

TESTING HYPOTHESIS FOR THE SHIFT PARAMETER IN DIFFUSION PROCESSES

GEJZA DOHNAL

The problem of testing hypotheses for diffusion type processes when the unknown parameter plays the role of shift is considered. The parametric family of distributions is LAMN in this model and there are no satisfactory results for asymptotic optimality of the tests in the literature. Using the conditional inference approach a test is constructed which is asymptotically optimal under the conditional distribution. Numerical results are presented to illustrate the problem.

1. INTRODUCTION

Dealing with the asymptotic inference for stochastic processes many authors distinguish two type of processes,

- (a) the ergodic one and
- (b) the non-ergodic one

(see [2] for definition). In particular for the diffusion stochastic processes it means the distinction between the processes satisfying

- (a') the local asymptotic normality (LAN) condition, and
- (b') the local asymptotic mixed normality (LAMN) condition (see [6], [7]).

This partition involves different statistical approaches. While the case (a) permits us to use some classical statistical methods and procedures (least-squares or maximum-likelihood estimates, maximum-likelihood ratio tests or tests based on a score statistic, etc.), the situation (b) is quite different. The main difficulties arise from the fact that the estimators and test statistics have non-standard limiting distributions.

The problem of estimation in non-ergodic models was treated by several authors (Basawa and Prakasa Rao [3], Basawa and Brockwell [1], Dohnal [6], Feigin [9], Heyde [16], Jeganathan [11], Swansen [13] and others).

This paper is concerned with a testing hypothesis problem in the LAMN (non-ergodic) case. For ergodic models there exists the uniformly most powerful test based on the log-likelihood ratio (see [3], [13]). Swansen in [13] showed that it is im-

possible to achieve any similar result in the non-ergodic model. The difficulties regarding the efficiency questions were treated by Basawa and Scott [4], Feigin [8] and Sweeting [14]. Basawa and Koul [2] derived the limit distributions of the score and likelihood-ratio statistics.

Some approaches to the investigation of the asymptotic properties of tests in non-ergodic models lead to the conditional inference. The works of Basawa and Brockwell [1] and of Feigin [9] represent the two directions in this way. In this paper the result of the former is used to construct a likelihood-ratio tests which is optimal under the conditional probability measure.

2. THE MODEL

Let the process $\{\xi_t, t \in [0, T]\}$ be defined by the stochastic differential equation

$$d\xi_t = a(\xi_t - \vartheta) dt + b(\xi_t - \vartheta) dW_t, \quad t \in [0, T],$$

where $\{W_t, t \in [0, T]\}$ is the standard Wiener process, ϑ is an unknown real parameter from an open set $\Theta \subset \mathbb{R}$. $a(x)$ and $b(x)$ are continuous real-valued functions with continuous derivatives a', a'', b', b'', b''' , $b(x) > 0$ for all $x \in \mathbb{R}$. Suppose that ξ_0 is a random variable with probability density $\pi(x - \vartheta)$.

Denote by P_ϑ the probability distribution of the process ξ with $\vartheta \in \Theta$, defined on (Ω, F) where $F = \sigma(\xi_t, t \in [0, T])$. Consider the sample $X^n = (X_0, X_1, \dots, X_n)$ from the process ξ such that $X_k = \xi_{kT/n}$, $k = 0, 1, \dots, n$. The chain X^n induces the probability measure P_ϑ^n on (Ω, F^n) , $F^n = \sigma(X_0, \dots, X_n)$. P_ϑ^n is a restriction of P_ϑ to F^n .

Theorem 1. In the model described above, the family $\{P_\vartheta^n, \vartheta \in \Theta\}_{n \geq 1}$ satisfies the local asymptotic mixed normality (LAMN) condition in $\vartheta \in \Theta$, i.e.,

- (i) there exist sequences $\{A_n(\vartheta)\}_{n \geq 1}$, $\{I_n(\vartheta)\}_{n \geq 1}$ of F^n -measurable random variables, $I_n > 0$ a.s. for $n = 1, 2, \dots$ so that
 - (1)
$$L_n(\vartheta_n^n, \vartheta) = \log(dP_{\vartheta_n^n}^n/dP_\vartheta^n) = h A_n(\vartheta) - \frac{1}{2} h^2 I_n(\vartheta) + o_p(1),$$
 as $n \rightarrow \infty$ under P_ϑ^n , where $\vartheta_n^n = \vartheta + hn^{-1/2}$, $h \in \mathbb{R}$, $n = 1, 2, \dots$
 - (ii) there are almost surely positive random variable $I(\vartheta)$ and a random variable $A(\vartheta) \sim N(0, 1)$ independent of $I(\vartheta)$ so that

$$(A_n(\vartheta), I_n(\vartheta)) \rightarrow (A(\vartheta) I^{1/2}(\vartheta), I(\vartheta))$$
 in distribution under P_ϑ^n as $n \rightarrow \infty$.

Moreover

- (iii)
$$\Gamma_n(\vartheta) \rightarrow \Gamma(\vartheta) = 2/T \int_0^T g^2(\xi_t - \vartheta) dt,$$
 in P_ϑ^n -probability as $n \rightarrow \infty$, where $g(x) = b'(x)/b(x)$.

Proof. Recall that

$$L_n(\vartheta_n^n, \vartheta) = \log \pi(X_0 - \vartheta_n^n) - \log \pi(X_0 - \vartheta) + L_n^0(\vartheta_n^n, \vartheta),$$

where $L_n^0(\vartheta_n^n, \vartheta)$ is the log-likelihood ratio in the model with fixed initial value $\xi_0 = x_0$.

Using the Taylor's expansion for logarithm we obtain

$$L_n(\vartheta_n^a, \vartheta) = L_n^0(\vartheta_n^a, \vartheta) + o_p(1), \quad n \rightarrow \infty.$$

Now, the assertions (i), (ii) of Theorem 1 follow immediately from Proposition 1 in [6].

To derive an explicit formula for Δ_n and Γ_n we employ the result of Dacunha-Castelle and Florens-Zmirou [5] which gives a very useful expansion of transition probability density for a transformed process

$$\eta_t = T_\vartheta(\xi_t) = \int_0^{\xi_t} \frac{dt}{b(X_t - \vartheta)}, \quad t \in [0, T]$$

This expansion has the form

$$q^a\left(\frac{T}{n}, Y_{k-1}^a, Y_k^a\right) = \left(\frac{n}{2\pi T}\right)^{1/2} \exp\left\{-\frac{n}{2T}(Y_k^a - Y_{k-1}^a) + \int_{Y_{k-1}^a}^{Y_k^a} f(y, \vartheta) dy - \frac{T}{2n} \left(\int_{Y_{k-1}^a}^{Y_k^a} \left(f(y, \vartheta) + \frac{\partial}{\partial y} f(y, \vartheta)\right) dy\right) (Y_k^a - Y_{k-1}^a)^{-1} + o_p\left(\frac{T}{n}\right)^{3/2}\right\},$$

where $f(y, \vartheta) = a(T_\vartheta^{-1}(y) - \vartheta)/b(T_\vartheta^{-1}(y) - \vartheta) - \frac{1}{2}b'(T_\vartheta^{-1}(y) - \vartheta)$. For ϑ_n^a and ϑ_0 we obtain

$$(2) \quad \frac{q^{s_n^n}\left(\frac{T}{n}, Y_{k-1}^{s_n^n}, Y_k^{s_n^n}\right)}{q^{s_0}\left(\frac{T}{n}, Y_{k-1}^{s_0}, Y_k^{s_0}\right)} = \exp\left\{\frac{n}{T} [g(X_k - \vartheta_0) h n^{-1/2} + (g'(X_k - \vartheta_0) - g^2(X_k - \vartheta_0)) h^2 n^{-1}] (W_k - W_{k-1})^2 + o_p(n^{-1/2})(W_k - W_{k-1}) + o_p(n^{-3/2})\right\}.$$

Notice that

$$\frac{dP_\vartheta^n}{dP_{\vartheta_0}^n} = \prod_{k=1}^n \frac{p_\vartheta\left(\frac{T}{n}, X_{k-1}, X_k\right)}{p_{\vartheta_0}\left(\frac{T}{n}, X_{k-1}, X_k\right)} = \prod_{k=1}^n \frac{b(T_{\vartheta_0}^{-1}(X_k^{s_0}) - \vartheta_0) q^s\left(\frac{T}{n}, Y_{k-1}^s, Y_k^s\right)}{b(T_\vartheta^{-1}(Y_k^s) - \vartheta) q^{s_0}\left(\frac{T}{n}, Y_{k-1}^{s_0}, Y_k^{s_0}\right)}.$$

Hence from (2) and using the Taylor's expansion for the logarithm of $b(x - \vartheta)$

$$L_n(\vartheta_n^a, \vartheta_0) = \sum_{k=1}^n \left[\frac{h}{\sqrt{n}} g(X_k - \vartheta_0) \left(\frac{n}{T} (\delta W_k)^2 - 1\right) - \frac{h^2}{n} g^2(X_k - \vartheta_0) \right] + R_n$$

where R_n is such that $P_{\vartheta_0}^n \text{-} \lim_{n \rightarrow \infty} R_n = 0$. This gives

$$(3) \quad \Gamma_n = 2/n \sum_{k=1}^n g^2(X_k - \vartheta_0)$$

$$(4) \quad A_n = \left[\sum_{k=1}^n n^{-1/2} g(X_k - \vartheta_0) (n(\delta W_k)^2 - 1) \right] \Gamma_n^{-1/2}.$$

(iii) of Theorem 1 follows from the convergence of sums (3) to the integral $2/T \cdot \int_0^T g^2(X_t - \vartheta) dt$. The expression of A_n will be used in the sequel. \square

Remark. If $\Gamma(\vartheta)$ is a non-degenerate random variable then the model is called non-ergodic. The case when $\Gamma(\vartheta)$ is nonrandom corresponds to the ergodic model and coincides with occurs of g is constant. Thus the model satisfy the LAN condition.

3. TESTING HYPOTHESIS

Let us consider the problem of testing hypothesis $H: \vartheta = \vartheta_0$ against the one-sided alternative $K: \vartheta > \vartheta_0$. The test is defined by the indicator function $\varphi_n = I(T_n(X^n) \in C)$, $C \in \mathbb{R}$, where $T_n(X^n)$ is a test statistic. The power function of the test φ_n is given by

$$\beta_{\varphi_n}(\vartheta) = E_{\vartheta} I(T_n(X^n) \in C).$$

The test φ_n is said to be of size α (or simply α -test) if $\beta_{\varphi_n}(\vartheta_0) = \alpha$, $\alpha \in [0, 1)$.

One way to compare tests in parametric families of distributions is to consider a sequence of testing problems $H: \vartheta = \vartheta_0$ against $K_n: \vartheta = \vartheta_h^n$, $h > 0$, $n = 1, 2, \dots$. We shall say that a sequence of α -tests is (locally asymptotically) optimal if there exists $h_0 > 0$ so that

$$\overline{\lim}_{n \rightarrow \infty} \beta_{\varphi_n}(\vartheta_h^n) \leq \lim_{n \rightarrow \infty} \beta_{\varphi_n}(\vartheta_h^n)$$

for all sequences of α -tests $\{\varphi_n\}_{n \geq 1}$ and for all $h \in [0, h_0]$.

In the LAMN model a uniformly most powerful test does not exist. From the Neyman-Pearson Lemma it follows that the most powerful α -test of hypothesis $H: \vartheta = \vartheta_0$ against $K_n: \vartheta = \vartheta_h^n$ we have in the form $\varphi_n = I(L_n(\vartheta_h^n, \vartheta_0) > k_n^*)$, $n = 1, 2, \dots$, what gives (see (1))

$$\varphi_n = I(h \Delta_n(\vartheta_0) - \frac{1}{2} h^2 \Gamma_n(\vartheta_0) + o_p(1) > k_n^*), \quad n = 1, 2, \dots$$

where k_n^* is defined by

$$E_{\vartheta_0} I(h \Delta_n(\vartheta_0) - \frac{1}{2} h^2 \Gamma_n(\vartheta_0) > k_n^*) = \alpha.$$

Swansen in [13] shows that there exists an optimal sequence of uniformly most powerful α -tests only if $\Gamma(\vartheta_0)$ is non-random.

4. CONDITIONALITY

Define a regular conditional probability measure $P_{\vartheta_0|\gamma}$ on (Ω, F) by

$$P_{\vartheta_0|\gamma}(A) = P_{\vartheta_0}(A \mid \Gamma(\vartheta) = \gamma^2), \quad A \in F.$$

Let $P_{\vartheta_0|\gamma}^n$ be a restriction of $P_{\vartheta_0|\gamma}$ on (Ω, F^n) . $\Gamma(\vartheta)$ is so called "mixing variable" and

the LAMN family $\{P_{\vartheta}^n, \vartheta \in \Theta\}_{n \geq 1}$ can be considered intuitively as a mixture of a LAN families of conditional measures $\{P_{\vartheta|\gamma}^n, \vartheta \in \Theta\}_{n \geq 1}$ in the following sense,

$$P_{\vartheta}^n = \int P_{\vartheta|\gamma} P_{\vartheta}(I(\vartheta) \in d\gamma^2).$$

When the mixing variable $I(\vartheta)$ is ancillary then it is natural to consider a conditional inference approach (see [10]). Basawa and Brockwell in [1] proved the next theorem.

Theorem 2. Suppose that the conditions (i), (ii), (iii) from Theorem 1 hold. Let $P_{\vartheta|\gamma}^n$ be the conditional probability measure under P_{ϑ} for X^n given $I(\vartheta) = \gamma^2$. Then under $P_{\vartheta|\gamma}$ and $n \rightarrow \infty$

- (I) $L_n(\vartheta_n^*, \vartheta) = h \Delta_n(\vartheta) - \frac{1}{2} h^2 \gamma^2 + o_p(1)$,
 - (II) $(\Delta_n(\vartheta), \Gamma_n(\vartheta)) \rightarrow (\Delta\gamma, \gamma^2)$ in distribution,
 - (III) $\Gamma_n(\vartheta) \rightarrow \gamma^2$ in probability
- (excepting $\gamma^2 \in N(\vartheta)$ such that $P_{\vartheta}(I(\vartheta) \in N(\vartheta)) = 0$).

See [1] for the proof.

In the case studied here we can show that $I(\vartheta)$ is an ancillary variable. We change the integral

$$I(\vartheta) = 2/T \int_0^T g^2(X_t - \vartheta) dt$$

into

$$I = 2/T \int_0^T g^2(Z_t) dt$$

by the substitution $Z_t = X_t - \vartheta$. $\{Z_t, t \in [0, T]\}$ is the process satisfying the stochastic differential equation

$$dZ_t = a(Z_t) dt + b(Z_t) dW_t, \quad t \in [0, T],$$

with initial density $\pi(z)$ for Z_0 . Hence I is independent of ϑ and thus it is ancillary.

Now, using Theorem 2 and the Neyman-Pearson Lemma we are able to construct a test φ_n which is optimal under $P_{\vartheta|\gamma}$. When $I = \gamma^2$ is fixed then from (1) and (4) we obtain the test statistic in the form

$$T_n(X^n) = \Delta_n(\vartheta_0) = \frac{1}{\gamma} \sum_{k=0}^{n-1} g(X_k - \vartheta_0) \frac{n[X_{k+1} - X_k - a(X_k - \vartheta_0)]^2 - b^2(X_k - \vartheta_0)}{\sqrt{(n) b^2(X_k - \vartheta_0)}}.$$

$T_n(X^n)$ has asymptotically the $N(0, 1)$ distribution under $P_{\vartheta_0|\gamma}$ and the α -test can be given by

$$I(T_n(X^n) > u_{\alpha})$$

or

$$I(|T_n(X^n)| > u_{\frac{\alpha}{2}})$$

as two-sided alternative test. u_{α} is α -percentile of $N(0, 1)$ distribution. Under $P_{\vartheta_0|\gamma}$ has $T_n(X^n)$ the $N(0, 1)$ distribution (see [15]).

To perform the conditional inference we cannot use the principle of ancillarity directly since I depends on ϑ . We don't know the true value of γ . Thus we replace

$T_n(X^n)$ by

$$\tilde{T}_n(X^n) = \frac{A_n}{\tilde{\gamma}_n}$$

where $A_n = A_n(\vartheta_0)$, $\tilde{\gamma}_n^2 = 2/n \sum_{k=0}^{n-1} g^2(X_k - \vartheta_0)$. We can show that the limiting distribution of $\tilde{T}_n(X^n)$ is the same as of $T_n(X^n)$, as $n \rightarrow \infty$, under $P_{\vartheta_0|17}$, $P_{\vartheta_n|17}$, respectively.

Lemma 1. $(T_n(X^n) - \tilde{T}_n(X^n)) \rightarrow 0$ in probability under $P_{\vartheta_0|17}$ and $P_{\vartheta_n|17}$, respectively as $n \rightarrow \infty$.

Proof. Let

$$T_n(X^n, \vartheta) = \frac{A_n}{\tilde{\gamma}_n(\vartheta)},$$

where $\tilde{\gamma}_n^2(\vartheta) = 2/n \sum_{k=0}^{n-1} g^2(X_k - \vartheta)$. First, we show that

$$(5) \quad T_n(X^n, \vartheta_n^h) = T_n(X^n, \vartheta_0) + O_P(n^{-1/2}), \quad n \rightarrow \infty.$$

Denoting

$$G_n(\vartheta) = \frac{1}{\tilde{\gamma}_n(\vartheta)} = \left(\frac{2}{n} \sum_{k=0}^{n-1} g_k^2(\vartheta) \right)^{-1/2}, \quad g_k(\vartheta) = g(X_k - \vartheta),$$

using Taylor's expansion we obtain

$$G_n(\vartheta_n^h) = G_n(\vartheta_0) + G'_n(\vartheta_0) \frac{h}{\sqrt{n}} + O_P(n^{-1/2}),$$

where

$$G'_n(\vartheta_0) = - \left(\frac{2}{n} \sum_{k=0}^{n-1} g'_k(\vartheta_0) g_k(\vartheta_0) \right) \left(\frac{2}{n} \sum_{k=0}^{n-1} g_k^2(\vartheta_0) \right)^{-3/2}.$$

From Hölder's inequality follows

$$\frac{1}{n} \sum_{k=0}^{n-1} g'_k g_k \leq O_P(1) \frac{1}{n} \sum_{k=0}^{n-1} |g'_k| \leq O_P(1) \left(\frac{1}{n} \sum_{k=0}^{n-1} g_k^2 \right)^{1/2}$$

for all n . Hence

$$G'_n(\vartheta_0) \leq O_P(1) \frac{\left(\frac{1}{n} \sum_{k=0}^{n-1} g_k^2(\vartheta_0) \right)^{1/2}}{\left(\frac{2}{n} \sum_{k=0}^{n-1} g_k^2(\vartheta_0) \right)^{3/2}} \leq O_P(1) \frac{1}{\tilde{\gamma}_n^2(\vartheta_0)}$$

and

$$G_n(\vartheta_n^h) = \frac{1}{\tilde{\gamma}_n(\vartheta_0)} + O_P(n^{-1/2}), \quad n \rightarrow \infty$$

which implies (5).

Note that $T_n(X^n, \vartheta_0) = \tilde{T}_n(X^n)$. Moreover (III) in Theorem 2 implies

$$(T_n(X^n, \vartheta) - T_n(X^n)) \rightarrow 0 \quad \text{as } n \rightarrow \infty$$

in P_{ϑ_1} probability. The assertion of Lemma 1 follows immediately. \square

5. NUMERICAL RESULTS

This section gives some numerical results obtained by simulation. The process ξ was simulated using the Heune's scheme (see [12]) for the equation

$$d\xi_t = -(\xi_t - \vartheta) dt + (1 + (\xi_t - \vartheta)^2) dW_t, \quad t \in [0, 1],$$

$$\pi(x - \vartheta) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}(x - \vartheta)^2\right).$$

In this case

$$g(X_k - \vartheta) = \frac{2(X_k - \vartheta)}{1 + (X_k - \vartheta)^2}.$$

ϑ_0	ϑ_1	n	m	m_A $\alpha = 0.05$	m_A $\alpha = 0.1$	\bar{T}_n	\bar{G}_n
1.0	0.5	100	20	10	3	2.112	1.352
1.0	1.0	100	20	18	18	0.929	0.253
1.0	1.1	100	20	20	18	0.8586	0.288
1.0	1.2	100	20	18	16	1.123	0.392
1.0	1.3	100	20	15	10	1.402	0.609
1.0	1.5	100	20	6	2	2.385	1.361
1.0	2.0	100	20	0	0	4.764	1.916
3.0	2.4	100	20	4	2	2.839	1.328
3.0	2.7	100	20	14	11	1.389	0.631
3.0	3.25	100	20	17	17	1.230	0.514

The test statistic was constructed for the hypothesis $H: \vartheta = \vartheta_0$. Hence its value and the value of \tilde{y}_n was computed with ϑ_0 , while the simulation of the process ξ was done with $\vartheta = \vartheta_1$. There were performed m simulations with the step n^{-1} for every ϑ_0, ϑ_1 . The number of acceptances of the hypothesis was denoted by m_A . \bar{T}_n, \bar{G}_n denote the average value of $|T_n|, G_n$, respectively.

(Received June 12, 1987.)

REFERENCES

- [1] I. V. Basawa and P. J. Brockwell: Asymptotic conditional inference for regular nonergodic models with an application to autoregressive processes. *Ann. Statist.* *12* (1984), 161–171.
- [2] I. V. Basawa and H. L. Koul: Asymptotic tests of composite hypothesis for non-ergodic type stochastic processes. *Stoch. Proc. Appl.* *9* (1979), 291–305.
- [3] I. V. Basawa and B. L. S. Prakasa Rao: Asymptotic inference for stochastic processes. *Stoch. Proc. Appl.* *10* (1980), 221–254.
- [4] I. V. Basawa and D. J. Scott: Efficient tests for stochastic processes. *Sankhyā Ser. A* *39* (1977), 21–31.
- [5] D. Dacunha-Castelle and D. Florens-Zmirou: Time-discretization effect for the estimation of a parameter of a differential stochastic equation. Prepublication, Univ. de Paris-Sud, 1984.
- [6] G. Dohnal: On estimating the diffusion coefficient. *J. Appl. Probab.* *24* (1987), 105–114.
- [7] G. Dohnal: Odhadý koeficientu difúze (Estimation of the Diffusion Coefficient). Ph. D. Thesis, Charles University, Prague 1985.
- [8] P. D. Feigin: The efficiency criteria problem for stochastic processes. *Stoch. Proc. Appl.* *6* (1978), 115–127.
- [9] P. D. Feigin: Asymptotic theory of conditional inference for stochastic processes. *Stoch. Proc. Appl.* *22* (1986), 89–102.
- [10] J. Hájek: On basis concept of statistics. *Proc. 5th Berkeley Symp. Math. Stat. Probab. I* (1967), 139–162.
- [11] P. Jeganathan: On the asymptotic theory of estimation when the limit of log-likelihood ratios is mixed normal. *Sankhyā Ser. A* *44* (1982), 173–212.
- [12] W. Rümelin: Numerical Treatment of Stochastic Differential Equations. *Forschungsschwerpunkt Dynamische Systeme*, Univ. Bremen, 1980.
- [13] A. R. Swansen: A note on asymptotic inference in a class of non-stationary processes. *Stoch. Proc. Appl.* *15* (1983), 181–191.
- [14] T. J. Sweeting: On efficient tests for branching processes. *Biometrika* *65* (1978), 123–127.
- [15] G. G. Roussas: *Contiguity of Probability Measures*. Cambridge Tracts in Math. and Math. Physics, Cambridge 1972.
- [16] C. C. Heyde: On an optimal property of the maximum likelihood estimator of a parameter from a stochastic process. *Stoch. Proc. Appl.* *1* (1978), 1–19.

RNDr. Gejza Dohnal, CSc., katedra matematiky a deskriptivní geometrie strojní fakulty ČVUT (Department of Mathematics and Descriptive Geometry, Faculty of Mechanical Engineering, Czech Technical University), Karlovo nám. 13, 120 00 Praha 2, Czechoslovakia.