Minimum Distance Estimation in AR(1)-processes

Inauguraldissertation
der Philosophisch-naturwissenschaftlichen Fakultät
der Universität Bern

vorgelegt von

Michel Piot

von Thierrens

Leiter der Arbeit: Prof. Dr. J. Hüsler
Institut für mathematische Statistik und Versicherungslehre

Minimum Distance Estimation in AR(1)-processes

Inauguraldissertation
der Philosophisch-naturwissenschaftlichen Fakultät
der Universität Bern

vorgelegt von

Michel Piot

von Thierrens

Leiter der Arbeit: Prof. Dr. J. Hüsler
Institut für mathematische Statistik und Versicherungslehre

Von der Philosophisch-naturwissenschaftlichen Fakultät angenommen.

Bern, den 22.10.1998

Der Dekan:
Prof. Dr. A. Pfiffner

Preface

After having studied mathematics and geography I was interested in combining these two sciences. So the common denominator was the Institute of Mathematical Statistics where I got the opportunity to write a thesis.

From a climatological viewpoint I was interested in analysing time series of temperature and pressure, which can for example be modeled as autoregressive processes.

On the statistical side a lot of research was done on autoregressive processes in the case of normal distributed errors, but in reality this is a very strong assumption for climatological time series .

It was Prof. J. Hüsler's idea to model climatological time series as first-order autoregressive processes, with contaminated normal distributed errors, which seems to be a more realistic approach.
In order to find an estimate of the autoregressive parameter ρ, the theory of minimum distance estimates turned out to be suitable.
Koul (1992) showed the way of finding an appropriate method to prove the asymptotic normality of the minimum distance estimator $\hat{\rho}$.
Thus the present thesis is a mathematical-statistical tool which enables us to analyse climatological time series in a more realistic way.

Acknowledgements

My special thanks go to Prof. J. Hüsler for his patient support, his numerous ideas to prove important details and his success in inviting Prof. H. Koul.

It was a great stroke of luck for me to meet Prof. Hira L. Koul from Michigan State University several times, since he is the expert in the domain of minimum distance estimates in linear regression and autoregression. The discussions with him led to great progress in my work. I would like to express my gratitude to him.

Furthermore I wish to thank Christoph Leuenberger for his help in the analysis domain, Peter Balsiger for solving informatics questions, Philipp Busato and Enkelejd Hashorva for re-reading this manuscript.

Last but not least I thank all people who helped me in any way during my education, especially my parents who encouraged and supported me all the years.

Contents

1 Introduction **1**

2 Contaminated normal distribution **4**
2.1 Model 4
2.2 Moments 5

3 Minimum Distance Estimation **13**
3.1 Description of the chapter 13
3.2 Definitions and Lemmata 15
3.3 Propositions 25
3.4 Theorems 33

4 MDE versus least square estimator **37**
4.1 Least-square estimator 37
4.2 Comparison 39

5 Application **42**
5.1 General Case 42
5.2 Logistic distribution 47

A Fourier Analysis **48**

B Probability **50**

References **54**

1 Introduction

Beran (1984) describes minimum distance estimates in an easy way: “Minimum distance estimates are estimates of the parameters chosen to minimize the distance between the data and the fitted model.”

Wolfowitz (1957) was first to consider minimum distance procedures, and he established consistency of minimum distance estimates under general conditions. Asymptotic distributions for the Cramér-von-Mises distance were derived by Parr and Schucany (1980) and Millar (1981).
For more detailed references prior to 1981 see the bibliography in Parr (1981).

Koul and DeWet (1983) provided suitable analogues of the Cramér-von-Mises type minimum distance estimators in a linear regression model. These estimators are obtained by minimizing an integral of squared difference between weighted empiricals of the residuals and their expectations with respect to a large class of integrating measures.

Some years later Koul (1986) gave a class of minimum L^2-distance estimators of the autoregression parameter in the first-order autoregression model for the case when the error distribution is unknown but symmetric. For this purpose he used random weighted empirical processes.

Finally, Koul (1992) extended the theory to $AR(p)$-models under the assumption that the error distribution is known or is unknown but symmetric.

In the present work Φ is used as model distribution. Using random weighted empirical processes in the same way as Koul (1986), it can be proved that the minimum distance estimator of the autoregression parameter of the first-order autoregression model is asymptotically normal distributed even if the error distribution is unknown but contaminated normal.

The outline of the thesis:

Consider the time series $\{X_{in}\}$ satisfying an $AR(1)$-process, e.g.

$$X_{in} = \rho X_{i-1,n} + \varepsilon_{in}, \qquad i \leq n.$$

Assume that the true distribution function H_n of ε_{in} is contaminted normal, e.g.

$$H_n = (1 - a_n)\Phi + a_n H_0,$$

where H_0 is an unknown distribution function with first moment μ_0 and $a_n \to 0$ as $n \to \infty$.

Define a class of minimum distance estimators taking Φ as the model distribution function in place of H_n:

$$M_n(t) := \int \left[n^{-1/2} \sum_{i=1}^{n} X_{i-1,n} \left\{ I(\varepsilon_{in} \leq x + (t-\rho)X_{i-1,n}) - \Phi(x) \right\} \right]^2 dG(x).$$

The question is whether this class of minimum distance estimators has nice properties; especially the asymptotic distribution is of interest.

The Main Theorem (Theorem 3.4.1) of this work states that the asymptotic distribution of $\sqrt{n}(\tilde{\rho} - \rho)$, where $\tilde{\rho} := \text{argmin}_t \ M_n(t)$, is normal, e.g.

$$\sqrt{n}(\tilde{\rho} - \rho) \xrightarrow{d} N\left(-\frac{k\mu_0(1+\rho)\int (H_0(x) - \Phi(x))\varphi(x)\,dG(x)}{\int \varphi^2(x)\,dG(x)}, \frac{\sigma_\psi^2(1-\rho^2)}{\left(\int \varphi^2(x)\,dG(x)\right)^2}\right).$$

In order to prove this theorem, the work is divided in two chapters:

Chapter 2 develops the results that are used in connection with the contaminated normal distributed errors ε_{in}. Since the minimum distance estimator depends on n, the time series $\{X_{in}\}$ has to be decomposed in two parts, e.g.

$$X_{in} = \sum_{j=0}^{\infty} \rho^j (1 - U_{i-j,n})\epsilon_{i-j} + \sum_{j=0}^{\infty} U_{i-j,n}\delta_{i-j}, \qquad i \leq n,$$

where ϵ_{i-j} has distribution function Φ, and δ_{i-j} obeys H_0.

Chapter 3 starts with a description of the steps of the proof of the Main Theorem (Theorem 3.4.1) and then states all definitions and lemmata used. To find the asymptotic distribution of $\sqrt{n}(\tilde{\rho} - \rho)$ define an auxiliary function $\hat{M}_n(t)$ with $\hat{\rho} := \text{argmin}_t \ \hat{M}_n(t)$.

It can be shown (Proposition 3.3.3) that

$$\sup |M_n(t) - \hat{M}_n(t)| = o_p(1),$$

where the supremum is taken over a compact set of $\mathbb{R}$.
Since $\sqrt{n}(\hat{\rho} - \tilde{\rho}) = o_p(1)$ (Theorem 3.4.4), and therefore

$$\sqrt{n}(\hat{\rho} - \rho) = \sqrt{n}(\tilde{\rho} - \rho) + \sqrt{n}(\hat{\rho} - \tilde{\rho}) = \sqrt{n}(\tilde{\rho} - \rho) + o_p(1),$$

the asymptotic distribution of $\sqrt{n}(\tilde{\rho} - \rho)$ and $\sqrt{n}(\hat{\rho} - \rho)$ must be the same.

Therefore it suffices to find the asymptotic distribution of $\sqrt{n}(\hat{\rho} - \rho)$, which is easier, because the properties of $\hat{M}_n(t)$ are less complicated than those of $M_n(t)$.

To achieve this goal the exact expression of $\sqrt{n}(\hat{\rho} - \rho)$ is derived. Then in the Main Theorem (Theorem 3.4.1) the asymptotic normality of $\sqrt{n}(\hat{\rho} - \rho)$ is proved.

The last two chapters are applications of the Main Theorem.

Chapter 4 compares the asymptotic distribution of this minimum distance estimator with the least square estimator. Proposition 4.2.1 shows the cases where the minimum distance estimator has smaller bias than the least square estimator.

Finally Chapter 5 discusses how to find the minimum distance estimator $\tilde{\rho}$ of $M_n(t)$ numerically, using Corollary 5.1.8 which explicitely states where the local extremes can be attained.

For the special case of the weighting function $g(x) \equiv x$ it is shown that the function $M_n(t)$ has a minimum (Proposition 5.1.9).

2 Contaminated normal distribution

In this chapter it is assumed that a_n is a sequence with $0 \leq a_n \leq 1$, for all n. Note that in this case $O(a_n) + O(a_n^k) = O(a_n)$, for all $k \geq 1$.

2.1 Model

Definition 2.1.1. *Let*

$$H_n = (1 - a_n)\Phi + a_n H_0 \Longleftrightarrow H_n - \Phi = a_n(H_0 - \Phi)$$

be the distribution function of the contaminated normal distribution. H_0 is a distribution function with density h_0, first moment μ_0 and second moment $\sigma_0^2 + \mu_0^2$.

We get

$$\frac{dH_n(x)}{dx} = h_n(x) = (1 - a_n)\varphi(x) + a_n h_0(x).$$

Definition 2.1.2. *For $i \leq n$ let $U_{in} \sim Bin(1, a_n)$, $\epsilon_i \sim N(0,1)$ and $\delta_i \sim H_0$ be independent random variables for all i.*

Lemma 2.1.3. *The random variable*

$$\varepsilon_{in} := (1 - U_{in})\epsilon_i + U_{in}\delta_i, \qquad i \leq n,$$

has distribution function H_n.

Proof. The proof is an easy consequence of the independence of the random variable U_{in}, ϵ_i and δ_i.

$$\begin{aligned} P(\varepsilon_{in} \leq x) &= P((1 - U_{in})\epsilon_i + U_{in}\delta_i \leq x) \\ &= P((1 - U_{in}) = 1)P(\epsilon_i \leq x) + P(U_{in} = 1)P(\delta_i \leq x) \\ &= (1 - a_n)\Phi(x) + a_n H_0(x) = H_n(x). \end{aligned}$$

□

Definition 2.1.4. *Let $X_{in} = \rho X_{i-1,n} + \varepsilon_{in}$, $i \leq n$, be an $AR(1)$-process with ε_{in} i.i.d. contaminated normal.*

Definition 2.1.5. *An $AR(p)$-process is said to be causal if there exists a sequence of constants $\{\psi_j\}$ such that $\sum_{j=0}^{\infty} |\psi_j| < \infty$ and*

$$X_{in} = \sum_{j=0}^{\infty} \psi_j \varepsilon_{i-j,n}, \qquad i \leq n.$$

Remark 2.1.6. A good review on causal processes is given in Brockwell and Davis (1993), p. 77ff.

Lemma 2.1.7. *For every fixed n the $AR(1)$-process $X_{in} = \rho X_{i-1,n} + \varepsilon_{in}$, $i \leq n$, is causal and stationary iff $|\rho| < 1$.*

Proof. Brockwell and Davis (1993), p. 85 and Priestley (1994), p. 121 ff.

Remark 2.1.8. For more details on causal AR-processes and stationarity see Anderson (1971), p. 166ff., p. 372ff. A whole chapter on second order properties, which play an important role in this context, can be found in Loève (1978), p. 121-159.

Definition 2.1.9. *Let X_{in}, $i \leq n$, be an $AR(1)$-process as in Definition 2.1.4. We write*

$$\begin{aligned} X_{in} = \sum_{j=0}^{\infty} \rho^j \varepsilon_{i-j,n} &= \sum_{j=0}^{\infty} \rho^j \left[(1 - U_{i-j,n})\epsilon_{i-j} + U_{i-j,n}\delta_{i-j}\right] \\ &= \sum_{j=0}^{\infty} \rho^j \epsilon_{i-j} + \sum_{j=0}^{\infty} \rho^j U_{i-j,n}(\delta_{i-j} - \epsilon_{i-j}) \\ &=: \tilde{X}_i + \hat{X}_{in}. \end{aligned}$$

2.2 Moments

In this section we are calculating the moments which are used in Chapter 3 and Chapter 4.

Lemma 2.2.1. *Let $X_{in} = \rho X_{i-1,n} + \varepsilon_{in}$, $i \leq n$, $0 < \rho < 1$, be an $AR(1)$-process with ε_{in} i.i.d. contaminated normal. Then, for $i \leq n$,*

1. $\mathrm{E}\varepsilon_{in} = a_n\mu_0 = O(a_n)$,
2. $\mathrm{E}\varepsilon_{in}^2 = 1 + O(a_n)$,
3. $\mathrm{E}\varepsilon_{in}^3 = O(a_n)$,
4. $\mathrm{E}\varepsilon_{in}^4 = 3 + O(a_n)$.

Proof. Use the definition $\varepsilon_{in} := (1 - U_{in})\epsilon_i + U_{in}\delta_i$.

1. $\mathrm{E}\varepsilon_{in} = \mathrm{E}(1 - U_{in})\,\mathrm{E}\epsilon_i + \mathrm{E}U_{in}\,\mathrm{E}\delta_i = a_n\mu_0$.
2. The second moment of ε_{in} is

$$\begin{aligned} \mathrm{E}\varepsilon_{in}^2 &= \mathrm{E}\left((1 - U_{in})\epsilon_i + U_{in}\delta_i\right)^2 = \mathrm{E}(1 - U_{in})^2\,\mathrm{E}\epsilon_i^2 + \mathrm{E}U_{in}^2\,\mathrm{E}\delta_i^2 \\ &= (1 - a_n) + a_n(\mu_0^2 + \sigma_0^2) = 1 + O(a_n). \end{aligned}$$

3. Third and fourth moment can be found similarly.

□

Remark 2.2.2. To calculate the exact expressions of the forth moments we use the equation

$$\sum_{i,j,k,l=0}^{\infty} c^i c^j c^k c^l = \sum c^{4i} + 4\sum_{i\neq j} c^{3i}c^j + 3\sum_{i\neq j} c^{2i}c^{2j} + 6\sum_{\substack{i\neq j\\ i\neq k\\ j\neq k}} c^{2i}c^j c^k + \sum_{\substack{i\neq j, i\neq k\\ i\neq l, j\neq k\\ j\neq l, k\neq l}} c^i c^j c^k c^k. \tag{2.1}$$

Lemma 2.2.3. *The moments of the random variable $\tilde{X}_i$, $i \leq n$, are:*

1. $\mathrm{E}\tilde{X}_i = 0$,
2. $\mathrm{E}\tilde{X}_i^2 = \dfrac{1}{1-\rho^2}$,
3. $\mathrm{E}\tilde{X}_i^3 = 0$,
4. $\mathrm{E}\tilde{X}_i^4 = \dfrac{3}{(1-\rho^2)^2}$.

Proof. Use Definition 2.1.9 to show that

1. $\mathrm{E}\tilde{X}_i = \sum_{j=0}^{\infty} \rho^j\, \mathrm{E}\epsilon_{i-j} = 0$,
2. $\mathrm{E}\tilde{X}_i^2 = \sum_{j=0}^{\infty} \rho^{2j}\, \mathrm{E}\epsilon_{i-j}^2 = \dfrac{1}{1-\rho^2}$.
3. Follows alike as in part 4.
4. We use the representation of $\tilde{X}_i^4$ and Remark 2.2.2 to show that

$$\begin{aligned}\mathrm{E}\tilde{X}_i^4 &= \frac{1}{1-\rho^4}\,\mathrm{E}\epsilon_i^4 + 3\left(\frac{1}{(1-\rho^2)^2} - \frac{1}{1-\rho^4}\right)(\mathrm{E}\epsilon_i^2)^2\\ &= \frac{3}{1-\rho^4} + \frac{3}{(1-\rho^2)^2} - \frac{3}{1-\rho^4} = \frac{3}{(1-\rho^2)^2}.\end{aligned}$$

Lemma 2.2.4. *Let $i \leq n$. Assume that $\mathrm{E}\delta_i^3$ and $\mathrm{E}\delta_i^4$ exist. The moments of $\hat{X}_{in}$ are:*

1. $\mathrm{E}\hat{X}_{in} = \dfrac{a_n\mu_0}{1-\rho} = O(a_n)$,

2. $\mathrm{E}\hat{X}_{in}^2 = O(a_n)$,

3. $\mathrm{E}\hat{X}_{in}^3 = O(a_n)$,

4. $\mathrm{E}\hat{X}_{in}^4 = O(a_n)$.

Proof. Use Definition 2.1.9 to show that

1. $\mathrm{E}\hat{X}_{in} = \sum_{j=0}^{\infty} \rho^j \,\mathrm{E}\left(U_{i-j,n}(\delta_{i-j} - \epsilon_{i-j})\right) = \frac{a_n \mu_0}{1-\rho}.$

2. Since U_{in}, δ_i and ϵ_i are independent random variables and each summand of

$$\mathrm{E}\hat{X}_{in}^k = \mathrm{E}\left(\sum_{j=0}^{\infty} \rho^j U_{i-j,n}(\delta_{i-j} - \epsilon_{i-j})\right)^k, \qquad k = 2, 3, 4,$$

contains the random variable U_{in} with $\mathrm{E}U_{in}^l = a_n$, $l \geq 1$, for all i, every expectation value has a least one factor a_n. The moments of ϵ_i and δ_i exist and since $0 < \rho < 1$ the sum exists, too. These arguments complete the proof. To find the exact expressions of the moments use Remark 2.2.2.

□

Corollary 2.2.5. *The moments of* X_{in}, $i \leq n$, *are*

1. $\mathrm{E}X_{in} = \frac{a_n \mu_0}{1-\rho} = O(a_n)$,

2. $\mathrm{E}X_{in}^2 = \frac{1}{1-\rho^2} + O(a_n)$,

3. $\mathrm{E}X_{in}^3 = O(a_n)$,

4. $\mathrm{E}X_{in}^4 = \frac{3}{(1-\rho^2)^2} + O(a_n)$.

Proof. We use the representation

$$X_{in} = \sum_{j=0}^{\infty} \rho^j \varepsilon_{i-j,n}.$$

Then the proof can easily be done with the help of Remark 2.2.2. □

Lemma 2.2.6. *Let* $i < j \leq n+1$. *Then*

1. $\mathrm{E}X_{i-1,n}X_{j-1,n} = \rho^{j-i}\,\mathrm{E}X_{0n}^2 + \frac{a_n^2 \mu_0^2}{1-\rho}\,\frac{1-\rho^{j-i}}{1-\rho}$, *2.* $\mathrm{E}X_{i-1,n}^2 X_{j-1,n}^2 = \dfrac{1+2\rho^{2(j-i)}}{(1-\rho^2)^2} + O(a_n),$

3. $\mathrm{E}\left(\sum_{i=1}^{n} X_{i-1,n}\right)^2 = O(n) + \dfrac{n^2 a_n^2 \mu_0^2}{(1-\rho)^2} = O(n) + O(n^2 a_n^2).$

Proof. 1. We use the expansion for $X_{j-1,n}$ with $i < j$:

$$\begin{aligned}
&\mathrm{E}X_{j-1,n}X_{i-1,n} \\
&\quad = \mathrm{E}\big[(\rho^{j-i}X_{i-1,n} + \rho^{j-i-1}\varepsilon_{in} + \rho^{j-i-2}\varepsilon_{i+1,n} + \ldots + \rho\varepsilon_{j-2,n} + \varepsilon_{j-1,n})X_{i-1,n}\big] \\
&\quad = \rho^{j-i}\,\mathrm{E}X_{i-1,n}^2 + \rho^{j-i-1}\,\mathrm{E}\varepsilon_{in}\,\mathrm{E}X_{i-1,n} + \ldots \\
&\qquad + \rho\,\mathrm{E}\varepsilon_{j-2,n}\,\mathrm{E}X_{i-1,n} + \mathrm{E}\varepsilon_{j-1,n}\,\mathrm{E}X_{i-1,n} \\
&\quad = \rho^{j-i}\,\mathrm{E}X_{0n}^2 + \mathrm{E}X_{0n}\,\mathrm{E}\varepsilon_{0n}[\rho^{j-i-1} + \rho^{j-i-2} + \ldots + \rho + 1] \\
&\quad = \rho^{j-i}\,\mathrm{E}X_{0n}^2 + \frac{a_n^2\mu_0^2}{1-\rho}\,\frac{1-\rho^{j-i}}{1-\rho}.
\end{aligned}$$

2. We use the same idea as in part 1. For $X_{i-1,n}$ with $i < j$ we get

$$\begin{aligned}
&\mathrm{E}X_{i-1,n}^2 X_{j-1,n}^2 = \mathrm{E}[X_{i-1,n}^2(\rho^{j-i}X_{i-1,n} + \rho^{j-i-1}\varepsilon_{in} + \ldots + \rho\varepsilon_{j-2,n} + \varepsilon_{j-1,n})^2] \\
&\quad = \mathrm{E}\Bigg(X_{i-1,n}^2[\rho^{2(j-i)}X_{i-1,n}^2 + 2\rho^{j-i}X_{i-1,n}\sum_{k=1}^{j-i}\rho^{j-i-k}\varepsilon_{i+k-1,n} \\
&\qquad + \sum_{k=1}^{j-i}\rho^{2(j-i-k)}\varepsilon_{i+k-1,n}^2 + 2\sum_{l=2}^{j-i}\sum_{k=1}^{l-1}\rho^{j-i-k}\rho^{j-i-l}\varepsilon_{i+k-1,n}\varepsilon_{i+l-1,n}]\Bigg) \\
&\quad = \rho^{2(j-i)}\,\mathrm{E}X_{i-1,n}^4 + 2\rho^{j-i}\,\mathrm{E}X_{i-1,n}^3\,\mathrm{E}\varepsilon_{0n}\sum_{k=1}^{j-i}\rho^{j-i-k} \\
&\qquad + \mathrm{E}X_{i-1,n}^2\,\mathrm{E}\varepsilon_{0n}^2\sum_{k=1}^{j-i}\rho^{2(j-i-k)} + 2\,\mathrm{E}X_{i-1,n}^2(\,\mathrm{E}\varepsilon_{0n})^2\sum_{l=2}^{j-i}\sum_{k=1}^{l-1}\rho^{j-i-k}\rho^{j-i-l} \\
&\quad = \frac{3\rho^{2(j-i)}}{(1-\rho^2)^2} + \mathrm{E}X_{0n}^2\,\mathrm{E}\varepsilon_{0n}^2[\rho^{2(j-i-1)} + \ldots + \rho^2 + 1] + O(a_n) \\
&\quad = \frac{2\rho^{2(j-i)}+1}{(1-\rho^2)^2} + O(a_n).
\end{aligned}$$

To show that $O(a_n)$ is uniform we need the inequality

$$\begin{aligned}
\sum_{l=2}^{j-i}\sum_{k=1}^{l-1}\rho^{j-i-k}\rho^{j-i-l} &= \frac{\rho - \rho^{j-i+1} - \rho^{(j-i)} + \rho^{2(j-i)}}{(1-\rho^2)(1-\rho)} \\
&\le \frac{2\rho}{(1-\rho^2)(1-\rho)}
\end{aligned}$$

and consider the expression

$$
\begin{aligned}
2\rho^{j-i}\,\mathrm{E}X_{0n}^3\,\mathrm{E}\varepsilon_{0n}\sum_{k=1}^{j-i}\rho^{j-i-k} &+ 2\,\mathrm{E}X_{0n}^2(\,\mathrm{E}\varepsilon_{0n})^2\sum_{l=2}^{j-i}\sum_{k=1}^{l-1}\rho^{j-i-k}\rho^{j-i-l}\\
&= 2O(a_n^2)\rho^{j-i}\sum_{k=1}^{j-i}\rho^{j-i-k} + 2O(a_n^2)\sum_{l=2}^{j-i}\sum_{k=1}^{l-1}\rho^{j-i-k}\rho^{j-i-l}\\
&\le 2O(a_n^2)\frac{\rho^{j-i}}{1-\rho} + 4O(a_n^2)\frac{\rho}{(1-\rho)(1-\rho^2)}\\
&\le 2O(a_n^2)\frac{1}{1-\rho} + 4O(a_n^2)\frac{\rho}{(1-\rho)(1-\rho^2)} = O(a_n^2).
\end{aligned}
$$

3. Using part 1 and having in mind that

$$\sum_{j=2}^{n}\sum_{i=1}^{j-1}\rho^{j-i} = O(n),$$

we find

$$
\begin{aligned}
\mathrm{E}\left(\sum_{i=1}^{n}X_{i-1,n}\right)^2 &= n\,\mathrm{E}X_{i-1,n}^2 + 2\sum_{j=2}^{n}\sum_{i=1}^{j-1}\mathrm{E}X_{j-1,n}X_{i-1,n}\\
&= n\,\mathrm{E}X_{0n}^2 + 2\sum_{j=2}^{n}\sum_{i=1}^{j-1}\left(\rho^{j-i}\,\mathrm{E}X_{i-1,n}^2 + \frac{a_n^2\mu_0^2}{1-\rho}\,\frac{1-\rho^{j-i}}{1-\rho}\right)\\
&= n\,\mathrm{E}X_{0n}^2 + O(n)\,\mathrm{E}X_{0n}^2 + \frac{n(n-1)a_n^2\mu_0^2}{(1-\rho)^2} + O(na_n^2)\\
&= O(n) + O(n^2a_n^2).
\end{aligned}
$$

□

Remark 2.2.7. We note that $O(a_n)$ is independent on i and j. In this case we will say that $O(a_n)$ is uniform.

To prove Proposition 4.1.4 we need the following lemma:

Lemma 2.2.8. *Let* $R := \frac{\mu_0}{\sqrt{n}}\sum_{i=1}^{n}\hat{X}_{i-1,n}U_{in}$. *Then*

$$1.\quad \mathrm{E}R = \frac{\mu_0^2\sqrt{n}a_n^2}{1-\rho}$$

and

$$2.\quad \mathrm{Var}\ R = O(a_n^2).$$

Proof. 1. This part is obvious.

2. The second moment of R is

$$\mathrm{E}\left(\frac{\mu_0}{\sqrt{n}}\sum_{i=1}^{n}\hat{X}_{i-1,n}U_{in}\right)^2 = \frac{\mu_0^2}{n}\,\mathrm{E}\left(\sum_{i=1}^{n}\hat{X}_{i-1,n}U_{in}\right)^2$$
$$= \frac{\mu_0^2}{n}\,\mathrm{E}\left(\sum_{i=1}^{n}\hat{X}_{i-1,n}^2U_{in} + 2\sum_{i<j}\hat{X}_{i-1,n}\hat{X}_{j-1,n}U_{in}U_{jn}\right).$$

(a) Since $\hat{X}_{jn}$, $\hat{X}_{in}$ and $U_{i+1,n}$ are not independent we first consider the term

$$\mathrm{E}\hat{X}_{in}\hat{X}_{jn}U_{i+1,n}U_{j+1,n}.$$

Assume $j > i$ and use Definition 2.1.9 to write

$$\hat{X}_{jn} = \sum_{k=0}^{j-i-2}\rho^k U_{j-k,n}(\delta_{j-k} - \epsilon_{j-k}) + \rho^{j-i-1}U_{i+1,n}(\delta_{i+1} - \epsilon_{i+1}) + \rho^{j-i}\hat{X}_{in}. \tag{2.2}$$

Using (2.2) we get

$$\begin{aligned}
&\mathrm{E}\hat{X}_{in}\hat{X}_{jn}U_{i+1,n}U_{j+1,n}\\
&= \mathrm{E}U_{j+1,n}\,\mathrm{E}U_{i+1,n}\,\mathrm{E}\hat{X}_{in}\sum_{k=0}^{j-i-2}\rho^k\,\mathrm{E}U_{j-k,n}(\delta_{j-k} - \epsilon_{j-k})\\
&\quad + \rho^{j-i-1}\,\mathrm{E}U_{j+1,n}\,\mathrm{E}U_{i+1,n}(\delta_{i+1} - \epsilon_{i+1})\,\mathrm{E}\hat{X}_{in}\\
&\quad + \rho^{j-i}\,\mathrm{E}U_{j+1,n}\,\mathrm{E}U_{i+1,n}\,\mathrm{E}\hat{X}_{in}^2\\
&= (\mathrm{E}U_{0n})^3\,\mathrm{E}\hat{X}_{0n}(\mathrm{E}\delta_0 - \mathrm{E}\varepsilon_0)\sum_{k=0}^{j-i-2}\rho^k + \rho^{j-i-1}(\mathrm{E}U_{0n})^2\,\mathrm{E}\hat{X}_{0n}(\mathrm{E}\delta_0 - \mathrm{E}\varepsilon_0)\\
&\quad + \rho^{j-i}(\mathrm{E}U_{0n})^2\,\mathrm{E}\hat{X}_{0n}^2\\
&= \frac{a_n^4\mu_0^2}{1-\rho}\,\frac{1-\rho^{j-i-1}}{1-\rho} + \rho^{j-i-1}\frac{a_n^3\mu_0^2}{1-\rho} + \rho^{j-i}O(a_n^3)\\
&= \frac{a_n^4\mu_0^2}{(1-\rho)^2} + \rho^{j-i}O(a_n^3).
\end{aligned} \tag{2.3}$$

Next we calculate the sum

$$2\sum_{i<j}\mathrm{E}\hat{X}_{in}\hat{X}_{jn}U_{i+1,n}U_{j+1,n} = 2\sum_{i<j}\left[\frac{a_n^4\mu_0^2}{(1-\rho)^2} + O(a_n^3)\rho^{j-i}\right]$$
$$= \frac{n(n-1)a_n^4\mu_0^2}{(1-\rho)^2} + O(na_n^3).$$

(b) Now we go back to the original expression and use (2.3):

$$\begin{aligned} \mathrm{E}R^2 &= \frac{\mu_0^2}{n}\,\mathrm{E}\left(\sum_{i=1}^{n} \hat{X}_{i-1,n}^2 U_{in} + 2\sum_{i<j} \hat{X}_{i-1,n}\hat{X}_{j-1,n}U_{in}U_{jn}\right) \\ &= \mu_0^2\,\mathrm{E}\hat{X}_{0n}^2\,\mathrm{E}U_{0n} + \frac{(n-1)a_n^4\mu_0^4}{(1-\rho)^2} + O(a_n^3) \\ &= \frac{na_n^4\mu_0^4}{(1-\rho)^2} + O(a_n^2). \end{aligned}$$

(c) Finally

$$\mathrm{Var}\ R = \mathrm{E}R^2 - (\mathrm{E}R)^2 = O(a_n^2).$$

□

Corollary 2.2.9. *If* $\lim_{n\to\infty} \sqrt{n}a_n^2 =: k < \infty$ *we find*

$$R \xrightarrow{P} \frac{\mu_0^2 k}{1-\rho}.$$

Proof. It is a consequence of Lemma 2.2.8 and Chebychev's inequality.

Lemma 2.2.10. *Note that* $\max_x \varphi(x) = \frac{1}{\sqrt{2\pi}} < 1$ *and therefore*

$$\int |H_0 - \Phi|\varphi\, dG < \int |H_0 - \Phi|\, dG.$$

1. *Assume that* $\int |H_0 - \Phi|\, dG < \infty$ *and* $\int \Phi(1-\Phi)\, dG < \infty$. *Then, uniformly in* $n \geq 1$,

$$\int H_n(1-H_n)\, dG < \infty.$$

2. *Let* $\int h_0^2\, dG < \infty$, $\int \varphi^2\, dG < \infty$. *Then, uniformly in* $n \geq 1$, $\int h_n^2\, dG < \infty$ *and*

$$\int h_n^2\, dG = \int \varphi^2\, dG + O(a_n^2).$$

3. *If* $\int |H_0 - \Phi| h_0\, dG < \infty$ *and* $\int |H_0 - \Phi|\, dG < \infty$ *then*

$$\int (H_0 - \Phi)h_n\, dG = \int (H_0 - \Phi)\varphi\, dG + O(a_n).$$

Proof. 1. Use Definition 2.1.1, the facts $\int |\Phi - H_0|\, dG > \int (\Phi - H_0)^2\, dG$ and $\int \Phi(\Phi - H_0)\, dG < \int |\Phi - H_0|\, dG$ to get

$$\begin{aligned} 0 &< \int H_n(1 - H_n)\, dG \\ &= -a_n^2 \int (\Phi - H_0)^2\, dG + \int \Phi(1 - \Phi)\, dG \\ &\quad + a_n \int (H_0 - \Phi)\, dG + 2a_n \int \Phi(\Phi - H_0)\, dG \\ &\le -a_n^2 \int (\Phi - H_0)^2\, dG + \int \Phi(1 - \Phi)\, dG + 3a_n \int |\Phi - H_0|\, dG \\ &\le a_n(3 - a_n) \int |\Phi - H_0|\, dG + \int \Phi(1 - \Phi)\, dG < \infty. \end{aligned}$$

2. Take Definition 2.1.1 to show that

$$\int h_n^2\, dG \le 2(1 - a_n)^2 \int \varphi^2\, dG + 2a_n^2 \int h_0^2\, dG \in L^2(G), \qquad \text{uniformly in } n.$$

The triangle inequality yields

$$\Big|\ \| h_n \|_G^2 - \| \varphi \|_G^2\ \Big| \le \| h_n - \varphi \|_G^2 = a_n^2 \| h_0 - \varphi \|_G^2 = O(a_n^2).$$

3. Consequence of the definition of $h_n(x)$.

□

3 Minimum Distance Estimation

3.1 Description of the chapter

The aim of this chapter is to prove that a certain minimum distance estimator $\tilde{\rho}$ of an $AR(1)$-process $X_{in} = \rho X_{i-1,n} + \varepsilon_{in}$ with ε_{in} contaminated normal distributed, i.e. $\varepsilon_{in} \sim H_n = (1 - a_n)\Phi + a_n H_0$, is asymptotically normal distributed under some assumptions.

Since several lemmata are used more than once we divide the chapter in the following way: definitions and lemmata (Chapter 3.2), propositions (Chapter 3.3) and theorems (Chapter 3.4).

The proof can be divided into four parts:

Part 1

We define the original distance function $M_n(t)$ as

$$\begin{aligned} M_n(t) &:= \int \left[n^{-1/2} \sum_{i=1}^{n} X_{i-1,n} \left\{ I(X_{in} - tX_{i-1,n} \le x) - \Phi(x) \right\} \right]^2 dG(x) \\ &= \int \left[n^{-1/2} \sum_{i=1}^{n} X_{i-1,n} \left\{ I(\varepsilon_{in} \le x + (t-\rho)X_{i-1,n}) - \Phi(x) \right\} \right]^2 dG(x) \end{aligned}$$

and approximate it by a function $\hat{M}_n(t)$, whose integrand is a quadratic form in t, i.e.

$$\begin{aligned} \hat{M}_n(t) &:= \int \left[n^{-1/2} \sum_{i=1}^{n} X_{i-1,n} \left\{ I(\varepsilon_{in} \le x) - \Phi(x) + X_{i-1,n} h_n(x)(t-\rho) \right\} \right]^2 dG(x) \\ &= \int \Big[n^{-1/2} \sum_{i=1}^{n} X_{i-1,n} \{ [I(\varepsilon_{in} \le x) - H_n(x)] \\ &\qquad + [H_n(x) - \Phi(x)] + X_{i-1,n} h_n(x)(t-\rho) \} \Big]^2 dG(x) \\ &:= \int [W_n(x,\rho) + B_n(x) + n^{-1/2} \sum X_{i-1,n}^2 h_n(x)(t-\rho)]^2 \, dG(x). \end{aligned}$$

Lemmata 2.2.10(2)/3.2.6/3.2.7/3.2.11 and Corollary 3.2.12 are used to show that this function $\hat{M}_n(t)$ is well defined. Since it is a quadratic form it has a derivative and we find a minimum, $\hat{\rho}$, i.e.

$$\sqrt{n}(\hat{\rho} - \rho) = -\frac{\int W_n(x,\rho) + B_n(x))h_n(x)\, dG(x)}{Y_n \int h_n^2(x)\, dG(x)}.$$

Part 2

In the second part we prove the asymptotic distribution of $\sqrt{n}(\hat{\rho}-\rho)$ (Theorem 3.4.1), which is a consequence of Proposition 3.3.9.

To prove this theorem we need

- $b_n \to \frac{k\mu_0 I}{1-\rho}$, (Corollary 3.2.9),
- $Y_n \to \frac{1}{1-\rho^2}$ with $Y_n := n^{-1}\sum_{i=1}^n X_{i-1,n}^2$, (Corollary 3.2.13),
- $\int W_n(x,\rho)h_n(x)\,dG(x) \approx \int W_n(x,\rho)\varphi(x)\,dG(x)$, (Lemma 3.2.6),
- $\int h_n^2(x)\,dG(x) \approx \int \varphi(x)^2\,dG(x)$, (Lemma 2.2.10 (2)).

With these lemmata we can show that

$$\sqrt{n}(\hat{\rho}-\rho) \approx -\frac{\int W_n(x,\rho)\varphi(x)\,dG(x)}{\frac{1}{1-\rho^2}\int \varphi^2(x)\,dG(x)} - \frac{\frac{k\mu_0 I}{1-\rho}}{\frac{1}{1-\rho^2}\int \varphi^2(x)\,dG(x)}.$$

Furthermore $\int W_n(x,\rho)\varphi(x)\,dG(x) = -\frac{1}{\sqrt{n}}\sum_{i=1}^n X_{i-1,n}\{\psi(\varepsilon_{in}) - \mathrm{E}\psi(\varepsilon_{in})\} =: \mathcal{S}_n$ and Lemma 3.2.16 shows that $\{\mathcal{S}_n, \mathcal{F}_{i-1,n}\}$ with $\mathcal{F}_{i-1,n} = \sigma(X_{0n}, \varepsilon_{0n}, \ldots, \varepsilon_{i-1,n})$ is a zero mean, square-integrable martingale array and so we can apply Corollary B.18/Theorem B.17 to show that $\mathcal{S}_n$ is asymptotically normal distributed.

Finally we need Proposition B.11 and Proposition B.12 to finish the proof.

Part 3

Now we consider the difference $|\hat{M}_n(t) - M_n(t)|$ and show that over a compact set $\left[\rho - \frac{B}{\sqrt{n}}, \rho + \frac{B}{\sqrt{n}}\right]$ and certain assumptions (Proposition 3.3.3)

$$\sup_{t \in \left[\rho - \frac{B}{\sqrt{n}}, \rho + \frac{B}{\sqrt{n}}\right]} |\hat{M}_n(t) - M_n(t)| = o_p(1). \tag{3.1}$$

Therefore $\hat{M}_n(t)$ is a good approximation for $M_n(t)$. Lemmata 3.2.6/3.2.7/3.2.11 are tools that are needed for this proof.

As special case we consider $G(x) \equiv x$ and show that one assumption of Proposition 3.3.3 is trivially implied (Proposition 3.3.4).

Part 4

In part 3 we have proved that $\hat{M}_n(t)$ is a good approximation for our original function $M_n(t)$ over a compact set $\left[\rho - \frac{B}{\sqrt{n}}, \rho + \frac{B}{\sqrt{n}}\right]$. To get the asymptotic distribution of $\tilde{\rho}$ we have to be sure that the approximation (3.1) can be applied if $t = \hat{\rho}$ and $t = \tilde{\rho}$, because we need the equation $o_p(1) = |\hat{M}_n(\hat{\rho}) - M_n(\tilde{\rho})|$. Therefore we must be sure, that $\hat{\rho}$ and $\tilde{\rho}$ are lying inside this compact interval, i.e.

$$|\sqrt{n}(\tilde{\rho} - \rho)| \leq B \text{ and } |\sqrt{n}(\hat{\rho} - \rho)| \leq B,$$

which is equivalent to check conditions (3.10) and (3.11) in Lemma 3.2.17. The proofs of (3.10) and (3.11) are the contents of Proposition 3.3.6 and Proposition 3.3.7.

Since $M_n(t)$ is well approximated by $\hat{M}_n(t)$, and we know that $\hat{\rho}$ and $\tilde{\rho}$ are lying inside the compact set $\left[\rho - \frac{B}{\sqrt{n}}, \rho + \frac{B}{\sqrt{n}}\right]$, Theorem 3.4.4 guarantees

$$\sqrt{n}(\hat{\rho} - \tilde{\rho}) = o_p(1).$$

But this is equal to the statement, that $\hat{\rho}$ and $\tilde{\rho}$ have the same asymptotic distributions. So $\tilde{\rho}$ has asymptotic a normal distribution, too.

3.2 Definitions and Lemmata

Definition 3.2.1. *Define the error integral*

$$M_n(t) := \int \left[n^{-1/2} \sum_{i=1}^{n} X_{i-1,n} \left\{ I(\varepsilon_{in} \leq x + (t-\rho)X_{i-1,n}) - \Phi(x) \right\} \right]^2 dG(x), \tag{3.2}$$

where $G(x)$ is a nondecreasing real valued function that is right continuous and has left limits. Denote the integrand of $M_n(t)$ as

$$K_n(x,t) := n^{-1/2} \sum_{i=1}^{n} X_{i-1,n} \left\{ I(\varepsilon_{in} \leq x + (t-\rho)X_{i-1,n}) - \Phi(x) \right\}.$$

Definition 3.2.2. *We define the minimum distance estimator $\tilde{\rho}$ of $M_n(t)$ as*

$$\tilde{\rho} = \tilde{\rho}_n := \operatorname*{argmin}_t \; M_n(t). \tag{3.3}$$

Since the distribution of the $\{\varepsilon_{in}\}$ still does not appear in the definition of $M_n(t)$, we expand (3.2) with

$$H_n(x + (t-\rho)X_{i-1,n}) - H_n(x) - (t-\rho)X_{i-1,n}h_n(x).$$

So $M_n(t)$ becomes

$$M_n(t) = \int \Bigg[n^{-1/2} \sum_{i=1}^{n} X_{i-1,n} \left\{ I(\varepsilon_{in} \leq x + (t-\rho)X_{i-1,n}) - H_n(x + (t-\rho)X_{i-1,n}) \right\}$$
$$+ n^{-1/2} \sum_{i=1}^{n} X_{i-1,n} \left\{ H_n(x + (t-\rho)X_{i-1,n}) - H_n(x) - (t-\rho)X_{i-1,n} h_n(x) \right\}$$
$$+ n^{-1/2} \sum_{i=1}^{n} X_{i-1,n} \left\{ H_n(x) - \Phi(x) \right\} + n^{-1/2} \sum_{i=1}^{n} X_{i-1,n}^2 (t-\rho) h_n(x) \Bigg]^2 dG(x). \tag{3.4}$$

The term

$$n^{-1/2} \sum_{i=1}^{n} X_{i-1,n} \left\{ I(\varepsilon_{in} \leq x + (t-\rho)X_{i-1,n}) - H_n(x + (t-\rho)X_{i-1,n}) \right\}$$

represents a random weighted empirical process,

$$n^{-1/2} \sum_{i=1}^{n} X_{i-1,n} \left\{ H_n(x + (t-\rho)X_{i-1,n}) - H_n(x) - (t-\rho)X_{i-1,n} h_n(x) \right\}$$

is the expansion of H_n up to order 2 around the point x,

$$n^{-1/2} \sum_{i=1}^{n} X_{i-1,n} \left\{ H_n(x) - \Phi(x) \right\}$$

is a random weighted difference between the true distribution H_n and the modeled distribution Φ and finally

$$n^{-1/2} \sum_{i=1}^{n} X_{i-1,n}^2 (t-\rho) h_n(x)$$

is the rest term.

Definition 3.2.3. *Write*

$$W_n(x,t) := n^{-1/2} \sum_{i=1}^{n} X_{i-1,n} \{ I(\varepsilon_{in} \leq x + (t-\rho)X_{i-1,n}) - H_n(x + (t-\rho)X_{i-1,n}) \},$$
$$B_n(x) := n^{-1/2} \sum_{i=1}^{n} X_{i-1,n} \{ H_n(x) - \Phi(x) \}.$$

Rewriting (3.4) with these substitutions and expanding with $W_n(x,\rho)$ we get

$$\begin{aligned} M_n(t) = & \int \Big[W_n(x,t) - W_n(x,\rho) \\ & + n^{-1/2} \sum X_{i-1,n} \left\{ H_n(x + (t-\rho)X_{i-1,n}) - H_n(x) - (t-\rho)X_{i-1,n}h_n(x) \right\} \\ & + W_n(x,\rho) + B_n(x) + n^{-1/2} \sum X_{i-1,n}^2 h_n(x)(t-\rho) \Big]^2 \, dG(x). \end{aligned} \tag{3.5}$$

Definition 3.2.4. *Define the auxiliary function*

$$\hat{M}_n(t) := \int \left[W_n(x,\rho) + B_n(x) + n^{-1/2} \sum_{i=1}^{n} X_{i-1,n}^2 h_n(x)(t-\rho) \right]^2 dG(x) \tag{3.6}$$

and let

$$\hat{K}_n(x,t) := \left[W_n(x,\rho) + B_n(x) + Y_n h_n(x) \sqrt{n}(t-\rho) \right]^2$$

be the integrand where

$$Y_n := n^{-1} \sum_{i=1}^{n} X_{i-1,n}^2. \tag{3.7}$$

Definition 3.2.5. *We define the minimum distance estimator $\hat{\rho}$ of $\hat{M}_n(t)$ as*

$$\hat{\rho} = \hat{\rho}_n := \operatorname*{argmin}_t \; \hat{M}_n(t). \tag{3.8}$$

In Lemma 3.2.6 - Lemma 3.2.16 the terms of (3.5) are investigated. To prove the propositions of Chapter 3.3 and the theorems of Chapter 3.4 it must be assured that each expression is at least bounded in probability.

Lemma 3.2.6. *Suppose $\int \Phi(1-\Phi)\, dG < \infty$, $\int |\Phi - H_0|\, dG < \infty$, $\int h_0^2\, dG < \infty$, $\int \varphi^2\, dG < \infty$. Then*

$$\int W_n^2(x,\rho)\, dG(x) = O_p(1)$$

and

$$\int W_n(x,\rho)h_n(x)\, dG(x) = \int W_n(x,\rho)\varphi(x)\, dG(x) + O_p(a_n).$$

Proof. Consider Definition B.16. Note that

- $\mathrm{E}I(\varepsilon_{in} \le x) = H_n(x)$ for all $i \le n$ and all $x \in \mathbb{R}$ and

- $\mathrm{E}X = \mathrm{E}(\mathrm{E}(X|\mathcal{F}))$, (Breiman (1992), p. 75).

Then with Lemma 2.2.10 we find

$$\mathrm{E}\int W_n^2(x,\rho)\,dG(x) = \frac{1}{n}\int \mathrm{E}\left(\sum_{i=1}^{n} X_{i-1,n}\left(I(\varepsilon_{in}\le x)-H_n(x)\right)\right)^2 dG(x)$$

$$= \mathrm{E}X_{0n}^2\int H_n(x)(1-H_n(x))\,dG(x)$$

$$+\frac{2}{n}\int\sum_{i<j}\mathrm{E}\left\{X_{i-1,n}X_{j-1,n}[I(\varepsilon_{in}\le x)-H_n(x)][I(\varepsilon_{jn}\le x)-H_n(x)]\right\}$$

$$= \mathrm{E}X_{0n}^2\int H_n(x)(1-H_n(x))\,dG(x)$$

$$+\frac{2}{n}\int\sum_{i<j}\mathrm{E}\left\{\mathrm{E}\left\{X_{i-1,n}X_{j-1,n}[I(\varepsilon_{in}\le x)-H_n(x)][I(\varepsilon_{jn}\le x)-H_n(x)]\big|\mathcal{F}_{j-1,n}\right\}\right\}$$

$$= \mathrm{E}X_{0n}^2\int H_n(x)(1-H_n(x))\,dG(x)$$

$$+\frac{2}{n}\int\sum_{i<j}\mathrm{E}\left\{X_{i-1,n}X_{j-1,n}[I(\varepsilon_{in}\le x)-H_n(x)]\right\}\mathrm{E}\left\{[I(\varepsilon_{jn}\le x)-H_n(x)]\big|\mathcal{F}_{j-1,n}\right\}$$

$$= \mathrm{E}X_{0n}^2\int H_n(x)(1-H_n(x))\,dG(x)<\infty,\quad \text{uniformly in } n\ge 1.$$

The claim is then a consequence of Markov's inequality (Corollary B.8).

For $\int W_n(x,\rho)h_n(x)\,dG(x)$ we need the Cauchy-Schwarz inequality and the first part of the lemma:

$$\begin{aligned}
&\left|\int W_n(x,\rho)\ h_n(x)\,dG(x)-\int W_n(x,\rho)\varphi(x)\,dG(x)\right|\\
&\qquad=\left|\int W_n(x,\rho)(h_n(x)-\varphi(x))\,dG(x)\right|\\
&\qquad\le\sqrt{\int W_n^2(x,\rho)\,dG(x)\int[h_n(x)-\varphi(x)]^2\,dG(x)}\\
&\qquad=a_n\sqrt{\int W_n^2(x,\rho)\,dG(x)}\sqrt{\int[h_0(x)-\varphi(x)]^2\,dG(x)}\\
&\qquad=O_p(a_n).
\end{aligned}$$

□

Lemma 3.2.7. *Assume* $\int(H_0-\Phi)^2\,dG<\infty$. *Then*

$$\int B_n^2(x)dG(x)=O_p(na_n^4)+O_p(a_n^2)$$

where $B_n(x)$ is as in Definition 3.2.3.

Proof. Use Markov's inequality (Corollary B.8) and Lemma 2.2.6(3) to show that

$$P\left(\frac{a_n}{\sqrt{n}}\left|\sum_{i=1}^{n} X_{i-1,n}\right| \geq \eta\right) \leq \frac{a_n^2\,\mathrm{E}\left(\sum_{i=1}^{n} X_{i-1,n}\right)^2}{n\varepsilon^2} = O(a_n^2) + O(na_n^4). \qquad (3.9)$$

With the assumption $\int (H_0 - \Phi)^2\, dG < \infty$ and (3.9) we find

$$\begin{aligned}\int B_n^2(x)\, dG(x) &= n^{-1}\left(\sum_{i=1}^{n} X_{i-1,n}\right)^2 \int [H_n(x) - \Phi(x)]^2\, dG(x) \\ &= \left[\frac{a_n}{\sqrt{n}}\sum_{i=1}^{n} X_{i-1,n}\right]^2 \int [H_0(x) - \Phi(x)]^2\, dG(x) \\ &= O_p(na_n^4) + O_p(a_n^2).\end{aligned}$$

□

Lemma 3.2.8. *Suppose $\int |\Phi - H_0|\, dG < \infty$, $\int |\Phi - H_0| h_0\, dG < \infty$. With $b_n := \int B_n(x) h_n(x)\, dG(x)$ we find*

$$P(|b_n - \mathrm{E}b_n| > \eta) \leq O(a_n^2).$$

Proof. We will prove this lemma with the Chebychev inequality (Lemma B.9). Remember the defintion of $B_n(x)$ (Definition 3.2.3). Define

$$\begin{aligned}I &:= \int (H_0(x) - \Phi(x))\varphi(x)\, dG(x), \\ J &:= \int (H_0(x) - \Phi(x))(\varphi(x) - h_0(x))\, dG(x), \\ K_n &:= a_n J - I = O(1).\end{aligned}$$

Then we get

$$b_n = \int B_n(x) h_n(x)\, dG(x) = \frac{a_n}{\sqrt{n}} K_n \sum_{i=1}^{n} X_{i-1,n}.$$

Now we calculate $\mathrm{E}b_n$, i.e.

$$\mathrm{E}b_n = \frac{a_n^2\sqrt{n}\mu_0}{1-\rho}\, K_n,$$

and $\mathrm{E}b_n^2$, i.e.

$$\begin{aligned}
\mathrm{E}b_n^2 &= \mathrm{E}\left(\int B_n(x)h_n(x)\,dG(x)\right)^2 = \mathrm{E}\left(\frac{a_n}{\sqrt{n}}\sum X_{i-1,n}K_n\right)^2 \\
&= \frac{a_n^2}{n}K_n^2\,\mathrm{E}\left(\sum X_{i-1,n}\right)^2 \\
&= \frac{a_n^2}{n}K_n^2\left(\frac{n}{1-\rho^2} + O(na_n) + 2\sum_{i<j}\left(\frac{\rho^{j-i}}{1-\rho^2} + \rho^{j-i}O(a_n) + \frac{a_n^2\mu_0^2}{(1-\rho)^2}\right)\right) \\
&= K_n^2\left(\frac{a_n^2}{1-\rho^2} + O(a_n^3) + O(a_n^2) + (n-1)\frac{a_n^4\mu_0^2}{(1-\rho)^2}\right) \\
&= K_n^2\left(\frac{na_n^4\mu_0^2}{(1-\rho)^2} + O(a_n^2)\right).
\end{aligned}$$

Therefore the variance of b_n is

$$\mathrm{Var}\ b_n = \mathrm{E}b_n^2 - (\mathrm{E}b_n)^2 = K_n^2 O(a_n^2) = O(a_n^2).$$

Putting together the three parts we find

$$P(|b_n - \mathrm{E}b_n| > \eta) \le \frac{1}{\eta^2}O(a_n^2).$$

□

Corollary 3.2.9. *If* $\lim_{n\to\infty}\sqrt{n}a_n^2 =: k < \infty$. *Then, with the notations of the proof of Lemma 3.2.8,*

$$b_n \xrightarrow{p} \frac{k\mu_0 I}{1-\rho}.$$

Proof. It is a consequence of the proof of Lemma 3.2.8.

Remark 3.2.10. Note, that $\int B_n^2(x)\,dG(x) = O_p(1)$ and $\mathrm{E}b_n = O(1)$ if $a_n = O(n^{-1/4})$.

Lemma 3.2.11. *Assume that* $\mathrm{E}\delta_0^3$, $\mathrm{E}\delta_0^4$ *exist. Then*

$$P(|Y_n - \mathrm{E}Y_n| > \eta) \le O(a_n) + O(n^{-1}).$$

Proof. We use the Chebyshev's inequality (Corollary B.9).
Remember that

$$\mathrm{E}Y_n = \frac{1}{n}\mathrm{E}\sum_{i=1}^{n}X_{i-1,n}^2 = \frac{1}{1-\rho^2} + O(a_n).$$

Note that

$$\sum_{j=2}^{n}\sum_{i=1}^{j-1}\rho^{2(j-i)} = \frac{-2\rho^2+\rho^4}{(-1+\rho^2)^2} - \frac{-\rho^2(n+1)+\rho^4(n+1)-\rho^{2(n+1)}}{(-1+\rho^2)^2} = O(n).$$

With the results of Lemma 2.2.6 we can show that

$$\begin{aligned}
\mathrm{E}Y_n^2 &= \frac{1}{n^2}\mathrm{E}\sum_{i=1}^{n}\sum_{j=1}^{n} X_{i-1,n}^2 X_{j-1,n}^2 \\
&= \frac{1}{n^2}\sum_{i=1}^{n}\mathrm{E}X_{i-1,n}^4 + \frac{1}{n^2}\mathrm{E}\sum_{i\neq j} X_{i-1,n}^2 X_{j-1,n}^2 \\
&= \frac{1}{n^2}\sum_{i=1}^{n}\mathrm{E}X_{i-1,n}^4 + \frac{1}{n^2}\sum_{j=2}^{n}\sum_{i=1}^{j-1}\frac{1+2\rho^{2(j-i)}}{(1-\rho^2)^2} \\
&\quad + \frac{1}{n^2}\sum_{i=2}^{n}\sum_{j=1}^{i-1}\frac{1+2\rho^{2(i-j)}}{(1-\rho^2)^2} + O(a_n) \\
&= O(n^{-1}) + \frac{n(n-1)}{n^2(1-\rho^2)^2} + O(a_n) \\
&= \frac{1}{(1-\rho^2)^2} + O(a_n) + O(n^{-1})
\end{aligned}$$

and

$$\mathrm{Var}\ Y_n = \mathrm{E}Y_n^2 - (\mathrm{E}Y_n)^2 = O(a_n) + O(n^{-1}).$$

□

Corollary 3.2.12. *Suppose that* $\mathrm{E}\delta_0^3$, $\mathrm{E}\delta_0^4$ *exist. Then*

$$\mathrm{E}Y_n^2 < \infty.$$

Proof. It is a consequence of the proof of Lemma 3.2.11.

Corollary 3.2.13. *Let* $a_n = O(n^{-1/4})$. *Then*

$$Y_n \xrightarrow{P} \frac{1}{1-\rho^2}.$$

Proof. It is a consequence of Lemma 3.2.11.

To prove Lemma 3.2.16 we introduce a new function ψ and the random variables ξ_{in}, Z_{in} and $\mathcal{S}_n$.

Definition 3.2.14. *Assume* $\varphi \in L^1(G)$ *and define*

$$\begin{aligned}\psi(x) &:= \int_{-\infty}^{x} \varphi(y)\, dG(y), \quad x \in \mathbb{R},\\ \xi_{in} &:= -n^{-1/2} X_{i-1,n}\{\psi(\varepsilon_{in}) - \mathrm{E}\psi(\varepsilon_{in})\},\\ Z_{in} &:= \psi(\varepsilon_{in}) - \mathrm{E}\psi(\varepsilon_{in}),\\ \mathcal{S}_n &:= \sum_{i=1}^{n} \xi_{in}.\end{aligned}$$

Lemma 3.2.15. *The random variable* Z_{in} *is uniformly bounded and has finite variance.*

Proof. Define $K := \psi(\infty)$ and note, that $\psi(x)$ is a nondecreasing, positive and bounded function. Then

$$\begin{aligned}0 &\leq \mathrm{E}\psi(\varepsilon_{in}) \leq \mathrm{E}\psi(\infty) = K < \infty,\\ 0 &\leq \mathrm{E}\psi^2(\varepsilon_{in}) \leq \mathrm{E}\psi^2(\infty) = K^2 < \infty.\end{aligned}$$

For Z_{in} we therefore find

$$\begin{aligned}EZ_{in} &= 0,\\ |Z_{in}| &= |\psi(\varepsilon_{in}) - \mathrm{E}\psi(\varepsilon_{in})| \leq K,\\ \mathrm{E}Z_{in}^2 &= \mathrm{E}\psi^2(\varepsilon_{in}) - (\mathrm{E}\psi(\varepsilon_{in}))^2 \leq K^2.\end{aligned}$$

Lemma 3.2.16. *Assume that* $a_n \to 0$ *as* $n \to \infty$. *Consider the notation of Definition 3.2.14. We find*

$$\mathcal{S}_n \xrightarrow{d} N\left(0, \frac{\sigma_\psi^2}{1-\rho^2}\right),$$

with $\sigma_\psi^2 := \mathrm{E}Z_{0n}^2$ *as* $n \to \infty$.

Proof. Consider the integral $\int W_n(x,\rho)\varphi(x)\, dG(x)$. We find

$$\begin{aligned}\int W_n(x,\rho)\varphi(x)\, dG(x) &= \int \left[n^{-1/2} \sum_i X_{i-1,n} \{I(\varepsilon_{in} \leq x) - H_n(x)\} \right] d\psi(x)\\ &= n^{-1/2} \sum_i X_{i-1,n} \left\{ \psi(\infty) - \psi(\varepsilon_{in}) - \int H_n(x)\, d\psi(x) \right\}\\ &= n^{-1/2} \sum_i X_{i-1,n} \left\{ -\psi(\varepsilon_{in}) + \int (1 - H_n(x))\, d\psi(x) \right\}\\ &= -n^{-1/2} \sum_i X_{i-1,n} \left\{ \psi(\varepsilon_{in}) - \int \psi(x)\, dH_n(x) \right\}\end{aligned}$$

$$= -n^{-1/2} \sum_i X_{i-1,n} \{\psi(\varepsilon_{in}) - E\psi(\varepsilon_{in})\}$$
$$=: \sum_i \xi_{in} =: \mathcal{S}_n.$$

Now $\{\mathcal{S}_n, \mathcal{F}_{i-1,n}\}$ with $\mathcal{F}_{i-1,n} = \sigma(X_{0n}, \varepsilon_{0n}, \dots, \varepsilon_{i-1,n})$ is a zero mean, square-integrable martingale array. To use Corollary B.18 we need to check the assumptions (B.4), (B.5), (B.6).

(B.4) ok.

(B.5) Let $Z_{in} = \psi(\varepsilon_{in}) - \mathrm{E}\psi(\varepsilon_{in})$. Lemma 3.2.15 showed that $\max_i |Z_{in}| =: L$ is bounded and $\mathrm{E}Z_{in}^2 = O(1)$. Then

$$\begin{aligned}
&\sum_i \mathrm{E}\left(\xi_{in}^2 I(|\xi_{in}| > \delta)\Big|\mathcal{F}_{i-1,n}\right)\\
&= n^{-1} \sum_i \mathrm{E}\left(X_{i-1,n}^2 Z_{in}^2 I(|X_{i-1,n} Z_{in}| > n^{1/2}\delta)\Big|\mathcal{F}_{i-1,n}\right)\\
&\le n^{-1} \sum_i X_{i-1,n}^2 \,\mathrm{E}\left(L^2 I(|X_{i-1,n} L| > n^{1/2}\delta)\Big|\mathcal{F}_{i-1,n}\right)\\
&= n^{-1} \sum_i X_{i-1,n}^2 L^2 I\left(|X_{i-1,n}| > \frac{n^{1/2}\delta}{L}\right)\\
&\to 0,
\end{aligned}$$

since $I\left(|X_{i-1,n}| > \frac{n^{1/2}\delta}{L}\right) \to 0$ as $n \to \infty$.

(B.6) First we calculate $\mathrm{E}Z_{in}^2$. Since the ε_{in} are i.i.d. contaminated normal we get

$$\begin{aligned}
\mathrm{E}Z_{in}^2 &= \mathrm{Var}\ \psi(\varepsilon_{in}) = \mathrm{E}\psi^2(\varepsilon_{in}) - [\mathrm{E}\psi(\varepsilon_{in})]^2\\
&= \int \psi^2 \varphi \, dx - a_n \int \psi^2 (\varphi - h_0)\, dx\\
&\qquad - \left(\int \psi\varphi\, dx - a_n \int \psi(\varphi - h_0)\, dx\right)^2\\
&= \int \psi^2 \, d\Phi - \left(\int \psi \, d\Phi\right)^2 + O(a_n).
\end{aligned}$$

The conditional expectation of ξ_{in}^2 given $\mathcal{F}_{i-1,n}$ is

$$\begin{aligned}
\sum_{i=1}^n \mathrm{E}(\xi_{in}^2|\mathcal{F}_{i-1,n}) &= \sum_{i=1}^n \mathrm{E}(n^{-1} X_{i-1,n}^2 Z_{in}^2 |\mathcal{F}_{i-1,n})\\
&= n^{-1} \sum_{i=1}^n X_{i-1,n}^2 \,\mathrm{E}Z_{0n}^2 \xrightarrow{p} \frac{\sigma_\psi^2}{1-\rho^2}.
\end{aligned}$$

So the assumptions of Theorem B.17 are satisfied and the statement is proved. □

Finally three lemmata are used to prove the Main Theorem (Theorem 3.4).

Lemma 3.2.17. *Let f_n be a sequence of stochastic functions and*

$$\gamma_n := \operatorname*{argmin}_t \; f_n(t).$$

The conditions

- *$\forall\, \eta > 0$, there exist a $0 < k_\eta < \infty$ and $N_1 = N_1(\eta)$, such that for all $n \geq N_1$,*

$$P(\,|\, f_n(\rho)\,|\; \leq k_\eta) \geq 1 - \eta, \tag{3.10}$$

- *$\forall\, 0 < \alpha < \infty$ and $\eta > 0$ there exist a $N_2 = N_2(\eta, \alpha)$ and a $b = b(\eta, \alpha)$, such that for all $n \geq N_2$,*

$$P\left(\inf_{t \in \left[\rho - \frac{b}{\sqrt{n}}, \rho + \frac{b}{\sqrt{n}}\right]^c} f_n(t) \geq \alpha \right) \geq 1 - \eta, \tag{3.11}$$

imply

$$|n^{1/2}(\gamma_n - \rho)| = O_p(1).$$

Proof. Suppose $\alpha > k_\eta$. Then

$$\begin{aligned}
\{n : |f_n(\rho)| \leq k_\eta\} \cap \{n : \inf_{|h|>b} f_n(\rho + n^{-1/2}h) \geq \alpha\} \\
&= \{n : |f_n(\rho)| \leq k_\eta, \inf_{|h|>b} f_n(\rho + n^{-1/2}h) \geq \alpha\} \\
&\subset \{n : \inf_{|h|\leq b} f_n(\rho + n^{-1/2}h) \leq k_\eta, \inf_{|h|>b} f_n(\rho + n^{-1/2}h) \geq \alpha\} \\
&\subset \{n : \inf_{|h|>b} f_n(\rho + n^{-1/2}h) > \inf_{|h|\leq b} f_n(\rho + n^{-1/2}h)\}.
\end{aligned}$$

So, for all $n \geq \max\{N_1, N_2\}$ the infimum outside the intervall $[-b, b]$ is greater than inside and therefore $|n^{1/2}(\gamma_n - \rho)| = O_p(1)$. □

Lemma 3.2.18. *Let f and g be bounded functions. Then*

$$|\inf_x f(x) - \inf_x g(x)| \leq \sup_x |f(x) - g(x)|.$$

Proof. Assume that $d := |\inf f - \inf g| > \sup |f - g|$. Without loss of generality $\inf f > \inf g$. Then there exist $\eta > 0$ with $d > \sup|f - g| + \eta$ and a y so that $|g(y) - \inf g| < \eta$ and $\inf f > g(y)$. Then

$$\begin{aligned}
f(y) - g(y) + \eta = |f(y) - g(y)| + \eta \leq \sup|f - g| + \eta < d \\
&= |\inf f - g(y) + g(y) - \inf g| \\
&\leq |\inf f - g(y)| + |g(y) - \inf g| \\
&\leq \inf f - g(y) + \eta.
\end{aligned}$$

So $f(y) < \inf f$, which is a contradition. □

Lemma 3.2.19. *Let U_i be a stationary sequence of random variables and* $\mathrm{E}U_0^2 < \infty$. *Then*

$$\frac{\max_{1\le i\le n}|U_i|}{\sqrt{n}} = o_p(1).$$

Proof. Since $\mathrm{E}U_0^2$ exists we get

$$\begin{aligned} P\left(\frac{\max_i |U_i|}{\sqrt{n}} > \vartheta\right) &\le \sum_{i=1}^{n} P(|U_i| > \vartheta\sqrt{n}) = nP(|U_0| > \vartheta\sqrt{n}) \\ &\le \frac{1}{\vartheta^2}\,\mathrm{E}U_0^2 I(|U_0| > \vartheta\sqrt{n}) \to 0, \qquad \text{as } n \to \infty. \end{aligned}$$

□

3.3 Propositions

Now we will compare the two functions $M_n(t)$ and $\hat{M}_n(t)$. With the results of Lemma 3.2.6 - Lemma 3.2.16 we are able to prove Proposition 3.3.3.

Definition 3.3.1. *We use following notations:*

$$A_n(x,t) := W_n(x,t) - W_n(x,\rho),$$

$$C_n(x,t) := n^{-1/2}\sum_{i=1}^{n} X_{i-1,n}\{H_n(x+(t-\rho)X_{i-1,n}) - H_n(x) - (t-\rho)X_{i-1,n}h_n(x)\},$$

$$D_n(x,t) := W_n(x,\rho) + n^{-1/2}\sum X_{i-1}^2 h_n(x)(t-\rho).$$

Definition 3.3.2. *Let*

$$V_n(\rho,B) := \left[\rho - \frac{B}{\sqrt{n}}, \rho + \frac{B}{\sqrt{n}}\right]$$

and

$$V_n^c(\rho,B) := \mathbb{R}\backslash V_n(\rho,B).$$

Proposition 3.3.3. *Assume that*

$$\begin{gathered} h_0, \varphi \in L^2(G), \\ (H_0 - \Phi), \Phi(1-\Phi) \in L^1(G), \\ \mathrm{E}\delta_0^4,\ \mathrm{E}\delta_0^3 \textit{ exist}, \\ 0 < B < \infty, \\ \sup_{t\in V_n(\rho,B)} \int C_n^2(x,t)\,dG(x) = o_p(1), \\ a_n = O(n^{-1/4}). \end{gathered}$$

Then

$$\sup_{t\in V_n(\rho,B)} |M_n(t) - \hat{M}_n(t)| = o_p(1), \tag{3.12}$$

or equivalent

$$\sup_{|u|\le B} |M_n(\rho + n^{-1/2}u) - \hat{M}_n(\rho + n^{-1/2}u)| = o_p(1). \tag{3.13}$$

Proof. Consider Definitions 3.2.3 and 3.3.1. Write

$$M_n(t) = \int (A_n(x,t) + C_n(x,t) + D_n(x,t) + B_n(x))^2 \, dG,$$
$$\hat{M}_n(t) = \int (D_n(x,t) + B_n(x))^2 \, dG.$$

Then

$$\begin{aligned} M_n(t) - \hat{M}_n(t) = &\int \left(A_n^2(x,t) + C_n^2(x,t)\right) \, dG(x) + \\ &+ 2\int A_n(x,t)(B_n(x) + C_n(x,t) + D_n(x,t)) \, dG(x) + \\ &+ 2\int C_n(x,t)(B_n(x) + D_n(x,t)) \, dG(x). \end{aligned}$$

In Koul (1992) (Lemma 7.4.3) it is shown, that $\sup_{V_n(\rho,B)} \int A_n^2(x,t)\, dG(x) = o_p(1)$. Lemma 3.2.7 with Remark 3.2.10 assures that $\int B_n^2(x)\, dG(x) = O_p(1)$. Since

$$\begin{aligned} \sup_{t\in V_n(\rho,B)} \int D_n^2(x,t)\, dG(x) \le 2\int W_n^2(x,\rho)\, dG(x) + \\ + 2 \sup_{t\in V_n(\rho,B)} |n^{1/2}(t-\rho)|^2\, Y_n^2 \int h_n^2(x)\, dG(x) \end{aligned} \tag{3.14}$$

is $O_p(1)$ (Lemma 3.2.6, Corollary 3.2.12 and $0 < B < \infty$), the proposition is a consequence of the Cauchy-Schwarz inequality. So we have

$$\sup_{t\in V_n(\rho,B)} |M_n(t) - \hat{M}_n(t)| = o_p(1).$$

□

Proposition 3.3.4. *Consider the case $G(x) \equiv x$. Assume that*

$$\int \left\{ \sup_{|v|\le s} |h_0(x+v) - h_0(x)| \right\}^2 dx \to 0, \quad as \quad s \to 0. \tag{3.15}$$

Then

$$\sup_{t\in V_n(\rho,B)} \int C_n^2(x,t)\, dG(x) = o_p(1).$$

Proof. Write $u = t - \rho$ and start with the expression

$$\begin{aligned}C_n(x, u+\rho) &= n^{-1/2}\sum_{i=1}^{n} X_{i-1,n}\left\{H_n(x+uX_{i-1,n}) - H_n(x) - uX_{i-1,n}h_n(x)\right\}\\ &= (1-a_n)n^{-1/2}\sum_{i=1}^{n} X_{i-1,n}\left\{\Phi(x+uX_{i-1,n}) - \Phi(x) - uX_{i-1,n}\varphi(x)\right\}\\ &\quad + a_n n^{-1/2}\sum_{i=1}^{n} X_{i-1,n}\left\{H_0(x+uX_{i-1,n}) - H_0(x) - uX_{i-1,n}h_0(x)\right\}\\ &=: (1-a_n)C_n^*(x,u+\rho) + a_n C_n^{**}(x,u+\rho).\end{aligned}$$

Now with the inequality $(a+b)^2 \leq 2a^2 + 2b^2$ we obtain

$$\int [C_n(x,u+\rho)]^2\,dx \leq 2(1-a_n)^2\int \left[C_n^*(x,u+\rho)\right]^2\,dx + 2a_n^2\int \left[C_n^{**}(x,u+\rho)\right]^2\,dx.$$

The next step consists in showing that both parts on the right side are converging to zero in probability when taking the supremum over $|n^{1/2}u| \leq B$.

Consider the second term first. To begin with we have the representation

$$H_0(x+tv) - H_0(x) = v\int_0^t h_0(x+sv)\,ds, \qquad \forall\, t>0,\ -\infty < x,\ v < \infty.$$

Also note that

$$\sup_{|u|\leq \frac{B}{\sqrt{n}}}\int (C_n^{**}(x,u+\rho))^2\,dx = \sup_{|t|\leq B}\int \left(C_n^{**}(x,tn^{-1/2}+\rho)\right)^2\,dx.$$

Now write, for a $t>0,\ -\infty<x<\infty$,

$$\begin{aligned}C_n^{**}&(x,tn^{-1/2}+\rho)\\ &= n^{-1/2}\sum_{i=1}^{n} X_{i-1,n}\left\{H_0(x+t\,n^{-1/2}X_{i-1,n}) - H_0(x) - t\,n^{-1/2}X_{i-1,n}h_0(x)\right\}\\ &= n^{-1}\sum_{i=1}^{n} X_{i-1,n}^2\int_0^t \left\{h_0(x+s\,n^{-1/2}X_{i-1,n}) - h_0(x)\right\}\,ds.\end{aligned}$$

Therefore, for all $0<t\leq B$ and for all real x,

$$|C_n^{**}(x,tn^{-1/2}+\rho)| \leq n^{-1}\sum_{i=1}^{n} X_{i-1,n}^2\int_0^B \left|h_0(x+s\,n^{-1/2}X_{i-1,n}) - h_0(x)\right|\,ds.$$

One obtains a similar bound for $-B\leq t\leq 0$. Putting these bounds together we obtain that for all $|t|\leq B$ and for all real x,

$$|C_n^{**}(x,tn^{-1/2}+\rho)| \leq n^{-1}\sum_{i=1}^{n} X_{i-1,n}^2\int_{-B}^B \left|h_0(x+s\,n^{-1/2}X_{i-1,n}) - h_0(x)\right|\,ds.$$

Now let $A_{n,\delta} := [\max_{1\le i\le n} n^{-1/2}|X_{i-1,n}| \le \delta]$. Note that on $A_{n,\delta}$,

$$|C_n^{**}(x, tn^{-1/2} + \rho)| \le 2B\ n^{-1} \sum_{i=1}^{n} X_{i-1,n}^2 \sup_{|v|\le \delta B} |h_0(x+v) - h_0(x)|,$$

so that

$$\begin{aligned}
&\sup_{|t|\le B} \int \left(C_n^{**}(x, tn^{-1/2} + \rho)\right)^2\ dx \\
&\quad\le 4B^2 \left[n^{-1} \sum_{i=1}^{n} X_{i-1,n}^2\right]^2 \int \left\{\sup_{|v|\le \delta B} |h_0(x+v) - h_0(x)|\right\}^2\ dx \\
&\quad= o_p(1),
\end{aligned}$$

by (3.15) and by Lemma 3.2.11 which guarantees that $n^{-1}\sum_{i=1}^n X_{i-1,n}^2 = O_p(1)$, and by the fact that for every $\delta > 0$, $P(A_{n,\delta}) \to 1$, which inturn follows from the fact

$$\frac{\max_{1\le i\le n} |X_{i-1,n}|}{\sqrt{n}} = o_p(1).$$

Now, in the case $h_0 = \varphi$, then the left hand side of (3.15) satisfies, in view of the Cauchy-Schwarz inequality and Fubini's Theorem,

$$\begin{aligned}
\int \left\{\sup_{|v|\le s} |\varphi(x+v) - \varphi(x)|\right\}^2\ dx &\le \int \left[\int_{-s}^{s} \varphi'(x+t)\ dt\right]^2\ dx \\
&\le 2s \int \int_{-s}^{s} [\varphi'(x+t)]^2\ dt\ dx \\
&= 4s^2 \int [\varphi'(x)]^2\ dx = o(1), \quad as \quad s \to 0.
\end{aligned}$$

So (3.15) is *a priori* satisfied by the normal density and hence by an argument similar to the above one obtains

$$\sup_{|t|\le B} \int \left(C_n^{*}(x, tn^{-1/2} + \rho)\right)^2\ dx = o_p(1).$$

Thus the proposition follows. □

Proposition 3.3.5. *With the assumptions of Proposition 3.3.3 we find*

$$\left|\inf_{t\in V_n(\rho,B)} \hat{M}_n(t) - \inf_{t\in V_n(\rho,B)} M_n(t)\right| = o_p(1).$$

Proof. It is a consequence of Lemma 3.2.18 and Proposition 3.3.3. □

Proposition 3.3.6. *With the assumptions of Proposition 3.3.3*

1. $\hat{M}_n(\rho)$ *and*

2. $M_n(\rho)$

satisfy the condition (3.10).

Proof. 1. It is a consequence of Lemmata 3.2.7 and 3.2.6.

2. It is a consequence of Proposition 3.3.3 and 1.

□

Proposition 3.3.7. *With the assumptions of Proposition 3.3.3*

1. $\hat{M}_n(t)$ *and*

2. $M_n(t)$

satisfy the condition (3.11).

Proof. 1. Define the function

$$\hat{S}_n(x,t) := W_n(x,\rho) + B_n(x) + Y_n h_n(x)\sqrt{n}(t-\rho)$$

which is nondecreasing in t since $Y_n \geq 0$. With the Cauchy-Schwarz inequality we find

$$\left(\int \hat{S}_n(x,t)\varphi(x)\,dG(x)\right)^2 \leq \int \hat{S}_n^2(x,t)\,dG(x)\int \varphi^2(x)\,dG(x)$$

and so

$$\hat{M}_n(t) \geq \frac{\left(\int \hat{S}_n(x,t)\varphi(x)\,dG(x)\right)^2}{\int \varphi^2(x)\,dG(x)}, \qquad \forall t.$$

Define $\hat{T}_n(t) := \int \hat{S}_n(x,t)\varphi(x)\,dG(x)$. It is obvious that $\hat{T}_n(t)$ is a

(a) nondecreasing function in t for every n,

(b) $\inf_t \hat{T}_n(t) < 0$ and $\sup_t \hat{T}_n(t) > 0$ for every n.

For the next steps we use the inequalities

$$|a+b| \geq \Big||a|-|b|\Big| \geq |b|-|a|.$$

As a consequence of Lemmata 3.2.6 and 3.2.7 for any $\eta > 0$ there exist a N'_η and a K_η such that

$$P\left(\left|\int [W_n(x,\rho)+B_n(x)]\varphi(x)\,dG(x)\right| \leq K_\eta\right) \geq 1-\eta/2, \qquad \forall n \geq N'_\eta.$$

Choose

$$B \geq \frac{K_\eta + (z \int \varphi^2 \, dG)^{1/2}}{Y_n \int h_n \varphi \, dG}$$

and define

$$E := \left\{ \rho - \frac{B}{\sqrt{n}}, \, \rho + \frac{B}{\sqrt{n}} \right\},$$
$$V_n^c(\rho, B) = \left[\rho - \frac{B}{\sqrt{n}}, \, \rho + \frac{B}{\sqrt{n}} \right]^c.$$

Then

$$\begin{aligned}
P\left(\frac{\inf_{t\in E}\{\hat{T}_n^2(t)\}}{\int \varphi^2 \, dG} \geq z \right) &= P\left(\inf_{t\in E} |\hat{T}_n(t)| \geq \left(z \int \varphi^2 \, dG \right)^{1/2} \right) \\
&\geq P\left(\inf_{t\in E} \left| \left| \int [W_n(x,\rho) + B_n(x)] \varphi(x) \, dG(x) \right| \right.\right. \\
&\quad \left.\left. - \left| \sqrt{n}(t-\rho) Y_n \int h_n(x)\varphi(x) \, dG(x) \right| \right| \geq \left(z \int \varphi^2 \, dG \right)^{1/2} \right) \\
&\geq P\left(\left| \left| \int [W_n(x,\rho) + B_n(x)] \varphi(x) \, dG(x) \right| \right.\right. \\
&\quad \left.\left. - \, B \, Y_n \int h_n(x)\varphi(x) \, dG(x) \right| \geq \left(z \int \varphi^2 \, dG \right)^{1/2} \right) \\
&\geq P\left(- \left| \int [W_n(x,\rho) + B_n(x)] \varphi(x) \, dG(x) \right| \right. \\
&\quad \left. \geq \left(z \int \varphi^2 \, dG \right)^{1/2} - B \, Y_n \int h_n(x)\varphi(x) \, dG(x) \right) \\
&= P\left(\left| \int [W_n(x,\rho) + B_n(x)] \varphi(x) \, dG(x) \right| \right. \\
&\quad \left. \leq B \, Y_n \int h_n(x)\varphi(x) \, dG(x) - \left(z \int \varphi^2 \, dG \right)^{1/2} \right) \\
&\geq P\left(\left| \int [W_n(x,\rho) + B_n(x)] \varphi(x) \, dG(x) \right| \leq K_\eta \right) \\
&\geq 1 - \eta/2, \qquad \forall n \geq N'_\eta.
\end{aligned}$$

Since $\hat{T}_n(t)$ is a nondecreasing function we conclude that

$$P\left(\frac{\inf_{t\in E}\{\hat{T}_n^2(t)\}}{\int \varphi^2 \, dG} \geq z \right) \leq P\left(\frac{\inf_{t\in V_n^c(\rho,B)}\{\hat{T}_n^2(t)\}}{\int \varphi^2 \, dG} \geq z \right).$$

2. We divide the proof in several parts.

(a) Consider the function

$$S_n(x,t) := n^{-1/2}\sum_{i=1}^{n} X_{i-1,n}\{I(\varepsilon_{in} \le x + (t-\rho)X_{i-1,n}) - \Phi(x)\}.$$

As a consequence of the Cauchy-Schwarz inequality we find

$$\left(\int S_n(x,t)\varphi(x)\,dG(x)\right)^2 \le \int S_n^2(x,t)\,dG(x)\int \varphi^2(x)\,dG(x)$$

and so

$$M_n(t) \ge \frac{\left(\int S_n(x,t)\varphi(x)\,dG(x)\right)^2}{\int \varphi^2(x)\,dG(x)}, \qquad \forall t.$$

(b) Define

$$\begin{aligned} T_n(t) :=& \int S_n(x,t)\varphi(x)\,dG(x) \\ =& n^{-1/2}\sum X_{i-1,n}\int\{I(\varepsilon_i \le x + (t-\rho)X_{i-1,n}) - \Phi(x)\}\varphi(x)\,dG(x) \\ =:& n^{-1/2}\sum X_{i-1,n}Q(t). \end{aligned}$$

Now we show that $T_n(t)$ has the following properties:

i. nondecreasing function in t for every n,
ii. $\inf_t T_n(t) < 0$ and $\sup_t T_n(t) > 0$ for every n,
iii. $\sup_{t\in V_n(\rho,B)} |\hat{T}_n(t) - T_n(t)| = o_p(1)$ for every n.

i. We decompose $T_n(t) := T_{1,n}(t) + T_{2,n}(t)$, where

$$\begin{aligned} T_{1,n}(t) &:= n^{-1/2}\sum X_{i-1,n}I(X_{i-1,n} \ge 0)Q(t), \\ T_{2,n}(t) &:= n^{-1/2}\sum X_{i-1,n}I(X_{i-1,n} < 0)Q(t). \end{aligned}$$

Assume $t_1 \le t_2$.

- $X_{i-1,n} \ge 0$: Then $t_1X_{i-1,n} \le t_2X_{i-1,n}$. So $T_{1,n}(t_1) \le T_{1,n}(t_2)$.
- $X_{i-1,n} < 0$. We find

$$\begin{aligned} & t_1X_{i-1,n} \ge t_2X_{i-1,n} \\ &\implies I(\varepsilon_{in} \le x + (t_1-\rho)X_{i-1,n}) - \Phi(x) \\ &\qquad\qquad \ge I(\varepsilon_{in} \le x + (t_2-\rho)X_{i-1,n}) - \Phi(x) \\ &\implies X_{i-1,n}(I(\varepsilon_{in} \le x + (t_1-\rho)X_{i-1,n}) - \Phi(x)) \\ &\qquad\qquad \le X_{i-1,n}(I(\varepsilon_{in} \le x + (t_2-\rho)X_{i-1,n}) - \Phi(x)) \\ &\implies T_{2,n}(t_1) \le T_{2,n}(t_2). \end{aligned}$$

So $T_n(t)$ is nondecreasing in t.

ii. We decompose again $T_n(t) = T_{1,n}(t) + T_{2,n}(t)$ as in the part 2(b)i and see that $\lim_{t\to-\infty} T_n(t) < 0$ and $\lim_{t\to\infty} T_n(t) > 0$.

iii. For any $0 < B < \infty$ and with the Cauchy-Schwarz inequality and proof of Proposition 3.3.3 we get

$$\begin{aligned}\sup_{t\in V_n(\rho,B)} |\hat{T}_n(t) - T_n(t)| &= \sup_{t\in V_n(\rho,B)} \left| \int [A_n(x,t) + C_n(x,t)]\, \varphi(x)\, dG(x) \right| \\ &\leq 2 \sup_{t\in V_n(\rho,B)} \int A_n^2(x,t)\varphi^2(x)\, dG(x) \\ &\quad + 2 \sup_{t\in V_n(\rho,B)} \int C_n^2(x,t)\varphi^2(x)\, dG(x) \\ &= o_p(1).\end{aligned}$$

Since $T_n(t)$ is nondecreasing in t there exists a N''_η such that for all $n \geq N''_\eta$

$$\begin{aligned}&P\left(\frac{\inf_{t\in V_n^c(\rho,B)}\{T_n^2(t)\}}{\int \varphi^2\, dG} \geq z\right) \geq P\left(\frac{\inf_{t\in E}\{T_n^2(t)\}}{\int \varphi^2\, dG} \geq z\right) \\ &\geq P\left(\frac{\inf_{t\in E}\{\hat{T}_n^2(t)\}}{\int \varphi^2\, dG} \geq z\right) - \eta/2 \\ &\geq 1 - \eta, \qquad \forall n \geq \max\{N'_\eta, N''_\eta\}.\end{aligned}$$

The last step is a consequence of 1.

□

Corollary 3.3.8. *The estimators $\tilde{\rho}$ and $\hat{\rho}$ are satisfying*

$$\begin{aligned}|\sqrt{n}(\tilde{\rho} - \rho)| &\leq B < \infty, \\ |\sqrt{n}(\hat{\rho} - \rho)| &\leq B < \infty.\end{aligned}$$

Proof. The functions $M_n(t)$ and $\hat{M}_n(t)$ satisfy (3.10) and (3.11), see Proposition 3.3.6 and Proposition 3.3.7. Therefore the Corollary is a consequence of Lemma 3.2.17. □

Proposition 3.3.9. *With the assumptions of Proposition 3.3.3 the argmin $\hat{\rho}$ of $\hat{M}_n(t)$ is given by*

$$\sqrt{n}(\hat{\rho} - \rho) = -\frac{\int (W_n(x,\rho) + B_n(x))h_n(x)\, dG(x)}{Y_n \int h_n^2(x)\, dG(x)}. \tag{3.16}$$

Proof. Note that

$$\int (B_n(x) + W_n(x,\rho))^2 \, dG(x) = O_p(1)$$

and

$$\int (B_n(x) + W_n(x,\rho)) h_n(x) \, dG(x) = O_p(1),$$

which are consequences of Lemmata 3.2.6, 3.2.7 and the Cauchy-Schwarz inequality. Now consider (3.6). We get

$$\begin{aligned} \hat{M}_n(t) &= \int \left[B_n(x) + W_n(x,\rho) + Y_n\sqrt{n}(t-\rho)h_n(x)\right]^2 \, dG(x) \\ &= \int [B_n(x) + W_n(x,\rho)]^2 \, dG(x) \\ &\quad + 2Y_n\sqrt{n}(t-\rho) \int [B_n(x) + W_n(x,\rho)] \, h_n(x) \, dG(x) \\ &\quad + (Y_n\sqrt{n}(t-\rho))^2 \int h_n^2(x) \, dG(x) \end{aligned} \tag{3.17}$$

and differentiating with respect to t

$$\begin{aligned} \frac{d\hat{M}_n(t)}{dt} &= 2Y_n\sqrt{n} \int (B_n(x) + W_n(x,\rho))h_n(x) \, dG(x) \\ &\quad + 2Y_n^2 n(t-\rho) \int h_n^2(x) \, dG(x) \stackrel{!}{=} 0, \end{aligned}$$

so the minimum is

$$\sqrt{n}(\hat{\rho} - \rho) = -\frac{\int (B_n(x) + W_n(x,\rho))h_n(x) \, dG(x)}{Y_n \int h_n^2(x) \, dG(x)}.$$

□

Remark 3.3.10. The right side of (3.6) is

$$\hat{M}_n(t) = \hat{M}_n(\rho) + \hat{M}_n'(\rho)(t-\rho) + \hat{M}_n''(\rho)\frac{(t-\rho)^2}{2}.$$

3.4 Theorems

Theorem 3.4.1. *With $k := \lim_{n\to\infty} \sqrt{n} a_n^2$, $\psi(x) = \int_{-\infty}^{x} \varphi(y) \, dG(y)$ and the assumptions of Proposition 3.3.3, the expression $\sqrt{n}(\hat{\rho}-\rho)$ converges in distribution, i.e.*

$$\sqrt{n}(\hat{\rho} - \rho) \xrightarrow{d} N\left(-\frac{k\mu_0(1+\rho)\int (H_0(x) - \Phi(x))\varphi(x) \, dG(x)}{\int \varphi^2(x) \, dG(x)}, \frac{\sigma_\psi^2(1-\rho^2)}{\left(\int \varphi^2(x) \, dG(x)\right)^2}\right),$$

where

$$\lim_{n\to\infty} \operatorname{Var} \psi(\varepsilon_{in}) = \sigma_\psi^2 = \int \psi^2 \, d\Phi - \left(\int \psi \, d\Phi\right)^2 .$$

Proof. Since $a_n = O(n^{-1/4})$ we are sure that $k = O(1)$. We have shown in Lemma 3.2.16 that

$$\int W_n(x,\rho)\varphi(x)\, dG(x) = \mathcal{S}_n \to N\left(0, \frac{\sigma_\psi^2}{1-\rho^2}\right).$$

Consider the term (3.16)

$$\begin{aligned}
\sqrt{n}(\hat\rho - \rho) &= -\frac{\int W_n(x,\rho)h_n(x)\, dG(x)}{Y_n \int h_n^2(x)\, dG(x)} - \frac{\int B_n(x)h_n(x)\, dG(x)}{Y_n \int h_n^2(x)\, dG(x)} \\
&= -\frac{\int W_n(x,\rho)\varphi(x)\, dG(x) + O_p(a_n)}{Y_n \int \varphi^2(x)\, dG(x) + O(a_n^2)} - \frac{\int B_n(x)h_n(x)\, dG(x)}{Y_n \int \varphi^2(x)\, dG(x) + O(a_n^2)} \\
&= -\frac{\mathcal{S}_n + O_p(a_n)}{Y_n \int \varphi^2(x)\, dG(x) + O(a_n^2)} - \frac{b_n}{Y_n \int \varphi^2(x)\, dG(x) + O(a_n^2)} \\
&= -\frac{\mathcal{S}_n}{Y_n \int \varphi^2(x)\, dG(x) + O(a_n^2)} \\
&\quad - \frac{b_n}{Y_n \int \varphi^2(x)\, dG(x) + O(a_n^2)} + O_p(a_n).
\end{aligned}$$

Since $a_n = O(n^{-1/4})$, we have $Y_n \xrightarrow{P} \frac{1}{1-\rho^2}$ (Corollary 3.2.13), $b_n \xrightarrow{P} \frac{k\mu_0 \int (H_0 - \Phi)\varphi\, dG}{1-\rho}$ (Corollary 3.2.9) which yields

$$\begin{aligned}
Y_n \int \varphi^2(x)\, dG(x) + O(a_n^2) &\xrightarrow{P} \frac{\int \varphi^2(x)\, dG(x)}{1-\rho^2}, \\
\frac{b_n}{Y_n \int \varphi^2(x)\, dG(x) + O(a_n^2)} &\xrightarrow{P} \frac{k\mu_0(1+\rho)\int (H_0 - \Phi)\varphi\, dG}{\int \varphi^2(x)\, dG(x)}.
\end{aligned}$$

The conclusion follows with Proposition B.11, Proposition B.12 and Remark B.10. □

Corollary 3.4.2. *If $G(x) \equiv x$ then*

$$\sqrt{n}(\hat\rho - \rho) \xrightarrow{d} N\left(-2\sqrt{\pi}k\mu_0(1+\rho)\int (H_0(x) - \Phi(x))\varphi(x)\, dx, \frac{\pi}{3}(1-\rho^2)\right).$$

Proof. Note that with $\psi(x) = \Phi(x)$ we find

$$\int \psi^2 \, d\Phi = \frac{1}{3}, \qquad \int \varphi^2(x)\, dx = \frac{1}{2\sqrt{\pi}},$$

$$\int \psi(x)\varphi(x)\, dx = \frac{1}{2}, \qquad \sigma_\psi^2 = \frac{1}{12}.$$

Then the Corollary is a consequence of Theorem 3.4.1. □

Corollary 3.4.3. *If* $a_n = o(n^{-1/4})$ *then*

$$\sqrt{n}(\hat{\rho} - \rho) \xrightarrow{d} N\left(0, \frac{\sigma_\psi^2(1-\rho^2)}{\left(\int \varphi^2\, dG\right)^2}\right).$$

Proof. If $a_n = o(n^{-1/4})$ then $k = o(1)$. The rest is a consequence of Theorem 3.4.1.

Theorem 3.4.4. *With the assumptions of Proposition 3.3.3 it follows that*

$$\sqrt{n}(\hat{\rho} - \tilde{\rho}) = o_p(1).$$

Proof. With Propositions 3.3.5 and 3.3.3 we get

$$\begin{aligned} |\hat{M}_n(\tilde{\rho}) - \hat{M}_n(\hat{\rho})| &\le |\hat{M}_n(\tilde{\rho}) - M_n(\tilde{\rho}) + M_n(\tilde{\rho}) - \hat{M}_n(\hat{\rho})| \\ &\le |\hat{M}_n(\tilde{\rho}) - M_n(\tilde{\rho})| + |M_n(\tilde{\rho}) - \hat{M}_n(\hat{\rho})| \\ &\le \sup_{t \in V_n(\rho,B)} |\hat{M}_n(t) - M_n(t)| + |M_n(\tilde{\rho}) - \hat{M}_n(\hat{\rho})| \\ &= o_p(1). \end{aligned} \tag{3.18}$$

Since the last inequality holds with probability $1-\delta$, $\delta > 0$, we must choose $B = B(\delta)$ sufficiently large.

Equation (3.16) shows that

$$\sqrt{n}(\hat{\rho} - \rho)\, Y_n \int h_n^2(x)\, dG(x) = -\int [W_n(x,\rho) + B_n(x)] h_n(x)\, dG(x). \tag{3.19}$$

Remember that $Y_n = \frac{1}{n}\sum_{i=1}^n X_{i-1,n}^2$. Consider (3.17) and use (3.19) to write

$$\begin{aligned} \hat{M}_n(t) &= \hat{M}_n(\rho) + 2\sqrt{n}(t-\rho) Y_n \int [W_n(x,\rho) + B_n(x)] h_n(x)\, dG(x) \\ &\qquad + Y_n^2\, n(t-\rho)^2 \int h_n^2(x)\, dG(x) \\ &= \hat{M}_n(\rho) - 2n(t-\rho)(\hat{\rho}-\rho) Y_n^2 \int h_n^2(x)\, dG(x) \\ &\qquad + n(t-\rho)^2 Y_n^2 \int h_n^2(x)\, dG(x). \end{aligned} \tag{3.20}$$

Now we start with (3.18) and use (3.20) to get

$$\begin{aligned}
o_p(1) &= \left| \hat{M}_n(\tilde{\rho}) - \hat{M}_n(\hat{\rho}) \right| \\
&= \Big| 2nY_n^2 \int h_n^2(x)\, dG(x) \left[-(\tilde{\rho}-\rho)(\hat{\rho}-\rho) + (\hat{\rho}-\rho)^2 \right] \\
&\qquad + nY_n^2 \int h_n^2(x)\, dG(x) \left[(\tilde{\rho}-\rho)^2 - (\hat{\rho}-\rho)^2 \right] \Big| \\
&= \left| nY_n^2 \int h_n^2(x)\, dG(x) \left[-2(\tilde{\rho}-\rho)(\hat{\rho}-\rho) + (\hat{\rho}-\rho)^2 + (\tilde{\rho}-\rho)^2 \right] \right| \\
&= \left(\sqrt{n}(\hat{\rho}-\tilde{\rho}) \right)^2 Y_n^2 \int h_n^2(x)\, dG(x) \\
&= \left(\sqrt{n}(\hat{\rho}-\tilde{\rho}) \right)^2 O_p(1).
\end{aligned}$$

Since $\hat{\rho}$ is asymptotically normal distributed we conclude that $\tilde{\rho}$ is asymptotically normal distributed, too. □

4 MDE versus least square estimator

We want to compare the least square estimator with the minimum distance estimator.

4.1 Least-square estimator

Lemma 4.1.1. *The least square estimator in an AR(1)-process is*

$$\tilde{\rho}_{LS} = \tilde{\rho}_{LS}(n) = \frac{\sum_{i=1}^{n} X_{in} X_{i-1,n}}{\sum_{i=1}^{n} X_{i-1,n}^2}.$$

Theorem 4.1.2. *Let $\tilde{\rho}_{LS}$ be the least square estimator in an AR(1)-process $X_i = \rho X_{i-1} + \varepsilon_i$, where ε_i are standard normal distributed. Then*

$$\sqrt{n}(\tilde{\rho}_{LS} - \rho) \xrightarrow{d} N(0, 1 - \rho^2).$$

Proof. Brockwell and Davis (1993), p. 259 states the result, the proof can be found on p. 386-396.

Corollary 4.1.3. *Let $X_i = \rho X_{i-1} + \varepsilon_i$ be an $AR(1)$-process, where ε_i are standard normal distributed. Then*

$$\frac{1}{\sqrt{n}} \sum_{i=1}^{n} X_{i-1} \varepsilon_i \xrightarrow{d} N\left(0, \frac{1}{1-\rho^2}\right).$$

Proof. Start with

$$\begin{aligned} \frac{1}{\sqrt{n}} \sum_{i=1}^{n} X_{i-1} \varepsilon_i &= \frac{1}{\sqrt{n}} \left(\sum_{i=1}^{n} X_i X_{i-1} - \rho \sum_{i=1}^{n} X_{i-1}^2 \right) \\ &= \frac{1}{\sqrt{n}} \left(\frac{\sum_{i=1}^{n} X_i X_{i-1}}{\sum_{i=1}^{n} X_{i-1}^2} \sum_{i=1}^{n} X_{i-1}^2 - \rho \sum_{i=1}^{n} X_{i-1}^2 \right) \\ &= \sqrt{n} \left(\frac{\sum_{i=1}^{n} X_i X_{i-1}}{\sum_{i=1}^{n} X_{i-1}^2} - \rho \right) \frac{1}{n} \sum_{i=1}^{n} X_{i-1}^2. \end{aligned}$$

Now

$$\frac{1}{n} \sum_{i=1}^{n} X_{i-1}^2 \xrightarrow{P} \frac{1}{1-\rho^2}$$

and

$$\sqrt{n} \left(\frac{\sum_{i=1}^{n} X_i X_{i-1}}{\sum_{i=1}^{n} X_{i-1}^2} - \rho \right) \xrightarrow{d} N(0, 1 - \rho^2).$$

The claim is then a consequence of Proposition B.11 and Proposition B.12. □

Proposition 4.1.4. *Suppose $a_n = O(n^{-1/4})$. Let $\tilde{\rho}_{LS}$ be the least square estimator in an AR(1)-process $X_{in} = \rho X_{i-1,n} + \varepsilon_{in}$, where ε_{in} are contaminated normal distributed. Then with $k := \lim_{n\to\infty} \sqrt{n} a_n^2$ we find*

$$\sqrt{n}(\tilde{\rho}_{LS} - \rho) \xrightarrow{d} N\left(k\mu_0^2(1+\rho), 1-\rho^2\right).$$

Proof. Consider

$$\sqrt{n}(\tilde{\rho}_{LS} - \rho) = \frac{n^{-1/2}\sum_{i=1}^n X_{i-1,n}\varepsilon_{in}}{n^{-1}\sum_{i=1}^n X_{i-1,n}^2} =: \frac{N_n}{D_n}.$$

First we examine the denominator

$$D_n = \frac{1}{n}\sum_{i=1}^n X_{i-1,n}^2 = Y_n.$$

Corollary 3.2.13 shows that

$$D_n \xrightarrow{p} \frac{1}{1-\rho^2}.$$

The nominator is

$$\begin{aligned} N_n &= \frac{1}{\sqrt{n}}\sum_{i=1}^n X_{i-1,n}\varepsilon_{in} \\ &= \frac{1}{\sqrt{n}}\sum_{i=1}^n \left[\tilde{X}_{i-1} + \hat{X}_{i-1,n}\right]\left[(1-U_{in})\epsilon_i + U_{in}\delta_i\right] \\ &= \frac{1}{\sqrt{n}}\Big[\sum \tilde{X}_{i-1}\epsilon_i - \sum \tilde{X}_{i-1}U_{in}\epsilon_i + \sum \tilde{X}_{i-1}U_{in}\delta_i + \sum \hat{X}_{i-1,n}\epsilon_i \\ &\quad - \sum \hat{X}_{i-1,n}U_{in}\epsilon_i + \sum \hat{X}_{i-1,n}U_{in}(\delta_i - \mu_0) + \sum \hat{X}_{i-1,n}U_{in}\mu_0\Big] \end{aligned}$$

We have to analyse all these terms and start with the most important:

(a) With Corollary 4.1.3 we see that

$$\frac{1}{\sqrt{n}}\sum_{i=1}^n \tilde{X}_{i-1}\epsilon_i \xrightarrow{d} N\left(0, \frac{1}{1-\rho^2}\right).$$

(b) Corollary 2.2.9 proves that

$$\frac{1}{\sqrt{n}}\sum \hat{X}_{i-1,n}U_{in} \xrightarrow{p} \frac{k\mu_0}{1-\rho}.$$

(c) Consider

$$R := \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \hat{X}_{i-1,n} U_{in} (\delta_i - \mu_0).$$

It is easy to see that $\mathrm{E}R = 0$.

In order to use Chebychev's inequality we have to determine the variance of R. Since δ_j is independent of $\hat{X}_{j-1,n}$ and U_{jn} the second moment is

$$\begin{aligned} &\mathrm{E}\left(\frac{1}{\sqrt{n}} \sum_{i=1}^{n} \hat{X}_{i-1,n} U_{in} (\delta_i - \mu_0)\right)^2 \\ &= \frac{1}{n} \mathrm{E}\left(\sum_{i=1}^{n} \hat{X}^2_{i-1,n} U_{in} (\delta_i - \mu_0)^2 + 2 \sum_{i<j} \hat{X}_{i-1,n} \hat{X}_{j-1,n} U_{in} U_{jn} (\delta_i - \mu_0)(\delta_j - \mu_0)\right) \\ &= \mathrm{E}\hat{X}^2_{0n} \, \mathrm{E}U_{in} (\delta_i - \mu_0)^2 = O(a_n^2). \end{aligned}$$

Putting together the parts we find $R \xrightarrow{p} 0$.

(d) All other expressions are converging to 0 in probability, too. The proofs follow similar steps.

Putting together parts (a) - (d) we get

$$N_n \xrightarrow{d} N\left(\frac{k\mu_0^2}{1-\rho}, \frac{1}{1-\rho^2}\right)$$

and hence the claim. □

Note that from Theorem 3.4.1 we may write

$$\sqrt{n}(\hat{\rho} - \rho) \xrightarrow{d} N\left(-\frac{k\mu_0(1+\rho)\int (H_0(x) - \Phi(x))\varphi(x)\, dG(x)}{\int \varphi^2(x)\, dG(x)}, \frac{\sigma_\psi^2 (1-\rho^2)}{\left(\int \varphi^2(x)\, dG(x)\right)^2}\right).$$

4.2 Comparison

Proposition 4.2.1. *The asymptotic bias of the least square estimator $\sqrt{n}(\tilde{\rho}_{LS} - \rho)$ is greater than that of the minimum distance estimator $\sqrt{n}(\hat{\rho} - \rho)$, if*

$$|\mu_0| > \frac{\left|\int (H_0(x) - \Phi(x))\varphi(x)\, dG(x)\right|}{\int \varphi^2(x)\, dG(x)}.$$

Proof. It is a consequence of Theorem 3.4.1 and Proposition 4.1.4. □

Corollary 4.2.2. *Consider the case $G(x) \equiv x$. Then*

$$\int \Phi(x)\varphi(x)\,dx = \frac{1}{2},$$
$$\int \varphi^2(x)\,dx = \frac{1}{2\sqrt{\pi}}.$$

In this case the asymptotic bias of the least square is greater than that of the minimum distance estimator if

$$|\mu_0| > 2\sqrt{\pi}\left|\int H_0(x)\varphi(x)\,dx - 1/2\right|. \tag{4.1}$$

Lemma 4.2.3. *If $b \neq 0$, $e > 0$ then*

$$\int \varphi(a + bx)\Phi(c + ex)\,dx = \frac{1}{|b|}\Phi\left(\frac{c/e - a/b}{(1/e^2 + 1/b^2)^{1/2}}\right).$$

Proof. Cain (1994) p. 124f.

Lemma 4.2.4. *The taylor expansion of $\Phi(x)$ is*

$$\Phi(x) = \frac{1}{2} + \frac{1}{\sqrt{2\pi}}\sum_{n=0}^{\infty}\frac{(-1)^n x^{2n+1}}{n!2^n(2n+1)}.$$

For $|x| < \sqrt{\pi}$ the sequence

$$|f_n(x)| := \left|\frac{x^{2n+1}}{n!2^n(2n+1)}\right|$$

is monoton nonincreasing and $\operatorname{sgn} f_1(x) = \operatorname{sgn} x$.

Proof. Abramowitz and Stegun (1968), p. 932.

The second part follows immediateley from

$$|\,f_{n+1}(x)\,| = \left|\frac{x^{2(n+1)+1}}{(n+1)!2^{n+1}(2n+3)}\right| = \left|\frac{x^{2n+1}}{n!2^n(2n+1)}\,\frac{(2n+1)x^2}{2(n+1)(2n+3)}\right|$$
$$< |\,f_n(x)\,|\,\frac{x^2}{2(n+1)} < |\,f_n(x)\,|\,\frac{\pi}{2(n+1)} < |\,f_n(x)\,|, \qquad \forall n.$$

Corollary 4.2.5. *Let $H_0(x)$ be a normal distribution function with mean μ_0 and variance σ_0^2. Then the bias of the asymptotic distribution of the minimum distance estimator is less than that of the least square estimator if*

- $|\mu_0| > \sqrt{\pi}$ *or*

- $\sigma_0 \geq 1$ *and* $0 \neq |\mu_0| \leq \sqrt{\pi}$.

Proof. Let us start with

$$\int H_0(x)\varphi(x)\,dx = \int \Phi\left(\frac{x-\mu_0}{\sigma_0}\right)\varphi(x)\,dx = \Phi\left(\frac{-\mu_0}{\sqrt{1+\sigma_0^2}}\right), \tag{4.2}$$

which is a consequence of Lemma 4.2.3.

Now we consider the inequality (4.1) and use expression (4.2):

$$|\mu_0| > 2\sqrt{\pi}\left|\Phi\left(\frac{-\mu_0}{\sqrt{\sigma_0^2+1}}\right) - \frac{1}{2}\right|. \tag{4.3}$$

Since $|1/2 - \Phi(x)| \leq 1/2$ for all x, the right side of (4.3) is smaller or equal $\sqrt{\pi}$.

- Therefore if $|\mu_0| > \sqrt{\pi}$ the conclusion remains true.
- Let $0 < \mu_0 \leq \sqrt{\pi}$. Then we have to satisfy the condition

$$\mu_0 - \sqrt{\pi} + 2\sqrt{\pi}\Phi\left(\frac{-\mu_0}{\sqrt{\sigma_0^2+1}}\right) > 0.$$

As a consequence of Lemma 4.2.4 we get

$$\begin{aligned}\Phi\left(\frac{-\mu_0}{\sqrt{\sigma_0^2+1}}\right) &= \frac{1}{2} + \frac{1}{\sqrt{2\pi}}\frac{-\mu_0}{\sqrt{\sigma_0^2+1}} + \frac{1}{\sqrt{2\pi}}\sum_{n=1}^{\infty}\frac{(-1)^n\left(\frac{-\mu_0}{\sqrt{\sigma_0^2+1}}\right)^{2n+1}}{n!2^n(2n+1)} \\ &\geq \frac{1}{2} + \frac{1}{\sqrt{2\pi}}\frac{-\mu_0}{\sqrt{\sigma_0^2+1}} \geq \frac{1}{2} + \frac{-\mu_0}{2\sqrt{\pi}},\end{aligned}$$

and the case $0 < \mu_0 \leq \sqrt{\pi}$ is proved.

Consider the case $-\sqrt{\pi} \leq \mu_0 < 0$. Then (4.3) yields

$$-\mu_0 - \sqrt{\pi} + 2\sqrt{\pi}\Phi\left(\frac{-\mu_0}{\sqrt{\sigma_0^2+1}}\right) > -\mu_0 > 0.$$

□

5 Application

In this chapter we suppose that n is a fixed number!

5.1 General Case

Definition 5.1.1. *Let*

$$f^{\pm}(x) := \frac{1}{2}\left(\lim_{\eta\to 0+} f(t_j+\eta) + \lim_{\eta\to 0+} f(t_j-\eta)\right),$$
$$g_{i-1} := g(X_{i-1}).$$

Definition 5.1.2. *Assume that*

- $c_1 < \min_j\{X_j\} < \max_i\{X_i\} < c_2$,
- *g is a bounded function with finite number of jumps and*
- *$F(x)$ is any distribution function with*

$$\int F(x)(1-F(x))\,dx < \infty$$

and a continuous density function $f(x)$.

We define the function

$$M_n(t) := \int \left[n^{-1/2}\sum_{i=1}^{n} g_{i-1}\left\{I(X_i \le x + tX_{i-1}) - F(x)\right\}\right]^2 dx =: \int S_n^2(x,t)\,dx. \tag{5.1}$$

Lemma 5.1.3. *$M_n(t)$ exists for all t.*

Proof.

$$M_n(t) = \int_{-\infty}^{\min_i\{X_i - tX_{i-1}\}} \left[n^{-1/2}\sum g_{i-1}F(x)\right]^2 dx + \int_{\min_i\{X_i - tX_{i-1}\}}^{\max_i\{X_i - tX_{i-1}\}} S_n^2(x,t)\,dx$$
$$+ \int_{\max_i\{X_i - tX_{i-1}\}}^{\infty} \left[n^{-1/2}\sum g_{i-1}(1-F(x))\right]^2 dx. \tag{5.2}$$

Each part of this integral is bounded, so $M_n(t)$ exists.

Lemma 5.1.4. *The function $M_n(t)$ is continuous for all n.*

Proof. Let $f(t) := n^{-1/2}\sum_k g_{k-1}(I(X_k \leq x + tX_{k-1}) - F(x))$. We choose a sequence t_m converging to t_0. With $|f(t_m) + f(t_0)| \leq 2n^{-1/2}\sum |g_{i-1}| := d$ we get

$$\begin{aligned}
|M_n(t_m) - M_n(t_0)| &\leq \int |f^2(t_m) - f^2(t_0)|\, dx \\
&= \int |f(t_m) + f(t_0)|\, |f(t_m) - f(t_0)|\, dx \\
&\leq d \int |f(t_m) - f(t_0)|\, dx \\
&\leq dn^{-1/2} \int \sum_k \Big| g_{k-1}(I(X_k \leq x + t_m X_{k-1}) \\
&\qquad - I(X_k \leq x + t_0 X_{k-1})\Big|\, dx \\
&= dn^{-1/2} \sum_k \Big| g_{k-1}(t_m - t_0) X_{k-1} \Big| \\
&= dn^{-1/2} |t_m - t_0| \sum_k |g_{k-1} X_{k-1}| \to 0.
\end{aligned}$$

hence for $t_m \to t_0$ it follows that $|M_n(t_m) - M_n(t_0)| \to 0$.

Lemma 5.1.5. *The function $M_n(t)$ is piecewise differentiable. The nondifferentiable points are of the form $t = \frac{X_i - X_k}{X_{i-1} - X_{k-1}}$ for all i, k with $X_{i-1} \neq X_{k-1}$.*

Proof. For the proof we use Proposition A.1. Consider equation (5.2). The first and the third term are special cases of the second one. So let $H_n(t)$ be the parameter integral over the finite intervall $[a, b]$, where $a := \min_l\{X_l - tX_{l-1}\}$, $b := \max_l\{X_l - tX_{l-1}\}$.
Note that for any i, j we have

$$\min_l\{X_l - tX_{l-1}\} \leq \max\{X_i - tX_{i-1}, X_j - tX_{j-1}\} \leq \max_l\{X_l - tX_{l-1}\}.$$

If $X_{i-1} = X_{j-1}$ then $\max\{X_i - tX_{i-1}, X_j - tX_{j-1}\} = \max\{X_i, X_j\} - tX_{i-1}$ and is therefore differentiable for all t. So let $X_{i-1} \neq X_{j-1}$.

$$\begin{aligned}
H_n(t) = \int_a^b S_n^2(x,t)\, dx &= \frac{1}{n}\int_a^b \left(\sum_{i=1}^n g_{i-1}(I(X_i \leq x + tX_{i-1}) - F(x))\right)^2 dx \\
= \frac{1}{n}\int_a^b \sum_{i,j=1}^n & g_{i-1}g_{j-1}(I(X_i \leq x + tX_{i-1})I(X_j \leq x + tX_{j-1}) \\
&- 2I(X_i \leq x + tX_{i-1})F(x) + F^2(x))\, dx.
\end{aligned}$$

Omiting the sums and the function g, which are independent of t, we find three different kinds of terms:

Term 1:

$$\begin{aligned}
&\int_a^b I(X_i \le x + tX_{i-1}) I(X_j \le x + tX_{j-1})\, dx \\
=& \int_a^b I(\max\{X_i - tX_{i-1}, X_j - tX_{j-1}\} \le x)\, dx \\
=& \int_{\max\{X_i - tX_{i-1}, X_j - tX_{j-1}\}}^b dx \\
=& \max_l \{X_l - tX_{l-1}\} - \max\{X_i - tX_{i-1}, X_j - tX_{j-1}\}
\end{aligned}$$

is differentiable for all

$$t \notin \left\{ \frac{X_i - X_j}{X_{i-1} - X_{j-1}} \right\}.$$

Term 2:

$$\begin{aligned}
&\int_a^b F(x) I(X_i \le x + tX_{i-1})\, dx \\
=& \int_{a+tX_{i-1}}^{b+tX_{i-1}} F(y - tX_{i-1}) I(X_i \le y)\, dy \\
=& \int_{X_i}^{b+tX_{i-1}} F(y - tX_{i-1})\, dy
\end{aligned}$$

which is also differentiable for the same values of t as in term 1. For the last equation we used the fact, that for all i

$$\max\{\min_l \{X_l - tX_{l-1}\} + tX_{i-1}, X_i\} = X_i.$$

Term 3: analogue. □

Theorem 5.1.6. *Assume*

$$G_t(x) := \sum_i g_{i-1} \left[I(X_i - tX_{i-1} \le x) - F(x) \right] \in L^2(\mathbb{R}^n)$$

and

$$F'(x) = f(x) \in L^2(\mathbb{R}^n).$$

Then the function

$$M_n(t) = \int \left[n^{-1/2} \sum_{i=1}^n g_{i-1} \left\{ I(X_i \le x + tX_{i-1}) - F(x) \right\} \right]^2 dx$$

is piecewise differentiable for all

$$t \notin \left\{ \frac{X_i - X_k}{X_{i-1} - X_{k-1}}, \forall i, k \right\} \text{ and } X_{i-1} \neq X_{k-1}$$

and its derivative is

$$M_n'(t) = \frac{2}{n} \sum_{i,k} g_{k-1} g_{i-1} X_{k-1} (I^{\pm}(X_i - X_k \leq t(X_{i-1} - X_{k-1})) - F(X_k - tX_{k-1})),$$

where

$$I^{\pm}(x \leq y) := \begin{cases} 1 & x < y \\ \frac{1}{2} & x = y \\ 0 & x > y. \end{cases}$$

Proof. We assumed $G_t, f \in L^2(\mathbb{R})$. Hence $\hat{G}_t$ exists and $\hat{G}_t \in L^2(\mathbb{R}, \mathbb{C})$. We have

$$\hat{G}_t(\omega) = \int_{\mathbb{R}} e^{-2\pi i \omega x} G_t(x)\, dx = \sum g_{i-1} \left(\frac{1}{2\pi i \omega} e^{-2\pi i \omega (X_i - tX_{i-1})} + \hat{f}(\omega) \right),$$

taking the derivative

$$\frac{d}{dt} \hat{G}_t(\omega) = \sum g_{i-1} X_{i-1} e^{-2\pi i \omega (X_i - tX_{i-1})}.$$

By Plancherel (Proposition A.5) we get

$$M_n(t) = \int_{\mathbb{R}} |G_t(x)|^2\, dx = \int_{\mathbb{R}} |\hat{G}_t(\omega)|^2\, d\omega.$$

As $M_n(t)$ is differentiable for $t \neq \frac{X_i - X_k}{X_{i-1} - X_{k-1}}$ and with Proposition A.6 and Theorem A.7 we get

$$\begin{aligned} \frac{d}{dt} M_n(t) &= \frac{d}{dt} \int_{\mathbb{R}} \hat{G}_t(\omega) \overline{\hat{G}_t(\omega)}\, d\omega \\ &= \int_{\mathbb{R}} \hat{G}_t(\omega) \frac{d}{dt} \overline{\hat{G}_t(\omega)}\, d\omega \\ &= 2\Re \int_{\mathbb{R}} \hat{G}_t(\omega) \overline{\frac{d}{dt} \hat{G}_t(\omega)}\, d\omega \\ &= 2 \sum g_{i-1} X_{i-1}\ \Re \int_{\mathbb{R}} e^{2\pi i (X_i - tX_{i-1})\omega} \hat{G}_t(\omega)\, d\omega \\ &= 2 \sum g_{i-1} X_{i-1} G_t^{\pm}(X_i - tX_{i-1}), \end{aligned}$$

where we have set

$$G_t^{\pm}(x) = \frac{1}{2} \left(\lim_{\eta \to 0+} G_t(x + \eta) + \lim_{\eta \to 0+} G_t(x - \eta) \right).$$

Remark 5.1.7. There exist at most n^2 jump points.

Corollary 5.1.8. *The function $M_n(t)$ has its extremes either in the differentiable intervals where the condition*

$$\sum_{i,k} g_{k-1}g_{i-1}X_{k-1}\left(I^{\pm}(X_i - X_k \leq t(X_{i-1} - X_{k-1})) - F(X_k - tX_{k-1})\right) = 0 \quad (5.3)$$

must be satisfied or at the jump points, which are of the form

$$t = \frac{X_i - X_k}{X_{i-1} - X_{k-1}} \quad \forall i, k.$$

Proof. It's a consequence of Theorem 5.1.6.

Proposition 5.1.9. *Let* $\operatorname{sgn} g(x) = \operatorname{sgn} x$*. Then* $M_n(t)$ *has a minimum.*

Proof. Consider the derivative

$$M_n'(t) = \frac{2}{n}\sum_{i,k} g_{k-1}g_{i-1}X_{k-1}\left(I^{\pm}(X_i - X_k \leq t(X_{i-1} - X_{k-1})) - F(X_k - tX_{k-1})\right)$$

and define

$$U(i,k) := g_{k-1}g_{i-1}X_{k-1}\left(I^{\pm}(X_i - X_k \leq t(X_{i-1} - X_{k-1})) - F(X_k - tX_{k-1})\right).$$

We have to check 6 cases:

$X_{i-1} - X_{k-1} > 0$:	a)	$X_{i-1} > X_{k-1} > 0$,
	b)	$X_{k-1} < 0$ and $X_{i-1} > 0$,
	c)	$X_{k-1} < X_{i-1} < 0$.
$X_{i-1} - X_{k-1} < 0$:	d)	$X_{k-1} > X_{i-1} > 0$,
	e)	$X_{k-1} > 0$ and $X_{i-1} < 0$,
	f)	$X_{k-1} < X_{k-1} < 0$.

In cases b), c), d), e) we find $\lim_{t\to\pm\infty} U(i,k) = 0$, a) yields

$$\lim_{t\to\pm\infty} U(i,k) = \pm g_{k-1}g_{i-1}X_{k-1} \gtrless 0$$

and f)

$$\lim_{t\to\pm\infty} U(i,k) = \mp g_{k-1}g_{i-1}X_{k-1} \gtrless 0.$$

Combining these cases we get the result

$$\lim_{t\to\infty} M_n'(t) = \frac{2}{n}\sum_{A} g_{k-1}X_{k-1}|g_{i-1}| > 0,$$

where $A = \{(i,j);\ \operatorname{sgn} X_{k-1}\ \operatorname{sgn}(X_{i-1} - X_{k-1}) = 1\}$, and analogous for

$$\lim_{t\to-\infty} M_n'(t) = -\frac{2}{n}\sum_A g_{k-1}X_{k-1}|g_{i-1}| < 0.$$

Since $M_n(t)$ is continuous we conclude that $\lim_{t\to\pm\infty} M_n(t) = \infty$. So $M_n(t)$ must have a minimum. □

Remark 5.1.10. The special case $g(x) = x$ is used for the simulation.

5.2 Logistic distribution

In this chapter we choose as a special case the logistic distribution, e.g

$$F(x) = \left(1 + \exp\left(-\frac{x}{b}\right)\right)^{-1},$$

because $\int F$ and $\int F^2$ have primitives. We find the expression

$$M_n(t) = n^{-1}\sum_{i,j} g_{i-1}g_{j-1}\left[-b - \max\{X_i - tX_{i-1}, X_j - tX_{j-1}\} + \right.$$
$$\left. +2b\ln(1 + \exp\{(X_i - tX_{i-1})/b\})\right].$$

We have $b = \sqrt{3}\sigma/\pi$ and $\mu = 0$. With the logistic distribution it is possible to find easily the minimum of the function $M_n(t)$ numerically.

A Fourier Analysis

Proposition A.1. *Let* $F(t) = \int_{\psi_1(t)}^{\psi_2(t)} f(t,x)\,dx$. *If* ψ_1 *and* ψ_2 *have continuous derivatives with respect to* t *throughout the interval* $a \leq t \leq b$ *and* $f(t,x)$ *is continuously differentiable in a region wholly enclosing the region* $R = \{(t,x); a \leq t \leq b, c \leq x \leq d\}$ *then we get*

$$F'(t) = \int_{\psi_1(t)}^{\psi_2(t)} f_t(t,x)\,dx - \psi_1'(t)f(t,\psi_1(t)) + \psi_2'(t)f(t,\psi_2(t)).$$

Proof. Courant (1945), p. 220.

Definition A.2 (Rudin (1987) p. 65). *Let* X *be a measure space with a positive measure* μ. *If* $0 < p < \infty$ *and if* f *is a complex measurable function on* X, *define*

$$\| f \|_p = \left[\int_X |f|^p \,d\mu \right]^{1/p}$$

and let $L^p(\mu)$ *consist of all* f *for which*

$$\| f \|_p < \infty.$$

We call $\| f \|_p$ *the* L^p*-norm of* f.

Definition A.3. *If* $f \in L^1(\mathbb{R}^n)$, *the Fourier transform of* f *is the function* $\hat{f}$ *defined by letting*

$$\hat{f}(x) = \int_{\mathbb{R}^n} f(t) e^{-2\pi i <x,t>}\,dt,$$

for all $x \in \mathbb{R}^n$.

Proposition A.4. *If both* f *and* $\hat{f}$ *are integrable then*

$$f(x) = \int_{\mathbb{R}^n} \hat{f}(t) e^{2\pi i <x,t>}\,dt$$

for almost every x.

Proof. Stein and Weiss (1971), p. 11.

Proposition A.5 (Plancherel Theorem). *If* $f \in L^1 \cap L^2$ *then*

$$\int_{\mathbb{R}^n} \hat{f}^2(\omega)\,d\omega = \int_{\mathbb{R}^n} f^2(x)\,dx.$$

Proof. Stein and Weiss (1971), p. 16.

Proposition A.6 (Parseval Formula). *If $f, g \in L^2$ then*

$$\int_{\mathbb{R}^n} \overline{\hat{f}(\omega)}\hat{g}(\omega)\, d\omega = \int_{\mathbb{R}^n} \overline{f(x)} g(x)\, dx.$$

Proof. Rudin (1987), p. 187.

Theorem A.7. *Suppose $f : \mathbb{R} \to \mathbb{C}$ is continuous except at a finite number of points $t_1, t_2 \ldots t_n$, where $\lim_{\eta \to 0+} f(t_j + \eta)$ and $\lim_{\eta \to 0+} f(t_j - \eta)$ exist. Then, if $\int_{\mathbb{R}} |f(t)|\, dt$ converges and $\hat{f}(\xi) = O(|\xi|^{-1})$ as $|\xi| \to \infty$, it follows that for $R \to \infty$*

$$\int_{-R}^{R} \hat{f}(\xi) e^{2\pi i \xi t}\, d\xi \to f(t) \qquad \text{for } t \notin \{t_1, t_2, \ldots t_n\},$$
$$\int_{-R}^{R} \hat{f}(\xi) e^{2\pi i \xi t_j}\, d\xi \to \frac{1}{2}\left(\lim_{\eta \to 0+} f(t_j + \eta) + \lim_{\eta \to 0+} f(t_j - \eta) \right).$$

Proof. Körner (1988), p. 300f.

B Probability

Definition B.1 (Agresti (1990), p. 418). *The symbol $o_p(z_n)$ denotes a random variable of smaller order than z_n for large n, in the sense that $o_p(z_n)/z_n$ converges in probability to 0; that is, for any fixed $\varepsilon > 0$*

$$P\left(\left|\frac{o_p(z_n)}{z_n}\right| \leq \varepsilon\right) \to 1, \qquad \text{as } n \to \infty.$$

The notation $O_p(z_n)$ represents a random variable such that for every $\varepsilon > 0$ there is a constant K and an integer n_0 such that

$$P\left(\left|\frac{O_p(z_n)}{z_n}\right| \leq K\right) \geq 1 - \varepsilon, \qquad \text{for all } n > n_0.$$

Definition B.2. *We say Z_n converges to Z in distribution ($Z_n \xrightarrow{d} Z$), if*

$$\lim_{n\to\infty} F_{Z_n}(t) = F_Z(t)$$

at each point t, where $F_Z(\cdot)$ is continuous.

Definition B.3. *We say Z_n converges to Z in probability ($Z_n \xrightarrow{p} Z$), if for every $\varepsilon > 0$*

$$\lim_{n\to\infty} P(|Z_n - Z| > \varepsilon) = 0.$$

Definition B.4. *We say Z_n converges almost surely (a.s.) to Z, iff*

$$P(\omega : Z_n(\omega) \to Z(\omega)) = 1.$$

Proposition B.5. *We have*

$$Z_n \xrightarrow{a.s.} Z \Longrightarrow Z_n \xrightarrow{p} Z \Longrightarrow Z_n \xrightarrow{d} Z_n.$$

Proof. Rohatgi (1993), p. 246 and p. 250

Proposition B.6 (Pointwise Ergodic Theorem). *Let $\{X_n\}_{n\geq 1}$ be a stationary ergodic sequence of random variables with $\mathrm{E}|X_n| < \infty$. Then*

$$\lim_{n\to\infty} \frac{1}{n}\sum_{i=1}^{n} X_i = \mathrm{E}X, \quad a.s.$$

Proof. Berger (1993), p. 140.

Proposition B.7. *Let $h(x)$ be a nonnegative Borel-measurable function of an r.v. X. If $\mathrm{E}h(X)$ exists, then, for every $\varepsilon > 0$,*

$$P(h(X) \geq \varepsilon) \leq \frac{\mathrm{E}h(X)}{\varepsilon}.$$

Proof. Rohatgi (1993), p. 100.

Corollary B.8 (Markov's inequality). *Let $r > 0$ and $K > 0$. Then*

$$P(|X| \geq K) \leq \frac{\mathrm{E}|X|^r}{K^r}.$$

Corollary B.9 (Chebychev's inequality). *Let* $\mathrm{E}X = \mu$ *and* $\mathrm{Var}\ X = \sigma^2$. *Then*

$$P(|X - \mu| \geq K\sigma) \leq \frac{1}{K^2}.$$

Remark B.10. If $Z \sim N(\mu, \sigma^2)$ then $\dfrac{Z}{c} \sim N\left(\dfrac{\mu}{c}, \dfrac{\sigma^2}{c^2}\right)$.

Proposition B.11 (Slutsky's Theorem). *If $X_n \xrightarrow{p} a$, $Y_n \xrightarrow{p} b$ and $g(\cdot)$ is a continuous function, then*

1. $X_n \pm Y_n \xrightarrow{p} a \pm b$,
2. $X_n Y_n \xrightarrow{p} ab$,
3. $X_n / Y_n \xrightarrow{p} a/b$, *provided* $b \neq 0$,
4. $g(X_n) \xrightarrow{p} g(a)$.

Proof. Rohatgi (1993), p. 244f.

Proposition B.12. *1. If $X_n \xrightarrow{d} X$ and $X_n - Y_n \xrightarrow{p} 0$, then $Y_n \xrightarrow{d} X$,*

2. *if $X_n - Y_n \xrightarrow{p} 0$, and $g(\cdot)$ is a continuous function, then $g(X_n) - g(Y_n) \xrightarrow{p} 0$,*
3. *if $X_n \xrightarrow{d} X$ and $Y_n \xrightarrow{p} c$, then $X_n \pm Y_n \xrightarrow{d} X \pm c$,*
4. *if $X_n \xrightarrow{d} X$ and $Y_n \xrightarrow{p} c$, then $X_n Y_n \xrightarrow{d} cX$, if $c \neq 0$,*
5. *if $X_n \xrightarrow{d} X$ and $Y_n \xrightarrow{p} c$, then $X_n / Y_n \xrightarrow{d} X/c$, if $c \neq 0$.*

Proof. Chow and Teicher (1978), p. 249f.

Definition B.13 (Kallenberg (1997), p. 44). *A family of random variables ξ_t, $t \in T$, is said to be uniformly integrable if*

$$\lim_{r \to \infty} \sup_{t \in T} \mathrm{E}[\,|\,\xi_t\,|;\ |\,\xi_t\,| > r] = 0.$$

Lemma B.14. *The random variables ξ_t, $t \in T$, are uniformly integrable iff* $\sup_t \mathrm{E}\,|\xi_t\,| < \infty$ *and*

$$\lim_{PA \to 0} \sum_{t \in T} \mathrm{E}[\,|\xi_t\,|; A] = 0.$$

Proof. Kallenberg (1997), p. 44.

Proposition B.15. *If $X_1, X_2, \ldots$ are i.i.d. r.v.'s with common law $\mathcal{L}(X)$ and $\mathrm{E}|X|^q \le \infty$ for some positive integer q, then*

$$\frac{\sum_{i=1}^n X_i^k}{n} \xrightarrow{p} \mathrm{E}X^k \qquad \text{for } 1 \le k \le q,$$

and $n^{-1}\sum_i^n X_i^k$ is consistent for $\mathrm{E}X^k$, $1 \le k \le q$.

Proof. Rohatgi (1993), p. 336.

Definition B.16. *Let $(\Omega, \mathcal{F}, P)$ be a probability space, $\{\mathcal{F}_{in}, 1 \le i \le n\}$ be an array of sub σ-fields such that $\mathcal{F}_{in} \subset \mathcal{F}_{i+1,n}, 1 \le i \le n-1$; X_{in} be a $\mathcal{F}_{in}$ measurable r.v. with $\mathrm{E}X_{in}^2 < \infty$, $\mathrm{E}(X_{in}|\mathcal{F}_{i-1,n}) = 0$ $(2 \le i \le n)$; and let $S_{jn} = \sum_{i \le j} X_{in}, 1 \le j \le n$. Then $\{S_{in}, \mathcal{F}_{in}; 1 \le i \le n, n \ge 1\}$ is called a zero-mean, square-integrable martingale array with differences $\{X_{in}; 1 \le i \le n, n \ge 1\}$.*

Theorem B.17. *Let $\{S_{in}, \mathcal{F}_{in}; 1 \le i \le k_n, n \ge 1\}$ be a zero-mean, square-integrable martingale array with differences X_{in}, and let η^2 be an a.s. finite r.v. Suppose that*

$$\max|X_{in}| \xrightarrow{p} 0, \tag{B.1}$$

$$\sum_i X_{in}^2 \xrightarrow{p} \eta^2, \tag{B.2}$$

$$\mathrm{E}\left(\max_i X_{in}^2\right) \qquad \text{is bounded in } n, \tag{B.3}$$

$$\text{the } \sigma\text{-fields are nested: } \mathcal{F}_{in} \subset \mathcal{F}_{i+1,n} \text{ for } 1 \le i \le k_n, n \ge 1. \tag{B.4}$$

Then $S_{k_n n} = \sum_i X_{in} \xrightarrow{d} Z$, where the r.v. Z has characteristic function $\mathrm{E}\exp(-\frac{1}{2}\eta^2 t^2)$.

Proof. Hall and Heyde (1980), p. 60.

Corollary B.18. *If (B.1) and (B.3) are replaced by the conditional Lindeberg condition*

$$\text{for all } \varepsilon > 0, \quad \sum_i \mathrm{E}[X_{in}^2 I(|X_{in}| > \varepsilon)|\mathcal{F}_{i-1,n}] \xrightarrow{p} 0, \tag{B.5}$$

and if (B.2) is replaced by an analogous condition on the conditional variance

$$V_{k_n n}^2 = \sum \mathrm{E}(X_{in}^2|\mathcal{F}_{i-1,n}) \xrightarrow{p} \eta^2, \tag{B.6}$$

and if (B.4) holds, then the conclusion of Theorem B.17 remains true.

Proof. Hall and Heyde (1980), p. 63.

Proposition B.19 (Hall and Heyde (1980), p. 46). *Conditional version of Chebyshev's inequality:*

$$P\left(|X_{in}| > \varepsilon \middle| \mathcal{F}_{i-1,n}\right) \leq \varepsilon^{-2}\,\mathrm{E}\left[X_{in}^2 I\left(|X_{in}| > \varepsilon\right) \middle| \mathcal{F}_{i-1,n}\right]. \tag{B.7}$$

Remark B.20. Two approximations to the normal distribution, Abramowitz and Stegun (1968):

$$\Phi(x) \geq 1 - \frac{1}{x}(2\pi)^{-1/2}\exp(-x^2/2) =: p_1(x), \qquad x > \frac{11}{5}, \tag{B.8}$$

$$\Phi(x) \leq 1 - \frac{(4+x^2)^{1/2} - x}{2}(2\pi)^{-1/2}\exp(-x^2/2) =: p_2(x), \qquad x > \frac{7}{5}. \tag{B.9}$$

References

Abramowitz, M. and I. A. Stegun (1968). *Handbook of mathematical functions.* New York: Dover Publications, Inc.

Agresti, A. (1990). *Categorical Data Analysis.* Probability and its Applications. New York: John Wiley & Sons.

Anderson, T. W. (1971). *The Statistical Analysis of Time Series.* New York: John Wiley & Sons.

Beran, R. (1984). Minimum distance procedures. In P. R. Krishnaiah and P. K. Sen (Eds.), *Handbook of Statistics*, Volume 4, pp. 741–754. Elsevier Science Publisher.

Berger, M. A. (1993). *An Introduction to Probability and Stochastic Processes.* Springer Texts in Statistics. New York: Springer Verlag.

Breiman, L. (1992). *Probability.* Philadelphia: Siam.

Brockwell, P. J. and R. A. Davis (1993). *Time Series: Theory and Methods.* Springer Series in Statistics. New York: Springer Verlag.

Cain, M. (1994). The moment-generating function of the minimum of bivariate normal random variables. *The American Statistician 48*(No. 2), 124–125.

Chow, Y. S. and H. Teicher (1978). *Probability Theory (Independence, Interchangeability, Martingale).* New York: Springer Verlag.

Courant, R. (1945). *Differential and Integral Calculus*, Volume 2. London: Blackie and Son Limited.

Hall, P. and C. C. Heyde (1980). *Martingale limit theory and its applications.* New York: Academic Press.

Kallenberg, O. (1997). *Foundations of Modern Probability.* Probability and its Applications. New York: Springer Verlag.

Körner, T. W. (1988). *Fourier Analysis.* Cambridge: Cambridge University Press.

Koul, H. L. (1986). Minimum distance estimation and goodness-of-fit tests in first-order autoregression. *The Annals of Statistics 14*(No. 3), 1194–1213.

Koul, H. L. (1992). *Weighted Empiricals and Linear Models.* Lecture Notes – Monograph Series. Hayward: Institute of Mathematical Statistics.

Koul, H. L. and T. DeWet (1983). Minimum distance estimation in a linear regression model. *The Annals of Statistics 11*(No. 3), 921–932.

Loève, M. (1978). *Probability Theory II.* Graduate Texts in Mathematics. New York: Springer Verlag.

Millar, P. W. (1981). Robust estimation via minimum distance methods. *Zeitschrift für Wahrscheinlichkeit und verwandte Gebiete 55*, 73–89.

Parr, W. C. (1981). Minimum distance estimation: A bibliography. *Communication in Statistics Theory & Methods 12*(No. A10), 1205–1224.

Parr, W. C. and W. R. Schucany (1980). Minimum distance and robust estimation. *Journal of the American Statistical Association 75*, 616–624.

Priestley, M. B. (1994). *Spectral analysis and time series.* Probability and mathematical statistics. London: Academic Press.

Rohatgi, V. K. (1993). *Probability Theory and Mathematical Statistics.* New Delhi: Wiley Eastern Limited.

Rudin, W. (1987). *Real and Complex Analysis.* Mathematics Series. New York: McGraw-Hill International Editions.

Stein, E. M. and G. Weiss (1971). *Fourier Analysis on Euclidean Spaces.* Princeton, New Jersey: Princeton University Press.

Wolfowitz, J. (1957). Minimum distance method. *Annals of Mathematical Statistics 28*, 75–88.

Lebenslauf

Als Sohn der Elisabeth und des Serge Piot-Tagmann wurde ich am 29. Dezember 1969 in Bern geboren. In Zollikofen besuchte ich Primar- und Sekundarschule, bevor ich 1984 ins Gymnasium Bern-Neufeld wechselte, wo ich 1988 die Maturität Typus C erlangte. Im Herbst 1989 begann ich das Studium der Mathematik an der Universität Bern, welches ich im April 1994 mit dem Nebenfach Geographie und Ergänzungsfach BWL abschloss. Seit 1995 bin ich als Assistent am Institut für mathematische Statistik und Versicherungslehre tätig.