마코프 누적 프로세스에서의 확률적 콘벡스성
(Stochastic Convexity in Markov Additive Processes)

윤 복 식

홍익대학교 기초과학과

ABSTRACT

Stochastic convexity(concavity) of a stochastic process is a very useful concept for various stochastic optimization problems. In this study we first establish stochastic convexity of a certain class of Markov additive processes through the probabilistic construction based on the sample path approach. A Markov additive process is obtained by integrating a functional of the underlying Markov process with respect to time, and its stochastic convexity can be utilized to provide efficient methods for optimal design or for optimal operation schedule of a wide range of stochastic systems. We also clarify the conditions for stochatic monotonicity of the Markov process, which is required for stochatic convexity of the Markov additive process. This result shows that stochastic convexity can be used for the analysis of probabilistic models based on birth and death processes, which have very wide application area. Finally we demonstrate the validity and usefulness of the theoretical results by developing efficient methods for the optimal replacement scheduling based on the stochastic convexity property.

1. 서 론

무작위적 특성을 가지고 시간에 따라 동적으로 변해가는 어떤 시스팀의 상태를 나타내는 스토캐스틱 프로세스 $\{X(t), t \geq 0\}$를 고려해보자. 시간 t에서 시스팀의 상태가 x일 때 f(x)의 비용이 소요된다면 [0,t] 사이에 소요되는 총비용은

$$Y(t) = \int_0^t f(X(u))\, du \tag{1}$$

과 같이 표현될 것이다. 만약 $\{X(t), t \geq 0\}$가 마코프 프로세스라면 $\{Y(t), t \geq 0\}$은 $\{X(t), t \geq 0\}$에서 유도된 마코프 누적 프로세스가 된다(Cinlar(1972a,b)). Y(t)가 의미하는 것은 시간 t까지의 누적된 비용이므로 그 비용을 최소화하는 형태의 스케쥴링 문제에 응용될 수 있어 $\{Y(t), t \geq 0\}$의 확률적 콘벡스성을 파악하면 최적 스케쥴링의 문제를 훨씬 쉽게 해결할 수 있을 것이다. 실제로 마코프 누적 프로세스는 마코프 프로세스로 모형화 할 수있는 광범위한 확률적 시스팀에서의 성능에 관련된 측도로 사용될 수 있기 때문에 본 연구의 실용적 가치는 매우 높다고 생각된다. 예를들면 fault-tolerant 콤퓨터 시스팀에서의 성능측도(Sumita et al.,1987), software reliability 분야에서의 실험 중단 시점(Ross,1987, Sumita와 Shanthikumar,1986), Maintenace 이론에서의 최적 교체 주기(Barlow와 Proschan,1975) 등의 문제에 직접적으로 적용될 수 있다. 결론적으로 본 연구는 확률적 콘벡스성에 관한 이론적인 기여는 물론 성능 평가나 최적 스케쥴링 문제에 매우 유용한 방법론을 제시할 수 있을 것이다.

2. 확률적 단조성과 콘벡스성

2.1. 확률적 단조성(stochastic monotonocity), 확률적 콘벡스성의 이론적 배경 - 샘플 경로 접근법(Sample path approach)

우선 SI에 대해서는 다음과 같은 샘플경로에 연관된 사실이 알려져

있다.

(정리 2.1) 만약 $\{X(\theta),\theta\in\Theta\}$가 SI이면, 모든 $\theta\in\Theta$에 대해

(i) $Z(\theta) \sim X(\theta)$ 이고

(ii) 만약 $\theta_1\le\theta_2$ 이면, $Z(\theta_1)\le Z(\theta_2)$ a.s.

를 만족하는 공통의 확률공간(common probability space)에서 정의된 확률변수들 $\{Z(\theta),\theta\in\Theta\}$가 존재한다.

(증명) $P\{X(\theta)>x\} = F^c(x,\theta)$이고, U는 uniform(0,1)의 확률변수라 하자. $(F^c)^{-1}(u,\theta) \equiv \inf\{x:F^c(x,\theta)\le u\}$ 라고 정의 하면 확률변수

$$Z(\theta) = (F^c)^{-1}(U,\theta),\ \theta\in\Theta \tag{1}$$

이 정리를 만족함을 알 수 있다. (cf.Kamae et.al.(1977)), Th1 참조.)

■

이와 유사하게 샘플경로에서의 확률적인 증가 콘벡스성, 콘케이브성을 각각 SICX(sp)(stochastically increasing and convex in sample path sense)와 SICV(sp)(stochastically increasing and concave in sample path sense)로 표현하면, 이들을 다음과 같이 정의할 수 있다.

(정의 2.1) $\theta_1\le\theta_2\le\theta_3\le\theta_4$이고 $\theta_2-\theta_1=\theta_4-\theta_3$를 만족하는 모든 $\theta_i\in\Theta$, i=1,2,3,4에 대해 아래의 조건 (i),(ii),(iii)를 만족하는 4개의 확률변수 Z_i, i=1,2,3,4가 동일한 확률공간에서 정의될 수 있으면 $\{X(\theta),\theta\in\Theta\}$를 SICX(sp)라고 말하고, (i),(iv),(v)를 만족하는 4개의 확률변수 Z_i, i=1,2,3,4가 동일한 확률공간에서 정의될 수 있으면 $\{X(\theta),\theta\in\Theta\}$를 SICV(sp)라고 말한다.

(i) $Z_i \sim X(\theta_i)$, i=1,2,3,4,

(ii) $Z_2-Z_1 \le Z_4-Z_3$ a.s.

(iii) $\max(Z_1,Z_2,Z_3) \le Z_4$ a.s.

(iv) $Z_2-Z_1 \ge Z_4-Z_3$ a.s.

(v) $Z_1 \le \min(Z_2,Z_3,Z_4)$ a.s.

(이때 ~ : stochastically equivalent, 즉 "같은 분포를 갖는")

SICX(sp)는 SICX보다 강한 확률적 성질이기 때문에 $\{X(\theta),\theta\in\Theta\}$가 SICX(sp)의 성질을 가지면 자동적으로 SICX의 성질을 갖게 된다.(Shaked & Shanthikumar(1988a),Th3.6) 따라서 $\{X(\theta),\theta\in\Theta\}$가 SICX 임을 보이기위해 SICX(sp)임을 보이면 충분한데, 이는 (정의 2.1)의 조건들을 만족시키는 확률변수들을 만들어 내는 구축적인 방법에 의해 비교적 간단하게 행해질 수 있다. 더우기 SICX(sp)는 SICX가 갖지못하는 매우 유용한 보존적 특성(예를들면 convolution이나 mixture에 닫혀있는 것, Shaked & Shanthikumar (1988a) 참조.) 들을 가지고 있기 때문에 많은 경우에 보다 편리하게 사용될 수 있는 개념이다.

1.2. 마코프 프로세스에서의 확률적 단조성

본 절에서는 θ가 시간일 때에 한정하여 매우 광범위하게 적용되고 있는 마코프 프로세스의 시간에 따른 단조성이나 콘벡스성에 대해 알아 보자. 우리가 고려할 마코프 프로세스는 상태공간(state space) S가 최소의 원소를 0으로 갖고 부분 순서화된(partially ordered) 공간일 경우인데, $S\in N^{+}=\{0,1,2,\ldots\}$인 경우나 양의 실수($\mathbb{R}^{+}$)인 경우를 포함하므로 거의 모든 실제적인 문제에 적용될 것이다. 시간은 연속시간 혹은 이산시간인 경우에 모두 고려되는데 이산시간의 마코프 프로세스를 마코프 체인이라고 부르기로 한다. 마코프 체인에 대해서는 다음과 같은 단조성의 충분 조건을 발견할 수 있다.

(정리 2.2) 마코프 체인 $\{X_n, n\in N^{+}\}$에서 $X_0=0$ a.s. 이고, 모든 $x\in S$에 대해 $P\{X_n>x\mid X_{n-1}=i\}$가 i의 증가함수이면 $\{X_n, n\in N^{+}\}$는 (n에 대해) SI이다.

(증명) Stoyan,1983 p64 (Th 4.2.4a) 참조 ■

연속시간의 마코프 프로세스의 경우에도 이와 유사하게 다음과 같은 단조성의 충분 조건을 확립할 수 있다.

(정리 2.3) 마코프 프로세스 $\{X(t), t\geq 0\}$에서 모든 $x \in S$ 에 대해 $P\{X(t)>x | X(0)=i\}$가 i의 증가함수이면, $X(0)=0$ a.s. 일때 $\{X(t), t\geq 0\}$는 (t에 대해) SI이다.

(증명) s<t 에 대해

$$\begin{aligned} P\{X(t)>x|X(0)=0\} &= \sum_i P\{X(t)>x|X(t-s)=i\}P\{X(t-s)=i|X(0)=0\} \\ &\geq \sum_i P\{X(t)>x|X(t-s)=0\}P\{X(t-s)=i|X(0)=0\} \\ &= P\{X(t)>x|X(t-s)=0\} \sum_i P\{X(t-s)=i|X(0)=0\} \\ &= P\{X(s)>x|X(0)=0\} \quad \blacksquare \end{aligned}$$

(정리 2.3)의 조건을 만족하는 마코프 프로세스 중에서 대표적인 것으로 대기이론등에 매우 광범위하게 응용되는 birth-death process를 들 수 있다.

(따름정리 2.4) $X(0)=0$ a.s.일 때 birth-death process $\{X(t), t\geq 0\}$는 (t에 대해) SI이다.

(증명) Ross(1983), p257 proposition 8.2.3에 의해 모든 $x \in S$ 에 대해 $P\{X(t)>x|X(0)=i\}$가 i의 증가함수 임을 알 수 있다. ■

마코프 체인의 단조성은 마코프 누적 프로세스의 확률적 콘벡스성을 확립하는 기초가 된다.

3. 마코프 누적 프로세스에서의 확률적 콘벡스성

앞에서 언급한 바와 같이 시스팀의 성능에 관련된 측도로서 마코프 누적 프로세스는 매우 유용한데 실제 모형화에 있어서 계산상의 난점때문에

이용에 어려움이 많다. 시작 시점부터 t까지의 마코프 프로세스의 과거 역사가 모두 고려되어야 하므로 이에 포함된 계산량이 매우 많게 되는 데 최적화 시점의 추적에 상당히 효율적인 방법을 사용하여야만 유용성이 손상되지 않을 것이다. 이러한 효율적인 방법을 찾아내는 이론적인 근거로서 확률적 콘벡스성은 특히 중요하다.

콘벡스 함수의 특성은 변화율이 점점 증가한다는 것이므로 (정리 2.1)을 이용하여 다음과 같은 유용한 결과를 얻을 수 있다.

(정리 3.1) 만약 $\{X(t), t \geq 0\}$이 SI 이면 마코프 누적프로세스 (1.1)은 다음과 같은 성질을 가지고 있다.

(a) 만약 f가 증가함수이면, $\{Y(t), t \geq 0\}$은 SICX(sp)이다.

(b) 만약 f가 감소함수이면, $\{Y(t), t \geq 0\}$은 SICV(sp)이다.

(증명) 우선 (a)의 경우에 $0 \leq t_1 \leq t_2 \leq t_3 \leq t_4$ 이고 $t_3 - t_1 = t_4 - t_2$ 를 만족하는 $t_i, i=1,2,3,4$를 선택하자. $\{X(t), t \geq 0\}$이 SI이므로 (정리 2.1)에 따라,

(i) 모든 $t \geq 0$ 에 대해 $Z(t) \sim X(t)$이고

(ii) $0 \leq u \leq t_2 - t_1$인 모든 $u \geq 0$에 대해 $Z(t_1+u) \leq Z(t_3+u)$ a.s.

를 만족하는 마코프 체인 $\{Z(t), t \geq 0\}$ 이 존재한다. 이제

$$Y(t) = \int_0^t f(Z(u))\, du \qquad (1)$$

로 놓으면

$$Y(t_i) \sim Y(t_i), \quad i=1,2,3,4 \qquad (2)$$

$$Y(t_1) + Y(t_4) \geq Y(t_2) + Y(t_3) \ \text{a.s.} \qquad (3)$$

$$\max\{Y(t_i), i=1,2,3\} \leq Y(t_4) \qquad (4)$$

이 됨을 쉽게 알 수 있으므로 (정의 2.2)에 의해 (a)가 성립함을 알 수 있다. (b)의 경우도 이와 유사하게 증명할 수 있다. ■

이 정리는 2.2절에서 언급한 마코프 프로세스의 단조성의 충분 조건과 결합하여 매우 광범위하게 응용될 수 있다. 그중에서 최적 대체 주기 결정에 관한 문제에 적용하여 효율적인 최적화 방법을 찾아낼 수 있을

것 이다.

4. 응용 예

어떤 기계의 작동 상태에 따라 가동비용(수리, 유지 비용 포함)과 생산율이 다른 경우를 고려해보자. 이때 작동 상태의 변화는 gracefully degrading system과 같은 fault-tolerant design 에 의해 일부 부품의 고장에 따른 기계의 생산 능력의 변동에서 기인할 수 도 있고(Sumita,1987) 여러번의 수리에 의한 기계 성능의 하락에 의한 것일 수도 있다(Stadje,1990).

$X(t) \in \{0,1,2,...,M\}$ 를 시스팀의 시간 t에서의 작동 상태라 하고 $X(t)$의 값이 클 수록 시스팀의 작동 조건이 악화되는 것을 의미한다고 하자. 예를 들면 시스팀이 M개의 동일한 부품을 사용한 병렬구조일때 고장이 난 부품의 갯수를 시스팀의 상태로 표현할 수 있을 것이다. 시스팀이 완전히 대체되는 시점 T까지는 부분품의 고장이나 수리를 통해 시스팀의 상태가 birth-death 프로세스형태로 변한다고 하자. 그때의 가동 비용과 생산수입을 각각 $c(X(t))$, $r(X(t))$라고 하고 시스팀대체 비용을 D라고 할때, 최초 대체 시점 T 까지의 이익은

$$R(T) = \int_0^T r(X(u)\ du - \int_0^T c(X(u))\}\ du \ - D \qquad (1)$$

으로 표현되고 $c(\cdot)$, $r(\cdot)$ 이 적분가능한 0보다 큰 함수라고 할때 기대 이익 $E[R(T)]$가 존재하고

$$E[R(T)] = \int_0^T E[r(X(u)]du - \int_0^T E[c(X(u)]du \ - D \qquad (2)$$

과 같이 표현될 수 있다.

위의 병렬 시스팀의 예에서와 같이 한 부품이 고장이 났을 경우 나

머지 부품들은 작동을 계속한다면 고장난 부품의 갯수가 적을수록 생산량은 많아질 것이고, 수리 비용은 적어질 것이다. 즉 비용은 증가함수, 수입은 감소함수로 가정한다면 현실적인 고려에서 타당할 것이다. 이를 가정할 때 2장및 3장의 결과를 이용하여 다음을 얻을 수 있다.

(정리 4.1) $X(0)=0$ a.s.이고, c를 증가 함수, r을 감소 함수라고 할 때 $E[R(T)]$는 T의 콘케이브 함수이다.

(증명)

우선 $\{X(t), t\geq 0\}$는 birth-death process 이므로 $X(0)=0$ a.s. 일때 SI이다. 주어진 조건을 가지고 (정리 3.1)을 적용하면

$$\{A(T) = \int_0^T r(X(u))du,\ T\geq 0\}$$ 은 SICV(sp)이고 ,

$$\{B(T) = \int_0^T c(X(u) du,\ T\geq 0\}$$ 은 SICX((sp)임을 알 수 있다.

따라서 $E[A(T)]$는 증가 콘케이브 함수이고, $E[B(T)]$는 증가 콘벡스 함수이며 따라서 전체 $E[R(T)]$는 콘케이브 함수이다. ■

$E[R(T)]$를 최대화하는 문제는 콘벡스함수 $E[R(T)]$의 최대점을 추적함으로써 비교적 간단히 구할 수 있다. $d(x)=r(x) - c(x)$, $x\in S$ 로 놓고 식 (2)를 다르게 쓰면

$$E[R(T)] + D = E \int_0^T d(X(u))du$$

$$= \int_0^T E[d(X(u)]du$$

$$= \int_0^T \sum d(x)P\{X(u)=x\}du$$

따라서 $h(T)=E[R(T)] + D$, $\mathbf{p}$를 마코프체인의 초기 분포(즉 $p(x) = P\{X(0)=x\}$), $P(t)$를 시간 t에서의 전이함수행렬 (즉 $P(t)_{ij} = P\{X(t)=j \mid X(0)=i\}$)라 하면, 벡터 표기법을 사용하여

$$h(T) = \int_0^T \mathbf{p}^T P(u)\, \mathbf{d}\, du \tag{3}$$

을 얻고 이를 미분하여

$$h'(T) = \mathbf{p}^T P(T)\, \mathbf{r} \tag{4}$$

실제로 (4)의 P(T)를 계산하는 것이 힘이 들므로 Yoon & Shanthikumar(1989)의 결과를 이용하여

$$h'(T) \cong v(T) \equiv \mathbf{p}^T \underline{P}^k\, \mathbf{r}, \qquad k=[\lambda T] \tag{5}$$

와 같이 근사 계산을 할 수 있다. 여기서 $\underline{P}$ 는 Yoon & Shanthikumar(1989)에 제시된 방법에 따라 적절히 취하면 되는 데 여기서는 주어진 마코프 체인 유한한 상태 공간을 가져 당연히 단일화가 가능하게 되므로 계산상 가장 효율적인 내부 단일화(internal uniformization) 기법을 사용하여

$$\underline{P} = I + (1/\lambda)\, R \tag{6}$$

로 취한다. 이때 I는 단위행렬, R은 마코프체인의 generator이고 $\lambda \geq \sup_i\{-R_{ii}\}$ 를 만족하여야 한다.

이제 (5)를 이용하여 v=0 이 되는 점을 추적하면 (정리 4.1)에 의해 최대점이 근사적으로 구해 진다. 이 근사점은 λ를 충분히 크게 하면 매우 실제 최대점에 근접하게 되는 데, 효과적인 계산을 위해 λ를 2의 멱수로

잡아야 하고 λ가 2^{10} 이상이면 실용적으로 충분하다 (Yoon & Shanthikumar(1989) 참조). 이때 h'의 단조성에 근거하여 sequential search를 하되 건너 뛰는 시간 단위를 2단계로 조정하여 계산상의 효율성을 높일 수 있는 데 이를 구체적으로 나타내면 아래의 알고리듬과 같다.

<Algorithm Search1>

0. λ를 2의 멱수로 충분히 크게 잡는다(2^{10} 이상). 대체시간의 최대 한계(=U)를 설정하여 MAX= $\lfloor \log_2(\lambda U) \rfloor$ 로 놓는다.

1. i=0,1,2,...,MAX로 증가시키면서 $t=2^i/\lambda$ 에 대해 (4)를 이용하여 h'(t)를 계산하고 h'(t)가 최초로 0보다 작거나 같게되는 i를 구하여 UB라고 놓는다.

2. 만약 UB=0 이면 최적시점 $t^* = 2^{UB}/\lambda$를 내주고 멈춘다. 그렇지 않으면 3으로 간다.

3. $i=2^{UB-1}$, $2^{UB-1}+1$, ..., 2^{UB}로 증가시키면서 $t=2^i/\lambda$에 대해 (4)를 이용하여 h'(t)를 계산하고 h'(t)가 최초로 0보다 작거나 같게되는 i를 구하여 i*라고 놓는다. 최적시점 $t^* = i^*/\lambda$를 내주고 멈춘다.

(예)

{X(t),t≥0}를 상태공간 S={1,2,3,4,5}를 갖고, generator

$$R = \begin{matrix} -5 & 5 & 0 & 0 & 0 \\ 1 & -6 & 5 & 0 & 0 \\ 0 & 2 & -7 & 5 & 0 \\ 0 & 0 & 3 & -8 & 5 \\ 0 & 0 & 0 & 4 & -4 \end{matrix}$$

인 birth-death 과정이라고 하자. 초기 분포를 $p^T = (1\ 0\ 0\ 0\ 0)$, $r(x) = e^{-x}$, $c(x) = \log(x)$ 라고 할 때(D는 임의의 값):

k = 2^{10} 일때 최소점 t = 94 / 2^{10} = 0.091796875

k = 2^{12} 일때 최소점 t = 376 / 2^{12} = 0.091796875

$k = 2^{15}$ 일때 최소점 $t = 3007 / 2^{15} = 0.091766357$

여기서 k가 증가함에 따라 정확성이 증가하는 것을 관찰할 수 있다.
아래 (그림4.1)은 일때 t에 관한 $k=2^{10}$ 로 놓고 v(t)로 근사화한 h'(t)의
변화를 보여 주는 데 h'의 단조 감소성을 확인할 수 있다.

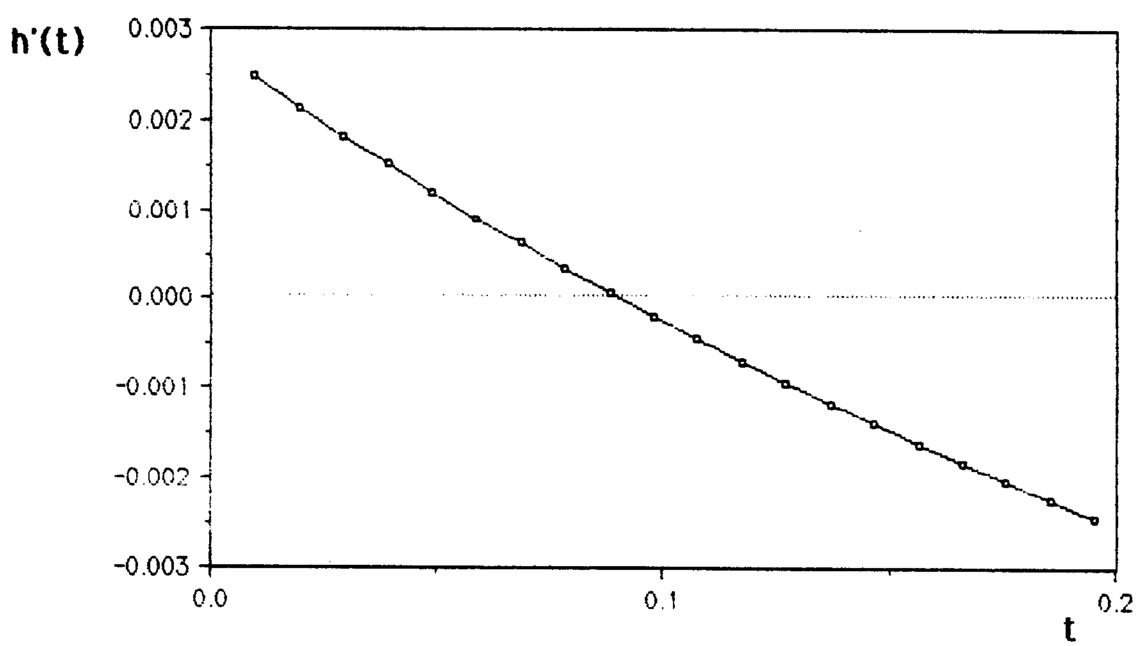

그림 4.1. h'(t) 의 변화

참고 문헌

[1] Barlow, R.E. and F.Proschan, **Statistical Theory of Reliability and Life Testing**, Holt,Rinehart and Wiston, N.Y., 1975.

[2] Cinlar, E.,"Markov Additive Processes: I and II", **Z. Wahrscheinlichkeitstheorie verw. Geb.**, v24, 85-121, 1972.

[3] Grassman,W., "The Convexity of the Mean Queue Size of the M/M/c Queue with respect to the Traffic Intensity," **J. Appl. Prob.**, v20, 916-919, 1983.

[4] Kamae,T.,U.Krengel and G.L.O'Brien, "Stochastic Inequalities on partially ordered spaces," Ann. Prob. v5, 899-912, 1977.

[5] Lee,H.L. and M.A.Cohen, "A Note on the Performance Measures of M/M/c Queueing Systems," **J. Appl. Prob.**, v20, 920-923, 1983.

[6] Ross.S.M., "Software Reliability:The Stopping Rule Problem," Tech. Rep., Dept. of IE&OR, University of California, Berkeley, 1987.

[7] Shaked M. and J.G.Shanthikumar, "Stochastic Convexity and Its Application," **Adv. Appl. Prob.**, v20, 427-446, 1988a.

[8] ________, "Temporal Stochastic Convexity and Concavity," **Stoch. Process. Appli.**, v27, 1-20, 1988b.

[9] ________, "Parametric Stochastic Convexity and Concavity of Stochastic Processes," Tech. Rep., School of Business Administration, University of California, Berkeley, 1988c.

[10] ________, "Convexity of a Set of Stochastically Ordered Random Variables," **Adv. Appli. Prob.**, v22, 160-177, 1990.

[11] Shanthikumar J.G. and Yao, "Second-Order Stochastic Properties in Queueing Systems," **Proceeding of IEEE**, v77, 162-170, 1989.

[12] Stoyan,S., **Comparison Methods for Queues and Other Stochastic Models**, Wiley, New York, 1983.

[13] Sumita,U and J.G.Shanthikumar, "A Software Reliability Model with Multiple-Error Introduction and Removal, **IEEE Trans. Reliab.**, v.R-35,459-462, 1986.

[14] Sumita,U., J.G.Shanthikumar and Y.Masuda, "Analysis of Fault Tolerant Computer Systems," **Microelectron. Reliab.**, v27, 65-78, 1987.

[15] Yoon,B.S., **Approximations for the Transient Behavior of Stochastic Processes: Discretization and Uniformization**, Ph.D. Dissertation, University of California, Berkeley, 1988.

[16] Yoon,B.S. and J.G.Shanthikumar, "Bounds and Approximations for the Trnasient Behavior of Continuous-time Markov Chains," **Probability in the Engineering and Informational Sciences**, v3, 175-198, 1989.