A mixed limit theorem for stable random fields.

By

T.V. Kurien\textsuperscript{1}

and

Jayaram Sethuraman\textsuperscript{2}

Department of Statistics
Florida State University

January 1991
Department of Statistics Technical Report Number M 832
U.S. Army Technical Report Number D 114
AFOSR Technical Report Number 91-251

Key Words: Mixed limit theorem, stable random field, pattern analysis.

\textsuperscript{1} Research supported by the Air Force Office of Scientific Research under Grant AFOSR91-0048.

\textsuperscript{2} Research supported by Army Research Office Grant DAAL03-90-G-010.
A mixed limit theorem for stable random fields.

By

T.V. Kurien

and

Jayaram Sethuraman

Department of Statistics
Florida State University

Abstract

A mixed distributional limit theorem for a stable random field of index $0 < \alpha < 2$ is derived. These random fields are of special interest in pattern analysis, in particular, in pattern synthesis. This paper considers the case when the underlying graph that the random field is defined on is linear. This result is encouraging insofar as it shows that the mixed limit theorems do exist in the stable case. The final limiting distribution can be written in terms of the stable process of index $\alpha$ in $D[0,1]$.

Introduction.

A picture or image can be considered to be a collection of gray levels located at the vertices of a set, that is to say that it can be represented by the collection $\{U_\alpha : \alpha \in V\}$ where $U_\alpha$ represents a gray level located at the site $\alpha$ of a set $V$. For instance, a raster image on a TV screen can be viewed as a collection of gray levels $U_{(i,j)}$ located at $(i,j) \in \{0, \ldots, n\} \times \{0, \ldots, n\}$. More structure can be given to the set of sites $V$ by considering a graph $G = (V,e)$ with vertex set $V$ and edge set $e$. The graph structure can be so chosen as to reflect the relations that we would expect to see between sites in the particular class of images that we are considering.
A random image can therefore be modelled by a collection of random variables \( U_\alpha \) located at the vertices \( \alpha \) of a graph \( G = (V, e) \). Gibbs distributions have been used to specify the distribution of random images because the graph structure \( G \) can be tailored to take into account the dependencies intrinsic to the structure of the images arising in practice.

The general pattern analysis problem can be broadly described as follows. There is a true image \( I \) which cannot be observed. However, a deformed version \( I^p \) of \( I \) can be observed. The physical process deforming the true image \( I \) is due to the process of observation itself. This physical process could be mathematically modelled as an additive noise or by some other process. The pattern analysis problem is to reconstruct the true image \( I \) based on the observed image \( I^p \) by using a statistical procedure that is optimal in some sense.

Bayesian Image analysis first formulates a prior distribution for \( I \) which incorporates what is known about the structure of the true image \( I \). Gibbs distributions have been used as prior distributions in Bayesian Image analysis [5]. The next step in Bayesian Image analysis is to update the prior notion of \( I \) using \( I^p \). This gives rise to a a posterior distribution on the set of images. Under fairly general conditions, the posterior distribution is a Gibbs distribution. The posterior mean or the posterior mode is usually used as an estimate of the true image \( I \).

Bayesian Image analysis has had a good deal of success in dealing with pattern analytic problems. The HANDS project at Brown University [8] is an application of Bayesian Image analysis to recognizing biological shapes like human hands. For Bayesian Image analysis, it is necessary to generate an observation from the posterior distribution, and this is commonly done using the Gibbs Sampler [5], which in turn uses the Metropolis algorithm [9]. In fact, we can simulate observations from the prior distribution using the same Gibbs Sampler, to evaluate its suitability for the problem at hand. This is called pattern synthesis. However these methods are not very fast and require a good deal of computer time.

A natural question that now arises is whether we can replace simulations with direct mathematical approximations of such distributions when the number of vertices (sites) of the graph is very large. The answer to this question is in the affirmative, for special graphs, provided that the number of vertices increases to infinity and the scale of gray levels present at each site increases at a certain rate. Results that describe this kind of behavior are called mixed limit theorems and have been studied previously by Grenander and Sethuraman [7] and also by Chow [4].

A graph of interest in pattern analysis is the linear connection graph \( G_1 \) defined below. \( G_1 \) has vertex set \( \{0, \ldots, n\} \) and edge set \( \{(i, i + 1) : 0 \leq i < n\} \).

We will introduce some notation in order to state the mixed limit theorem for the graph \( G_1 \). Let \( A(\cdot) \) be a symmetric density and \( Q(\cdot) \) be a density function on \( \mathbb{R} \). Let \( (U_0, \ldots, U_n) \) which reside at the vertices of \( G_1 \) have a joint density proportional to

\[
\prod_{i=0}^{n-1} A(\sqrt{n}(u_{i+1} - u_i)) \prod_{i=0}^n Q(u_i).
\]

The function \( A \) is called the acceptor function in the pattern analysis literature. The acceptor function models the local dependence between the \( U_i \)'s. The situation when the local
dependence specified by the pattern analytic problem at hand should make neighboring \( U_i \) alike can be modelled by taking the acceptor function \( A \) to be a symmetric density which is decreasing on \( \mathbb{R}^+ \). The situation when the local dependence should make neighboring \( U_i \)'s different can be modelled by taking \( A \) to be a symmetric density which has bounded support and is increasing on \( \mathbb{R}^+ \). In the equation above, the factor of \( \sqrt{n} \) represents an increase in the scale of gray levels present at each site as alluded to earlier.

Define
\[
X_i = \sqrt{n} U_i
\]
and \( Y_i = X_i - X_0, i = 0, \ldots, n. \)

Note that \( Y_0 = 0 \) and that the distributions of the \( X_i \)'s do depend on \( n \). We shall not explicitly indicate this dependence for the rest of this paper. The case when \( A(\cdot) \) has a second moment and \( Q(x) \) is essentially of the form \( \exp(-x^2) \) was the first situation to be considered. Mixed limit theorems for the distribution in \( C[0,1] \) of the process formed by linearly interpolating the \( X_i \)'s have been derived by Chow [4]. The limiting process in this case is a Gaussian process with a known covariance function. This corresponds to the situation in which the acceptor function \( A \) has thin tails. The case where the acceptor function \( A \) has thick tails is the situation that we will study in this paper. Let \( A(\cdot) \) be a stable density with index \( \alpha, 0 < \alpha < 2 \). Then \( A \) has thicker than normal tails. Theorem 2.2 is a mixed limit theorem for the distribution of a family of stochastic processes on \([0,1]\) based on the \( X_i \)'s. The processes that we will study are related to the stable processes (see Breiman [10]). Since stable processes are defined on the space \( D[0,1] \), this is the natural topological space to prove distributional limit theorems for the processes under study.

**Section 2.**

Construct the stochastic processes \( G_n(t) \) and \( X_n(t) \) on \([0,1]\) as follows:

\[
X_n(t) = n^{-1/\alpha} X_{[nt]}, \\
G_n(t) = n^{-1/\alpha} Y_{[nt]}.
\]

The random variables \( X_i \) and \( Y_i \) are as defined in equation (1.1). The processes \( X_n(\cdot) \) and \( G_n(\cdot) \) are random variables in \( D[0,1] \). A theorem that establishes the weak convergence in \( D[0,1] \) of the stochastic processes \( G_n(t) \) and \( X_n(t) \) will be referred to as a mixed limit theorem.

For the rest of this paper we shall assume that \( Q(x) = \exp(-x^2) \) and that \( A(\cdot) \) is a symmetric stable density of index \( \alpha, 0 < \alpha < 2 \).

Let \( f_\alpha(\cdot) \) denote the symmetric stable density with index \( \alpha \). Let \( T_0 \) have density \( f_\alpha(\cdot) \). Then \( f_\alpha(\cdot) \) is a bounded density and \( P\{T_0 > x\} \sim x^\alpha \). We will make use of these facts in the proof of Lemma 2.1.

Now \((X_0, Y_1, \ldots, Y_n)\) has a joint density proportional to

\[
\prod_{i=0}^{n-1} f_\alpha(y_{i+1} - y_i) \exp(-n^{-1} \sum_{i=1}^{n} (y_i + x_0)^2 - x_0^2/n).
\]
In Theorem 2.1 we will show that $G_n(\cdot)$ converges weakly to a process $G(\cdot)$ in $D[0, 1]$. Corollary 2.1 of this theorem will show that $X_n(t)$ will also converge weakly in $D[0, 1]$ to the same process $G(\cdot)$.

The method of proof of Theorem 2.1 is as follows. We will show that the distribution of $G_n(\cdot)$ conditional on $X_0 = x_0$ converges weakly in $D[0, 1]$. We then show that the distribution of $X_0$ converges strongly on $\mathbb{R}$. A result of Sethuraman [3] which is stated as Lemma 2.2 in the proof of Theorem 2.1 will establish the weak convergence of $G_n(\cdot)$. Lemma 2.1 is a weak convergence result for stable laws which will be used in the proof of Theorem 2.2.

Let $X$ and $Y$ be random variables. Then $X \overset{d}{=} Y$ denotes that $X$ and $Y$ are equal in distribution. Let $\{X_n, n = 1, \ldots\}$ be a sequence of random variables on a complete separable metric space $S$ and let $X$ and $X_0$ be random variables in $S$. Then $X_n \overset{w}{\rightarrow} X$ denotes that $\mathcal{L}(X_n)$ converges weakly to $\mathcal{L}(X)$ in $S$. $\mathcal{L}(X_n|X_0 = x_0)$ denotes the conditional distribution of $(X_n$ given that $X_0 = x_0)$. Let $R(t)$ be a stationary and independent increment process in $D[0, 1]$ with $R(t) \overset{d}{=} t^{1/\alpha}T_0$. Lemma 2.1 below explicitly constructs such a process $R(\cdot)$ in $D[0, 1]$.

**Lemma 2.1**

Let $0 < \alpha < 2$ and let $T_0, T_1, \ldots$ be i.i.d. with density $f_\alpha$. Consider the partial sums $S_i = \sum_{j=1}^i T_j$, $i = 1, 2, \ldots$, $S_0 = 0$ and let the process $(R_n(t), 0 \leq t \leq 1)$ be defined by

$$R_n(t) = n^{-1/\alpha}S_{[nt]}.$$  

Then: $R_n(\cdot) \overset{w}{\rightarrow} R(\cdot)$ in $D[0, 1]$.

**Proof**

The method of proof is the usual one, i.e., we shall first prove finite-dimensional convergence and then verify the appropriate tightness conditions.

Let us calculate the characteristic function $\phi_n(\cdot, t)$ of $R_n(t), t \in [0, 1]$.

$$\phi_n(u, t) = E(\exp(iuR(t)))$$

$$= E(\exp(iun^{-1/\alpha}\sum_{1 \leq j \leq [nt]} T_j))$$

$$= \exp(-|u|^\alpha d_n(t))$$  \hspace{1cm} (2.3)$$

where $d_n(t) = (\frac{\inf(T_j)}{n})^{1/\alpha} \rightarrow t^{1/\alpha}$ as $n \rightarrow \infty$ for $t \in [0, 1]$.

Let $t_0 = 0 < t_1 < \cdots < t_k < t_{k+1} = 1$, and let $f_\alpha(\cdot)$ be the symmetric stable density with index $\alpha$.

Let

$$X_{i,n} = R_n(t_i) - R_n(t_{i-1}), \quad 1 \leq i \leq k + 1, \quad n = 1, 2, \ldots$$
Note that \( \{X_{i,n}, i = 1, \ldots, k+1\} \) are mutually independent random variables. Let \( f_{i,n} \) be the density of \( X_{i,n} \). Then by (2.3) it is clear that \( f_{i,n} \) is given by

\[
f_{i,n}(x) = f_{\alpha} \left( \frac{x}{d_n(t_j - t_{j-1})} \right) / d_n(t_j - t_{j-1}).
\]

Since \( X_{i,n} \) are mutually independent and \( R_n(t_i) = \sum_{j=0}^{i} X_{j,n}, i = 1, \ldots, k+1 \), the density of \( (R_n(t_1), \ldots, R_n(t_k)) \) is given by

\[
\prod_{i=1}^{k} f_{i,n}(y_i - y_{i-1}) \text{ where } y_0 = 0.
\]

This density converges to the density of \( (R(t_1), \ldots, R(t_k)) \). Hence the finite dimensional distributions converge. We will verify tightness next.

Let \( 0 = t_0 < t_1 < t_2 < t_3 < t_4 = 1, \varepsilon > 0 \) and let \( U, V \) be stable symmetric random variables with index \( \alpha \). Then by the mutual independence of the increments \( X_{i,n} \), and the facts about the tail probabilities of stable symmetric random variables, it follows that

\[
P\{|R_n(t_3) - R_n(t_2)| \geq \varepsilon, |R_n(t_2) - R_n(t_1)| \geq \varepsilon\}
\]

\[
= P\{|R_n(t_3) - R_n(t_2)| \geq \varepsilon\} \cdot P\{|R_n(t_2) - R_n(t_1)| \geq \varepsilon\}
\]

\[
= P\{|U| \geq \frac{\varepsilon}{d_n(t_2 - t_1)}\} \cdot P\{|V| \geq \frac{\varepsilon}{d_n(t_3 - t_2)}\}
\]

\[
\leq \frac{M \cdot (t_3 - t_1)^2}{\varepsilon^{2\alpha}}.
\]

Using Theorem (15.6) in Billingsley [1], tightness of the distributions is immediate. This proves the lemma. \( \square \)

Define \( \nu(x_0, x), f(u) \) and \( \mu(u) \) as follows

\[
\nu(x_0, x) = \exp\left( - \int_{0}^{1} (x(t) + x_0)^2 dt \right) \left( \int_{D[0,1]} \exp\left( - \int_{0}^{1} (x(t) + x_0)^2 dt \right) dP(x) \right)^{-1}
\]

\[
f(u) = \int_{D[0,1]} \exp\left( - \int_{0}^{1} (x(t) + u)^2 dt \right) dP(x)
\]

\[
\mu(u) = f(u) / \int_{\mathbb{R}} f(y) dy.
\]

**Note** The finiteness of these integrals becomes clear in the proof of Theorem 2.1 below.

In order to state the mixed limit theorem for \( G_n \) we introduce the following notation.

Let \( Q_n = \mathcal{L}(G_n), P_n = \mathcal{L}(R_n), P = \mathcal{L}(R) \). For \( x_0 \in \mathbb{R} \) let \( Q_{n,x_0} = \mathcal{L}(G_n|X_0 = x_0) \). These are probability measures on \( D[0,1] \).
Theorem 2.1

Let $A(\cdot)$ be a stable density of index $\alpha$, $0 < \alpha < 2$. Then $Q_n$ converges to $Q$ weakly in $D[0,1]$ where the density of $Q$ is given by

$$dQ(x) = \frac{\exp\left(-\int_0^1 (x(t) - \int_0^1 x(u)du)^2 dt\right)}{\int_{D[0,1]} \exp\left(-\int_0^1 (x(t) - \int_0^1 x(u)du)^2 dt\right) dP(x)} dP(x).$$

Proof

If we omit the exponential term in (2.2), then $G_n(\cdot)$ and $R_n(\cdot)$ have the same distribution.

To study $Q_n$, we shall first show that $Q_n,Z_0$ converges weakly in $D[0,1]$. We will then show that the densities of $X_0$ converge. This proves strong convergence for the densities of $X_0$ via Scheffe's theorem [2]. This will prove a weak convergence result for the distribution of $Q_n(\cdot)$ on $D[0,1]$, via a result of Sethuraman [3] stated below.

Lemma 2.2

Let $\Lambda_n$ be a sequence of probability measures on $V \times W$ where $V$ and $W$ are topological spaces. Let $\mu_n$ be the marginal distribution of $\Lambda_n$ on $V$ and $\nu_n(v,\cdot)$ be the conditional p. m. of $\Lambda_n$ on $W$. Suppose that $\mu_n(A) \to \mu(A) \ \forall A \subseteq V$ and $\nu_n(v,\cdot) \to \nu(v,\cdot)$, weakly for almost all $v$ (w. r. t $\mu$). Then, $\Lambda_n \to \Lambda$ weakly where

$$\Lambda(A \times B) = \int_A \nu(v, B) \ d\mu(v)$$

for each measurable rectangle $A \times B \in V \times W$.

We shall apply this theorem where $W = D[0,1]$, $V = \mathbb{R}$, $\mu_n$ is the distribution of $X_0$, and $\nu_n(x_0,\cdot) = \mathcal{L}(Q_n(\cdot)|X_0 = x_0)$.

We will now show the weak convergence of $\mathcal{L}(Q_n(\cdot)|X_0 = x_0)$ by considering it's Radon-Nikodym derivative with respect to $P_n$ and using the fact that $P_n \overset{w}{\to} P$. We shall denote a function $x(\cdot)$ in $D[0,1]$ by $x$ from now on. Let $\frac{dQ_n,z_0}{dP_n}(x)$ denote the Radon-Nikodym derivative of $Q_n,z_0$ with respect to $P_n(\cdot)$ at $x$ for $x$ in $D[0,1]$. For $x$ in $D[0,1]$ let

$$k_n(x) = \exp\left(-n^{-1} \sum_{i=1}^n (x(i/n) + x_0)^2\right)$$

and $k(x) = \exp\left(-\int_0^1 (x(t) + x_0)^2 dt\right)$.

Then

$$\frac{dQ_n,z_0}{dP_n}(x) = \frac{k_n(x)}{\int_{D[0,1]} k_n(y) dP_n(y)}.$$
Let \( f \) be a bounded continuous function from \( D[0,1] \) to \( \mathbb{R} \). Then
\[
\int_{D[0,1]} f(x) dQ_{n,x_0} = \int_{D[0,1]} f(x) k_n(x) dP_n / \int_{D[0,1]} k_n(x) dP_n. \tag{2.4}
\]
The weak convergence of \( Q_{n,x_0} \) will be proved by showing that
\[
\int_{D[0,1]} f(x) k_n(x) dP_n(x) \rightarrow \int_{D[0,1]} f(x) k(x) dP(x) \quad \text{and}
\int_{D[0,1]} k_n(x) dP_n(x) \rightarrow \int_{D[0,1]} k(x) dP(x). \tag{2.5}
\]
This will prove that \( Q_{n,x_0} \xrightarrow{w} Q_{x_0} \) and that
\[
\frac{dQ_{x_0}}{dP}(z) = k(x) / \int_{D[0,1]} k(y) dP(y).
\]
Lemma 2.3 will help us to verify this.

**Lemma 2.3**
Let \( f : \mathbb{R} \rightarrow \mathbb{R} \) be continuous and \( x_n \rightarrow x \in D[0,1] \). Then,
\[
n^{-1} \sum_{i=1}^{n} f(x_n(i/n)) \rightarrow \int_0^1 f(x(t)) \, dt.
\]

**Proof**
Without loss of generality, we can assume that \( f(x) = x \) since the map \( x \rightarrow f(x) \) induced on \( D[0,1] \) by \( f \) is a continuous map from \( D[0,1] \rightarrow D[0,1] \). Hence, it suffices to show that
\[
n^{-1} \sum_{i=1}^{n} x_n(i/n) \rightarrow \int_0^1 x(t) \, dt.
\]
Let us define
\[
y_n(t) = \begin{cases} x_n(i/n) & \text{if } (i-1)/n < t \leq i/n \\ x_n(0) & \text{if } t = 0. \end{cases}
\]
Then, \( n^{-1} \sum_{i=1}^{n} x_n(i/n) = \int_0^1 y_n(t)dt \). Since \( x_n \rightarrow x \) in \( D[0,1] \), there exist \( \lambda_n \) which are strictly increasing, continuous and surjective mappings of \( [0,1] \rightarrow [0,1] \) such that
\[
x_n(\lambda_n(t)) \rightarrow x(t) \text{ uniformly in } t
\]
and \( \lambda_n(t) \rightarrow t \) uniformly in \( t \). \( \tag{2.6} \)
Let $s \in [0,1]$ be a continuity point of $x(\cdot)$. The following argument shows that $y_n(s) \to x(s)$.

There exists $i_n$ such that $(i_n - 1)/n < s \leq i_n/n$. Let $s_n^* = \lambda_n^{-1}(i_n/n)$. Then $y_n(s) = x_n(\lambda_n s_n^*)$. Hence
\[
\begin{align*}
|y_n(s) - x(s)| &\leq |x_n(\lambda_n s_n^*) - x(s)| + |x(s_n^*) - x(s)| \\
&\leq \sup_{t \in [0,1]} |x_n(\lambda_n t) - x(t)| + |x(s_n^*) - x(s)|. \tag{2.7}
\end{align*}
\]

Now $s_n^* \to s$ and since $s$ is a continuity point of $x(\cdot)$, $|x(s_n^*) - x(s)| \to 0$ as $n \to \infty$. The first term in (2.7) converges to 0 by (2.6). Hence $y_n(s) \to x(s)$, as asserted. Since the number of discontinuity points of $x(\cdot)$ is countable, $y_n(s) \to x(s)$ a. s. (Lebesgue). By the Dominated Convergence Theorem, and because $x_n$ and hence $y_n$ are uniformly bounded,
\[
\int_0^1 y_n(t)dt \to \int_0^1 x(t)dt.
\]

This proves Lemma 2.3.

Hence it follows that
\[
\begin{align*}
n^{-1} \sum_{i=1}^n x_n^2(i/n) &\to \int_0^1 x^2(t)dt \\
n^{-1} \sum_{i=1}^n x_n(i/n) &\to \int_0^1 x(t)dt. \tag{2.8}
\end{align*}
\]

Thus $\exp(-n^{-1} \sum_{i=1}^n (x(i/n) + x_0)^2) \to \exp(-\int_0^1 (x(t) + x_0)^2 dt)$ by (2.8) since $\exp(-t)$ is a bounded continuous function on $[0, \infty]$, $k_n(x_n) \to k(x)$ and $f(x_n)k_n(x_n) \to f(x)k(x)$ as $x_n \to x$ in $D[0,1]$. Also note that $k_n(x)$ and $k(x)$ are bounded.

Then by using Theorem 5.5 of Billingsley [1] it follows that both terms in (2.5) converge as stated.

Thus $Q_{n,x_0} \overset{w}{\to} Q_{x_0}$ as $n \to \infty$ and
\[
\frac{dQ_{x_0}}{dP} = \exp(-\int_0^1 (x(t) + x_0)^2 dt) \left(\int_{D[0,1]} \exp(-\int_0^1 (x(t) + x_0)^2 dt) dP(x)\right)^{-1} \tag{2.9}
\]

Now, we will show that the distribution of $X_0$ converges strongly by using Scheffe's theorem [2].

Define
\[
f_n(x_0) = \exp(-x_0^2/n) \int_{D[0,1]} \exp(-n^{-1} \sum_{i=1}^n (x(i/n) + x_0)^2) dP_n(x).
\]

Then, the density of $X_0$ is given by $f_n(x_0) \int_{-\infty}^\infty f_n(x_0)dx_0$. By (2.5) we know that $f_n(x_0) \to f(x_0)$ as $n \to \infty$ where
\[
f(x_0) = \int_{D[0,1]} \exp(-\int_0^1 (x(t) + x_0)^2 dt) dP(x). \tag{2.10}
\]
In order to successfully apply Schéffe’s theorem, we need to check that \( \int_{\mathbb{R}} f_n(x_0)dx_0 \rightarrow \int_{\mathbb{R}} f(x_0)dx_0 \).

By Fubini’s Theorem and elementary algebra we know that

\[
\int_{\mathbb{R}} f_n(x_0)dx_0 = \int_{\mathbb{R}} \exp\left(-\frac{x_0^2}{n}\right) \int_{D[0,1]} \exp\left(-n^{-1} \sum_{i=1}^{n} (x(i/n) + x_0)^2\right) dP_n(x)dx_0
\]

\[
= \int_{D[0,1]} \exp\left(-n^{-1} \sum_{i=1}^{n} x^2(i/n) + \left[\frac{1}{n(n+1)} \sum_{i=1}^{n} x(i/n)\right]^2\right) \times \int_{\mathbb{R}} \exp\left(-n/(n+1)(x_0 - \frac{1}{n(n+1)} \sum_{i=1}^{n} x(i/n))^2\right) dx_0 \ dP_n(x)
\]

\[
\rightarrow \sqrt{\pi} \int_{D[0,1]} \exp\left(-x^2(t) + \left(\int_{0}^{1} x(t)dt\right)^2\right) dP(x).
\]

This limit is equal to \( \int_{\mathbb{R}} f(x_0)dx_0 \), because

\[
\int_{\mathbb{R}} f(x_0)dx_0 = \int_{\mathbb{R}} \int_{D[0,1]} \exp\left(-\int_{0}^{1} (x(t) + x_0)^2dt\right) dP(x)dx_0
\]

\[
= \sqrt{\pi} \int_{D[0,1]} \exp\left(-\int_{0}^{1} x^2(t)dt + \left(\int_{0}^{1} x(t)dt\right)^2\right) dP(x).
\] (2.11)

Thus the distribution of \( X_0 \) converges strongly to a distribution with density function given by

\[
d\mu(x_0) = f(x_0) / \int_{\mathbb{R}} f(x_0)dx_0.
\] (2.12)

Hence, by Lemma 2.2 it follows that \( Q_n \) converges weakly to a distribution \( Q \) on \( D[0,1] \) with density given by

\[
dQ(x) = \left(\int_{\mathbb{R}} \nu(x, x_0)d\mu(x_0)\right) dP(x).
\] (2.13)

The integral in (2.13) can be further simplified by using (2.10) and (2.11).

\[
\int_{\mathbb{R}} \nu(x_0, x)d\mu(x_0) = \frac{\int_{\mathbb{R}} \exp\left(-\int_{0}^{1} (x(t) + x_0)^2dt\right) dx_0}{\int_{\mathbb{R}} f(x_0)dx_0}
\]

\[
= \frac{\exp\left(-\int_{0}^{1} x(t)dt - \int_{0}^{1} x(u)du^2dt\right)}{\int_{D[0,1]} \exp\left(-\int_{0}^{1} x(t)dt - \int_{0}^{1} x(u)du^2dt\right) dP(x)}.
\]

This completes the proof of Theorem 2.1

The following corollary is a mixed limit theorem for \( X_n(t) \). This shows that mixed limit theorems exist for the stable case and the limits are Gibbs distributions.
Corollary 2.1

$X_n(t)$ converges weakly in $D[0,1]$ to the process $G(\cdot)$ which has distribution $Q$.

Proof

Since $X_n(t) = n^{-1/\alpha + 1/2} X_0 + G_n(t)$, $0 < \alpha < 2$, and the distributions of $X_0$ converge strongly as $n \to \infty$ it follows from Slutsky's Theorem that the distributions of $X_n(\cdot)$ converges weakly in $D[0,1]$ to $Q(\cdot)$. □

Remark. The next problem that has to be addressed is that of explicitly evaluating the integral $\int_{D[0,1]} \exp(-\int_0^1 (x(t) - \int_0^1 x(u)du)^2 dt) dP(x)$. This will be necessary in order to simulate observations from the limiting distribution $Q$. 
References


