A purely martingale version of Lindeberg's central limit theorem

V. G. Vovk*
9-3-451 ulitsa Ramenki, Moscow 117607 Russia
`vovk@ium.ac.msk.su`

June 1995

Abstract

We prove Lindeberg's central limit theorem in Glenn Shafer's event-tree framework for probability theory.

1 Theorem

Our definitions follow Shafer's [6] Chapters 11 (event trees) and 12 (martingale probability). For the first time the idea of martingale probability seems to occur in Dawid [2] (for details, see [8], Section 3).

An *event tree* is a partially ordered set Γ that satisfies

Axiom E1 *Any two elements s and t of Γ are either ordered (i.e., $s \le t$ or $t \le s$) or divergent (i.e., there is no $u \in \Gamma$ such that $s \le u$ and $t \le u$).*

Axiom E2 *Γ has an initial element $\Box$ (that is, $\Box \le s$ for all $s \in \Gamma$).*

We call the elements of Γ *situations*, and we take $s \le t$ to mean that t cannot happen unless s happens earlier.

---

*The research described in this publication was made possible in part by Grant No. MRS000 from the International Science Foundation.

A *path* is a maximal set of ordered situations. We will consider only event trees for which all paths are finite; let Γ be such an event tree. We identify the paths in Γ with the terminal situations of Γ in a natural way: a path is identified with its terminal situation.

The *sample space* is the set of all paths. A *variable* is a real-valued function on the sample space.

A *process* is a real-valued function on the event tree. A *cut* (of a situation s) is a maximal set of divergent situations (after s). A *martingale* is a process μ which satisfies

Axiom M *If $s \in \Gamma$ and Ξ is a cut of s, then there exist situations $t_1, t_2 \in \Xi$ such that $\mu(t_1) \leq \mu(s) \leq \mu(t_2)$.*

A *catalog* is a linear space M of martingales[1]. A catalog M is *Doob* if it satisfies the following axioms:

Axiom V0 *M contains the constant martingales.*

Axiom V1 *If $\mu \in M$ and Ξ is a cut, then $\mathbf{E}_{\Xi}(\mu) \in M$.*

By $\mathbf{E}_{\Xi}(\mu)$ we mean the process μ' defined by

$$\mu'(s) := \mu(\Xi(s)),$$

where

$$\Xi(s) := \begin{cases} s, & \text{if } s \text{ is before } \Xi, \\ t \in \Xi \text{ such that } t \leq s, & \text{otherwise.} \end{cases}$$

We will consider only catalogs which are Doob and satisfy the following condition: if $\{\mu_\theta \,|\, \theta \in \Theta\}$ is a *locally finite* family of martingales (in the sense that for each situation s there exist only finitely many θ such that $\mu_\theta(s) \neq 0$), then $\sum_{\theta \in \Theta} \mu_\theta$ is a martingale.

Let M be a catalog on Γ; we say that (Γ, M) is a *martingale tree.* An *M-martingale difference* is a process of the form $s \mapsto \mu(s) - \mu(s^-)$, where s^- is the mother of s (except $\Box^- := \Box$) and $\mu \in M$. If $U = U(\omega)$ is a variable, we define the *upper* and *lower expectation* as follows:

$$\mathcal{E}^+[U \mid M] := \inf\{\mu(\Box) \mid \mu \in M,\ \mu(\omega) \geq U(\omega),\ \text{for all terminal } \omega \in \Gamma\},$$

---

[1] Note that under this definition we do not need de Finetti's assumption of coherence (which is about the same as the financial-theoretic requirement of nonexistence of arbitrage opportunities).

$$\mathcal{E}^-[U \mid M] := \sup\{\mu(\Box) \mid \mu \in M,\ \mu(\omega) \le U(\omega),\ \text{for all terminal } \omega \in \Gamma\}.$$

If s is a nonterminal situation and X is a partial process defined at least on the daughters of s, we define the *one-step-ahead expectations*

$$\mathbf{E}_s^+[X \mid M] := \inf\{\mu(s) \mid \mu \in M,\ \mu(t) \ge X(t),\ \text{for all daughters } t \text{ of } s\},$$

$$\mathbf{E}_s^-[X \mid M] := \sup\{\mu(s) \mid \mu \in M,\ \mu(t) \le X(t),\ \text{for all daughters } t \text{ of } s\}.$$

If U is the indicator of an event (*event* is a subset of the sample space), the definition of $\mathcal{E}^+$ and $\mathcal{E}^-$ specializes to the definition of *upper* and *lower probability*:

$$\mathcal{P}^+[E \mid M] := \mathcal{E}^+[\mathbf{I}(E) \mid M],$$

$$\mathcal{P}^-[E \mid M] := \mathcal{E}^-[\mathbf{I}(E) \mid M],$$

where E is an event and $\mathbf{I}$ stands for the indicator. Note that always

$$\mathcal{E}^-[U \mid M] = -\mathcal{E}^+[-U \mid M],$$

$$\mathcal{P}^-[E \mid M] = 1 - \mathcal{P}^+[E^c \mid M].$$

Let (Γ_n, M_n) be a sequence of martingale trees and, for each n, $U_n = U_n(\omega)$ be a variable in (Γ_n, M_n). We say that U_n *tends to* 0 *in probability*, $U_n(\omega) \xrightarrow{M_n} 0$, as $n \to \infty$, if

$$\forall \delta > 0\colon \lim_{n\to\infty} \mathcal{P}^+[|U_n| > \delta \mid M_n] = 0.$$

Notice that the definition of catalog implies that the lower expectation never exceeds the upper expectation.

The following theorem, which is the main result of this paper, is analogous to the martingale central limit theorem in Shiryaev [7], Theorem VII.8.4, and Liptser and Shiryaev [5], Theorem 5.5.9. We let $\mathcal{N}(du \mid m, \sigma^2)$ stand for the normal distribution in $\mathbb{R}$ with mean m and variance σ^2. Each expression $A(\omega)$ containing ω is interpreted as the variable $\omega \mapsto A(\omega)$, ω ranging over the paths.

Theorem 1 *Let (Γ_n, M_n) be a sequence of martingale trees and, for each n, X_n be an M_n-martingale difference in (Γ_n, M_n) and Ξ_n be a cut in Γ_n. Lindeberg's condition*

$$\forall \delta > 0\colon \sum_{s < \Xi_n(\omega)} \mathbf{E}_s^+[X_n^2 \mathbf{I}(|X_n| \ge \delta) \mid M_n] \xrightarrow{M_n} 0 \ \ (n \to \infty) \tag{1}$$

and the condition

$$\sum_{s<\Xi_n(\omega)} \mathbf{E}_s^{\pm}[X_n^2 \mid M_n] \xrightarrow{M_n} \sigma^2 \ \ (n \to \infty), \tag{2}$$

where σ is a positive constant, imply that, for every bounded continuous function $f\colon \mathbb{R} \to \mathbb{R}$,

$$\mathcal{E}^{\pm}\left[f\left(\sum_{s\le\Xi_n(\omega)} X_n(s)\right) \mid M_n\right] \longrightarrow \int_{-\infty}^{\infty} f(u)\,\mathcal{N}(du \mid 0,\sigma^2) \ \ (n\to\infty). \tag{3}$$

2 Proof

In the proof of Theorem 1 we shall use Lindeberg's method; the simplest variants of this proof appeared in [8] (the proof of Theorem 5) and [9] (the proof of Theorem 6).

First we notice that it suffices to prove that (1) and (2) imply that, for any bounded continuous f,

$$\limsup_{n\to\infty} \mathcal{E}^{+}\left[f\left(\sum_{s\le\Xi_n(\omega)} X_n(s)\right) \mid M_n\right] \le \int_{-\infty}^{\infty} f(u)\,\mathcal{N}(du \mid 0,\sigma^2). \tag{4}$$

(This is because this inequality will continue to hold if we replace f by $-f$.) Fix bounded continuous f and assume (1) and (2); we are required to prove (4). Without loss of generality we assume that each Ξ_n is the terminal cut of Γ_n; therefore, we can replace $\Xi_n(\omega)$ by ω in (4).

First we assume that f is a smooth function that is constant outside a compact interval. We introduce the function

$$g(h) := \sup_x \left| f(x+h) - f(x) - f'(x)h - \frac{1}{2}f''(x)h^2 \right|$$

and note that

$$\left| [f(x+h) - f(x)] - \left[f'(x)h + \frac{1}{2}f''(x)h^2\right] \right| \le g(h) \tag{5}$$

and, for some constant K,

$$g(h) \le K|h|^3,\ g(h) \le Kh^2,\ \forall h. \tag{6}$$

We shall suppose that $K > 2$.

Fix arbitrarily small $\delta \in \,]0,1[$. Let $\epsilon \in \,]0,1[$ be so small that, for all $v \in \mathbb{R}$ and all $\Delta \in \,]0, 2\epsilon]$,

$$\left| \int_{-\infty}^{\infty} f(u) \mathcal{N}(du \mid v, \Delta) - f(v) \right| \le \delta \tag{7}$$

and

$$\left| \int_{-\infty}^{\infty} f(u) \mathcal{N}(du \mid 0, \sigma^2 + \epsilon) - \int_{-\infty}^{\infty} f(u) \mathcal{N}(du \mid 0, \sigma^2) \right| \le \delta. \tag{8}$$

Let us check that such ϵ indeed exists. The existence of ϵ satisfying (7) follows from f being uniformly continuous and bounded—see, e.g., Feller [3], Lemma VII.1.1. To prove the existence of ϵ satisfying (8), it suffices to prove that the variation distance between the normal distributions $\mathcal{N}(du|0,\sigma^2)$ and $\mathcal{N}(du|0,\sigma^2+\epsilon)$ tends to 0 as $\epsilon \to 0$. By Shiryaev [7], Theorem III.9.1, we can here replace the variation distance with $1 - H$, where H is the corresponding Hellinger integral. An easy calculation shows that indeed $H \to 1$ as $\epsilon \to 0$.

Fix n so large that the lower probability of the conjunction of the three inequalities

$$\left| \sum_{s<\omega} \mathbf{E}_s^{\pm}[X_n^2 \mid M_n] - \sigma^2 \right| \le \epsilon, \tag{9}$$

$$\sum_{s<\omega} \mathbf{E}_s^{+}[X_n^2 \mathbf{I}(|X_n| \ge \delta^2) \mid M_n] \le \delta^4 \tag{10}$$

is at least $1 - \delta$. Inequality (10) implies

$$\sum_{s<\omega} \mathbf{E}_s^{+}[X_n^2 \mathbf{I}(|X_n| \ge \delta) \mid M_n] \le \delta \tag{11}$$

and, for all $s < \omega$,

$$\mathbf{E}_s^{+}[X_n^2 \mid M_n] \le 2\delta^4. \tag{12}$$

We shall show that for some martingale $\overline{F} \in M_n$ we have

$$\overline{F}(\Box) = \int_{-\infty}^{\infty} f(u) \mathcal{N}(du \mid 0, \sigma^2 + \epsilon) + (2 \sup |f''| + 8K\sigma^2 + 13K)\delta + \delta \tag{13}$$

and, for all ω satisfying (9) and (10),

$$\overline{F}(\omega) \ge f\left(\sum_{t \le \omega} X_n(t) \right). \tag{14}$$

Since ω satisfies (9) and (10) with lower probability at least $1-\delta$, our goal will be achieved (recall (8) and that δ can be taken arbitrarily small).

Put, for each situation $s \in \Gamma_n$,

$$F(s) := \int_{-\infty}^{\infty} f(S(s)+u)\,\mathcal{N}\left(du \mid 0, \sigma^2 + \epsilon - \sum_{t<s} \mathbf{E}_t^{\mathrm{mean}}[X_n^2 \mid M_n]\right),$$

where

$$S(s) := \sum_{t \le s} X_n(t), \;\; \mathbf{E}_s^{\mathrm{mean}} := \frac{1}{2}(\mathbf{E}_s^{+} + \mathbf{E}_s^{-}),$$

and $\mathcal{N}(du \,|\, m, v)$ is interpreted as $\mathcal{N}(du \,|\, m, 0)$ when $v<0$. The idea behind the construction of $\overline{F}$ is to "approximate", in some sense, F.

For each $s \in \Gamma_n$ and $v \in \mathbb{R}$ put

$$F_s(v) := \int_{-\infty}^{\infty} f(v+u)\,\mathcal{N}\left(du \mid 0, \sigma^2 + \epsilon - \sum_{t \le s} \mathbf{E}_t^{\mathrm{mean}}[X_n^2 \mid M_n]\right).$$

When t is a daughter of s and $t \le \omega$ for an ω satisfying (9) and (10), we find (using Taylor's formula and the fact[2] that (5) continues to hold when f is replaced by F_s):

$$\begin{aligned}
& F(t) - F(s) \\
&= F_s(S(s)+X_n(t)) - \textstyle\int_{-\infty}^{\infty} F_s(S(s)+u)\,\mathcal{N}\left(du \mid 0, \mathbf{E}_s^{\mathrm{mean}}[X_n^2 \mid M_n]\right) \\
&\le \Big[F_s(S(s)) + F_s'(S(s))X_n(t) + \tfrac{1}{2}F_s''(S(s))X_n^2(t) + g(X_n(t))\Big] \\
&\quad - \Big[F_s(S(s)) + \tfrac{1}{2}F_s''(S(s))\mathbf{E}_s^{\mathrm{mean}}[X_n^2 \mid M_n] \\
&\quad + \textstyle\int_{-\infty}^{\infty} g(u)\,\mathcal{N}\left(du \mid 0, \mathbf{E}_s^{\mathrm{mean}}[X_n^2 \mid M_n]\right)\Big] \\
&= F_s'(S(s))X_n(t) + \tfrac{1}{2}F_s''(S(s))\left(X_n^2(t) - \mathbf{E}_s^{\mathrm{mean}}[X_n^2 \mid M_n]\right) \\
&\quad + \Big\{g(X_n(t)) + \textstyle\int_{-\infty}^{\infty} g(u)\,\mathcal{N}\left(du \mid 0, \mathbf{E}_s^{\mathrm{mean}}[X_n^2 \mid M_n]\right)\Big\}.
\end{aligned} \tag{15}$$

---

[2] This fact follows from Kolmogorov's [4] Theorem IV.5.1: we can average (5) with respect to x distributed as

$$\mathcal{N}\left(du \mid v, \sigma^2 + \epsilon - \sum_{t \le s} \mathbf{E}_t^{\mathrm{mean}}[X_n^2 \mid M_n]\right).$$

Let us estimate the "remainder" in the braces in (15): using the inequality

$$g(h) \le K|h|^3 \,\mathbf{I}(|h| \le \delta) + Kh^2 \,\mathbf{I}(|h| > \delta) \le K\delta h^2 + Kh^2 \,\mathbf{I}(|h| > \delta), \qquad (16)$$

which follows from (6), we obtain

$$g(X_n(t)) \le K\delta X_n^2(t) + KX_n^2(t) \,\mathbf{I}(|X_n(t)| > \delta);$$

and using (16) and (12), we further obtain

$$\begin{gathered}
\textstyle\int_{-\infty}^{\infty} g(u) \,\mathcal{N}\left(du \mid 0, \mathbf{E}_s^{\mathrm{mean}}[X_n^2 \mid M_n]\right) \\
\le K\delta \textstyle\int_{-\infty}^{\infty} u^2 \,\mathcal{N}\left(du \mid 0, \mathbf{E}_s^{\mathrm{mean}}[X_n^2 \mid M_n]\right) \\
+K \textstyle\int_{-\infty}^{\infty} u^2 \,\mathbf{I}(|u| > \delta) \,\mathcal{N}\left(du \mid 0, \mathbf{E}_s^{\mathrm{mean}}[X_n^2 \mid M_n]\right) \\
\le K\delta \mathbf{E}_s^{\mathrm{mean}}[X_n^2 \mid M_n] + \tfrac{K}{\delta} \textstyle\int_{-\infty}^{\infty} |u|^3 \,\mathcal{N}\left(du \mid 0, \mathbf{E}_s^{\mathrm{mean}}[X_n^2 \mid M_n]\right) \\
\le K\delta \mathbf{E}_s^{\mathrm{mean}}[X_n^2 \mid M_n] + \tfrac{K}{\delta} \left(\mathbf{E}_s^{\mathrm{mean}}[X_n^2 \mid M_n]\right)^{3/2} \textstyle\int_{-\infty}^{\infty} |u|^3 \,\mathcal{N}(du \mid 0, 1) \\
\le K\delta \mathbf{E}_s^{\mathrm{mean}}[X_n^2 \mid M_n] + \tfrac{K}{\delta} \mathbf{E}_s^{\mathrm{mean}}[X_n^2 \mid M_n] 2\delta^2 3 \\
= 7K\delta \mathbf{E}_s^{\mathrm{mean}}[X_n^2 \mid M_n].
\end{gathered}$$

Therefore, we obtain from (15):

$$\begin{gathered}
F(t) - F(s) \\
\le F_s'(S(s)) X_n(t) + \tfrac{1}{2} F_s''(S(s)) \left(X_n^2(t) - \mathbf{E}_s^{\mathrm{mean}}[X_n^2 \mid M_n]\right) \\
+K\delta X_n^2(t) + KX_n^2(t) \,\mathbf{I}(|X_n(t)| > \delta) + 7K\delta \mathbf{E}_s^{\mathrm{mean}}[X_n^2 \mid M_n].
\end{gathered} \qquad (17)$$

For each $s \in \Gamma_n$ define the *level* of s to be

$$\mathrm{lev}(s) := \#\{t \mid t < s\}.$$

If $s \in \Gamma_n$ and μ is a process, the *s-local component* $\mathbf{Loc}_s(\mu)$ of μ is the only s-local process (a process is *s-local* if $\mu(\Box) = 0$ and $\mu(t_1) = \mu(t_2)$ whenever $t_1 \ne s$ and t_2 is a daughter of t_1) which satisfies

$$\mathbf{Loc}_s(\mu)(t) - \mathbf{Loc}_s(\mu)(s) = \mu(t) - \mu(s), \ t \in \mathbf{D}(s),$$

$\mathbf{D}(s)$ being the daughter set of s. Note that μ can be recovered from $\mathbf{Loc}_s(\mu)$, $s \in \Gamma_n$:

$$\mu = \mu(\Box) + \sum\{\mathbf{Loc}_s(\mu) \mid s \text{ is a nonterminal situation}\}.$$

For each $s \in \Gamma_n$ fix three s-local martingales $\mu_s, \mu_s^+, \mu_s^- \in M_n$ such that, on $\mathbf{D}(s)$,

$$\mu_s \geq X_n^2 \,\mathbf{I}(|X_n| > \delta) - \mathbf{E}_s^+[X_n^2 \,\mathbf{I}(|X_n| > \delta) \mid M_n] - \delta 2^{-\operatorname{lev}(s)},$$

$$\mu_s^+ \geq X_n^2 - \mathbf{E}_s^+[X_n^2 \mid M_n] - \delta 2^{-\operatorname{lev}(s)},$$

$$\mu_s^- \leq X_n^2 - \mathbf{E}_s^-[X_n^2 \mid M_n] + \delta 2^{-\operatorname{lev}(s)}.$$

Now we can define the martingale $\overline{F}$. The initial value $\overline{F}(\Box)$ is given by (13), and for each nonterminal $s \in \Gamma_n$ the s-local component of $\overline{F}$ is defined as the following weighted sum of μ_s, μ_s^+, μ_s^-, and the s-local component of X_n:

$$\mathbf{Loc}_s(\overline{F}) = F_s'(S(s))\,\mathbf{Loc}_s(X_n) + K\delta\mu_s^+ + K\mu_s + \frac{1}{2}F_s''(S(s))\mu_s^{\pm},$$

where

$$\mu_s^{\pm} := \begin{cases} \mu_s^+, & \text{if } F_s''(S(s)) > 0, \\ \mu_s^-, & \text{if } F_s''(S(s)) < 0. \end{cases}$$

We deduce from (17):

$$\begin{gathered} F(t) - F(s) \leq \overline{F}(t) - \overline{F}(s) \\ +\tfrac{1}{2}\sup|f''|\left(\tfrac{\mathbf{E}_s^+[X_n^2|M_n] - \mathbf{E}_s^-[X_n^2|M_n]}{2} + \delta 2^{-\operatorname{lev}(s)}\right) \\ +K\delta\left(\mathbf{E}_s^+[X_n^2 \mid M_n] + \delta 2^{-\operatorname{lev}(s)}\right) \\ +K\left(\mathbf{E}_s^+[X_n^2\,\mathbf{I}(|X_n| > \delta) \mid M_n] + \delta 2^{-\operatorname{lev}(s)}\right) + 7K\delta\mathbf{E}_s^{\mathrm{mean}}[X_n^2 \mid M_n]. \end{gathered}$$

Therefore, for each ω satisfying (9) and (10) we obtain (without loss of generality we assume $\epsilon \leq \delta$):

$$\begin{gathered} F(\omega) - F(\Box) \leq \overline{F}(\omega) - \overline{F}(\Box) \\ +\tfrac{1}{2}\sup|f''|(\epsilon + 2\delta) + K\delta(\sigma^2 + \epsilon + 2\delta) + K(\delta + 2\delta) + 7K\delta(\sigma^2 + \epsilon) \\ \leq \overline{F}(\omega) - \overline{F}(\Box) + (2\sup|f''| + 8K\sigma^2 + 13K)\,\delta \end{gathered}$$

(see (9) to (11)). So we obtain:

$$\begin{gathered} \overline{F}(\omega) \geq \overline{F}(\Box) + F(\omega) - F(\Box) - (2\sup|f''| + 8K\sigma^2 + 13K)\,\delta \\ = F(\omega) + \delta = \textstyle\int_{-\infty}^{\infty} f(u)\,\mathcal{N}(du \mid S(\omega), \Delta) + \delta \geq f(S(\omega)) \end{gathered}$$

(see (7)), where $\Delta \in [0, 2\epsilon]$. Hence we established (14), which completes the proof under the assumption that f is smooth and constant outside a finite interval.

Lemma 2 *For any $\sigma > 0$, $\epsilon > 0$, and bounded continuous $f\colon \mathbb{R} \to \mathbb{R}$, there exists a smooth and constant outside a finite interval function $\tilde{f}$ such that $f \le \tilde{f}$ and*

$$\int_{-\infty}^{\infty} \tilde{f}(u)\,\mathcal{N}(du \mid 0,\sigma^2) \le \int_{-\infty}^{\infty} f(u)\,\mathcal{N}(du \mid 0,\sigma^2) + \epsilon. \tag{18}$$

Proof There are an increasing finite sequence $t_1, \dots, t_k$ of real numbers and a function $g\colon \mathbb{R} \to \mathbb{R}$ which is constant on each of the intervals

$$\left]-\infty, t_1\right], \left]t_1, t_2\right], \dots, \left]t_{k-1}, t_k\right], \left]t_k, \infty\right[$$

and satisfies $g(-\infty) = g(\infty)$ and

$$f \le g, \ \int_{-\infty}^{\infty} g(u)\,\mathcal{N}(du \mid 0,\sigma^2) \le \int_{-\infty}^{\infty} f(u)\,\mathcal{N}(du \mid 0,\sigma^2) + \frac{\epsilon}{2}.$$

It is well known that there exists a smooth function $\phi\colon \mathbb{R} \to [0,1]$ such that $\phi(t) = 1$ for $t \le 0$ and $\phi(t) = 0$ for $t \ge 1$ (see, e.g., (7.1) of Billingsley [1]). Choose $\delta > 0$ which does not exceed any of the values $\frac{t_2 - t_1}{2}, \dots, \frac{t_k - t_{k-1}}{2}$. Using the function ϕ, one can easily construct a smooth function $\tilde{f}$ such that:

- $g \le \tilde{f}$,
- $\tilde{f}(t) = g(t)$ for each $t \notin \bigcup_{i=1}^{k} \left]t_i - \delta, t_i + \delta\right[$;
- $\tilde{f}(t) \in [g(t_i-), g(t_i+)]$ when $t \in \left]t_i - \delta, t_i + \delta\right[$.

Choosing sufficiently small δ, we ensure

$$\int_{-\infty}^{\infty} \tilde{f}(u)\,\mathcal{N}(du \mid 0,\sigma^2) \le \int_{-\infty}^{\infty} g(u)\,\mathcal{N}(du \mid 0,\sigma^2) + \frac{\epsilon}{2}.$$

$\square$

Now we can easily finish the proof. We are required to establish (4) for any bounded continuous f. By Lemma 2, there exists a smooth and constant outside a finite interval function $\tilde{f}$ such that $f \le \tilde{f}$ and (18) holds. Since (4) holds for $\tilde{f}$, we have:

$$\begin{aligned} &\limsup_{n\to\infty} \mathcal{E}^+[f(S) \mid M_n] \le \limsup_{n\to\infty} \mathcal{E}^+[\tilde{f}(S) \mid M_n] \\ &\le \textstyle\int_{-\infty}^{\infty} \tilde{f}(u)\,\mathcal{N}(du \mid 0,\sigma^2) \le \int_{-\infty}^{\infty} f(u)\,\mathcal{N}(du \mid 0,\sigma^2) + \epsilon. \end{aligned}$$

Since ϵ can be arbitrarily small, this proves (4).

References

[1] P. Billingsley. *Convergence of probability measures.* New York: Wiley (1968)

[2] A. P. Dawid. Calibration-based empirical probability (with discussion). *Ann. Statist.* **13**, 1251–1285 (1985)

[3] W. Feller. *An introduction to probability theory and its applications.* Vol. 2. 2nd ed. New York: Wiley (1971)

[4] A. N. Kolmogorov. *Foundations of the theory of probability.* New York: Chelsea (1950)

[5] R. S. Liptser, A. N. Shiryaev. *Theory of martingales.* Dordrecht: Kluwer (1989)

[6] G. Shafer. *The art of causal conjecture*, to appear

[7] A. N. Shiryaev. *Veroyatnost'.* 2nd ed. Moscow: Nauka (1989). English translation of the 1st edition: *Probability.* Berlin: Springer (1984).

[8] V. G. Vovk. A logic of probability, with application to the foundations of statistics (with discussion). *J. R. Statist. Soc.* B **55**, 317–351 (1993)

[9] V. G. Vovk. Forecasting point and continuous processes: prequential analysis. *Test* **2**, 189–217 (1993)

[10] V. G. Vovk. A purely martingale version of Kolmogorov's strong law of large numbers. *Probab. Theory Appl.*, to appear