnson NL, Kotz S, Kemp AW (1992) Univariate discrete distributions. J. Wiley & Sons, New York
[15] Klar B (1998) Klassische und neue statistische Anpassungstests. Dissertation, Universität Karlsruhe
[16] Kocherlakota S, Kocherlakota K (1986) Goodness of fit tests for discrete distributions. Commun. Statist.-Theor. Meth. 15:815–829
[17] Müller A (1996) Ordering of risks: A comparative study via stop-loss transforms. Insurance: Mathematics and Economics 17:215–222
[18] Nakamura M, Pérez-Abreu V (1993) Use of an empirical probability generating function for testing a Poisson model. Canad. J. Statist. 21:149–156
[19] Rayner JCW, Best DJ (1989) Smooth tests of goodness of fit. & Oxford University Press, New York
[20] Shorak GR, Wellner JA (1986) Empirical Processes with applications to statistics. Wiley, New York
[21] Spinelli JJ, Stephens MA (1997) Cramér-von Mises tests of fit for the Poisson distribution. Canad. J. Statist. 25:257–268