

## BINARY STATISTICAL EXPERIMENTS WITH PERSISTENT NONLINEAR REGRESSION

UDC 519.21

D. V. KOROLIUK

ABSTRACT. A sequence of binary stochastic experiments with persistent nonlinear regression is considered. The regression is defined as a product of a linear directed force and nonlinear term changing the directed force in a neighborhood of boundary points. A stochastic approximation for the sequence of stationary experiments is constructed with the help of an autoregressive process with normal perturbation. A stochastic approximation of the sequence of exponential stochastic experiments is also constructed with the help of an exponential autoregressive process.

### 1. INTRODUCTION

The preceding paper [1] is devoted to the study of the asymptotic behavior of stochastic experiments constructed with the help of a linear regression function of the increments.

In the current paper, the directed forces (see [1, Assumption]) include a nonlinear factor and this makes the asymptotic analysis of stochastic experiments much more complicated. At the same time, only the variance of the normal autoregressive process is influenced by the nonlinear component (see [1, Proposition 5.1]).

### 2. SETTING OF THE PROBLEM

Stochastic experiments are defined in terms of a sequence of collections of jointly independent random variables  $(\delta_r(k), 1 \leq r \leq N)$  being identically distributed for  $r \in [1, N]$  and assuming only two values  $\pm 1$ .

Binary stochastic experiments are given by

$$(1) \quad S_N(k) := \frac{1}{N} \sum_{r=1}^N \delta_r(k), \quad k \geq 0.$$

Binary stochastic experiments can also be represented in terms of frequencies,

$$(2) \quad S_N(k) = S_N^+(k) - S_N^-(k), \quad k \geq 0,$$

$$(3) \quad S_N^\pm(k) := \frac{1}{N} \sum_{r=1}^N I\{\delta_r(k) = \pm 1\}, \quad k \geq 0.$$

As usual,  $I(A)$  denotes the indicator of a random event  $A$ . Here  $S_N^+(k)$  and  $S_N^-(k)$  mean the density of 1's and  $-1$ 's in the  $k^{\text{th}}$  sample, respectively.

Note that

$$(4) \quad S_N^+(k) + S_N^-(k) = 1.$$

---

2010 *Mathematics Subject Classification.* Primary 62F05, 60J70, 62M05.

*Key words and phrases.* Binary statistical experiment, persistent regression, equilibrium state, stochastic approximation, exponential stochastic experiment, exponential normal autoregressive process.

Therefore binary stochastic experiments (1)–(2) uniquely define the frequencies in stochastic experiments and vice versa:

$$(5) \quad S_N^\pm(k) := [1 \pm S_N(k)]/2, \quad S_N(k) = 2S_N^+(k) - 1 = 1 - 2S_N^-(k).$$

We study the asymptotic behavior as  $N \rightarrow \infty$  of the increments of stochastic experiments over a unit of time:

$$(6) \quad \Delta S_N^\pm(k) := S_N^\pm(k+1) - S_N^\pm(k), \quad k \geq 0,$$

and

$$(7) \quad \Delta S_N(k) := S_N(k+1) - S_N(k), \quad k \geq 0.$$

In what follows we denote by  $p_\pm$  the corresponding frequencies in stochastic experiments,

$$p_+ + p_- = 1, \quad 0 \leq p_\pm \leq 1,$$

and by  $\bar{p} := (p_+, p_-)$  the frequency for the corresponding two component stochastic experiments

$$(8) \quad \bar{S}_N(k) := (S_N^+(k), S_N^-(k)).$$

In the current paper, the regression function of increments is given by the product of two components, namely of the linear directed force  $[V_-p_- - V_+p_+]$  defined in [1] and the factor  $p_+p_-$  that changes the directed force in a neighborhood of boundary values  $-1$  and  $+1$ . Such a nonlinear factor appears, for example, in the Wright–Fisher model (see, for example, [2, Chapter 10]).

**Assumption 1** (Main assumption). *The regression function of increments of frequency components of stochastic experiments possesses the property of a persistent nonlinear regression determined by the following nonlinear directed force:*

$$(9) \quad C_0^\pm(\bar{p}) := \mathbb{E} [\Delta S_N^\pm(k) \mid \bar{S}_N(k) = \bar{p}] = \pm p_+p_- [V_-p_- - V_+p_+],$$

where the parameters of  $V_\pm$  are such that

$$(10) \quad 0 < V_\pm < 1.$$

According to relation (9), the regression function of the increments of binary stochastic experiments

$$(11) \quad C_0(s) := \mathbb{E} [\Delta S_N(k) \mid S_N(k) = s]$$

admits the following representation:

$$(12) \quad C_0(s) = C_0^+(\bar{p}) - C_0^-(\bar{p}) = -\frac{1}{4}V(1-s^2)(s-\rho), \quad |s| \leq 1.$$

It is clear that the directed parameters and their frequency components are such that

$$(13) \quad V_0 := V_- - V_+, \quad V := V_- + V_+,$$

and

$$(14) \quad s = p_+ - p_-, \quad p_\pm = \frac{1}{2}(1 \pm s), \quad \rho = V_0/V.$$

The balance condition for the directed force is obvious, namely

$$(15) \quad C_0^+(\bar{p}) + C_0^-(\bar{p}) = 0.$$

*Remark 2.1.* The main assumption means that the increment of frequencies of stochastic experiments (6) evaluated at the moment  $k+1$  increases or decreases proportionally to the additional frequency with coefficients  $V_+$  or  $V_-$ , respectively.

*Remark 2.2.* Another consequence of the main assumption is that the regression functions of stochastic experiments (9)–(12) are given by

$$(16) \quad C^\pm(\bar{p}) := \mathbf{E} \left[ \bar{S}_N^\pm(k+1) \mid \bar{S}_N(k) = \bar{p} \right] = p_\pm + C_0^\pm(\bar{p}),$$

$$(17) \quad C(s) := \mathbf{E}[S_N(k+1) \mid S_N(k) = s] = s + C_0(s), \quad |s| \leq 1.$$

*Remark 2.3.* Representation (2) and equality (4) imply that

$$(18) \quad C(s) = C^+(\bar{p}) - C^-(\bar{p}), \quad C^+(\bar{p}) + C^-(\bar{p}) = 1.$$

Using (14), condition (15) determines the frequency components (16):

$$(19) \quad C^\pm(\bar{p}) = \frac{1}{2}[1 \pm C(s)], \quad s = p_+ - p_-, \quad p_+ + p_- = 1.$$

### 3. EQUILIBRIUM STATE OF STOCHASTIC EXPERIMENTS

The equilibrium state for the probability frequencies  $\rho_\pm$  is defined by invariant states of the directed force

$$(20) \quad C_0^\pm(\bar{\rho}) = 0, \quad \bar{\rho} := (\rho_+, \rho_-).$$

Therefore the equilibrium state of the directed force is determined by the equations

$$(21) \quad V_- \rho_- - V_+ \rho_+ = 0, \quad \rho_- + \rho_+ = 1.$$

The directed force in the case of binary stochastic experiments (1) also has an equilibrium state (see (12), (14))

$$(22) \quad C_0(\rho) = 0, \quad \rho := \rho_+ - \rho_-.$$

The equilibrium frequencies are determined by a solution of equations (21). This solution is given by

$$(23) \quad \rho_\pm = V_\mp / V, \quad V = V_- + V_+.$$

The regression function of increments of frequencies admits a representation in terms of frequency fluctuations.

**Corollary 3.1.** *The directed forces are such that*

$$(24) \quad C_0^\pm(\bar{p}) = -V_0(\bar{p})(p_\pm - \rho_\pm), \quad C_0(s) = -V(s)(s - \rho),$$

$$(25) \quad V_0(\bar{p}) := V p_+ p_- = V p_\pm (1 - p_\pm), \quad V(s) = \frac{1}{4} V (1 - s^2).$$

*This follows from (9), (12), and (21)–(23).*

A feature of the directed force (24) as a function of fluctuations is that the increments of stochastic experiments (6)–(7) are acting (in the mean sense) in the opposite direction to the fluctuation  $s - \rho$ ; that is, the dynamics of stochastic experiments is directed (in the mean sense) toward the point of equilibrium  $\rho$ .

Therefore the directed force statistically describes the “power” that governs the dynamics of stochastic experiments and leads to the equilibrium state. This fact is described mathematically in the following result.

**Theorem 3.1.** *Assume that at least one of the following three initial states of stochastic experiments converges with probability one:*

$$(26) \quad S_N^\pm(0) \xrightarrow{P1} \rho_\pm, \quad S_N(0) \xrightarrow{P1} \rho = \rho_+ - \rho_-$$

as  $N \rightarrow \infty$ . Then their frequency components converge with probability one, as well:

$$(27) \quad S_N^\pm(k) \xrightarrow{P1} \rho_\pm \quad \text{and} \quad S_N(k) \xrightarrow{P1} \rho$$

as  $N \rightarrow \infty$  for every finite  $k > 0$ .

*Proof of Theorem 3.1.* In what follows we denote the fluctuations of frequencies and binary stochastic experiments by

$$(28) \quad \begin{aligned} \widehat{S}_N^\pm(k) &:= S_N^\pm(k) - \rho_\pm, & \widehat{S}_N(k) &:= S_N(k) - \rho, \\ \widehat{s} &:= s - \rho, & \rho &= \rho_+ - \rho_-, & \widehat{\rho}_\pm &:= \rho_\pm - \rho, & k &\geq 0. \end{aligned}$$

According to (1)–(5), it is sufficient to prove either the convergence of stochastic experiments (1) or that of one of their frequency components (2). Below we prove the convergence of stochastic experiments (1).

Consider the martingale difference

$$(29) \quad \widehat{\mu}_N(k+1) := \Delta \widehat{S}_N(k) - C_0(S_N(k)) = \Delta \widehat{S}_N(k) + V(S_N(k)) \widehat{S}_N(k), \quad k \geq 0.$$

The main assumption (9) implies that

$$(30) \quad \mathbb{E} [\widehat{\mu}_N(k+1) \mid S_N(k)] = 0.$$

Further we need the following:

**Lemma 3.1.** *The quadratic characteristics of the martingale difference (29) are given by*

$$(31) \quad \begin{aligned} \mathbb{E} [\widehat{\mu}_N^2(k+1) \mid S_N(k)] &= B(S_N(k))/N, \\ B(s) &:= 1 - C^2(s), & C(s) &= s + C_0(s). \end{aligned}$$

Similarly, the quadratic characteristics of the frequency martingale difference

$$(32) \quad \widehat{\mu}_N^\pm(k+1) := \Delta \widehat{S}_N^\pm(k) - C_0^\pm(S_N(k)), \quad k \geq 0,$$

are given by

$$(33) \quad \mathbb{E} [(\widehat{\mu}_N^\pm(k+1))^2 \mid \overline{S}_N(k)] = C^\pm(\overline{S}_N(k)) [1 - C^\pm(\overline{S}_N(k))] / N.$$

*Proof of Lemma 3.1.* As proved in [1, Lemma 2.1],

$$(34) \quad \text{Var} [S_N^\pm(k+1) \mid \overline{S}_N(k) = \bar{p}] = \frac{1}{N} C^\pm(\bar{p}) [1 - C^\pm(\bar{p})], \quad k \geq 0.$$

This implies that

$$\mathbb{E} [(\widehat{\mu}_N^\pm(k+1))^2 \mid \overline{S}_N(k)] = \mathbb{E} [(\mu_N^\pm(k+1))^2 \mid \overline{S}_N(k)] = \text{Var} [S_N^\pm(k+1) \mid \overline{S}_N(k)].$$

The relations for  $\widehat{\mu}_N(k)$  are proved similarly. □

Consider the sum of martingale differences (29),

$$(35) \quad \widehat{M}_N(n) := \sum_{k=0}^n \widehat{\mu}_N(k+1), \quad n \geq 0.$$

Relations (31) imply that the quadratic characteristics of martingales (35) tend to zero with probability one as  $N \rightarrow \infty$  if  $n \geq 0$  is fixed. In particular, if  $n = 0$ , then

$$(36) \quad \begin{aligned} \widehat{M}_N(0) &= \widehat{\mu}_N(1) = \Delta \widehat{S}_N(0) + V(S_N(0)) \widehat{S}_N(0) \\ &= \widehat{S}_N(1) - \widehat{S}_N(0) \cdot [1 - V(S_N(0))]. \end{aligned}$$

Thus

$$(37) \quad \widehat{S}_N(1) = \widehat{\mu}_N(1) + [1 - V(S_N(0))] \cdot \widehat{S}_N(0) \xrightarrow{P1} 0, \quad N \rightarrow \infty,$$

according to initial conditions (26) stated in Theorem 3.1.

Now we prove the almost sure convergence (27) as well as (28) by induction if  $k \geq 1$  is fixed. Theorem 3.1 is proved.  $\square$

4. AN APPROXIMATION OF STOCHASTIC EXPERIMENTS  
BY A NORMAL AUTOREGRESSIVE PROCESS

The existence of equilibrium states (20), (22) of stochastic experiments (1)–(2) together with the convergence proved in Theorem 3.1 allows us to use another approximation of stochastic experiments with the help of the asymptotic behavior of the martingale component in the representation of increments of stochastic experiments (29) characterized by the first two moments (see (30)–(31)),

$$(38) \quad \begin{aligned} \mathbb{E} [\widehat{\mu}_N(k+1) \mid S_N(k)] &= 0, \\ \mathbb{E} [\widehat{\mu}_N^2(k+1) \mid S_N(k)] &= B(S_N(k))/N, \quad k \geq 0. \end{aligned}$$

The asymptotic behavior of the martingale difference in (29) can be obtained under the Lindeberg condition (see [5])

$$(39) \quad \sum_{k=0}^{n-1} \mathbb{E} [N\widehat{\mu}_N^2(k+1)I \{|\widehat{\mu}_N^2(k+1)| \geq \varepsilon/N\} \mid S_N(k)] \xrightarrow{P1} 0.$$

**Theorem 4.1.** *Assume that the conditions of Theorem 3.1 hold. Then conditions (38)–(39) imply that the sums of normalized martingale difference (29) converge in distribution:*

$$(40) \quad \sqrt{N} \sum_{k=0}^{n-1} \widehat{\mu}_N(k+1) \xrightarrow{d} n\sigma W, \quad N \rightarrow \infty, \quad \forall n \geq 1,$$

and

$$(41) \quad \sqrt{N} \sum_{k=0}^{n-1} \widehat{\mu}_N^\pm(k+1) \xrightarrow{d} n\check{\sigma}W, \quad N \rightarrow \infty, \quad \forall n \geq 1.$$

Here  $W$  is a standard normal random variable and the variances  $\sigma^2$  and  $\check{\sigma}^2$  are given by

$$\check{\sigma} = \rho_+\rho_-, \quad \sigma^2 = 4\check{\sigma}^2 = 1 - \rho^2, \quad \rho = \rho_+ - \rho_-.$$

*Remark 4.1.* It is known (see [5]) that the uniform convergence in probability of the martingale difference

$$\mathbb{E} \left[ \max_{1 \leq k \leq n} |\widehat{\mu}_N^\pm(k)| \right] \xrightarrow{P} 0, \quad N \rightarrow \infty,$$

implies Lindeberg’s condition (39).

**Corollary 4.1.** *The convergence of the sum of martingale differences (29) and their frequency components (32) implies the following convergence in distribution for every fixed  $k \geq 0$ :*

$$(42) \quad \begin{aligned} \sqrt{N} \left[ \Delta \widehat{S}_N(k) + V(S_N(k))\widehat{S}_N(k) \right] &\xrightarrow{d} \sigma \widehat{W}(k+1), \quad N \rightarrow \infty, \quad k \geq 0, \\ \sqrt{N} \left[ \Delta \widehat{S}_N^\pm(k) + V(S_N(k))\widehat{S}_N^\pm(k) \right] &\xrightarrow{d} \check{\sigma} \widehat{W}(k+1), \quad N \rightarrow \infty, \quad k \geq 0. \end{aligned}$$

**Proposition 4.1.** *Limit relations (42) allow us to use the approximation of stochastic experiments (1)–(2) by a normal autoregressive process defined by*

$$(43) \quad \begin{aligned} \widehat{\widehat{S}}_N(k+1) &= V(\widetilde{S}_N(k))\widehat{\widehat{S}}_N(k) + (\sigma/\sqrt{N})\widehat{W}(k+1), \\ \widehat{\widehat{S}}_N^\pm(k+1) &= V(\widetilde{S}_N^\pm(k))\widehat{\widehat{S}}_N^\pm(k) + (\check{\sigma}/\sqrt{N})\widehat{W}^\pm(k+1). \end{aligned}$$

In (42) and (43) as well as in (40)–(41),  $\widehat{W}(k+1)$  and  $\widehat{W}^\pm(k+1)$  are standard normal random variables.

*Remark 4.2.* As a function of fluctuations, the regression function of increments of stochastic experiments admits the following representation:

$$(44) \quad C_0(s) = -\frac{1}{4}V[\sigma^2 - (\widehat{s} + 2\rho)\widehat{s}] \widehat{s}, \quad \widehat{s} = s - \rho.$$

Equality (44) follows from the representation of the nonlinear factor (24)–(25) as a function of fluctuations:

$$1 - s^2 = \sigma^2 - (\widehat{s} + 2\rho)\widehat{s}.$$

Proposition 4.1 and representation (44) imply the following approximation of stochastic experiments by a linear normal autoregressive process.

**Proposition 4.2.** *Limit relations (42) imply the approximation of stochastic experiments (1)–(2) by a linear normal autoregressive process defined by*

$$(45) \quad \begin{aligned} \Delta\widehat{\widehat{S}}_N(k+1) &= -\frac{V}{4}\sigma^2\widehat{\widehat{S}}_N(k) + (\sigma/\sqrt{N})\widehat{W}(k+1), \\ \Delta\widehat{\widehat{S}}_N^\pm(k+1) &= -V\check{\sigma}^2\widehat{\widehat{S}}_N^\pm(k) + (\check{\sigma}/\sqrt{N})\widehat{W}^\pm(k+1). \end{aligned}$$

Stochastic experiments  $\widehat{\widehat{S}}_N(k)$  and  $\widehat{\widehat{S}}_N^\pm(k)$ ,  $k \geq 0$ , are well defined by difference stochastic equations (45) under given initial values.

### 5. PROOF OF THEOREM 4.1

Consider fluctuations of stochastic experiments centered with an equilibrium state  $\rho$  and normalized by factor  $\sqrt{N}$ :

$$(46) \quad \zeta_N(k) := \sqrt{N}\widehat{S}_N(k) = \sqrt{N}[S_N(k) - \rho].$$

Further we use the following representation of the martingale difference (29) (also see equalities (24)–(25)) with a separated linear component:

$$(47) \quad \sqrt{N}\widehat{\mu}_N(k+1) = \Delta\zeta_N(k) - \frac{1}{4}V\sigma^2\zeta_N(k) + \frac{1}{\sqrt{N}}R(\widehat{S}_N(k))\zeta_N^2(k),$$

where

$$R(\widehat{s}) = -\frac{1}{4}V(\widehat{s} + 2\rho).$$

The remainder term on the right-hand side of (47) tends to zero in probability as  $N \rightarrow \infty$  since the product of functions  $R(\widehat{S}_N(k))\zeta_N^2(k)$  is bounded.

Consider the sum of martingale differences and its asymptotic representation

$$(48) \quad M_N(n) := \sum_{k=0}^{n-1} \sqrt{N}\widehat{\mu}_N(k+1) = \sum_{k=0}^{n-1} [\zeta_N(k+1) + V(S_N(k))\zeta_N(k)].$$

The quadratic characteristics of the martingale  $M_N(n)$  are given by

$$(49) \quad \langle M_N \rangle_n = \sum_{k=0}^{n-1} [1 - C^2(S_N(k))], \quad n \geq 0$$

(see Lemma 3.1). The second term on the right-hand side of (49) tends to zero in probability as  $N \rightarrow \infty$ . By Theorem 4.1, the quadratic characteristics converge with probability one:

$$(50) \quad \langle M_N \rangle_n \xrightarrow{P1} n(1 - \rho^2) = n\sigma^2 \quad \text{as } N \rightarrow \infty, \quad n \geq 0.$$

The central limit theorem for martingale differences (see [6], for example) implies the convergence in distribution of martingale (48) to the sum of normal random variables  $W(k)$ ; that is, (40) holds.

Theorem 4.1 is proved.

Taking into account the asymptotic representation of the martingale

$$(51) \quad M_N(n) = \sum_{k=0}^{n-1} \left[ \Delta\zeta_N(k) - \frac{1}{4}V\sigma^2\zeta_N(k) \right] + \frac{1}{\sqrt{N}} \sum_{k=0}^{n-1} R(\widehat{S}_N(k))\zeta_N^2(k),$$

we obtain the following result.

**Corollary 5.1.** *For every fixed  $k \geq 1$ ,*

$$(52) \quad \sum_{k=0}^{n-1} \left[ \Delta\zeta_N(k+1) - \frac{1}{4}V\sigma^2\zeta_N(k) \right] \xrightarrow{d} \sigma \sum_{k=0}^{n-1} W(k+1), \quad N \rightarrow \infty.$$

Using limit relation (42) one can introduce a new approximating model of normalized fluctuations of binary stochastic experiments in the following form, which is suitable for several applications.

**Proposition 5.1.** *Limit relations (52) imply that the normal autoregressive process  $\zeta_N(k)$ ,  $k \geq 0$ , is a solution of the following difference stochastic equation:*

$$(53) \quad \Delta\zeta_N(k) = -V(\rho)\zeta_N(k) + \sigma\Delta W(k),$$

where

$$V(\rho) := \frac{1}{4}V(1 - \rho^2) = \frac{1}{4}V\sigma^2, \quad \Delta W(k) := W(k+1) - W(k), \quad \sigma^2 = 1 - \rho^2.$$

The existence of two models in the case of the persistent regression for binary stochastic experiments described in Propositions 4.1 and 4.2 allow one to solve a problem of the statistical analysis for these models by using initial values of stochastic experiments normalized by sums of sample values (1) satisfying the main assumption (9).

### 6. EXPONENTIAL STOCHASTIC EXPERIMENTS

Like the preceding paper [1], we apply the limit results of Theorems 3.1 and 4.1 for *exponential stochastic experiments* defined by the symmetric exponential statistics

$$(54) \quad \Pi_N(\lambda, k) := \prod_{r=1}^N [1 + \lambda\delta_r(k)], \quad k \geq 0,$$

and corresponding frequency components

$$(55) \quad \Pi_N^\pm(\lambda, k) := \prod_{r=1}^N [1 \pm \lambda I\{\delta_r(k) = \pm 1\}], \quad k \geq 0.$$

It is clear that

$$\Pi_N(\lambda, k) = \Pi_N^+(\lambda, k) \cdot \Pi_N^-(\lambda, k), \quad k \geq 0.$$

Further we establish the asymptotic behavior as  $N \rightarrow \infty$  of exponential stochastic experiments defined by (54). Similar results hold for frequency exponential stochastic experiments (55), as well.

7. STATIONARY REGIME OF EXPONENTIAL STOCHASTIC EXPERIMENTS

Consider an exponential martingale with normalization  $\lambda_N = \lambda/N$ :

$$(56) \quad \mu_N^e(\lambda/N, k + 1) = \Pi_N(\lambda/N, k + 1) / \bar{\Pi}_N(\lambda/N, k), \quad k \geq 0,$$

where  $\bar{\Pi}_N(\lambda/N, k)$  is the conditional expectation of an exponential stochastic experiment:

$$(57) \quad \bar{\Pi}_N(\lambda, k) := \mathbb{E}[\Pi_N(\lambda, k + 1) \mid S_N(k)] = [1 + \lambda C(S_N(k))]^N.$$

The martingale property of  $\mu_N^e(\lambda/N, k)$  follows from

$$(58) \quad \mathbb{E}[\mu_N^e(\lambda/N, k + 1) \mid S_N(k)] = 1, \quad k \geq 0.$$

The existence of a stationary regime for exponential stochastic experiments is established in the following result.

**Theorem 7.1.** *Assume that condition (26) holds; that is, the initial values of stochastic experiments converge with probability one. Then the exponential stochastic experiment (53) converges in probability,*

$$(59) \quad \Pi_N(\lambda/N, k) \xrightarrow{P} \exp(\lambda\rho), \quad N \rightarrow \infty, \quad k \geq 0,$$

and its conditional expectation (57) converges in probability, as well,

$$(60) \quad \bar{\Pi}_N(\lambda/N, k) \xrightarrow{P} \exp(\lambda\rho), \quad N \rightarrow \infty, \quad k \geq 0.$$

The proof of Theorem 7.1 follows along the lines of the proof of an analogous result in the preceding paper (see [1, Theorem 7.1]) by using the first-order Le Cam approximation for the logarithmic function.

**Corollary 7.1.** *If condition (26) holds, then*

$$(61) \quad \mu_N^e(\lambda/N, k + 1) \xrightarrow{P} 1, \quad N \rightarrow \infty.$$

8. APPROXIMATION OF EXPONENTIAL STOCHASTIC EXPERIMENTS BY A NORMAL AUTOREGRESSIVE PROCESS

Now we consider exponential stochastic experiments and their conditional expectations for the new scale  $\lambda_N = \lambda/\sqrt{N}$ :

$$(62) \quad \Pi_N(\lambda/\sqrt{N}, k + 1) = \prod_{r=1}^N [1 + \lambda\delta_r(k + 1)/\sqrt{N}], \quad k \geq 0,$$

$$\bar{\Pi}_N(\lambda/\sqrt{N}, k) = [1 + \lambda C(S_N(k))/\sqrt{N}]^N, \quad k \geq 0.$$

The exponential martingale with a new normalization,

$$(63) \quad \mu_N^e(\lambda_N, k + 1) := \Pi_N(\lambda/\sqrt{N}, k + 1) / \bar{\Pi}_N(\lambda/\sqrt{N}, k), \quad k \geq 0,$$

obviously possesses the martingale property,

$$\mathbb{E}[\mu_N^e(\lambda_N, k + 1) \mid S_N(k)] = 1, \quad k \geq 0.$$

The following limit result allows one to approximate exponential stochastic experiments.



**Theorem 8.1.** *Let condition (26) hold; that is, the initial values of stochastic experiments converge with probability one. Then we have the convergence in distribution, namely*

$$(64) \quad \mu_N^e(\lambda/\sqrt{N}, k + 1) \xrightarrow{d} \exp [\lambda\sigma W(k + 1) - \lambda^2\sigma^2/2], \quad N \rightarrow \infty, k \geq 0.$$

*Proof of Theorem 8.1.* As in the proof of an analogous theorem in the preceding paper (see [1, Theorem 7.1]), we use the equality  $\Pi = \exp \ln \Pi$  together with the second-order Le Cam approximation (see [4, Lemma 6.3.1]). □

*Remark 8.1.* The limit in (64) is an exponential martingale

$$(65) \quad \mathbb{E} \exp [\lambda\sigma W(k + 1) - \lambda^2\sigma^2/2] = 1, \quad k \geq 0.$$

According to (63), we have

$$(66) \quad \Pi_N(\lambda/\sqrt{N}, k + 1) = \bar{\Pi}_N(\lambda/\sqrt{N}, k) \mu_N^e(\lambda/\sqrt{N}, k + 1).$$

Since the conditional expectation  $\bar{\Pi}_N(\lambda/\sqrt{N}, k)$  converges in probability,

$$\left\{ \bar{\Pi}_N(\lambda/\sqrt{N}, k) \exp[-\lambda\sqrt{N}C(S_N(k))] \right\} \xrightarrow{P} \exp[-\lambda^2\rho^2/2], \quad N \rightarrow \infty, k \geq 0$$

(this is proved in [1, Section 8]), we obtain the following asymptotic representation:

$$(67) \quad \bar{\Pi}_N(\lambda/\sqrt{N}, k) = \exp[\lambda\sqrt{N}C(S_N(k)) - \lambda^2\rho^2/2] e^{R_N}, \quad k \geq 0.$$

Here and in what follows the remainder term  $R_N$  tends to zero in probability as  $N \rightarrow \infty$ .

Considering the regression functions (17) and (24)–(25), the conditional expectation (67) is represented as a function of fluctuations  $\widehat{S}_N(k) := S_N(k) - \rho$ :

$$(68) \quad \bar{\Pi}_N(\lambda/\sqrt{N}, k) = \exp \left\{ \lambda\sqrt{N}[\rho + (1 - V(S_N(k)))\widehat{S}_N(k)] - \lambda^2\rho^2/2 \right\} e^{R_N}, \\ k \geq 0.$$

The asymptotic representation of an exponential stochastic experiment (68) together with Theorem 8.1 and representation (66) implies the approximation of an exponential stochastic experiment (62) by a normal autoregressive process.

In Propositions 8.1 and 8.2, we consider exponential stochastic experiments with normal exponential autoregression that generate an approximation of an exponential stochastic experiment presented in Theorem 8.1.

**Proposition 8.1.** *Limit relations (64) and (68) imply that the exponential stochastic experiments with a nonlinear normal autoregression are given by*

$$(69) \quad \widetilde{\Pi}_N(\lambda/\sqrt{N}, k + 1) = \exp \left\{ \lambda\sqrt{N} \left[ \rho + [1 - V(\widetilde{S}_N(k))]\widehat{\widetilde{S}}_N(k) \right] - \lambda^2\rho^2/2 \right\} \\ \times \exp [\lambda\sigma W(k + 1) - \lambda^2\sigma^2/2],$$

or, in other words,

$$(70) \quad \widetilde{\Pi}_N(\lambda/\sqrt{N}, k + 1) = \exp \left\{ \sqrt{N}\lambda C(\widehat{\widetilde{S}}_N(k)) + \sigma W(k + 1) - \lambda^2/2 \right\}.$$

*Remark 8.2.* Important for the approximation by a normal autoregressive process (69) or (70) is the property that the conditional expectation of a normal autoregressive process (69) (or (70)) asymptotically coincides with the *regression function* (conditional expectation) of defining exponential stochastic experiments (62), namely

$$(71) \quad \mathbb{E} \left[ \widetilde{\Pi}_N(\lambda/\sqrt{N}, k + 1) \right] = \exp \left[ \sqrt{N}\lambda C(\widetilde{S}_N(k)) - \lambda^2\rho^2/2 \right] = \bar{\Pi}_N(\lambda/\sqrt{N}, k) e^{R_N}.$$

**Proposition 8.2.** *Limit relations (64) and (68) together with representation (44) imply that exponential stochastic experiments with a linear normal autoregressive process are given by*

$$(72) \quad \begin{aligned} & \tilde{\Pi}_N(\lambda/\sqrt{N}, k+1) \\ &= \exp \left\{ \lambda\sqrt{N} \left[ \rho - \frac{1}{4}V\sigma^2\widehat{S}_N(k) \right] - \lambda^2\rho^2/2 \right\} \exp [\lambda\sigma W(k+1) - \lambda^2\sigma^2/2] \end{aligned}$$

or, in other words,

$$(73) \quad \tilde{\Pi}_N(\lambda/\sqrt{N}, k+1) = \exp \left\{ \sqrt{N}\lambda \left[ \rho - \frac{1}{4}V\sigma^2\widehat{S}_N(k) \right] + \sigma W(k+1) - \lambda^2/2 \right\}.$$

## 9. CONCLUDING REMARKS

The six models considered above can be used in the statistical analysis of compound systems with cooperative interactions. More precisely, we obtained approximations by a nonlinear normal autoregression (43) and by linear normal autoregression (45) for the model of stochastic experiments (1). For the model of exponential stochastic experiments (62), we established approximations by an exponential normal autoregression (69) and linear normal autoregression (72). These models will be used elsewhere for the analysis of statistical data in some applied problems.

## 10. ACKNOWLEDGEMENT

The author is indebted to an anonymous referee for valuable remarks that helped to improve the presentation of the results of the paper.

## BIBLIOGRAPHY

1. D. V. Koroliuk, *Two component binary statistical experiments with persistent linear regression*, *Teor. Imovirnost. Math. Statist.* **90** (2014), 91–101; English transl. in *Theor. Probability and Math. Statist.* **90** (2015), 103–114. MR3242023
2. S. N. Ethier and T. G. Kurtz, *Markov Processes: Characterization and Convergence*, Wiley, New York, 1986. MR838085 (88a:60130)
3. A. V. Skorokhod, *Asymptotic Methods in the Theory of Stochastic Differential Equations*, “Naukova dumka”, Kiev, 1987; English transl., American Mathematical Society, Providence, RI, 2009. MR913305 (88m:60164)
4. Yu. V. Borovskikh and V. S. Korolyuk, *Martingale Approximation*, VSP, 1997. MR1640099 (99f:60001)
5. A. N. Shiryaev, *Probability-2*, MCNMO, Moscow, 2004. (Russian)
6. A. N. Shiryaev, *Essentials of Stochastic Finance: Facts, Models, Theory* “Fazis”, Moscow, 1998; English transl., World Scientific Pub. Co. Inc., Singapore, 1999. MR1695318 (2000e:91085)
7. Yu. S. Mishura and G. M. Shevchenko, *Mathematics of Finance*, Kyiv University Publishing House, Kyiv, 2011. (Ukrainian)
8. M. Abundo, L. Accardi, L. Stella, and N. Rosato, *A stochastic model for the cooperative relaxation of proteins, based on a hierarchy of interactions between amino acidic residues*, *Math. Models and Methods in Appl. Sciences* (1998) no. 8, 327–358. MR1618474 (99g:92020)
9. V. S. Korolyuk and D. Koroliuk, *Diffusion approximation of stochastic Markov models with persistent regression*, *Ukrain. Matem. Zh.* **47** (1995), no. 7, 928–935; English transl. in *Ukrain. Math. J.* **47** (1995), no. 7, 1065–1073. MR1367948 (97j:60040)

INSTITUTE OF TELECOMMUNICATIONS AND GLOBAL INFORMATION SPACE OF NATIONAL ACADEMY OF SCIENCE OF UKRAINE, CHOKOLOVSKIĀ BLVD., 13, KYIV, 03110, UKRAINE  
E-mail address: dimitri.koroliuk@ukr.net

Received 26/APR/2013  
Translated by V. SEMENOV