ASYMPTOTIC DISTRIBUTION OF COORDINATES ON HIGH DIMENSIONAL SPHERES

M. C. SPRUILL
School of Mathematics, Georgia Institute of Technology, Atlanta, GA 30332-016
email: spruill@math.gatech.edu

Submitted January 5, 2007, accepted in final form July 19, 2007
AMS 2000 Subject classification: 60F17, 52A40, 28A75
Keywords: empirical distribution, dependent arrays, micro-canonical ensemble, Minkowski area, isoperimetry.

Abstract
The coordinates \( x_i \) of a point \( x = (x_1, x_2, \ldots, x_n) \) chosen at random according to a uniform distribution on the \( \ell_2(n) \)-sphere of radius \( n^{1/2} \) have approximately a normal distribution when \( n \) is large. The coordinates \( x_i \) of points uniformly distributed on the \( \ell_1(n) \)-sphere of radius \( n \) have approximately a double exponential distribution. In these and all the \( \ell_p(n), 1 \leq p \leq \infty \), convergence of the distribution of coordinates as the dimension \( n \) increases is at the rate \( \sqrt{n} \) and is described precisely in terms of weak convergence of a normalized empirical process to a limiting Gaussian process, the sum of a Brownian bridge and a simple normal process.

1 Introduction
If \( Y_n = (Y_{1n}, \ldots, Y_{nn}) \) is chosen according to a uniform distribution on the sphere in \( n \) dimensions of radius \( \sqrt{n} \) then, computing the ratio of the surface area of a polar cap to the whole sphere, one finds that the marginal probability density of \( Y_{jn} / \sqrt{n} \) is

\[
 f_n(s) = \kappa_n (1 - s^2)^{(n-3)/2} I_{(-1,1)}(s),
\]

where \( \kappa_n = \frac{\Gamma(n/2)}{\sqrt{\pi} \Gamma(n/4)} \). Stirling’s approximation shows

\[
 \lim_{n \to \infty} \kappa_n \frac{(1 - \frac{v^2}{n})^{(n-3)/2} I_{(-\sqrt{n}, \infty)}(v)}{v^{-3/2}} = \frac{1}{\sqrt{2\pi}} e^{-v^2/2},
\]

so appealing to Scheffe’s theorem (see [3]) one has

\[
 \lim_{n \to \infty} P[Y_{jn} \leq t] = \lim_{n \to \infty} \kappa_n \int_{-\sqrt{n}}^{t} (1 - \frac{v^2}{n})^{(n-3)/2} \frac{dv}{\sqrt{n}} = \Phi(t)
\]

and \( Y_{jn} \) is asymptotically standard normal as the dimension increases. This is an elementary aspect of a more comprehensive result attributed to Poincaré; that the joint distribution of the
first $k$ coordinates of a vector uniformly distributed on the sphere $S_{2,n}(\sqrt{n})$ is asymptotically that of $k$ independent normals as the dimension increases. Extensions have been made by Diaconis and Freedman \[9\], Rachev and Ruschendorff \[13\], and Stam in \[10\] to convergence in variation norm allowing also $k$ to grow with $n$. In \[9\] the authors study $k = o(n)$ and relate some history of the problem. Attribution of the result to Poincare was not supported by their investigations; the first reference to the theorem on convergence of the first $k$ coordinates they found was in the work of Borel \[4\]. Borel’s interest, like ours, centers on the empirical distribution (edf)

$$\mathbb{F}_n(t) = \frac{\# \{ Y_{i,n} \leq t ; i = 1, \ldots, n \}}{n}. \quad (1)$$

The proportion of coordinates $Y_{i,n}$ less than or equal to $t$ in $(-\infty, \infty)$ is $\mathbb{F}_n(t)$.

As pointed out in \[4\], answers to Borel’s questions about Maxwell’s theorem are easy using modern methods.

If $Z_1, Z_2, \ldots$ are iid $N(0,1)$ and $R_n = \frac{1}{n} \sum_{i=1}^n Z_i^2$ then it is well known that $R_n^{-1/2}(Z_1, \ldots, Z_n)$ is uniform on $S_{2,n}(n^{1/2})$, so that the edf of $Z_1, Z_2, \ldots, Z_n$ is $G_n$ then since $nG_n(t)$ is binomial, the weak law of large numbers shows that $G_n(t) \xrightarrow{} \Phi(t)$. By continuity of square-root and $\Phi$ and $R_n \xrightarrow{} 1$ it follows, as indicated,

$$\mathbb{F}_n(t) - \Phi(t) = G_n(R_n^{1/2}t) - \Phi(t) \xrightarrow{} 0 + 0.$$ 

that the right-most term of the right hand side converges to 0 in probability. Finally, by the Glivenko-Cantelli lemma (see equation (13.3) of \[3\]) it follows that the left-most term on the right hand side tends to zero in probability. The argument yields asymptotic normality and, assuming continuity, an affirmative answer to the classical statistical mechanical question of equivalence of ensembles: does one have equality of the expectations $E[G[k(Y)] = \int k(y)dG(y)$ and $E[k(Y)] = \int k(y)dU(y)$ where, corresponding to the micro-canonical ensemble, $U$ is the uniform distribution on $\{y : H(y) = c^2\}$, and $G$ is the Gibbs’ distribution satisfying $dG(y) = e^{-\alpha H(y)}dy$ with $a$ such that $E_G[H(Y)] = \int H(y)dG(y) = c^2$, and $H(y)$ the Hamiltonian? For $H(x) = cx^2$, if the functional $g_k(F) = \int k(y)df(y)$ is continuous, then the two are equivalent modulo the choice of constants.

More generally, what can be said about the error in approximating the functional $g(F)$’s value by $g(\mathbb{F}_n)$? In the case of independence there are ready answers to questions about the rate of convergence and the form of the error; for the edf $Q_n$ determined from $n$ independent and identically distributed univariate observations from $Q$, it is well known that the empirical process $D_n(t) = \sqrt{n}(Q_n(t) - Q(t)), t \in (-\infty, \infty)$ converges weakly ($D_n \Rightarrow B \circ Q$) to a Gaussian process as the sample size $n$ increases. Here $B$ is a Brownian bridge and it is seen that the rate of convergence is $\sqrt{n}$ with a Gaussian error. If the functional $g$ is differentiable (see Serfling \[15\]), then $\sqrt{n}(g(Q_n) - g(Q)) \Rightarrow Dg(L)$, where $Dg$ is the differential of $g$ and $L = B \circ Q$ is the limiting error process. The key question in the case of coordinates constrained to the sphere is: does the process $\sqrt{n}(\mathbb{F}_n(t) - \Phi(t))$ converge weakly to a Gaussian process? The answer will be shown to be yes as will the answers to the analogous questions in each of the spaces $\ell_p(n)$ if $\Phi$ is replaced in each case by an appropriate distribution. Even though the random variables are dependent, convergence to a Gaussian process will occur at the rate $\sqrt{n}$. The limiting stochastic process $L(t) = B(F_p(t)) + \frac{dL}{dp}Z$ differs from the limit in the iid case.

To state our result, for $1 \leq p < \infty$, let $\frac{1}{p} + \frac{1}{q} = 1$ and introduce the family of distributions $F_p$
on \((-\infty, \infty)\) whose probability densities with respect to Lebesgue measure are
\[
f_p(t) = \frac{p^{1/2}e^{-|t|^p/p}}{2\Gamma(1/p)},
\]
(2)
The space \(\ell_p(n)\) is \(\mathbb{R}^n\) with the norm \(\|x\|_p = (\sum_{j=1}^n |x_j|^p)^{1/p}\) where \(x = (x_1, \ldots, x_n)\). The sphere of “radius” \(r\) is \(S_{p,n}(r) = \{x \in \mathbb{R}^n : \|x\|_p = r\}\). The ball of radius \(r\) is \(B_{p,n}(r) = \{x \in \mathbb{R}^n : \|x\|_p \leq r\}\). The convergence indicated by \(D_n \Rightarrow D\) is so-called weak convergence of probability measures defined by \(\lim_{n \to \infty} E[h(D_n)] = E[h(D)]\) for all bounded continuous \(h\) and studied in, for example, [3]. The following will be proven, where uniformly distributed in the statement refers to \(\sigma_{p,n}\) defined in section 3.

**Theorem 1.** Let \(p \in [1, \infty)\) and \(Y_n = (Y_{1n}, \ldots, Y_{nn})\) be uniformly distributed according to \(\sigma_{p,n}\) on the sphere \(S_{p,n}(n^{1/p})\). There is a probability space on which are defined a Brownian bridge \(B\) and a standard normal random variable \(Z\) so that if \(\mathbb{F}_n\) is as defined in (1) then
\[
\sqrt{n}(\mathbb{F}_n(t) - F_p(t)) \Rightarrow B(F_p(t)) + \frac{tf_p(t)}{\sqrt{p}} Z,
\]
as \(n \to \infty\), where the indicated sum on the right hand side is a Gaussian process and
\[
\text{cov}(B(F_p(t)), Z) = -tf_p(t).
\]

## 2 Idea of the proof of the theorem

Let \(X_n = (X_1, \ldots, X_n)\) where \(X_1, X_2, \ldots\) are iid \(F_p\) random variables. Then the uniform random vector \(Y_n\) on the \(n\)-sphere of radius \(n^{1/p}\) has the same distribution as \(\frac{n^{1/p}X_n}{\|X_n\|_p}\). Let
\[
\psi_p(X_n) = \left(\frac{\sum_{j=1}^n |x_j|^p}{n}\right)^{1/p}
\]
and \(G_n\) be the usual empirical distribution formed from the \(n\) iid random variables \(\{X_i\}_{i=1}^n\). Then the process of interest concerning (1) can be expressed probabilistically as
\[
\sqrt{n}(\mathbb{F}_n(t) - F_p(t)) \overset{d}{=} \sqrt{n}(G_n(t\psi_p(X_n)) - F_p(t\psi_p(X_n))) + (F_p(t\psi_p(X_n)) - F_p(t))
\]
(5)
It is well known that the process \(\mathbb{F}_n(t) - F_p(t)\) converges weakly to \(B(F_p(t))\), where \(B\) is a Brownian bridge process. Noting that \(\psi_p(X_n) \overset{p}{\to} 1\) as \(n \to \infty\) and that a simple Taylor’s expansion of the second term yields that \(\sqrt{n}(F_p(t\psi_p(X_n)) - F_p(t))\) converges weakly to the simple process \(\frac{tf_p(t)}{\sqrt{p}} V\), where \(V\) is a standard normal random variable, it can be seen that the process in question, the empirical process based on an observation uniform on the \(n^{1/p}\)-sphere in \(\ell_p(n)\), the ensphereal process defined by the left hand side of (5), converges weakly to a zero mean Gaussian process
\[
B(F_p(t)) + V\frac{tf_p(t)}{\sqrt{p}}
\]
as the dimension \(n\) increases. The covariance of the two Gaussian summands will be shown to be
\[
\text{cov}(B(F_p(t)), \frac{sf_p(s)}{\sqrt{p}} V) = \frac{sf_p(s)}{\sqrt{p}} (-tf_p(t)).
\]
Details of the uniform distribution \(\sigma_{p,n}\) of Theorem 1 on the spheres in \(\ell_p(n)\) are given next.
3 Uniform distribution and \( F_p \)

The measure \( \sigma_{p,n} \) of Theorem 11 assigns to measurable subsets of \( S_{p,n}(1) \) their Minkowski surface area, an intrinsic area in that it depends on geodesic distances on the surface. See 6. The measure \( \sigma_{p,n} \) coincides on \( S_{p,n}(1) \), with measures which have appeared in the literature (see 2, 13, and 14) in conjunction with the densities \( f_p \). In particular, it is shown that it coincides with the measure \( \mu_{p,n} \) defined below (see 11) which arose for Rachev and Ruschedorf 13 in the disintegration of \( V_n \).

3.1 The isoperimetric problem and solution

Let \( K \subset R^n \) be a centrally symmetric closed bounded convex set with 0 as an internal point. Then \( \rho_K(x) = \inf \{t : x \in tK, t > 0\} \) defines a Minkowski norm \( \|x\|_K = \rho_K(x) \) on \( R^n \). The only reasonable (Busemann 4) n-dimensional volume measure in this Minkowski space is translation invariant and must coincide with the (Lebesgue) volume measure \( V_n \). One choice for surface area is the Minkowski surface area \( \sigma_K \), defined for smooth convex bodies \( D \) by

\[
\sigma_K(D) = \lim_{\epsilon \to 0} \frac{V_n(D + \epsilon K) - V_n(D)}{\epsilon}.
\]

For a more general class of sets \( M \) (see, for example, equation (18) of 14 for details) the Minkowski surface area can be shown to satisfy

\[
\sigma_K(\partial M) = \int_{\partial M} \|u\|_{K^n} d\sigma_2(u),
\]

where \( \sigma_2 \) is Euclidean surface area, \( u \) is the (Euclidean) unit normal to the surface \( \partial M \), and \( \| \cdot \|_{K^n} \) is the norm in the dual space, also a Minkowski normed space in which the unit ball is the polar reciprocal \( K^0 = \{x^* \in R^n : \langle x^*, x \rangle \leq 1 \forall x \in K\} \) of \( K \). Here \( \langle x, y \rangle = \sum_{i=1}^{n} x_i y_i \).

It follows from the work of Busemann 4 that among all solids \( M \) for which the left hand side of 6 is fixed, the solid maximizing the volume \( V_n \) is the polar reciprocal \( C^0 \) of the set \( C \) of points \( u \) of \( K^n \). The latter is the unit sphere \( S_{K^n}(1) \) of the dual space (see also 8). It follows from (\( \partial K^0 \) \( 0 = K \) that \( C^0 = B_K(1) = K \), the unit ball. This solution also agrees in the case of smooth convex sets with that from Minkowski’s first inequality (see (15) of 11); the solution is the unit ball \( B_K(1) \).

In the case of interest here \( \ell_p(n), 1 < p < \infty \); take \( K = B_{p,n}(1) \) and denote \( \sigma_K \) by \( \sigma_p \). For the sphere \( S_{p,n}(r) \) the Minkowski surface area satisfies

\[
\sigma_p(S_{p,n}(r)) = \lim_{\epsilon \to 0} \frac{V_n(B_{p,n}(r + \epsilon)) - V_n(B_{p,n}(r))}{\epsilon}.
\]

By homogeneity \( V_n(B_{p,n}(r)) = r^n V_n(B_{p,n}(1)) \) so one has \( \sigma_p(S_{p,n}(r)) = V_n(B_{p,n}(1)) \frac{d\Gamma}{d\Gamma} \). By a formula due to Dirichlet (see 11) the volume of \( B_{p,n}(1) \) is \( V_n(B_{p,n}(1)) = \frac{2^n \Gamma(n/2)}{n \pi^{n/2}} \), so the Minkowski surface area of the radius \( r \) sphere in \( \ell_p(n) \) is

\[
\sigma_p(S_{p,n}(r)) = r^{n-1} \frac{2^n \Gamma(n/2)}{p^{n-1} \Gamma(p/2)}.
\]

The simple formula 9 for \( \sigma_p(S_{p,n}(r)) \) should be contrasted with the Euclidean surface area \( \sigma_2(S_{p,n}(r)) \) for which there is no simple closed form. See 8.
3.2 Disintegration of $V_n$ and Minkowski surface area

If $f$ is smooth and $D = \{ x : f(x) \leq c \}$ is a compact convex centrally symmetric set with 0 as an internal point and if $g$ is a measurable function on $\partial D$ then by (7) and $\| \cdot \|_{K^p} = \| \cdot \|_q$, one has $\int_{\partial D} g(x) d\sigma_{p,n}(x) = \int_{\partial D} g(x) \sigma(x) d\sigma_2(x)$. So $d\sigma_{p,n}/d\sigma_2 = \| \nabla f(x) \|_q/\| \nabla f(x) \|_2$. In particular, for the surface $\partial B_{p,n}(r) = S_{p,n}(r) = \{ x \in \mathbb{R}^n : f(x) = r^p \}$, where $f(x) = \sum_{j=1}^n |x_j|^p$, one has a.e. $\sigma_2$, $\frac{\partial f(x)}{\partial x_j} = p \text{sgn}(x_j)|x_j|^{p-1} = p \text{sgn}(x_j)|x_j|^{p/q}$, so for a.e. $x \in S_{p,n}(r)$

$$
\frac{d\sigma_{p,n}}{d\sigma_2}(x) = \frac{p(\sum_{j=1}^n |x_j|^{p/q})^{1/q}}{p(\sum_{j=1}^n |x_j|^{2p/q})^{1/2}} = \frac{r^{p/q}}{\sqrt{\sum_{j=1}^n |x_j|^{2p/q}}} \tag{9}
$$

For $r > 0$ fixed, define the mapping $T_r$ by $T_r(v_1, \ldots, v_{n-1}) = (v_1, \ldots, v_{n-1}, (r^p - \sum_{j=1}^{n-1} v_j^p)^{1/p})$. This maps the region $v_1 > 0, \sum_{j=1}^{n-1} v_j^p < r^p$ into the sphere $S_{p,n}(r)$. It follows that

$$
d\sigma_2(v_1, \ldots, (r^p - \sum_{j=1}^{n-1} v_j^p)^{1/p}) = |\frac{\partial}{\partial v_1} T_r \wedge \frac{\partial}{\partial v_2} T_r \wedge \cdots \wedge \frac{\partial}{\partial v_{n-1}} T_r| dv_1 \ldots dv_{n-1}.
$$

Since $\frac{\partial}{\partial v_j} T_r = e_j + c_j e_n$, where $c_j = -(v_j^{p-1} / (r^p - \sum_{i=1}^{n-1} v_i^p))^{1/p-1} = - \left( \frac{v_j}{v_n} \right)^{p-1}$ and

$$(e_1 + c_1 e_n) \wedge (e_2 + c_2 e_n) \wedge \cdots \wedge (e_{n-1} + c_{n-1} e_n) = e_{1,2,\ldots,n-1} + c_1 e_{n,2,\ldots,n-1} + c_2 e_{1,n,3,\ldots,n-1} + \cdots + c_{n-1} e_{1,2,\ldots,n-2,n},$$

it is seen that

$$
|\frac{\partial}{\partial v_1} T_r \wedge \cdots \wedge \frac{\partial}{\partial v_{n-1}} T_r| = \sqrt{\sum_{j=1}^{n-1} c_j^2} = \sqrt{\sum_{j=1}^n |v_j|^{2p/q}} / (r^p - \sum_{i=1}^{n-1} |v_i|^p)^{1/q}. \tag{10}
$$

From (10) and (8) it follows that the measure $\sigma_{p,n}$ coincides with Rachev and Ruschendorf’s measure $\mu_{p,n}$ defined (see their equation (3.1)) on the portion of $S_{p,n}(1)$ with all $v_i > 0$ and analogously elsewhere by

$$
\mu_{p,n}(A) = \int_U I_A(v_1, \ldots, v_{n-1}, (1 - \sum_{j=1}^{n-1} v_j^p)^{1/p}) \frac{1}{(1 - \sum_{j=1}^{n-1} v_j^p)^{(p-1)/p}} dv_1 \ldots dv_{n-1}, \tag{11}
$$

where $U = \{(v_1, \ldots, v_{n-1}) : v_i \geq 0, \sum_{j=1}^{n-1} v_j^p < 1 \}$, and $A$ is any measurable subset of $S_{p,n}(1)$.

3.3 Minkowski uniformity under $F_p$

The probability $P$ is uniform with respect to $\mu$ if $P$ is absolutely continuous with respect to $\mu$ and the R-N derivative $f = \frac{dP}{d\mu}$ is constant. The probability measure $P$ is uniform on the sphere $S_{p,n}(1)$ if $f$ is constant and the measure $\mu$ is surface area. If $X_1, \ldots, X_n$ are iid $F_p$ and

$$
R = \frac{1}{(\sum_{j=1}^n |X_j|^p)^{1/p}}(X_1, \ldots, X_n),
$$

then

$$
\mu(X_1, \ldots, X_n) = \frac{1}{(\sum_{j=1}^n |X_j|^p)^{1/p}}(X_1, \ldots, X_n)
$$
then \( n^{1/p}R \) is distributed uniformly with respect to Minkowski surface area on the sphere \( S_p,n(n^{1/p}) \). This follows from the literature and our calculations above but for a self contained proof consider for \( g : R^n_+ \rightarrow R \) measurable the integral \( I = \int g(v)dv \). Let \( T(v) = \left( \frac{v_1}{\sum_{i=1}^{n} v_i^{1/p}}, \ldots, \frac{v_n}{\sum_{i=1}^{n} v_i^{1/p}} \right) \), \( \sum_{i=1}^{n} v_i^{1/p} \). Here the domain of \( T \) is the region \( \sum_{i=1}^{n} v_i^{1/p} \leq t^p \). The range of \( T \) is \( \{(u_1, \ldots, u_{n-1}, r) : u_i \geq 0, \sum_{i=1}^{n-1} u_i^{1/p} \leq 1, r \geq 0 \} \). Then \( T \) is invertible with inverse \( T^{-1}(u_1, \ldots, u_{n-1}, r) = (ru_1, \ldots, ru_{n-1}, r(1 - \sum_{i=1}^{n-1} u_i^{1/p})^{1/p}) \). Therefore

\[
I = \int \cdots \int g(v_1, \ldots, v_n)dv_1 \cdots dv_n
= \int \cdots \int g(ru_1, \ldots, ru_{n-1}, r(1 - \sum_{i=1}^{n-1} u_i^{1/p})^{1/p})|J(u_1, \ldots, u_{n-1}, r)|du_1 \cdots du_{n-1}dr
= \int_0^\infty \int_U g(ru_1, \ldots, ru_{n-1}, r(1 - \sum_{j=1}^{n-1} u_j^{1/p})^{1/p})r^{n-1}d\mu_{p,n}(u)dr
\]

since \( J = \frac{1}{(1 - \sum_{i=1}^{n-1} u_i^{1/p})^{1/p}} \). In particular, if \( f \) is the joint density of \( X_1, \ldots, X_n \) with respect to \( V_n \) and \( M \) is a measurable subset of \( S_p,n(1) \), then letting \( A = R^{-1}(M) \), one has the probability

\[
P[R \in M] = P[(X_1, \ldots, X_n) \in A] = \int_0^\infty \int_M f(ru_1, \ldots, ru_n)r^{n-1}d\mu_{p,n}(u)dr
= \frac{p^{n/q}}{(2\Gamma(1/p))^n} \int_M \int_0^\infty r^{n-1}e^{-r^p/p}drd\sigma_{p,n}(u)
= \frac{p^{n-1}\Gamma(\frac{n}{p})}{2^n\Gamma(\frac{p}{p})^n} \sigma_{p,n}(M).
\]

Therefore, if \( X_1, \ldots, X_n \) are iid \( F_p \) and \( R \) is given in \([12] \), then the density of \( R \) is uniform with respect to \( \sigma_{p,n} \).

4 Proof of the theorem for \( \ell_p(n) \), \( 1 \leq p < \infty \)

The techniques of Billingsley \([8] \) on weak convergence of probability measures and uniform integrability will be employed to prove Theorem \([4] \).

Let \( (\Omega, \mathcal{A}, P) \) denote a probability space on which is defined the sequence \( U_j \sim U(0, 1) \), \( j = 1, 2, \ldots \), endowed with independent random variables, identically distributed uniformly on the unit interval. Fixing \( p \in [1, \infty) \), one has that the iid \( F_p \)-distributed sequence of random variables \( X_1, X_2, \ldots \) can be expressed as \( X_j = F_p^{-1}(U_j) \). The usual empirical distribution based on the iid \( X_j \) is then

\[
G_n(t) = \frac{1}{n} \# \{ X_j \leq t \} = \frac{1}{n} \# \{ U_j \leq F_p(t) \} = U_n(F_p(t)),
\]

where \( U_n \) is the empirical distribution, edf, of the iid uniforms. Suppressing the dependence on \( \omega \) in both, define for each \( n = 1, 2, \ldots \) the empirical process \( \Delta_n(u) = \sqrt{n}(U_n(u) - u) \) for \( u \in [0, 1] \) and (see also \([11] \))

\[
V_n = \sqrt{n} \left( \frac{1}{n} \sum_{j=1}^{n} |F_p^{-1}(U_j)|^p - 1 \right).
\]
The metric $d_0$ of $[3]$ (see Theorem 14.2) on $D[0,1]$ is employed. It is equivalent to the Skorohod metric generating the same sigma field $\mathcal{D}$ and $D[0,1]$ is a complete separable metric space under $d_0$.

The processes of basic interest are $\sqrt{n}(F_n(t) - F_p(t)), t \in (-\infty, \infty)$. As commonly utilized in the literature, the alternative parametrization relative to $u \in [0,1]$ is sometimes adopted below in terms of which the basic process is expressed as

$$\sqrt{n}(F_p(F_p^{-1}(u)) - u).$$

(13)

In terms of this parametrization the processes concerning us are $E_n(u) = \sqrt{n}(G_n(F_p^{-1}(u)\psi_p(X_n)) - F_p(F_p^{-1}(u)))$; these generate the same measures on $(D[0,1], \mathcal{D})$ as the processes. Weak convergence of the processes $E_n$ will be proven.

Introduce for $c > 0$ the mappings $\phi(c, \cdot)$ defined by $\phi(c, u) = F_p(cF_p^{-1}(u)), 0 < u < 1, \phi(c, 1) = 1,$ and $\phi(c, 0) = 0$. Then if

$$E_n^{(1)}(u) = \Delta_n(\phi((\frac{V_n}{\sqrt{n}} + 1)^{1/p}, u)), \quad (14)$$

and

$$E_n^{(2)}(u) = \sqrt{n}\left(\phi((\frac{V_n}{\sqrt{n}} + 1)^{1/p}, u) - \phi(1, u)\right)$$

(15)

one observes that

$$E_n(u) = E_n^{(1)}(u) + E_n^{(2)}(u).$$

The following concerning product spaces will be used repeatedly. Take the metric $d$ on the product space $M_1 \times M_2$, as

$$d((x_1, y_1), (x_2, y_2)) = \max\{d_1(x_1, x_2), d_2(y_1, y_2)\}, \quad (16)$$

where $d_i$ is the metric on $M_i$.

**Proposition 1.** If $(X_n(\omega), Y_n(\omega))$ are $(\Omega, \mathcal{A}, P)$ to $(M_1 \times M_2, M_1 \times M_2)$ measurable random elements in a product $M_1 \times M_2$ of two complete separable metric spaces then weak convergence of $X_n \Rightarrow X$ and $Y_n \Rightarrow Y$ entails relative sequential compactness of the measures $\nu_n(\cdot) = P[(X_n, Y_n) \in \cdot]$ on $(M_1 \times M_2, M_1 \times M_2)$ with respect to weak convergence.

**Proof:** By assumption and Prohorov’s theorem (see Theorem 6.2 of $[3]$) it follows that the sequences of marginal measures $\nu_n^X, \nu_n^Y$ are both tight. Let $\epsilon > 0$ be arbitrary, $K_X \in \mathcal{M}_1$ be compact and satisfy $P[\omega \in \Omega : X_n(\omega) \in K_X] \geq 1 - \epsilon/2$ for all $n$ and $K_Y \in \mathcal{M}_2$ be such that $P[\omega \in \Omega : Y_n(\omega) \in K_Y] \geq 1 - \epsilon/2$ for all $n$. Then $K_X \times K_Y \in M_1 \times M_2$ is compact (since it is is clearly complete and totally bounded under the metric when - as they do here - those properties of the sets $K_X$ and $K_Y$ hold) and since

$$P[(X_n \in K_X) \cap (Y_n \in K_Y)] = 1 - P[(X_n \in K_X)^c \cup (Y_n \in K_Y)^c]$$

and $P[(X_n \in K_X)^c \cup (Y_n \in K_Y)^c] \leq 2 \cdot \epsilon/2$, one has for all $n$

$$\nu_n(K_X \times K_Y) = P[(X_n, Y_n) \in K_X \times K_Y] \geq 1 - \epsilon.$$

Thus the sequence of measures $\nu_n$ is tight and by Prohorov’s theorem (see Theorem 6.1 of $[3]$) it follows that there is a probability measure $\hat{\nu}$ on $(M_1 \times M_2, M_1 \times M_2)$ and a subsequence $n'$ so that $\nu_{n'} \Rightarrow \hat{\nu}$. $\square$

It is shown next (see $[3]$) that $\sqrt{n}(G_n(t\psi_p(X_n)) - F_p(t\psi_p(X_n))) \Rightarrow B(F_p(t))$. 

Lemma 1. Let $1 \leq p < \infty$. Then (see [2])
\[
\mathbb{E}_n^{(1)} \Rightarrow B,
\]
where $B$ is a Brownian bridge process on $[0,1]$.

Proof: The random time change argument of Billingsley [3], page 145 is used. There, the set $D_n \subset D[0,1]$ of non-decreasing functions $\phi : [0,1] \rightarrow [0,1]$ is employed and here it is first argued that the functions $\phi(c,-)$, for $c > 0$ fixed are in $D_n$. For $u_0 \in (0,1)$ one calculates the derivative
\[
\frac{d}{du} \phi(c,u)|_{u=u_0} = \phi_u(c,u_0) = \frac{cF_p(cF_p^{-1}(u_0))}{f_p(F_p^{-1}(u_0))}
\]
from which continuity of $\phi(c,-)$ on $(0,1)$ follows. Consider $u_n \rightarrow 1$. Let $1 > \epsilon > 0$ be arbitrary and $a \in (-\infty,\infty)$ be such that $F_p(t) > 1 - \epsilon$ for $t > a/2$. Let $N < \infty$ be such that $n > N$ entails $F_p^{-1}(u_n) > a/c$ Then for $n > N$ one has $\phi(c,u_n) \geq F_p(a) > 1 - \epsilon = \phi(c,1) - \epsilon$. Since $\phi(c,-)$ is plainly increasing on $(0,1)$, for $n > N$ one has $|\phi(c,1) - \phi(c,u_n)| < \epsilon$. Thus $\phi(c,-)$ is continuous at 1 and a similar argument shows it to be continuous at 0. It is therefore a member of $D_n$.

Next, consider the distance $d_0(\phi(c,-),\phi(1,-))$. Details of its definition are in [3] in material surrounding equation (14.17), but the only feature utilized here is that for $x,y \in C[0,1]$, $d_0(x,y) \leq \|x-y\|_{\infty}$. Denoting $\frac{\partial}{\partial \alpha}\phi(c,u)|_{\alpha=a}$ by $\phi_{\alpha}(a,u)$ one has for some $\xi = \xi_u$ between $c$ and 1
\[
\phi(c,u) - \phi(1,u) = \phi_{\xi}(\xi,u)(c-1) = f_p(\xi F_p^{-1}(u))F_p^{-1}(u)(c-1)
\]
and since uniformly on compact sets $c \in [a,b] \subset (0,\infty)$ one has $\sup_{-\infty<x<\infty}|xf_p(cx)| < B$ for some $B < \infty$ it follows that for $\epsilon - 1 < \delta < 1$ one has
\[
\|\phi(c,-) - \phi(1,-)\|_{\infty} \leq B\delta.
\]
Therefore, if $C_n \overset{p}{\rightarrow} 1$ then $d_0(\phi(C_n,-),\phi(1,-)) \overset{p}{\rightarrow} 0$. Since if $X \sim F_p$ then $|X|^p \sim G(1/p,p)$, the gamma distribution with mean 1 and variance $p^2/p = p$, it follows from the ordinary CLT that $\frac{1}{\sqrt{n}}V_n \overset{d}{\rightarrow} N(0,1)$. Thus the $D$-valued random element $\Phi_n = \phi(\frac{V_n}{\sqrt{n}} + 1)^{1/p,-}$ satisfies $\Phi_n \Rightarrow \phi(1,-) = c(\cdot)$, the identity. As is well known, $\Delta_n \Rightarrow B$, so if $(\Delta_n,\Phi_n) \overset{D}{\rightarrow} (B,e)$ then as shown in [3] (see material surrounding equation (17.7) there) and consulting [4], $\mathbb{E}_n^{(1)} = \Delta_n o \Phi_n \Rightarrow B o e = B$.

Consider the measures $\nu_\epsilon$ on $D \times D$ whose marginals are $(\Delta_n,\Phi_n)$ and let $n'$ be any subsequence. It follows from Proposition [4] that there is a probability measure $\bar{\nu}$ on $D \times D$ and a further subsequence $n''$ such that $\nu_{n''} \Rightarrow \bar{\nu}$. Here $\nu_{n''}$ has marginals $(\Delta_{n''},\Phi_{n''})$ and so $\bar{\nu}$ must be a measure whose marginals are $(B,e)$; so $(\Delta_{n''},\Phi_{n''}) \overset{D}{\rightarrow} (B,e)$. It follows that $\mathbb{E}_n^{(1)} \Rightarrow B$. Since every subsequence has a further subsequence converging weakly to $B$, it must be that $\mathbb{E}_n^{(1)} \Rightarrow B$. □

Lemma 2. Let $1 \leq p < \infty$. Then (see [5])
\[
\mathbb{E}_n^{(2)} \Rightarrow Z \frac{F_p^{-1}(\cdot) f_p(F_p^{-1}(\cdot))}{\sqrt{p}},
\]
where $Z \sim N(0,1)$.
Proof: One has, for $1 \leq p < \infty$
\[
\phi(c, u) - \phi(1, u) = [\phi_c(1, u) + \epsilon(c, u)](c - 1),
\]
where for fixed $u \in (0, 1)$, $\epsilon(c, u) \to 0$ as $c \to 1$ and for $\delta$ sufficiently small and uniformly on $|c - 1| < \delta$, $\|\epsilon(c, \cdot)\|_\infty < A$ for some $A < \infty$. With $C_n = (\frac{1}{\sqrt{n}} + 1)^{1/p}$ it follows that
\[
\mathbb{E}_n^{(2)}(u) = \phi_c(1, u)\sqrt{n}(\frac{V_n}{\sqrt{n}} + 1)^{1/p} - 1 + o_p(1)
\]
\[
= \frac{\phi_c(1, u)}{p}V_n + o_p(1)
\]
\[
\overset{d}{\to} \frac{\phi_c(1, u)}{\sqrt{p}} Z,
\]
where $Z \sim N(0, 1)$. □

Denote by $\mu_n$ the joint probability measure on $D \times D$ of $(\mathbb{E}_n^{(1)}, \mathbb{E}_n^{(2)})$. Applying Proposition 1 as in Lemma 1 there is a subsequence $\mu_{n_i}$ and a probability measure $\bar{\mu}$ on $D \times D$ whose marginals, in light of Lemmas 1 and 2, must be $(B, \frac{\phi_c(1, u)}{\sqrt{p}} Z)$. It will be shown next that for any such measure $\bar{\mu}$, one has
\[
\text{cov}(B(u), Z) = -F_p^{-1}(u)f_p(F_p^{-1}(u)).
\]

An arbitrary sequence $\{V_n\}_{n \geq 1}$ of random variables is uniformly integrable (ui) if
\[
\lim_{\alpha \to \infty} \sup_{n} \int_{|V_n| > \alpha} |V_n(\omega)|dP(\omega) = 0.
\]
The fact that if $\sup_n E[|V_n|^{1+\epsilon}] < \infty$ for some $\epsilon > 0$ then $\{V_n\}$ is ui will be employed as will Theorem 5.4 of [3] which states that if $\{V_n\}$ is ui, and $V_n \to V$ then $\lim_{n \to \infty} E[V_n] = E[V]$. It is well known that in a Hilbert space $(L_2(\Omega, \mathcal{A}, P))$ here a set is weakly sequentially compact if and only if it is bounded and weakly closed (see Theorem 4.10.8 of [4]).
In the following it is more convenient to deal with the original $X_i$. It is assumed, without loss of generality and for ease of notation, that the subsequence is the original $n$ so $\mu_n \Rightarrow \bar{\mu}$.

Lemma 3. For $\bar{\mu}$
\[
\text{cov}(B \circ F_p(t), Z) = -t f_p(t).
\]

Proof: Fix $t \in (-\infty, \infty)$ and let $C_n = \sqrt{n}(G(t) - F_p(t))$ and $D_n = \sqrt{n}(W_n - 1)$, where $W_n = \frac{1}{n} \sum_{i=1}^{n} |X_i|^p$. The expectations $E[|C_nD_n|^2]$ will be computed and it will be shown that the supremum over $n$ is finite. In particular, it will be demonstrated that $E[C_n^2D_n^2] = n^{-2}(K_1n^2 + K_2n)$ so that $C_nD_n$ is ui. Define $A_i = |X_i|^p - 1$ and $B_i = I(\infty, t)(X_i) - F_p(t)$. Note that $E[A_i] = E[B_i] = 0$, $i = 1, \ldots, n$ that $A_i$’s for different indexes are independent and the same applies to $B_i$’s. Furthermore, $E[A_i^2] = \frac{1}{n}p^2 = p$ and $E[B_i^2] = F_p(t)(1 - F_p(t))$. One has $(C_nD_n)^2 = \frac{1}{n^2}(\sum_{i=1}^{n} A_i)^2(\sum_{i=1}^{n} B_i)^2$ so that $C_n^2D_n^2$ is the sum of four terms $S_1, S_2, S_3, S_4$ where
\[
S_1 = \sum_{i=1}^{n} A_i^2 \sum_{i=1}^{n} B_i^2,
S_2 = \sum_{i=1}^{n} B_i^2 \sum_{i \neq j \neq k} A_i A_j B_k B_l,
S_3 = \sum_{i \neq j} A_i^2 \sum_{i \neq j} B_i A_j B_k,
S_4 = \sum_{i \neq j} A_i A_j \sum_{i \neq j} B_i B_l.
\]
Consider first $S_2$. A typical term in the expansion will be $A_i^2 B_n C_v$, where $u \neq v$. Only the ones for which $i$ equals $u$ or $v$ have expectations possibly differing from 0, but if $i = u$ then since $B_n$ is independent and 0 mean it too has expectation 0. Thus $E[S_2] = 0$. The same argument applies to $E[S_4]$. In $S_4$ we'll have, using similar arguments, $E[S_4] = \sum_{i \neq j} E[A_i B_i] E[A_i B_j] = (n^2 - n) E[A_1 B_1] E[A_2 B_2]$. In the case of $S_1$ one has

$$E[S_1] = E\left[ \sum_{i=1}^{n} A_i^2 B_i^2 + \sum_{u \neq v} A_u^2 B_v^2 \right]
= nE[A_i^2 B_i^2] + (n^2 - n)E[A_i^2 E[B_i^2]].$$

Therefore

$$\sup_n E[|C_n D_n|^2] = \sup_n n^{-2}(K_1 n^2 + K_2 n) < \infty,$$

where

$$K_1 = E[A_1 B_1] E[A_2 B_2] + E[A_i^2 E[B_i^2]]$$
and

$$K_2 = E[A_i^2 B_i^2] - E[A_1 B_1] E[A_2 B_2] - E[A_i^2 E[B_i^2]].$$

It follows that $C_n D_n$ is ui and $\lim_{n \to \infty} E[C_n D_n] = E[B \circ F_p(t) Z_1]$ where $Z_1 \sim N(0, p)$. Noting that for some $K < \infty$

$$\sup_{w \geq 0} \left| \frac{F_p(t w^{1/p}) - F_p(t) - p^{-1} t f_p(t) (w - 1)}{(w - 1)^2} \right| < K$$

one has

$$E[n \left( F_p(t W_n^{1/p}) - F_p(t) - p^{-1} t f_p(t) (W_n - 1) \right)^2] \leq n E[(W_n - 1)^4] = \frac{3p^2}{n} + \frac{6p^3}{n^2} \to 0$$
and it is seen that $\|\sqrt{n}(F_p(t W_n^{1/p}) - F_p(t) - p^{-1} t f_p(t) \sqrt{n}(W_n - 1))\|_2 \to 0$. It follows now from $\|C_n\|_2 = F_p(t)(1 - F_p(t))$ and weak sequential compactness by passing to subsequences, that

$$\lim_{n \to \infty} E[C_n \sqrt{n}(F_p(t W_n^{1/p}) - F_p(t))] = E[B \circ F_p(t) Z].$$

On the other hand, by a direct computation,

$$E[\sqrt{n}(G_n(t) - F_p(t))(\sqrt{n}(W_n - 1))] = n E[G_n(t)(W_n - 1)]
= \frac{n}{n^2} \sum_{i=1}^{n} \sum_{j=1}^{n} E[I_{(-\infty, t)}(X_i)(|X_j|^p - 1)]
= \frac{1}{n} \sum_{i=1}^{n} E[I_{(-\infty, t)}(X_i)(|X_i|^p - 1)]
= E[I_{(-\infty, t)}(X_1)(|X_1|^p - 1)]
= \int_{-\infty}^{t} |x|^p \frac{1}{2\Gamma(1/p)} e^{-|x|^{p}/p} dx - F_p(t),$$

so that letting $u = x$ and $dv = x^{p-1} e^{-x^p/p} dx$ one has $\int_{0}^{t} x^p e^{-x^p/p} dx = -x e^{-x^p/p} |_{0}^{t} + \int_{0}^{t} e^{-x^p/p} dx$ and hence

$$E[\sqrt{n}(G_n(t) - F_p(t))(\sqrt{n}(W_n - 1))] = -t f_p(t) + F_p(t) - F_p(t) = -t f_p(t).$$
Therefore,

\[ E[B \circ F_p(t)Z] = -tf_p(t). \]

\[ \square \]

A plot of a portion of the covariance function close to 0 appears in Figure 1 and a comparison of variances on the same scale in Figure 2.

**Figure 1:** Comparison of covariance functions; empiric is Brownian bridge

**Figure 2:** Comparison of variance functions for \( p = 2 \): solid is Brownian bridge

**Lemma 4.** Let \( 1 \leq p < \infty \) be fixed and \( E_n(u) = E_n^{(1)}(u) + E_n^{(2)}(u), 0 \leq u \leq 1 \) (see equations 14 and 15). Then there is a Gaussian process \( E(u) = B(u) + \frac{F_p^{-1}(u)F_p(F_p^{-1}(u))}{\sqrt{p}}Z \) satisfying 17 for which \( E_n \Rightarrow E \).
Proof: From what has been done so far it follows that for an arbitrary subsequence \( n' \) of \( n \) the measures \( \mu_{n'} \) on \( D \times D \) which are the joint distributions of \((E^{(1)}_n, E^{(2)}_n)\) have a further subsequence \( n'' \) and there is a probability measure \( \tilde{\mu} \) on \( D \times D \) for which \( \mu_{n''} \Rightarrow \tilde{\mu} \). This measure has marginals \((B, \tilde{\omega}^{(1)}_{n''})\) and the covariance of \( B(u) \) and \( Z \) is given by (17). Since \( \tilde{\mu} \) concentrates on \( C \times C \) and \( \theta(x, y) = x + y \) is continuous thereon, one has a probability measure \( \eta \) on \( D \) defined for \( A \in D \) by \( \eta(A) = \tilde{\mu}(\theta^{-1}A) \) and the support of \( \eta \) is contained in \( C \). It will now be argued that this measure \( \eta \) is Gaussian. It is convenient to do this in terms of the original \( X_j \)’s. Let \( X_1, X_2, \ldots \), be iid \( F_p \), fix \( -\infty < t_1 < t_2 < t_k < \infty \), and consider the random vectors \( W^{(n)}(t) = (W_n(t_1), \ldots W_n(t_k)), \) where

\[
W_n(t) = \sqrt{n} \frac{1}{n} \sum_{i=1}^{n} (I_{-\infty, t_i}(\frac{X_v}{\psi_p(X_n)}) - F_p(t)).
\]

Since \( E^{(n'') \iff (E^{(p')}_n(t_1), \ldots, E^{(p')}_n(t_k))) \Rightarrow (E(F_p(t_1)), \ldots, E(F_p(t_k))) = W \) and since \( E \) is continuous wp 1 and \( \psi_p(X_n) \rightarrow 1 \) one has also \( W^{(n'')}(t/\psi_p(X_n)) \Rightarrow W \). Noting that

\[
W_n(t_j/\psi_p(X_n)) = \sqrt{n} (G(t_j) - F_p(t_j)) - (F_p(t_j/\psi_p(X_n)) - F_p(t_j)))
\]

it is seen that \( W \), being the limit in law of sums of iid well-behaved vectors, is a multivariate normal. Furthermore, the limiting finite dimensional marginals do not depend on the subsequence. Therefore, the measure \( \eta \) is unique and Gaussian and the claim has been proven. \( \square \)

5 \( \ell_\infty(n) \)

Convergence also holds in the case \( p = \infty \), where one can arrive at the correct statement and conclusion purely formally by taking the limit as \( p \rightarrow \infty \) in the statement of Theorem 11 so \( F_\infty \) is the uniform on \([-1, 1] \), the random vector \( Y_n = (Y_{1n}, \ldots, Y_{nn}) \in S_{\infty, n}(1), \) and for \( t \in [-1, 1] \)

\[
\sqrt{n}(F_n(t) - \frac{1 + t}{2} I_{[-1, 1]}(t)) \Rightarrow B \circ F_\infty(t).
\]

This follows from:

1. If \( \psi_\infty(X_n) = \max\{|X_1|, \ldots, |X_n|\} \), then \( \psi_\infty(X_n) \in [0, 1] \) and one has for \( 1 > v > 0 \), that \( P[\psi_\infty(X_n) \leq v] = (\int_{-v}^{v} \frac{1}{2} dx)^n = v^n \) so \( \psi_\infty(X_n) \overset{P}{\rightarrow} 1 \) and

\[
P[n(\psi_\infty(X_n) - 1) \leq -v] = (1 - \frac{v}{n})^n \rightarrow e^{-v},
\]

the term in the limit process additional to the Brownian bridge part (the right-most term in (13)) washes out and one has as limit simply the Brownian bridge \( B(\frac{1 + t}{2} I_{[-1, 1]}(t)) \).

Furthermore (see also [13]) the measure \( \sigma_\infty, n \) on \( S_{\infty, n}(1) \) coincides with ordinary Euclidean measure.
6 Acknowledgment

Leonid Bunimovich introduced me to the question of coordinate distribution in $\ell_2$. Important modern references resulted from some of Christian Houdré’s suggested literature on the isoperimetry problem in $\ell_p$. Thanks also are hereby expressed to the referees and editors of this journal for their careful attention to my paper and valuable comments.

References


