ARO 25839.8-MA

(2)

MARKOV CHAIN MOMENT FORMULAS FOR REGENERATIVE SIMULATION

by

James M. Calvin

AD-A210 684

TECHNICAL REPORT No. 35

June 1989

Prepared under the Auspices
of
U.S. Army Research Contract
DAAL03-88-K-0063

DEPARTMENT OF OPERATIONS RESEARCH
STANFORD UNIVERSITY
STANFORD, CALIFORNIA

89 7 31 206

UNCLASSIFIED MASTER COPY - FOR REPRODUCTION PURPOSES

SECURITY CLASSIFICATION OF THIS PAGE

REPORT DOCUMENTATION PAGE

1a. REPORT SECURITY CLASSIFICATION Unclassified	1b. RESTRICTIVE MARKINGS
2a. SECURITY CLASSIFICATION AUTHORITY	3. DISTRIBUTION/AVAILABILITY OF REPORT Approved for public release; distribution unlimited.
2b. DECLASSIFICATION/DOWNGRADING SCHEDULE	
4. PERFORMING ORGANIZATION REPORT NUMBER(S) Technical Report No. 35	5. MONITORING ORGANIZATION REPORT NUMBER(S) ARO 25839.8-MA
6a. NAME OF PERFORMING ORGANIZATION Dept. of Operations Research 6b. OFFICE SYMBOL (If applicable)	7a. NAME OF MONITORING ORGANIZATION U. S. Army Research Office
6c. ADDRESS (City, State, and ZIP Code) Stanford, CA 94305-4022	7b. ADDRESS (City, State, and ZIP Code) P. O. Box 12211 Research Triangle Park, NC 27709-2211
8a. NAME OF FUNDING/SPONSORING ORGANIZATION U. S. Army Research Office 8b. OFFICE SYMBOL (If applicable)	9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER DAAL03-88-K-0063
8c. ADDRESS (City, State, and ZIP Code) P. O. Box 12211 Research Triangle Park, NC 27709-2211	10. SOURCE OF FUNDING NUMBERS PROGRAM ELEMENT NO. / PROJECT NO. / TASK NO. / WORK UNIT ACCESSION NO.

11. TITLE (Include Security Classification)
Markov Chain Moment Formulas for Regenerative Simulation

12. PERSONAL AUTHOR(S)
James M. Calvin

13a. TYPE OF REPORT	13b. TIME COVERED	14. DATE OF REPORT (Year, Month, Day)	15. PAGE COUNT
Technical	FROM ____ TO ____	June 1989	13

16. SUPPLEMENTARY NOTATION
The view, opinions and/or findings contained in this report are those of the author(s) and should not be construed as an official Department of the Army position, policy, or decision, unless so designated by other documentation.

17. COSATI CODES			18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number)
FIELD	GROUP	SUB-GROUP	Regenerative simulation, moment formulas, Markov chains.

19. ABSTRACT (Continue on reverse if necessary and identify by block number)

Let $\{X_n : n \geq 0\}$ be a regenerative Markov chain on a general state space, and f a real-valued bounded function. Let τ and Z be random variables that have the distribution of a regeneration cycle length and the sum of $f(X_k)$ over a cycle, respectively. This paper derives expressions for moments of the form $E(\tau^j Z^k)$, which are then used to gain insight into the qualities of regenerative estimators based on different regeneration points.

20. DISTRIBUTION/AVAILABILITY OF ABSTRACT ☐ UNCLASSIFIED/UNLIMITED ☐ SAME AS RPT ☐ DTIC USERS	21. ABSTRACT SECURITY CLASSIFICATION Unclassified
22a. NAME OF RESPONSIBLE INDIVIDUAL	22b. TELEPHONE (Include Area Code) / 22c. OFFICE SYMBOL

8. References

[1] Billingsley, P., "Convergence of Probability Measures," John Wiley and Sons, 1968.

[2] Chung, K.L., "Markov Chains with Stationary Transition Probabilities," Springer-Verlag, Berlin-New York-Heidelberg, 1960.

[3] Cogburn, R., "A Uniform Theory for Sums of Markov Chain Transition Probabilities," *The Annals of Probability* **3**, No.2, 1975 191-214.

[4] Cogburn, R., "The Central Limit Theorem for Markov Processes," *Proc. Sixth Berkeley Symp. Math. Statist. Prob.* **2** 485-512.

[5] Glynn, P.W. and D.L. Iglehart, "A Joint Central Limit Theorem for the Sample Mean and Regenerative Variance Estimator," *Annals of Operations Research* **8**, 1987, 41-55.

[6] Malinovskii, V.K., "Limit Theorems for Harris Chains, I,", *Siam Journal Probability*.

[7] Orey, S. ,"Recurrent Markov Chains," *Pacific J. Math* **9** (1959), 805-827.

[8] Orey, S. ,"Lecture Notes on the Limit Theorems for Markov Chain Transition Probabilities," ,Van Nostrand, New York, 1971.

[9] Dellacherie, C. , and P. Meyer, "Probabilities and Potential C", North-Holland, 1988.

[10] Selby, S. (ed.), "Standard Mathematical Tables, 23^{rd} edition," CRC Press.

Markov Chain Moment Formulas for Regenerative Simulation

Abstract

Let $\{X_n : n \geq 0\}$ be a regenerative Markov chain on a general state space, and f a real-valued bounded function. Let τ and Z be random variables that have the distribution of a regeneration cycle length and the sum of $f(X_k)$ over a cycle, respectively. This paper derives expressions for moments of the form $E(\tau^j Z^k)$, which are then used to gain insight into the qualities of regenerative estimators based on different regeneration points.

Keywords

Regenerative simulation, moment formulas, Markov chains.

1. Introduction

The regenerative method of simulation output analysis uses the fact that the interblocks of a regenerative stochastic process are independent and identically distributed to construct a consistent estimator of the variance constant used to derive confidence intervals. If a process has more than one regeneration point, the estimator will have the same limiting value no matter which point is used to block the observations. While all such estimators have the same limit, different regeneration points may yield variance estimators with different variances. A common rule of thumb for obtaining an estimator with low variance is to choose the regeneration point that has the least mean regeneration time.

Glynn and Iglehart ([5]) proved a bivariate central limit theorem for the regenerative point estimator and the standard deviation estimator. Numerical calculations presented in the paper showed that the off-diagonal element in the covariance matrix appeared to be independent of the return state used to delimit regenerative cycles. The purpose of this paper is to derive an expression for the covariance matrix that appears in the central limit theorem in the case of a class of Markov chains. The expressions derived show that the off-diagonal term is independent of the return state. Some insight is gained into the nature of the variance of the variance estimators for different return states. An example is given where the state that yields the least variable variance estimator has the greatest mean regeneration time.

2. Notation

Let $(S, \mathcal{S})$ be a measurable space, where the σ-field $\mathcal{S}$ is countably generated (for example, S could be a metric space and $\mathcal{S}$ its Borel σ-field). Let P be a Markov kernel on $(S, \mathcal{S})$, and put $P^0(x, \cdot) = \delta_x$, and for $n > 1$,

$$P^n(x, A) = \int_S P^{n-1}(x, dy) P(y, A).$$

For any initial probability φ, the Markov kernel P determines a probability measure P_φ on the product measurable space $\prod_{n=0}^{\infty}(S^n, \mathcal{S}^n)$, where each $(S^n, \mathcal{S}^n) = (S, \mathcal{S})$, through the relations

$$P_\varphi(X_0 \in A_0, \cdots, X_n \in A_n) = \int_{A_0} \varphi(dx_0) \int_{A_1} P(x_0, dx_1) \cdots \int_{A_{n-1}} P(x_{n-2}, dx_{n-1}) P(x_{n-1}, A_n)$$

where X_k is the projection of $\prod_{n=0}^{\infty}(S^n, \mathcal{S}^n)$ onto $(S^k, \mathcal{S}^k)$. We write E_φ for the expectation with respect to the probability P_φ, and if $\varphi = \delta_x$, then we write P_x and E_x.

For $A \in \mathcal{S}$, s_A will denote the first return time of the chain to the set A and τ_A will denote the first hitting time of the set A (s and τ coincide unless the chain starts in A, in which case $s_A = 0$). We will write s_x and τ_x instead of $s_{\{x\}}$ and $\tau_{\{x\}}$.

For $A \in \mathcal{S}$ and $x \in S$ let

$$P^n_{xz}(A) = P_x\{X_n \in A; X_k \neq z, 0 < k < n\},$$

and define the kernels

$$Q_k(x, A) = \sum_{n=1}^{\infty} n^k P^n_{xz}(A), \quad k = 0, 1, \cdots.$$

$Q_0(x, A)$ is the expected amount of time the chain spends in A before absorption at z.

We will consider only chains that are ***uniformly φ-recurrent***; that is, there exists a σ-finite measure φ such that if $\varphi(A) > 0$, then $P_x[\tau_A \geq n] \to 0$ uniformly in $x \in S$. Uniformly φ-recurrent chains have a unique invariant probability measure, which will be denoted by π. We assume that the chain is aperiodic.

Throughout the paper z will denote a fixed return state with $\pi(\{z\}) > 0$. When subscripts are omitted the state will be understood to be z: e.g. $E(\tau) = E_z(\tau_z)$.

A consequence of the uniform recurrence is that for every $x \in S$, $E_x(\tau_z^2) < \infty$, and therefore $Q_0(x, \cdot)$ and $Q_1(x, \cdot)$ are finite measures for each x, with respective variations

$$Q_0(x, S) = \sum_{n=1}^{\infty} P_x\{\tau_z \geq n\} = E_x(\tau_z)$$

and

$$Q_1(x, S) = \sum_{n=1}^{\infty} n P_x\{\tau_z \geq n\} = \sum_{n=1}^{\infty} \left(\frac{n + n^2}{2}\right) P_x\{\tau_z = n\}$$

$$= \frac{1}{2} E_x(\tau_z) + \frac{1}{2} E_x(\tau_z^2).$$

Note that $Q_1(z, \{z\}) = E(\tau)$.

For the chain to be Harris recurrent (also called φ-recurrent) requires only that there exist a σ-finite measure φ such that

$$\varphi(A) > 0 \Rightarrow P_x[\tau_A < \infty] = 1$$

for every $x \in S$. Thus we are considering a sub-class of Harris recurrent Markov chains.

Let $F(S)$ denote the Banach space of bounded measurable functions from S to R with supremum norm. and denote by $M(S)$ the Banach space of finite signed measures on $(S, \mathcal{S})$ with total variation norm. We will use the same notation $\|\cdot\|$ for both norms, the context making clear which one applies.

There is a natural bilinear functional connecting $F(S)$ and $M(S)$, given by

$$(\mu, f) \rightarrow \mu f \triangleq \int_S f(x)\mu(dx)$$

wh -e $\mu \in M(S)$ and $f \in F(S)$. If N is a kernel, $\mu \in M(S)$, and $f \in F(S)$, we obtain a measure μN and a bounded function Nf given by

$$\mu N(A) = \int_S \mu(dx)N(x, A),$$

and

$$Nf(x) = \int_S N(x, dy)f(y).$$

An expression such as μNf is unambiguous, since

$$\mu(Nf) = (\mu N)f.$$

If N and M are kernels, we define the composition of M and N, MN, by

$$MN(x, A) = \int_S M(x, dy)N(y, A).$$

Composition is associative, so we can write an expression such as $\mu NMQf$ without parentheses. To avoid excessive use of parentheses we adopt the convention that functions are multiplied first in an expression: for example, if $f, g, h \in F(S)$, then

$$\mu NfMgh = \mu N(f(M(gh))).$$

Let f be a bounded real valued function on S, with $E_\pi f(X) = 0$.

3. Transition Probability Lemmas

The purpose of the lemmas in this section is to give expressions for Q_k in terms of the n-step transition probabilities for the Markov chain.

We will make use of the following fact (Theorem 6.1 in [8]).

Lemma 1: Consider a chain on $(S, \mathcal{S})$ which is uniformly φ-recurrent. If the chain is aperiodic, there exist $a < \infty$ and $\rho < 1$ such that

$$\| (\lambda_1 - \lambda_2) P^n \| \leq a\rho^n \| \lambda_1 - \lambda_2 \|$$

for any two probability measures λ_1 and λ_2 on $(S, \mathcal{S})$.

Proof: See [8]. ∎

Define the kernels

$$\Pi(x, dy) \triangleq \pi(dy), \quad G \triangleq \sum_{n=1}^{\infty} (P^n - \Pi), \quad H_k(x, A) \triangleq E_x\{\tau_z^k\} P(z, A), \quad k \geq 0,$$

$$H_0(x, A) \triangleq P(z, A) - P(x, A),$$

and let

$$g(x) \triangleq E_x(\tau_z).$$

The function g is measurable and by assumption is bounded. We will use the notation G_z to denote the measure $G(z, \cdot)$.

By Lemma 1, expressions such as

$$(\mu - \nu)G, \quad (\mu - \nu)Q_1$$

denote finite measures for any probability measures μ and ν.

Lemma 2:

$$Q_k = Q_k \Pi - \left(\sum_{j=1}^{k} \binom{k}{j} (-1)^j Q_{k-j} + H_k \right) (I + G).$$

In particular, for all $A \in \mathcal{S}$ and $x \in S$

$$Q_0(x, A) = g(x)\pi(A) + G(x, A) - G(z, A)$$

and

$$Q_1(x, A) = Q_1(x, S)\pi(A) + G(x, A) - G(z, A) + GG(x, A) - GG(z, A) - g(x)G_z(A).$$

Note that $Q_0(z, \cdot) = E(\tau)\pi(\cdot)$.

Proof: By definition

$$\int_{S \setminus \{z\}} n^k P_{xz}^n(dy) P(y, A) = n^k P_{xz}^{n+1}(A)$$

$$= (n+1)^k P_{xz}^{n+1}(A) + \sum_{j=1}^{k} \binom{k}{j} (-1)^j (n+1)^{k-j} P_{xz}^{n+1}(A).$$

Summing over n gives

$$\int_{S\setminus\{z\}} Q_k(x,dy)P(y,A) = Q_k(x,A) - P_{xz}(A) + \sum_{j=1}^{k}\binom{k}{j}(-1)^j\{Q_{k-j}(x,A) - P_{xz}(A)\}$$

$$= Q_k(x,A) + \sum_{j=1}^{k}\binom{k}{j}(-1)^j Q_{k-j}(x,A) - P_{xz}(A)1_{\{k=0\}}$$

and so

$$\int_S Q_k(x,dy)P(y,A) = Q_k(x,A) + \sum_{j=1}^{k}\binom{k}{j}(-1)^j Q_{k-j}(x,A) + Q_k(x,\{z\})P(z,A) - P_{xz}(A)1_{\{k=0\}}.$$

We can write the last equation as

$$Q_kP = Q_k + \sum_{j=1}^{k}\binom{k}{j}(-1)^j Q_{k-j} + H_k,$$

or

$$Q_k = Q_kP - \sum_{j=1}^{k}\binom{k}{j}(-1)^j Q_{k-j} - H_k.$$

Iterating the last equation gives

$$Q_k = Q_kP^n - \sum_{j=1}^{k}\binom{k}{j}(-1)^j Q_{k-j}\sum_{l=0}^{n}P^l - H_k\sum_{l=0}^{n}P^l.$$

Letting $n \to \infty$ and using bounded convergence gives

$$Q_k = Q_k\Pi - \left(\sum_{j=1}^{k}\binom{k}{j}(-1)^j Q_{k-j} + H_k\right)(I+G),$$

which is the result. ∎

4. Moment Calculations

Let τ and Z be random variables that have the same distributions as τ_z and

$$\sum_{n=1}^{\tau_z} f(X_n),$$

respectively, under P_z.

In this section we will derive expressions for moments of the form $E(\tau^i Z^j)$. These moments will be expressed in terms of the following quantities:

$$\chi_1 = \pi f g,$$

$$\chi_2 = \pi f^2 g + 2\pi f G f g,$$

$$\eta_1 = -(\delta_z + G_z) f,$$

and

$$\eta_2 = -(\delta_z + G_z) f^2 - 2(\delta_z + G_z) f G f.$$

We begin by determining expressions for $E(Z^k)$. Below we define several quantities that depend on the transition probability of the chain, but not on the return state z. For consistency of notation, we will use m_2 to denote the quantity usually called σ^2 (both notations will be used). Let

$$\sigma^2 = m_2 \triangleq \pi f^2 + 2\pi f G f$$

$$= E_\pi f(X_0)^2 + 2\sum_{n=1}^{\infty} E_\pi[f(X_0)f(X_n)],$$

$$m_3 = E_\pi f(X_0)^3 + 3\sum_{n=1}^{\infty} E_\pi[f(X_0)^2 f(X_n)] + 3\sum_{n=1}^{\infty} E_\pi[f(X_0)f(X_n)^2]$$

$$+6\sum_{n=1}^{\infty}\sum_{m=1}^{\infty} E_\pi[f(X_0)f(X_n)f(X_{n+m})],$$

and

$$m_4 = E_\pi f(X_0)^4 + 4\sum_{n=1}^{\infty} E_\pi[f(X_0)^3 f(X_n)] + 4\sum_{n=1}^{\infty} E_\pi[f(X_0)f(X_n)^3]$$

$$+6\sum_{n=1}^{\infty}\{E_\pi[f(X_0)^2 f(X_n)^2] - E_\pi[f(X_0)^2]E_\pi[f(X_n)^2]\}$$

$$+12\sum_{n=1}^{\infty}\sum_{m=1}^{\infty}\{E_\pi[f(X_0)^2 f(X_n)f(X_{n+m})] - E_\pi[f(X_0)^2]E_\pi[f(X_n)f(X_{n+m})]\}$$

$$+12\sum_{n=1}^{\infty}\sum_{m=1}^{\infty} E_\pi[f(X_0)f(X_n)^2 f(X_{n+m})] + 12\sum_{n=1}^{\infty}\sum_{m=1}^{\infty} E_\pi[f(X_0)f(X_n)f(X_{n+m})^2]$$

$$+24\sum_{n=1}^{\infty}\sum_{m=1}^{\infty}\sum_{k=1}^{\infty} E_\pi[f(X_0)f(X_n)f(X_{n+m})f(X_{n+m+k})],$$

and in general

$$m_n = \sum_{k=1}^{n} \sum_{p_1, \cdots, p_k} \binom{n}{p_1} \binom{n-p_1}{p_2} \cdots \binom{n-p_1-\cdots-p_{k-1}}{p_k} \pi f^{p_1} G f^{p_2} G \cdots G f^{p_k},$$

where the second sum is over all positive p_i that sum to n.

Lemma 3: If the series in the definitions of m_2, m_3, and m_4 converge absolutely, then

$$E\left(Z^2\right) = E(\tau) m_2,$$

$$E(Z^3) = E(\tau)\left(m_3 + 3m_2(\chi_1 + \eta_1)\right),$$

and

$$E\left(Z^4\right) = E(\tau)\Big(m_4 + 4m_3(\chi_1 + \eta_1) + 6m_2\left((\chi_1 + \eta_1)^2 + \chi_1^2 + \eta_1^2 + \chi_2 + \eta_2\right)\Big).$$

Proof: Add to S a new state Δ with $f(\Delta) = 0$, and define random variables $\{\xi_n\}$ taking values in $S \cup \{\Delta\}$ by

$$\xi_n = \begin{cases} \Delta, & \text{if } X_k = z \text{ for some } 0 < k < n, \\ X_n, & \text{otherwise.} \end{cases} \tag{1}$$

According to the definition, $\xi_0 = X_0$ and $\xi_{\tau_z} = z$.

We will establish the third moment result. Using the random variables defined in (1),

$$E_z \left(\sum_{n=1}^{\tau_z} f(X_n) \right)^3 = E_z \left(\sum_{n=1}^{\infty} f(\xi_n) \right)^3$$

$$= E_z \Bigg(\sum_{n=1}^{\infty} f(\xi_n)^3 + 3\sum_{n=1}^{\infty}\sum_{m=1}^{\infty} f(\xi_n)^2 f(\xi_{n+m}) + 3\sum_{n=1}^{\infty}\sum_{m=1}^{\infty} f(\xi_n) f(\xi_{n+m})^2$$

$$+6\sum_{n=1}^{\infty}\sum_{m=1}^{\infty}\sum_{l=1}^{\infty} f(\xi_n) f(\xi_{n+m}) f(\xi_{n+m+l}) \Bigg)$$

$$= \sum_{n=1}^{\infty} E_z\left[f(\xi_n)^3\right] + 3\sum_{n=1}^{\infty}\sum_{m=1}^{\infty} E_z\left[f(\xi_n)^2 f(\xi_{n+m})\right] + 3\sum_{n=1}^{\infty}\sum_{m=1}^{\infty} E_z\left[f(\xi_n) f(\xi_{n+m})^2\right]$$

$$+6\sum_{n=1}^{\infty}\sum_{m=1}^{\infty}\sum_{l=1}^{\infty} E_z\left[f(\xi_n) f(\xi_{n+m}) f(\xi_{n+m+l})\right]$$

where the interchange is allowed because of the assumed absolute convergence of m_3

$$= \sum_{n=1}^{\infty} \int_S P_{zz}^n(dx) f(x)^3 + 3\sum_{n=1}^{\infty}\sum_{m=1}^{\infty} \int_{S\setminus\{z\}} P_{zz}^n(dx) f(x)^2 \int_S P_{xz}^m(dy) f(y)$$

$$+3\sum_{n=1}^{\infty}\sum_{m=1}^{\infty} \int_{S\setminus\{z\}} P_{zz}^n(dx) f(x) \int_S P_{xz}^m(dy) f(y)^2$$

$$+6\sum_{n=1}^{\infty}\sum_{m=1}^{\infty}\sum_{l=1}^{\infty} \int_{S\setminus\{z\}} P_{zz}^n(dx) f(x) \int_{S\setminus\{z\}} P_{xz}^m(dy) f(y) \int_S P_{yz}^l(dv) f(v)$$

$$= Q_0(z,\cdot)f^3 + 3Q_0(z,\cdot)f^2Q_0f + 3Q_0(z,\cdot)fQ_0f^2 + 6Q_0(z,\cdot)fQ_0fQ_0f$$

$$-3f(z)Q_0(z,\cdot)f^2 - 6f(z)Q_0(z,\cdot)fQ_0f$$

$$= E(\tau)\left[\pi f^3 + 3\pi f^2Q_0f + 3\pi fQ_0f^2 + 6\pi fQ_0fQ_0f\right]$$

$$-3f(z)Q_0(z,\cdot)f^2 - 6f(z)Q_0(z,\cdot)fQ_0f$$

$$= E(\tau)m_3 + 3m_2(\chi_1 + \eta_1).$$

and the result follows from Lemma 2.

The second and fourth moment results follow from similar arguments. ∎

The next lemma gives expressions for some of the mixed moments.

Lemma 4: If

$$E_z\left(\tau_z \sum_{n=1}^{\tau_z} |f(X_n)|\right) < \infty$$

then

$$E[\tau Z] = E(\tau)\,(\chi_1 + \eta_1)\,.$$

Also

$$E\left[\tau Z^2\right] = \frac{1}{2}\sigma^2\left[E(\tau) + E(\tau^2)\right] + 2E(\tau)\pi fGf + +2E(\tau)\pi fGGf$$

$$+E(\tau)\Big((\chi_1+\eta_1)^2 + \chi_1^2 + \eta_1^2 + \chi_2 + \eta_2\Big).$$

Proof: Using the random variables defined by (1),

$$E_z\left(\tau_z \sum_{n=1}^{\tau_z} f(X_n)\right) = E_z\left(\sum_{n=1}^{\infty} \tau_z f(\xi_n)\right)$$

$$= \sum_{n=1}^{\infty} E_z\left(\tau_z f(\xi_n)\right)$$

(where the interchange is allowed by the lemma's hypothesis)

$$= \sum_{n=1}^{\infty} E_z\left[f(\xi_n)\,(n + s_{\xi_n z})\right] = \sum_{n=1}^{\infty} n\cdot E_z[f(\xi_n)] + \sum_{n=1}^{\infty} E_z[f(\xi_n)s_{\xi_n z}]$$

$$= \sum_{n=1}^{\infty} n\cdot E_z[f(\xi_n)] + \sum_{n=1}^{\infty} E_z E[f(\xi_n)s_{\xi_n z}|\xi_n] = \sum_{n=1}^{\infty} nE_z[f(\xi_n)] + \sum_{n=1}^{\infty} E_z[f(\xi_n)E[s_{\xi_n z}|\xi_n]]$$

$$= \sum_{n=1}^{\infty} n\int_S P_{zz}^n(dx)f(x) + \sum_{n=1}^{\infty}\int_S P_{zz}^n(dx)f(x)E_x(s_z)$$

$$= \int_S \sum_{n=1}^{\infty} nP_{zz}^n(dx)f(x) + \int_S \sum_{n=1}^{\infty} P_{zz}^n(dx)f(x)E_x(s_z)$$

$$= \int_S Q_1(z,dx)f(x) + \int_S Q_0(z,dx)f(x)E_x(s_z)$$

(the interchange is justified since $\sum_{n=1}^{\infty} nP_{zz}^n(\cdot)$ converges absolutely)

$$= Q_1(z,\cdot)f + Q_0(z,\cdot)fg - Q_0(z,\{z\})f(z)E_z(\tau_z)$$

$$= -g(z)G_z f + g(z)\pi fg - g(z)f(z)$$

$$= E(\tau)(\eta_1 + \chi_1),$$

which is the desired result.

Proceeding similarly for the second equation,

$$E(\tau Z^2) = \sum_{n=1}^{\infty} E_z\left[f^2(\xi_n)(n + s_{\xi_n z})\right] + 2\sum_{n=1}^{\infty}\sum_{m=1}^{\infty} E_z\left[(n+m+s_{\xi_{n+m} z})f(\xi_n)f(\xi_{n+m})\right]$$

$$= \sum_{n=1}^{\infty} nE_z\left[f^2(\xi_n)\right] + \sum_{n=1}^{\infty} E_z\left[f^2(\xi_n)s_{\xi_n z}\right]$$

$$+2\sum_{n=1}^{\infty}\sum_{m=1}^{\infty} nE_z[f(\xi_n)f(\xi_{n+m})]$$

$$+2\sum_{n=1}^{\infty}\sum_{m=1}^{\infty} mE_z[f(\xi_n)f(\xi_{n+m})] + 2\sum_{n=1}^{\infty}\sum_{m=1}^{\infty} E_z[f(\xi_n)f(\xi_{n+m})s_{\xi_n z}]$$

$$= \sum_{n=1}^{\infty} n\int_S P_{zz}^n(dx)f^2(x) + \sum_{n=1}^{\infty}\int_S P_{zz}^n(dx)f(x)^2 E_x(s_z) + 2\sum_{n=1}^{\infty}\sum_{m=1}^{\infty} n\int_{S\setminus\{z\}} P_{zz}^n(dx)f(x)\int_S P_{xz}^m(dy)f(y)$$

$$+2\sum_{n=1}^{\infty}\sum_{m=1}^{\infty} m\int_{S\setminus\{z\}} P_{zz}^n(dx)f(x)\int_S P_{xz}^m(dy)f(y)$$

$$+2\sum_{n=1}^{\infty}\sum_{m=1}^{\infty}\int_{S\setminus\{z\}} P_{zz}^n(dx)f(x)\int_S P_{xz}^m(dy)f(y)E_y(s_z)$$

$$= Q_1(z,\cdot)f^2 + Q_0(z,\cdot)f^2 g - g(z)f(z)^2 + 2Q_1(z,\cdot)fQ_0 f + 2Q_0(z,\cdot)fQ_1 f - 2f(z)Q_1(z,\cdot)f$$

$$+2Q_0(z,\cdot)fQ_0 fg - 2f(z)g(z)\pi fg + 2f(z)^2 g(z)$$

$$= [Q_1(z,S)\pi - g(z)G_z]f^2 + g(z)\pi f^2 g + 2[Q_1(z,S)\pi - g(z)G_z]f[g\pi + G - G_z]f$$

$$+2g(z)\pi f[Q_1(x,S)\pi + G - G_z + GG - GG_z - gG_z]f + 2g(z)\pi f[g\pi + G - G_z]fg$$

$$+g(z)f(z)^2 - 2f(z)g(z)\pi fg - 2f(z)[Q_1(z,S)\pi - g(z)G_z]f$$

$$= Q_1(z,S)\pi f^2 - g(z)G_z f^2 + g(z)\pi f^2 g + 2Q_1(z,S)\pi fGf + 2g(z)(G_z f)^2 - 2g(z)G_z fGf$$

$$+2g(z)\pi fGf + 2g(z)\pi fGGf - 2g(z)(\pi fg)(G_z f) + 2g(z)(\pi fg)^2 + 2g(z)\pi fGfg$$

$$+g(z)f(z)^2 - 2f(z)g(z)\pi fg + 2f(z)g(z)G_z f$$

$$= Q_1(z,S)m_2 + 2E(\tau)\pi fGf + 2E(\tau)\pi fGGf$$

$$+E(\tau)[-G_z f^2 + \pi f^2 g + 2(G_z f)^2 - 2G_z fGf - 2(\pi fg)(G_z f) + 2(\pi fg)^2$$

$$+2\pi fGfg + f(z)^2 - 2f(z)\pi fg + 2f(z)(G_z f)]$$

$$= Q_1(z,S)\sigma^2 + 2E(\tau)\pi fGf + 2E(\tau)\pi fGGf + E(\tau)\left[\eta_2 + 2\eta_1^2 + 2\chi_1\eta_1 + 2\chi_1^2 + \chi_2\right],$$

as desired. ∎

5. Estimator Covariance Matrix

Let $S_0 = 0$, and $S_n = f(X_0) + \cdots + f(X_{n-1})$. Under the assumptions on the chain given in section 2,

$$\frac{1}{n} E\left(S_n^2\right) \to \sigma^2$$

as $n \to \infty$, and

$$n^{-1/2} S_n \Rightarrow \mathcal{N}(0, \sigma^2)$$

as $n \to \infty$. We are interested in estimating σ^2 in order to obtain confidence intervals. In the regenerative method, S_n is divided up into independent blocks by starting a new block whenever a regeneration point is reached. If Z_i is the ith block and K_n is the number of regenerative cycles in the first n observations, then

$$\sigma^2 = \lim_{n \to \infty} \frac{1}{n} E\left(Z_1 + \cdots + Z_{K_n}\right)^2 = \lim_{n \to \infty} \frac{1}{n} \sum_{j=1}^{K_n} E\left(Z_j^2\right)$$

as $n \to \infty$. Choosing different regeneration points will in general give different estimator variances.

Let $r(n)$ and $s(n)$ denote the regenerative mean and standard deviation estimators, respectively, based on observation of the chain up to time n:

$$r(n) = \frac{1}{n} \sum_{j=1}^{K_n} f(X_j), \quad s(n)^2 = \frac{1}{n} \sum_{j=1}^{K_n} \left(Z_j - r(n)\tau_j\right)^2.$$

It is shown in [5] (for general regenerative processes) that

$$n^{1/2}\left(r(n) - r, s(n) - \sigma\right) \Rightarrow \mathcal{N}(0, D),$$

where

$$D_{11} = E(Z^2)/E(\tau),$$

$$D_{12} = \frac{E(Z^3) - 3\sigma^2 E(\tau Z)}{2\sigma E(\tau)},$$

and

$$D_{22} = \left(E(Z^4) - 2\sigma^2 E(\tau Z^2) + \sigma^4 E(\tau^2) - E(\tau)^{-1}\left(4E(\tau Z)E(Z^3) - 8\sigma^2[E(\tau Z)]^2\right)\right) / \left(4\sigma^2 E(\tau)\right).$$

Using the formulas from Lemmas 3 and 4, the covariance matrix can be written

$$D = \begin{bmatrix} \sigma^2 & \frac{m_3}{2\sigma} \\ & \\ \frac{m_3}{2\sigma} & \begin{array}{c} c + \chi_1^2 + \eta_1^2 \\ +\chi_2 + \eta_2 \end{array} \end{bmatrix},$$

where

$$c \triangleq \frac{m_4}{4\sigma^2} - \frac{\sigma^2}{4} - \pi f G f - \pi f G G f$$

is independent of the return state. Notice that the diagonal term is also independent of the return state z, since as previously mentioned, σ^2 and m_3 are independent of the return state.

Let

$$\sqrt{n}(r(n) - r) \Rightarrow a,$$

$$\sqrt{n}(s(n) - \sigma) \Rightarrow b,$$

wh--e we view a and b as elements of the Hilbert space L_2 with inner product

$$(x, y) = E(xy).$$

Then

$$(a,a) = \sigma^2, \quad (b,b) = D_{22}, \quad (a,b) = \frac{m_3}{2\sigma}.$$

It follows that we can write

$$b = \frac{m_3}{2\sigma^3}a + q,$$

where $(a,q) = 0$ and

$$(q,q) = \frac{1}{4\sigma^2 E(\tau)} E\left(Z^2 - \sigma^2\tau\right)^2 - \frac{1}{4\sigma^2(E(\tau))^2}\left(E(Z^3 - \sigma^2\tau Z)\right)^2.$$

For a random variable X with finite third moment define the *coefficient of momental skewness* ([10]) of X by

$$\frac{E(X - EX)^3}{2\text{var}(X)^{3/2}}.$$

Let κ_n be the coefficient of momental skewness for the random variable $\sqrt{(n)}S_n$ under the initial distribution π. Clearly κ_n is defined independently of the return state, and

$$\kappa \triangleq \lim_{n\to\infty} \kappa_n = \frac{m_3}{2\sigma^3}.$$

With this notation, $D_{12} = \kappa\sigma^2$, and the orthogonal decomposition is

$$b = \kappa a + q.$$

For any symmetrical chain, for example a birth and death process on $\{-N, \cdots, 0, \cdots, N\}$ for which the birth and death parameters as well as the values of the function f are symmetrical about 0, $\kappa = 0$ and so a and b are orthogonal.

Choosing a return state to minimize variance of the standard deviation estimator is equivalent to choosing a return state to maximize correlation between the estimators for the mean and the standard deviation.

6. Example

Consider the Markov chain with state space $S = \{1, 2, 3\}$ and transition matrix

$$P = \begin{bmatrix} 1-\epsilon & \epsilon & 0 \\ 1/2 & 0 & 1/2 \\ 0 & \epsilon & 1-\epsilon \end{bmatrix},$$

for $0 < \epsilon < 1$. The stationary distribution is

$$\pi = \left(\frac{1}{2+2\epsilon}, \frac{2\epsilon}{2+2\epsilon}, \frac{1}{2+2\epsilon} \right).$$

Let $f = (-M, 0, M)$ for some $M > 0$; then $E_\pi f = 0$. The values of the quantities that vary with state are given below (the values for state 3 are the same as for state 1).

State	η_1^2	χ_1^2	η_2	λ_2
1,3	$\frac{M^2}{\epsilon^2}$	$\frac{M^2(1-\epsilon)^2}{\epsilon^2}$	$\frac{M^2(2-\epsilon)}{\epsilon^2}$	$\frac{M^2\epsilon(2-\epsilon)}{(1+\epsilon)^2}$
2	0	0	$\frac{M^2(2-\epsilon)}{\epsilon^2(1+\epsilon)}$	$-\frac{M^2(2-\epsilon)}{(1+\epsilon)^2}$

The difference in variances is

$$D_{22}(1) - D_{22}(2) = 2\left(\frac{M}{\epsilon}\right)^2,$$

while

$$E_1(\tau_1) = E_3(\tau_3) = 2 + 2\epsilon \to 2, \qquad E_2(\tau_2) = \frac{1+\epsilon}{\epsilon} \to \infty$$

as $\epsilon \downarrow 0$. Therefore, while the mean regeneration time for state 2 grows without bound as $\epsilon \downarrow 0$, it gives the least variable estimator, with the difference going to ∞ as $\epsilon \downarrow 0$. Essentially all of the difference is accounted for by the η_1^2 and χ_1^2 terms.

In this example, the kurtosis of S_n increases as M increases or ϵ decreases, so the variance of all 3 standard deviation estimators increases as $\epsilon \downarrow 0$.

7. Conclusion

The covariance matrix that appears in the central limit theorem for the regenerative mean and standard deviation estimators has been expressed in a form so that several conclusions could be reached. First, the off-diagonal term is independent of the return state chosen for blocking. Second, the expression for the variance of the standard deviation estimator shows that the variance is increased by kurtosis in the partial sum process. The variance does depend on the return state used for blocking, and an example showed that the state with the shortest mean return time can give the most variable standard deviation estimator.