

Statistical Skorohod embedding problem: optimality and asymptotic normality

Denis Belomestny¹ and John Schoenmakers²

March 24, 2015

Abstract

Given a Brownian motion B , we consider the so-called statistical Skorohod embedding problem of recovering the distribution of an independent random time T based on i.i.d. sample from B_T . Our approach is based on the genuine use of the Mellin transform. We propose a consistent estimator for the density of T , derive its convergence rates and prove their optimality. Moreover we address the question of asymptotic normality.

Keywords: Skorohod embedding problem, Mellin transform, multiplicative deconvolution, volatility density estimation.

1 Introduction

The so called Skorohod embedding (SE) problem or Skorohod stopping problem was first stated and solved by Skorohod in 1961. This problem can be formulated as follows.

Problem 1.1 (Skorohod Embedding Problem). *For a given probability measure μ on \mathbb{R} , such that $\int |x|d\mu(x) < \infty$ and $\int x d\mu(x) = 0$, find a stopping time T such that $B_T \sim \mu$ and $B_{T \wedge t}$ is a uniformly integrable martingale.*

The SE problem has recently drawn much attention in the literature, see e.g. Oblój, [3], where the list of references consists of more than 100 items. In fact, there is no unique solution to the SE problem and there are currently more than 20 different solutions available. This means that from a statistical point of view, the SE problem is not well posed. In this paper we study what we call *statistical Skorohod embedding* (SSE) problem.

Problem 1.2 (Statistical Skorohod Embedding Problem). *Based on i.i.d. sample X_1, \dots, X_n from the distribution of B_T consistently estimate the distribution of the random time $T \geq 0$, where B and T are assumed to be independent.*

The independence of B and T is needed to ensure the identifiability of the distribution of T from the distribution of B_T . In fact, due to the well-known scaling properties of the Brownian motion, the SSE problem is closely related to the multiplicative deconvolution problem and the problem of volatility estimation (see, e.g. Van Es et al [7], Van Es et al [6] and Van Es et al [5]). A standard approach to such type of problems is to first make a log-transformation and then solve the resulting additive deconvolution problem by means of the standard kernel density deconvolution technique. Our approach is different

¹D. Belomestny, Duisburg-Essen University Thea-Leymann-Str. 9, D-45127 Essen, Germany, denis.belomestny@uni-due.de

Partially supported by the Deutsche Forschungsgemeinschaft through the SFB 823 “Statistical modeling of nonlinear dynamic processes”

²Weierstrass Institute for Applied Analysis and Stochastics, Mohrenstr. 39, 10117 Berlin, Germany, schoenma@wias-berlin.de

and makes use of the Mellin transform device, which, in view of the well known properties of the Mellin transform, seems to be appropriate here. We construct a consistent estimator for the density of T and derive its convergence rates in different norms. Furthermore, we show that the obtained rates are optimal in minimax sense. The asymptotic normality of the proposed estimator is addressed as well. Let us stress that the results on optimality of the rates and asymptotic normality can not be derived from the known results on additive deconvolution problems.

The paper is organised as follows. In Section 2 we discuss some properties of the Mellin transform and introduce our main estimation procedure. Section 3 is devoted to the convergence properties of the estimator constructed in the previous section. In particular, we prove upper bounds for the pointwise estimation risk and show that these upper bounds are optimal in minimax sense. In Section 4 asymptotic normality of the proposed estimator is analysed. Some numerical examples can be found in Section 5. Finally, all proofs are collected in Section 6.

2 Construction of the estimator for p_T

Let B be a Brownian motion and let a random variable $T \geq 0$ be independent of B . We then have,

$$(1) \quad X := B_T \sim \sqrt{T} B_1$$

and the problem of reconstructing T is related to a multiplicative deconvolution problem. While for additive deconvolution problems the Fourier transform plays an important role, here we can conveniently use the Mellin transform.

Definition 2.1. Let ξ be a non-negative random variable with a probability density p_ξ , then the *Mellin transform* of p_ξ is defined via

$$(2) \quad \mathcal{M}[p_\xi](z) := \mathbb{E}[\xi^{z-1}] = \int_0^\infty p_\xi(x) x^{z-1} dx$$

for all $z \in \mathcal{S}_\xi$ with $\mathcal{S}_\xi = \{z \in \mathbb{C} : \mathbb{E}[\xi^{\operatorname{Re}z-1}] < \infty\}$.

Since p_ξ is a density, it is integrable and so at least $\{z \in \mathbb{C} : \operatorname{Re}(z) = 1\} \subset \mathcal{S}_\xi$. Under mild assumptions on the growth of p_ξ near the origin, one obtains

$$\{z \in \mathbb{C} : 0 \leq a_\xi < \operatorname{Re}(z) < b_\xi\} \subset \mathcal{S}_\xi$$

for some $0 \leq a_\xi < 1 \leq b_\xi$. Then the Mellin transform (2) exists and is analytic in the strip $a_\xi < \operatorname{Re} z < b_\xi$. For example, if p_ξ is essentially bounded in a right-hand neighbourhood of zero, we may take $a_\xi = 0$ and $b_\xi = 1$. The role of the Mellin transform in probability theory is mainly related to the product of independent random variables: in fact it is well-known that the probability density of the product of two independent random variables is given by the Mellin convolution of the two corresponding densities. Due to (1), the SSE problem is closely connected to the Mellin convolution. Suppose that the random time T has a density p_T which is essentially bounded in a right-hand neighbourhood of zero and fulfills

$$(3) \quad \int_0^\infty x^\delta p_T(x) dx < \infty$$

for some $\delta > 0$. Since $\mathcal{S}_{|B_1|} \supset \{z \in \mathbb{C} : \operatorname{Re}(z) > 0\}$, we derive for $0 < \operatorname{Re}(z) \leq 1$,

$$\begin{aligned} \mathcal{M}[p_{|X|}](z) &= \mathbb{E}[|B_1|^{z-1}] \mathbb{E}[T^{(z-1)/2}] \\ &= \mathcal{M}[p_{|B_1|}](z) \mathcal{M}[p_T]((z+1)/2) = \frac{2^{(z-1)/2}}{\sqrt{\pi}} \Gamma(z/2) \mathcal{M}[p_T]((z+1)/2). \end{aligned}$$

As a result

$$\mathcal{M}[p_T](z) = \frac{\sqrt{\pi}}{2^{z-1}} \frac{\mathcal{M}[p_{|X|}](2z-1)}{\Gamma(z-1/2)}, \quad 1/2 < \operatorname{Re}(z) \leq 1$$

and the Mellin inversion formula yields

$$(4) \quad \begin{aligned} p_T(x) &= \frac{1}{2\pi} \int_{\gamma-i\infty}^{\gamma+i\infty} x^{-\gamma-iv} \mathcal{M}[p_T](\gamma+iv) dv \\ &= \frac{1}{\sqrt{\pi}} \int_{-\infty}^{\infty} x^{-\gamma-iv} \frac{\mathcal{M}[p_{|X|}](2(\gamma+iv)-1)}{2^{\gamma+iv} \Gamma(\gamma+iv-1/2)} dv \quad \text{for } 1/2 < \gamma \leq 1, \quad x > 0. \end{aligned}$$

Furthermore, the Mellin transform of $p_{|X|}$ can be directly estimated from the data X_1, \dots, X_n via the empirical Mellin transform:

$$(5) \quad \mathcal{M}_n[p_{|X|}](z) := \frac{1}{n} \sum_{k=1}^n |X_k|^{z-1}, \quad 1/2 < \operatorname{Re}(z) \leq 1,$$

where the condition $\operatorname{Re}(z) > 1/2$ guarantees that the variance of the estimator (5) is finite. Note however that the integral in (4) may fail to exist if we replace $\mathcal{M}[p_{|X|}]$ by $\mathcal{M}_n[p_{|X|}]$. We so need to regularize the inverse Mellin operator. To this end, let us consider a kernel $K(\cdot) \geq 0$ supported on $[-1, 1]$ and a sequence of bandwidths $h_n > 0$ tending to 0 as $n \rightarrow \infty$. Then we define, in view of (5), for some $3/4 < \gamma \leq 1$,

$$(6) \quad p_{n,\gamma}(x) := \frac{1}{\sqrt{\pi}} \int_{-\infty}^{\infty} x^{-\gamma-iv} K(vh_n) \frac{\mathcal{M}_n[p_{|X|}](2(\gamma+iv)-1)}{2^{\gamma+iv} \Gamma(\gamma-1/2+iv)} dv.$$

3 Convergence

For our convergence analysis, we will henceforth take the simplest kernel

$$K(y) = 1_{[-1,1]}(y),$$

but note that in principle other kernels may be considered as well. The next theorem states that $p_{n,\gamma}$ converges to p_T at a polynomial rate, provided the Mellin transform of p_T decays exponentially fast. We shall use throughout the notation $a_n \lesssim b_n$ if a_n is bounded by a constant multiple of b_n , independently of the parameters involved, that is, in the Landau notation $a_n = O(b_n)$, $n \rightarrow \infty$.

Theorem 3.1. *For any $\beta > 0$ and $L > 0$, introduce the class of functions*

$$\mathcal{C}(\beta, L) := \left\{ f : \int_{-\infty}^{\infty} |\mathcal{M}[f](1+iv)| e^{\beta|v|} dv \leq L \right\}.$$

Assume that $p_T \in \mathcal{C}(\beta, L)$ for some $\beta > 0$ and $L > 0$, and that (3) holds. Then for some constant C_L depending on L only, it holds

$$(7) \quad \sup_{x \geq 0} \mathbb{E} \left[\{x|p_T(x) - p_{n,1}(x)|\}^2 \right] \leq C_L \left[e^{-2\beta/h_n} + \frac{1}{n} e^{\pi/h_n} \right].$$

By next choosing

$$(8) \quad h_n = (\pi + 2\beta)/\log n$$

we arrive at the rate

$$(9) \quad \sup_{x \geq 0} \sqrt{\mathbb{E} \left[\{x|p_T(x) - p_{n,1}(x)|\}^2 \right]} \lesssim n^{-\frac{\beta}{\pi+2\beta}}$$

as $n \rightarrow \infty$.

$\mathcal{C}(\beta, L)$	$\mathcal{D}(\beta, L)$
$n^{-\frac{\beta}{\pi+2\beta}}$	$\log^{-\beta}(n)$

Table 1: Minimax rates of convergence for the classes $\mathcal{C}(\beta, L)$ and $\mathcal{D}(\beta, L)$.

Let us turn now to some examples.

Example 3.2. Consider the class of Gamma densities

$$p_T(x; \alpha) = \frac{x^{\alpha-1} \cdot e^{-x}}{\Gamma(\alpha)}, \quad x \geq 0$$

for $\alpha > 0$. Since

$$\mathcal{M}[p_T](z) = \frac{\Gamma(z + \alpha - 1)}{\Gamma(\alpha)}, \quad \operatorname{Re}(z) > 0,$$

we derive that $p_T \in \mathcal{C}(\beta, L)$ for all $0 < \beta < \pi/2$ and some $L = L(\beta)$ due to the asymptotic properties of the Gamma function (see Lemma 6.5 in Appendix). As a result, Theorem 3.1 implies

$$\sup_{x \geq 0} \sqrt{\mathbb{E} \left[\left\{ x |p_T(x) - p_{n,1}(x)| \right\}^2 \right]} \lesssim n^{-\rho}, \quad n \rightarrow \infty$$

for any $\rho < 1/4$.

If $\mathcal{M}[p_T]$ decays polynomially fast, we get the following result.

Theorem 3.3. *Consider the class of functions*

$$\mathcal{D}(\beta, L) = \left\{ f : \int_{-\infty}^{\infty} |\mathcal{M}[f](1+iv)| (1+|v|^\beta) dv \leq L \right\},$$

and assume that $p_T \in \mathcal{D}(\beta, L)$ for some $\beta > 0$ and $L > 0$. If (3) holds for some $\delta > 0$, then for some constant D_L ,

$$(10) \quad \sup_{x \geq 0} \mathbb{E} \left[\left\{ x |p_T(x) - p_{n,1}(x)| \right\}^2 \right] \leq D_L \left[h_n^{2\beta} + \frac{1}{n} e^{\pi/h_n} \right].$$

By choosing

$$(11) \quad h_n = \frac{\pi}{\log n - 2\beta \log \log n},$$

we arrive at

$$(12) \quad \sup_{x \geq 0} \sqrt{\mathbb{E} \left[\left\{ x |p_T(x) - p_{n,1}(x)| \right\}^2 \right]} \lesssim \log^{-\beta}(n), \quad n \rightarrow \infty.$$

The rates of Theorem 3.1 and Theorem 3.3 summarized in Table 1 are in fact optimal (up to a logarithmic factor) in minimax sense for the classes $\mathcal{C}(\beta, L)$ and $\mathcal{D}(\beta, L)$, respectively.

Theorem 3.4. Fix some $\beta > 1$. There are $\varepsilon > 0$ and $x > 0$ such that

$$\liminf_{n \rightarrow \infty} \inf_{p_n} \sup_{p_T \in \mathcal{C}(\beta, L)} \mathbf{P}_{p_T}^{\otimes n} \left(|p_T(x) - p_n(x)| \geq \varepsilon n^{-\frac{\beta}{\pi+2\beta}} \log^{-\rho}(n) \right) > 0,$$

$$\liminf_{n \rightarrow \infty} \inf_{p_n} \sup_{p_T \in \mathcal{D}(\beta, L)} \mathbf{P}_{p_T}^{\otimes n} \left(|p_T(x) - p_n(x)| \geq \varepsilon \log^{-\beta}(n) \right) > 0,$$

for some $\rho > 0$, where the infimum is taken over all estimators (i.e. all measurable functions of X_1, \dots, X_n) of p_T and $\mathbf{P}_{p_T}^{\otimes n}$ is the distribution of the i.i.d. sample X_1, \dots, X_n with $X_1 \sim W_T$ and $T \sim p_T$.

4 Asymptotic normality

In the case of $K(v) = 1_{[-1,1]}(v)$, the estimate $p_{n,\gamma}(x)$ can be written as

$$p_{n,\gamma}(x) := \frac{1}{\sqrt{\pi}} \int_{-1/h_n}^{1/h_n} \left[\frac{1}{n} \sum_{k=1}^n |X_k|^{2(\gamma+iv-1)} \right] \frac{x^{-\gamma-iv}}{2^{\gamma+iv} \Gamma(\gamma-1/2+iv)} dv$$

$$= \frac{1}{n} \sum_{k=1}^n Z_{n,k},$$

where

$$Z_{n,k} := \int_{-1/h_n}^{1/h_n} |X_k|^{2(\gamma+iv-1)} \frac{x^{-\gamma-iv}}{2^{\gamma+iv} \Gamma(\gamma-1/2+iv)} dv.$$

The following theorem holds

Theorem 4.1. Suppose that for some $\gamma > 0$,

$$\left. \frac{d}{du} (\Gamma(2\gamma-3/2+iu) \mathcal{M}[p_T](2\gamma-1+iu)) \right|_{u=0} \neq 0,$$

and

$$\int_{-\infty}^{\infty} |\mathcal{M}[p_T](2\gamma-1+iu)| du < \infty,$$

then

$$\rho_n^{-1} (p_{n,\gamma}(x) - \mathbb{E}[p_{n,\gamma}(x)]) \xrightarrow{\mathcal{D}} \mathcal{N}(0, \sigma^2)$$

for some $\sigma^2 > 0$, where $\rho_n = n^{-1/2} h_n^{2(\gamma-1)} \log^{-2}(1/h_n) \exp[\pi/h_n]$ and $h_n \asymp c \log^{-1}(n)$ for some $c > 0$.

5 Numerical examples

Barndorff-Nielsen et al. [1] consider a class of variance-mean mixtures of normal distributions which they call generalized hyperbolic distributions. The univariate and symmetric members of this family appear as normal scale mixtures whose mixing distribution is the generalized inverse Gaussian distribution with density

$$(13) \quad p_T(v) = \frac{(\varkappa/\delta)^\lambda}{2K_\lambda(\delta\varkappa)} v^{\lambda-1} \exp\left(-\frac{1}{2} \left(\varkappa^2 v + \frac{\delta^2}{v}\right)\right), \quad v > 0,$$

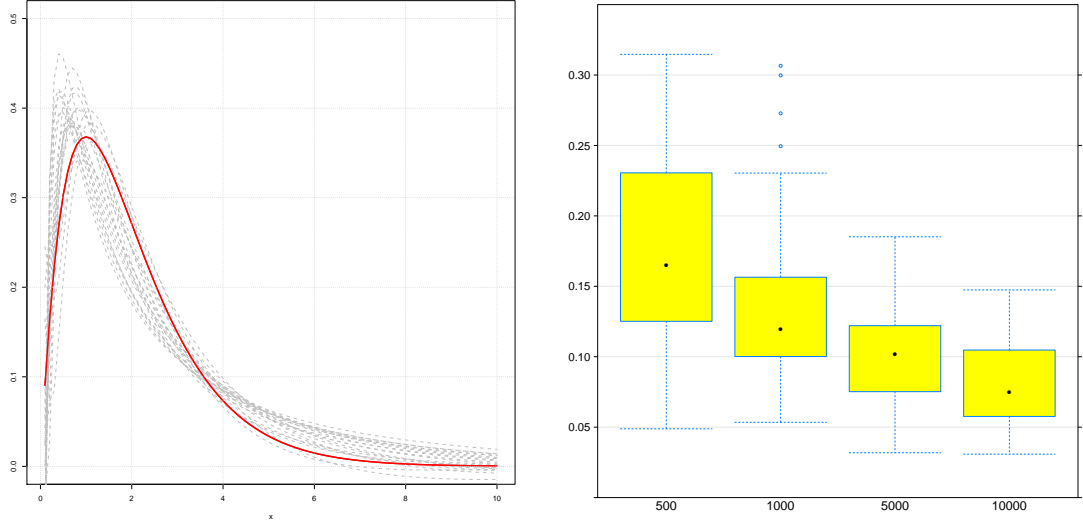


Figure 1: Left: the Gamma density (red) and its 50 estimates (grey) for the sample size $n = 1000$. Right: the box plots of the loss $\sup_{x \in [0, 10]} \{|p_{n, \gamma}(x) - p_T(x)|\}$ for different sample sizes.

for some \varkappa , $\delta \geq 0$ and $\lambda > 0$, where K is a modified Bessel function. The resulting normal scale mixture has probability density function

$$p_X(x) = \frac{\varkappa^{1/2}}{(2\pi)^{1/2} \delta \lambda} K_\lambda(\delta \varkappa) (\delta^2 + x^2)^{\frac{1}{2}(\lambda - \frac{1}{2})} K_{\lambda - \frac{1}{2}}(\varkappa (\delta^2 + x^2)^{1/2}).$$

Let us start with a simple example, Gamma density $p_T(x) = x \exp(-x)$, $x \geq 0$, which is a special case of (13) for $\delta = 0$, $\lambda = 2$ and $\varkappa = \sqrt{2}$. We simulate a sample of size n from the distribution of X , and construct the estimate (6) with the bandwidth h_n given (up to a constant not depending on n) by (8) and $\gamma = 0.9$. In Figure 1 (left), one can see 50 estimated densities based on 50 independent samples from W_T of size $n = 1000$, together with p_T in red. Next we estimate the distribution of the loss $\sup_{x \in [0, 10]} \{|p_{n, \gamma}(x) - p_T(x)|\}$ based on 100 independent repetitions of the estimation procedure. The corresponding box plots for different n are shown in Figure 1 (right).

6 Proofs

6.1 Proof of Theorem 3.1

First let us estimate the bias of $p_{n, \gamma}$. We have

$$\begin{aligned} \mathbb{E}[p_{n, 1}(x)] &= \frac{1}{\sqrt{\pi}} \int_{-\infty}^{\infty} x^{-1-iv} K(vh_n) \frac{\mathcal{M}[p_{|X|}](2(1+iv) - 1)}{2^{1+iv} \Gamma(1/2 + iv)} dv \\ &= \frac{1}{2\pi} \int_{-1/h_n}^{1/h_n} x^{-1-iv} \mathcal{M}[p_T](1+iv) dv. \end{aligned}$$

Hence

$$p_T(x) - \mathbb{E}[p_{n, 1}(x)] = \frac{1}{2\pi} \int_{\{|v| \geq 1/h_n\}} \mathcal{M}[p_T](1+iv) x^{-1-iv} dv$$

and we then have the estimate,

$$\begin{aligned}
\sup_{x \geq 0} \{x |\mathbb{E}[p_{n,1}(x)] - p_T(x)|\} &\leq \frac{1}{2\pi} \int_{\{|v| \geq 1/h_n\}} |\mathcal{M}[p_T](1+iv)| dv \\
&\leq \frac{e^{-\beta/h_n}}{2\pi} \int_{\{|v| \geq 1/h_n\}} e^{-\beta|v|} |\mathcal{M}[p_T](1+iv)| e^{\beta|v|} dv \\
(14) \qquad \qquad \qquad &\leq L \frac{e^{-\beta/h_n}}{2\pi}.
\end{aligned}$$

As to the variance, by the simple inequality $\text{Var}(\int f_t dt) \leq \left(\int \sqrt{\text{Var}[f_t]} dt\right)^2$, which holds for any random function f_t with $\int \mathbb{E}[f_t^2] dt < \infty$, we get

$$\begin{aligned}
\text{Var}[x p_{n,1}(x)] &= \text{Var} \left[\frac{1}{\sqrt{\pi}} \int_{-\infty}^{\infty} x^{-iv} K(vh_n) \frac{\mathcal{M}_n[p_{|X|}](2(1+iv)-1)}{2^{1+iv}\Gamma(1/2+iv)} dv \right] \\
&\leq \frac{1}{\pi 2^2} \left[\int_{-1/h_n}^{1/h_n} \frac{\sqrt{\text{Var}(\mathcal{M}_n[p_{|X|}](2(1+iv)-1))}}{|\Gamma(1/2+iv)|} dv \right]^2 \\
&\leq \frac{1}{2n\pi} \left[\int_{-1/h_n}^{1/h_n} \frac{\sqrt{\text{Var}(|X|^{2iv})}}{|\Gamma(1/2+iv)|} dv \right]^2 \\
(15) \qquad \qquad \qquad &\leq \frac{1}{2n\pi} \left[\int_{-1/h_n}^{1/h_n} \frac{1}{|\Gamma(1/2+iv)|} dv \right]^2.
\end{aligned}$$

We obtain from (15) due to Corollary 6.6 (see Appendix),

$$\text{Var}[x p_{n,\gamma}(x)] \leq \frac{C_L}{n} e^{\pi/h_n}$$

and so (7) follows. Finally, by plugging (8) into (7) we get (9) and the proof is finished.

6.2 Proof of Theorem 3.3

The proof is analog to the one of Theorem 3.1, the only difference is the bias estimate (14) that now becomes

$$\sup_{x \geq 0} \{x |\mathbb{E}[p_{n,1}(x)] - p_T(x)|\} \leq \frac{L}{2\pi} h_n^\beta.$$

With the choice (11) we obtain the logarithmic rate (12).

6.3 Proof of Theorem 3.4

Our construction relies on the following basic result (see [4] for the proof).

Theorem 6.1. *Suppose that for some $\varepsilon > 0$ and all $n \in \mathbb{N}$ there are two densities $p_{0,n}, p_{1,n} \in \mathcal{G}$ for some class \mathcal{G} such that*

$$d(p_{0,n}, p_{1,n}) > 2\varepsilon v_n,$$

where d is some metric on \mathcal{G} . If the observations in model n follow the product law $\mathbb{P}_{p,n} = \mathbb{P}_p^{\otimes n}$ under the density $p \in \mathcal{G}$ and

$$\chi^2(p_{1,n} | p_{0,n}) := \int \frac{(p_{1,n}(x) - p_{0,n}(x))^2}{p_{0,n}(x)} dx \leq n^{-1} \log(1 + (2 - 4\delta)^2)$$

holds for some $\delta \in (0, 1/2)$, then the following lower bound holds for all density estimators \hat{p}_n based on observations from model n :

$$\inf_{\hat{p}_n} \sup_{p \in \mathcal{G}} \mathbb{P}_p^{\otimes n} (d(\hat{p}_n, p) \geq \varepsilon v_n) \geq \delta.$$

If the above holds for fixed $\varepsilon, \delta > 0$ and all $n \in \mathbb{N}$, then the optimal rate of convergence in a minimax sense over \mathcal{G} is not faster than v_n .

We only present the proof of a lower bound for the class $\mathcal{C}(\beta, L)$. The proof for the class $\mathcal{C}(\beta, L)$ is similar. Let us start with the construction of the densities $p_{0,n}$ and $p_{1,n}$. Define for any $\nu > 1$ and $M > 0$ two auxiliary functions

$$q(x) = \frac{\nu \sin(\pi/\nu)}{\pi} \frac{1}{1+x^\nu}, \quad x \geq 0$$

and

$$\rho_M(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{\log^2(x)}{2}} \frac{\sin(M \log(x))}{x}, \quad x \geq 0.$$

The properties of the functions q and ρ_M are collected in the following lemma.

Lemma 6.2. *The function q is a probability density on \mathbb{R}_+ with the Mellin transform*

$$\mathcal{M}[q](z) = \frac{\sin(\pi/\nu)}{\sin(\pi z/\nu)}, \quad \operatorname{Re}[z] > 0.$$

The Mellin transform of the function ρ_M is given by

$$(16) \quad \mathcal{M}[\rho_M](u+iv) = \frac{1}{2} \left[e^{(u-1+i(v+M))^2/2} - e^{(u-1+i(v-M))^2/2} \right].$$

Hence

$$\int_0^\infty \rho_M(x) dx = \mathcal{M}[\rho_M](1) = 0.$$

Set now for any $M > 0$

$$q_{0,M}(x) := q(x), \quad q_{1,M}(x) := q(x) + (q \vee \rho_M)(x),$$

where $f \vee g$ stands for the multiplicative convolution of two functions f and g on \mathbb{R}_+ defined as

$$(f \vee g)(x) := \int_0^\infty \frac{f(t)g(x/t)}{t} dt, \quad x \geq 0.$$

The following lemma describes some properties of $q_{0,M}$ and $q_{1,M}$.

Lemma 6.3. *For any $M > 0$ the function $q_{1,M}$ is a probability density satisfying*

$$\|q_{0,M} - q_{1,M}\|_\infty = \sup_{x \in \mathbb{R}_+} |q_{0,M}(x) - q_{1,M}(x)| \gtrsim \exp(-M\pi/\nu), \quad M \rightarrow \infty.$$

Moreover, $q_{0,M}$ and $q_{1,M}$ are in $\mathcal{C}(\beta, L)$ for all $0 < \beta < \pi/\nu$ and L depending on γ .

Proof. First note that

$$\int_0^\infty q_{1,M}(x) dx = 1 + \int_0^\infty (q \vee \rho_M)(x) dx = 1 + \mathcal{M}[q](1)\mathcal{M}[\rho_M](1) = 1.$$

Furthermore, due to the Parseval identity

$$\begin{aligned}
(q \vee \rho_M)(y) &= \int_0^\infty \frac{1}{\sqrt{2\pi}} e^{-\frac{\log^2(x)}{2}} \frac{\sin(M \log(x))}{x^2} \frac{1}{1 + (y/x)^\nu} dx \\
&= \int_{-\infty}^\infty \frac{1}{\sqrt{2\pi}} e^{-\frac{v^2}{2}} \sin(Mv) \frac{e^{-v}}{1 + e^{-\nu(v-y_i)}} dv \\
&= e^{-\log(y)} \int_{-\infty}^\infty \frac{1}{\sqrt{2\pi}} e^{-\frac{v^2}{2}} \sin(Mv) \frac{e^{(y_i-v)}}{1 + e^{\nu(\log(y)-v)}} dv \\
&= \frac{e^{-\log(y)}}{2\pi} \int_{-\infty}^\infty \frac{1}{\sqrt{2\pi}} e^{-\frac{v^2}{2}} \sin(Mv) \frac{e^{(\log(y)-v)}}{1 + e^{\nu(\log(y)-v)}} dv \\
&= \frac{e^{-\log(y)}}{2\pi} \int_{-\infty}^\infty e^{-iu \log(y)} \left[\frac{H(u+M) - H(u-M)}{2} \right] \mathcal{F}[R](u) du,
\end{aligned}$$

where $R(x) = \frac{e^x}{1+e^{\nu x}}$ and $H(x) = e^{-x^2/2}$. Note that

$$\mathcal{F}[R](u) = \int_{-\infty}^\infty \frac{e^{x+iu x}}{1 + e^{\nu x}} dx = \frac{1}{\nu} \int_{-\infty}^\infty \frac{e^{v/\nu+iu v/\nu}}{1 + e^v} dx = \frac{1}{\nu} \Gamma\left(\frac{1+iu}{\nu}\right) \Gamma\left(1 - \frac{1+iu}{\nu}\right).$$

Hence due to (19)

$$\sup_{y \in \mathbb{R}_+} |q_{0,M}(y) - q_{1,M}(y)| = \sup_{y \in \mathbb{R}_+} |(q \vee \rho_M)(y)| \gtrsim \exp(-M\pi/\nu), \quad M \rightarrow \infty.$$

The second statement of the lemma follows from Lemma 6.2 and the fact that $\mathcal{M}[q \vee \rho_M] = \mathcal{M}[q] \mathcal{M}[\rho_M]$. \square

Let $T_{0,M}$ and $T_{1,M}$ be two random variables with densities $q_{0,M}$ and $q_{1,M}$, respectively. Then the density of the r.v. $|W_{T_{i,M}}|$, $i = 0, 1$, is given by

$$p_{i,M}(x) := \frac{2}{\sqrt{2\pi}} \int_0^\infty \lambda^{-1/2} e^{-\frac{x^2}{2\lambda}} q_{i,M}(\lambda) d\lambda \quad i = 0, 1.$$

For the Mellin transform of $p_{i,M}$ we get

$$\begin{aligned}
\mathcal{M}[p_{i,M}](z) &= \mathbb{E}[|W_1|^{z-1}] \mathbb{E}[T_{i,M}^{(z-1)/2}] \\
&= \mathbb{E}[|W_1|^{z-1}] \mathcal{M}[q_{i,M}]((z+1)/2) \\
(17) \quad &= \frac{2^{z/2}}{\sqrt{2\pi}} \Gamma(z/2) \mathcal{M}[q_{i,M}]((z+1)/2), \quad i = 0, 1.
\end{aligned}$$

Lemma 6.4. *The χ^2 -distance between the densities $p_{0,M}$ and $p_{1,M}$ fulfills*

$$\chi^2(p_{1,M}|p_{0,M}) = \int \frac{(p_{1,M}(x) - p_{0,M}(x))^2}{p_{0,M}(x)} dx \lesssim e^{-M\pi(1+2/\nu)}, \quad M \rightarrow \infty.$$

Fix some $\kappa \in (0, 1/2)$. Due to Lemma 6.4, the inequality

$$n\chi^2(p_{1,M}|p_{0,M}) \leq \kappa$$

holds for M large enough, provided

$$M = \frac{1 + \varepsilon}{\pi(1 + 2/\nu)} (\log(n) + (\nu - 1) \log \log(n))$$

for arbitrary small $\varepsilon > 0$. Hence Lemma 6.3 and Theorem 6.1 imply

$$\inf_{\hat{p}_n} \sup_{p \in \mathcal{C}(\beta, L)} \mathbb{P}_{p,n}(\|\hat{p}_n - p\|_\infty \geq cv_n) \geq \delta.$$

for any $\beta < \pi/\nu < \pi$, some constants $c > 0$, $\delta > 0$ and $v_n = n^{-\beta/(\pi+2\beta)} \log^{-\frac{\pi-\beta}{\pi+2\beta}}(n)$.

6.4 Proof of Proposition 4.1

Thus in order to prove asymptotic normality of $\sqrt{n}(p_{n,\gamma}(x) - \mathbb{E}[p_{n,\gamma}(x)])$, we need to check the Lyapounov condition, i.e. for some $\delta > 0$

$$\frac{\mathbb{E}|Z_{n,1} - \mathbb{E}Z_{n,1}|^{2+\delta}}{n^{\delta/2}[\text{Var}(Z_{n,1})]^{1+\delta/2}} \rightarrow 0, \quad n \rightarrow \infty.$$

For this we need a lower bound for $\text{Var}(Z_{n,1})$. Since $\mathbb{E}Z_{n,1} \rightarrow (2x\sqrt{\pi}) \cdot p_T(x)$, it suffices to show that $\text{Var}(Z_{n,1}) \gtrsim h_n^a \exp(2bh_n^{-1})$ and $\mathbb{E}|Z_{n,1}|^{2+\delta} \lesssim h_n^{-c} \exp((2+\delta)bh_n^{-1})$ for some positive constants a, b and c . Then Lyapounov's condition holds for any $\delta > 0$ by choosing $h_n \sim c(\log^{-1}(n))$. We have

$$\begin{aligned} \text{Var}(Z_{n,1}) &= \frac{1}{\pi} \int_{-1/h_n}^{1/h_n} \int_{-1/h_n}^{1/h_n} \frac{x^{-2\gamma-i(v-u)}}{2^{2\gamma+i(v-u)}} \frac{\text{Cov}(|X_1|^{2(\gamma+iv-1)}, |X_1|^{2(\gamma+iu-1)})}{2^{2\gamma+i(v-u)}\Gamma(\gamma-1/2+iv)\Gamma(\gamma-1/2-iu)} du dv \\ &= \frac{1}{\pi} \int_{-1/h_n}^{1/h_n} \int_{-1/h_n}^{1/h_n} \frac{1}{(2x)^{2(\gamma-1)+i(v-u)}} \frac{\mathcal{M}[p_{|X_1|}](4\gamma-3+2i(v-u))}{\Gamma(\gamma-1/2+iv)\Gamma(\gamma-1/2-iu)} dv du \\ &\quad - \frac{1}{\pi} \left| \int_{-1/h_n}^{1/h_n} \frac{1}{(2x)^{(\gamma+iv-1)}} \frac{\mathcal{M}[p_{|X_1|}](2\gamma-1+2iv)}{\Gamma(\gamma-1/2+iv)} dv \right|^2 = R_1 - R_2. \end{aligned}$$

Note that

$$\begin{aligned} R_2 &\leq \frac{1}{(2x)^{2(\gamma-1)}} \left(\int_{-1/h_n}^{1/h_n} \left| \frac{\mathcal{M}[p_{|X_1|}](2\gamma-1+2iv)}{\Gamma(\gamma-1/2+iv)} \right| dv \right)^2 \\ &= \frac{1}{(2x)^{2(\gamma-1)}} \left| \int_{-1/h_n}^{1/h_n} |\mathcal{M}[p_T](\gamma+iv)| dv \right|^2 < C < \infty \end{aligned}$$

and furthermore

$$\begin{aligned} R_1 &= \frac{1}{x^{2(\gamma-1)}\pi} \int_{-1/h_n}^{1/h_n} \int_{-1/h_n}^{1/h_n} \frac{1}{x^{i(v-u)}} \frac{\Gamma(2\gamma-3/2+i(v-u))\mathcal{M}[p_T](2\gamma-1+i(v-u))}{\Gamma(\gamma-1/2+iv)\Gamma(\gamma-1/2-iu)} dv du \\ &= \frac{1}{x^{2(\gamma-1)}\pi} \times I_n. \end{aligned}$$

Without loss of generality we may take $x = 1$ (for $x \neq 1$ the proof is similar). Observe that

$$\begin{aligned} |\Gamma(\gamma-1/2+iv)| &\geq C_1 1_{|v|\leq 2} + C_2 1_{|v|>2} |v|^{\gamma-1} e^{-\pi|v|/2}, \\ |\Gamma(\gamma-1/2-iu)| &\geq C_1 1_{|u|\leq 2} + C_2 1_{|u|>2} |u|^{\gamma-1} e^{-\pi|u|/2} \end{aligned}$$

for some constants $C_1 > 0$, $C_2 > 0$, and that

$$|\Gamma(2\gamma-3/2+i(v-u))| \leq D_1 1_{|u-v|\leq 2} + 1_{|u-v|>2} D_2 |u-v|^{2(\gamma-1)} e^{-\pi|u-v|/2}$$

for some $D_1 > 0$, $D_2 > 0$. Using these estimate, one can easily derive that the integral

$$I_{1,n,\rho} := \int_{-1/h_n}^{1/h_n} \int_{-1/h_n}^{1/h_n} 1_{|v-u|\geq \rho} \frac{\Gamma(2\gamma-3/2+i(v-u))\mathcal{M}[p_T](2\gamma-1+i(v-u))}{\Gamma(\gamma-1/2+iv)\Gamma(\gamma-1/2-iu)} dv du$$

can be bounded from above as

$$|I_{1,n,\rho}| \lesssim h_n^{-3|1-\gamma|} e^{\pi(\frac{1}{2h_n} - (\rho/2)\vee 1)} + h_n^{-2|1-\gamma|-1} e^{\pi(\frac{1}{h_n} - \rho/2)} + h_n^{-4|1-\gamma|-1} e^{\pi(\frac{1}{h_n} - \rho\vee 2)}$$

for $n \rightarrow \infty$. Similarly

$$\int_{-1/h_n}^{1/h_n} \int_{-1/h_n}^{1/h_n} \mathbf{1}_{|u| \leq \frac{1}{h_n} - \rho} \mathbf{1}_{|v-u| \leq \rho} \frac{\Gamma(2\gamma - 3/2 + i(v-u)) \mathcal{M}[p_T](2\gamma - 1 + i(v-u))}{\Gamma(\gamma - 1/2 + iv) \Gamma(\gamma - 1/2 - iu)} dv du = O\left(h_n^{-l} e^{\pi(\frac{1}{h_n} - \rho)}\right)$$

and

$$\int_{-1/h_n}^{1/h_n} \int_{-1/h_n}^{1/h_n} \mathbf{1}_{|v| \leq \frac{1}{h_n} - \rho} \mathbf{1}_{|v-u| \leq \rho} \frac{\Gamma(2\gamma - 3/2 + i(v-u)) \mathcal{M}[p_T](2\gamma - 1 + i(v-u))}{\Gamma(\gamma - 1/2 + iv) \Gamma(\gamma - 1/2 - iu)} dv du = O\left(h_n^{-l} e^{\pi(\frac{1}{h_n} - \rho)}\right)$$

for some $l > 0$. Hence

$$\begin{aligned} I_{2,n,\rho} &:= \int_{-1/h_n}^{1/h_n} \int_{-1/h_n}^{1/h_n} \mathbf{1}_{|v-u| \leq \rho} \frac{\Gamma(2\gamma - 3/2 + i(v-u)) \mathcal{M}[p_T](2\gamma - 1 + i(v-u))}{\Gamma(\gamma - 1/2 + iv) \Gamma(\gamma - 1/2 - iu)} dv du \\ &= \int_{-1/h_n}^{1/h_n} \int_{-1/h_n}^{1/h_n} \mathbf{1}_{|u| \geq \frac{1}{h_n} - \rho} \mathbf{1}_{|v| \geq \frac{1}{h_n} - \rho} \mathbf{1}_{|v-u| \leq \rho} \\ &\quad \times \frac{\Gamma(2\gamma - 3/2 + i(v-u)) \mathcal{M}[p_T](2\gamma - 1 + i(v-u))}{\Gamma(\gamma - 1/2 + iv) \Gamma(\gamma - 1/2 - iu)} dv du + O\left(h_n^{-l} e^{\pi(\frac{1}{h_n} - \rho)}\right) \\ &= I_{3,n,\rho} + O\left(h_n^{-l} e^{\pi(\frac{1}{h_n} - \rho)}\right). \end{aligned}$$

Now let us study the asymptotic behaviour of the integral $I_{3,n,\rho}$. To this end, we will use the Stirling formulas

$$\begin{aligned} \Gamma(\gamma - 1/2 + iv) &= (\gamma - 1/2 + iv)^{\gamma - 1 + iv} e^{-\gamma + 1/2 - iv} \sqrt{2\pi} (1 + O(|v|^{-1})), \\ \Gamma(\gamma - 1/2 - iu) &= (\gamma - 1/2 - iu)^{\gamma - 1 - iu} e^{-\gamma + 1/2 + iu} \sqrt{2\pi} (1 + O(|u|^{-1})). \end{aligned}$$

First consider the case $u, v \rightarrow +\infty$, where

$$\begin{aligned} \Gamma(\gamma - 1/2 + iv) \Gamma(\gamma - 1/2 - iu) &= 2\pi \exp[iv \log v - iu \log u - i(v-u)] \\ &\quad \times \exp\left[-\frac{\pi}{2}(u+v) + (\gamma-1)(\log v + \log u)\right] (1 + O(1/u) + O(1/v)). \end{aligned}$$

Let $\rho_n = h_n^{-\alpha}$ for $0 < \alpha < 1/2$. Then on the set

$$\left\{|u| \geq \frac{1}{h_n} - \rho_n\right\} \cap \left\{|v| \geq \frac{1}{h_n} - \rho_n\right\} \cap \{|v-u| \leq \rho_n\} \cap \{v \geq 0, u \geq 0\}$$

we define $u = 1/h_n - r$, $v = 1/h_n - s$ with $0 < r, s < \rho_n$, $|r-s| < \rho_n$ to get

$$\begin{aligned} \Gamma(\gamma - 1/2 + iv) \Gamma(\gamma - 1/2 - iu) &= 2e^{i(1/h_n - s) \log(1/h_n - s) - i(1/h_n - r) \log(1/h_n - r) - i(r-s)} \\ &\quad \times h_n^{-2(\gamma-1)} \exp[-\pi/h_n] \exp[(r+s)\pi] \\ &\quad \times (1 + O(h_n))(1 + O(\rho_n h_n)). \end{aligned}$$

Using the asymptotic decomposition

$$(1/h_n - s) \log(1/h_n - s) - (1/h_n - r) \log(1/h_n - r) - (r-s) = (r-s) \log(1/h_n) + O(\rho_n^2 h_n),$$

we derive

$$\begin{aligned}\Gamma(\gamma - 1/2 + iv)\Gamma(\gamma - 1/2 - iu) &= 2\pi h_n^{-2(\gamma-1)} \exp[-\pi/h_n] \exp[(r+s)\pi] \\ &\quad \times \exp[i(r-s)\log(1/h_n)] (1 + O(\rho_n^2 h_n)).\end{aligned}$$

Analogously, on the set

$$\left\{|u| \geq \frac{1}{h_n} - \rho_n\right\} \cap \left\{|v| \geq \frac{1}{h_n} - \rho_n\right\} \cap \{|v-u| \leq \rho_n\} \cap \{v \leq 0, u \leq 0\}$$

we define $u = -1/h_n + r$, $v = -1/h_n + s$, with $0 < r, s < \rho_n$, $|r-s| < \rho_n$, to get

$$\begin{aligned}\Gamma(\gamma - 1/2 + iv)\Gamma(\gamma - 1/2 - iu) &= 2\pi h_n^{-2(\gamma-1)} \exp[-\pi/h_n] \exp[(r+s)\pi] \\ &\quad \times \exp[-i(r-s)\log(1/h_n)] (1 + O(\rho_n^2 h_n)).\end{aligned}$$

Hence the integral I_{3,n,ρ_n} can be decomposed as follows

$$I_{3,n,\rho_n} =: \frac{h_n^{2(\gamma-1)}}{2\pi} \exp[\pi/h_n] \left\{ \operatorname{Re}[I_{4,n,\rho_n}] + O(\rho_n^2 h_n) \right\},$$

where

$$\begin{aligned}I_{4,n,\rho_n} &= \int \int \mathbf{1}_{0 \leq r \leq \rho_n} \mathbf{1}_{0 \leq s \leq \rho_n} \mathbf{1}_{|r-s| \leq \rho_n} \exp[-(r+s)\pi] \Gamma(2\gamma - 3/2 + i(r-s)) \\ &\quad \times \mathcal{M}[p_T](2\gamma - 1 + i(r-s)) \exp[i(s-r)\log(1/h_n)] dr ds \\ &= \int_0^{\rho_n} e^{-2v\pi} R_n(v) dv\end{aligned}$$

with

$$R_n(v) = \int \mathbf{1}_{0 \leq u \leq \rho_n - v} e^{-u\pi} \Gamma(2\gamma - 3/2 + iu) \mathcal{M}[p_T](2\gamma - 1 + iu) e^{iu \log(1/h_n)} du.$$

Using the well known results on the asymptotic of Fourier transform, it is easy to show that

$$\begin{aligned}R_n(v) &= e^{i\pi/2} \Gamma(2\gamma - 3/2) \mathcal{M}[p_T](2\gamma - 1) \log^{-1}(1/h_n) \\ &\quad + e^{i\pi} \left[\frac{d}{du} (\Gamma(2\gamma - 3/2 + iu) \mathcal{M}[p_T](2\gamma - 1 + iu)) \Big|_{u=0} \right] \log^{-2}(1/h_n) + O(\log^{-3}(1/h_n))\end{aligned}$$

uniformly in v . As a result

$$\operatorname{Re}[I_{4,n,\rho_n}] = \left[\frac{d}{du} (\Gamma(2\gamma - 3/2 + iu) \mathcal{M}[p_T](2\gamma - 1 + iu)) \Big|_{u=0} \right] \log^{-2}(1/h_n) + O(\log^{-3}(1/h_n)).$$

Combining all above estimates, we finally get

$$\begin{aligned}(18) \quad \operatorname{Var}(Z_{n,1}) &= \frac{h_n^{2(\gamma-1)}}{\pi^2} \log^{-2}(1/h_n) \exp[\pi/h_n] \\ &\quad \times \left\{ \left[\frac{d}{du} (\Gamma(2\gamma - 3/2 + iu) \mathcal{M}[p_T](2\gamma - 1 + iu)) \Big|_{u=0} \right] \right. \\ &\quad \left. + O(\log^{-1}(1/h_n)) + O(\rho_n^2 h_n \log^2(1/h_n)) + O\left(e^{-\pi\rho_n/2} \log^2(1/h_n)\right) \right\}.\end{aligned}$$

Using the decomposition (18), the Lyapounov condition for some $\delta > 0$

$$\frac{\mathbb{E}|Z_{n,1} - \mathbb{E}Z_{n,1}|^{2+\delta}}{n^{\delta/2} [\operatorname{Var}(Z_{n,1})]^{1+\delta/2}} \rightarrow 0, \quad n \rightarrow \infty$$

is easy to verify, since $\mathbb{E}Z_{n,1} \rightarrow p_T(x)$.

Lemma 6.5. For any $\alpha \geq -2$, there exist positive constants C_1 and $C_2(\alpha)$ such that uniformly for $|\beta| \geq 2$,

$$(19) \quad C|\beta|^{\alpha-1/2}e^{-|\beta|\pi/2} \leq |\Gamma(\alpha + i\beta)| \leq C_\alpha|\beta|^{\alpha-1/2}e^{-|\beta|\pi/2}.$$

Corollary 6.6. For all $0 < \alpha < 1/2$ and all $U > 2$, it holds

$$(20) \quad \int_{-U}^U \frac{d\beta}{|\Gamma(\alpha + i\beta)|} \leq CU^{1/2-\alpha}e^{U\pi/2}$$

for a constant $C > 0$. For $\alpha > 1/2$, we have

$$(21) \quad \int_{-U}^U \frac{d\beta}{|\Gamma(\alpha + i\beta)|} \leq C_1(\alpha) + C_2e^{U\pi/2}$$

where C_2 does not depend on α .

References

- [1] O Barndorff-Nielsen, John Kent, and Michael Sørensen. Normal variance-mean mixtures and z distributions. *International Statistical Review*, 50(2):145–159, 1982.
- [2] Alexander Meister. *Deconvolution problems in nonparametric statistics*, volume 193. Springer, 2009.
- [3] Jan Oblój et al. The Skorokhod embedding problem and its offspring. *Probability Surveys*, 1:321–392, 2004.
- [4] Alexandre B Tsybakov. *Introduction to nonparametric estimation*, volume 11. Springer, 2009.
- [5] Bert Van Es, Shota Gugushvili, Peter Spreij, et al. A kernel type nonparametric density estimator for decompounding. *Bernoulli*, 13(3):672–694, 2007.
- [6] Bert Van Es and Peter Spreij. Estimation of a multivariate stochastic volatility density by kernel deconvolution. *Journal of Multivariate Analysis*, 102(3):683–697, 2011.
- [7] Bert Van Es, Peter Spreij, and Harry Van Zanten. Nonparametric volatility density estimation. *Bernoulli*, 9(3):451–465, 2003.