Selected Transl. in Math.
Statist. and Probaility
Volume 14, 1978

ON THE STATISTICAL EXAMINATION OF CONTINUOUS STATE MARKOV PROCESSES. II*

MÁTYÁS ARATÓ

Introduction

The first part [2] of this paper dealt with continuous-time one-dimensional stationary Gaussian Markov processes, although the observations are usually made at discrete points; therefore we must now also investigate the discrete case from the point of view of statistics. Contrary to continuous processes, the number of unknown parameters now equals three: m, ρ, σ_ξ^2 or m, ρ, σ_ζ^2, where

$$m = \mathsf{E}\,\xi(n), \quad \sigma_\xi^2 = \operatorname{Var}\,\xi(n), \quad \varrho = \frac{\mathsf{E}\,\{\xi(n)-m\}\{\xi(n-1)-m\}}{\sigma_\xi^2}, \tag{1}$$
$$\sigma_\zeta^2 = (1-\varrho^2)\,\sigma_\xi^2, \qquad (n=1, 2, \ldots).$$

In this paper we shall use the results obtained for the continuous case, and therefore references to part I will often be made.

It should be mentioned that the case of a known ρ and σ_ξ^2 has been investigated by Luvsanceren [8], [9], the case of a known σ_ξ^2 and m by Linnik [7], and the case of a known σ_ξ^2 also by Luvsanceren [8], [9].

In this paper we shall assume that all three parameters are unknown, and we shall examine the behavior of various estimators.

I have already mentioned that a number of papers (see, for example, T. W. Anderson [1] and the references given there) have dealt with the estimation of the parameter ρ; a survey of these results will be included in the present paper. It should also be mentioned that misformulation of the problem has often in the past led to unsatisfactory results.

Some of the results obtained in this paper were included in my dissertation [3], but the theorems and their proofs are published here for the first time. Since the completion of my dissertation, new results have required further elaboration and publication (see §6).

AMS (MOS) subject classifications (1970). Primary 62F10, 62M05, 62-02; Secondary 62F20.

*Translation of Magyar Tud. Akad. Mat. Fiz. Oszt. Közl. **14** (1964), 137-159.

Time Discrete Stationary Normal One-Dimensional Case

§1. Distribution of observations and possible estimators

Theorem 1, proved in the previous paper [2], implies that the correlation function $R(n)$ of a stationary Gaussian Markov process is $R(n) = \rho^n$, $n = 0, \pm 1, \dots$, where $\rho = \mathbf{E}\{(\xi(k) - m)(\xi(k-1) - m)\}/(\mathrm{Var}\ \xi(k))$. On the other hand (see, for example, Rosenblatt [11]), $\xi(n)$ satisfies the difference equation

(1.1) $$\xi(n+1) - \varrho\xi(n) = \zeta(n+1)$$

assuming that $\mathbf{E}\xi(n) = 0$. Here $\zeta(n)$ is a sequence of independent normally distributed random variables, and $\zeta(n)$ is independent of $\xi(n-1)$. When $\mathrm{Var}\ \zeta(n) = \sigma_\xi^2$, the following formula evidently holds:

(1.2) $$\sigma_\zeta^2 = (1-\sigma^2)\sigma_\xi^2.$$

From (1.1) it follows that for $\mathbf{E}\xi(k) = m$ the density function of the random vector $\xi(1), \dots, \xi(n)$ has the following form:

(1.3)
$$p_{\xi(1),\dots,\xi(n)}(x_1, \dots, x_n) = (2\pi)^{-n}\sigma_\xi^{-n}(1-\varrho^2)^{-\frac{n-1}{2}} \cdot \exp\left\{-\frac{1}{2\sigma_\xi^2(1-\varrho^2)}\left[(x_1-m)^2(1-\varrho^2) + \sum_{i=2}^n (x_i - \varrho x_{i-1} - m(1-\varrho))^2\right]\right\}$$
$$= (2\pi)^{-n}\sigma_\zeta^{-n}(1-\varrho^2)^{1/2}\exp\left\{-\frac{1}{2\sigma_\zeta^2}\left[(x_1-m)^2(1-\varrho^2) + \sum_{i=2}^n (x_i - \varrho x_{i-1} - m(1-\varrho^2))^2\right]\right\},$$

or in matrix notation

$$p_{\xi(1),\dots,\xi(n)}(x_1, \dots, x_n) = (2\pi)^{-n}\sigma_\xi^{-n}(1-\varrho^2)^{-\frac{n-1}{2}} \cdot \exp\left\{-\frac{1}{2\varrho_\xi^2(1-\varrho^2)}(X-m)R_n^{-1}(X-m)^*\right\},$$

where

$$R_n^{-1} = \begin{pmatrix} 1 & -\varrho & 0 & 0 & \dots & 0 \\ -\varrho & 1+\varrho^2 & -\varrho & 0 & \dots & 0 \\ 0 & -\varrho & 1+\varrho^2 & -\varrho & \dots & 0 \\ \vdots & & \ddots & & & \\ \vdots & & & \ddots & & \\ 0 & \dots & \dots & & 1+\varrho^2 & -\varrho \\ 0 & \dots & \dots & & -\varrho & 1 \end{pmatrix},$$

$$X - m = (x_1 - m, \dots, x_n - m), \quad (X-m)^* = \begin{pmatrix} x_1 - m \\ \vdots \\ x_n - m \end{pmatrix}.$$

The conditional density function of the random variables $\xi(2), \dots, \xi(n)$ under the condition that $\xi(1) = x_1$ is

(1.4)
$$p(x_2, \dots, x_n | \xi(1) = x_1) = (2\pi)^{-\frac{n-1}{2}}\sigma_\zeta^{-(n-1)}\exp\left\{-\frac{1}{2\sigma_\zeta^2}\sum_2^n (x_i - \varrho x_{i-1} - m(1-\varrho))^2\right\}.$$

Before investigating maximum likelihood estimators, let us deal with the logarithmic derivative of the density. Let the unknown parameters be m, σ_ξ^2 and ρ, and let us introduce the following notation:

(1.5)
$$R_n^{(1)} = \frac{\partial \log p}{\partial m} = \frac{1}{\sigma_\xi^2}\left\{(x_1 - m) + \frac{1}{1+\varrho}\sum_{i=2}^n (x_i - m - \varrho(x_{i-1} - m))\right\},$$
$$R_n^{(2)} = \frac{\partial \log p}{\partial \sigma_\xi^2} = -\frac{n}{2\sigma_\xi^2} + \frac{1}{2\sigma_\xi^4}\left\{(x_1 - m)^2 + \frac{1}{1-\varrho^2}\sum_2^n (x_i - m - \varrho(x_{i-1} - m))^2\right\},$$
$$R_n^{(3)} = \frac{\partial \log p}{\partial \varrho} = \frac{(n-1)\varrho}{1-\varrho^2} - \frac{\varrho}{\sigma_\xi^2(1-\varrho^2)}\sum_2^n [x_i - m - \varrho(x_{i-1} - m)]^2 + \frac{1}{\sigma_\xi^2(1-\varrho^2)}\sum_2^n [x_i - m - \varrho(x_{i-1} - m)](x_{i-1} - m).$$

In the case of the unknown parameters m, σ_ζ^2 and ρ the corresponding derivatives will have the following form:

$$H_n^{(1)} = \frac{\partial \log p}{\partial m} = \frac{1-\varrho}{\sigma_\zeta^2}\left\{(1+\varrho)(x_1-m) + \sum_2^n (x_i - m - \varrho(x_{i-1}-m))\right\},$$

(1.6)

$$H_n^{(2)} = \frac{\partial \log p}{\partial \sigma_\xi^2} = -\frac{n}{2\sigma_\zeta^2} + \frac{1}{2\sigma_\zeta^4}\left\{(1-\varrho^2)(x_1-m)^2 + \sum_2^n (x_i - m - \varrho(x_{i-1}-m))^2\right\}, \quad H_n^{(3)}$$

$$= \frac{\partial \log p}{\partial \varrho} = -\frac{\varrho}{1-\varrho^2} + \frac{1}{\sigma_\zeta^2}\left\{\varrho(x_1-m)^2 + \sum_2^n (x_i - m - \varrho(x_{i-1}-m))(x_{i-1}-m)\right\}.$$

Using the above-mentioned formula $\mathsf{E}(\xi(i)-m)(\zeta(i+k)-m) = 0$ for $k \geqslant 1$ and the characteristic property of normally distributed variables (provided $\mathsf{E}\xi_i = 0$)

$$\mathsf{E}\xi_1\xi_2\xi_3\xi_4 = \mathsf{E}\xi_1\xi_2\,\mathsf{E}\xi_3\xi_4 + \mathsf{E}\xi_1\xi_3\,\mathsf{E}\xi_2\xi_4 + \mathsf{E}\xi_1\xi_4\,\mathsf{E}\xi_2\xi_3,$$

and after simple but lengthy calculations, we obtain

(1.7)

$$\mathrm{Var}\ R_n^{(1)} = \mathsf{E}R_n^{(1)}R_n^{(1)} = \frac{2+(n-2)(1-\varrho)}{(1+\varrho)\sigma_\xi^2},$$

$$\mathrm{Var}\ R_n^{(2)} = \mathsf{E}R_n^{(2)}R_n^{(2)} = \frac{n}{2\sigma_\xi^4},$$

$$\mathrm{Var}\ R_n^{(3)} = \mathsf{E}R_n^{(3)}R_n^{(3)} = \frac{(n-1)(1+\varrho^2)}{(1-\varrho^2)^2},$$

$$\mathsf{E}R_n^{(1)}R_n^{(2)} = \mathsf{E}R_n^{(1)}R_n^{(3)} = 0,$$

$$\mathsf{E}R_n^{(2)}R_n^{(3)} = -\frac{(n-1)\varrho}{\sigma_\xi^2(1-\varrho^2)},$$

and similarly

(1.8)

$$\mathrm{Var}\ H_n^{(1)} = \frac{(1-\varrho)[2+(n-2)(1-\varrho)]}{\sigma_\zeta^2},$$

$$\mathrm{Var}\ H_n^{(2)} = \frac{n}{2\sigma_\zeta^4},$$

$$\mathrm{Var}\ H_n^{(3)} = \frac{n-1}{1-\varrho^2} + \frac{2\varrho^2}{(1-\varrho^2)^2},$$

$$\mathsf{E}H_n^{(1)}H_n^{(2)} = \mathsf{E}H_n^{(1)}H_n^{(3)} = 0,$$

$$\mathsf{E}H_n^{(2)}H_n^{(3)} = \frac{\varrho}{\sigma_\zeta^2(1-\varrho^2)}, \quad \frac{\mathsf{E}H_n^{(2)}H_n^{(3)}}{\sqrt{\mathrm{Var}\ H_n^{(2)}}\sqrt{\mathrm{Var}\ H_n^{(3)}}} = \frac{\varrho\sqrt{2}}{\sqrt{n[(n-1)(1-\varrho^2)+2\varrho^2]}}.$$

As can be seen, when all three parameters m, σ_ξ^2, ρ or m, σ_ζ^2, ρ are unknown the determination of the maximum likelihood estimators is very time-consuming, and a successful investigation of their asymptotic behavior can hardly be expected. Instead, we shall use an idea first arrived at by Wald, according to which we start with a study of the asymptotic behavior of the quantities $R_n^{(1)}$, $R_n^{(2)}$, $R_n^{(3)}$ or $H_n^{(1)}$, $H_n^{(2)}$, $H_n^{(3)}$ and only subsequently demonstrate that the solution of the system of equations (when normalized by the corresponding variances) has the same distribution, uniformly in unknown parameters, as the quantities $R_n^{(1)}/\sqrt{\mathrm{Var}\ R_n^{(1)}}$, $R_n^{(2)}/\sqrt{\mathrm{Var}\ R_n^{(2)}}$ and $R_n^{(3)}/\sqrt{\mathrm{Var}\ R_n^{(3)}}$. The normalizing factors, which are to be multiplied by, will be precisely $\sqrt{\mathrm{Var}\ R_n^{(1)}}$, $\sqrt{\mathrm{Var}\ R_n^{(2)}}$ and $\sqrt{\mathrm{Var}\ R_n^{(3)}}$ respectively.

When considering the variances in (1.7) and (1.8) we immediately obtain the result that for ρ close to one the maximum likelihood estimation of m is not consistent in the case of (1.7), while in the case of (1.8) the variance does not remain finite.

As we are mainly interested in the asymptotic behavior of estimates, it is sufficient to consider estimates following from the conditional density function (1.4). In this case, for example,

(1.9)

$$\bar{H}_n^{(1)} = \frac{\partial \log p(\ldots|.)}{\partial m} = \frac{1-\varrho}{\sigma_\zeta^2}\sum_2^n (x_i - m - \varrho(x_{i-1}-m)),$$

$$\bar{H}_n^{(2)} = \frac{\partial \log p(\ldots|.)}{\partial \sigma_\zeta^2} = -\frac{n-1}{2\sigma_\zeta^2} + \frac{1}{2\sigma_\zeta^4}\sum_2^n (x_i - m - \varrho(x_{i-1}-m))^2,$$

$$\bar{H}_n^{(3)} = \frac{\partial \log p(\ldots|.)}{\partial \varrho} = \frac{1}{\sigma_\zeta^2}\sum_2^n (x_i - m - \varrho(x_{i-1}-m))(x_{i-1}-m).$$

Thus the corresponding system of maximum likelihood equations is simplified. From the equations $\bar{H}_n^{(1)} = 0$, $\bar{H}_n^{(2)} = 0$ and $\bar{H}_n^{(3)} = 0$ we obtain the following relations between $\hat{m}$, $\hat{\sigma}_\xi^2$ and $\hat{\rho}$:

$$\hat{m} = \frac{x_n - \hat{\varrho} x_1 + (1-\hat{\varrho}) \sum_2^n x_i}{(n-1)(1-\hat{\varrho})},$$

$$\hat{\sigma}_\zeta^2 = \frac{1}{n-1} \sum_2^n (x_i - \hat{m} - \hat{\varrho}(x_{i-1} - \hat{m}))^2,$$

$$\hat{\varrho} = \frac{\sum_2^n (x_i - \hat{m})(x_{i-1} - \hat{m})}{\sum_2^n (x_{i-1} - \hat{m})^2}.$$

If we assume that our sequence of observation is cyclical (i.e. $x_n = x_1$), it follows from the above equations that m can be estimated by the arithmetic mean, while for the estimation of ρ we obtain a so-called serial correlation coefficient. These simplifications are not always permissible;(1) for instance for $\rho \to 1$ the best estimate for m would be $(x_1 + x_n)/2$, as we shall see below.

From the form of the density function (1.3) we can obtain a sufficient system of statistics belonging to such parameters in the form

$$\left\{x_1 + x_n,\ \sum_2^{r-1} x_i,\ x_1^2 + x_n^2,\ \sum_2^{n-1} x_i^2,\ \sum_2^n x_i x_{i-1}\right\}.$$

§2. Estimates of single parameters and their distributions

When only a single parameter is unknown we obtain the following estimates.

Let m be unknown. Its maximum likelihood estimate (cf. (2.6) in [2]) will be

$$\hat{m} = \frac{x_1 + x_n + (1-\varrho) \sum_2^n x_i}{2 + (1-\varrho)(n-2)}, \tag{2.1}$$

where m is normally distributed with parameters

$$\left(m, \sigma_\xi \sqrt{\frac{1+\varrho}{2+(n-2)(1-\varrho)}}\right).$$

Let $m = 0$; then σ_ξ^2 has the following maximum likelihood estimate:

$$\hat{\sigma}_\xi^2 = \frac{1}{n(1-\varrho^2)}\left\{(1-\varrho^2)x_1^2 + \sum_2^n (x_i - \varrho x_{i-1})^2\right\}. \tag{2.2}$$

(1)Papers dealing with statistics of processes often, however, contain much simplifications (see, for example, Anderson [1] and the references mentioned there), and it is therefore necessary to show the interconnections with the present paper.

The estimate $\hat{\sigma}_\xi^2$ has a χ^2 distribution with expectation σ_ξ^2 and characteristic function

$$f(\alpha) = \left(1 - \frac{2\sigma_\zeta^2 i\alpha}{n(1-\varrho^2)}\right)^{-\frac{n}{2}} = \left(1 - \frac{2\sigma_\xi^2 i\alpha}{n}\right)^{-\frac{n}{2}}. \tag{2.3}$$

Let $m = 0$; let σ_ξ^2 (or σ_ζ^2) be known and ρ unknown. To obtain a maximum likelihood estimate we must solve a cubic equation, while on the basis of conditional density function we obtain the following estimate from (1.9):

$$\hat{\varrho} = \frac{\sum_2^n x_i x_{i-1}}{\sum_1^{n-1} x_i^2}. \tag{2.4}$$

The ergodic theorem yields

$$\frac{1}{n-1} \sum_1^{n-1} \xi^2(i) \to \sigma_\xi^2 \tag{2.5}$$

in the mean square sense and with probability 1, while the random variable

$$\sum_2^n \xi(i)\xi(i-1) - \varrho \sum_1^{n-1} \xi^2(i) = \sum_2^n \xi(i-1)(\xi(i) - \varrho\xi(i-1)) = \sum_2^n \xi(i-1)\zeta(i) \tag{2.6}$$

has the variance $(n-1)(1-\rho^2)\sigma_\xi^2 = (n-1)\sigma_\zeta^2$. Hence it follows that the variance of the random variable

$$\sqrt{n-1}(\hat{\varrho} - \varrho) = \sqrt{n-1}\, \frac{\sum_2^n \xi(i)\xi(i-1) - \varrho \sum_1^{n-1} \xi^2(i)}{\sum_1^{n-1} \xi^2(i)}$$

$$= \frac{\sum_2^n \xi(i)\xi(i-1) - \varrho \sum_1^{n-1} \xi^2(i)}{\sum_1^{n-1} \frac{\xi^2(i)}{n-1}} \cdot \frac{1}{\sqrt{n-1}}$$

asymptotically equals $1 - \rho^2$.

The estimate $\hat{\rho}$ has a distribution which is asymptotically normal for any fixed value of ρ; this follows e.g. from the results obtained by Volkonskiĭ and Rozanov [13]. This uniform asymptotic normality, however, holds true only for the interval $-1 + \epsilon < \rho < 1 - \epsilon$ (arbitrary $\epsilon > 0$). Thus confidence intervals (upper and lower estimates) for ρ can only be constructed in an open interval $(-1, 1)$. Linnik [7] has dealt with the problem of estimators for ρ.

2. In case of two unknown parameters we shall mention only the case dealing with m and ρ. When $\sigma_\xi^2 = 1$ (this problem has been investigated by

Luvsanceren in [8] and [9]), it has been shown that maximum likelihood estimates have an asymptotically normal distribution uniformly in the interval $-\infty < m < \infty, -1 < \rho < 1$ with a covariance matrix

$$\begin{pmatrix} \sigma_\xi \sqrt{\dfrac{1+\varrho}{2+(n-2)(1-\varrho)}} & 0 \\ 0 & \dfrac{1-\varrho^2}{\sqrt{(n-1)(1+\varrho^2)}} \end{pmatrix}.$$

The proof is based on the fact that the quantities $R_n^{(1)}$ and $R_n^{(2)}$ have asymptotically a normal distribution uniformly in the corresponding interval. However, when $\sigma_\xi^2 = 1$ (this case is closer to physical reality as well as to the continuous case) the uniform asymptotic normality of the distribution in the relevant interval $-\infty < m < \infty, -1 < \rho < 1$ does not hold (see below, §5). It should be mentioned that in a narrower interval $-\infty < m < \infty, -1 + \epsilon < \rho < 1 - \epsilon$ uniform asymptotic normality follows from general results (see, for example, Volkonskiĭ and Rozanov [13]).

The distinction between the cases $\sigma_\xi^2 = 1$ and $\sigma_\xi^2 = 1$ becomes clear when we compare the corresponding variances in (1.7) and (1.8).

§3. Distribution of the derivative of a liklihood function

To investigate the asymptotic behavior of maximum likelihood estimators, let us first investigate the properties of the distribution of the random vector $(R_n^{(1)}, R_n^{(2)}, R_n^{(3)})$ for $n \longrightarrow \infty$. We shall use the following notation:

$$\widetilde{R}_n^{(1)} = \frac{R_n^{(1)}}{\sqrt{\operatorname{Var} R_n^{(1)}}}, \quad \widetilde{R}_n^{(2)} = \frac{R_n^{(2)}}{\sqrt{\operatorname{Var} R_n^{(2)}}}, \quad \widetilde{R}_n^{(3)} = \frac{R_n^{(3)}}{\sqrt{\operatorname{Var} R_n^{(3)}}}.$$

THEOREM 3.1. *The characteristic function $f_n(t_1, t_2, t_3)$ of the random vector $(\widetilde{R}_n^{(1)}, \widetilde{R}_n^{(2)}, \widetilde{R}_n^{(3)})$ converges as $n \longrightarrow \infty$ for any t_1, t_2 and t_3 uniformly to the characteristic function of the normal distribution with expectation* (0, 0, 0) *and with correlation matrix*

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & -\varrho\sqrt{\dfrac{2}{1+\varrho^2}} \\ 0 & -\varrho\sqrt{\dfrac{2}{1+\varrho^2}} & 1 \end{pmatrix};$$

convergence is uniform when $-\infty < m < \infty$, $0 < \sigma_\xi^2 \leqslant K < \infty$ and $-1 < \rho < 1$, where K is an arbitrary fixed constant.

PROOF. The form of the matrix (1.3) and formulas (1.5) and (1.7) give

the following characteristic function of $(\widetilde{R}_n^{(1)}, \widetilde{R}_n^{(2)}, \widetilde{R}_n^{(3)})$:

$$f_n(t_1, t_2, t_3) = \mathsf{E} \exp\{it_1 \tilde{R}_n^{(1)} + it_2 \tilde{R}_n^{(2)} + it_3 \tilde{R}_n^{(3)}\}$$

$$(3.1) \qquad = c_n \int \ldots \int \exp\left\{-\frac{1}{2\sigma_\xi^2(1-\varrho^2)}(X-m)R_n^{-1}(X-m)^* + it_1 \tilde{R}_n^{(1)} + it_2 \tilde{R}_n^{(2)} + it_3 \tilde{R}_n^{(3)}\right\} dx_1 \ldots dx_n$$

$$= c_n \exp\left\{-it_2\sqrt{\frac{n}{2}} + it_3 \varrho\sqrt{\frac{n-1}{1+\varrho^2}}\right\} \int \ldots \int \exp\left\{-\frac{1}{2}[YA_nY^* - YA^*]\right\} dy_1 \ldots dy_n,$$

where

$$(3.2) \qquad c_n = (2\pi)^{-\frac{n}{2}} \sigma_\xi^{-n} (1-\varrho^2)^{-\frac{n-1}{2}},$$

$$Y = X - m,$$

$$(3.3) \qquad A_n = \begin{pmatrix} a_1 & b & 0 & 0 & \cdots & & 0 \\ b & a & b & 0 & \cdots & & 0 \\ 0 & b & a & b & \cdots & & 0 \\ \vdots & & & & \ddots & & \\ 0 & & & & & a & b \\ 0 & & & & & b & a_1 \end{pmatrix}$$

and

$$a_1 = \frac{1}{\sigma_\xi^2(1-\varrho^2)}\left[1 - \frac{2it_2}{\sqrt{2n}} + \frac{2it_3\varrho}{\sqrt{(n-1)(1+\varrho^2)}}\right],$$

$$a = \frac{1}{\sigma_\xi^2(1-\varrho^2)}\left[(1+\varrho^2)\left(1 - \frac{2it_2}{\sqrt{2n}}\right) + \frac{4it_3\varrho}{\sqrt{(n-1)(1+\varrho^2)}}\right],$$

$$(3.4) \qquad b = \frac{-1}{\sigma_\xi^2(1-\varrho^2)}\left[\varrho\left(1 - \frac{2it_2}{\sqrt{2n}}\right) + \frac{it_3(1+\varrho^2)}{\sqrt{(n-1)(1+\varrho^2)}}\right],$$

$$A^* = \begin{pmatrix} c \\ (1-\varrho)c \\ \vdots \\ (1-\varrho)c \\ c \end{pmatrix}, \quad c = \frac{2it_1}{\sigma_\xi\sqrt{(1+\varrho)[2+(n-2)(1-\varrho)]}}.$$

Let numbers d_i $(i = 1, \ldots, n)$ be chosen so that

$$\begin{aligned} a_1 d_1 + b d_2 &= \frac{c}{2}, \\ b d_3 + a d_2 + b d_1 &= (1-\varrho)\frac{c}{2}, \\ &\vdots \\ b d_{n-1} + a d_{n-2} + b d_{n-3} &= (1-\varrho)\frac{c}{2}, \\ b d_{n-1} + a_1 d_n &= \frac{c}{2}. \end{aligned} \tag{3.5}$$

Under the transformation $y_i = z_i - d_i$, the expression $YA_nY^* - Y\Lambda^*$ becomes

$$(z_1 - d_1, \ldots, z_n - d_n) A_n \begin{pmatrix} z_1 - d_1 \\ \vdots \\ z_n - d_n \end{pmatrix} - D_n, \tag{3.6}$$

where

$$\begin{aligned} D_n &= a_1 d_1^2 + a(d_2^2 + \ldots + d_{n-1}^2) + a_1 d_n^2 + 2b(d_1 d_2 + \ldots + d_{n-1} d_n) \\ &= \frac{c}{2}(d_1 + d_n) + \frac{(1-\varrho)c}{2} \sum_{2}^{n-1} d_i. \end{aligned} \tag{3.7}$$

The general solution of (3.5) is

$$d_i = d + \theta_1 u_1^i + \theta_2 u_2^i, \qquad (i = 1, \ldots, n), \tag{3.8}$$

where

$$d = \frac{(1-\varrho)c}{2(a+2b)}, \tag{3.9}$$

and u_1 and u_2 are zeros of the equation $bu^2 + au + b = 0$. The quantities θ_1 and θ_2 can be determined using the first and the last equations of (3.5):

$$\begin{aligned} \theta_1 &= \left(\frac{c}{2} - d(a_1 + b)\right) \frac{a_1 u_2^n + b u_2^{n-1} - (a_1 u_2 + b u_2^2)}{(a_1 u_1 + b u_1^2)(a_1 u_2^n + b u_2^{n-1}) - (a_1 u_2 + b u_2^2)(a_1 u_1^n + b u_1^{n-1})}, \\ \theta_2 &= \left(\frac{c}{2} - d(a_1 + b)\right) \frac{a_1 u_1 + b u_1^2 - (a_1 u_1^n + b u_1^{n-1})}{(a_1 u_1 + b u_1^2)(a_1 u_2^n + b u_2^{n-1}) - (a_1 u_2 + b u_2^2)(a_1 u_1^n + b u_1^{n-1})}. \end{aligned} \tag{3.10}$$

On the basis of a new solution, say $z_i - d_i = x_i$, the characteristic function (3.1) becomes

$$\begin{aligned} f_n(t_1, t_2, t_3) = c_n \exp\left\{-it_2\sqrt{\frac{n}{2}} + it_3\varrho\sqrt{\frac{n-1}{1+\varrho^2}} + \frac{D_n}{2}\right\}. \\ \cdot \int \ldots \int \exp\left\{-\frac{1}{2} X A_n X^*\right\} dx_1 \ldots dx_n. \end{aligned} \tag{3.11}$$

Since (see, for example, Cramér [5], p. 136)

$$\int \ldots \int \exp\left\{-\frac{1}{2} X A_n X^*\right\} dx_1 \ldots dx_n = (2\pi)^{n/2} |A_n|^{-1/2},$$

it follows that

$$f_n(t_1, t_2, t_3) = (2\pi)^{n/2} c_n |A_n|^{-1/2} \exp\left\{-it_2\sqrt{\frac{n}{2}} + it_3\varrho\sqrt{\frac{n-1}{1+\varrho^2}} + \frac{D_n}{2}\right\}. \tag{3.12}$$

From (3.3) it becomes apparent that

$$|A_n| = a_1 |\tilde{A}_{n-2}| - 2b^2 a_1 |\tilde{A}_{n-3}| + b^4 |\tilde{A}_{n-4}|, \tag{3.13}$$

where $|\tilde{A}_n|$ satisfies the difference equation

$$|\tilde{A}_n| = a|\tilde{A}_{n-1}| - b^2 |\tilde{A}_{n-2}|. \tag{3.14}$$

It can now easily be shown that

$$|\tilde{A}_n| = \alpha_1 v_1^n + \alpha_2 v_2^n, \tag{3.15}$$

where v_1 and v_2 are roots of the equation $v^2 - av + b^2 = 0$, while from the conditions $|\tilde{A}_1| = a$ and $|\tilde{A}_2| = a^2 - b^2$ we get

$$\alpha_1 = \frac{v_1}{v_1 - v_2}, \quad \alpha_2 = -\frac{v_2}{v_1 - v_2}. \tag{3.16}$$

By substitution in (3.15) and (3.13) we get the following two expressions:

$$|\tilde{A}_n| = \frac{v_1^{n+1} - v_2^{n+1}}{v_1 - v_2}, \tag{3.17}$$

$$\begin{aligned} |A_n| &= \frac{1}{v_1 - v_2}\left[v_1^{n-3}(a_1 v_1 - b^2)^2 - v_2^{n-3}(a_1 v_2 - b^2)^2\right] \\ &= \frac{v_1^{n-3}(a_1 v_2 - b^2)^2}{v_1 - v_2}\left[1 - \left(\frac{v_2}{v_1}\right)^{n-3} \frac{(a_1 v_2 - b^2)^2}{(a_1 v_2 - b^2)^2}\right]. \end{aligned} \tag{3.18}$$

To simplify calculations, let us write

$$f_n^{(1)}(t_1, t_2, t_3) = \exp\left\{\frac{D_n}{2}\right\},$$

$$f_n^{(2)}(t_1, t_2, t_3) = |A_n|^{-\frac{1}{2}} \sigma_\xi^{-n} (1-\varrho^2)^{-\frac{n-1}{2}} \exp\left\{-it_2\sqrt{\frac{n}{2}} + it_3\varrho\sqrt{\frac{n-1}{1+\varrho^2}}\right\},$$

and let us deal with the asymptotic behavior of the two functions separately. In what follows, M_i and $\bar{M}_i$ are constants, while the variables t_1, t_2 and t_3 belong to an arbitrary finite interval $T_1 \times T_2 \times T_3$ which is uniformly bounded for $\rho \in (-1, 1)$ and $\sigma_\xi^2 \in (0, K]$.

From (3.4) we obtain

(3.19)
$$v_1 = \frac{1}{\sigma_\xi^2(1-\varrho^2)}\left[(1+\varrho^2)\left(1-\frac{2it_2}{\sqrt{2n}}\right) + \frac{4it_3\varrho}{\sqrt{(n-1)(1+\varrho^2)}} + (1-\varrho^2)\sqrt{\left(1-\frac{2it_2}{\sqrt{2n}}\right)^2 + \frac{4t_3^2}{(n-1)(1+\varrho^2)}}\right],$$
$$v_2 = \frac{1}{\sigma_\xi^2(1-\varrho^2)}\left[(1+\varrho^2)\left(1-\frac{2it_2}{\sqrt{2n}}\right) + \frac{4it_2\varrho}{\sqrt{(n-1)(1+\varrho^2)}} - (1-\varrho^2)\sqrt{\left(1-\frac{2it_2}{\sqrt{2n}}\right)^2 + \frac{4t_3^2}{(n-1)(1+\varrho^2)}}\right].$$

For a sufficiently large n we obtain

$$\left(1-\frac{2it_2}{\sqrt{2n}}\right)^{-2} = 1 + \frac{4it_2}{\sqrt{2n}} + \frac{M_1}{n},$$

and

$$\left[1+\frac{4t_3^2}{(n-1)(1+\varrho^2)}\left(1+\frac{4it_2}{\sqrt{2n}}+\frac{M_1}{n}\right)\right]^{1/2} = 1 + \frac{2t_3^2}{(n-1)(1+\varrho^2)} + \frac{M_2}{n^{3/2}},$$

and therefore (3.19) can be written as

(3.19′)
$$v_1 = \frac{1}{\sigma_\xi^2(1-\varrho^2)}\left[1-\frac{2it_2}{\sqrt{2n}} + \frac{2it_3\varrho}{\sqrt{(n-1)(1+\varrho^2)}} + \frac{t_3^2(1-\varrho^2)}{(n-1)(1+\varrho^2)} + \frac{M_2}{n^{3/2}}\right],$$
$$v_2 = \frac{1}{\sigma_\xi^2(1-\sigma^2)}\left[1-\frac{2it_2}{\sqrt{2n}} + \frac{2it_3\varrho}{\sqrt{(n-1)(1+\varrho^2)}} - \frac{t_3^2(1-\varrho^2)}{(n-1)(1+\varrho^2)} + \frac{M_2}{n^{3/2}}\right].$$

Using the series $\log(1+x) = x - x^2/2 + \theta x^3/3$, where $|\theta| < 1$ for $|x| < ½$, on the basis of (3.19′) we obtain

$$v_1^{-\frac{n-1}{2}} = \sigma_\xi^{-(n-1)}(1-\varrho^2)^{-\frac{n-1}{2}} \exp\left\{-\frac{n-1}{2}\left[-\frac{2it_2}{\sqrt{2n}} + \frac{2\varrho it_3}{\sqrt{(n-1)(1+\varrho^2)}} + \frac{t_3^2(1-\varrho^2)}{(n-1)(1+\varrho^2)} - \frac{4\varrho t_2 t_3}{\sqrt{2n(n-1)(1+\varrho^2)}} + \frac{1}{2}\left(\frac{2t_2^2}{n} + \frac{4\varrho^2 t_3^2}{(n-1)(1+\varrho^2)}\right) + \frac{M_3'}{n^{3/2}}\right]\right\}$$

and therefore

(3.20)
$$v_1^{-\frac{n-1}{2}} = \sigma_\xi^{-(n-1)}(1-\varrho^2)^{-\frac{n-1}{2}} \exp\left\{it_2\frac{n-1}{\sqrt{2n}} - it_3\varrho\sqrt{\frac{n-1}{1+\varrho^2}} - \frac{t_3^2(1-\varrho^2)}{2(1+\varrho^2)} + \frac{2\varrho t_2 t_3}{\sqrt{\frac{2n}{n-1}(1+\varrho^2)}} - \frac{t_2^2}{2}\frac{n-1}{n} - \frac{\varrho^2 t_3^2}{1+\varrho^2}\right\}\left(1+\frac{M_3}{\sqrt{n}}\right).$$

It can easily be calculated that

$$a_1 v_1^2 - b^2 = \frac{1}{\sigma_\xi^4(1-\varrho^2)}\left[\left(1-\frac{2it_2}{\sqrt{2n}}\right)^2 + 2it_3\left(1-\frac{2it_2}{\sqrt{2n}}\right)\frac{\varrho}{\sqrt{(n-1)(1+\varrho^2)}} + \frac{M_4}{n}\right],$$

(3.21)
$$v_1 - v_2 = \frac{1}{\sigma_\xi^2}\left[1-\frac{2it_2}{\sqrt{2n}} + \frac{M_5}{n}\right],$$
$$v_1(v_1-v_2)^{1/2} = \frac{1}{\sigma_\xi^3(1-\varrho^2)}\left[1+\frac{M_6}{\sqrt{n}}\right],$$
$$a_1 v_2^2 - b^2 = \frac{1}{\sigma_\xi^4(1-\varrho^2)}\cdot\frac{M_7}{n}.$$

From (3.21), (3.20) and (3.18) we get

$$f_n^{(2)}(t_1, t_2, t_3) = \exp\left\{-it_2\sqrt{\frac{1}{2n}} - \frac{t_2^2}{2}\frac{n-1}{n} - \frac{t_3^2}{2} + \frac{2\varrho t_2 t_3}{\sqrt{2\frac{n}{n-1}(1+\varrho^2)}}\right\}\left(1+\frac{M_8}{\sqrt{n}}\right).$$

(3.22)

The asymptotic behavior of $f_n^{(1)}(t_1, t_2, t_3)$ can be discussed as follows. From (3.7)–(3.10) we obtain

$$D_n = \frac{c}{2}(2d + \theta_1 u_1 + \theta_2 u_2 + \theta_1 u_1^n + \theta_2 u_2^n) + \frac{(1-\varrho)c}{2}\left[(n-2)d + \theta_1 u_1^2\frac{1-u_1^{n-2}}{1-u_1} + \theta_2 u_2^2\frac{1-u_2^{n-2}}{1-u_2}\right] = \frac{c^2(1-\varrho)}{4(a+2b)}[2+(n-2)(1-\varrho)]$$
$$+ \frac{c^2}{4(a+2b)}[a+2b-(1-\varrho)(a_1+b)]\cdot g_n(t_1, t_2, t_3),$$

where

$$g_n(t_1, t_2, t_3) = \left\{\left[1+u_1^{n-1} + (1-\varrho)\frac{1-u_1^{n-2}}{1-u_1}\right]\frac{\theta_1}{\frac{c}{2}-d(a_1+b)}\right.$$
$$\left. + \left[1+u_2^{-(n-1)} + (1-\varrho)u_2^{-n+2}\frac{1-u_2^{n-2}}{1-u_2}\right]\frac{\theta_2}{\frac{c}{2}-d(a_1+b)}\right. .$$

Thus

$$D_n = -t_1^2\left(1+\frac{\bar{M}_1}{\sqrt{n}}\right) - \frac{t_1^2}{2+(1-\varrho)(n-2)}\left(1+\frac{\bar{M}_1}{\sqrt{n}}\right)\frac{2it_3}{\sqrt{(n-1)(1+\varrho^2)}}\cdot \tag{3.23}$$

$$\cdot\frac{1}{\sigma_\xi^2(1+\varrho)}\cdot g_n(t_1, t_2, t_3).$$

Because u_1 and u_2 satisfy the relations

$$u_1 = -\frac{v_2}{b}, \quad u_2 = -\frac{v_1}{b}, \quad u_1, u_2 = 1, \quad |u_1| \leqq 1, \quad |u_2| \geqq 1 \tag{3.24}$$

it is easy to show that

$$a_1u_2+b = \frac{1}{b}\frac{1}{\sigma_\xi^2(1-\varrho^2)}\left[\left(1-\frac{2it_2}{\sqrt{2n}}\right)^2 + 2it_3\left(1-\frac{2it_2}{\sqrt{2n}}\right)\frac{\varrho}{\sqrt{(n-1)(1+\varrho^2)}}+\frac{\bar{M}_2}{n}\right],$$

$$a_1u_1+b = \frac{1}{b}\frac{1}{\sigma_\xi^4(1-\varrho^2)}\cdot\frac{\bar{M}_3}{n},$$

$$a_1+bu_2 = a_1-v_1 = \frac{\bar{M}_4}{\sigma_\xi^2\cdot n}, \tag{3.25}$$

$$a_1+bu_1 = a_1-v_2 = \frac{1}{\sigma_\xi^2}\left[1-\frac{2it_2}{\sqrt{2n}}+\frac{\bar{M}_5}{n}\right],$$

$$\frac{v_2}{v_1} = \left[\varrho-\frac{\bar{M}_6}{\sqrt{n}}\right],$$

$$b = \frac{1}{\sigma_\xi^2(1-\varrho^2)}\left(-\varrho+\frac{\bar{M}_7}{\sqrt{n}}\right), \quad \text{where} \quad \bar{M}_7 \neq 0 \quad \text{for} \quad \varrho = 0.$$

Let the uniformly bounded quantity $1/\sigma_\xi^4(1-\rho^2)(b^2-v_1a_1)$ be denoted by $\bar{M}_8$ and the quantity $n\sigma_\xi^4(1-\rho^2)(b^2-v_2a_1)$ by $\bar{M}_3$; then

$$\frac{(a_1u_2+b)-(a_1+bu_2)u_2^{-(n-1)}}{(a_1+bu_1)(a_1u_2+b)-\frac{u_1^{n-2}}{u_2^{n-2}}(a_1+bu_2)(a_1u_1+b)} \tag{3.26}$$

$$= \sigma_\xi^2\frac{1-\bar{M}_{10}\left(-\varrho+\frac{\bar{M}_7}{\sqrt{n}}\right)\left(\frac{\bar{M}_4}{n}\right)\bar{M}_8}{1-\frac{\bar{M}_5}{\sqrt{n}}-\bar{M}_9\frac{\bar{M}_4}{n}\cdot\frac{\bar{M}_3}{n}\cdot\bar{M}_8} = \sigma_\xi^2\cdot\bar{M}_{11},$$

where

$$\left|\frac{u_1}{u_2}\right|^{n-1} = \bar{M}_9 \leqq 1, \quad |u_2|^{-(n-1)} = \bar{M}_{10} \leqq 1.$$

Similarly

$$u_2\frac{a_1+bu_1-(a_1u_1+b)u_1^{n-2}}{(a_1+bu_1)(a_1u_2+b)-\frac{u_1^{n-2}}{u_2^{n-2}}(a_1+bu_2)(a_1u_1+b)} = \sigma_\xi^2\cdot\bar{M}_{14},$$

which follows from the relations

$$|u_1| = \bar{M}_{13} \leqq 1, \quad v_1 = \frac{1}{\sigma_\xi^2(1-\varrho^2)}\left[1+\frac{\bar{M}_{12}}{\sqrt{n}}\right].$$

From equations (3.24) for u_1 and u_2 we obtain the quantities

$$\frac{1}{1+\varrho}\frac{1+u_1^{n-1}+(1-\varrho)u_1\frac{1-u_1^{n-2}}{1-u_1}}{2+(n-2)(1-\varrho)} = \bar{M}_{15}, \tag{3.27}$$

$$\frac{1}{1+\varrho}\frac{1+u_2^{-(n-1)}+(1-\varrho)u_2^{-n-2}\frac{1-u_2^{n-2}}{1-u_2}}{2+(n-2)(1-\varrho)} = \bar{M}_{16},$$

which are uniformly bounded. Thus from (3.27), (3.26) and (3.23) we get

$$D_n = -t_1^2\left(1+\frac{\bar{M}_1}{\sqrt{n}}\right) \tag{3.28}$$

$$-t_1^2\left(1+\frac{\bar{M}_1}{\sqrt{n}}\right)\frac{2it_3}{\sqrt{(n-1)(1+\varrho^2)}}[\bar{M}_{15}\cdot\bar{M}_{11}+\bar{M}_{16}\cdot\bar{M}_{14}],$$

and finally

$$f_n^{(1)}(t_1, t_2, t_3) = \exp\left\{\frac{D_n}{2}\right\} = \exp\left\{-\frac{t_1^2}{2}\right\}\left(1+\frac{\bar{M}_{17}}{\sqrt{n}}\right). \tag{3.29}$$

From (3.12), (3.22) and (3.29) we obtain the following result for the characteristic function of $(\tilde{R}_n^{(1)}, \tilde{R}_n^{(2)}, \tilde{R}_n^{(3)})$:

$$f_n(t_1, t_2, t_3) = f_n^{(1)}(t_1, t_2, t_3)\cdot f_n^{(2)}(t_1, t_2, t_3) \tag{3.30}$$

$$= \exp\left\{-\frac{t_1^2}{2}-\frac{t_2^2}{2}-\frac{t_3^2}{2}+\frac{2\varrho t_2t_3}{\sqrt{2(1+\varrho^2)}}\right\}\left(1+\frac{M}{\sqrt{n}}\right).$$

Hence as $n \longrightarrow \infty$ the functions $f_n(t_1, t_2, t_3)$ converge in any finite interval of the Cartesian product $T_1 \times T_2 \times T_3$ to the characteristic function of the normal distribution, uniformly in $-\infty < m < \infty$, $-1 < \rho < 1$, $0 < \sigma_\xi^2 \leqslant K < \infty$. The proof of the theorem is complete.

By rearranging the corresponding formulas the rate of convergence can also be determined.

§4. Asymptotic distribution of maximum likelihood estimates

Using Theorem 3.1, we can now investigate the asymptotic behavior of the maximum likelihood estimates as given by the solutions of the equations

$$R_n^{(1)}=0, \quad R_n^{(2)}=0, \quad R_n^{(3)}=0. \tag{4.1}$$

Convergence of distributions is understood in the following as weak convergence (see Gnedenko and Kolmogorov [6]). We now prove

THEOREM 4.1. *The system of equations* (4.1) *has almost surely as* $n \to \infty$ *a solution* $\hat{m}(\xi_1, \ldots, \xi_n)$, $\hat{\sigma}_\xi^2(\xi_1, \ldots, \xi_n)$, $\hat{\rho}(\xi_1, \ldots, \xi_n)$ *such that the distribution of the random vector*

$$(\hat{m}_n - m)\sqrt{\frac{2+(n-2)(1-\varrho)}{\sigma_\xi^2(1+\varrho)}}, \quad (\hat{\sigma}_{\xi,n}^2 - \sigma_\xi^2)\frac{1}{\sigma_\xi^2}\sqrt{\frac{n}{2}}, \quad (\hat{\varrho}_n - \varrho)\frac{\sqrt{(n-1)(1+\varrho^2)}}{1-\varrho^2}.$$

converges to the distribution of the random vector $(\widetilde{R}_n^{(1)}, \widetilde{R}_n^{(2)}, \widetilde{R}_n^{(3)})$ *as* $n \to \infty$, *and this convergence is uniform in the region* $-\infty < m < \infty$, $0 < \sigma_\xi^2 \leqslant K < \infty$, $-1 < \rho < 1$.

PROOF. Let us take

$$\begin{aligned} L_n^{(1)} &= (1+\varrho)\sigma_\xi^2 R_n^{(1)} = 0, \\ L_n^{(2)} &= 2(1-\varrho)\sigma_\xi^4 R_n^{(2)} = 0, \\ L_n^{(3)} &= (1-\varrho^2)\sigma_\xi^2 R_n^{(3)} = 0, \end{aligned} \tag{4.2}$$

which is evidently equivalent to the system (4.1). The left-hand sides in (4.2) are polynomials in the variables m, σ_ξ^2 and ρ; their Taylor series about the true values m_0, $\sigma_{\xi,0}^2$ and ρ_0 of the parameters are as follows:

$$\begin{aligned} &L_n^{(1)}(m_0, \sigma_0^2, \varrho_0) + \left.\frac{\partial L_n^{(1)}}{\partial m}\right|_{m_0,\sigma_0^2,\varrho_0}(m-m_0) + \left.\frac{\partial L_n^{(1)}}{\partial \varrho}\right|_{m_0,\sigma_0^2,\varrho_0}(\varrho-\varrho_0) + \ldots = 0, \\ &L_n^{(2)}(m_0, \sigma_0^2, \varrho_0) + \left.\frac{\partial L_n^{(2)}}{\partial m}\right|_{m_0,\sigma_0^2,\varrho_0}(m-m_0) + \ldots = 0, \\ &L_n^{(3)}(m_0, \sigma_0^2, \varrho_0) + \left.\frac{\partial L_n^{(3)}}{\partial m}\right|_{m_0,\sigma_0^2,\varrho_0}(m-m_0) + \ldots = 0. \end{aligned} \tag{4.3}$$

We could calculate with no difficulty the derivatives in (4.3), but we shall not write down the corresponding formulas. With the substitution

$$m - m_0 = x\sqrt{\frac{\sigma_0^2(1+\varrho_0)}{2+(n-2)(1-\varrho_0)}},$$

$$\sigma_\xi^2 - \sigma_0^2 = y\sigma_0^2\sqrt{\frac{2}{n}},$$

$$\varrho - \varrho_0 = z\frac{1-\varrho_0^2}{\sqrt{(n-1)(1+\varrho_0^2)}},$$

and after dividing the first equation (4.3) by $\sigma_0\sqrt{1+\rho_0}\sqrt{2+(n-2)(1-\rho_0)}$, the second by $(1-\rho_0^2)\sigma_0^2\sqrt{2n}$ and the third by $(1-\rho_0^2)\sigma_0^2\sqrt{(n-1)(1+\rho_0^2)}$, we obtain the following system:

$$\begin{aligned} &\tilde{R}_n^{(1)} + x\left.\frac{\partial L_n^{(1)}}{\partial m}\right|_0 \frac{1}{2+(n-2)(1-\varrho_0)} + \ldots = 0, \\ &\tilde{R}_n^{(2)} + y\left.\frac{\partial L_n^{(2)}}{\partial \sigma_\xi^2}\right|_0 \frac{1}{n(1-\varrho_0^2)} + \ldots = 0, \\ &\tilde{R}_n^{(3)} + z\left.\frac{\partial L_n^{(3)}}{\partial \varrho}\right|_0 \frac{1}{\sigma_0^2(n-1)(1+\varrho_0^2)} + \ldots = 0. \end{aligned} \tag{4.4}$$

It can easily be shown that

$$\left.\frac{\partial L_n^{(1)}}{\partial m}\right|_0 \cdot \frac{1}{2+(n-2)(1-\varrho_0)} = -1,$$

$$\left.\frac{\partial L_n^{(2)}}{\partial \sigma_\xi^2}\right|_0 \cdot \frac{1}{n(1-\varrho_0^2)} = -1,$$

$$\left.\frac{\partial L_n^{(3)}}{\partial \varrho}\right|_0 \cdot \frac{1}{\sigma_0^2(n-1)(1+\varrho_0^2)} \to -1,$$

where convergence should be understood as convergence almost surely as well as uniform convergence on the set $-\infty < m_0 < \infty$, $0 < \sigma_0^2 \leqslant K < \infty$, $-1 < \rho_0 < 1$. The remaining terms in (4.4) tend to zero almost surely and uniformly on the above-mentioned set in the true-parameter region.

For large n the quantities $|R_n^{(i)}|$ and $|\widetilde{R}_n^{(i)}|/|R_n^{(j)}|$, $i, j = 1, 2, 3$, are bounded in probability and uniformly in the same set, as follows from Theorem 3.1. Therefore the system

$$
(4.5)\qquad \begin{aligned} 1-\varepsilon_1+\ldots &= 0,\\ 1-\varepsilon_2+\ldots &= 0,\\ 1-\varepsilon_3+\ldots &= 0, \end{aligned}
$$

where

$$
\varepsilon_1=\frac{x}{\tilde R_n^{(1)}},\quad \varepsilon_2=\frac{y}{\tilde R_n^{(2)}},\quad \varepsilon_3=\frac{z}{\tilde R_n^{(3)}},
$$

for large n with probability arbitrarily close to 1 has a solution $\epsilon_1^{(n)}, \epsilon_2^{(n)}, \epsilon_3^{(n)}$ which uniformly on the set $-\infty < m_0 < \infty$, $0 < \sigma_0^2 \leqslant K < \infty$, $-1 < \rho < 1$ belongs to the interval $(1-\delta, 1+\delta)$ (with arbitrary δ). Hence the limit distribution of the variables

$$
x_n=\varepsilon_1^{(n)}\cdot\tilde R_n^{(1)},\quad y_n=\varepsilon_2^{(n)}\cdot\tilde R_n^{(2)},\quad z=\varepsilon_3^{(n)}\cdot\tilde R_n^{(3)}
$$

for $n\longrightarrow\infty$ coincides with the distributions of $\tilde R_n^{(1)}$, $\tilde R_n^{(2)}$ and $\tilde R_n^{(3)}$, because $\epsilon_i^{(n)}=1$ $(i=1,2,3)$ as $n\longrightarrow\infty$ uniformly in the region $-\infty < m_0 < \infty$, $0 < \sigma_0^2 \leqslant K < \infty$, $-1 < \rho_0 < 1$. The proof of Theorem 4.1 is complete.

The relations

$$
x_n=(\hat m_n-m_0)\sqrt{\frac{2+(n-2)(1-\varrho_0)}{\sigma_0^2(1-\varrho_0)}},
$$

$$
y_n=(\hat\sigma_n^2-\sigma_0^2)\frac{1}{\sigma_0^2}\sqrt{\frac{n}{2}},
$$

$$
z_n=(\hat\varrho_n-\varrho_0)\frac{\sqrt{(n-1)(1+\varrho_0^2)}}{1-\varrho_0^2}
$$

show that the estimators $\hat\rho_n$ and $\hat\sigma_n^2$ are uniformly consistent, which is not true for $\hat m_n$.

Using Theorem 4.1 and Cramér's theorem ([5], p. 281), confidence intervals for m, σ_ξ^2 and ρ can be constructed.

§5. Results obtained for discrete analogues of the continuous-time case

As we mentioned above, we have a case corresponding to the continuous-time one when the parameters m, σ_ξ^2 and ρ are unknown. Here the assertion concerning uniform asymptotic normality of quantities $H_n^{(i)}$ (see (1.6)) is not true. Nevertheless, two theorems concerning $H_n^{(i)}$ can be proved; these theorems correspond to Theorems 3.1 and 4.1 when $\kappa=(1-\rho^2)n\longrightarrow\infty$.

THEOREM 5.1. *The distribution of the random variables* $\tilde H_n^{(i)} = H_n^{(i)}/\mathrm{Var}\,H_n^{(i)}$ $(i=1,2,3)$ *tends to the normal distribution with parameters*

$$
\left(0,0,0;\ \begin{vmatrix}1&0&0\\0&1&0\\0&0&1\end{vmatrix}\right)
$$

as $\kappa\longrightarrow\infty$.

The proof proceeds in the same manner as in §3. The quantities v_1 and v_2 in (3.19) will, however, have the following form:

$$
(5.1)\qquad v_1=\frac{1}{2\sigma_\zeta^2}\left[(1+\varrho^2)\left(1-\frac{2it_2}{\sqrt{2(n-1)}}\right)+\frac{it_3 2\varrho\sqrt{1-\varrho^2}}{\sqrt{n-1}}\right.
$$

$$
\left.+(1-\varrho^2)\left\{\left(1-\frac{2it_2}{\sqrt{2(n-1)}}\right)^2-\frac{4i\varrho t_3}{\sqrt{(n-1)(1-\varrho^2)}}\left(1-\frac{2it_2}{\sqrt{2(n-1)}}\right)+\frac{4(1-\varrho^2)t_3^2}{n-1}\right\}^{1/2}\right],
$$

$$
v_2=\frac{1}{2\sigma_\zeta^2}\left[(1+\varrho^2)\left(1-\frac{2it_2}{\sqrt{2(n-1)}}\right)+\frac{it_3 2\varrho\sqrt{1-\varrho^2}}{\sqrt{n-1}}\right.
$$

$$
\left.-(1-\varrho^2)\left\{\left(1-\frac{2it_2}{\sqrt{2(n-1)}}\right)^2-\frac{4i\varrho t_3}{\sqrt{(n-1)(1-\varrho^2)}}\left(1-\frac{2it_2}{\sqrt{2(n-1)}}\right)+\frac{4(1-\varrho^2)t_3^2}{n-1}\right\}^{1/2}\right].
$$

When considering (3.20) it can be seen from (5.1) that the normality of the asymptotic distribution holds only for $\kappa\longrightarrow\infty$.

For $\kappa\longrightarrow\infty$ we obtain estimators equivalent to the maximal likelihood estimators when considering the following:

$$
(5.2)\qquad \hat m=\frac{1}{n}\sum_1^n\xi(k),\quad \hat\sigma_\xi^2=(1-\hat\varrho^2)\hat s_\xi^2,
$$

$$
\hat\varrho=\frac{1}{(n-1)\hat s_\xi^2}\sum_2^n\eta(k)\eta(k-1),
$$

where

$$
(5.3)\qquad \eta(k)=\xi(k)-\hat m,\quad \hat s_\xi^2=\frac{1}{n}\sum_{k=1}^n\eta^2(k).
$$

Simple calculations give us

$$
(5.4)\qquad \begin{aligned}
&\mathsf{E}\hat m=m,\quad \mathrm{Var}\,\hat m=\frac{\sigma_\xi^2}{n}+\frac{2\varrho}{1-\varrho}\frac{\sigma_\xi^2}{n}+o\left(\frac{1}{n}\right),\\
&\mathsf{E}(\hat\varrho-\varrho)=O\left(\frac{1}{n}\right),\quad \mathrm{Var}\,\hat\varrho=\frac{1-\varrho^2}{n}+o\left(\frac{1}{n}\right),\\
&\mathsf{E}(\hat\sigma_\zeta^2-\sigma_\zeta^2)=O\left(\frac{1}{n}\right),\quad \mathrm{Var}\,\sigma_\zeta^2=\frac{2(1-\varrho^2)^2}{n}\sigma_\xi^4+o\left(\frac{1}{n}\right).
\end{aligned}
$$

Similarly, the following theorem can be proved:

THEOREM 5.2. *As $\kappa \to \infty$, the estimators $\hat{m} \sim m$, $\hat{\sigma}_\xi^2 \sim \sigma_\xi^2$ and $\hat{\rho} \sim \rho$ are asymptotically efficient, and the distribution of the random vector*

$$\frac{\hat{m}-m}{\sqrt{\frac{1+\varrho}{1-\varrho}\frac{\sigma_\xi^2}{n}}}, \quad \frac{\hat{\sigma}_\zeta^2-\sigma_\zeta^2}{\sqrt{\frac{2(1-\varrho^2)^2\sigma_\xi^2}{n}}}, \quad \frac{\hat{\varrho}-\varrho}{\sqrt{\frac{1-\varrho^2}{n}}} \tag{5.5}$$

tends to the normal distribution with parameters

$$\left(0, 0, 0, \begin{Vmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{Vmatrix}\right).$$

The following theorem can be proved in an even easier manner.

THEOREM 5.3. *As $n \to \infty$ the estimator $\hat{\sigma}_\xi^2 \sim \sigma_\xi^2$ obtained from (5.2) is asymptotically efficient, and the distribution of the ratio*

$$\frac{\hat{\sigma}_\zeta^2-\sigma_\zeta^2}{\sigma_\zeta^2\sqrt{\frac{2}{n}}} \tag{5.6}$$

tends to (0, 1) *normal distribution.*

PROOF. We have seen in §2, (2.3), that the characteristic function of the random variable

$$\zeta_n = -\sqrt{\frac{n}{2}}+\frac{1}{2\sigma_\zeta^2}\sqrt{\frac{2}{n}}\Big\{(1-\varrho^2)(\xi(1)-m)^2 + \sum_2^n [\xi(i)-m-\varrho(\xi(i-1)-m)]^2\Big\} \tag{5.7}$$

has the form

$$\left(1-\frac{2it}{\sqrt{2n}}\right)^{-\frac{n}{2}} \exp\left(-it\sqrt{\frac{n}{2}}\right).$$

Hence ζ_n is asymptotically normally distributed as $n \to \infty$. On the other hand, $\hat{\rho} \to \rho$ and $\hat{m} \to m$ in probability, and therefore according to Cramér's theorem [5] the asymptotic distribution of the random variable

$$\zeta_n^* = -\sqrt{\frac{n}{2}}+\frac{1}{2\sigma_\zeta^2}\sqrt{\frac{2}{n}}\Big\{(1-\hat{\varrho}^2)(\xi(1)-\hat{m})+\frac{1}{2\sigma_\zeta^2}\sqrt{\frac{2}{n}}\sum_2^n[\xi(i)-\hat{m}- -(\xi(i-1)-\hat{m})\hat{\varrho}]^2\Big\} = -\sqrt{\frac{n}{2}}+\frac{1}{2\sigma_\zeta^2}\sqrt{\frac{2}{n}}(1-\hat{\varrho}^2)\hat{s}_\xi^2+\frac{(\xi(n)-\hat{m})^2+(\xi(1)-\hat{m})^2}{\sqrt{2n}\cdot\sigma_\zeta^2}$$

coincides with that of ζ_n for $n \to \infty$. The solution of the equation $\zeta_n^* = 0$ yields the estimator $\hat{\sigma}_\xi^2$ when one neglects a term of order $O(1/\sqrt{n})$, which completes the proof.

Theorems 3.3 and 3.4 in the previous paper [2] remain true also in the discrete case when parameters m, σ_ξ^2 and ρ are considered to be unknown. To verify this we only need to prove that continuous functionals of trajectories of a discrete stationary Gaussian Markov process converge in probability to functionals of trajectories of a process continuous at the corresponding points, and that this convergence is uniform in the parameter space.

Let $\xi_n(t)$ $(0 \leqslant t \leqslant T)$ be the polygonal function associated with the process $\xi(t)$, i.e.

$$\xi_n(t) = \xi\left(\frac{kT}{n}\right)+\frac{n}{kT}\left(t-\frac{kT}{n}\right)\left[\xi\left(\frac{k+1}{n}T\right)-\xi\left(\frac{kT}{n}\right)\right], \tag{5.8}$$

for

$$\frac{kT}{n} \leqq t \leqq \frac{k+1}{n}T, \qquad k=0, 1, \ldots, n-1.$$

The following lemma is true.

LEMMA 5.1. *Let $\xi(t)$ be a stationary Gaussian Markov process. Then uniformly in $-\infty < m < \infty$, $0 < \lambda \leqslant \lambda_0$ (and $2\lambda\sigma_\xi^2 = \sigma_\zeta^2 =$ constant) the following inequality is valid:*

$$P\{\sup_{|t'-t''|<\delta} |\xi(t')-\xi(t'')|>\varepsilon\} \leqq \frac{2\sigma_\zeta^2\delta+\sigma_\zeta^2\cdot\lambda_0\delta^2}{\varepsilon^2}. \tag{5.9}$$

PROOF. Formula (1.1) in [2] shows that

$$\xi(t')-\xi(t'') = -\lambda\int_{t''}^{t'}\xi(s)\,ds+\int_{t''}^{t'}d\zeta(s)$$

and therefore

$$\mathsf{E}\{\sup_{|t'-t''|<\delta}|\xi(t')-\xi(t'')|^2\} \leqq 2\,\mathsf{E}\left\{\sup_{|t'-t''|<\delta}\left|\lambda\int_{t''}^{t'}\xi(s)\,ds\right|^2\right\}$$

$$(5.10)\qquad +2\,\mathsf{E}\left\{\sup_{|t'-t''|<\delta}\left|\int_{t''}^{t'}d\zeta(s)\right|^2\right\}$$

$$\leqq 2\delta\int_{t''}^{t'}\mathsf{E}(\lambda\xi(s))^2\,ds+2\sigma_\zeta^2\cdot\delta \leqq \sigma_\zeta^2\cdot\lambda\cdot\delta^2+2\sigma_\zeta^2\cdot\delta.$$

Using the Čebyšev inequality, we obtain (5.9) from (5.8). The case of $\lambda \longrightarrow \infty$ would require separate discussion; but, as the results in [2] show, in this case confidence intervals can be constructed, and therefore we can now leave it aside.

Lemma 5.1 ensures that the assumptions of Maruyama's theorem [10] are satisfied; this means that the following theorem is valid:

THEOREM 5.4. *Let $\xi(t)$ be a stationary Gaussian Markov process, and let $\xi_n(t)$ be the corresponding polygonal function* (5.8); *further, let $f(t)$ and $g(t)$ be continuous functions on the interval $0 \leqslant t \leqslant T$ such that $f(0) < \xi(0) < g(0)$. Then*

$$(5.11)\qquad \lim_{n\to\infty} P\{f(t)\leqq\xi_n(t)\leqq g(t),\ 0\leqq t\leqq T\}=P\{f(t)\leqq\xi(t)\leqq g(t),\ 0\leqq t\leqq T\}$$

uniformly in $-\infty < m < \infty$, $0 < \lambda \leqslant \lambda_0$.

From Theorem 5.4 immediately follows

COROLLARY. *Let $\xi(t)$ be a stationary Gaussian Markov process, let $\overline{h}(\xi(t))$ and $\underline{h}(\xi(t))$ $(0 \leqslant t \leqslant T)$ be continuous functionals and let ϵ be a positive real number such that*

$$(5.12)\qquad P\{\underline{h}(\xi(t)) < m < \overline{h}(\xi(t))\} > 1-\varepsilon.$$

Then for any $\epsilon_1 > 0$ there exists (uniformly for $-\infty < m < \infty$ and $0 < \lambda \leqslant \lambda_0$) an integer n, which depends on ϵ and ϵ_1 only, such that

$$(5.13)\qquad P\{\underline{h}(\xi_n(t)) < m < \overline{h}(\xi_n(t))\} > 1-\varepsilon-\varepsilon_1.$$

This result, considered in conjunction with Theorem 3.3 of the previous paper, means that in the discrete case no finite confidence interval can be constructed.

§6. The problem of condensing observation points

We have seen in §2 that the maximum likelihood estimator for m is the weighted average of the statistics $\xi(1) + \xi(n)$ and $\Sigma_2^{n-1}\,\xi(i)$. We could therefore expect that in estimating the expected value of the process $\xi(t)$, observed in the interval $[0, T]$, by

$$(6.1)\qquad \overline{m}_n = \frac{1}{n+1}\sum_{k=0}^{n}\xi\left(\frac{Tk}{n}\right)$$

a value n exists such that the variance of $\overline{m}_n$ is minimal. But condensation of observation points over all limits is of no use. This can be illustrated as follows: Let $T = 1$ and $\sigma_\xi^2 = 1$, and let m be estimated by $m_1 = (\xi(0) + \xi(1))/2$ and by $m_2 = \int_0^1 \xi(t)\,dt$. It can now be clearly seen that (see, for example, formula (2.15) in [2])

$$(6.2)\qquad \mathrm{Var}\ m_1 = \frac{1+e^{-\lambda}}{4\lambda},\quad \mathrm{Var}\ m_2 = \frac{\lambda+e^{-\lambda}-1}{\lambda^3}.$$

We have $\mathrm{Var}\ m_1 < \mathrm{Var}\ m_2$ for $0 < \lambda < 2$ and $\mathrm{Var}\ m_2 < \mathrm{Var}\ m_1$ for $\lambda > 2$. Therefore, when $T = 1$, $\sigma_\xi^2 = 1$ and $\lambda < 2$, the estimator m_1 is better than m_2 (both are equally good for $\lambda = 2$).

The problem of estimating the variance of the process can be put in a similar way. Let us assume $m = 0$, and let us estimate the variance by

$$(6.3)\qquad s_n^2 = \frac{1}{n+1}\sum_{k=0}^{n}\xi^2\left(\frac{kT}{n}\right).$$

What is the value of n with the minimum variance in this context? Let us again take $T = 1$ and $\sigma_\xi^2 = 1$, and let us compare the variances of

$$(6.4)\qquad s_1^2 = \frac{\xi^2(0)+\xi^2(1)}{2},\quad s^2 = \int_0^1 \xi^2(t)\,dt.$$

We easily obtain

$$(6.5)\qquad \begin{aligned}\mathrm{Var}\ s_1^2 &= \sigma^4(1+e^{-2\lambda}) = \sigma^4\left(1+e^{-\frac{1}{\sigma^2}}\right)\\ \mathrm{Var}\ s^2 &= \frac{\sigma^4}{\lambda^2}(e^{-2\lambda}+2\lambda-1) = 4\sigma^8\left(e^{-\frac{1}{\sigma^2}}+\frac{1}{\sigma^2}-1\right).\end{aligned}$$

Simple calculations show that $\mathrm{Var}\ s_1^2 < \mathrm{Var}\ s^2$ for $0 < \lambda < 1$ (or for $\sigma^2 > ½$). Difficulties arise here, since σ^2 is also an unknown parameter, and, moreover, the variables s_1^2 and s^2 do not have normal distributions; therefore the above estimates can only be used as approximations.

When we have to choose, for a given λ, the length of the interval for which the extimator

$$\overline{m}_1 = \frac{\xi(0)+\xi(T)}{2}$$

or the estimator

$$\overline{m}_2 = \frac{\xi(0) + \xi\left(\frac{T}{2}\right) + \xi(T)}{3}$$

is better, the corresponding variances

$$\operatorname{Var} \overline{m}_1 = \frac{\sigma^2}{2}(1+e^{-\lambda T})$$

$$\operatorname{Var} \overline{m}_2 = \frac{\sigma^2}{3}\left(1+\frac{2}{3}e^{-\lambda T}+\frac{4}{3}e^{-\lambda T/2}\right)$$

give us the following inequality for T:

(6.6) $$3+5e^{-\lambda T} \leq 8e^{-\frac{\lambda T}{2}}.$$

In the general case, let $R(\tau)$ be a twice differentiable correlation function of the process $\xi(t)$, and let its second derivative be bounded on the interval $(0, T)$. Then the following theorem holds true.

THEOREM 6.1. *If, in addition, the correlation function $R(\tau)$ of the process $\xi(t)$ satisfies the condition*

$$\int_0^T R(\tau)\left(1-\frac{2\tau}{T}\right) d\tau > 0,$$

then among the estimates (6.1) *there exists one* (*for finite n*) *which is of minimum variance.*

It would be useful to find similar conditions also for variance estimates.

BIBLIOGRAPHY

1. T. W. Anderson, *On asymptotic distributions of estimates of parameters of stochastic difference equations*, Ann. Math. Statist. **30** (1959), 676–687. MR **21** #6072.

2. Mátyás Arató, *On the statistical examination of continuous state Markov processes*. I, Magyar Tud. Akad. Mat. Fiz. Oszt. Közl. **14** (1964), 13–34; English transl., Selected Transl. Math. Statist. and Probability, vol. 14, Amer. Math. Soc., Providence, R. I., 1978, pp. 203–225. MR 37 #3645.

3. ———, *Some statistical questions on stationary Gaussian Markov processes*, Dissertation, Moscow State Univ., Moscow, 1962. (Russian)

4. ———, *Estimation of the parameters of a stationary Gaussian Markov process*, Dokl. Akad. Nauk SSSR **145** (1962), 13–16 = Soviet Math. Dokl. **3** (1962), 905–909. MR **26** #834.

5. Harold Cramér, *Mathematical methods of statistics*, Princeton Univ. Press, Princeton, N. J., 1946. MR **8**, 39.

6. B. V. Gnedenko and A. N. Kolmogorov, *Limit distributions for sums of independent random variables*, GITTL, Moscow, 1949; English transl., Addison-Wesley, Reading, Mass., 1954; rev. ed., 1968. MR **12**, 839; **16**, 52; **38** #1722.

7. Ju. V. Linnik, *On a question of the statistics of dependent events*, Izv. Akad. Nauk SSSR Ser. Mat. **14** (1950), 501–522. (Russian) MR **12**, 512.

8. Š. Luvsanceren, *Maximum likelihood estimates and confidence regions for unknown parameters of a stationary Gaussian process of Markov type*, Candidate's Dissertation, Moscow State Univ., Moscow, 1954. (Russian) RŽ Mat. **1954** #4525.

9. ———, *Maximum likelihood estimates and confidence regions for unknown parameters of a stationary Gaussian process of Markov type*, Dokl. Akad. Nauk SSSR **98** (1954), 723–726. (Russian) MR **16**, 385.

10. Gisirô Maruyama, *Continuous Markov processes and stochastic equations*, Rend. Circ. Mat. Palermo (2) **4** (1955), 48–90. MR **17**, 166.

11. Murray Rosenblatt, *Random processes*, Oxford Univ. Press, New York, 1962. MR **24** #A3686.

12. S. Ja. Vilenkin, *On estimating the mean of stationary processes*, Teor. Verojatnost. i Primenen. **4** (1959), 451–453 = Theor. Probability Appl. **4** (1959), 415–416.

13. V. A. Volkonskiĭ and Ju. A. Rozanov, *Some limit theorems for random functions*. I, Teor. Verojatnost. i Primenen **4** (1959), 186–207 = Theor. Probability Appl. **4** (1959), 178–197. MR **21** #4477.

Translated by J. GREGOR