On the statistical properties related to an algorithm of constructing the probability density functions

By

Kensuke TANAKA

(Received May 1, 1972)

1. Introduction and summary

In this paper, we shall make use of an algorithm which was introduced by T. KITAGAWA [7] as a method of statistical treatment of the controlled system and afterward was developed by S. KANŌ [3], [4], [5] and [6]. In these days, this algorithm is used in many practical applications.

The algorithm, which may be called the statistical control, is a method of constructing the prescribed processes from the unknown original stochastic processes by an infinite iteration of a certain linear transformation on the basis of observations of the unknown processes at each stage. Moreover, in [5] S. KANŌ showed that the prescribed probability distributions are successively constructed from the unknown probability distributions of the original processes by the above algorithm.

Especially, this paper is concerned with [5] in which the unknown original stochastic process is assumed to have a finite state space. But, in this place, we shall treat the problem of constructing some prescribed continuous probability density function from a continuous probability density function of an original stochastic process defined on R^1. In order to solve the problem with respect to an independent original stochastic process, we shall use an algorithm of transforming linearly the unknown continuous probability density function by using the estimate of the unknown continuous probability density function in [10] and the prescribed continuous probability density function at each stage. Then, we shall discuss the asymptotically statistical properties of a limit distribution constructed by an infinite iteration of the linear transformation. Furthermore, in the case when the unknown original stochastic process is a simple Markoff process with stationary transition probabilities, we shall discuss the asymptotically statistical properties related to the same algorithm of constructing the prescribed transition probability density functions from the unknown transition probability density functions.

This paper consists of three sections. In Section 2, we shall treat the case of an independent original stochastic process and in Section 3 the case of an original Markoff process with stationary transition probabilities.

The author wishes to express his hearty thanks to Professors of Kyushu University, S. KANŌ and N. FURUKAWA for their kind advices.

2. The statistical properties in the case of an independent original stochastic process

Let an original stochastic process $\{X_t, t=0, 1, 2, \cdots\}$ be an independent process with an unknown probability density function $p_0(x)$ defined on R^1 satisfying the conditions:

(A. 1) $p_0(x)>0$ for all $x \in R^1$,

(A. 2) $\sup_{x \in R^1} p_0(x)=M_0<\infty$,

(A. 3) $p_0(x)$ is uniformly continuous in x.

The purpose of this section is to transform an original stochastic process with $p_0(x)$ into a process with the prescribed probability density function $q(x)$ defined on R^1 satisfying the conditions:

(B. 1) $q(x)>0$ for all $x \in R^1$,

(B. 2) $\sup_{x \in R^1} q(x)=M<\infty$,

(B. 3) $q(x)$ is uniformly continuous in x.

To solve this problem, we make use of the following algorithm. At the first stage, random samples of size n_0 are drawn from the population expressed by the random variable X_0 and an estimate of $p_0(x)$ for each x,

(2. 1) $$\hat{p}_0(x)=\frac{1}{n_0 h_0(n_0)} \sum_{i=1}^{n_0} K_0\left(\frac{x-x_i(0)}{h_0(n_0)}\right)$$

is calculated, where $x_i(0)$, $i=1, 2, \cdots n_0$, are the n_0 sample values, $\{h_0(n), n=1, 2, \cdots\}$ is a sequence of positive real numbers satisfying the conditions:

(C_0 1) $1 \geqq h_0(1) \geqq h_0(2) \cdots\cdots$ and $\lim_{n\to\infty} h_0(n)=0$,

(C_0 2) $\lim_{n\to\infty} n h_0(n)=\infty$,

and $K_0(x)$ is a real-valued function defined on R^1 satisfying the conditions:

(K_0 1) $K_0(x) \geqq 0$ for all $x \in R^1$ and $\int_{-\infty}^{\infty} K_0(x)dx=1$,

(K_0 2) $\sup_{x \in R^1} K_0(x)=K_0<\infty$,

(K_0 3) $\int_{-\infty}^{+\infty} K_0^2(x)\,dx=a_0<\infty$,

(K_0 4) $K_0(x)$ is uniformly continuous in x.

Then, we transform the probability density function $p_0(x)$ into $p_1(x) \equiv p_0(x)+q(x)-\widehat{p_0}(x)$. This fact means that the stochastic process $\{X_t,\ t=1, 2, \cdots\}$ was transformed into an independent process $\{X_{1t},\ t=1, 2, \cdots\}$ with the unknown probability density function $p_1(x)$. At the second stage, again random samples of size n_1 are drawn from the population expressed by the random variable X_{11} and an estimate of $p_1(x)$ for each x,

(2. 2) $$p_1(x)=\frac{1}{n_1 h_1(n_1)}\sum_{i=1}^{n_1} K_1\left(\frac{x-x_i(1)}{h_1(n_1)}\right)$$

is calculated, where $x_i(1)$, $i=1, 2, \cdots, n_1$, are the n_1 sample values, $\{h_1(n),\ n=1, 2, \cdots\}$ is a sequence of positive real numbers satisfying the conditions:

(C_1 1) $1 \geqq h_1(1) \geqq h_1(2) \geqq \cdots\cdots$ and $\lim_{n\to\infty} h_1(n)=0$,

(C_1 2) $\lim_{n\to\infty} n h_1(n)=\infty$,

and $K_1(x)$ is a real-valued function defined on R^1 satisfying the conditions:

(K_1 1) $K_1(x) \geqq 0$ for all $x \in R^1$ and $\int_{-\infty}^{\infty} K_1(x)dx=1$,

(K_1 2) $\sup_{x\in R^1} K_1(x)=K_1<\infty$,

(K_1 3) $\int_{-\infty}^{\infty} K_1^2(x)\,dx=a_1<\infty$

(K_1 4) $K_1(x)$ is uniformly continuous in x.

Then, we transform the probability density function $p_1(x)$ into $p_2(x) \equiv p_1(x)+q(x)-\widehat{p_1}(x)$. This fact means that the stochastic process $\{X_{1t},\ t=2, 3, \cdots\}$ was transformed into an independent process $\{X_{2t},\ t=2, 3, \cdots\}$ with the unknown probability density function $p_2(x)$. Generally, at the m-th stage, random samples of size n_{m-1} are drawn from the population expressed by the random variable $X_{m-1,\ m-1}$ and an estimate of $p_{m-1}(x)$ for each x,

(2. 3) $$\hat{p}_{m-1}(x)=\frac{1}{n_{m-1}h_{m-1}(n_{m-1})}\sum_{i=1}^{n_{m-1}} K_{m-1}\left(\frac{x-x_i(m-1)}{h_{m-1}(n_{m-1})}\right)$$

is calculated, where $x_i(m-1)$, $i=1, 2, \cdots, n_{m-1}$, are the n_{m-1} sample values, $\{h_{m-1}(n),\ n=1, 2, \cdots\}$ is a sequence of positive real numbers satisfying the conditions:

(C_{m-1}1) $1 \geqq h_{m-1}(1) \geqq h_{m-1}(2) \geqq \cdots$ and $\lim_{n\to\infty} h_{m-1}(n)=0$,

(C_{m-1}2) $\lim_{n\to\infty} n h_{m-1}(n)=\infty$,

and $K_{m-1}(x)$ is a real-valued function defined on R^1 satisfying the conditions:

(K_{m-1}1) $K_{m-1}(x) \geqq 0$ for all $x \in R^1$ and $\int_{-\infty}^{\infty} K_{m-1}(x)dx=1$,

(K_{m-1}2) $\sup_{x\in R^1} K_{m-1}(x)=K_{m-1}<\infty$,

$(K_{m-1}3)$ $\int_{-\infty}^{\infty} K^2{}_{m-1}(x)dx = a_{m-1} < \infty,$

$(K_{m-1}4)$ $K_{m-1}(x)$ is uniformly continuous in x.

Then, we transform $p_{m-1}(x)$ into $p_m(x) \equiv p_{m-1}(x) + q(x) - \widehat{p}_{m-1}(x)$. This fact means again that the stochastic process $\{X_{m-1,t},\ t=m,\ m+1, \cdots\}$ was transformed into an independent process $\{X_{m,t},\ t=m,\ m+1, \cdots\}$ with the unknown probability density function $p_m(x)$.

Next, we mention without proof the lemma given by E. Parzen [10] in order to prove main results in the paper.

Lemma 1. *Let $K(x)$ be a real-valued function defined on R^1 satisfying the following conditions:*

(2. 4) $K(x) \geqq 0$ *for all* $x \epsilon R^1$,

(2. 5) $\sup_{x \epsilon R^1} K(x) = K < \infty,$

and $f(x)$ be a real-valued function defined on R^1 satisfying the following conditions:

(2. 6) $\sup_{x \epsilon R^1} |f(x)| < \infty$ *and* $\int_{-\infty}^{\infty} |f(x)| dx < \infty,$

(2. 7) $f(x)$ *is uniformly continuous in* x.

Then, it holds that

(2. 8) $\lim_{n \to \infty} f_n(x) = p_0(x) \int_{-\infty}^{\infty} K(y)dy$ *uniformly in* x,

where

(2. 9) $f_n(x) = \frac{1}{h(n)} \int_{-\infty}^{\infty} K\left(\frac{x-y}{h(n)}\right) p_0(y)dy$

and $\{h(n),\ n=1, 2, \cdots\}$ is a sequence of positive real numbers satisfying the condition:

(2. 10) $1 \geqq h(1) \geqq h(2) \geqq \cdots$ *and* $\lim_{n \to \infty} h(n) = 0.$

The two statistical properties related to the algorithm are stated in the following theorems.

Theorem 2. 1

(2. 11) $E[p_{m+1}(x)] \approx q(x)$ *for sufficiently large* n_m,

where $\approx$ denotes asymptotic equality.

Proof. By using Lemma 1, we can obtain

$$E[p_{m+1}(x)] = E\{E[p_{m+1}(x) | p_m]\}$$
$$= E\{E[(p_m(x) - \widehat{p}_m(x)]) | p_m]\} + q(x)$$

$$= \mathrm{E}\Big[p_m(x) - \frac{1}{h_m(n_m)}\int_{-\infty}^{\infty} K_m\Big(\frac{x-y}{h_m(n_m)}\Big)p_m(y)\,dy\Big] + q(x)$$

$$\approx q(x) \qquad \text{as } n_m \longrightarrow \infty,$$

where $\mathrm{E}[\,\cdot\,|p_m]$ denotes the conditional expectation given the probability density function p_m at the m-th stage.

THEOREM 2. 2 *For sufficiently large* n_l, $l=1, 2, \cdots, m$,

(2. 12)
$$\mathrm{E}[(p_{m+1}(x)-q(x))^2] \approx$$
$$\approx \sum_{i=0}^{m-1}(-1)^i\Big(\prod_{j=0}^{i}\frac{1}{n_{m-j}}\Big)\Big(\frac{a_{m-i}}{h_{m-i}(n_{m-i})}-q(x)\Big)q(x)$$
$$+\prod_{j=0}^{m-1}\Big(\frac{-1}{n_{m-j}}\Big)\frac{1}{n_0}\Big(\frac{a_0}{h_0(n_0)}-p_0(x)\Big)p_0(x).$$

If, for $i=0, 1, \cdots, m$, $n_i=n$,

$$h_i(l)=h(l),\ l=1, 2, \cdots,$$

and $K_i(x)=K(x)$ *for all* $x\in R^1$, *then*

(2. 13)
$$\lim_{n\to\infty}\mathrm{E}[(p_{m+1}(x)-q(x))^2] \approx \frac{1}{n}\Big(\frac{a}{h(n)}-q(x)\Big)q(x),$$

where $a_m=\int_{-\infty}^{\infty}K_m^2(x)dx$ and $a=\int_{-\infty}^{\infty}K^2(x)dx$.

PROOF. From $p_{m+1}(x)=p_m(x)+q(x)-\hat{p}_m(x)$ and Lemma 1, we have

(2. 14)
$$\mathrm{E}[(p_{m+1}(x)-q(x))^2] = \mathrm{E}\{\mathrm{E}[(p_m(x)-\hat{p}_m(x))^2|p_m]\}$$
$$= \mathrm{E}\Big[\mathrm{E}\Big[\Big\{\frac{1}{n_m}\sum_{i=1}^{n_m}\Big(p_m(x)-\frac{1}{h_m(n_m)}K_m\Big(\frac{x-x_i(m)}{h_m(n_m)}\Big)\Big)\Big\}^2|p_m\Big]\Big]$$
$$\approx \mathrm{E}\Big[\frac{(p_m(x))^2}{n_m}-\frac{2p_m(x)}{n_m}\int_{-\infty}^{\infty}\frac{1}{h_m(n_m)}K_m\Big(\frac{x-y}{h_m(n_m)}\Big)p_m(y)dy$$
$$+\frac{1}{n_m}\int_{-\infty}^{\infty}\Big\{\frac{1}{h_m(n_m)}K_m\Big(\frac{x-y}{h_m(n_m)}\Big)\Big\}^2 p_m(y)\,dy\Big]$$
$$\approx \mathrm{E}\Big[\Big(\frac{a_m}{n_m h_m(n_m)}-\frac{p_m(x)}{n_m}\Big)p_m(x)\Big].$$

Repeating this calculation of conditional expectation, (2. 14) is reduced to

(2. 15)
$$\mathrm{E}\Big[\Big(\frac{a_m}{n_m h_m(n_m)}-\frac{p_m(x)}{n_m}\Big)p_m(x)\Big]$$
$$= \mathrm{E}\Big[\Big\{\frac{a_m}{n_m h_m(n_m)}-\frac{q(x)+(p_{m-1}(x)-\hat{p}_{m-1}(x))}{n_m}\Big\}\{q(x)$$
$$+(p_{m-1}(x)-\hat{p}_{m-1}(x))\}\Big]$$

$$\approx \frac{1}{n_m}\Big(\frac{a_m}{h_m(n_m)}-q(x)\Big)q(x)-\frac{1}{n_m}\mathrm{E}[(p_{m-1}(x)-\widehat{p}_{m-1}(x))^2]$$

$$\approx \frac{1}{n_m}\Big(\frac{a_m}{h_m(n_m)}-q(x)\Big)q(x)-\frac{1}{n_m n_{m-1}}\Big(\frac{a_{m-1}}{h_{m-1}(n_{m-1})}-q(x)\Big)q(x)$$

$$+\frac{1}{n_m n_{m-1}}\mathrm{E}\,[(p_{m-2}(x)-\widehat{p}_{m-2}(x))^2].$$

Hence, inductively we can obtain (2. 13) and the theorem is proved.

REMARK. In this paper, we may consider that $p_{m+1}(x)$ defined on R^1 is non-negative for sufficiently large n_m because of

$$\lim_{n\to\infty}\mathrm{E}[\int_{-\infty}^{\infty}(p_m(x)-\widehat{p}_m(x))^2dx\,|\,p_m]=0.$$

3. The statistical properties in the case of an original Markoff process with stationary transition probabilities

Let an original stochastic process $\{X_t,\ t=0,\ 1,\ \cdots\}$ be a simple Markoff process with transition probability density functions $p_0(x\,|\,x')$ defined on $R^1\times R^1$ and an initial probability density function $p_0(x)$ defined on R^1 satisfying the conditions:

(A'. 1) $p_0(x)\geqq 0$ for all $x\epsilon R^1$ and $\int_{-\infty}^{\infty}p_0(x)dx=1$,

(A'. 2) $\sup_{x\epsilon R^1}p_0(x)=M_0'<\infty$,

(A'. 3) $p_0(x)$ is uniformly continuous in x,

(A'. 4) $p_0(x\,|\,x')>0$ for all $x,\ x'\epsilon R^1$ and $\int_{-\infty}^{\infty}p_0(x\,|\,x')dx=1$ for all $x'\epsilon R^1$,

(A'. 5) $\sup_{x\epsilon R^1,\ x'\epsilon R^1}p_0(x\,|\,x')=M''_0<\infty$,

(A'. 6) $p_0(x\,|\,x')$ is uniformly continuous in x and x'.

Here, we shall consider to transform an original Markoff process into a process with the prescribed stationary transition probability density function $q(x\,|\,x')$ defined on $R^1\times R^1$ satisfying the conditions:

(B'. 1) $q(x\,|\,x')>0$ for all $x,\ x'\epsilon R^1$ and $\int_{-\infty}^{\infty}q(x\,|\,x')dx=1$, for all $x'\epsilon R^1$,

(B'. 2) $\sup_{x\epsilon R^1,\ x'\epsilon R^1}q(x\,|\,x')=M<\infty$,

(B'. 3) $q(x\,|\,x')$ is uniformly continuous in x and x'.

We make use of the following algorithm in order to solve this problem. At the first stage, sequences of random samples of size n_0 are drawn from the stochastic process $\{X_t,\ t=0,\ 1$

$\cdots, 2T-1\}$ and an estimate of $p_0(x|x')$ for each x and x',

$$(3.1)\qquad \widehat{p_0}(x|x')=\frac{\sum_{t=1}^{T}\frac{1}{n_0h_0^2(n_0)}\sum_{i=1}^{n_0}K_0\left(\frac{x-x_i(2t-1)}{h_0(n_0)}\right)K_0\left(\frac{x'-x_i(2t-2)}{h_0(n_0)}\right)}{\sum_{t=1}^{T}\frac{1}{n_0h_0(n_0)}\sum_{i=1}^{n_0}K_0\left(\frac{x'-x_i(2t-2)}{h_0(n_0)}\right)},$$

is calculated, where $x_i(t)$, $i=1, 2, \cdots, n_0$, are the n_0 sample values at time t, $\{h_0(n), n=1, 2, \cdots\}$ is a sequence of positive real numbers satisfying (C_0 1) and (C_0 2) in Section 2 and $K_0(x)$ is a real-valued function defined on R^1 satisfying the following conditions:

(K_0' 1) $\qquad K_0(x)>0 \qquad$ for all $x\epsilon R^1$

and (K_0 2)—(K_0 4) in Section 2.

Then, we transform the transition probability density function $p_0(x|x')$ into $p_0(x|x')-\widehat{p_0}(x|x')+q(x|x')\equiv p_1(x|x')$. At the second stage, sequences of random samples of size n_1 are drawn from the stochastic process $\{X_{1t}, t=2T, 2T+1, \cdots, 4T-1\}$ expressed by the transition probability density function $p_1(x|x')$ and an estimate of $p_1(x|x')$ for each x and x',

$$(3.2)\qquad \widehat{p_1}(x|x')=\frac{\sum_{t=1}^{T}\frac{1}{n_1h_1^2(n_1)}\sum_{i=1}^{n_1}K_1\left(\frac{x-x_i(2T+2t-1)}{h_1(n_1)}\right)K_1\left(\frac{x'-x_i(2T+2t-2)}{h_1(n_1)}\right)}{\sum_{t=1}^{T}\frac{1}{n_1h_1(n_1)}\sum_{i=1}^{n_1}K_1\left(\frac{x'-x_i(2T+2t-2)}{h_1(n_1)}\right)},$$

is calculated, where $x_i(t)$, $i=1, 2, \cdots, n_1$, are the n_1 sample values at time t, $\{h_1(n), n=1, 2, \cdots\}$ is a sequence of positive real nmmbers satisfying (C_1 1) and (C_1 2) in Section 2 and $K_1(x)$ is a real-valued function defined on R^1 satisfying the following conditions:

(K'_1 1) $\qquad K_1(x)>0 \qquad$ for all $x\epsilon R^1$

and (K_1 2)—(K_1 4) in Section 2.

Then, we transform the transition probability density function $p_1(x|x')$ into $p_1(x|x')-\widehat{p_1}(x|x')+q(x|x')\equiv p_2(x|x')$. Generally, at the m-th stage, sequences of random samples of size n_{m-1} are drawn from the stochastic process $\{X_{m-1,t}, t=2(m-1)T, \cdots, 2mT-1\}$ expressed by the transition probability density function $p_{m-1}(x|x')$ and an estimate of $p_{m-1}(x|x')$ for each x and x',

(3.3) $\widehat{p_{m-1}}(x|x')$

$$=\frac{\sum_{t=1}^{T}\frac{1}{n_{m-1}h_{m-1}^2(n_{m-1})}\sum_{i=1}^{n_{m-1}}K_{m-1}\left(\frac{x-x_i(2(m-1)T+2t-1)}{h_{m-1}(n_{m-1})}\right)K_{m-1}\left(\frac{x'-x_i(2(m-1)T+2t-2)}{h_{m-1}(n_{m-1})}\right)}{\sum_{t=1}^{T}\frac{1}{n_{m-1}h_{m-1}(n_{m-1})}\sum_{i=1}^{n_{m-1}}K_{m-1}\left(\frac{x'-x_i(2(m-1)T+2t-2)}{h_{m-1}(n_{m-1})}\right)}$$

is calculated, where $x_i(t)$, $i=1, 2, \cdots, n_{m-1}$, are the n_{m-1} sample values at time t, $\{h_{m-1}(n), n=1, 2, \cdots\}$ is a sequence of positive real numbers satisfying (C_{m-1}1) and (C_{m-1}2) in Section 2 and $K_{m-1}(x)$ is a real-valued function defined on R^1 satisfying the following conditions:

(K$'_{m-1}$1) $K_{m-1}(x)>0$ for all $x\in R^1$

and (K$_{m-1}$2)—(K$_{m-1}$4) in Section 2.

In order to prove the theorems in this section, we need the following lemmas.

LEMMA 2. *At the first stage,* $\widehat{p_0}(x|x')-p_0(x|x')$ *has the same limit distribution as*

(3. 4)

$$\frac{\sum_{t=1}^{T}\frac{1}{n_0}\sum_{i=1}^{n_0}\left[\frac{1}{h^2{}_0(n_0)}K_0\left(\frac{x-x_i(2t-1)}{h_0(n_0)}\right)K_0\left(\frac{x'-x_i(2t-2)}{h_0(n_0)}\right)-\frac{p_0(x|x')}{h_0(n_0)}K_0\left(\frac{x'-x_i(2t-2)}{h_0(n_0)}\right)\right]}{\sum_{t=1}^{T}p_{2t-2}(x')},$$

where $p_t(x)$ *is the probability density function at time* t.

PROOF. By the law of large numbers and Lemma 1, we can obtain

$$P\lim_{n_0\to\infty}\sum_{t=1}^{T}\frac{1}{n_0}\sum_{i=1}^{n_0}\frac{1}{h_0(n_0)}K_0\left(\frac{x'-x_i(2t-2)}{h_0(n_0)}\right)=$$

$$=\lim_{n_0\to\infty}\sum_{t=1}^{T}\frac{1}{n_0}\sum_{i=1}^{n_0}\mathrm{E}\left[\frac{1}{h_0(n_0)}K_0\left(\frac{x'-x_i(2t-2)}{h_0(n_0)}\right)\right]$$

$$=\sum_{t=1}^{T}p_{2t-2}(x').$$

Thus, using the Cramér's result, the lemma is completed (see P. 254 in [2]).

LEMMA 3. *At the first stage, if* n_0 *is sufficiently large then we have*

(3. 5) $\mathrm{E}[\widehat{p_0}(x|x')]\approx p_0(x|x')$ *for all* x *and* $x'\in R^1$,

(3. 6) $$\mathrm{E}[(p_0(x|x')-\widehat{p_0}(x|x'))^2]\approx\frac{\frac{1}{n_0}\left(\frac{a_0}{h_0(n_0)}\right)\left\{\frac{a_0}{h_0(n_0)}-p_0(x|x')\right\}p_0(x|x')}{A_0(x')}$$

for all x *and* $x'\in R^1$,

where $A_0(x')=\sum_{t=1}^{T}p_{2t-2}(x')$ *and* $a_0=\int_{-\infty}^{\infty}K_0^2(y)dy$.

PROOF. First we shall prove (3. 5). By the definition of $\widehat{p_0}(x|x')$ and Lemma 3, we have

(3. 7) $\widehat{p_0}(x|x')-p(x|x')$

$$\approx\frac{\sum_{t=1}^{T}\frac{1}{n_0}\sum_{i=1}^{n_0}\left[\frac{1}{h_0^2(n_0)}K_0\left(\frac{x-x_i(2t-1)}{h_0(n_0)}\right)K_0\left(\frac{x'-x_i(2t-2)}{h_0(n_0)}\right)-\frac{p_0(x|x')}{h_0(n_0)}K_0\left(\frac{x'-x_i(2t-2)}{h_0(n_0)}\right)\right]}{A_0(x')}.$$

In order to show $E[\widehat{p}_0(x|x')]\approx p_0(x|x')$, it is sufficient to prove

$$E\Big[\frac{1}{n_0}\sum_{i=1}^{n_0}\Big\{\frac{1}{h^2_0(n_0)}K_0\Big(\frac{x-x_i(2t-1)}{h_0(n_0)}\Big)K_0\Big(\frac{x'-x_i(2t-2)}{h_0(n_0)}\Big)-\frac{p_0(x|x')}{h_0(n_0)}K_0\Big(\frac{x'-x_i(2t-2)}{h_0(n_0)}\Big)\Big\}\Big]\approx 0.$$

This is done, using Lemma 1, as follows:

(3. 8)
$$E\Big[\frac{1}{n_0}\sum_{i=1}^{n_0}\Big\{\frac{1}{h^2_0(n_0)}K_0\Big(\frac{x-x_i(2t-1)}{h_0(n_0)}\Big)K_0\Big(\frac{x'-x_i(2t-2)}{h_0(n_0)}\Big)$$
$$-\frac{p_0(x|x')}{h_0(n_0)}K_0\Big(\frac{x'-x_i(2t-2)}{h_0(n_0)}\Big)\Big\}\Big]$$
$$=\int_{-\infty}^{\infty}\frac{1}{h_0(n_0)}K_0\Big(\frac{x'-y_2}{h_0(n_0)}\Big)\Big[\int_{-\infty}^{\infty}\Big\{\frac{1}{h_0(n_0)}K_0\Big(\frac{x-y_1}{h_0(n_0)}\Big)$$
$$-p_0(x|x')\Big\}p_0(y_1|y_2)dy_1\Big]p_{2t-2}(y_2)dy_2$$
$$\approx\int_{-\infty}^{\infty}\frac{1}{h_0(n_0)}K_0\Big(\frac{x'-y_2}{h_0(n_0)}\Big)(p_0(x|y_2)-p_0(x|x'))p_{2t-2}(y_2)dy_2$$
$$=\int_{-\infty}^{\infty}\frac{1}{h_0(n_0)}K_0\Big(\frac{x'-y_2}{h_0(n_0)}\Big)p_0(x|y_2)p_{2t-2}(y_2)dy_2$$
$$-p_0(x|x')\int_{-\infty}^{\infty}\frac{1}{h_0(n_0)}K_0\Big(\frac{x'-y_2}{h_0(n_0)}\Big)p_{2t-2}(y_2)dy_2$$
$$\approx p_0(x|x')p_{2t}(x')-p_0(x|x')p_{2t-2}(x')=0.$$

Secondly, we shall prove (3. 6). We calculate for the first time the following equations:

(3. 9)
$$E\Big[\frac{1}{n_0^2}\sum_{i=1}^{n_0}\Big\{\frac{1}{h^2_0(n_0)}K_0\Big(\frac{x-x_i(2t-1)}{h_0(n_0)}\Big)K_0\Big(\frac{x'-x_i(2t-2)}{h_0(n_0)}\Big)$$
$$-\frac{p_0(x|x')}{h_0(n_0)}K_0\Big(\frac{x'-x_i(2t-2)}{h_0(n_0)}\Big)\Big\}^2\Big]$$
$$=\frac{1}{n_0}E\Big[\Big\{\frac{1}{h_0(n_0)}K_0\Big(\frac{x-x_i(2t-1)}{h_0(n_0)}\Big)\Big\}^2\Big\{\frac{1}{h_0(n_0)}K_0\Big(\frac{x'-x_i(2t-2)}{h_0(n_0)}\Big)\Big\}^2$$
$$-2p_0(x|x')\Big\{\frac{1}{h_0(n_0)}K_0\Big(\frac{x-x_i(2t-1)}{h_0(n_0)}\Big)\Big\}\Big\{\frac{1}{h_0(n_0)}K_0\Big(\frac{x'-x_i(2t-2)}{h_0(n_0)}\Big)\Big\}^2$$
$$+p_0^2(x|x')\Big\{\frac{1}{h_0(n_0)}K_0\Big(\frac{x'-x_i(2t-2)}{h_0(n_0)}\Big)\Big\}^2\Big]$$
$$=\frac{1}{n_0}\Big[\int_{-\infty}^{\infty}\Big\{\frac{1}{h_0(n_0)}K_0\Big(\frac{x'-y_2}{h_0(n_0)}\Big)\Big\}^2\Big\{\int_{-\infty}^{\infty}\Big(\frac{1}{h_0(n_0)}K_0\Big(\frac{x-y_1}{h_0(n_0)}\Big)\Big)^2p_0(y_1|y_2)dy_1\Big\}p_{2t-2}(y_2)dy_2$$
$$-2p_0(x|x')\int_{-\infty}^{\infty}\Big\{\frac{1}{h_0(n_0)}K_0\Big(\frac{x'-y_2}{h_0(n_0)}\Big)\Big\}^2\Big\{\int_{-\infty}^{\infty}\frac{1}{h_0(n_0)}K_0\Big(\frac{x-y_1}{h_0(n_0)}\Big)p_0(y_1|y_2)dy_1\Big\}p_{2t-2}(y_2)dy_2$$

$$+p^2_0(x|x')\int_{-\infty}^{\infty}\left\{\frac{1}{h_0(n_0)}K_0\left(\frac{x'-y_2}{h_0(n_0)}\right)\right\}^2 p_{2t-2}(y_2)dy_2\Big]$$

$$\approx\frac{1}{n_0}\Big[\frac{1}{h_0(n_0)}\int_{-\infty}^{\infty}K_0^2(y)dy\int_{-\infty}^{\infty}\left\{\frac{1}{h_0(n_0)}K_0\left(\frac{x'-y_2}{h_0(n_0)}\right)\right\}^2 p_0(x|y_2)p_{2t-2}(y_2)dy_2$$

$$-2p_0(x|x')\int_{-\infty}^{\infty}\left\{\frac{1}{h_0(n_0)}K_0\left(\frac{x'-y_2}{h_0(n_0)}\right)\right\}^2 p_0(x|y_2)p_{2t-2}(y_2)dy_2$$

$$+p_0^2(x|x')\left(\frac{1}{h_0(n_0)}\int_{-\infty}^{\infty}K_0^2(y)dy\right)p_{2t-2}(x')$$

$$\approx\frac{1}{n_0}\Big[\left(\frac{a_0}{h_0(n_0)}\right)^2 p_0(x|x')p_{2t-2}(x')-2\left(\frac{a_0}{h_0(n_0)}\right)p_0^2(x|x')p_{2t-2}(x')$$

$$+\left(\frac{a_0}{h_0(n_0)}\right)p_0^2(x|x')p_{2t-2}(x')\Big]$$

$$=\frac{a_0}{n_0h_0(n_0)}\left\{\frac{a_0}{h_0(n_0)}-p_0(x|x')\right\}p_0(x|x')p_{2t-2}(x')$$

and

$$(3.10)\quad \mathrm{E}\Big[\left\{\frac{1}{h_0^2(n_0)}K_0\left(\frac{x-x_i(2t-1)}{h_0(n_0)}\right)K_0\left(\frac{x'-x_i(2t-2)}{h_0(n_0)}\right)-\frac{p_0(x|x')}{h_0(n_0)}K_0\left(\frac{x'-x_i(2t-2)}{h_0(n_0)}\right)\right\}\cdot$$

$$\cdot\left\{\frac{1}{h_0^2(n_0)}K_0\left(\frac{x-x_i(2t'-1)}{h_0(n_0)}\right)K_0\left(\frac{x'-x_i(2t'-2)}{h_0(n_0)}\right)-\frac{p_0(x|x')}{h_0(n_0)}K_0\left(\frac{x'-x_i(2t-2)}{h_0(n_0)}\right)\right\}\Big]$$

$$\approx 0 \qquad \text{for } t\neq t'.$$

Hence, from (3. 9) and (3. 10), we can obtain

$$(3.11)\quad \mathrm{E}[(p_0(x|x')-\widehat{p_0}(x|x'))^2]$$

$$\approx\frac{\mathrm{E}\Big[\left\{\sum_{t=1}^{T}\frac{1}{n_0}\sum_{i=1}^{n_0}\left(\frac{1}{h^2_0(n_0)}K_0\left(\frac{x-x_i(2t-1)}{h_0(n_0)}\right)K_0\left(\frac{x-x_i(2t-2)}{h_0(n_0)}\right)-\frac{p_0(x|x')}{h_0(n_0)}K_0\left(\frac{x'-x_i(2t-2)}{h_0(n_0)}\right)\right)\right\}^2\Big]}{A_0^2(x')}$$

$$\approx\frac{\frac{1}{n_0}\sum_{t=1}^{T}\mathrm{E}\Big[\left\{\frac{1}{h_0(n_0)}K_0\left(\frac{x-x_i(2t-1)}{h_0(n_0)}\right)K_0\left(\frac{x'-x_i(2t-2)}{h_0(n_0)}\right)-\frac{p_0(x|x')}{h_0(n_0)}K_0\left(\frac{x'-x_i(2t-2)}{h_0(n_0)}\right)\right\}^2\Big]}{A_0^2(x')}$$

$$\approx\frac{\frac{a_0}{n_0h_0(n_0)}\left\{\frac{a_0}{h_0(n_0)}-p_0(x|x')\right\}p_0(x|x')\sum_{t=1}^{T}p_{2t-2}(x')}{A_0^2(x')}$$

$$=\frac{\frac{a_0}{n_0h_0(n_0)}\left\{\frac{a_0}{h_0(n_0)}-p_0(x|x')\right\}p_0(x|x')}{A_0(x')}$$

Thus, the proof of the lemma was completed.

THEOREM 3. 1

(3. 12) $\quad \mathrm{E}[p_{m+1}(x|x')] \approx q(x|x')$ *for sufficiently large* n_m.

This theorem is easily proved by Lemma 3.

THEOREM 3. 2 *For sufficiently large* n_l, $l=1, 2, \cdots, m$

(3. 13) $\quad \mathrm{E}[(p_{m+1}(x|x')-q(x|x'))^2]$

$$\approx \sum_{i=0}^{m-1}(-1)^i \prod_{j=0}^{i} \frac{1}{n_{m-j}A_{m-j}(x')}\left(\frac{a_{m-j}}{h_{m-j}(n_{m-j})}\right)\left(\frac{a_{m-i}}{h_{m-i}(n_{m-i})}-q(x|x')\right)q(x|x')$$

$$+\left\{\prod_{j=0}^{m-1}\left(\frac{-1}{n_{m-j}A_{m-j}(x')}\right)\left(\frac{a_{m-j}}{h_{m-j}(n_{m-j})}\right)\right\}\left(\frac{1}{n_0A_0(x')}\right)\left(\frac{a_0}{h_0(n_0)}-p_0(x|x')\right)p_0(x|x'),$$

where $A_l(x')=\sum_{t=1}^{T} p_{2lT+2t-2}(x')$.

If, for $i=0, 1, 2, \cdots, m$, $n_i=n$,

$h_i(l)=h(l)$, $l=1, 2, \cdots$,

and $K_i(x)=K(x)$ *for all* $x \in R^1$, then

(3. 14) $\quad \mathrm{E}[(p_{m+1}(x|x')-q(x|x'))^2]$

$$\approx\left\{\sum_{i=0}^{m-1}(-1)^i\left(\prod_{j=0}^{i}\frac{1}{A_{m-j}(x')}\right)\left(\frac{1}{n}\right)^{i+1}\right\}\left(\frac{a}{h(n)}\right)\left(\frac{a}{h(n)}-q(x|x')\right)q(x|x')$$

$$+\left\{\prod_{j=0}^{m}\left(\frac{-1}{A_{m-j}(x')}\right)\right\}\left(\frac{1}{n}\right)^{m+1}\left(\frac{a}{h(n)}\right)\left(\frac{a}{h(n)}-p_0(x|x')\right)p_0(x|x'),$$

where $\quad a=\int_{-\infty}^{\infty} K^2(y)dy.$

PROOF. By the definition of $p_{m+1}(x|x')$, $\mathrm{E}[(p_{m+1}(x|x')-q(x|x'))^2]$ is written as

(3. 15) $\quad \mathrm{E}[(p_m(x|x')-\widehat{p}_m(x|x'))^2]$.

If we take the conditional expectation of $(p_m(x|x')-\widehat{p}_m(x|x'))^2$, then we can use the result of Lemma 3. So that

(3. 16) $\quad \mathrm{E}[\mathrm{E}\{(p_m(x|x')-\widehat{p}_m(x|x'))^2|p_m\}]$

$$\approx \mathrm{E}\left[\frac{1}{n_mA_m(x')}\left(\frac{a_m}{h_m(n_m)}\right)\left(\frac{a_m}{h_m(n_m)}-p_m(x|x')\right)p_m(x|x')\right].$$

Repeating this calculation of conditional expectation, (3. 16) is reduced to

(3. 17) $\quad \mathrm{E}[(p_m(x|x')-\widehat{p}_m(x|x'))^2]$

$$\approx \mathrm{E}\left[\frac{1}{n_mA_m(x')}\left(\frac{a_m}{h_m(n_m)}\right)\left\{\frac{a_m}{h_m(n_m)}-q(x|x')-(p_{m-1}(x|x')\right.\right.$$

$$\left.\left.-\widehat{p}_{m-1}(x|x'))\right\}\left\{q(x|x')+(p_{m-1}(x|x')-\widehat{p}_{m-1}(x|x'))\right\}\right]$$

$$\approx \frac{1}{n_m A_m(x')}\left(\frac{a_m}{h_m(n_m)}\right)\left(\frac{a_m}{h_m(n_m)}-q(x|x')\right)q(x|x')$$

$$-\frac{1}{n_m A_m(x')}\left(\frac{a_m}{h_m(n_m)}\right)\mathrm{E}\left[(p_{m-1}(x|x')-\hat{p}_{m-1}(x|x'))^2\right]$$

$$\approx \frac{1}{n_m A_m(x')}\left(\frac{a_m}{h_m(n_m)}\right)\left(\frac{a_m}{h_m(n_m)}-q(x|x')\right)q(x|x')$$

$$-\frac{1}{n_m A_m(x')}\left(\frac{a_m}{h_m(n_m)}\right)\left\{\frac{1}{n_{m-1}A_{m-1}(x')}\left(\frac{a_{m-1}}{h_{m-1}(n_{m-1})}\right)\left(\frac{a_{m-1}}{h_{m-1}(n_{m-1})}\right.\right.$$

$$\left.-q(x|x')\right)q(x|x')-\frac{1}{n_{m-1}A_{m-1}(x')}\left(\frac{a_{m-1}}{h_{m-1}(n_{m-1})}\right)\mathrm{E}[p_{m-2}(x|x')$$

$$\left.-\hat{p}_{m-2}(x|x'))^2]\right\}.$$

Hence, inductively we can obtain (3. 13) and the theorem is proved.

NIIGATA UNIVERSITY

References

[1] ANDERSON, T. W. and GOODMAN Leo, A.: *Statistical inference about Markov chains.* Ann. Math. Stat., 28 (1957), 89–110.

[2] CRAMÉR, H.: "Mathematical methods of statistics," Princeton University Press, Princeton, 1946.

[3] KANŌ, S.: *Linear controlled stochastic processes.* Science Reports of Kagoshima Univ., No. 8 (1959), 1–14.

[4] KANŌ, S.: *Multi-dimensional linear controlled stochastic processes with discrete parameter.* Science Reports of Kagoshima Univ., No. 10 (1961), 11–20.

[5] KANŌ, S.: *On the statistical control of probability distribution.* Science Reports of Kagoshima Univ., No. 11 (1962), 1–6.

[6] KANŌ, S.: *On the statistical controls of poisson processes.* Science Reports of Kagoshima Univ., No. 15 (1966), 1–5.

[7] KITAGAWA, T.: *Successive processes of statistical controls (1).* Mem. Fac., Sci., Kyushu Univ., Ser. A, 7 (1952), 13–28.

[8] KITAGAWA, T.: *Successive processes of statistical controls (2).* Mem. Fac. Sci., Kyushu Univ., Ser. A. 13 (1959), 1–16.

[9] KITAGAWA, T.: *Successive processes of statistical controls (3).* Mem. Fac. Sci., Kyushu Univ., Ser. A, 14 (1960), 1–33.

[10] PARZEN, E.: *On estimation of a probability density and mode.* Ann. Math. Stat., 33 (1962), 1065–1076.

[11] ROUSSAS, G. G.: *Nonparametric estimation of the transition distribution function of a Markov process.* Ann. Math. Stat., 40 (1969), 1386–1400.