KYBERNETIKA — VOLUME 25 (1989), NUMBER 3

TEST FOR SUBMODEL IN GIBBS-MARKOV BINARY RANDOM SEQUENCE

MARTIN JANŽURA

The Gibbs-Markov random sequences (as studied in frame of statistical physics) are convenient as probability models for sequences of dependent binary data. Thus, the model is given by a system of interactions which may be understood and estimated as a vector parameter. Setting some of the interactions equal to zero, we obtain a submodel. A test for the submodel is derived in the present paper, and a numerical example with simulated data is included.

1. INTRODUCTION

Random sequences with binary state space occur in many problems of decision-making and control. The state space, in general, may be given by any dichotomy, but for the sake of simplicity we shall use the $\{0, 1\}$ representation.

In order to obtain reasonable results from the statistical procedures we suppose the generating random sequence to satisfy the "bilateral" R-Markov property with fixed positive integer R. We assume the bilateral Markov property rather than the unilateral one since we intend to apply the Gibbsian description of the distribution.

The Gibbsian description being applied, we obtain a parametrization which brings a deeper insight into the dependence structure. Thus, we may deal not only with the "transition probabilities" but also with their inner structure given by a system of interactions.

Naturally, some of the interactions may vanish, i.e. they may be equal to zero. Since the model is given by a system of non-zero interactions, we obtain a hierarchical class of models. And the aim of the present paper is to derive a method for testing a submodel. More precisely, if we know that a higher model takes place we ask whether the data can be explained with the aid of a given submodel.

The proposed method may be considered as a generalization of the method for testing submodel in contingency tables (cf. e.g. [1]). The fact is not surprising since the Gibbsian models for Markov random sequences mean a generalization of the log-linear models used for the contingency tables.

The Gibbs random sequences represent a special case of the Gibbs random fields, studied in frame of statistical physics. Thus, we use many results obtained in general for the random fields. As the main basic references we follow [8] and [10].

2. GIBBS-MARKOV BINARY RANDOM SEQUENCES

Let us denote by $X = \{0, 1\}$ the state space, by $\mathscr{F} = \exp X$ the σ-algebra of all its subsets, and by $\mathscr{Z}$ the set of all integers.

Let us fix positive integer R. By R-Markov binary random sequence (b.r.s.) we mean any probability measure μ on the product space $(X, \mathscr{F})^{\mathscr{Z}}$ satisfying

$$\mu(x_j \mid x_{\mathscr{Z}\setminus\{j\}}) = \mu(x_j \mid x_{\{j-R,\dots,j-1,j+1,\dots,j+R\}})$$

for every $j \in \mathscr{Z}$ and a.e. $x \in X^{\mathscr{Z}}[\mu]$. (We use the abbreviated notation $\mu(x_j \mid x_{\mathscr{Z}\setminus\{j\}})$ for the conditional distribution.)

We denote by $[R] = \{1, \dots, R\}$ the set of all positive integers not greater than R, and by $\mathscr{E} = \exp[R]$ the system of all subsets of $[R]$. For $A \in \mathscr{E}$ we put $\bar{A} = A \cup \{0\}$.

For a fixed non-void subsystem $\mathscr{A} \subset \mathscr{E}$ let us suppose a real-valued function

$$U\colon \mathscr{A} \to \mathscr{R}$$

to be given which is often called the potential (of range R). The assumed values

$$U(\mathscr{A}) = \{U_A\}_{A\in\mathscr{A}} \in \mathscr{R}^{\mathscr{A}}$$

are called the (finite range) interactions.

An R-Markov b.r.s. μ is called Gibbs b.r.s. with respect to the potential U (or, equivalently, with respect to the system of interaction $\{U_A\}_{A\in\mathscr{A}}$) if

$$\mu(x_j \mid x_{\{j-R,\dots,j-1,j+1,\dots,j+R\}}) =$$
$$= \left[\exp\left\{\sum_{A\in\mathscr{A}} U_A \sum_{t\in\bar{A}} \left(\prod_{k\in\bar{A}} x_{k-t+j}\right)\right\}\right] \cdot \left[1 + \exp\left\{\sum_{A\in\mathscr{A}} U_A \sum_{t\in\bar{A}} \left[\prod_{k\in\bar{A}\setminus\{t\}} x_{k-t+j}\right)\right\}\right]^{-1}$$

Thus, the interactions may be understood as parameters of the model.

For every potential $U \in \mathscr{R}^{\mathscr{A}}$ there exists exactly one Gibbs b.r.s. μ_U (cf. e.g. Theorem 5.6.2 in [10]), which is stationary and ergodic (cf. Proposition 5.4 and Proposition 4.1. in [8]).

Further, let $\mathbf{M}_U$ be the so called transfer matrix, i.e. the strictly positive-valued $(2^R \times 2^R)$-matrix with elements given by the formula

$$\mathbf{M}_U(x_{[R]}, z_{]R]}) = \exp\left\{\sum_{B\in\mathscr{A}} U_B \sum_{t=1}^{R} \left\{\prod_{k\in(B+t)\cap[R]} x_k \prod_{l\in(B+t-R)\cap[R]} z_l\right\}\right\}$$

for every $x_{[R]}, z_{[R]} \in X^{[R]}$.

By $\lambda_{\max}(\mathbf{M}_U)$ we denote the uniquely defined (due to the well-known Perron-Frobenius theorem) strictly positive eigenvalue of $\mathbf{M}_U$, which is greater in absolute value than all other eigenvalues.

Due to Theorem 5.6.2 in [10] the function

$$p(U) = R^{-1} \log \lambda_{\max}(\mathbf{M}_U)$$

(called the pressure) is a real analytic function on $\mathscr{R}^{\mathscr{A}}$.

Explicitly, denoting $C_B = \{x \in X^{\mathscr{Z}}, x_0 \prod_{i \in B} x_i = 1\}$ for every $B \in \mathscr{E}$, it holds that

$$\nabla p(U) = \left\{\frac{\partial p}{\partial U_A}(U)\right\}_{A \in \mathscr{A}} = \{\mu_U(C_A)\}_{A \in \mathscr{A}},$$

and

$$\nabla^2 p(U) = \left\{\frac{\partial^2 p}{\partial U_A \, \partial U_B}(U)\right\}_{A,B \in \mathscr{A}} = \left\{\sum_{k=-\infty}^{\infty} [\mu_U(C_B \cap T^{-k} C_A) - \mu_U(C_A)\, \mu_U(C_B)]\right\}_{A,B \in \mathscr{A}}$$

cf. Section 3 in [2] and Theorem 5.1 in [7]) are continuous functions of the interactions. Here $T: X^{\mathscr{Z}} \to X^{\mathscr{Z}}$ is the shift defined through $T(x)_k = x_{k+1}$ for every $k \in \mathscr{Z}$, $x \in X^{\mathscr{Z}}$.

Moreover, the function p is strictly convex (cf. Lemma 8.6 and Lemma 8.7 in [8]) wherefrom a one-to-one correspondence between U and μ_U follows, and even strongly convex (cf. [3]) wherefrom it especially follows that

$$\nabla^2 p(U) > 0 \quad \text{(i.e. positive definite) for every } U \in \mathscr{R}^{\mathscr{A}}.$$

Therefore, we may introduce the transform

$$\Phi: \mathscr{R}^{\mathscr{A}} \to \mathscr{R}^{\mathscr{A}}$$

given by

$$\Phi(U) = \{\mu_U(C_A)\}_{A \in \mathscr{A}} = \nabla p(U) \quad \text{for every } U \in \mathscr{R}^{\mathscr{A}},$$

and we observe that Φ is a one-to-one differentiable mapping with the Jacobi matrix

$$\nabla \Phi(U) = \left(\frac{\partial \mu_U(C_A)}{\partial U_B}\right)_{A,B \in \mathscr{A}} = \nabla^2 p(U)$$

continuous as a function of U and positive definite for every fixed U.

Thus, the inverse mapping Φ^{-1} exists on the open set $\Phi(\mathscr{R}^{\mathscr{A}}) \subset \mathscr{R}^{\mathscr{A}}$, being also differentiable with the Jacobi matrix

$$\nabla \Phi^{-1}(\theta) = (\nabla^2 p(\Phi^{-1}(\theta)))^{-1} > 0$$

inverse to the Jacobi matrix of Φ.

We finish this section with the appropriate version of the variational principle for Gibbs distributions (cf. e.g. Proposition 8.1 in [8]):

For every $U^0 \in \mathscr{R}^{\mathscr{A}}$ is holds

$$\min_{U \in \mathscr{R}^{\mathscr{A}}} (P(U) - \langle U, \Phi(U^0)\rangle) = p(U^0) - \langle U^0, \Phi(U^0)\rangle .$$

3. ESTIMATION

A finite sequence $(x_1, \ldots, x_N) \in \{0, 1\}^N$ of binary data is supposed to be generated by a Gibbs b.r.s. μ_{U^0} with $U^0 \in \mathscr{R}^{\mathscr{A}}$. Considering the interactions as a vector parameter we obtain a parameter estimation problem.

In the same way as in [5] we estimate at first the transformed parameter

$$\theta^0 = \Phi(U^0)$$

with the aid of the usual "frequency" estimator, i.e.

$$\hat{\theta}_A^N = (N - r(A))^{-1} \sum_{i=1}^{N-r(A)} \prod_{j \in \bar{A}} x_{j+i} \quad \text{for every } A \in \mathscr{A}$$

where $r(A) = \max \{j \in \bar{A}\}$.

Then we define the estimate of the interactions as the inverse transform, i.e.

$$\hat{U}^N = \Phi^{-1}(\hat{\theta}^N) \quad \text{for} \quad \hat{\theta}^N \in \Phi(\mathscr{R}^{\mathscr{A}})$$
$$\hat{U}^N = 0 \qquad \text{for} \quad \hat{\theta}^N \notin \Phi(\mathscr{R}^{\mathscr{A}}) .$$

(In fact, for $\hat{\theta}^N \notin \Phi(\mathscr{R}^{\mathscr{A}})$ we may define the estimate arbitrarily.)

Proposition 3.1. For every $U^0 \in \mathscr{R}^{\mathscr{A}}$ the estimate $\hat{\theta}^N$ of $\theta^0 = \Phi(U^0)$ is consistent and asymptotically normal with the covariance matrix given by $\nabla^2 p(U^0)$, i.e.

$$\hat{\theta}^N \to \theta^0 \quad \text{a.s.} \quad [\mu_{U^0}]$$

and

$$N^{1/2}(\hat{\theta}^N - \theta^0) \to \mathbb{N}(0, \nabla^2 p(U^0)) \quad \text{in distribution} \quad [\mu_{U^0}] .$$

The estimate $\hat{U}^N$ of U^0 is consistent and asymptotically normal with the covariance matrix $(\nabla^2 p(U^0))^{-1}$.

Proof. Cf. proofs of Theorem 3.1 and Theorem 3.2 in [5]. □

Remark. Due to the variational principle introduced in the preceding section we obtain the estimate $\hat{U}^N$ by minimization of

$$p(U) - \langle U, \hat{\theta}^N \rangle .$$

4. TEST FOR SUBMODEL

Suppose we want to test the hypothesis that some of the interactions are equal to zero. Thus, if we denote

$$\mathscr{R}^{\mathscr{A}|\mathscr{B}} = \{\alpha \in \mathscr{R}^{\mathscr{A}}; \alpha_B = 0 \quad \text{for} \quad B \in \mathscr{A} \setminus \mathscr{B}\}$$

where $\mathscr{B} \subset \mathscr{A}$ is a given subsystem, we may formulate the hypothesis as follows

$$\mathsf{H}_0 \colon U^0 \in \mathscr{R}^{\mathscr{A}|\mathscr{B}} .$$

We try to construct such a test statistics which does not depend on the actual values of the remaining interactions U_B, $B \in \mathscr{B}$.

We shall use the concepts introduced in the preceding sections, but with added subscripts $\mathscr{A}$ or $\mathscr{B}$ in order to indicate which subsystem is considered to be fixed at the moment. Namely, the idea of the test consists in comparison of the attained minima of the function

$$p_{\mathscr{A}}(U) - \langle U, \hat{\theta}^N \rangle ,$$

when minimizing over $\mathscr{R}^{\mathscr{A}}$ and $\mathscr{R}^{\mathscr{A}|\mathscr{R}}$, respectively. And the minimization over $\mathscr{R}^{\mathscr{A}|\mathscr{B}}$ may be identified with the minimization of the function

$$p_{\mathscr{B}}(U) - \langle U, \mathrm{Pr}_{\mathscr{B}}^{\mathscr{A}}(\hat{\theta}^N) \rangle$$

over $\mathscr{R}^{\mathscr{B}}$ (here $\mathrm{Pr}_{\mathscr{B}}^{\mathscr{A}} : \mathscr{R}^{\mathscr{A}} \to \mathscr{R}^{\mathscr{B}}$ is the projection), i.e. with the estimation of the parameters of the submodel.

Lemma 4.1. If $\theta \in \Phi_{\mathscr{A}}(\mathscr{R}^{\mathscr{A}})$ then $\mathrm{Pr}_{\mathscr{B}}^{\mathscr{A}}(\theta) \in \Phi_{\mathscr{B}}(\mathscr{R}^{\mathscr{B}})$.

Proof. The condition $\theta \in \Phi_{\mathscr{A}}(\mathscr{R}^{\mathscr{A}})$ means that the strictly convex function

$$F(U) = p_{\mathscr{A}}(U) - \langle U, \theta \rangle$$

assumes its minimum at a unique point $U^* \in \mathscr{R}^{\mathscr{A}}$.

Therefore the function F is coercive. Namely, we have

$$F(U) \geqq F(U^*) + \varepsilon \| U - U^* \|$$

for $U \in \mathscr{R}^{\mathscr{A}}$ satisfying $\| U - U^* \| \geqq 1$, where

$$\varepsilon = \min \{ F(U - F(U^*); \| U - U^* \| = 1 \} > 0 .$$

Hence the restriction of the function F to the subspace $\mathscr{R}^{\mathscr{A}|\mathscr{R}}$ remains strictly convex and coercive, and according to Proposition 1.2 in [4], it assumed its minimum as well. □

Thus we may define a function $\mathfrak{G} : \Phi_{\mathscr{A}}(\mathscr{R}^{\mathscr{A}}) \to \mathscr{R}$ through the formula

$$\mathfrak{G}(\theta) = p_{\mathscr{B}}(\Phi_{\mathscr{B}}^{-1}(\mathrm{Pr}_{\mathscr{B}}^{\mathscr{A}}(\theta))) - \langle \Phi_{\mathscr{B}}^{-1}(\mathrm{Pr}_{\mathscr{B}}^{\mathscr{A}}(\theta)), \mathrm{Pr}_{\mathscr{B}}^{\mathscr{A}}(\theta) \rangle - \\ - [p_{\mathscr{A}}(\Phi_{\mathscr{A}}^{-1}(\theta)) - \langle \Phi_{\mathscr{A}}^{-1}(\theta), \theta \rangle] .$$

Lemma 4.1. It holds

$$\nabla \mathfrak{G}(\theta) = \Phi_{\mathscr{A}}^{-1}(\theta) - \begin{pmatrix} \Phi_{\mathscr{B}}^{-1}(\mathrm{Pr}_{\mathscr{B}}^{\mathscr{A}}(\theta)) \\ 0_{\mathscr{A} \setminus \mathscr{B}} \end{pmatrix},$$

$$\nabla^2 \mathfrak{G}(\theta) = [\nabla^2 p_{\mathscr{A}}(\Phi_{\mathscr{A}}^{-1}(\theta)]^{-1} - \begin{pmatrix} [\nabla^2 p_{\mathscr{B}}(\Phi_{\mathscr{B}}^{-1}(\mathrm{Pr}_{\mathscr{B}}^{\mathscr{A}}(\theta)))]^{-1} & 0_{\mathscr{B}, \mathscr{A} \setminus \mathscr{B}} \\ 0_{\mathscr{A} \setminus \mathscr{B}, \mathscr{B}} & 0_{\mathscr{A} \setminus \mathscr{B}, \mathscr{A} \setminus \mathscr{B}} \end{pmatrix}.$$

Especially, for $\theta \in \Phi_{\mathscr{A}}(\mathscr{R}^{\mathscr{A}|\mathscr{B}})$ it holds

$$\nabla \mathfrak{G}(\theta) = 0 ,$$

$$\nabla^2 \mathfrak{G}(\theta) = [\nabla^2 p_{\mathscr{A}}(\Phi_{\mathscr{A}}^{-1}(\theta))]^{-1} - \begin{pmatrix} [\nabla^2 p_{\mathscr{B}}(\mathrm{Pr}_{\mathscr{B}}^{\mathscr{A}}(\Phi_{\mathscr{A}}^{-1}(\theta)))]^{-1} & 0_{\mathscr{B}, \mathscr{A} \setminus \mathscr{B}} \\ 0_{\mathscr{A} \setminus \mathscr{B}, \mathscr{B}} & 0_{\mathscr{A} \setminus \mathscr{B}, \mathscr{A} \setminus \mathscr{B}} \end{pmatrix}.$$

Proof. All the results can be obtained by direct calculations. □

Due to the consistency of the estimate $\hat{\theta}^N$ we know that

$$\mu_{U^0}(\hat{\theta}^N \in \Phi_{\mathscr{A}}(\mathscr{R}^{\mathscr{A}})) \to 1 \quad \text{as} \quad N \to \infty .$$

Thus, for $\hat{\theta}^N \notin \Phi_{\mathscr{A}}(\mathscr{R}^{\mathscr{A}})$ we may define $\mathfrak{G}(\hat{\theta}^N)$ arbitrarily.

Now, we may formulate the main result of the paper.

Theorem 4.2. For every $U^0 \in \mathscr{R}^{\mathscr{A}|\mathscr{B}}$ it holds

$$2.N.\,\mathfrak{G}(\hat{\theta}^N) \to \chi_f^2 \text{ in distribution } [\mu_{U^0}] ,$$

where $f = \text{card}\,(\mathscr{A}) - \text{card}\,(\mathscr{B})$.

Proof. Let $\theta^0 = \Phi_{\mathscr{A}}(U^0)$. For $\hat{\theta}^N \in \Phi_{\mathscr{A}}(\mathscr{R}^{\mathscr{A}})$ we use the well-known Taylor formula to obtain

$$\mathfrak{G}(\hat{\theta}^N) = \mathfrak{G}(\theta^0) + \langle \nabla\mathfrak{G}(\theta^0), \hat{\theta}^N - \theta^0 \rangle +$$
$$+ \tfrac{1}{2}(\hat{\theta}^N - \theta^0)^{\mathrm{T}} \nabla^2\mathfrak{G}(\theta^0 + \gamma_N(\hat{\theta}^N - \theta^0))\,(\hat{\theta}^N - \theta^0)$$

for some $\gamma_N \in [0, 1]$. Since $\mathfrak{G}(\theta^0) = 0$ and $\nabla\mathfrak{G}(\theta^0) = 0$ we set for $\hat{\theta}^N \notin \Phi_{\mathscr{A}}(\mathscr{R}^{\mathscr{A}})$ directly

$$\mathfrak{G}(\hat{\theta}^N) = \tfrac{1}{2}(\hat{\theta}^N - \theta^0)^{\mathrm{T}} \nabla^2\mathfrak{G}(\theta^0)\,(\hat{\theta}^N - \theta^0) .$$

By Proposition 3.1 it holds

$$\hat{\theta}^N \to \theta^0 \quad \text{a.s.} \quad [\mu_{U^0}]$$

and

$$N^{1/2}(\hat{\theta}^N - \theta^0) \to \mathbb{N}(0, \nabla^2 p(U^0)) \quad \text{in distribution} \quad [\mu_{U^0}] .$$

Therefore it follows from the known limit theorem (cf. e.g. 2c. 4(x) in [9]) that

$$2.N.\mathfrak{G}(\hat{\theta}^N) \to \mathfrak{L}(\zeta^{\mathrm{T}} \nabla^2\mathfrak{G}(\theta^0)\,\zeta) \quad \text{in distribution } [\mu_{U^0}] ,$$

where $\mathfrak{L}(\zeta) = \mathbb{N}(0, \nabla^2 p_{\mathscr{A}}(U^0))$. And since

$$\nabla^2\mathfrak{G}(\theta^0)\,\nabla p_{\mathscr{A}}(U^0) = \begin{pmatrix} 0_{\mathscr{B},\mathscr{B}} & 0_{\mathscr{B},\mathscr{A}\setminus\mathscr{B}} \\ 0_{\mathscr{A}\setminus\mathscr{B},\mathscr{B}} & I_{\mathscr{A}\setminus\mathscr{B},\mathscr{A}\setminus\mathscr{B}} \end{pmatrix}$$

is idempotent with the trace equal to $f = \text{card}\,(\mathscr{A}) - \text{card}\,(\mathscr{B})$, we finally obtain by 3b.4 (II) in [9] that

$$\mathfrak{L}(\zeta^{\mathrm{T}} \nabla^2\mathfrak{G}(\theta^0)\,\zeta) = \chi_f^2 . \qquad \square$$

Thus, if we want to test the hypothesis

$$\mathsf{H}_0 \colon U^0 \in \mathscr{B}^{\mathscr{A}|\mathscr{B}}$$

against the alternative

$$\mathsf{H}_1 \colon U^0 \in \mathscr{B}^{\mathscr{A}} \setminus \mathscr{B}^{\mathscr{A}|\mathscr{B}}$$

we reject the hypothesis whenever the statistics

$$2 \,.\, N \,.\, \mathfrak{G}(\hat{\theta}^N)$$

exceeds the corresponding quantile of the χ^2-distribution with $\text{card}\,(\mathscr{A}) - \text{card}\,(\mathscr{B})$ degrees of freedom.

There is an important aspect which is worth to be mentioned.

For a pair μ, ν of b.r.s.'s let us define the information gain of μ with respect to ν by

$$H(\mu \mid \nu) = \lim_{n\to\infty} n^{-1} \int \log \frac{\mu(x_{[n]})}{\nu(x_{[n]})} \, \mathrm{d}\mu(x) ,$$

providing the expressions make sense and the limit exists. According to Theorem 8.2 and Proposition 8.1 in [8] we have

$$\mathfrak{G}(\hat{\theta}^N) = H(\mu_{\hat{U}^N_{\mathscr{A}}} \mid \mu_{\hat{U}^N_{\mathscr{B}}})$$

where

$$\hat{U}^N_{\mathscr{A}} = \Phi^{-1}_{\mathscr{A}}(\hat{\theta}^N)$$

and

$$\hat{U}^N_{\mathscr{B}} = \Phi^{-1}_{\mathscr{B}}(\mathrm{Pr}^{\mathscr{A}}_{\mathscr{B}}(\hat{\theta}^N)) .$$

Therefore the test based on the statistics

$$2 \,.\, N \,.\, \mathfrak{G}(\hat{\theta}^N)$$

may be understood as a direct generalization of the test known for the contingency tables (cf. e.g. [1]).

5. EXAMPLE

Let us introduce a single example in order to illustrate the derived method.

We simulated "unilateral" 3-Markov sequences with the transition probabilities (four various cases) given by the following table. (We employed the "unilateral" Markov sequences since their simulation is easier and the R-Markov b.r.s.'s obey the "unilateral" R-Markov property as well – cf. e.g. Section 5 in [8].)

			$\log [\mathbf{P}_i(1 \mid x_1, x_2, x_3)/\mathbf{P}_i(0 \mid x_1, x_2, x_3)]$			
x_1	x_2	x_3	$i=0$	$i=1$	$i=2$	$i=3$
0	0	0	0	−0·5	−1	−1·5
0	0	1	0	−0·5	−1	−0·5
0	1	0	0	−0·5	0	−0·5
0	1	1	0	−0·5	0	0·5
1	0	0	0	0·5	0	−0·5
1	0	1	0	0·5	0	0·5
1	1	0	0	0·5	1	0·5
1	1	1	0	0·5	1	1·5

Thus, we have $R = 3$, $\mathscr{A} = \exp\{1, 2, 3\}$, card $\mathscr{A} = 8$.

We consider the following subsystems

$$\mathscr{B}_0 = \{0\} \qquad \text{card } \mathscr{B}_0 = 1$$
$$\mathscr{B}_1 = \{0, \{1\}\} \qquad \text{card } \mathscr{B}_1 = 2$$

$$\mathscr{B}_2 = \{0, \{1\}, \{2\}\} \qquad \text{card}\, \mathscr{B}_2 = 3$$
$$\mathscr{B}_3 = \{0, \{1\}, \{2\}, \{3\}\} \qquad \text{card}\, \mathscr{B}_3 = 4\,.$$

Let us denote

$$\mathsf{H}_i\colon\ \boldsymbol{U}^0 \in \mathscr{R}^{\mathscr{A}|\mathscr{B}_i} \quad \text{for} \quad i = 0, \ldots, 3$$

the hypotheses to be tested.

If H_0 takes place the sequence consists of independent random variables. If H_i, $i = 1, \ldots, 3$, holds the sequence is i-Markov with only pair-wise interactions.

For every case $\mathbf{P}_i$, $i = 0, \ldots, 3$ we repeated the simulation three times with various initial values and fixed length $N = 1000$ of every sequence. The results are arranged in the following table. We denote the test statistics $2\,.\,N\,.\,\mathbb{G}^{\mathscr{A}}_{\mathscr{B}_i}(\hat{\theta}^N)$ simply by $\hat{\chi}^2(i)$ for $i = 0, \ldots, 3$.

	$\mathbf{P}_0$			$\mathbf{P}_1$			$\mathbf{P}_2$			$\mathbf{P}_3$			f	$\chi^2_{f\,(0{\cdot}05)}$
$\hat{\chi}^2(0)$	5·26	7·44	6·97	81·65	54·97	85·38	177·98	157·85	119·68	206·06	249·56	302·99	7	14·07
$\hat{\chi}^2(1)$	3·02	6·88	4·64	8·53	3·95	3·11	109·59	48·21	48·71	127·55	73·43	117·83	6	12·59
$\hat{\chi}^2(2)$	2·66	6·88	4·52	5·00	3·26	1·26	7·98	2·07	2·92	54·76	34·00	53·34	5	11·07
$\hat{\chi}^2(3)$	2·41	6·88	4·40	4·87	3·06	1·22	4·45	1·09	1·11	6·75	2·71	4·14	4	9·49

The transition probabilities $\mathbf{P}_i$, $i = 0, \ldots, 3$, were chosen in order to correspond (at least approximately) to the hypotheses H_i, $i = 0, \ldots, 3$, respectively. The above results agree quite well with the assumptions since for the data simulated with the aid of $\mathbf{P}_i$ the test rejects just the hypotheses H_j, $j < i$.

6. CONCLUDING REMARK

The proposed method does not depend on the assumption of binary state space. It could be applied to a more general state space as well. The only problem will be with the numerical calculation of the pressure (cf. Section 2). The same problem occurs if we generalize the method for random fields instead of random sequences (cf. [6]).

(Received June 22, 1988.)

REFERENCES

[1] Y. M. M. Bishop, S. E. Fienberg and P. W. Holland: Discrete Multivariate Analysis: Theory and Practice. MIT Press, Cambridge, Mass. 1975.

[2] R. L. Dobrushin: Conditions of absence of phase transitions for one-dimensional classical model (in Russian). Mat. Sbornik *93* (1974), 29—49.

[3] R. L. Dobrushin and B. S. Nahapetian: Strong convexity of the pressure for lattice systems of classical statistical physics (in Russian). Teoret. Mat. Fiz. *20* (1974), 223—234.

[4] I. Ekeland and R. Temam: Convex Analysis and Variational Problems. North-Holland, Amsterdam 1976.

[5] M. Janžura: Estimating interactions in binary data sequences. Kybernetika *22* (1986), 277—284.

[6] M. Janžura: Estimating interactions in binary lattice data with nearest-neighbor property. Kybernetika *23* (1987), 136—142.
[7] H. Künsch: Decay of correlations under Dobrushin's uniqueness condition and its applications. Commun. Math. Phys. *84* (1982), 207—222.
[8] C. Preston: Random Fields. (Lecture Notes in Mathematics 534.) Springer-Verlag, Berlin—Heidelberg—New York 1976.
[9] C. R. Rao: Linear Statistical Inference and its Applications. John Wiley and Sons, New York 1973.
[10] D. Ruelle: Statistical Mechanics. Rigorous Results. Benjamin, New York 1969.

RNDr. Martin Janžura, CSc., Ústav teorie informace a automatizace ČSAV (Institute of Information Theory and Automation — Czechoslovak Academy of Sciences), Pod vodárenskou věží 4, 182 08 Praha 8. Czechoslovakia.