DOI 10.1007/s10559-015-9755-4
Cybernetics and Systems Analysis, Vol. 51, No. 4, July, 2015

CLASSIFICATION OF BINARY DETERMINISTIC STATISTICAL EXPERIMENTS WITH PERSISTENT REGRESSION

D. V. Koroliouk

UDC 519.24

Abstract. *For models of binary deterministic statistical experiments defined recursively by solutions of deterministic difference equations for probabilities of alternatives of binary states, a classification is investigated that depends on the parameter values of a guiding action that determine regression functions of probability increments for such alternatives. The classification is substantiated by limit properties of solutions to deterministic difference equations generating probabilities of alternatives.*

Keywords: *statistical experiment, deterministic difference equation for probabilities of alternatives of binary states, persistent regression, equilibrium, attracting state, repulsive state, dominance of alternatives.*

In [1], the dynamics of recursive statistical experiments with persistent regression is investigated in a scheme of series with the series parameter N (sample size), $N \to \infty$. In this case, a condition providing the presence of an equilibrium state determined by the equilibrium of a regression function is essentially used.

This article considers the dynamics of statistical experiments (SEs) with increasing the amount of time $k \to \infty$. Taking into account the mathematical models of population genetics from [2, 3], the regression function of increments in probabilities of alternatives is described by the product of the linear component defined by the parameters of the guiding action $V_\pm$ and the following nonlinear component providing the correctness of the definition of the regression function:

$$C_0(P) = P_+ P_- [V_- P_- - V_+ P_+], \; 0 \le P_\pm \le 1. \tag{1}$$

In this case, the problem arises as to the classification of SE models depending on values of the parameters of a guiding action $V_\pm$ that determine the regression function of SE increments.

A similar problem of classification of deterministic SE models was investigated in population genetics [2, 4, 5] with the use of differential equations for probabilities of alternative genotypes.

PROBLEM STATEMENT

A sequence of SEs with persistent regression is specified by average values of a sample $\delta(k) := (\delta_r(k), 1 \le r \le N)$, $k \ge 0$, that are independent in the aggregate when k are fixed and random quantities $\delta_r(k)$ are equally distributed with respect to r and assume the binary values -1 or 0; such a sequence is of the form

$$S_N(k) := \frac{1}{N} \sum_{r=1}^{N} \delta_r(k), \; k \ge 0. \tag{2}$$

Institute of Telecommunications and Global Information Space, National Academy of Sciences of Ukraine, Kyiv, Ukraine, *dimitri.koroliouk@ukr.net.* Translated from Kibernetika i Sistemnyi Analiz, No. 4, pp. 163–168, July–August, 2015. Original article submitted July 25, 2014.

 1060-0396/15/5104-0644

Values of binary quantities $\delta_r(k)$ are interpreted as a choice between two alternatives designated by the symbols + (the positive alternative $\delta_r(k)=1$) or − (the negative alternative $\delta_r(k)=0$).

As is obvious, SE (2) specifies the average value for the choice of the positive alternative at the kth stage. Similarly, the average value of the negative alternative is defined as

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

Here, by definition, $I(A)$ is the indicator of a random event A. The following identity obviously holds:

$$S_N(k)+S_N^-(k)=1.$$

We introduce the following probabilities of alternatives:

$$P_+(k):=P\{\delta_r(k)=1\}=1-P\{\delta_r(k)=0\}=1-P_-(k),\ k\ge 0. \tag{3}$$

By definition, we have

$$P_+(k)+P_-(k)=1,\ k\ge 0.$$

The dynamics of probabilities of alternatives is specified by the regression function of increments in probabilities (1).

Assumption 1. The probabilities of alternatives are specified by solutions of the deterministic difference equations (DDEs)

$$\Delta P_\pm(k+1)=\pm P_+(k)P_-(k)[V_-P_-(k)-V_+P_+(k)],\ k\ge 0, \tag{4}$$

$$\Delta P_\pm(k+1):=P_\pm(k+1)-P_\pm(k).$$

The numerical parameters of regression functions of probability increments satisfy the additional conditions

$$|V_\pm|<1,\ |V_+ + V_-|>0 \tag{5}$$

providing the correctness of the specification of probabilities of alternatives of DDEs (4).

Note that the regression function in Eqs. (4) has two absorbing states $\rho_\pm^0=0$ and $\rho_\mp^0=1$ and also the equilibrium state (equilibrium) (ρ_+,ρ_-) specified by the relationship

$$V_-\rho_- = V_+\rho_+. \tag{6}$$

Thus, the basic equilibrium is defined by the relationships

$$\rho_\pm = V_\mp / V,\ V=V_+ + V_-,\ |V|>0. \tag{7}$$

We introduce fluctuations of probabilities of alternatives as follows:

$$\hat{P}_\pm(k):=P_\pm(k)-\rho_\pm,\ k\ge 0. \tag{8}$$

Assumption 2. Fluctuations of probabilities of SE alternatives (3) are specified by solutions of the DDE

$$\hat{P}_\pm(k+1)-\hat{P}_\pm(k)=-VP_+(k)P_-(k)\hat{P}_\pm(k),\ k\ge 0. \tag{9}$$

As in Assumption 1, probabilities of alternatives are also determined by the following two parameters (see relationships (6) and (7)): V and $\rho_\pm$ $(\rho_+ + \rho_- = 1)$.

CLASSIFICATION OF DETERMINISTIC SE MODELS

The behavior of probabilities of the alternatives defined by solutions of DDEs (4) or (9) under constraints (5) essentially depends on values of the parameters of the guiding action $V_\pm$.

A classification of SE models with probabilities of alternatives $P_\pm(k)$, $k\ge 0$, specified by solutions of DDE (4) is determined by the following statement.

Proposition 1. The following parameters of the guiding action $V_{\pm}$ that satisfy admissible values of inequalities (5) fix the classification of SE models:

$\mathbb{M}1$ attracting: $V>0$ and $V_{\pm} \geq 0$;
$\mathbb{M}2$ repulsive: $V<0$ and $V_{\pm} \leq 0$;
$\mathbb{M}3+$ dominance +: $V \neq 0$ and $V_{+} \leq 0 \leq V_{-}$;
$\mathbb{M}3-$ dominance –: $V \neq 0$ and $V_{-} \leq 0 \leq V_{+}$.

A similar classification of SE models in which fluctuations of probabilities of alternatives (8) are specified by solutions of DDE (9) takes place with allowance for values of the equilibria $\rho_{\pm}$ and also the parameter $V = V_{+} + V_{-}$.

Proposition 2. Equilibrium values (equilibria) of alternatives $\rho_{\pm}$ determine the following classification of SE models that are specified by solutions of DDE (9):

$\mathbb{M}1$ attracting: $V>0$ and $0<\rho_{\pm}<1$;
$\mathbb{M}2$ repulsive: $V<0$ and $0<\rho_{\pm}<1$;
$\mathbb{M}3+$ dominance +: $V \neq 0$ and $V\rho_{-} \leq 0 \leq V\rho_{+}$;
$\mathbb{M}3-$ dominance –: $V \neq 0$ and $V\rho_{+} \leq 0 \leq V\rho_{-}$.

At the same time, SE models can be classified with allowance for the values of the equilibrium of the difference between alternatives $\rho = \rho_{+} - \rho_{-}$.

Proposition 3. The equilibrium values (equilibria) of the difference between alternatives $\rho = \rho_{+} - \rho_{-}$ determine the following classification of SE models:

$\mathbb{M}1$ attracting: $|\rho| \leq 1$ and $V>0$;
$\mathbb{M}2$ repulsive: $|\rho| \leq 1$ and $V<0$;
$\mathbb{M}3+$ dominance +: $|\rho| \geq 1$ and $V\rho>0$;
$\mathbb{M}3-$ dominance –: $|\rho| \geq 1$ and $V\rho<0$.

Remark. The classification of SE models specified by probabilities of alternatives $P_{\pm}(t)$, $t \geq 0$ (in continuous time), with the regression function $Q(x) = x(1-x)(ax+b)$ is presented in [5] (see also [2]) and corresponds to the classification given in Proposition 3.

SUBSTANTIATION OF THE CLASSIFICATION OF SE MODELS

The SE classification given in Propositions 1–3 is based on the limit behavior of probabilities of SE alternatives, which is described by the following theorem.

THEOREM 1. The probabilities of alternatives $P_{\pm}(k)$, $k \geq 0$, determined by solutions of DDEs (4) or (9) have the following asymptotic behavior:

- in the model M1 (attracting),

$$\lim_{k\to\infty} P_{\pm}(k) = \rho_{\pm}; \tag{10}$$

- in the model M2 (repulsive),

$$\lim_{k\to\infty} P_{\pm}(k) = 1 \text{ or } \lim_{k\to\infty} P_{\mp}(k) = 0 \tag{11}$$

under the initial conditions

$$P_{\pm}(0) > \rho_{\pm} \text{ or } P_{\mp}(0) < \rho_{\mp}; \tag{12}$$

- in the model M3± (dominance ±)

$$\lim_{k\to\infty} P_{\pm}(k) = 1 \text{ and } \lim_{k\to\infty} P_{\mp}(k) = 0. \tag{13}$$

Of interest are the interpretation of different SE models and also underlying motivation for the use of them in practical applications, for example, in population genetics [4, 5], economy, and in behavior models [6, 7].

Consider now the interpretation of the classification of SE models from the viewpoint of behavior models.

Taking into account regression functions of increments in alternatives (1), the classification of SE models is determined by the values of the parameters of the guiding action $V_{\pm}$ that characterize stimuli and suppressions of increments in probabilities of alternatives [7].

In the model $\mathbb{M}1$ (attracting), where $V > 0$ and $V\pm \geq 0$, the probabilities of the positive alternative $P_{+}(k)$, $k \geq 0$, increase (are stimulated) in proportion to the probability of the negative alternative with the parameter V_{-} and decrease in proportion to the probability of the positive alternative with the parameter V_{+}. This characterization of stimuli and suppressions leads to the presence of the stationary mode determined by the equilibrium $\rho_{\pm}$ of the regression function of increments.

In the model $\mathbb{M}2$ (repulsive), where $V < 0$ and $V\pm \leq 0$, the characterization of stimuli and suppressions exerts a back action on the probabilities of alternatives. As a result, the equilibrium of the regression function of increments becomes repulsive, i.e., the probabilities of alternatives tend to the absorbing states $\rho_{\pm} = 0$ or 1. The repulsive equilibrium state $\rho_{\pm}$ acts as a restrictive threshold. If the initial probabilities of alternatives are less than this threshold, then subsequent probabilities decrease up to the absorbing state $\rho = 0$, and if they are larger than the threshold value $\rho_{\pm}$, then they increase up to the absorbing state $\rho = 1$.

In the model $\mathbb{M}3\pm$ (dominance ±), where $|\rho_{+} - \rho_{-}| \geq 1$, the equilibrium of the regression function of increments $\rho_{\pm}$ is located outside the interval $(0, 1)$. Thus, the probabilities of alternatives tend to equilibria and, in the issue of attracting or repulsing, reach the absorbing states $\rho_{\pm} = 0$ or 1. At the same time, in the model $\mathbb{M}3+$ (dominance +), the probabilities of positive alternatives tend to 1, whereas those of negative ones tend to 0, and, in the model $\mathbb{M}3-$ (dominance –), the probabilities of negative alternatives tend to 1.

In this connection, the following statistical problems arise: the verification of hypotheses concerning some SE model or other in real experiments and also estimation of the parameters of the SE guiding action that are specified by DDEs (4) or (9).

PROOF OF THEOREM 1

The main idea of the proof of Theorem 1 lies in the fact that the probabilities of alternatives of an SE that are specified by solutions of DDEs (4) or (9) are bounded and monotone with respect to $k \to \infty$. This stipulates the existence of limits. It is required to make sure that the limit relationships for the probabilities of the frequencies obtained from DDEs (4) or (9) provide the truth of the statement of the theorem.

The proof is underlain by the following basic DDE (see (9)):

$$\Delta\hat{P}_{\pm}(k+1) = -VP_{+}(k)P_{-}(k)\hat{P}_{\pm}(k), \quad k \geq 0. \tag{14}$$

Model $\mathbb{M}1$. We first consider the case of $\mathbb{M}1+$, where $\rho_{+} < P_{+}(0) < 1$. Then, because of fluctuations (8), we have

$$0 < \hat{P}_{+}(0) < \rho_{-}. \tag{15}$$

When $V > 0$, basic DDE (14) implies the monotonicity of the sequence $\hat{P}_{+}(k)$,

$$\hat{P}_{+}(k+1) < \hat{P}_{+}(k), \quad k \geq 0. \tag{16}$$

We rewrite DDE (14) in the form

$$\hat{P}_{+}(k+1) = \hat{P}_{+}(k)[1 - VP_{+}(k)P_{-}(k)], \quad k \geq 0. \tag{17}$$

Taking into account conditions (5), we have the inequality

$$|VP_{+}P_{-}| \leq \frac{1}{2}. \tag{18}$$

According to inequalities (15), (16), and (18) and DDE (17), we obtain the following estimate for DDE (14):

$$\hat{P}_{+}(k+1) \geq \frac{1}{2}\hat{P}_{+}(k) \geq 0, \quad k \geq 0. \tag{19}$$

The monotonicity of probabilities (16) together with estimate (19) provide the existence of the following limits:

$$\hat{P}_{\pm}^{*} := \lim_{k\to\infty} \hat{P}_{\pm}(k), \; P_{\pm}^{*} := \lim_{k\to\infty} P_{\pm}(k) \tag{20}$$

for which the equation

$$P_{+}^{*} P_{-}^{*} \hat{P}_{\pm}^{*} = 0 \tag{21}$$

holds, whence we obtain limit (10),

$$\hat{P}_{\pm} = 0, \text{ i.e., } P_{+}^{*} = \rho_{+} \text{ and } P_{-}^{*} = \rho_{-}.$$

In the case of $\mathbb{M}1-$, where $0 < P_{+}(0) < \rho_{+}$, by virtue of the relationship

$$\hat{P}_{+}(k) + \hat{P}_{-}(k) = 0, \; k \ge 0, \tag{22}$$

the dual problem arises for $\hat{P}_{-}(k)$, namely, the following inequality (cf. inequalities (15)) holds:

$$\rho_{-} < P_{-}(0) < 1, \text{ i.e., } 0 < \hat{P}_{-}(0) < \rho_{+}.$$

Thus, the above reasoning is true for $\hat{P}_{-}(0)$ instead of $\hat{P}_{+}(0)$. In this case, we obtain the following limit values:

$$\hat{P}_{-}^{*} = 0, \text{ i.e., } P_{-}^{*} = \rho_{-} \text{ and } P_{+}^{*} = \rho_{+},$$

which completes the proof of limit (10) of Theorem 1 for the model $\mathbb{M}1$.

Model $\mathbb{M}2$. We first consider the case of $\mathbb{M}2+$, where $\rho_{+} < P_{+}(0) < 1$. From DDE (9), we obtain

$$0 < \hat{P}_{+}(k) < \rho_{-}.$$

According to DDE (14), when $V < 0$, we have the monotonicity of the sequence $\hat{P}_{+}(k)$,

$$\hat{P}_{+}(k+1) > \hat{P}_{+}(k), \; k \ge 0.$$

Relationships (14) and (18) with allowance for $V < 0$ yield the inequality

$$\hat{P}_{+}(k+1) \le \frac{3}{2} \hat{P}_{+}(k), \; k \ge 0. \tag{23}$$

By induction, we obtain the following uniform estimate from inequality (23):

$$\hat{P}_{+}(k+1) \le \frac{3}{2} \rho_{-}.$$

Hence, there are limits (20) for which Eq. (21) holds, which implies limits (11), i.e.,

$$P_{+}^{*} = 1, \quad P_{-}^{*} = 0.$$

In the case of $\mathbb{M}2-$, where $0 < P_{+}(0) < \rho_{+}$, by virtue of relationship (22), the dual problem arises for $P_{-}(0)$, namely, the following inequality holds:

$$\rho_{-} < P_{-}(0) < 1, \quad \text{i.e.,} \quad 0 < \hat{P}_{-}(0) < \rho_{+}.$$

The above reasoning concerning the model $\mathbb{M}2+$ is also true for $\hat{P}_{-}(0)$. Therefore, in this case, we obtain the following limit values for initial conditions (12):

$$P_{-}^{*} = 1, \quad P_{+}^{*} = 0,$$

which completes the proof of Theorem 1 for the model $\mathbb{M}2$.

Model $\mathbb{M}$ 3. We first consider the case of $\mathbb{M}\,3+\Pi$, where $V>0$ and $\rho_+\geq 1$. The following estimate takes place:

$$-\rho_+ < \hat{P}_+(k) < \rho_- < 0\,,\ k\geq 0. \tag{24}$$

The monotone increase in the sequence

$$\hat{P}_+(k+1) > \hat{P}_+(k),\ k\geq 0,$$

follows from basic equation (14) and also from the inequality for fluctuations (24).

The boundedness of

$$\hat{P}_+(k+1) \leq \frac{3}{2}\rho_-$$

from above follows from inequalities (23) and (24). Therefore, there are limits (20), and limit equation (21) holds.

Note that $\hat{P}_+^* \neq 0$ since it is strictly negative by virtue of estimate (24). We also have $P_+^* \neq 0$ since $P_+(k+1) > P_+(k) > 0$. Hence, $P_-^* = 0$, i.e., $\lim_{k\to\infty} P_+(k) = 1$, which completes the proof of limits (13) for $\mathbb{M}\,3+\Pi$.

In the case of $\mathbb{M}\,3-\mathrm{O}$, where $V<0$ and $\rho_+\geq 1$, the dual problem arises (see (22)), namely, the following inequalities hold

$$-\rho_- \leq \hat{P}_-(k) \leq \rho_+\,.$$

Thus, the entire above reasoning concerning the probabilities of the negative alternative remains true, whence the following relationships hold:

$$P_-^* = 1\,,\quad P_+^* = 0,$$

which completes the proof of limits (13) of Theorem 1 for the model $\mathbb{M}\,3-\mathrm{O}$.

Similarly, Theorem 1 is proved in the case of $\mathbb{M}\,3+\mathrm{O}$ and in the dual case of $\mathbb{M}\,3-\Pi$.

Theorem 1 is proved.

REFERENCES

1. D. V. Korolyuk, “Binary statistical experiments with persistent nonlinear regression,” Theory of Probability and Mathematical Statistics, No. 91, 64–73 (2014).
2. F. Hoppensteadt, Mathematical Methods of Population Biology, Cambridge Univ. Press, Cambridge (1982).
3. S. N. Ethier and T. G. Kurtz, Markov Processes: Characterization and Convergence, Wiley, New York (1986).
4. Yu. M. Svirezhev and V. P. Pasekov, Foundations of Mathematical Genetics [in Russian], Nauka, Moscow (1982).
5. A. V. Skorokhod, F. Hoppensteadt, and H. Salehi, Random Perturbation Methods with Applications in Science and Engineering, Springer, New York (2000).
6. R. Bush and F. Mosteller, “A stochastic model with applications to learning,” The Annals of Mathematical Statistics, **24**, No. 4, 559–585 (1953).
7. N. M. Amosov, Algorithms of the Mind [in Russian], Naukova Dumka, Kiev (1979).