A note on the Bickel-Rosenblatt test in autoregressive time series

Dirk Bachmann
Ruhr-Universität Bochum
Fakultät für Mathematik
44780 Bochum, Germany
e-mail: dirk.bachmann@ruhr-uni-bochum.de

Holger Dette
Ruhr-Universität Bochum
Fakultät für Mathematik
44780 Bochum, Germany
e-mail: holger.dette@ruhr-uni-bochum.de
FAX: +49 2 34 32 14 559

February 23, 2004

Abstract

In a recent paper Lee and Na (2001) introduced a test for a parametric form of the distribution of the innovations in autoregressive models, which is based on the integrated squared error of the nonparametric density estimate from the residuals and a smoothed version of the parametric fit of the density. They derived the asymptotic distribution under the null-hypothesis, which is the same as for the classical Bickel-Rosenblatt (1973) test for the distribution of i.i.d. observations. In this note we first extend the results of Bickel and Rosenblatt to the case of fixed alternatives, for which asymptotic normality is still true but with a different rate of convergence. As a by-product we also provide an alternative proof of the Bickel and Rosenblatt result under substantially weaker assumptions on the kernel density estimate. As a further application we derive the asymptotic behaviour of Lee and Na's statistic in autoregressive models under fixed alternatives. The results can be used for the calculation of the probability of the type II error of the Bickel-Rosenblatt test for the parametric form of the error distribution and for testing interval hypotheses in this context.

AMS Subject classification: 62M07

Keywords and phrases: Autoregressive process, goodness-of-fit test, nonparametric density estimation, asymptotic distribution under fixed alternatives

1 Introduction

The goodness-of-fit testing problem for the distribution of the innovations is of particular importance in time series analysis. In particular the hypothesis of Gaussian errors is of interest, because under this additional assumption inference simplifies substantially, and many statistical procedures in time series are based on the assumption of normality [see e.g. Brockwell and Davis (1991) or Fan and Yao (2002)]. In a recent paper Lee and Na (2002) considered the problem of testing the hypothesis

$$H_0 : f = f_0 \quad H_1 : f \neq f_0 \tag{1.1}$$

in the first-order-autoregressive process

$$X_j = \varphi X_{j-1} + Z_j, \tag{1.2}$$

where f_0 is a given density, Z_j are i.i.d. random variables with density f, mean 0 and variance $\sigma^2 > 0$. Their work was motivated by the fact that the limit distribution of tests based on functionals of the empirical process of the residuals $\hat{Z}_j = X_j - \hat{\varphi} X_{j-1}$ depends on the parameter estimates involved in the empirical process and is no longer a functional of the standard Brownian bridge [see e.g. Boldin (1982) or Koul and Levanthal (1989)]. Lee and Na (2002) proposed to use the Bickel-Rosenblatt test based on the residuals $\hat{Z}_1, \dots, \hat{Z}_n$ for the hypotheses (1.1) and proved asymptotic normality of the corresponding test statistic under the null hypothesis $H_0 : f = f_0$.
It is the purpose of the present paper to provide a more refined analysis of the Bickel-Rosenblatt test by a discussion of the asymptotic behaviour of the test statistic under fixed alternatives of the form

$$d(f, f_0) = \int (f - f_0)^2(x)dx > 0. \tag{1.3}$$

In Section 2 we show that under the alternative (1.3) a standardized version of the statistic of Bickel and Rosenblatt (1973) based on i.i.d. observations is still asymptotically normal distributed but with a different rate of convergence. This result allows a simple calculation of the probability of the type II error of the Bickel-Rosenblatt test, which is of particular importance if the null hypothesis cannot be rejected [see Berger and Delampady (1987) or Sellke, Bayarri and Berger (2001)]. The asymptotic distribution of the test statistic under fixed alternatives can also be used for the calculation of critical values in the problem of testing precise hypotheses of the form

$$H_0 : d(f, f_0) > \pi \quad H_1 : d(f, f_0) \leq \pi, \tag{1.4}$$

where π is a given bound in which the experimenter would denote deviations from the assumed density f_0 as not relevant. Note that the formulation of the hypotheses (1.4) allows the experimenter to test that the density f is approximately equal to f_0 (i.e. $d(f, f_0) \leq \pi$) at a controlled type I error.
In Section 3 we consider the statistic of Lee and Na (2002) under the alternative (1.3) and show that it has the same asymptotic behaviour as Bickel and Rosenblatt's statistic in the i.i.d. case which was derived in Section 2. It is also demonstrated that this result holds for composite hypothesis

$$H_0 : f \in \mathcal{F} \quad H_1 : f \notin \mathcal{F} \tag{1.5}$$

where

$$\mathcal{F} = \Big\{ \frac{1}{\sigma} f_0\Big(\frac{\cdot - \mu}{\sigma}\Big) \mid \mu \in \mathbb{R}; \sigma > 0 \Big\} \tag{1.6}$$

is a local scale family and f_0 is a given density. Finally, some of the proofs are given in an appendix in Section 4.

2 The test of Bickel and Rosenblatt revisited

Let $Z_1, Z_2, \ldots, Z_n$ denote independent identically distributed random variables with two times continuously differentiable density f with bounded second derivative and $K : \mathbb{R} \to \mathbb{R}$ be a continuous bounded symmetric kernel with compact support satisfying

$$\int K(x)dx = 1; \quad \int x^2 K(x)dx < \infty; \quad \int K^2(x)dx < \infty. \tag{2.1}$$

We consider the kernel estimator

$$f_n(x) = \frac{1}{n} \sum_{i=1}^{n} K_h(x - Z_i) \tag{2.2}$$

where $K_h(\cdot) = \frac{1}{h} K(\frac{\cdot}{h})$ is the scaled kernel and $h > 0$ denotes a bandwidth satisfying

$$nh^2 \to \infty, h \to 0 \tag{2.3}$$

if $n \to \infty$. For the problem of testing the hypothesis (1.1) Bickel and Rosenblatt (1973) proposed to reject the null-hypothesis for large values of the statistic

$$T_n = \int [f_n - K_h * f_0]^2(x)dx \tag{2.4}$$

where $f_1 * f_2$ denotes the convolution of the functions f_1 and f_2. Under the null hypothesis these authors showed asymptotic normality of T_n, namely

$$n\sqrt{h} \Big\{ T_n - \frac{1}{nh} \int K^2(t)dt \Big\} \xrightarrow{\mathcal{D}} \mathcal{N}(0, \tau^2), \tag{2.5}$$

where the asymptotic variance is given by

$$\tau^2 = 2 \int f_0^2(x)dx \int (K * K)^2(x)dx. \tag{2.6}$$

The following result now establishes asymptotic normality of an appropriately standardized version of T_n under fixed alternatives.

Theorem 2.1. *If the assumptions (2.1) - (2.3) are satisfied and the alternative (1.3) is valid we have*

$$\sqrt{n} \Big[T_n - \int (K_h * (f - f_0))^2(x)dx \Big] \xrightarrow{\mathcal{D}} \mathcal{N}(0, 4\varrho^2), \tag{2.7}$$

where the asymptotic variance is given by

$$\varrho^2 = \mathrm{Var}[(f - f_0)(Z_i)]. \tag{2.8}$$

In the Appendix we provide an alternative proof of the statement (2.5) based on a central limit theorem for degenerate U-statistics, which is of its own interest and particulary helpful to identify the limit distribution in the proof of the second part of Theorem 2.1. Moreover, with this technique the statement (2.5) can be proved under substantially weaker assumptions than imposed by Bickel and Rosenblatt (1973), who derived this result using an approximation of the normalized and centered sample distribution function by an appropriate Brownian process on a convenient probability space. It is also interesting to note that the centered version of T_n is of different order under the null hypothesis and alternative, namely

$$\begin{aligned} T_n - E[T_n] &\overset{H_0}{\sim} O_p\Big(\frac{1}{n\sqrt{h}}\Big) \\ T_n - E[T_n] &\overset{H_1}{\sim} O_p\Big(\frac{1}{\sqrt{n}}\Big) \ . \end{aligned} \tag{2.9}$$

Note that Theorem 2.1 can be used for the calculation of the probability of the type II error of the test, which rejects the null hypothesis $H_0 : f = f_0$, whenever

$$n\sqrt{h}\Big\{T_n - \frac{1}{nh}\int K^2(t)dt\Big\} > \tau u_{1-\alpha}, \tag{2.10}$$

where $u_{1-\alpha}$ is the $(1-\alpha)$ quantile of the standard normal distribution. A straightforward calculation gives under the alternative (1.3) for the probability of rejection the approximation

$$P(\text{“rejection”}) \approx \Phi\Big(\frac{\sqrt{n}}{2\varrho}d(f, f_0) - \frac{\tau}{2\varrho}\frac{u_{1-\alpha}}{\sqrt{nh}}\Big) \approx \Phi\Big(\frac{\sqrt{n}}{2\varrho}d(f, f_0)\Big).$$

A further application of Theorem 2.1 consists in the calculation of critical values of the test for the precise hypotheses defined in (1.4). Here the null hypothesis is rejected for small values of the statistic T_n, namely

$$\sqrt{n}\frac{T_n - \pi}{2\hat{\varrho}} \leq u_\alpha, \tag{2.11}$$

where $\hat{\varrho}$ is an appropriate estimator of the asymptotic variance ϱ in Theorem 2.1. Note that the test of the form (2.11) decides in favor of the alternative $H_1 : d(f, f_0) \leq \pi$ at a controlled type I error of size α. In other words if we decide that the “true” density is approximately equal to f_0, the probability of a possible error is α. We finally note that it is important to control this probability because subsequent data analysis will be performed under the assumption $f = f_0$ if the null hypothesis in (1.4) is rejected.

3 A goodness-of-fit test in autoregressive models

Consider the first order autoregressive model, where we are interested in testing the hypothesis (1.1) for the distribution of the innovations Z_i. Because these values are unobservable, we replace them by the residuals $\hat{Z}_i = X_i - \hat{\varphi} X_{i-1}$, where $\hat{\varphi}$ is a $\sqrt{n}$-consistent estimator of the parameter φ. Let

$$\hat{f}_n(x) = \frac{1}{n} \sum_{i=1}^{n} K_h(x - \hat{Z}_i) \tag{3.1}$$

denote the kernel density estimate based on the residuals $\hat{Z}_1, \ldots, \hat{Z}_n$ and define the statistic $\hat{T}_n$ as the analogue of T_n, where the random variable f_n defined in (2.2) is replaced by $\hat{f}_n$. Lee and Na (2002) showed that under the additional assumptions on the kernel K

$$K''' \text{exists}, \ \ K'' \text{is bounded} \tag{3.2}$$

$$\int |K^{(j)}(x)| dx < \infty, j = 1, 2, 3, \int |K^{(j)}(x)|^2 dx < \infty, j = 1, 2 \tag{3.3}$$

and on the bandwidth

$$nh^4 \to \infty \tag{3.4}$$

the statistics T_n and $\hat{T}_n$ are asymptotically equivalent, i.e.

$$n\sqrt{h}\,[\hat{T}_n - T_n] = o_P(1) \ , \tag{3.5}$$

and derived as a consequence the asymptotic normality of $\hat{T}_n$. The following results show that statements of this form remain true under fixed alternatives.

Theorem 3.1. *Assume that* $|\varphi| < 1$. *If the assumptions (2.1) - (2.3), (3.2) - (3.4) are satisfied and the alternative (1.3) is valid, then*

$$\sqrt{n}\,[\hat{T}_n - \int (K_h * (f - f_0))^2(x) dx] \xrightarrow{\mathcal{D}} N(0, 4\varrho^2), \tag{3.6}$$

where the asymptotic variance is given in (2.8).

Theorem 3.2. *Assume that* $|\varphi| > 1$ *and that the assumptions (2.1) - (2.3) and (3.4) are satisfied. If additionally the kernel* K *in the density estimate (3.1) is bounded such that there exits a constant* $B > 0$ *with*

$$\int |K(x + \delta) - K(x)| dx \leq B\delta \tag{3.7}$$

for all $\delta > 0$, *then the assertion (3.6) holds.*

Remark 3.3. Theorem 3.1 and 3.2 are also valid for testing the composite hypothesis (1.6) of a location scale family. To be precise consider the first-order autoregressive model

$$X_t = \mu + \rho X_{t-1} + Z_t, \tag{3.8}$$

where we are interested in testing the hypothesis

$$H_0 : M(f, f_0) = 0 \quad H_1 : M(f, f_0) > 0 \tag{3.9}$$

or the corresponding precise hypotheses of the form (1.4), where

$$M(f, f_0) = \min_{\sigma>0} \int \Big(f(x) - \frac{1}{\sigma} f_0\Big(\frac{x}{\sigma}\Big)\Big)^2 dx \tag{3.10}$$

is the best approximation of the density f by elements from the scale family

$$\mathcal{F} = \Big\{\frac{1}{\sigma} f_0\Big(\frac{\cdot}{\sigma}\Big) \mid \sigma > 0\Big\}.$$

We assume that the minimum in (3.10) exists and is attained at a unique point, say $\sigma_0 > 0$. If $\hat\mu, \hat\varphi, \hat\sigma$ are $\sqrt{n}$-consistent estimates of μ, φ, σ, respectively, $\hat{f}_n$ is the density estimate (3.1) from the residuals $\hat{Z}_i = X_i - \hat\mu - \hat\varphi X_{i-1}$ Lee and Na (2002) showed for the statistic

$$\bar{T}_n = \int \Big\{\hat{f}_n(x) - \Big(K_h * \frac{1}{\hat\sigma} f\Big(\frac{\cdot}{\hat\sigma}\Big)\Big)\Big\}^2 (x) dx \tag{3.11}$$

the asymptotic normality

$$n\sqrt{h}\Big(\bar{T}_n - \frac{1}{nh} \int K^2(x) dx\Big) \xrightarrow{\mathcal{D}} \mathcal{N}(0, \tau^2) \tag{3.12}$$

under the null hypothesis (3.9), where τ^2 is defined in (2.6). Combining these arguments with the arguments given for the proof of Theorem 2.1, 3.1 and 3.2 it can be shown that under any fixed alternative $M(f, f_0) > 0$ it follows

$$\sqrt{n}(\bar{T}_n - M(f, f_0)) \xrightarrow{\mathcal{D}} \mathcal{N}(0, 4\bar\rho^2)\ , \tag{3.13}$$

where

$$\bar\rho = \operatorname{Var}\Big(\Big\{f(Z_i) - \frac{1}{\sigma_0} f_0\Big(\frac{Z_i}{\sigma_0}\Big)\Big\}^2\Big)$$

and σ_0 is the unique minimizer in (3.10). The details are omitted for the sake of brevity.

4 Appendix: Proofs

Proof of Theorem 2.1. Let f denote the "true" density of the random variables Z_i. Recalling the definition of the statistic T_n and the density estimate f_n we obtain the following decomposition

$$\begin{aligned} T_n &= \int [f_n - K_h * f_0]^2(x) dx \\ &= \int [f_n - K_h * f]^2(x) dx + 2\int [f_n - K_h * f](x) g_h(x) dx + \int g_h^2(x) dx \end{aligned} \tag{4.1}$$

$$= \frac{2}{n^2} \sum_{i<j} \int [K_h(x - Z_i) - e_h(x)][K_h(x - Z_j) - e_h(x)]dx$$
$$+ \frac{2}{n} \sum_{i=1}^{n} \Big[(K_h * g_h)(Z_i) - E[(K_h * g_h)(Z_i)]\Big] + \frac{1}{n^2} \sum_{i=1}^{n} \int [K_h(x - Z_i) - e_h(x)]^2 dx$$
$$+ \int g_h^2(x)dx,$$

where the functions e_h and g_h are defined by $e_h := K_h * f$ and $g_h := K_h * (f - f_0)$, respectively. A straightforward calculation shows

$$\frac{1}{n^2} \sum_{i=1}^{n} \int [K_h(x - Z_i) - e_h(x)]^2 dx = \frac{1}{nh} \int K^2(x)dx + O_P(\frac{1}{n}), \tag{4.2}$$

and consequently the statistic T_n can be written as

$$T_n - \frac{1}{nh} \int K^2(x)dx - \int [K_h * (f - f_0)]^2(x)dx = \frac{2}{n^2} \sum_{i<j} H_n(Z_i, Z_j) + \frac{2}{n} \sum_{i=1}^{n} Y_i + O_P(\frac{1}{n}), \tag{4.3}$$

where the random variables $H_n(Z_i, Z_j)$ and Y_i are defined by

$$\begin{aligned} H_n(Z_i, Z_j) &= \int [K_h(x - Z_i) - e_h(x)][K_h(x - Z_j) - e_h(x)]dx, \\ Y_i &= (K_h * g_h)(Z_i) - E[K_h * g_h(Z_i)], \end{aligned} \tag{4.4}$$

respectively. Under the null hypothesis $H_0 : f = f_0$ we have $Y_i \equiv 0$ and obtain the stochastic expansion

$$T_n - \frac{1}{nh} \int K^2(x)dx = \frac{2}{n^2} \sum_{i<j} H_n(Z_i, Z_j) + O_P(\frac{1}{n}) \ = \ U_n + O_P(\frac{1}{n}), \tag{4.5}$$

where the last equality defines the statistic U_n. The asymptotic normality of the statistic

$$n\sqrt{h}\Big\{T_n - \frac{1}{nh} \int K^2(x)dx\Big\}$$

now follows from the corresponding statement for the random variable $n\sqrt{h}U_n$. In order to establish the weak convergence of this statistic we apply a central limit theorem for degenerate U-statistics. More precisely we will check the conditions of Theorem 1 in Hall (1984). Obviously, H_n is symmetric, $E[H_n(Z_1, Z_2) \mid Z_1] = 0$, and $E[H_n^2(Z_1, Z_2)] < \infty$ for each $n \in \mathbb{N}$. Moreover, a straightforward but tedious calculation shows

$$\lim_{n\to\infty} \text{Var}(\sqrt{h}H_n(Z_i, Z_j)) = \lim_{n\to\infty} E[hH_n^2(Z_i, Z_j)] = \int (K * K)^2(x)dx \int f_0^2(x)dx, \tag{4.6}$$

which gives for the variance of $n\sqrt{h}U_n$

$$\begin{aligned}
\mathrm{Var}(n\sqrt{h}U_n) &= E\Big[\frac{4h}{n^2}\sum_{\substack{i<j\\ i'<j'}} H_n(Z_i,Z_j)H_n(Z_{i'},Z_{j'})\Big] \\
&= E\Big[2h\frac{n-1}{n}H_n^2(Z_i,Z_j)\Big] + o(1) = \tau^2 + o(1)\ ,
\end{aligned}$$

where τ^2 is defined in (2.6). The final condition (2.1) of Hall's (1984) Theorem 2.1 is more difficult to check. First note that it follows from (4.6) that $E[H_n^2(Z_i,Z_j)] = O\Big(\frac{1}{h}\Big)$ (uniformly with respect to $i,j \in \{1,\ldots,n\}$). A similar calculation gives

$$E[H_n^4(Z_i,Z_j)] = \frac{1}{h^3}\int f_0^2(x)dx \cdot \int (K*K)^4(x)dx + O\Big(\frac{1}{h^2}\Big) = O\Big(\frac{1}{h^3}\Big). \tag{4.7}$$

Finally, we have to consider the quantity

$$G_n(Z_1,Z_2) = E[H_n(Z_1,Z_3)H_n(Z_3,Z_2) \mid Z_1,Z_2]$$

and obtain

$$\begin{aligned}
E[G_n^2(Z_i,Z_j)] &= \frac{1}{h^2}E\Big[\Big\{\int (K*K)(w)(K*K)\Big(w-\frac{Z_i-Z_j}{h}\Big)f(Z_i)dw\Big\}^2\Big] + O\Big(\frac{1}{h}\Big) \\
&= \frac{1}{h}\int\int\Big\{(K*K)*(K*K)\Big\}^2(s)dsf^4(v)dv + O\Big(\frac{1}{h}\Big) \\
&= O\Big(\frac{1}{h}\Big).
\end{aligned}$$

This gives

$$\frac{E[G_n^2(Z_i,Z_j)] + \frac{1}{n}E[H_n^4(Z_i,Z_j)]}{(E[H_n^2(Z_i,Z_j)])^2} = O\Big(h + \frac{1}{nh}\Big) = o(1)$$

and establishes condition (2.1) of Hall's (1984) Theorem 2.1. We therefore obtain the weak convergence

$$n\sqrt{h}U_n \xrightarrow{\mathcal{D}} \mathcal{N}(0,\tau^2).$$

Finally, the assertion (2.5) follows from (4.5).

For a proof of asymptotic normality of T_n under a fixed alternative of the form (1.3) we note that it follows from (2.3), (2.5), (4.3) and (4.5) that

$$T_n - \int [k_n * (f-f_0)]^2(x)dx = \frac{2}{n}\sum_{i=1}^n Y_i + o_p\Big(\frac{1}{\sqrt{n}}\Big)\ ,$$

where the random variables Y_i are defined by (4.4). A straightforward but tedious calculation shows

$$\mathrm{Var}(Y_i) = \mathrm{Var}((f-f_0)(Z_i)) + O(h^2) = \varrho^2 + O(h^2),$$

with the variance ϱ^2 given in 2.8 so that

$$\mathrm{Var}(\frac{2}{\sqrt{n}}\sum_{i=1}^{n} Y_i) = 4\varrho^2 + o(1),$$

while

$$E[Y_i^4] = O(1),$$

uniformly with respect to $i = 1, \dots, n$. The asymptotic normality of Theorem 2.1 now follows from Ljapunoff's theorem, which completes the proof of Theorem 2.1.

□

Proof of Theorem 3.1 and 3.2. We only consider the case $|\varphi| < 1$, the proof of Theorem 3.2 can be obtained by similar arguments. Obviously, the assertion follows from the estimate

$$\sqrt{n}(\hat{T}_n - T_n) = o_p(1) \tag{4.8}$$

For a proof of this estimate we will proceed as in Lee and Na (2002) who obtained the estimate

$$\int (\hat{f}_n - f_n)^2(x)dx = O_P(n^{-2}h^{-4}). \tag{4.9}$$

On the other hand Theorem 2.1 shows that under a fixed alternative

$$\int (f_n - K_h * f_0)^2(x)dx = O_P(1) \ , \tag{4.10}$$

and a straightforward calculation [using condition (3.4)] gives

$$\begin{aligned} |\hat{T}_n - T_n| \ &\leq \ \int (\hat{f}_n - f_n)^2(x)dx + 2[\int (\hat{f}_n - f_n)^2(x)dx]^{\frac{1}{2}}[\int (f_n - K_h * f_0)^2(x)dx]^{\frac{1}{2}} \\ &= \ O_P(\frac{1}{nh^2}) \ = \ o_P(\frac{1}{\sqrt{n}}), \end{aligned}$$

which proves the assertion of Theorem 3.1.

□

Acknowledgements: The support of the Deutsche Forschungsgemeinschaft (SFB 475, Komplexitätsreduktion in multivariaten Datenstrukturen, Teilprojekt B1) is gratefully acknowledged. Parts of the paper were written during a visit of the second author at the Institute of Statistics in Louvain-la-Neuve and this author would like to thank the institute for its hospitality. The authors are also grateful to Isolde Gottschlich, who typed parts of this paper with considerable technical expertise.

References

[1] **Berger, J. O.; Delampady, M.** (1987). *Testing precise hypotheses. With comments and a rejoinder by the authors.* Stat. Sci. 2, No.3, 317-352.

[2] **Bickel, P.J.; Rosenblatt, M.** (1973). *On some global measures of the deviations of density function estimates.* Ann. Stat. 1, 1071-1095.

[3] **Boldin, M.V.** (1982). *Estimation of the distribution of noise in an autoregression scheme.* Theory Probab. Appl. 27, 866-871.

[4] **Brockwell, P. J.; Davis, R. A.** (1991). *Time series: theory and methods.* Second edition. Springer Series in Statistics. Springer-Verlag, New York.

[5] **Fan, J. Yao, Q.** (2003). *Nonlinear Time Series.* Springer Series in Statistics. Springer-Verlag, New York Berlin Heidelberg.

[6] **Hall, P.** (1984). *Central Limit Theorem for integrated square error of multivariate nonparametric density estimators.* Journal of Multivariate Analysis 14, 1-16

[7] **Koul, H. L.; Levental, S.** (1989). *Weak convergence of the residual empirical process in explosive autoregression.* Ann. Stat. 17, No.4, 1784-1794

[8] **Lee, S.; Na, S.** (2002). *On the Bickel-Rosenblatt test for first-order autoregressive models.* Statist. Probab. Lett. 56, No. 1, 23–35.

[9] **Sellke, T.; Bayarri, M. J.; Berger, J. O.** (2001). *Calibration of p values for testing precise null hypotheses.* Amer. Statist. 55, No. 1, 62–71.