The Annals of Probability
1991, Vol. 19, No. 3, 1311–1337

APPROXIMATE INDEPENDENCE OF DISTRIBUTIONS ON SPHERES AND THEIR STABILITY PROPERTIES[1]

BY S. T. RACHEV AND L. RÜSCHENDORF

University of California, Santa Barbara and University of Münster

Let ζ be chosen at random on the surface of the p-sphere in $\mathbb{R}^n$, $0_{p,n} := \{x \in \mathbb{R}^n: \sum_{i=1}^n |x_i|^p = n\}$. If $p = 2$, then the first k components $\zeta_1, \dots, \zeta_k$ are, for k fixed, in the limit as $n \to \infty$ independent standard normal. Considering the general case $p > 0$, the same phenomenon appears with a distribution F_p in an exponential class $\mathscr{E}$. F_p can be characterized by the distribution of quotients of sums, by conditional distributions and by a maximum entropy condition. These characterizations have some interesting stability properties. Some discrete versions of this problem and some applications to de Finetti-type theorems are discussed.

1. Introduction. In a recent paper, Diaconis and Freedman [9] proved that the first k components of a point uniformly distributed on the 2-sphere in $\mathbb{R}^n$ are close with respect to the variation distance to k independent standard normal random variables, the distance being of the exact order k/n. For the 1-sphere a similar result was obtained for the exponential distribution (restricting the sphere to $\mathbb{R}_+^n$).

In this paper we consider this problem for any $p > 0$ and a meaningful extension to $p = \infty$. Again the distribution of the first k components can be approximated by the distribution of k independent random variables with a distribution F_p. The class of distributions F_p arising in this way is an exponential class of distributions connecting the exponential distribution ($p = 1$) with the normal distribution ($p = 2$) and the uniform distribution ($p = \infty$).

While for $p = 1, 2, \infty$ the meaning of the notion "uniform distribution on the p-sphere" is obvious from invariance considerations, this is not the case for other p values. In this paper we call the distribution arising from a disintegration of the Lebesgue measure λ^n on the spheres the "uniform distribution on the sphere." This choice of uniform distribution is well motivated by statistical mechanics. It gives the right measure in Liouville's theorem. An alternative and interesting case would be to consider the distribution on the p-sphere corresponding to the geometric surface measure. The geometric surface measure coincides with our uniform distribution only for the cases $p = 1, 2, \infty$. Its treatment seems to need a completely different proof than the proof for the uniform distribution given in this paper.

Our distributions F_p can be characterized also in some other ways as, for example, the maximum entropy distributions with fixed pth moment equal to

Received September 1988; revised April 1990.
[1]Research supported by a DFG Grant and NATO Grant CRG900798.
AMS 1980 *subject classifications*. Primary 60E05; 62B20.
Key words and phrases. de Finetti's theorem, characterization of distributions, stability.

one. There are some immediate applications to de Finetti-type theorems. We consider also some discrete versions of the problem and we investigate the stability properties of the characterizations of F_p in the final section. By the method of probability metrics we also obtain some quantitative results for the inverse characterizations.

2. Characterizations by beta-distributions. Motivated by the paper of Diaconis and Freedman [9] we started with the following question: Let $\zeta_1, \zeta_2, \dots$ be a sequence of iid positive r.v.'s with d.f. F satisfying the normalization $E\zeta_1^p = 1$, $\infty > p > 0$, and define

$$X_{k,n,p} = \sum_{j=1}^{k} \zeta_j^p \Big/ \sum_{j=1}^{n} \zeta_j^p, \qquad 1 \le k \le n,\ n \in \mathbb{N}. \tag{2.1}$$

Does there exist a (unique?) d.f. $F = F_p$ such that $X_{k,n,p}$ has a beta $B(k/p, (n-k)/p)$-distribution for any $k \le n$, $n \in \mathbb{N}$? It is well known that F_1 is the standard exponential distribution and F_2 is the absolute value of a standard normal r.v. (see, e.g., Cramér [4], Section 18 and Diaconis and Freedman [9]).

THEOREM 2.1. *For any $0 < p < \infty$, there exists exactly one distribution $F = F_p$, such that for all $k \le n$, $n \in \mathbb{N}$, $X_{k,n,p}$ has a $B(k/p, (n-k)/p)$-distribution. F_p has the density*

$$f_p(x) = \frac{p^{1-1/p}}{\Gamma(1/p)} \exp\left(-\frac{x^p}{p}\right), \qquad x \ge 0. \tag{2.2}$$

PROOF. Let the r.v.'s $(\zeta_i)_{i\in\mathbb{N}}$ have the common density f_p. Then $f_{\zeta_1^p}(x) = 1/p^{1/p}\Gamma(1/p)x^{1/p-1}e^{-(x/p)}$, $x \ge 0$, is the $\Gamma(1/p, 1/p)$-density and hence $\sum_{i=1}^k \zeta_i^p$ is $\Gamma(1/p, k/p)$-distributed. Usual calculations show that

$$f_\kappa(x) = B\left(\frac{k}{p}, \frac{n-k}{p}\right) \frac{x^{k/p-1}}{(x+1)^{n/p}}, \qquad x > 0,$$

where $\kappa = \sum_{i=1}^k \zeta_i^p / \sum_{i=k+1}^n \zeta_i^p$, which leads to the $B(k/p, (n-k)/p)$ distribution of $X_{k,n,p}$.

On the other hand, assuming that $X_{1,n,p}$ has a $B(1/p, (n-1)/p)$-distribution for all $n \in \mathbb{N}$, by the (SLLN) $nX_{1,n,p} \to \zeta_1^p$ a.s. Furthermore, the density of $(nX_{1,n,p})^{1/p}$, given by

$$\frac{\Gamma(n/p)}{\Gamma(1/p)\Gamma((n-1)/p)} \left(\frac{x^p}{n}\right)^{-1+1/p} \left(1 - \frac{x^p}{n}\right)^{(n-1)/p-1} \frac{p}{n} x^{p-1}, \qquad 0 < x,$$

converges pointwise to $f_p(x)$, since $(p/n)^{1/p}\Gamma(n/p)/\Gamma((n-1)/p) \to 1$ as $n \to \infty$ (see, e.g., Abramowitz and Stegun [1], page 257). Thus $f_{\zeta_1} = f_p$ as required. □

REMARK 2.1. We consider throughout this paper the case of nonnegative random variables. One obvious extension is to the case that $\tilde{\zeta}_1, \tilde{\zeta}_2, \dots$ are iid r.v.'s on the whole real line satisfying the conditions $E\tilde{\zeta}_1 = 0$, $E|\tilde{\zeta}_1|^p = 1$. Then

$$\tilde{X}_{k,n,p} = \sum_{j=1}^{k} |\tilde{\zeta}_j|^p \Big/ \sum_{j=1}^{n} |\tilde{\zeta}_j|^p$$

is $B(k/p, (n-k)/p)$-distributed iff the density $\tilde{f}_p$ of $\tilde{\zeta}_1$ satisfies $\tilde{f}_p(x) + \tilde{f}_p(-x) = 2 f_p(|x|)$. In this way, one gets for $p = 2$ the normal distribution and for $p = 1$ the Laplace distribution. Uniqueness can be obtained by the additional assumption of symmetry of F.

To get a meaningful result for $p = \infty$, let β be a $B(k/p, (n-k)/p)$-distributed r.v. and define $\gamma_{k,n,p} =: \beta^{1/p}$; then $\gamma_{k,n,p}$ has a density given by

$$f_{\gamma_{k,n,p}}(x) = B\left(\frac{k}{p}, \frac{n-k}{p}\right) p x^{k-1}(1 - x^p)^{(n-1)/p}, \qquad 0 \le x \le 1.$$

Let $\gamma_{k,n,\infty}$ be the weak limit of $\gamma_{k,n,p}$ as $p \to \infty$, that is,

$$P(\gamma_{k,n,\infty} \le x) = \begin{cases} \dfrac{n-k}{n} x^k & \text{if } 0 \le x < 1, \\ 1 & \text{if } x \ge 1. \end{cases} \tag{2.3}$$

In the sequel F_∞ stands for the uniform distribution on $[0, 1]$.

THEOREM 2.2. *Let $\zeta_1, \zeta_2, \dots$ be a sequence of positive iid r.v.'s and let $X_{k,n,\infty} = \bigvee_{i=1}^{k} \zeta_i / \bigvee_{i=1}^{n} \zeta_i$ ($\bigvee \zeta_i :=$ max ζ_i). Then $X_{k,n,\infty}$ and $\gamma_{k,n,\infty}$ are equally distributed for any $k \le n$, $n \in \mathbb{N}$, iff ζ_1 is uniformly distributed on* $[0, 1]$.

PROOF. Assuming that ζ_1 is F_∞-distributed, the d.f. of $X_{k,n,\infty}$ has the form $P(X \le x(X \vee Y))$, where X and Y are independent with $F_X(t) = t^k$, $F_Y(t) = t^{n-k}$, $0 \le t \le 1$. Therefore, for $0 \le x \le 1$,

$$\begin{aligned} F_{X_{k,n,\infty}}(x) &= \int_0^x P(t \le x(t \vee Y))\, dt^k \\ &= \int_0^x P(t \le xY, Y > t)\, dt^k + \int_0^x P(t \le xt, Y \le t)\, dt^k \\ &=: I_1(x) + I_2(x). \end{aligned}$$

Now $I_1(x) = (1 - k/n)x^k$ for $x \in [0, 1]$ and $I_2(x) = 0$ for $0 < x < 1$, $I_2(1) = k/n$. This implies that $X_{k,n,\infty}$ has a distribution given by (2.3).

On the other hand, if $X_{1,n,\infty}$ has the same distribution as $\gamma_{1,n,\infty}$, then letting $n \to \infty$, the distribution of $\bigvee_{i=1}^{n} \zeta_i$ converges weakly to 1 and therefore, the limit of $F_{X_{1,n,\infty}}(x)$ is x for all $x \in [0, 1]$. □

3. Characterization by conditional distributions and maximum entropy. Let $S_{p,s,n} := \{x \in \mathbb{R}_+^n: \sum_{i=1}^n x_i^p = s\}$ denote the p-sphere of radius s in $\mathbb{R}_+^n$, $0 < p < \infty$. Let for random variables $X, Y, P^{X|Y=s}$ denote the regular conditional distribution of X given $Y = s$ and let P^X denote the distribution of X.

For $\mu = \lambda_+^n$, let $\mu_s := \mu(\cdot|\sum_{i=1}^n x_i^p = s)$ denote a disintegration w.r.t. the p-spheres, then $\mu = \int \mu_s \times \mu^{S_{n,p}}(ds)$, where $S_{n,p}(x) = \sum x_i^p$. We call μ_s the *uniform distribution* on the p-sphere of radius s (up to a norming constant) throughout this paper. Let $U = \{t \in \mathbb{R}_+^{n-1}; \|t\|_p < 1\}$ and let $\phi: U \times \mathbb{R}_+ \to \mathbb{R}_+^n$, $\phi(t,s) = (st, s(1-\|t\|_p^p)^{1/p})$. Then for any measurable function $f: \mathbb{R}_+^n \to \mathbb{R}_+$, we have

$$\begin{aligned}\int_{\mathbb{R}_+^n} f(x)\,d\lambda^n(x) &= \int_{U\times\mathbb{R}_+} f(\phi(t,s)) \det D\phi(t,s)\,dt\,ds \\ &= \int_{\mathbb{R}_+}\int_U f\Big(st, s\big(1-\|t\|_p^p\big)^{1/p}\Big)s^{n-1}\big(1-\textstyle\sum t_i^p\big)^{1/p-1}\,dt\,ds,\end{aligned} \tag{3.1}$$

with $D\phi$ the Jacobian. On the other hand, for the surface measure S_s on $M_s = \{x \in \mathbb{R}_+^n; \|x\|_p = s\}$, we have:

$$\begin{aligned}\int_{M_s} f(x)\,dS_s(x) &= \int_{\tilde{M}_s} f(t, F(t))\sqrt{1+\|\nabla F(t)\|^2}\,dt \\ &= \int_U f\Big(st, s\big(1-\|t\|_p^p\big)^{1/p}\Big)s^{n-1}h(t)\,dt,\end{aligned} \tag{3.2}$$

where $\tilde{M}_s = \{t \in \mathbb{R}_+^{n-1}; \|t\|_p < s\}$, $F(t) = (s^p - \sum t_i^p)^{1/p}$ for $t \in \tilde{M}_s$ and $h(t) = [1 + (\sum t_i^{2p-2})(1-\|t\|_p^p)^{(2/p)-2}]^{1/2}$. Comparing (3.1) and (3.2), we see that the uniform distribution μ_s coincides with the surface measure S_s only for $p = 1, 2, \infty$, but is generally different.

LEMMA 3.1. *Let $\zeta_1, \dots, \zeta_n$ be iid according to F_p, where $0 < p < \infty$. Then the regular conditional distribution for $(\zeta_1, \dots, \zeta_n)$ given*

$$\sum_{i=1}^n \zeta_i^p = s, \qquad P_s = P^{(\zeta_1,\dots,\zeta_n)|\sum_{i=1}^n \zeta_i^p = s}$$

is uniform on the p-sphere of radius s, $S_{p,s,n}$.

PROOF. Let $\mu = \lambda_+^n$ be the Lebesgue measure on $\mathbb{R}_+^n$ and let $\mu_s := \mu(\cdot|\sum x_i^p = s)$ be the uniform distribution on $S_{p,s,n}$ (up to a norming constant) and $\mu = \int \mu_s \times \mu^{S_{n,p}}(ds)$, where $S_{n,p}(x) = \sum_{i=1}^n x_i^p$. For any probability measure $P = f\mu$ with density f w.r.t. μ the relation $P = \int P_s \times P^{S_{n,p}}(ds)$ holds similarly. From some well-known formulas for conditional distributions, one gets

$$h(s) = \frac{dP^{S_{n,p}}}{d\mu^{S_{n,p}}}(s) = \int f(y)\mu_s(dy).$$

Furthermore,

$$P_s \ll \mu_s[\mu^{S_{n,p}}] \quad \text{and} \quad \frac{dP_s}{d\mu_s}(x) = \frac{f(x)}{h(s)} = \frac{f(x)}{h(S_{n,p}(x))},$$

μ_s a.s. for x with $S_{n,p}(x) = s$. Applying this to $f = f_p^{(n)}$, $f_p(x) = ce^{-x^p/p}$, $x \geq 0$, one gets that

$$\frac{dP_s}{d\mu_s}(x) = \frac{\prod_{i=1}^n f_p(x_i)}{h(s)} = c^n \frac{e^{-s/p}}{h(s)}$$

is constant on $S_{p,s,n}$ and therefore P_s is the uniform distribution on $S_{p,s,n}$ for almost all s w.r.t. the distribution $P^{S_{n,p}}$. □

The corresponding result for $p = \infty$ is well known.

LEMMA 3.2. *If $\zeta_1, \ldots, \zeta_n$ are iid, F_∞-distributed, then*

$$P_s = P^{(\zeta_1, \ldots, \zeta_n) \mid \vee_{i=1}^n \zeta_i = s}$$

is uniform on $S_{\infty,s,n} = \{x \in \mathbb{R}_+^n \colon \vee_{i=1}^n x_i = s\}$ for almost all (a.a.) $s \in [0,1]$.

The converse of Lemma 3.1 can be formulated in an even stronger form.

LEMMA 3.3. *Let $X_i = \zeta_i^p$, $i = 1, \ldots, n$, with (ζ_i) iid distributed according to F_p, $0 < p < \infty$. If $Y_1, \ldots, Y_n$ are iid, $EY_1 = 1$ and if for some $n \in \mathbb{N}$, $n \geq 2$,*

$$P^{X_1 \mid \Sigma_{i=1}^n X_i = s} = P^{Y_1 \mid \Sigma_{i=1}^n Y_i = s} \quad \text{for } P^{\Sigma X_i} \text{ a.a. } s, \tag{3.3}$$

then Y_1 has the same distribution as X_1.

PROOF. For the proof we use Theorem 1 of Berk [3] stating: If X, Y are independent random variables such that $H(x|s) = P(X \leq x | X + Y = s)$ has a Lebesgue density $h(\cdot|s)$ on an interval $I \subset \mathbb{R}^1$, then X, Y have Lebesgue densities f, g and any other independent pair X^*, Y^* with the same conditional distribution has densities f^*, g^* such that $f^*(x)g^*(y) = e^{\alpha x + \alpha y - c} f(x)g(y)$ a.s. for some $\alpha \in I$.

Consider now at first the case $n = 2$. Then

$$f_{X_1}(x) = f_{\zeta_1^p}(x) = \frac{1}{p^{1/p}\Gamma(1/p)} x^{1/p - 1} e^{-x/p}, \qquad x \geq 0.$$

If $P^{Y_1 \mid Y_1 + Y_2 = s} = P^{X_1 \mid X_1 + X_2 = s}$, then by Berk's theorem, Y_1 has a density g with $g(x) = ce^{\alpha x} f(x)$. Therefore, $g(x) = \tilde{c} x^{1/p - 1} e^{-x(1/p - \alpha)}$, $x \geq 0$, is the density of a $\Gamma(1/p, (1/p) - \alpha)$-distribution. From the normalization $EY_1 = 1$ follows

$$\int x g(x)\, dx = \frac{1}{p} \frac{1}{(1/p) - \alpha} = \frac{1}{1 - p\alpha} = 1,$$

that is, $\alpha = 0$, implying Lemma 3.3 for the case $n = 2$.

In the case $n \geq 2$ let $f \sim P^{X_1}$ and let $g \sim P^{\Sigma_{i=1}^n X_i}$. Then we can argue by Berk's theorem as in the case $n = 2$. □

LEMMA 3.4. *If $\eta_1, \dots, \eta_n$ are iid with support* $[0,1]$ *and if for some $n \in \mathbb{N}$, $n \geq 2$,*

$$P^{\eta_1 \mid \vee_{i=1}^n \eta_i = s} = P^{\zeta_1 \mid \vee_{i=1}^n \zeta_i = s}, \qquad \forall\, s,$$

with (ζ_i) as in Lemma 3.2, *then $P^{\eta_1} = P^{\zeta_1}$.*

PROOF. For the proof note that in the case $n = 2$, $P^{\zeta_1 \mid \zeta_1 \vee \zeta_2 = s} = \frac{1}{2}\varepsilon_{\{s\}} + \frac{1}{2}U_{[0,s]}$. Therefore P^{η_1} has a density f and $f(x) = 1$ for all $x \in [0,1)$. □

If (as in Remark 2.1) we consider p-spheres in $\mathbb{R}^n$ defined by $\{x \in \mathbb{R}^n\colon \Sigma |x_i|^p = s\}$, then for real r.v's Y_i satisfying (3.3) we obtain similarly that $|Y_1|$ has the same distribution as X_1.

Since $P^{\zeta_1^p \mid \Sigma_{i=1}^n \zeta_i^p = s}$ uniquely determines $P^{\zeta_1 \mid \Sigma \zeta_i^p = s}$ for $\zeta_i \geq 0$, we obtain as a corollary:

THEOREM 3.4. *If $0 < p \leq \infty$ and if $\zeta_1, \dots, \zeta_n$ are iid on $\mathbb{R}_+$ with $E\zeta_1^p = 1$ for $p < \infty$ and with support* $[0,1]$ *if $p = \infty$, then $P^{(\zeta_1, \dots, \zeta_n) \mid \Sigma \zeta_i^p = s}$ is uniform on $S_{p,s,n}$ for a.a. s iff $F_{\zeta_1} = F_p$.*

The conditional uniformity suggests the following third characterization.

THEOREM 3.5. (a) *For $0 < p < \infty$, F_p is the distribution on $\mathbb{R}_+$ with maximum entropy in the class of all distributions F on $\mathbb{R}_+$ with $\int x^p\, dF(x) = 1$.*

(b) *For $p = \infty$, F_∞ is the distribution on $\mathbb{R}_+$ with maximum entropy and with support in* $[0,1]$.

PROOF. (a) The minimum discrimination information theorem of Kullback [12] (cf. also [5] and [18]) implies that necessary and sufficient conditions for maximum entropy in the class of distributions with given pth moment are the conditions $P \ll \lambda_+^1$ and $\ln(dP/d\lambda_+^1) = a + bx^p$, which by the normalization conditions imply that P has the distribution function F_p.

(b) The proof is obvious. □

4. Asymptotic independence and de Finetti's theorem. This section is inspired by the paper of Diaconis and Freedman [9]. In this paper the main two continuous examples are the following: Let ζ be chosen at random on the surface of the 2-sphere $O_{2,n,n} := \{x \in \mathbb{R}^n\colon \Sigma x_i^2 = n\}$ of radius n (resp., the simplex $S_{1,n,n} := \{x \in \mathbb{R}_+^n\colon \Sigma x_i = n\}$). Then $\zeta_1, \dots, \zeta_k$ are for k fixed, in the limit as $n \to \infty$ independent standard normals (resp., exponentials). Diaconis and Freedman obtained a right order bound on the variation distance between the law of $(\zeta_1, \dots, \zeta_k)$ and the law of k independent standard normals (resp., exponentials). We will extend Diaconis' and Freedman's results considering ζ being chosen at random on the surface of the 'p-sphere' $O_{p,n,n} = \{x \in \mathbb{R}^n: \sum_{i=1}^n |x_i|^p = n\}$ for $0 < p < \infty$ and $O_{\infty,n,n} = \{x \in \mathbb{R}^n, \bigvee_{i=1}^n |x_i| = n\}$ (respectively, the positive part of $O_{p,n}$, namely $S_{p,n,n} = \{x \in O_{p,n,n}; x_i \geq 0\}$). From the considerations in Sections 2 and 3, it will be sufficient in the following to restrict to the positive p-spheres $S_{p,n,n}$. We start with the case $p = \infty$. Let $\zeta_1, \dots, \zeta_n$ be iid $U(0,1)$-distributed. Then by Lemma 3.2, $(\zeta_1, \dots, \zeta_n)$ is conditionally, given $\bigvee_{i=1}^n \zeta_i = s$, uniform on $S_{\infty,s,n}$. Let $P_\sigma^{n,\infty}$ for $\sigma > 0$ be the law of$(\sigma\zeta_1, \dots, \sigma\zeta_n)$ and let $Q_{n,s,k}^{(\infty)}$ be the law of $(\eta_1, \dots, \eta_k)$, where $\eta = (\eta_1, \dots, \eta_n)$ is uniform on $S_{\infty,s,n}$. In the next theorem we shall evaluate the deviation between $Q_{n,s,k}^{(\infty)}$ and $P_s^{k,\infty}$ in terms of the total variation distance $\|Q_{nsk}^{(\infty)} - P_s^{k,\infty}\| = \sup_{A \in \mathscr{B}^k} |Q_{nsk}^{(\infty)}(A) - P_s^{k,\infty}(A)|$, $\mathscr{B}^k$ being the Borel sets in $\mathbb{R}^k$.

THEOREM 4.1.

$$\|Q_{nsk}^{(\infty)} - P_s^{k,\infty}\| = \frac{k}{n}. \tag{4.1}$$

PROOF. Without loss of generality, we may assume $s = 1$. Let $\tilde{Q}$ be the law of $\eta_1 \vee \cdots \vee \eta_k$ determined by Q_{nsk} and $\tilde{P}$ be the law of $\zeta_1 \vee \cdots \vee \zeta_k$. Then with

$$\gamma_{k,n,\infty} = \bigvee_{i=1}^k \zeta_i \Big/ \bigvee_{i=1}^n \zeta_i,$$

$\tilde{Q} = P^{\gamma_{k,n,\infty}}$ and $\tilde{Q}$ has a d.f. given by (2.3). On the other hand, $\tilde{P}((-\infty, x]) = x^k$, $0 \leq x \leq 1$. Hence $\tilde{Q} = ((n-k)/n)\tilde{P} + (k/n)\delta_1$ is the mixture of $\tilde{P}$ and δ_1. Since $\bigvee_{i=1}^n X_i$ is sufficient for $\{\tilde{Q}, \tilde{P}\}$, it follows from the well-known sufficiency theorem (cf. Csiszar [5]) that $\|Q_{n,1,k}^{(\infty)} - P_1^{k,\infty}\| = \|\tilde{Q} - \tilde{P}\| = k/n\|\delta_1 - \tilde{P}\| = k/n$. $\square$

Let C_n be the class of distributions of $X = (X_1, \dots, X_n)$ on $\mathbb{R}_+^n$ which share with the iid uniforms the property that given $M := \bigvee_{i=1}^n X_i = s$, the conditional joint distribution of X is uniform on $S_{\infty,s,n}$.

Clearly, $P_\sigma^{n,\infty} \in C_n$. Set $P_{\mu,n} = \int P_\sigma^{n,\infty}\mu(d\sigma)$ for any probability μ on $(0,\infty)$. As a consequence of Theorem 4.1, we get the finite form of the de Finetti theorem:

COROLLARY 4.2. *If $P \in C_n$, then there is a μ such that for all $k < n$,*

$$\|P_k - P_{\mu k}\| \leq \frac{k}{n}, \tag{4.2}$$

where P_k is the P-law of the first k-coordinates $(X_1, \dots, X_k)$.

PROOF. Define

$$\mu = P^{\bigvee_{i=1}^n X_i},$$

the distribution of $\bigvee_{i=1}^n X_i$, then $P_k = \int Q_{n,s,k}^{(\infty)}\mu(ds)$ and, therefore, $\|P_k - P_{\mu k}\| = \|\int(Q_{nks}^{(\infty)} - P_s^{k,\infty})\mu(ds)\| \le \int\|Q_{nsk}^{(\infty)} - P_s^{k,\infty}\|\mu(ds) = k/n$. □

In particular, one gets the infinite de Finetti-type characterization of scale mixtures of iid uniform variables (cf. Diaconis and Freedman [8], Example 2.5).

COROLLARY 4.3. *Let P be a probability on $\mathbb{R}_+^\infty$ with P_n being the P-law of the first n coordinates. Then P is a uniform scale mixture of iid uniform variables, iff $P_n \in C_n$ for every n.*

Following the same idea we will now consider the case $p \in (0,\infty)$. Let $\zeta_1, \zeta_2, \dots$ be iid r.v.'s with d.f. F_p given by Theorem 2.1. Then by Lemma 3.1, the conditional distribution of $(\zeta_1,\dots,\zeta_n)$ given $\sum_{i=1}^n \zeta_i^p = s$ is $Q_{n,s,n}^{(p)}$, where $Q_{n,s,k}^{(p)}$ is the distribution of the first k coordinates of a random vector $(\eta_1,\dots,\eta_n)$ uniformly distributed on the p-sphere of radius s, $S_{p,s,n}$ in $\mathbb{R}_+^n$. Let $P_\sigma^{n,p}$ be the law of the vector $(\sigma\zeta_1,\dots,\sigma\zeta_n)$. The next result shows that $Q_{n,s,k}^{(p)}$ is close to $P_{(s/n)^{1/p}}^{k,p}$ w.r.t. the total variation distance, that is, the distribution of $(\eta_1,\dots,\eta_k)$ is close to the distribution of

$$\left((s/n)^{1/p}\zeta_1,\dots,(s/n)^{1/p}\zeta_k\right).$$

In comparison with the case $p = \infty$ in Theorem 4.1, we have to use the rescaling $(s/n)^{1/p}$ for the approximating distribution $P_s^{k,p}$ here.

THEOREM 4.4. *Let $0 < p < \infty$, then for $n > k + p$*

$$\|Q_{n,s,k}^{(p)} - P_{(s/n)^{1/p}}^{k,p}\|$$

$$\le \Bigg[kp + \frac{k}{2} + \left(\frac{e^{1/12}}{12} + 1\right)p + \left(\frac{kp}{2} + \frac{p^2e^{1/12}}{12}\left(1 + \frac{k}{2n}\right)\right)\frac{1}{n-k} + \frac{pe^{1/12}}{24}\frac{k}{n}\Bigg]\frac{1}{n-k-p}. \tag{4.3}$$

PROOF. Obviously, $\|Q_{nsk}^{(p)} - P_{(s/n)^{1/p}}^{k,p}\| = \|Q_{n,n,k}^{(p)} - P_1^{k,p}\|$, so it sufficses to take $s = n$.

Let $\tilde{Q}_k$ be the $Q_{n,n,k}^{(p)}$-law of $\eta_1^p + \cdots + \eta_k^p$ and $\tilde{P}_k$ be the $P_1^{k,p}$-law of $\zeta_1^p + \cdots + \zeta_k^p$. Then $\|Q_{n,n,k}^{(p)} - P_1^{k,p}\| = \|\tilde{Q}_k - \tilde{P}_k\|$ as in the proof of Theorem 4.1. By Lemma 3.1, we consider $Q_{n,n,k}^{(p)}$ as the law of $(\zeta_1/R,\dots,\zeta_k/R)$, where $R^p = (1/n)\sum_{i=1}^n \zeta_i^p$. So $\tilde{Q}_k$ is the law of

$$\sum_{i=1}^k \left(\frac{\zeta_i}{R}\right)^p = n\frac{\sum_{i=1}^k \zeta_i^p}{\sum_{i=1}^n \zeta_i^p},$$

which by Theorem 2.1 is the law of $nB(k/p,(n-k)/p)$, that is, $\tilde{Q}_k$ has a density

$$(4.4)\qquad f(x) = \frac{1}{n}\frac{\Gamma(n/p)}{\Gamma(k/p)\Gamma((n-k)/p)}\left(\frac{x}{n}\right)^{k/p-1}\left(1-\frac{x}{n}\right)^{(n-k)/p-1}$$

for $0 \le x \le n$ and $f(x) = 0$ for $x > n$. On the other hand, $\tilde{P}_k$ has a gamma $(1/p, k/p)$-density

$$(4.5)\qquad g(x) = \frac{1}{p^{k/p}\Gamma(k/p)}e^{-x/p}x^{(k/p)-1} \quad \text{for } 0 \le x < \infty.$$

The following representation of the sup distance will be used further:

$$(4.6)\qquad \|\tilde{Q}_k - \tilde{P}_k\| = \int_0^\infty \max\left(0, \frac{f(x)}{g(x)} - 1\right)g(x)\,dx.$$

By (4.4) and (4.5), $f/g = Ah$, where

$$A = \left(\frac{p}{n}\right)^{k/p}\frac{\Gamma(n/p)}{\Gamma((n-k)/p)} \quad \text{and} \quad h(x) = e^{x/p}\left(1-\frac{x}{n}\right)^{((n-k)/p)-1}$$

for $x \in [0, n]$ and $h(x) = 0$ for $x > n$. We have

$$\log h(x) = \frac{x}{p} + \left(\frac{n-k}{p} - 1\right)\log\left(1 - \frac{x}{n}\right)$$

and $(\partial/\partial x)\log h(x) \ge 0$ iff $x \le k + p$. Hence, if $k + p \le n$,

$$(4.7)\qquad \log h(x) \le \frac{k+p}{p} + \left(\frac{n-k}{p} - 1\right)\log\left(1 - \frac{k+p}{n}\right).$$

We use the following consequence of the Stirling expansion of the gamma function (cf. Abramowitz and Stegun [1], page 257)

$$(4.8)\qquad \Gamma(x) = e^{-x}x^{x-1/2}(2\pi)^{1/2}e^{\theta/12x}, \qquad 0 < \theta < 1.$$

This implies that

$$A = \left(\frac{n}{n-k}\right)^{((n-k)/p)+1/2}e^{-k/p}\tilde{\theta},$$

with

$$\tilde{\theta} = e^{p/12((\theta_1/n)-\theta_2/(n-k))} \le e^{p/12n}.$$

Hence,

$$\begin{aligned}
Ah &\le e\left(\frac{n}{n-k}\right)^{((n-k)/p)+1/2}\left(\frac{n-k-p}{n}\right)^{((n-k)/p)-1}\tilde{\theta} \\
&= e\left(\frac{n-k-p}{n-k}\right)^{((n-k)/p)}\frac{n}{n-k-p}\left(\frac{n}{n-k}\right)^{1/2}\tilde{\theta} \\
&= e\left(1 - \frac{p}{n-k}\right)^{((n-k)/p)}\frac{n}{n-k-p}\left(\frac{1}{1-k/n}\right)^{1/2}\tilde{\theta}.
\end{aligned}$$

From Kumar and Pathak [13] (page 42), we use the following estimate:

$$(4.9)\qquad \sup_{0\le x<a}\left|e^{-x}-\left(1-\frac{x}{a}\right)^{a}\right|\le\frac{c}{a}\quad\text{with}\quad c=\sup_{0\le x<n}xe^{-x}=\frac{1}{e},\quad a>1,$$

implying that

$$\left|e\left(1-\frac{p}{n-k}\right)^{((n-k)/p)}-1\right|\le\frac{p}{n-k}.$$

Furthermore, we use the estimates

$$\left(1-\frac{k}{n}\right)^{-1/2}\le 1+\frac{k}{2n}\quad\text{and}\quad\tilde{\theta}\le e^{p/12n}\le 1+\frac{p}{12n}e^{1/12}$$

to obtain

$$Ah\le\left(1+\frac{p}{n-k}\right)\frac{n}{n-k-p}\left(1+\frac{k}{n}\right)\left(1+\frac{pe^{1/12}}{12n}\right),$$

implying that $Ah-1$ is bounded by the right-hand side of (4.3). □

REMARK 4.1. (a) Following the proof of Diaconis and Freedman [9], one can get, under the assumption that $p=p_1/p_2$ is rational, that is, $p_i\in\mathbb{N}$, the following bound:

$$(4.10)\qquad \begin{aligned}\|Q^{(p)}_{n,s,k}-P^{k,p}_{(s/n)^{1/p}}\|&\le\begin{cases}\dfrac{k+p}{n-k-p} & \text{if } \dfrac{k}{p}\in\mathbb{N},\ k+p<n,\\[2ex] \dfrac{mp_1+p}{n-p_1-p} & \text{if } k\in((m-1)p,mp),\end{cases}\\ &\le\frac{k+p_1+p}{n-k-p_1-p}\quad\text{for } k<n-p_1-p,\end{aligned}$$

which is for some cases better than the bound (4.3). Diaconis and Freedman [9] obtain in the case $p=2$ the bound $(k+3)/(n-k-3)$; in the case $p=1$, $(k+1)/(n-k+1)$.

(b) Similarly to Corollaries 4.2 and 4.3, one can consider the corresponding finite de Finetti-type theorems and the corresponding infinite characterization of scale mixtures of iid F_p distributed random variables.

There are some general results on the asymptotic independence of

$$P^{(X_1,\dots,X_k)\mid\Sigma_{i=1}^{n}X_i}$$

based on Edgeworth expansions but without explicit finite bounds (cf. Zabell [22], Diaconis and Freedman [9]). Diaconis and Freedman ([9]) have a precise result for exponential families. This implies the following theorem.

THEOREM 4.5. *Let $0 < p < \infty$ and $k, n \to \infty$ such that $k/n \to 0$, then the following holds*:

$$\sup_{s>0} \|Q^{(p)}_{n,s,k} - P^{k,p}_{(s/n)^{1/p}}\| = \tfrac{1}{4} E|1 - N^2_{0,1}| \frac{k}{n} + o\left(\frac{k}{n}\right), \tag{4.11}$$

where $N_{0,1}$ is a standard normal random variable.

PROOF. Let $\zeta_1, \zeta_2, \ldots$ be iid r.v.'s with d.f. F_p for some $0 < p < \infty$ and let $X_i = \zeta_i^p$, which are gamma $(1/p, 1/p)$ r.v.'s with density

$$h_p(x) = \frac{(1/p)^{1/p}}{\Gamma(1/p)} x^{(1/p)-1} e^{-x/p}, \qquad x > 0.$$

Consider the exponential family generated by h_p, that is,

$$\tilde{f}_\theta(x) = \frac{1}{a(\theta)} e^{\theta x} h_p(x) = \frac{(1/p)^{(1/p)-\theta}}{\Gamma((1/p) - \theta)} x^{(1/p)-1} e^{-x((1/p)-\theta)}, \qquad x > 0 \tag{4.12}$$

for $\theta \in (-\infty, 1/p) = \Theta$, that is, $\tilde{f}_\theta$ is a gamma $(1/p, (1/p) - \theta)$-distribution. If $\tilde{X}_1, \tilde{X}_2, \ldots$ are iid according to $\tilde{f}_\theta$, then the first moment

$$\mu(\theta) = E_\theta \tilde{X}_1 = \frac{1}{p} \frac{1}{1/p - \theta} = \frac{1}{1 - p\theta},$$

implying that $\mu(\theta) = s/n$ iff $\theta = (1 - (n/s)) \cdot 1/p$. It is easy to see that the exponential family $\mathscr{P} = \{P_\theta = \tilde{f}_\theta \mathbb{X}^1; \theta \in \Theta\}$ satisfies the four regularity conditions (6)–(9) of Theorem 10 of Diaconis and Freedman [9], that is:

(i) Θ is maximal.

(ii) $$\sup_{\theta \in \Theta} E_\theta \frac{\left(\tilde{X}_1 - \mu(\theta)\right)^4}{\sigma(\theta)^4} < \infty \quad \text{where } \sigma(\theta) = \operatorname{Var}_\theta\left(\tilde{X}_1\right).$$

(iii) Smoothness:

$$\sup_\theta \sup_{|t|>\delta} \left| \varphi_\theta\left(\frac{t}{\sigma(\theta)}\right) \right| < 1 \quad \text{where } \varphi_\theta(t) = E_\theta e^{it\tilde{X}_1}.$$

(iv) Integrability: $\sup_\theta \int |\varphi_\theta(t/\sigma(\theta))|^v \, dt < \infty$ for some $v \geq 1$.

By Theorem 10 of Diaconis and Freedman ([9]), with

$$Q_{n,s,k} = P_\theta^{(\tilde{X}_1, \ldots, \tilde{X}_k) | \sum_{i=1}^n \tilde{X}_i = s},$$

we obtain

$$\sup_{s>0} \|Q_{nsk} - P^{(\tilde{X}_1, \ldots, \tilde{X}_k)}_{(1-n/s) \cdot 1/p}\| = \frac{1}{4} \frac{k}{n} E|1 - N^2_{0,1}| + o\left(\frac{k}{n}\right). \tag{4.13}$$

Now we observe that the exponential centered distributions in (4.12) have the same conditional distributions as the noncentered distributions have (this is the main point of this proof):

$$(4.14)\qquad Q_{n,s,k} = P_\theta^{(\tilde X_1,\dots,\tilde X_k)\mid\sum_{i=1}^n \tilde X_i = s} = P^{(X_1,\dots,X_k)\mid\sum_{i=1}^n X_i = s}.$$

Furthermore, with $\theta_s = (1-(n/s))1/p$ we obtain that $P_{\theta_s}^{\sqrt[p]{\tilde X_1}}$ has the density

$$(4.15)\qquad f_{\sqrt[p]{\tilde X_1}}(x) = \tilde f_{\theta_s}(x^p)px^{p-1} = f_p\left(\left(\frac{n}{s}\right)^{1/p} x\right),$$

that is, $P_{\theta_s}^{\sqrt[p]{\tilde X_1}} = P^{1,p}_{(s/n)^{1/p}}$, the law of $(s/n)^{1/p}\zeta_1$, where $\zeta_1 \sim_d F_p$. The invariance of the sup distance w.r.t. monotonic transformations applied to $G(x) = (\sqrt[p]{x_1},\dots,\sqrt[p]{x_k})$ using $Q^{(p)}_{n,s,k} = Q^G_{n,s,k}$ implies the result. □

Let $C_{n,p}$ be the set of all probability measures on $(\mathbb{R}^n_+, \mathscr{B}^n_+)$ such that the conditional distribution given $\sum_{i=1}^n x_i^p = s$ is uniform for all $s > 0$. From theorem (12) of Diaconis and Freedman ([9]), the following corollary is immediate.

COROLLARY 4.6. *Let $k, n \to \infty$, $k/n \to 0$, let for $P \in C_{n,p}$, P_k be the distribution of the first k components and let $\mu_p = P^{\sum_{i=1}^n x_i^p}$ and $P_{\mu k} = \int P^{k,p}_{(s/n)^{1/p}}\, d\mu_p(s)$. Then $\sup_{P\in C_{n,p}}\|P_k - P_{\mu k}\|/\frac{1}{4}(k/n)E|1-N^2_{0,1}| \to 1$.*

5. Stability of the characterization of the exponential class $\{F_p, 0 < p \le \infty\}$. In Section 2 we have characterized F_p by means of the beta-distribution $B(k/p,(n-k)/p)$ of the corresponding quotients $X_{k,n,p} = \sum_{j=1}^n \zeta_j^p/\sum_{j=1}^n \zeta_j^p$, $0 < p < \infty$, respectively, by the distribution of $X_{k,n,\infty} = \bigvee_{i=1}^k \zeta_i/\bigvee_{i=1}^n \zeta_i$. In this section we want to show that this characterization remains stable in both directions, that is, the mapping from the distribution F of the ζ_i to the distribution of $X_{k,n,p}$ is continuous in both directions, and we shall give a quantitative version of this continuity. The method to prove this stability, which is formulated in terms of the distance ρ [cf. (5.1)], is to prove this uniform continuity in the first step for a different metric ζ for which this is easy to see. Then in the second step, we pass via some inequalities from ζ to ρ. The question of stability of the characterization of F_p by $X_{k,n,p}$ is motivated by the results of Sections 2 and 3, where the essential idea of the proofs was that the uniform distribution on the p-sphere $S_{p,s,n}$ is given by the distribution of $s^{1/p}(\zeta_1,\dots,\zeta_n)/(\sum_{i=1}^n \zeta_i^p)^{1/p}$.

We first consider a perturbed sequence $\tilde\zeta_1, \tilde\zeta_2,\dots$ of iid nonnegative r.v.'s with common d.f. $\tilde F_p$ close to F_p in the sense that the uniform distance

$$(5.1)\qquad \rho := \rho\left(\tilde\zeta_1,\zeta_1\right) = \rho\left(\tilde F_p, F_p\right)$$

[here $\rho(X,Y) = \sup_x|F_X(x) - F_Y(x)|$] is close to zero. The next theorem says that the distribution of $\tilde X_{k,n,p} = \sum_{i=1}^k \tilde\zeta_i^p/\sum_{i=1}^n \tilde\zeta_i^p$ is close to the beta $B(k/p,(n-k)/p)$-distribution w.r.t. the uniform distance. In the sequel, c denotes absolute constants which may be different in different places and

$c(\cdots)$ denotes quantities depending only on the arguments in the parentheses.

THEOREM 5.1. *For any* $0 < p < \infty$ *and* $(\tilde{\zeta}_i)$ *iid with* $E\tilde{\zeta}_1^p = 1$ *and* $\tilde{m}_\delta := E\tilde{\zeta}_1^{(2+\delta)p} < \infty$ $(\delta > 0)$, *the following estimate holds*:

$$\sup_{k,n} \rho\big(X_{k,n,p}, \tilde{X}_{k,n,p}\big) \le c(\delta, \tilde{m}_\delta, p)\rho^{\delta/(3(2+\delta))}. \tag{5.2}$$

PROOF. We start with the first claim:

CLAIM 1.

$$\begin{aligned}\rho\big(X_{k,n,p}, \tilde{X}_{k,n,p}\big) &\le \rho\left(\sum_{i=1}^k \zeta_i^p, \sum_{i=1}^k \tilde{\zeta}_i^p\right) + \rho\left(\sum_{i=k+1}^n \zeta_i^p, \sum_{i=k+1}^n \tilde{\zeta}_i^p\right)\\ &\le n\rho\big(\zeta_1, \tilde{\zeta}_1\big).\end{aligned} \tag{5.3}$$

To prove (5.3), observe that

$$X_{k,n,p} = \frac{X_1}{X_1 + X_2}, \qquad \tilde{X}_{k,n,p} = \frac{\tilde{X}_1}{\tilde{X}_1 + \tilde{X}_2},$$

where

$$X_1 = \sum_{i=1}^k \zeta_i^p, \qquad X_2 = \sum_{i=k+1}^n \zeta_i^p,$$

$$\tilde{X}_1 = \sum_{i=1}^k \tilde{\zeta}_i^p, \qquad \tilde{X}_2 = \sum_{i=k+1}^n \tilde{\zeta}_i^p.$$

Since $\varphi(t) = t/(1+t)$ is strictly monotone and $X_{k,n,p} = \varphi(X_1/X_2)$, we have that $\rho(X_{k,n,p}, \tilde{X}_{k,n,p}) = \rho(X_1/X_2, \tilde{X}_1/\tilde{X}_2)$. Choosing $\tilde{X}_1^* =_d X_1$, $\tilde{X}_1^*$ independent of $\tilde{X}_2$, we obtain

$$\begin{aligned}\rho\left(\frac{X_1}{X_2}, \frac{\tilde{X}_1}{\tilde{X}_2}\right) &\le \rho\left(\frac{X_1}{X_2}, \frac{\tilde{X}_1^*}{\tilde{X}_2}\right) + \rho\left(\frac{\tilde{X}_1^*}{\tilde{X}_2}, \frac{\tilde{X}_1}{\tilde{X}_2}\right)\\ &= \sup_{x\ge 0}\left|\int_0^\infty \left[P\left(\frac{y}{X_2} \le x\right) - P\left(\frac{y}{\tilde{X}_2} < x\right)\right] dF_{X_1}(y)\right|\\ &\quad + \sup_{x \ge 0}\left|\int_0^\infty \left[P\left(\frac{X_1}{y} \le x\right) - P\left(\frac{\tilde{X}_1}{y} \le x\right)\right] dF_{\tilde{X}_2}(y)\right|\\ &\le \int_0^\infty \sup_{x\ge 0}\left|P\left(X_2 \ge \frac{y}{x}\right) - P\left(\tilde{X}_2 \ge \frac{y}{x}\right)\right| dF_{X_1}(y)\\ &\quad + \int_0^\infty \sup_{x \ge 0}\big|P(X_1 \le xy) - P\big(\tilde{X}_1 \le xy\big)\big| dF_{\tilde{X}_2}(y)\\ &= \rho\big(X_1, \tilde{X}_1\big) + \rho\big(X_1, \tilde{X}_2\big).\end{aligned}$$

The second part of (5.3) follows from the smoothing inequality, $\rho(X+Z, Y+Z) \le \rho(X,Y)$ for Z independent of X, Y.

In order to estimate the ρ-distance between the sums of ζ_i^p's and $\tilde{\zeta}_i^p$'s, we use the following two-stage approach. First, choose an appropriate metric for comparison of the sums. In our case this will be the ζ-metric

$$\zeta(X,Y) = \sup\{|Ef(X) - Ef(Y)|: |f''(x)| \le 1 \text{ a.e.}\}.$$

One can easily check (see, e.g., [23]) that

$$\zeta\left(\frac{1}{\sqrt{n}}\sum_{i=1}^n Z_i, \frac{1}{\sqrt{n}}\sum_{i=1}^n \tilde{Z}_i\right) \le \zeta(Z_1, \tilde{Z}_1) \tag{5.4}$$

for independent, identically distributed r.v.'s (Z_i) and $(\tilde{Z}_i)$. In the second stage of this metric approach, we compare the ζ-metric with the metric ρ. By this comparison we obtain an estimate for ρ via the estimate for ζ.

CLAIM 2. Let $n > p$, $E\zeta_1^p = E\tilde{\zeta}_1^p$, $\sigma_p^2 = \mathrm{Var}(\zeta_1^p)$, $\tilde{\sigma}_p^2 = \mathrm{Var}(\tilde{\zeta}_1^p) < \infty$. Then

$$\begin{aligned}\rho\left(\sum_{i=1}^n \zeta_i^p, \sum_{i=1}^n \tilde{\zeta}_i^p\right) &\le 3\sigma_p^{2/3}\left(2\pi\left(1-\frac{p}{n}\right)\right)^{-1/3} \\ &\quad\times \zeta^{1/3}\left(\frac{1}{\sqrt{n}}\sum_{i=1}^n Z_i, \frac{1}{\sqrt{n}}\sum_{i=1}^n \tilde{Z}_i\right),\end{aligned} \tag{5.5}$$

where

$$Z_i := \frac{\zeta_i^p - 1}{\sigma}, \qquad \tilde{Z}_i := \frac{\tilde{\zeta}_i^p - E\tilde{\zeta}_i^p}{\sigma}$$

and

$$\zeta(X,Y) = \int_{-\infty}^{\infty}\left|\int_{-\infty}^{x}(F_X(t) - F_Y(t))\,dt\right|dx.$$

PROOF. For any $n = 1, 2, \dots,$ we have

$$\rho\left(\sum_{i=1}^n \zeta_i^p, \sum_{i=1}^n \tilde{\zeta}_i^p\right) = \rho\left(\frac{1}{\sqrt{n}}\sum_{i=1}^n Z_i, \frac{1}{\sqrt{n}}\sum_{i=1}^n \tilde{Z}_i\right).$$

From Baxter and Rachev [[2], (2.8)], we have

$$\rho(X,Y) \le 3M^{2/3}(\zeta(X,Y))^{1/3}, \tag{5.6}$$

where $M = \sup_{x\in\mathbb{R}^1} f_X(x)$, and the density of X is assumed to exist. We have

$$f_{(1/\sqrt{n})\sum_{i=1}^n Z_i}(x) = \sigma_p\sqrt{n}\, f_{\sum_{i=1}^n \zeta_i^p}\left(\sqrt{n}\,\sigma_p x + 1\right),$$

$$f'_{\sum_{i=1}^n \zeta_i^p}(x) = \frac{1}{p^{n/p}\Gamma(n/p)}\left[\left(\frac{n}{p} - 1\right)x^{n/p-2}e^{-x/p} - \frac{1}{p}x^{n/p-1}e^{-x/p}\right] = 0$$

iff $n/p - 1 = (1/p)x$, that is, iff $n - p = x$. Thus for $n > p$, we get

$$f_{\sum_{i=1}^n \zeta_i^p}(x) \le \frac{p^{(n-p)/p}(n/p - 1)^{(n-p)/p} e^{-(n-p)/p}}{p^{n/p}((n/p) - 1)\left[((n/p) - 1)^{(n/p)-3/2} e^{-(n/p)+1}\sqrt{2\pi}\right]}$$

$$\left[\text{using } \Gamma(z) \ge z^{z-1/2} e^{-z}\sqrt{2\pi}\right].$$

This implies that

$$\begin{aligned}\sigma_p\sqrt{n}\, f_{\sum_{i=1}^n \zeta_i^p}(x) &\le \sigma_p \frac{\sqrt{n}}{p^{n/p}((n/p) - 1)^{1/2}\sqrt{2\pi}} = \sigma_p \frac{\sqrt{n}}{\sqrt{2\pi}\sqrt{n-p}} \\ &= \sigma_p\left(2\pi\left(1 - \frac{p}{n}\right)\right)^{-1/2}.\end{aligned} \tag{5.7}$$

(5.6) and (5.7) together imply (5.5).

CLAIM 3. If $m_\delta < \infty$, then

$$\zeta\left(Z_1, \tilde{Z}_1\right) \le c(\delta, m_\delta, p)\rho^{\delta/(2+\delta)}. \tag{5.8}$$

PROOF. For r.v.'s X, Y with $E(X - Y) = 0$,

$$\begin{aligned}\zeta(X, Y) &\le \int_{-\infty}^{\infty} |x||F_X(x) - F_Y(x)|\,dx \\ &\le N^2\rho(X, Y) + \frac{1}{2}EX^2 1_{\{|X| > N\}} + \frac{1}{2}EY^2 1_{\{|Y| > N\}} \\ &\le N^2\rho(X, Y) + \frac{1}{2N^\delta}\left(E|X|^{2+\delta} + E|Y|^{2+\delta}\right).\end{aligned}$$

Minimizing the right-hand side over $N > 0$, we get (5.8). Combining Claims 2 and 3 we get $\rho(\sum_{i=1}^n \zeta_i^p, \sum_{i=1}^n \tilde{\zeta}_i^p) \le c(\delta, m_\delta, p)\rho^{\delta/(3(1+\delta))}$ if $p/n < \frac{1}{2}$. From Claim 1 we then obtain

$$\rho\left(X_{k,n,p}, \tilde{X}_{k,n,p}\right) \le \begin{cases} 2p\rho & \text{if } p \ge \frac{n}{2}, \\ 2p\rho + c\rho^{(\delta/3(2+\delta))} & \text{if } p \ge \frac{k}{2}, p < \frac{n}{2}, \\ c\rho^{\delta/(3(2+\delta))} & \text{if } p < \frac{k}{2}, p < \frac{n}{2}, \end{cases} \tag{5.9}$$

which proves the theorem. □

REMARK 5.1. (a) Claim 1 of the proof of Theorem 5.1 remains also true for the total variation metric. But ρ seems to be the appropriate metric for this problem since ρ is related to the ideal metric ζ [see (5.6)] while the total variation metric is too strong to be estimated from above by ζ.

(b) In particular, (5.2) implies the following result on qualitative stability:

$$\text{If } \tilde{\zeta}_1 \to_D \zeta_1, \quad m_\delta < \infty, \quad \text{then } \tilde{X}_{k,n,} \to_D X_{k,n,}. \tag{5.10}$$

For the stability of the opposite direction we prove the following result.

THEOREM 5.2. *For any* $0 < p < \infty$ *and any iid sequences* $(\zeta_i), (\tilde{\zeta}_i)$ *with* $E\zeta_1^p = E\tilde{\zeta}_1^p = 1$ *and* $\zeta_1, \tilde{\zeta}_1$ *having continuous distribution functions we have*

$$\rho(\zeta_1, \tilde{\zeta}_1) \le \sup_{k,n} \rho\left(X_{k,n,p}, \tilde{X}_{k,n,p}\right). \tag{5.11}$$

PROOF. Denote $X_i = \zeta_i^p$, $\tilde{X}_i = \tilde{\zeta}_i^p$. Then

$$\begin{aligned}
&\sup_{k,n} \rho\left(\sum_{i=1}^k \zeta_i^p \Big/ \sum_{i=1}^n \zeta_i^p, \sum_{i=1}^k \tilde{\zeta}_i^p \Big/ \sum_{i=1}^n \tilde{\zeta}_i^p \right) \\
&\quad \ge \sup_n \rho\left(\frac{X_1}{\sum_{i=1}^n X_i}, \frac{\tilde{X}_1}{\sum_{i=1}^n \tilde{X}_i} \right) = \sup_n \rho\left(\frac{X_1}{(1/n)\sum_{i=2}^{n+1} X_i}, \frac{\tilde{X}_1}{(1/n)\sum_{i=2}^{n+1} \tilde{X}_i} \right) \\
&\quad \ge \rho\left(X_1, \tilde{X}_1\right) - \limsup \rho\left(\frac{X_1}{(1/n)\sum_{i=2}^{n+1} X_i}, X_1 \right) \\
&\qquad - \limsup \rho\left(\frac{\tilde{X}_1}{(1/n)\sum_{i=2}^{n+1} \tilde{X}_i}, \tilde{X}_1 \right).
\end{aligned}$$

By the SLLN and the assumption $EX_1 = E\tilde{X}_1 = 1$,

$$\frac{X_1}{(1/n)\sum_{i=2}^{n+1} X_i} \to X_1 \text{ a.s.}, \qquad \frac{\tilde{X}_1}{(1/n)\sum_{i=2}^{n+1} \tilde{X}_i} \to \tilde{X}_1 \text{ a.s.} \tag{5.12}$$

Since X_1 and $\tilde{X}_1$ have continuous d.f.'s, the convergence in (5.12) is valid w.r.t. the uniform distance ρ. Hence $\sup_{k,n} \rho(X_{k,n,p}, \tilde{X}_{k,n,p}) \ge \rho(X_1, \tilde{X}_1) = \rho(\zeta_1, \tilde{\zeta}_1)$ as required. □

REMARK 5.2. By (5.3) and (5.11), we have that

$$\begin{aligned}
&\rho\left(\zeta_1, \tilde{\zeta}_1\right) \le \sup_{k,n} \rho\left(X_{k,n,p}, \tilde{X}_{k,n,p}\right), \\
&\sup_{1 \le k \le n} \rho\left(X_{k,n,p}, \tilde{X}_{k,n,p}\right) \le n\rho\left(\zeta_1, \tilde{\zeta}_1\right).
\end{aligned} \tag{5.13}$$

For the case $p = \infty$ and $\tilde{X}_{k,n,\infty} = \bigvee_{i=1}^{k} \tilde{\zeta}_i / \bigvee_{i=1}^{n} \tilde{\zeta}_i$, we use the following condition:

CONDITION 1. There exists a nondecreasing continuous function $\varphi(t) = \varphi_{\tilde{\zeta}_1}(t)$: $[0,1] \mapsto [0,\infty)$, $\varphi(0) = 0$ and such that

$$\varphi(t) \geq \sup_{1-t \leq x \leq 1} (-\log x)^{-1} |F_{\tilde{\zeta}_1}(x) - x|.$$

Obviously, Condition 1 is satisfied for $\tilde{\zeta}_1 =_d \zeta_1$ (which is uniformly distributed on [0, 1]) with $\varphi(t) = t^p$ (with any $p > 0$). Let $\psi(t) = (-\log(1-t)) \cdot \varphi(t)$ and let ψ^{-1} be the inverse of ψ.

THEOREM 5.3. (i) *If Condition* 1 *holds and if* $F_{\tilde{\zeta}_1}(1) = 1$, *then*

$$\Delta := \sup_{k,n} \rho\left(X_{k,n,\infty}, \tilde{X}_{k,n,\infty}\right) \leq c\sqrt{\varphi \circ \psi^{-1}(\rho)}, \quad \textit{where } \rho := \rho\left(\zeta_1, \tilde{\zeta}_1\right).$$

(ii) *If* $\tilde{\zeta}_1$ *has a continuous d.f., then* $\Delta \geq \rho$.

PROOF OF (i).

CLAIM 1. For any $1 \leq k \leq n$,

$$\rho\left(X_{k,n,\infty}, \tilde{X}_{k,n,\infty}\right) \leq \rho\left(\bigvee_{i=1}^{k} \zeta_i, \bigvee_{i=1}^{k} \tilde{\zeta}_i\right) + \rho\left(\bigvee_{i=k+1}^{n} \zeta_i, \bigvee_{i=k+1}^{n} \tilde{\zeta}_i\right) \leq n\rho. \tag{5.14}$$

PROOF.

$$X_{k,n,\infty} = \frac{X_1}{X_1 \vee X_2}, \qquad \tilde{X}_{k,n,\infty} = \frac{\tilde{X}_1}{\tilde{X}_1 + \tilde{X}_2},$$

where

$$X_1 = \bigvee_{i=1}^{k} \zeta_i, \qquad X_2 = \bigvee_{i=k+1}^{n} \zeta_i,$$

$$\tilde{X}_1 = \bigvee_{i=1}^{k} \tilde{\zeta}_i, \qquad \tilde{X}_2 = \bigvee_{i=k+1}^{n} \tilde{\zeta}_i.$$

Following the proof of (5.3), we may assume (X_1, X_2) is independent of $(\tilde{X}_1, \tilde{X}_2)$ and thus

$$\begin{aligned}\rho\left(X_{k,n,\infty}, \tilde{X}_{k,n,\infty}\right) &= \rho\left(\frac{X_1}{X_1 \vee X_2}, \frac{\tilde{X}_1}{\tilde{X}_1 \vee \tilde{X}_2}\right) = \rho\left(1 \vee \frac{X_2}{X_1}, 1 \vee \frac{\tilde{X}_2}{\tilde{X}_1}\right) \\ &\leq \rho\left(\frac{X_2}{X_1}, \frac{\tilde{X}_2}{\tilde{X}_1}\right) \leq \rho\left(\frac{X_2}{X_1}, \frac{\tilde{X}_2}{X_1}\right) + \rho\left(\frac{\tilde{X}_2}{X_1}, \frac{\tilde{X}_2}{\tilde{X}_1}\right) \\ &\leq \rho\left(X_2, \tilde{X}_2\right) + \rho\left(X_1, \tilde{X}_1\right),\end{aligned}$$

the last inequality following by taking conditional expectations.

CLAIM 2. Let $\rho_* = \rho_*(\zeta_1, \tilde{\zeta}_1) = \sup_{0 \le x \le 1}(-\log x)^{-1}|F_{\zeta_1}(x) - F_{\tilde{\zeta}_1}(x)|$. Then

$$\rho\left(\bigvee_{i=1}^{n} \zeta_i, \bigvee_{i=1}^{n} \tilde{\zeta}_i\right) \le c\sqrt{\rho_*}. \tag{5.15}$$

PROOF. Consider the transformation $f(t) = (-\log t)^{-1/\alpha}$ $(0 < t < 1)$. Then

$$\rho\left(\bigvee_{i=1}^{n} \zeta_i, \bigvee_{i=1}^{n} \tilde{\zeta}_i\right) = \rho\left(f\left(\bigvee_{i=1}^{n} \zeta_i\right), f\left(\bigvee_{i=1}^{n} \tilde{\zeta}_i\right)\right) = \rho\left(\bigvee_{i=1}^{n} X_i, \bigvee_{i=1}^{n} \tilde{X}_i\right), \tag{5.16}$$

where $X_i = f(\zeta_i)$, $\tilde{X}_i = f(\tilde{\zeta}_i)$. Since X_1 has an extreme value distribution with parameter α, so does $Z_n := n^{-1/\alpha} \bigvee_{i=1}^{n} X_i$. The density of Z_n is given by $f_{Z_n}(x) = (d/dx)e^{-x^{-\alpha}} = \alpha x^{-\alpha-1}e^{-x^{-\alpha}}$ and thus

$$C_n := \sup_{x>0} f_{Z_n}(x) = \alpha\left(\frac{\alpha+1}{\alpha}\right)^{(\alpha+1)/\alpha} \exp\left\{-\frac{\alpha+1}{\alpha}\right\}. \tag{5.17}$$

Let ρ_α be the weighted uniform distance

$$\rho_\alpha(X, Y) = \sup_{x>0} x^\alpha |F_X(x) - F_Y(x)|. \tag{5.18}$$

Then by de Haan and Rachev [[7], (2.38)],

$$\rho(X, Y) \le \Lambda_\alpha A^{\alpha/(1+\alpha)} \rho_\alpha^{1/(1+\alpha)}(X, Y), \tag{5.19}$$

where $\Lambda_\alpha := (1+\alpha)\alpha^{-\alpha(1+\alpha)}$ and $A := \sup_{x>0} F'_Y(x)$ (the existence of one density being assumed). Hence, by (5.16), (5.17) and (5.19),

$$\rho\left(\bigvee_{i=1}^{n} \zeta_i, \bigvee_{i=1}^{n} \tilde{\zeta}_i\right) = \rho(Z_n, \tilde{Z}_n) \le \Lambda_\alpha C_n^{\alpha/(1+\alpha)} \rho_\alpha^{1/(1+\alpha)}(Z_n, \tilde{Z}_n), \tag{5.20}$$

where $\tilde{Z}_n = n^{-1/\alpha} \bigvee_{i=1}^{n} \tilde{X}_i$. The metric ρ_α is an ideal metric of order α w.r.t. the maxima scheme for iid r.v.'s (see Zolotarev [23]) and in particular

$$\begin{aligned} \rho_\alpha(Z_n, \tilde{Z}_n) &= \rho_\alpha\left(n^{-1/\alpha} \bigvee_{i=1}^{n} X_i, n^{-1/\alpha} \bigvee_{i=1}^{n} \tilde{X}_i\right) \\ &\le n^{-1} \rho_\alpha\left(\bigvee_{i=1}^{n} X_i, \bigvee_{i=1}^{n} \tilde{X}_i\right) \\ &\le \rho_\alpha(X_1, \tilde{X}_1) = \rho_*(\zeta_1, \tilde{\zeta}_1). \end{aligned} \tag{5.21}$$

From Condition 1, we now obtain:

CLAIM 3.

$$\rho_* \le \varphi \circ \psi^{-1}(\rho). \tag{5.22}$$

PROOF. For any $0 \le t \le 1$,

$$\rho_* = \max\left\{ \sup_{0\le x\le 1-\varepsilon} (-\log x)^{-1}|F_{\zeta_1}(x) - x|, \right.$$

$$(5.23) \qquad \left. \sup_{1-\varepsilon\le x\le 1} (-\log x)^{-1}|F_{\zeta_1}(x) - x| \right\}$$

$$\le \max\left((-\log(1-\varepsilon))^{-1}\rho, \varphi(\varepsilon)\right).$$

Choosing ε by $\varphi(\varepsilon) = (-\log(1-\varepsilon))^{-1}\rho)$, that is, $\rho = \psi(\varepsilon)$, one obtains the claim.

From Claims 1, 2 and 3 we obtain

$$(5.24) \qquad \rho\left(X_{k,n,p}, \tilde{X}_{k,n,p}\right) \le \min\left(n\rho, c\sqrt{\varphi\circ\psi^{-1}(\rho)}\right).$$

PROOF OF (ii). For the proof of (ii), observe that

$$F_{\bigvee_{i=1}^n \tilde{\zeta}_i}(x) = F_{\tilde{\zeta}_i}^n(x) \to 1$$

for any x with $F_{\tilde{\zeta}_1}(x) > 0$. As in the proof of Theorem 5.2, we then obtain

$$\sup_{k,n} \rho\left(\frac{\bigvee_{i=1}^k \zeta_i}{\bigvee_{i=1}^n \zeta_i}, \frac{\bigvee_{i=1}^k \tilde{\zeta}_i}{\bigvee_{i=1}^n \tilde{\zeta}_i} \right) \ge \limsup_n \rho\left(\frac{\zeta_1}{\bigvee_{i=1}^n \zeta_i}, \frac{\tilde{\zeta}_1}{\bigvee_{i=1}^n \tilde{\zeta}_i} \right)$$

$$\ge \rho\left(\zeta_1, \tilde{\zeta}_1\right) - \limsup \rho\left(\frac{\zeta_1}{\bigvee_{i=1}^n \zeta_i}, \zeta_1 \right)$$

$$-\limsup \rho\left(\frac{\tilde{\zeta}_1}{\bigvee_{i=1}^n \tilde{\zeta}_i}, \tilde{\zeta}_1 \right) = \rho\left(\zeta_1, \tilde{\zeta}_1\right),$$

since $\zeta_1, \tilde{\zeta}_1$ have continuous d.f.'s.

REMARK 5.3. In (i), the constant c depends on $\alpha > 0$. So, one can optimize c by choosing α appropriately in (5.20).

With respect to the total variation $D(X, Y) = \sup_{A\in\mathscr{A}}|P(X \in A) - P(Y \in A)|$, we can formulate the following stability results.

THEOREM 5.5. (a) *For* $0 < p < \infty$,

$$(5.25) \quad D\left(X_{n,k,p}, \tilde{X}_{n,k,p}\right) \le D\left(\sum_{i=1}^k \zeta_i^p, \sum_{i=1}^k \tilde{\zeta}_i^p \right) + D\left(\sum_{i=k+1}^n \zeta_i^p, \sum_{i=k+1}^n \tilde{\zeta}_i^p \right).$$

If $E\zeta_i^{pj} = E\tilde{\zeta}_i^{pj} = EN_i^j$, $i = 1, 2$, $j = 1, 2$, *for some independent normal r.v.'s* N_i *and if for some* $r > 2$, *the pseudomoments* $\nu_r = \nu_r(\zeta_i, N_i) = \int |x|^r |F_{\zeta_i^p}(x) -$

$F_{N_i}(x)|\,dx \le a < \infty$, *then*

$$\sup_{k \le n} D\left(X_{k,n,p}, \tilde{X}_{k,n,p}\right) \le A(a) n^{-1/2}. \tag{5.26}$$

PROOF. By a well-known representation of the total variation metric,

$$\begin{aligned}
&D\left(X_{n,k,p}, \tilde{X}_{n,k,p}\right)\\
&= \inf\left\{P(X \ne Y); X =_d \sum_{i=1}^k \zeta_i^p \Big/ \sum_{i=1}^n \zeta_i^p, Y =_d \sum_{i=1}^k \tilde{\zeta}_i^p \Big/ \sum_{i=1}^n \tilde{\zeta}_i^p\right\}\\
&\le \inf\left\{P\left(\frac{X_1}{X_1 + X_2} \ne \frac{Y_1}{Y_1 + Y_2}\right);\right.\\
&\qquad X_1, X_2 \text{ independent}, X_1 =_d \sum_{i=1}^k \zeta_i^p, X_2 =_d \sum_{i=k+1}^n \zeta_i^p,\\
&\qquad\qquad \left. Y_1, Y_2 \text{ independent}, Y_1 =_d \sum \tilde{\zeta}_i^p, Y_2 =_d \sum \zeta_i^p\right\}\\
&\le \inf\{P(X_1 \ne Y_1) + P(X_2 \ne Y_2); \cdots\}\\
&= \inf\left\{P(X_1 \ne Y_1): X_1 =_d \sum_{i=1}^k \zeta_i^p, Y_1 =_d \sum_{i=1}^k \tilde{\zeta}_i^p\right\}\\
&\quad + \inf\left\{P(X_2 \ne Y_2); X_2 =_d \sum_{i=k+1}^n \zeta_i^p, Y_2 =_d \sum_{i=k+1}^n \tilde{\zeta}_i^p\right\}\\
&= D\left(\sum_{i=1}^k \zeta_i^p, \sum_{i=1}^k \tilde{\zeta}_i^p\right) + D\left(\sum_{i=k+1}^n \zeta_i^p, \sum_{i=k+1}^n \tilde{\zeta}_i^p\right).
\end{aligned}$$

From the triangle inequality and smoothing inequality for D we infer

$$D\left(\sum_{i=1}^k \zeta_i^p, \sum_{i=1}^k \tilde{\zeta}_i^p\right) \le kD\left(\zeta_1, \tilde{\zeta}_1\right)$$

and

$$D\left(\sum_{i=1}^n \zeta_i^p, \sum_{i=1}^n \tilde{\zeta}_i^p\right) \le D\left(\sum_{i=1}^n \zeta_i^p, \sum_{i=1}^n N_i\right) + D\left(\sum_{i=1}^n N_i, \sum_{i=1}^n \tilde{\zeta}_i^p\right),$$

which is bounded above by $A(a)n^{-1/2}$ by a result of Senatov [20], Sazonov [19] and Rachev and Yukich [16]. □

In the case $p = \infty$ one gets for the total variation distance:

THEOREM 5.6. *If* $(0 < \tilde{\zeta}_i < 1)$ *and if* $-1/\log \tilde{\zeta}_i$ *has a d.f. F satisfying for some* $s > 1$, $\int_0^\infty x^{1+s}|F'(x) - (1/x^2)e^{-1/x}|\,dx < \infty$, *then*

$$\sup_{k \le n} D\big(X_{k,n,\infty}, \tilde{X}_{k,n,\infty}\big) = O(n^{1-s}). \tag{5.27}$$

PROOF.

$$D\left(\frac{\bigvee_{i=1}^k \zeta_i}{\bigvee_{i=1}^n \zeta_i}, \frac{\bigvee_{i=1}^k \tilde{\zeta}_i}{\bigvee_{i=1}^n \tilde{\zeta}_i}\right) \le D\left(\bigvee_{i=1}^k \zeta_i, \bigvee_{i=1}^k \tilde{\zeta}_i\right) + D\left(\bigvee_{i=1}^n \zeta_i, \bigvee_{i}^n \tilde{\zeta}_i\right) = D_k + D_n.$$

With $f(t) = -1/\log t$, $0 < t < 1$, $f(\zeta_i) = X_i$ has the extreme value distribution with density $F_{X_1}(x) = e^{-1/x}$. Therefore,

$$D_n = D\left((1/n)\bigvee_{i=1}^n X_i, (1/n)\bigvee_{i=1}^n Y_i\right) = D\left(X_1, (1/n)\bigvee_{i=1}^n Y_i\right),$$

where $Y_i = f(\tilde{\zeta}_i)$. Thus we have to estimate the rate of convergence of the normalized maxima $(1/n)\bigvee_{i=1}^n Y_i$ to X_1 in terms of the total variation distance. Following Omey and Rachev [15] and Omey [14], one can prove that the condition $\int_0^\infty x^{1+s}|F'(x) - F'_{X_1}(x)|\,dx < \infty$, where F is the d.f. of Y_1, implies $\limsup n^{s-1}D((1/n)\bigvee_{i=1}^n Y_i, X_1) < \infty$. □

6. The discrete case. If $\zeta_1, \dots, \zeta_n$ are iid geometrically distributed with parameter $\theta \in (0,1)$, that is, $P_\theta^{\zeta_1}(j) = (1-\theta)\theta^j$, $j = 0, 1, \dots,$ then

$$Q_{n,s,n}^{(1)} = P_\theta^{(\zeta_1,\dots,\zeta_n)|\sum_{i=1}^n \zeta_i = s}$$

is uniform on the simplex $S_{1,s,n}^+ = \{x \in \mathbb{N}_0^n : \sum_{i=1}^n x_i = s\}$ for $s \in \mathbb{N}_0$ and Diaconis and Freedman [9] proved that

$$\|Q_{n,s,k}^{(1)} - P_{s/(n+s)}^{k,1}\| \le 2\left\{\frac{n^2}{(n-k-1)(n-k-2)} - 1\right\}, \tag{6.1}$$

where

$$P_\theta^{k,1} = P_\theta^{(\zeta_1,\dots,\zeta_k)} \quad \text{and} \quad Q_{n,s,k}^{(1)} = P_\theta^{(\zeta_1,\dots,\zeta_k)|\sum_{i=1}^n \zeta_i = s}.$$

We at first consider the corresponding result for $p = \infty$, that is, we assume that $\zeta_1, \dots, \zeta_n$ are iid uniformly distributed on the finite lattice $\{1, 2, \dots, m\}$.

LEMMA 6.1. *If* $\zeta_1, \dots, \zeta_n$ *are iid uniform on* $\{1, \dots, m\}$, *then the conditional distribution* $Q_{n,s,n}^{(\infty)} = P^{(\zeta_1,\dots,\zeta_n)|\max \zeta_i = s}$ *is for* $1 \le s \le m$ *uniformly distributed on* $S_{\infty,s,n} = \{x \in \mathbb{N}^n : \max x_i = s\}$.

PROOF. For any integers $k_1, \dots, k_n$ with $\max_{1 \le i \le n} k_i = s \le m$ we have:

$$P\Big(\zeta_1 = k_1, \dots, \zeta_n = k_n \Big/ \max_{1 \le i \le n} \zeta_i = s\Big)$$

$$= \frac{P(\zeta_1 = k_1, \dots, \zeta_n = k_n)}{P(\max_{1 \le i \le n} \zeta_i = s)} \tag{6.2}$$

$$= \frac{1}{m^n P(\max \zeta_i = s)} = \frac{1}{m^n(s^n - s^{n-1})},$$

since there are $s^n - (s-1)^n$ n-tuples $(r_1, \dots, r_n)$ with $\max_{1 \le i \le n} r_i = s$. □

THEOREM 6.2. *If $\zeta_1, \dots, \zeta_n$ are iid uniform on $\{1, \dots, m\}$ and if $m = s$, $1 \le k \le m$, then*

$$\|Q_{n,s,k}^{(\infty)} - P_m^k\| \le \frac{1}{2}\left(\frac{k}{n-k} + \frac{k}{n}\right), \tag{6.3}$$

P_m being the uniform distribution on $\{1, \dots, m\}$.

PROOF. There are $s^n - (s-1)^n$ n-tuples $(r_1, \dots, r_n)$ with $\max r_i = s$. To determine $Q_{n,s,k}^{(\infty)}$, we consider two cases.

CASE 1. $t = j_1 \vee \cdots \vee j_k < s$; then

$$Q_{n,s,k}^{(\infty)}(j_1, \dots, j_k) = \frac{s^{n-k} - (s-1)^{n-k}}{s^n - (s-1)^n}. \tag{6.4}$$

CASE 2. $t = j_1 \vee \cdots \vee j_k = s$; then

$$Q_{n,s,k}^{(\infty)}(j_1, \dots, j_k) = \frac{s^{n-k}}{s^n - (s-1)^n}. \tag{6.5}$$

Consider at first Case 1, where $t = j_1 \vee \cdots \vee j_j < s = m$. Then

$$I_1 = \frac{Q_{n,s,k}^{(\infty)}(j_1, \dots, j_k)}{P_m^k(j_1, \dots, j_k)} = s^k \frac{s^{n-k} - (s-1)^{n-k}}{s^n - (s-1)^n} = \frac{1 - (1 - (1/s))^{n-k}}{1 - (1 - (1/s))^n} \le 1.$$

$I_1 \ge \inf_{0 \le x < 1}(1 - x^\alpha)/(1 - x) =: \varphi(x)$, where $\alpha = (n-k)/n$. But $\varphi(x)$ is monotonically decreasing on $[0, 1]$, so $I_1 \ge \varphi(1) = \alpha$. Thus

$$0 \le \frac{1}{I_1} - 1 \le \frac{n}{n-k} - 1 = \frac{k}{n-k}$$

and hence

$$J_1 = \sum_{\substack{(j_1,\dots,j_k)\\ j_1 \vee \cdots \vee j_k < s}} |P_m^k(j_1,\dots,j_k) - Q_{nsk}^{(\infty)}(j_1,\dots,j_k)|$$

$$(6.6) \qquad = \sum_{\substack{(j_1,\dots,j_k)\\ j_1 \vee \cdots \vee j_k < s}} \left(\frac{1}{I_1} - 1\right) Q_{nsk}^{(\infty)}(j_1,\dots,j_k)$$

$$\le \frac{k}{n-k} Q_{n,s,k}^{(\infty)}\{(j_1,\dots,j_k): j_1 \vee \cdots \vee j_k < s\}.$$

In the second case, where $t = j_1 \vee \cdots \vee j_k = s$, we obtain

$$I_2 = \frac{Q_{nsk}(j_1,\dots,j_k)}{P_m^k(j_1,\dots,j_k)} = s^k \frac{s^{n-k}}{s^n - (s-1)^n} = \frac{1}{1 - (1 - (1/s))^n} \ge 1.$$

Hence

$$J_2 = \sum_{\substack{(j_1,\dots,j_k)\\ j_1 \vee \cdots \vee j_k = s}} |P_m^k(j_1,\dots,j_k) - Q_{nsk}^{(\infty)}(j_1,\dots,j_k)|$$

$$= \sum_{j_1 \vee \cdots \vee j_k = s} \left(\frac{s^{n-k}}{s^n - (s-1)^n} - s^{-k}\right)$$

$$(6.7) \qquad = \left(s^k - (s-1)^k\right)\left(\frac{s^{n-k}}{s^n - (s-1)^n} - s^{-k}\right)$$

$$= \left(1 - \left(1 - \frac{1}{s}\right)^k\right)\left(\frac{1}{1 - (1 - (1/s))^n} - 1\right) = \frac{1 - x^\beta}{1 - x} - 1 + x^\beta,$$

where $x = (1 - (1/s))^n$, $\beta = k/n$. Since $\varphi(x) = (1 - x^\beta)/(1 - x) - 1 + x^\beta \le \varphi(1) = \beta$, we obtain $J_2 \le k/n$ and therefore

$$(6.8) \qquad \|Q_{n,s,k}^{(\infty)} - P_m^k\| = \frac{1}{2}(J_1 + J_2) \le \frac{1}{2}\left(\frac{k}{n-k} + \frac{k}{n}\right).$$

REMARK 6.1. (a) If

$$s \ge \frac{1}{1 - (1 - k/(n-k))^{1/n}},$$

then the proof of Theorem 6.2 gives the improved bound

$$(6.9) \qquad \|Q_{n,s,k}^{(\infty)} - P_m^k\| \le \frac{1}{2}\frac{k}{n-k}.$$

(b) The estimate in Theorem 6.2 is uniform in s and therefore (6.3) implies a finite and also infinite de Finetti-type theorem. (For an analysis approach to this infinite de Finetti theorem cf. also Ressel [17], Example 4).

Consider now the discrete p-spheres of radius $s \in \mathbb{N}$:

$$S_{p,s,n} = \left\{ x \in \mathbb{N}^n \colon \sum_{i=1}^{n} x_i^p = s \right\}$$

and

$$S_{p,s,n}^{+} = \left\{ x \in \mathbb{N}_0^n \colon \sum_{i=1}^{n} x_i^p = s \right\}, \qquad p \in \mathbb{N}.$$

As in Lemma 6.1, we get:

Lemma 6.3. *If $\zeta_1, \dots, \zeta_n$ are iid uniformly distributed on $\{1, \dots, m\}$ (resp., $\{0, \dots, m\}$) and if $S_{p,s,n} \neq \varnothing$ (resp., $S_{psn}^{+} \neq \varnothing$), then the conditional distribution $Q_{n,s,n}^{(p)} = P^{(\zeta_1, \dots, \zeta_n) | \Sigma \zeta_i^p = s}$ is uniform on $S_{p,s,n}$ (resp., $S_{p,s,n}^{+}$).*

Remark 6.2. The question arises whether $S_{p,s,n}^{+} \neq \varnothing$ is related to the famous Waring problem of number theory. The main result of Hilbert (proved in 1909) says that for any $p \in \mathbb{N}$, there exists an $n \in \mathbb{N}$ such that $S_{p,s,n}^{+} \neq \varnothing$ for all $s \in \mathbb{N}$. Furthermore, a famous theorem of Hardy and Littlewood (cf. Vinogradov [21] and Ellison [11]) states:

If $n \geq 2^p + 1$, then

$$\begin{aligned} r_{p,n}(s) &= |S_{p,s,n}^{+}| \\ &= \frac{\Gamma(1 + (1/p))^n}{\Gamma(n/p)} s^{(n/p)-1} G(s,p,n) + o(s^{(n/p)-1}) \end{aligned} \tag{6.10}$$

as $s \to \infty$,

where

$$\begin{aligned} G(s,p,n) &= G(s) \\ &= \sum_{Q=1}^{\infty} \sum_{\substack{0 \leq a < Q \\ (a,Q)=1}} (S_{a,Q}/Q)^n e^{-2\pi i (a/Q)s}, \end{aligned}$$

$$S_{a,Q} = \sum_{0 \leq l < Q} e^{2\pi i (a/Q) l^p}$$

and $\infty > C(p,n) \geq G(s) \geq c(p,n) > 0$ for all s.

For $z \in (0,1)$, $p \in \mathbb{N}$, we can define in analogy to the continuous case distributions

$$P_{p,z}(j) = a_p z^{j^p}, \qquad j \in \mathbb{N}_0, \qquad a_p = a_p(z), \tag{6.11}$$

where $1/a_p = \sum_{j=0}^{\infty} z^{j^p}$. Let $\mu_p(z) = a_p \sum_{j=0}^{\infty} j z^{j^p} = \sum_j j P_{p,z}(j)$.

LEMMA 6.4. *Let $z \in (0,1)$, $p \in \mathbb{N}$ and $\zeta_1, \dots, \zeta_n$ be iid distributed according to $P_{p,z}$. Then the following hold:*

(i) $P_{p,z}^{(\zeta_1,\dots,\zeta_n)\mid \sum_{i=1}^n \zeta_i^p = s}$ *is uniformly distributed on $S_{p,s,n}^+$, if $S_{p,s,n}^+ \neq \varnothing$.*
(ii) $P_{p,z}$ *is the unique distribution with support $\mathbb{N}_0$ such that* (i) *holds.*
(iii) $P_{p,z}$ *is the distribution with support $\mathbb{N}_0$ with given first moment $\mu_p(z)$ and with maximum value of the entropy $-\sum_{j=0}^\infty \ln P(j) = H(P)$.*

REMARK 6.3. From (6.10), we obtain for $Q_{k,s,n}^{(p)} = P_{p,z}^{(\zeta_1,\dots,\zeta_k)\mid \sum_{i=1}^n \zeta_i^p = s}$,

$$\begin{aligned} \frac{Q_{k,s,n}^{(p)}(j_1,\dots,j_k)}{P_{p,z}(j_1,\dots,j_k)} &= \frac{r_{p,n-k}(s-t)}{r_{p,n}(s)a_p^k z^t} \quad \text{where } t = \sum_{i=1}^k j_i^p \\ &= \Big\{\Gamma(1+(1/p))^{n-k}(s-t)^{((n-k)/p)-1} \\ &\qquad \times G(s,n-k,p) + o\big((s-t)^{((n-k)/p)-1}\big)\Big\} \\ &\quad \times \Big\{\Gamma((n-k)/p)\big(\Gamma(1+(1/p))^n/\Gamma(n/p)\big)s^{(n/p)-1} \\ &\quad \times G(s,n,p)a_p^k z^t + o(s^{(n/p)-1})\Big\}^{-1} \\ &= \frac{(n/p)}{((n-k)/p)(1+(1/p))^k}\left(\frac{s-t}{s}\right)^{((n-k)/p)-1} \\ &\quad \times \frac{G(s,n-k,p)}{G(s,n,p)} \frac{1+o\big(((s-t)/s)^{((n-k)/p)-1}\big)}{s^{k/p}a_p^k z^t + o(s^{k/p})} \quad \text{as } s \to \infty. \end{aligned} \tag{6.12}$$

For a more explicit evaluation of the supremum distance one would need more information on the function $G(s,n,p)$ as a function of n, s.

From Corollary 3.3 of Zabell [22], we infer:

THEOREM 6.5. *Let $c_n = o(n)$, $k \geq 1$ and $\alpha_n = n\mu_p(z) + c_n \in \mathbb{N}$, then*

$$\|Q_{k,\alpha_n,n}^{(p)} - P_{p,z}^k\| = O\left(\frac{1+|c_n|}{n}\right), \tag{6.13}$$

where $P_{p,z}^k$ is the k-fold product of $P_{p,z}$.

In addition to the number theoretic result (6.10) concerning the case $s \to \infty$, n fixed, one gets from (6.13),

$$\sum_{t=0}^{\alpha_n} \left| \frac{r_{p,n-k}(\alpha_n - t)}{r_{p,n}(\alpha_n)} - a_p^k z^t \right| = O\left(\frac{1+|c_n|}{n}\right) \quad \text{as } n \to \infty. \tag{6.14}$$

For the analysis, the big problem is contained in the normalization $\varphi(z) = \sum_{j=0}^{\infty} z^{j^p} = 1/a_p$. By elementary calculations $(\varphi(z))^n = \sum_{s=1}^{\infty} r_{p,n}(s) z^s$ and therefore $r_{p,n}(s) = \int_0^1 [\varphi(e^{2\pi i\alpha})]^n e^{-2\pi i s\alpha}\,d\alpha$.

Acknowledgments. This research was done while Rachev was Visiting Professor at the University of Münster. He wants to acknowledge gratefully the grant received from the Deutsche Forschungsgemeinschaft (DFG). We thank the referees for their helpful comments. In particular the distinction of the uniform measure used in this paper and the geometric surface measure is due to the comments of a referee.

REFERENCES

[1] ABRAMOWITZ, M. and STEGUN, I. A. (1970). *Handbook of Mathematical Functions*. Dover, New York.

[2] BAXTER, L. A. and RACHEV, S. T. (1990). A note on the stability of the estimation of the exponential distribution. *Statist. Probab. Lett.* **10** 37–41.

[3] BERK, R. H. (1977). Characterizations via conditional distributions. *J. Appl. Probab.* **14** 806–816.

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

[5] CSISZAR, I. (1967). Information-type measures of difference of probability distributions and indirect observations. *Studia Sci. Math. Hungar.* **2** 229–318.

[6] CSISZAR, I. (1975). I-divergence geometry of probability distributions and minimization problems. *Ann. Probab.* **3** 146–158.

[7] DE HAAN, L. and RACHEV, S. T. (1989). Estimates of the rate of convergence for max-stable processes. *Ann. Probab.* **17** 651–677.

[8] DIACONIS, P. and FREEDMAN, D. (1984). Partial exchangeability and sufficiency. In *Statistics: Applications and New Directions*. Proceedings of the Indian Statistical Institute Golden Jubilee International Conference (J. K. Ghosh and J. Roy, eds.) 205–236. Indian Statistical Institute, Calcutta.

[9] DIACONIS, P. and FREEDMAN, D. (1987). A dozen de Finetti-style results in search of a theory. *Ann. Inst. H. Poincaré Probab. Statist.* **23** 397–423.

[10] DIACONIS, P. and FREEDMAN, D. (1987). A finite version of de Finetti's theorem for exponential families with uniform asymptotic estimates. Technical Report 91, Univ. California.

[11] ELLISON, W. J. (1971). Waring's problem. *Amer. Math. Monthly* **78** 10–36.

[12] KULLBACK, S. (1959). *Information Theory and Statistics*. Wiley, New York.

[13] KUMAR, A. and PATHAK, P. K. (1977). Sufficiency and tests of goodness of fit. *Scand. J. Statist.* **4** 39–43.

[14] OMEY, E. (1988). Rates of convergence for densities in extreme value theory. *Ann. Probab.* **16** 479–486.

[15] OMEY, E. and RACHEV, S. T. (1988). On the rate of convergence in extreme value theory. *Theory Probab. Appl.* **33** 560–565.

[16] RACHEV, S. T. and YUKICH, J. E. (1989). Rates for the CLT via new ideal metrics. *Ann. Probab.* **17** 775–788.

[17] RESSEL, P. (1985). De Finetti-type theorems: an analytic approach. *Ann. Probab.* **13** 898–922.

[18] RÜSCHENDORF, L. (1984). On the minimum discrimination information theorem. *Statist. Decisions* **1** 263–283.

[19] SAZONOV, V. V. (1981). *Normal Approximation—Some Recent Advances. Lecture Notes in Math.* **879**. Springer, New York.

[20] SENATOV, V. E. (1980). Uniform estimates of the rate of convergence in the multidimensional central limit theorem. *Theory Probab. Appl.* **25** 745–759.

[21] Vinogradov, I. M. (1975). *Trigonometrical Sums in Number Theory*. Statistical Publishing Society, Calcutta.

[22] Zabell, S. L. (1980). Rates of convergence for conditional expectations. *Ann. Probab.* **8** 928–941.

[23] Zolotarev, V. M. (1983). Probability metrics. *Theory Probab. Appl.* **28** 278–302.

Department of Statistics
and Applied Probability
University of California
Santa Barbara, California 93106

Westfälische Wilhelms-Universität
Institut für Math. Statistik
Einsteinstr. 62
D-4400 Münster
Germany