ABSTRACT

1. It has been shown recently [1] that the unique solution of the basic statistical problem of the multinomial process which has been obtained as part of the so called Bayes-Laplace-statistics (BL-statistics) can be used to determine the a priori unknown transition parameters of Markov chains by proper evaluation of measured transition counts.

2. The work reported in [1] is continued in the present paper in order to find the empirical stationary distribution function of Markov chains with unknown transition parameters and of processes with correlated random sequences whose behaviour can be described by a Markov chain. For this purpose the BL-statistics is developed with respect to the following tasks.

2.1 First we consider a Markov chain with two node only, whose transition parameters are unknown. We measure $n_0$ transitions from node 0 of which $y_0$ lead to node 0 and $n_0-y_0$ to node 1; and $n_1$ transitions from node 1 of which $y_1$ lead to node 1 and $n_1-y_1$ to node 0. It is shown that the posterior knowledge of the stationary probability $P_0$ for the chain being in state 0 is uniquely described by a density $f(P_0|y_0,n_0,y_1,n_1)$ whose moments $M_1$ and $M_2$ are given by weighted sums of Gauss hypergeometric functions $F(a,b;c;x)$ [2] and express the average stationary probability $P_0 = M_1$ resp. the relative error when stating $P_0$ by $d_0 = [M_2/M_1^2 - 1]^{1/2}$.

2.2 Next we consider a Markov chain whose $k+1$ nodes ($k=1,2,\ldots$) are interconnected by an arbitrary transition matrix with unknown transition parameters $P_{ij} = P(j|i)$; $i,j = 0,1,\ldots,k$. We may divide this chain in two parts, the left part consisting of the nodes $0,1,\ldots,r$ and the right part of the nodes $r+1, r+2, \ldots, k$. Then for any $r = 0,1,\ldots,k-1$ the left resp. right part can be attributed to node 0 resp. 1 of a 2-node Markov chain, called the "r-chain", and we may measure the number of transitions $n_r(r)$, $y_r(r)$ resp. $n_{g(r)}, y_{g(r)}(r)$ for every r-chain as defined in section 2.1. Using the results obtained in section 2.1 we may compute the average stationary probability $M_1(r) = F_r(r)$ as well as the second moment $M_2(r)$ and then the desired empirical stationary d.f. $F_r(x)$ and its relative error $d_r(x)$ of the $(k+1)$-node Markov chain:

$$F_r(x) = M_1(r); \quad d_r(x) = [M_2/M_1^2(x) - 1]^{1/2}; \quad r < x \leq r+1; \quad r = 0,1,\ldots,k-1.$$ (1)

2.3 Finally we consider a continuous random process with an unknown stationary d.f. $F(x)$. In $n+1$ trials we measure and store the chronological vector $(x_r) = (x_0,x_1,\ldots,x_r)$ expressing the information which value $x_{t+1}$ followed value $x_t$, $t = 0,1,\ldots,n-1$. After sorting the ordered vector $(x_r) = (x_0,x_1,\ldots,x_r)$, $x_{r+1} > x_r$ can be attributed to a $(n+1)$-node Markov chain with one transition only from each node except no transition from the "final node" belonging to the last measured value $x_r$, $t = n$. Since the transition counts $n_r(r)$, $y_r(r)$ and $n_{g(r)}, y_{g(r)}(r)$ of the 2-node r-chain as defined in section 2.2 can be calculated for every $r = 0,1,\ldots,n-1$ we may obtain from Eq.(1) the empirical stationary d.f. $F_r(x)$ and associated error $d_r(x)$ as step functions with one step at each point $x_r$. This result should be compared with the empirical d.f. $F_r(x)$ in case of independent x-values [3].

3. The solution of section 2 for determining the stationary d.f. of Markov chains resp. of continuous processes with correlation effects is considered to belong to the class of objective solutions having principal importance within the BL-statistics. In order to make this solution available for practical work (e.g. for the evaluation of simulated teletraffic data) methods will be discussed for reducing the necessary computertime e.g. by collecting the measured x-values in intervals and/or by suitable approximations of the exact formulae for $F_r(x)$ and $d_r(x)$. Moreover it seems desirable to introduce here too an equivalent to the LRE-algorithm [4].

REFERENCES


