THE OPTIMAL SAMPLING PROCEDURE FOR ESTIMATING THE MEAN OF STATIONARY MARKOV PROCESSES

YASUSHI TAGA

(Received Dec. 17, 1964; revised Feb. 1, 1965)

Summary

In estimating the mean of a stationary Markov process, the average of observations is used, which are spaced by time intervals distributed independently and identically according to a d.f. $F(x)$. The asymptotic variance of the estimate thus obtained is given, and the optimal sampling procedure, i.e., the extremal d.f. $F_0(x)$ is determined so as to minimize that variance.

1. Introduction

Let $\{x(t, \omega); -\infty<t<\infty\}$ be a strictly stationary stochastic process (real or complex) such that

$$(1)\qquad \begin{aligned} &E\{x(t)\}=m \\ &E\,|x(t)-m|^2\equiv v<\infty \\ &E\{(x(t+\tau)-m)(\bar{x}(t)-\bar{m})\}=v\cdot r(\tau)\,, \end{aligned}$$

where $\bar{x}(\cdot)$ and $\bar{m}$ are the complex conjugates of $x(\cdot)$ and m, respectively. Here we assume that $r(\tau)$ is continuous at $\tau=0$, i.e., $\{x(t, \omega)\}$ is continuous in the mean square.

Let $\{\Delta\tau_n(\omega);\ n=0, \pm 1, \pm 2, \cdots\}$ be a strictly stationary stochastic process defined on the same ω space as $x(t, \omega)$ process. Define $\tau_n(\omega)$ such that $\tau_0(\omega)=\varepsilon(\omega)$ and $\tau_n(\omega)-\tau_{n-1}(\omega)=\Delta\tau_{n-1}(\omega)$, where $\varepsilon(\omega)$ is a random variable. Assume further that almost all sample functions of $\{x(t, \omega)\}$ have finite limits from the right, i.e.,

$$\lim_{h\to 0+} x(t+h, \omega)=x(t+, \omega) \qquad \text{for all } t\,.$$

Then we can define a sequence of ω functions $\{x_{\tau, n}(\omega);\ n=0, \pm 1, \pm 2, \cdots\}$ as follows:

$$x_{\tau, n}(\omega)=\begin{cases} x(\tau_n(\omega), \omega) & \text{if } P\{\omega';\ \tau_n(\omega')=\tau_n(\omega)\}>0 \\ x(\tau_n(\omega)+, \omega) & \text{otherwise.} \end{cases}$$

Then it can be proved in the same way as is shown in [2] that $x_{\tau,n}(\omega)$'s are measurable ω functions and that $\{x_{\tau,n}(\omega);\ n=0, \pm 1, \pm 2, \cdots\}$ forms a weakly stationary stochastic process.

Hereafter, we shall restrict our attention to the case where $\{\Delta\tau_n(\omega)\}$ is a sequence of mutually independent and identically distributed (according to a d.f. $F(\tau)$) non-negative random variables and is independent of the process $\{x(t, \omega)\}$. Then it can be shown that

$$E\{x_{\tau,n}(\omega)\}=m$$

and

$$\rho(k)\equiv\frac{1}{v}E\{(x_{\tau,n+k}(\omega)-m)(\bar{x}_{\tau,n}(\omega)-\bar{m})\}$$

$$=\int_0^\infty r(\tau)dF^{(k)}(\tau) \qquad \text{for } k=1, 2, \cdots,$$

where $F^{(k)}(\tau)$ denotes the kth convolution of $F(\tau)$ with itself ($F^{(0)}(\tau)\equiv 1$).

In estimating the mean m of the original process $\{x(t, \omega)\}$, we shall use the estimate

$$\hat{m}=\frac{1}{n}\sum_{j=1}^{n} x_{\tau,j}(\omega). \tag{2}$$

Then, from the above considerations, it can be seen that

$$E(\hat{m})=m$$

and

$$V(\hat{m})=\frac{v}{n}\left\{1+\sum_{k=1}^{n-1}\left(1-\frac{k}{n}\right)\int_0^\infty (r(\tau)+\bar{r}(\tau))dF^{(k)}(\tau)\right\}, \tag{3}$$

where $\bar{r}(\tau)$ is the complex conjugate of $r(\tau)$. It is well-known in the renewal theory that the series $\sum_{k=1}^{\infty} F^{(k)}(\tau)$ and $\sum_{k=1}^{\infty} kF^{(k)}(\tau)$ are convergent for all $\tau\geqq 0$ if $F(0)<1$ (see [3]). We define $H_1(\tau)=\sum_{k=1}^{\infty} F^{(k)}(\tau)$ and $H_2(\tau)=\sum_{k=1}^{\infty} kF_k(\tau)$. Then $V(\hat{m})$ can be asymptotically represented by

$$V(\hat{m})\sim\frac{v}{n}\left\{1+\int_0^\infty (r(\tau)+\bar{r}(\tau))dH_1(\tau)-\frac{1}{n}\int_0^\infty (r(\tau)+\bar{r}(\tau))dH_2(\tau)\right\}, \tag{4}$$

if the integrals $\int_0^\infty |r(\tau)|dH_1(\tau)$ and $\int_0^\infty |r(\tau)|dH_2(\tau)$ are finite. Further, in these cases, the third term in the bracket of the right-hand side of

(4) is asymptotically negligible, and we obtain the final result

$$V(\hat{m}) \sim \frac{v}{n}\left\{1+\int_0^\infty (r(\tau)+\bar{r}(\tau))dH_1(\tau)\right\} . \tag{5}$$

For example, when $F(\tau)=1-e^{-\lambda\tau}$, we obtain

$$H_1(\tau)=\sum_{k=1}^\infty F_k(\tau)=\lambda\tau\ , \ H_2(\tau)=\sum_{k=1}^\infty kF_k(\tau)=\frac{1}{2}\lambda^2\tau^2+\lambda\tau\ ,$$

and

$$\begin{aligned} V(\hat{m}) &\sim \frac{v}{n}\left\{1+\lambda\left(1-\frac{1}{n}\right)\int_0^\infty [r(\tau)+\bar{r}(\tau)]d\tau-\frac{2\lambda^2}{n}\int_0^\infty \tau[r(\tau)+\bar{r}(\tau)]d\tau\right\} \\ &\sim \frac{v}{n}\left\{1+\lambda\int_0^\infty [r(\tau)+\bar{r}(\tau)]d\tau\right\}. \end{aligned} \tag{6}$$

This result is identical to the one obtained by Gebhard [1].

2. The optimal procedures for Markov processes

In the cases of stationary Markov processes, it is well-known that $r(\tau)$ is given by

$$r(\tau)=e^{-(\theta-i\alpha)\tau}, \qquad \text{for } \tau\geqq 0,\ \theta\geqq 0\,. \tag{7}$$

When $x(t)$ is real, $r(\tau)$ is also real and is given by $r(\tau)=e^{-\theta\tau}$, i.e., $\alpha=0$ (see [4]). If $\theta=0$, then it occurs that $x(t)=e^{2\pi i\alpha t}x(0)$ with probability 1. In the following, let us assume that $\theta>0$. It is easily seen from (3) and (7) that

$$V(\hat{m})=\begin{cases} \dfrac{v}{n}\left[1+2\mathscr{R}\left\{\dfrac{f(\theta,\alpha)}{1-f(\theta,\alpha)}\right\}-\dfrac{2}{n}\mathscr{R}\left\{\dfrac{1-f^n(\theta,\alpha)}{(1-f(\theta,\alpha))^2}\right.\right. \\ \qquad\left.\left.+\dfrac{f^n(\theta,\alpha)}{1-f(\theta,\alpha)}\right\}\right], & \text{if } |f(\theta,\alpha)|<1\,, \\ v\,, & \text{if } |f(\theta,\alpha)|=1\,, \end{cases} \tag{8}$$

where
$$f(\theta,\alpha)=\int_0^\infty e^{-(\theta-i\alpha)\tau}dF(\tau)$$

is the Laplace-Stieltjes transform of $F(\tau)$. It is noted that $|f(\theta,\alpha)|=1$ occurs if and only if $F(\tau)$ degenerates at $\tau=0$. Therefore, $V(\hat{m})$ can be represented asymptotically as

$$V(\hat{m})\sim\frac{v}{n}\left[1+2\mathscr{R}\left\{\frac{f(\theta,\alpha)}{1-f(\theta,\alpha)}\right\}\right], \qquad \text{if } |f(\theta,\alpha)|<1. \tag{8'}$$

Putting $f(\theta, \alpha)=u_F(\theta, \alpha)+iw_F(\theta, \alpha)$, we can rewrite (8) as follows:

$$V(\hat{m}) \sim \frac{v}{n}\left\{1+2\frac{u_F(1-u_F)-w_F^2}{(1-u_F)^2+w_F^2}\right\}, \tag{9}$$

where

$$\begin{aligned} u_F(\theta, \alpha)&=\int_0^\infty u(\tau)dF(\tau), \qquad u(\tau)=e^{-\theta\tau}\cos\alpha\tau, \\ w_F(\theta, \alpha)&=\int_0^\infty w(\tau)dF(\tau), \qquad w(\tau)=e^{-\theta\tau}\sin\alpha\tau. \end{aligned} \tag{10}$$

In order to minimize the asymptotic variance, it is sufficient to determine the d.f. $F(\tau)$ which minimizes the second member in the bracket of the right hand side of (9). To do so, let us consider the function

$$\varphi(u, w)=\frac{u(1-u)-w^2}{(1-u)^2+w^2} \tag{11}$$

in the unit circle, i.e., $u^2+w^2=1$.

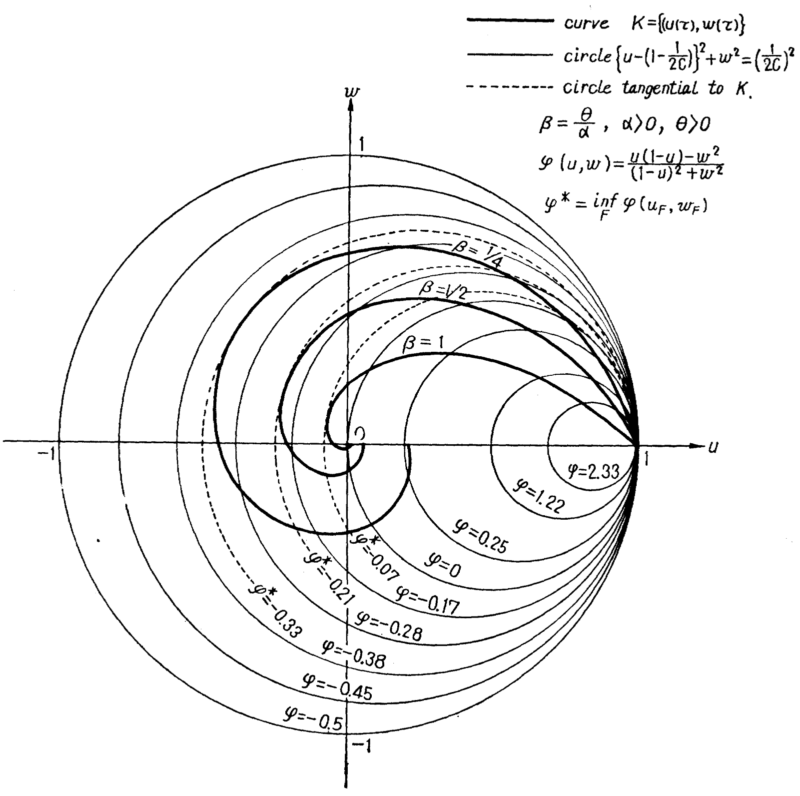

Fig. 1

Putting $\varphi(u, w)\equiv c-1$, we obtain the contour line

(12) $$\left\{u-\left(1-\frac{1}{2c}\right)\right\}^2+w^2=\left(\frac{1}{2c}\right)^2,$$

where $c=\varphi(u, w)+1=\dfrac{1-u}{(1-u)^2+w^2}\geqq 0$ (see Fig. 1).

Now, we try to obtain the infimum of the function $\varphi(u_F, w_F)$ under certain conditions of $F(\tau)$.

THEOREM. *Let $\mathscr{F}$ be the family of all distribution functions on the half open interval $[0, \infty)$, and let u_F, w_F and $\varphi(u, w)$ be given as in (10) and (11). Then, $\inf_{F\in\mathscr{F}} \varphi(u_F, w_F)$ exists, and in the case where there exists the extremal distribution function $F_0(\tau)$ which makes $\varphi(u_F, w_F)$ minimum, $F_0(\tau)$ must degenerate at one point.*

PROOF. Since $u^2(\tau)+w^2(\tau)=e^{-2\theta\tau}\leqq 1$, the curve $K=\{(u(\tau), w(\tau));\ 0\leqq \tau<\infty\}$ lies in the unit circle $u^2+w^2=1$ except one point $(1, 0)$ corresponding to $\tau=0$. Then the point (u_F, w_F) must lie also in the unit circle only if $F(\tau)$ does not degenerate at $\tau=0$, because (u_F, w_F) is the weighted mean of the points on K. Further, it is easily seen that the function $\varphi(u, w)$ is bounded from below in the unit circle and $\inf_{u^2+w^2\leqq 1} \varphi(u, w)=-\dfrac{1}{2}$ is attained on its boundary. From the above considerations, we can conclude that the infimum of $\varphi(u_F, w_F)$ exists and

(13) $$\inf_{F\in\mathscr{F}} \varphi(u_F, w_F)\geqq -\frac{1}{2}\,.$$

Now, we try to obtain the value of $\inf_{F\in\mathscr{F}} \varphi(u_F, w_F)$ and determine the extremal d.f. $F_0(\tau)$.

Case 1. If $x(t)$ is real, then $r(\tau)=u(\tau)=e^{-\theta\tau}$ and $w(\tau)=0$. Since the curve $K=\{(u(\tau), w(\tau));\ 0\leqq\tau<\infty\}$ becomes the half open interval $[0, 1)$, it is clear from Fig. 1 that

(14) $$\inf_{F\in\mathscr{F}} \varphi(u_F, w_F)=0\,,$$

i.e., $\varphi(u_F, w_F)$ can be made indefinitely small as time intervals are made longer in equal interval sampling. However, the extremal distribution to attain the infimum does not exist. If the family of distribution functions is restricted to $\mathscr{F}(\mu)$ which is the family of those with the common mean μ, then

(15) $$\varphi(u_F, w_F)=\frac{u_F}{1-u_F}=\frac{\int_0^\infty e^{-\theta\tau}dF(\tau)}{1-\int_0^\infty e^{-\theta\tau}dF(\tau)}\geqq\frac{e^{-\theta\mu}}{1-e^{-\theta\mu}}\,,$$

since $e^{-\theta\tau}$ is convex. Therefore, it holds that

$$\inf_{F\in\mathscr{F}(\mu)} \varphi(u_F, w_F) \geqq \frac{e^{-\theta\mu}}{1-e^{-\theta\mu}}, \tag{16}$$

and the extremal distribution $F_0(x)$ exists and is uniquely determined as the one which degenerates at $\tau=\mu$.

Case 2. If $x(t)$ is complex, then $r(\tau)$ is given by (7). The curve K starts from the point (1, 0) along the tangential line $w=-\dfrac{\theta}{\alpha}(u-1)$. Rotating around the origin, it descends the slope of contour lines of $\varphi(u, w)$ and reaches at the lowest point, say, $(u(\tau^*), w(\tau^*))$. Thereafter the curve K ascends the slope (see Fig. 1). Therefore, we can see that

$$\inf_{F\in\mathscr{F}} \varphi(u_F, w_F)=\varphi(u(\tau^*), w(\tau^*))\equiv\varphi^* \tag{17}$$

and the extremal d.f. $F_0(\tau)$ must degenerate at $\tau=\tau^*$.

The analytic determination of the value of τ^* can be done in the following way:

Let $C(\delta)$ be a circle centered at the point $(1-\delta, 0)$ with radius δ and through a point $(u(\tau), w(\tau))$ on K. The radius δ satisfies the equation

$$\{u(\tau)-(1-\delta)\}^2+w^2(\tau)=\delta^2. \tag{18}$$

Solving (18), we obtain δ as a function of τ such that

$$\delta(\tau)=\frac{u^2(\tau)+w^2(\tau)-2u(\tau)+1}{2(1-u(\tau))}=\frac{e^{-\beta\tau}-2\cos\tau+e^{\beta\tau}}{2(e^{\beta\tau}-\cos\tau)}, \tag{19}$$

where we have replaced $\alpha\tau$ by τ and β by $\dfrac{\theta}{\alpha}$ in (10). By differentiation we get

$$\begin{aligned}\frac{d\delta}{d\tau}&=\frac{e^{\beta\tau}(\beta\cos\tau+\sin\tau)+e^{-\beta\tau}(\beta\cos\tau-\sin\tau)-2\beta}{2(e^{\beta\tau}-\cos\tau)^2}\\&=\frac{\sqrt{1+\beta^2}\,\{e^{\beta\tau}\cos(\tau-\tau_0)+e^{-\beta\tau}\cos(\tau+\tau_0)\}-2\beta}{2(e^{\beta\tau}-\cos\tau)^2},\end{aligned} \tag{20}$$

where

$$\cos\tau_0=\frac{\beta}{\sqrt{1+\beta^2}} \qquad \left(0<\tau_0<\frac{\pi}{2}\right),$$

assuming without loss of generality that $\alpha>0$, i.e., $\beta>0$. From (19) and (20), we can see that

$$\lim_{\tau\to 0+}\delta(\tau)=\lim_{\tau\to 0+}\frac{-\beta e^{-\beta\tau}+2\sin\tau+\beta e^{\beta\tau}}{2(\beta e^{\beta\tau}+\sin\tau)}=0\ ,$$

$$\delta(\tau)=\frac{\{1-u(\tau)\}^2+w^2(\tau)}{2(1-u(\tau))}>0 \qquad \text{for all } \tau>0\ ,$$

and

$$\lim_{\tau\to 0+}\frac{d\delta}{d\tau}=\frac{\beta^2+1}{2\beta}(>0)\ .$$

It can be seen that the numerator of the right-hand side of (20) is monotone increasing (or decreasing) for $0<\tau<\frac{\pi}{2}\left(\text{or } \frac{\pi}{2}<\tau<\frac{3}{2}\pi\right)$, since its derivative is $(1+\beta^2)(e^{\beta\tau}-e^{-\beta\tau})\cos\tau$. Moreover, it becomes zero or negative for $\tau=0$ or $\tau=\frac{\pi}{2}+\tau_0$. Then we can conclude that there exists a value τ_1 in $\left(\frac{\pi}{2}\ ,\ \frac{\pi}{2}+\tau_0\right)$ such that

$$\text{(21)}\qquad \frac{d\delta}{d\tau}\begin{cases} >0 & \text{for } 0<\tau<\tau_1 \\ =0 & \text{for } \tau=\tau_1 \\ <0 & \text{for } \tau_1<\tau<\frac{3}{2}\pi\ . \end{cases}$$

Therefore, $\delta(\tau)$ attains its maximal value $\delta(\tau_1)$ at $\tau=\tau_1$. Next, we shall prove that $\delta(\tau_1)$ is the maximum of $\delta(\tau)$ in $[0,\infty)$. For that, it is sufficient to prove that $\delta(\tau)<\delta(\pi)(\leqq\delta(\tau_1))$ for $\pi<\tau<\infty$, and for this, it is sufficient to show that all points $(u(\tau),\ w(\tau))$ for $\pi<\tau<\infty$ lie in the circle $C(\delta(\pi))$ corresponding to $\tau=\pi$. In fact, since $u^2(\tau)+w^2(\tau)=e^{-\beta\tau}<e^{-\beta\pi}$ for $\pi<\tau<\infty$, all points $(u(\tau),\ w(\tau))$ are included inside the circle $u^2+w^2=e^{-\beta\pi}$. Since this circle is included in $C(\delta(\pi))$, all points $(u(\tau),\ w(\tau))$ for $\pi<\tau<\infty$ are included in $C(\delta(\pi))$. Hence $\delta(\tau)$ attains its maximum at $\tau=\tau_1$ $\left(\tau=\tau^*\equiv\frac{\tau_1}{\alpha}\text{ in the original scale}\right)$. Comparing (12) with (18), we can see that $\varphi(u(\tau),\ w(\tau))=\frac{1}{2\delta(\tau)}-1$ and that $\varphi(u(\tau),\ w(\tau))$ attains its minimum value $\frac{1}{2\delta(\tau^*)}-1$ at $\tau=\tau^*$. Therefore, the extremal d.f. $F_0(\tau)$, for time sampling, exists and degenerates at $\tau=\tau^*$.

3. Conclusion

In the case when the given process is a real-valued strictly stationary Markov process, the optimal sampling procedure exists and the extremal d.f. $F_0(\tau)$ degenerates at $\tau=\mu$ in the class $\mathscr{F}(\mu)$ of d.f.'s with mean μ, while the optimal sampling procedure does not exist in the class $\mathscr{F}$ of all d.f.'s.

In the case when the given process is a complex-valued strictly stationary Markov process, the optimal sampling procedure always exists and the extremal d.f. $F_0(\tau)$ degenerates at $\tau=\tau^*$ in $\mathscr{F}$.

In the general cases (including stationary non-Markovian processes), the time sampled process $\{x_{\tau,n}(\omega)\}$ can be proved to be a weakly stationary process with the spectral distribution function $P(\lambda)$. (Detailed forms of $P(\lambda)$ are stated in [2].)

Therefore, it is seen that

$$V(\hat{m})=\Delta P(0)+\frac{1}{n^2}\int_{-\pi}^{\pi}\frac{\sin^2\frac{n\lambda}{2}}{\sin^2\frac{\lambda}{2}}dP_0(\lambda)\,,$$

where $\Delta P(0)=P(0)-P(0-)$, $P_0(\lambda)=P(\lambda)$ if $\lambda<0$, and $P_0(\lambda)=P(\lambda)-\Delta P(0)$ if $\lambda\geqq 0$ (see section 1.7 in [5]). The term $\Delta P(0)$ vanishes when the original process $\{x(t)\}$ has a spectral density function. Since the stationary Markov processes have spectral density functions, the results obtained in this paper may be considered as the asymptotic evaluations of the second term in (22), i.e., those of $\left[\frac{dP}{d\lambda}\right]_{\lambda=0}$.

Acknowledgment

The author expresses his gratitude to the referee for his helpful suggestions, by which the paper was improved.

THE INSTITUTE OF STATISTICAL MATHEMATICS

REFERENCES

[1] R. F. Gebhard, "A limit theorem for random interval sampling of a stochastic process," *Ann. Math. Statist.*, 35 (1964), 866-868.

[2] H. Akaike, "Effect of timing-error on the power spectrum of sampled-data," *Ann. Inst. Stat. Math.*, 11 (1960), 145-165.

[3] D. R. Cox, *Renewal Theory, Matheuen's Monographs on Applied Probability and Statistics*, London, 1962.

[4] J. L. Doob, *Stochastic Processes*, John Wiley, New York, 1953, 518-526.

[5] U. Grenander and M. Rosenblatt, *Statistical Analysis of Stationary Time Series*, John Wiley, New York, 1957.