4.1. Stopping Times

- 아래를 가정하자.

$$\begin{aligned}&\Omega = \{(\omega_1, \omega_2, \dots,) : \omega_i \in S\}\\&\mathcal{F} = \mathcal{S} \times \mathcal{S} \times \dots\\&P = \mu \times \mu \times \dots \quad \mu \text{ is the distribution of } X_i.\\&X_n(\omega) = \omega_n\end{aligned}$$

- 맵핑 $\pi : \mathbb{N} \to \mathbb{N}$ 가 *finite permutation*이라는 것은 유한개의 i에 대해서면 $\pi(i) \neq i$가 성립한다는 의미이다. 따라서 퍼뮤테이션은 인덱스의 순서를 (유한개) 바꿔주는 함수라 볼 수 있다.

- 만약에 π가 $\mathbb{N}$의 *finite permutation*이고 $\omega \in S^{\mathbb{N}}$이라면 아래와 같이 정의할 수 있다.

$$i\text{-th element of } \pi\omega = \omega\text{의 } \pi(i)\text{-th element}$$

따라서

$$\big(\pi\omega\big)_i = \omega_{\pi(i)}$$

와 같이 쓸 수 있다. 이 경우 array of realization:

$$\omega_1, \omega_2, \dots$$

이 π에 의해서 재정렬된다고 해석할 수 있다. (Note that $\omega_1 = X_1(\omega)$)

- 따라서 array of realization 즉 어떠한 확률시행의 결과에 π를 취한다는 의미는 확률변수의 결과 중 유한개의 순서를 임의로 바꾼다는 것을 의미한다.

(def) 어떠한 사건 A가 *permutable*하다는 것은 사건 A가 아래를 식을 만족한다는 의미이다.

$$A = \pi^{-1}(A)$$

여기에서 $\pi^{-1}(A) = \{\omega : \pi\omega \in A\}$이다. 이 정의는 이해하기 그렇게 쉽지 않다. 아래의 설명들을 참고하여 보자.

- 우선 사건 A란 무엇인지 다시 생각해보자. 사건이란 확률시행의 결과를 재해석하여 구성한 어떠한 이벤트를 의미한다. 따라서 A는 예를들어 "주사위를 5번 던졌을때 짝수가 3번이상 나올 사건"과 같이 정의할 수 있다.

- 이제 퍼뮤터블의 의미를 살펴보자. 어떠한 사건 A 퍼뮤터블하다는 것은 확률변수의 결과 중 유한개의 순서를 임의로 바꾸어도 사건 A를 일관적으로 정의할 수 있다는 것을 의미한다.

example: 예를들어 보자. $X_1, X_2, \dots$이 1과 -1중 하나가 나오는 베르누이 시행이라고 하자. 만약

$$A : \text{랜덤변수들의 총합 즉 } X_1 + X_2 + X_3 + \dots \text{ 가 음수일 사건}$$

이라 정의한다고 하자. 직관적으로 A는 확률변수들의 순서를 바꾸어서 상관없으므로 이럴 경우 사건 A를 퍼뮤터블 하다고 말한다. 이 예제를 좀 더 수학적으로 표현하여 보자.

$$A = \{\omega : X_1(\omega) + X_2(\omega) + \dots < 0\}$$

π를 인덱스 1과 2를 바꾸는 변환이라고 하자. 그러면

$$\pi\omega = \pi(\omega_1, \omega_2, \dots) = (\omega_2, \omega_1, \dots)$$

따라서

$$\omega \in A \iff X_1(\omega) + X_2(\omega) + X_3(\omega) + \dots < 0$$

에 대응하는 것은

$$\pi\omega \in A \iff X_2(\omega) + X_1(\omega) + X_3(\omega) + \dots < 0$$

결국

$$\pi^{-1}(A) = \{\omega : \pi\omega \in A\} = \{\omega : \omega \in A\} = A.$$

example: 편의상 X_1, X_2, X_3을 베르누이시행에서 관측하였다고 하자.

$$A\text{: } X_1 + X_2 + X_3 = 1 \text{ 일 사건}$$

그러면

$$A = \{\omega : \boldsymbol{X}(\omega) = (1,0,0) \text{ or } \boldsymbol{X}(\omega) = (0,1,0) \text{ or } \boldsymbol{X}(\omega) = (0,0,1)\}$$

이다. 단, $\boldsymbol{X}(\omega) = \big(X_1(\omega), X_2(\omega), X_3(\omega)\big)$. 이때 A는 퍼뮤터블하다. 아래와 같은 collection을 고려하여 보자.

$$\mathcal{E} = \{\emptyset, A, A^c, \Omega\}$$

$\mathcal{E}$는 시그마필드이고 모든 원소가 퍼뮤터블하다.

(def) $\mathcal{E}$를 *exchangeable* 시그마필드라고 말한다. 즉 어떠한 시그마필드를 구성하는 모든 사건이 퍼뮤터블하면 익스체인지어블하다고 말한다.

example:

- 퍼뮤터블이벤트에 대한 예제를 좀더 살펴보자. (교재에 있는 예제이다.) 아래의 사건은 퍼뮤터블하다.

$$\{\omega : S_n(\omega) \in B,\ i.o.\}$$

위의 사건이 의미하는 것은 유한개의 n을 제외하고 모두

$$S_n \in B^c$$

라는 의미이다. 왜냐하면

$$\begin{aligned}&S_n \in B,\ i.o.\\&\iff S_n \in B^c,\ a.b.f.\\&\iff S_n \in B^c,\ \text{for all } n \geq N_0\end{aligned}$$

이기 때문이다.

- 이제 편의상 확률변수열 $\{X_n\}$에서 유한개의 인덱스를 서로 바꾸어 확률변수열 $\{\tilde{X}_n\}$을 만들었다고 가정하자. 그리고

$$S_n = X_1 + \dots + X_n = X_1(\omega) + \dots + X_n(\omega)$$

이라고 하고

$$\tilde{S}_n = \tilde{X}_1 + \dots + \tilde{X}_n = X_1(\pi\omega) + \dots + X_n(\pi\omega)$$

이라고 하자. 유한개의 n을 제외하면

$$S_n \equiv S_n(\omega) = S_n(\pi\omega) \equiv \tilde{S}_n$$

임을 주장할 수 있다. 즉

$$\exists N_0 \quad s.t. \quad \forall n \geq N_0 : \quad S_n = \tilde{S}_n$$

가 성립함을 보일 수 있다. 그도 그럴것이 π는 유한개의 인덱스를 서로 바꾸는 역할을 하므로 적당히 큰 수 N_0이상으로는 그 유한개의 바꿈이 없다고 보일수 있다. (이때 큰 수 N_0의 존재성은 "유한"개의 바꿈이라는 조건때문에 증명된다.) 예를들어 100을 1로 1을 100으로 바꾼다고 하자. 그러면 $N_0 = 100$이 존재하여

$$\forall n \geq 100 : \quad S_n = \tilde{S}_n$$

라고 주장할 수 있다.

- 이제

$$\exists N_0 \quad s.t. \quad \forall n \geq N_0 : \quad S_n = \tilde{S}_n$$

를 이용하면 아래를 얻을 수 있다. (풀이라기보다 표현들을 익숙하게 하기 위해 한번 쓴 것)

$$\begin{aligned}&\{S_n(\omega) \in B,\ i.o.\}\\&= \{S_n(\omega) \in B^c,\ a.b.f.\}\\&= \{\tilde{S}_n(\omega) \in B^c,\ a.b.f.\}\\&= \{S_n(\pi\omega) \in B^c,\ a.b.f.\}\\&= \{S_n(\pi\omega) \in B,\ i.o.\}\end{aligned}$$

따라서 사건 $A = \{\omega : S_n(\omega) \in B,\ i.o.\}$에 대하여

$$\omega \in A \iff \pi\omega \in A$$

이다.

note: $S_n(\omega) = S_n(\pi\omega)$은 그냥 모든 n에 대하여 성립하는것 아닌가? 하는 착각을 하지 않기를 바란다. 물론

$$S = X_1 + X_2 + X_3 + \dots$$

은 항상

$$S(\omega) = S(\pi\omega)$$

가 성립하지만 S_n은 그렇지 않다. 예를들어서 π를 1과 100의 인덱스를 서로 바꾸는 규칙이라고 하자.

$$\begin{aligned}&S_1(\omega) = X_1 \text{ and } S_1(\pi\omega) = X_{100}\\&S_2(\omega) = X_1 + X_2 \text{ and } S_2(\pi\omega) = X_{100} + X_2\\&\dots\end{aligned}$$

이므로 $n < 100$에 대하여서는 $S_n(\omega) \neq S_n(\pi\omega)$이다.

example: 아래의 사건도 퍼뮤터블하다.

$$\left\{\omega : \limsup_{n\to\infty} S_n(\omega)/c_n \leq 1\right\}$$

이유는 적당히 큰 n에 대하여 $S_n(\omega) = S_n(\pi\omega)$라고 주장할 수 있기 때문이다.

(thm) 테일-시그마필드에 속하는 모든 이벤트는 퍼뮤터블하다. 이때 *tail-σ-field*는 아래와 같이 정의되는 시그마필드 $\mathcal{T}$ 이다.

$$\mathcal{T} = \bigcap_{n=1}^{\infty} \mathcal{F}'_n \quad \text{where } \mathcal{F}'_n = \sigma(X_n, X_{n+1}, X_{n+2}, \dots)$$

note: 어떠한 이벤트 A가 $A \in \mathcal{T}$라는 것은 임의의 유한개의 realization의 결과를 **몰라도(=삭제하여도, 아무값이나 넣어도)** event A가 동일하게 정의될 수 있음을 의미한다.

- 위의 정리의 역은 성립하지 않는다. 즉 퍼뮤터블한 사건이 항상 테일시그마 필드의 원소는 아니다. 아까 소개한 바 있는 아래의 사건

$$\{\omega : S_n(\omega) \in B,\ i.o.\}$$

은 $\mathcal{E}$에 속하지만(=퍼뮤터블하지만) $\mathcal{T}$에 속하지 않는(=테일시그마필드의 원소는아닌) 사건이다. 이를 이해하기 위해 좀 더 구체적인 예로 생각해보자. 아래와 같은 구조에서 확률변수가 생성된다고 하자.

$$\begin{aligned}&X_1 \sim Ber(p)\\&X_2 = -X_1\\&X_n = 0 \text{ for } n \geq 3.\end{aligned}$$

관측가능한 확률변수열 $\{X_n\}$은 아래의 2경우 뿐이다.

$$-1, 1, 0, 0, 0, 0, \dots$$

$$1, -1, 0, 0, 0, 0, \dots$$

사건 A_n를 아래와 같이 정의하자.

$$A_n = \{S_n(\omega) \in B,\ i.o.\} \quad B = \{0\}$$

이라고 하자. 사건 A_n는 퍼뮤터블하다. 왜냐하면 확률변수열의 순서를 임의로 바꾸어도 적당히 큰 n에 대하여 (구체적으로는 $n \geq 2$에 대하여)

$$0 = S_n = \tilde{S}_n = 0$$

이 성립하기 때문이다. 즉 $X_1 + X_2 + \dots + X_n = 0$이라는 사실은 확률변수열의 순서를 아무리 바꾸어도 $n \geq 2$에서 항상 성립한다. 하지만 사건 A_n는 테일시그마필드의 원소가 아니다. 왜냐하면 A_n가 테일시그마필드의 원소라면 특정 값을 삭제해도 그 결과가 균일하게 정의되어야 하는데 첫번째 값 X_1을 삭제한다면

$$\begin{aligned}&S_n = X_2 + X_3 + \dots + X_n = X_2\\&\Longrightarrow P(A_n) = P(S_n \in B,\ i.o.) = P(X_2 = 0,\ i.o.) = 0\end{aligned}$$

X_1, X_2를 모두 삭제하면

$$\begin{aligned}&S_n = X_3 + X_4 + \dots + X_n = 0\\&\Longrightarrow P(A_n) = P(S_n \in B,\ i.o.) = P(S_n = 0,\ i.o.) = 1\end{aligned}$$

이 되어서 A_n을 균일하게 정의할 수 없다.

note: 물론 사건 A_n이 테일시그마필드의 원소라면 특정 확률변수를 삭제하는 것이 아니라 아예 $\mathbb{R}$의 임의의 값으로 바꾸어친다쳐도 항상 A_n이 균일하게 정의 할 수 있어야 한다. 처음부터 이렇게 생각했으면 A_n이 테일시그마필드의 원소가 아님을 더 쉽게 보일 수 있다. 세번째 값 X_3을 100으로 바꿔친다면

$$S_n = X_1 + X_2 + X_3 + \dots + X_n = 100.$$

세번째 값 X_3을 200으로 바꾸면

$$S_n = X_1 + X_2 + X_3 + \dots + X_n = 200.$$

따라서 A_n을 균일하게 정의할 수 없다.

- 참고로 위와 동일한 예제에서

$$A_n = \{S_n \in (-\infty, \infty)\}$$

와 같이 정의하면 사건은 퍼뮤터블하고 테일시그마필드의 원소가 된다.

- 만약에 확률변수열 $X_1, X_2, \dots$이 iid 라면 $\mathcal{E}$와 $\mathcal{T}$는 차이가 없다. 이게 바로 **휴이트-새비지**의 정리이다.

(Theorem 4.1.1. Hewitt-Savage 0-1 law.) (1) $X_1, X_2, \dots$ 이 i.i.d.이고 (2) $A \in \mathcal{E}$ 이라면

$$P(A) \in \{0, 1\}$$

이다.

(Theorem 4.1.2.) 랜덤워크 S_n은 아래의 4개의 가능성밖에 없다.

(i) $S_n = 0$ for all n.

(ii) $S_n \to \infty$.

(iii) $S_n \to -\infty$.

(iv) $-\infty = \liminf S_n < \limsup S_n = \infty$. 이중에서 (iv)의 경우는 S_n이 $-\infty$와 ∞사이를 끝없이 진동하는 경우이다. 즉 $(-2)^n$와 같은 경우이다.

(def) $\mathcal{F}_n = \sigma(X_1, \dots, X_n)$이라고 하자. 그리고 N을 $\{1, 2, \dots, \} \cup \{\infty\}$ 중 하나의 값을 가지는 확률변수라고 하자. 만약에 아래가 성립한다면 N을 *stopping time* 이라고 부른다.

$$\text{for all } n < \infty,\ \{N = n\} \in \mathcal{F}_n.$$

- 어떠한 확률변수 N이 stopping time인지 보려면 (1) N이 가질수 있는 값이 (무한대를 포함한) 자연수인지 체크하고 (2) $N = n$임을 판단할때 $X_1, X_2, \dots, X_n$ 까지의 정보만 있으면 되는지 체크하면 된다.

example: 아래와 같이 정의되는 확률변수 N은

$$N = \inf_n \{S_n \in A\}$$

stopping time이다. 왜냐하면 $N = n$ 인지 판단하기 위해서는 $X_1, \dots, X_n$ 까지만 관측하면 충분하기 때문이다. 이런식으로 정의되는 N은 특별히 *hitting time of A* 라고 부른다.

- 각각의 stopping time N에 대하여 $\mathcal{F}_N$을 아래와 같이 정의하자.

$$\mathcal{F}_N = \{A : A \cap \{N = n\} \in \mathcal{F}_n \text{ for all } n < \infty\}$$

$N = n$이 고정되었을 경우 n시점의 정보 $X_1, \dots, X_n$을 가지고 명확하게 정의할 수 있는 사건 A들의 집합을 $\mathcal{F}_N$이라고 한다.

example: N이 stopping time이라고 할때 $\mathcal{F}_N$의 가능한 한 예는 아래와 같다.

$$\mathcal{F}_N = \{N \leq n\} = \{N = 1\} \cup \{N = 2\} \cup \dots \cup \{N = n\}$$

이것을 따지는 방법은 아래와 같다.

(1) $n = 100$ 이라고 하자.

(2) $\mathcal{F}_N$의 임의의 원소 예를들어 사건 $\{N = 34\}$는 $X_1, \dots, X_{100}$으로 명확하게 알 수 있다.

(3) 따라서 $\{N = n\} = \mathcal{F}_N$이라고 볼 수 있다.

(Theorem 4.1.3.) (1) $X_1, X_2, \dots,$는 $i.i.d.$ 이고 (2) N이 $P(N < \infty) > 0$ 인 stopping time 이라면 아래가 성립한다.

6.1. Definitions

(def) $(S, \mathcal{S})$를 measurable space라고 하자. $X_n : (\Omega, \mathcal{F}) \to (S, \mathcal{S})$이라고 하자. 편하게

$$(S, \mathcal{S}) = (\mathbb{R}, \mathcal{R})$$

이라고 생각해도 무방하다. 어떠한 확률변수열 $\{X_n\}$이 filtration $\mathcal{F}_n := \sigma(X_0, \dots, X_n)$에서 정의되어 있다고 하자. 확률변수열 $\{X_n\}$이 아래의 조건을 만족할때 $\{X_n\}$을 $\mathcal{F}_n$에 대한 *Markov-chain*이라고 부른다.

$$\text{for all } B \in \mathcal{S}: \quad P(X_{n+1} \in B | \mathcal{F}_n) = P(X_{n+1} \in B | X_n)$$

note: 확률변수열 $X_1, X_2, \dots$의 값이 바로 이전의 값에 의해서만 결정되면 마코프체인이라고 한다. 즉 X_2의 값을 알기 위해서는 X_1의 값에 대한 정보만 있으면 되고 X_3의 값을 알기 위해서는 X_2에 대한 정보만 있으면 될때 $X_1, X_2, \dots$을 마코프체인이라고 한다.

note: $V = \{1, 2, 3, \dots, 100\}$이라고 하자. $\{X_n\}$을 trace of snow라고 하자. 그러면 $\{X_n\}$은 마코프체인이다.

- 4×4 그리드 세계를 가정하자.

$$\Omega = \{(1,1), \dots, (4,4)\}$$

이고

$$S = \{1, 2, 3 \dots, 16\}$$

이라고 하자. 확률변수 $X_1 : (\Omega, \mathcal{F}) \to (S, \mathcal{S})$은 아래와 같이 정의할 수 있는 맵핑이라고 하자.

$$\begin{aligned}&X_1\big((1,1)\big) = 1\\&X_2\big((1,2)\big) = 2\\&\dots\\&X_1\big((4,4)\big) = 16\end{aligned}$$

따라서

$$X_1 = 1, X_2 = 2$$

가 의미하는 것은 처음에는 $(1,1)$의 위치에 있다가 그다음에는 $(1,2)$의 위치로 이동하였다는 것을 의미한다. 이제 $(1,1)$의 위치에서 $(1,2)$의 위치로 이동하는 transition probability를 p라고 정의하자.

$$p : (S, \mathcal{S}) \to \mathbb{R}$$

기호로는 아래와 같이 쓴다.

$$p(x, A)$$

여기에서 $x \in S$, $A \in \mathcal{S}$이다.

example: $x = 1$, $A = \{1, 2, 5\}$이라고 하자.

$$\begin{aligned}&x = 1 \Leftrightarrow X(\omega) = 1 \Leftrightarrow \omega = (1,1)\\&A = \{1, 2, 5\} \Leftrightarrow \{\omega : X(\omega) \in A\} = \{(1,1), (1,2), (2,1)\}\end{aligned}$$

임을 주목하라. 따라서

$$p(x, A)$$

는 점 $(1,1)$에서 출발했는데 점 $(1,1), (1,2), (2,1)$중 하나에 도착할 확률이므로

$$p(x, A) = 1$$

이라고 볼 수 있다.

(def) transition probability의 정의를 사용하면 아래를 만족하는 확률변수열 $\{X_n\}$을 마코프체인이라 정의할 수 있다.

$$P(X_{n+1} \in B|\mathcal{F}_n) = p_n(X_n, B)$$

여기에서 p_n은 n번째에 어떠한 위치 X_n에서 B의 부분집합중 하나의 위치로 이동할 확률을 의미한다.

(결론1,2,3의 가정) 만약에 ($C1$) $(S, \mathcal{S})$이 *nice space* 이고 ($C2$) $\{p_n\}$이 잘 정의되며 ($C3$) $(S, \mathcal{S})$에서의 initional distribution μ가 잘 정의된다고 하자.

(결론1) 일단 유한개의 확률변수열 $\{X_n\}$에 대하여 consistence set of finite dimensional distribution을 아래와 같이 잘 정의할 수 있다.

$$\begin{aligned}&P(X_1 \in B_1, X_2 \in B_2, \dots, X_n \in B_n)\\&= \int_{B_0} \mu(dx_0) \int_{B_1} p_0\big(x_0, \mu(dx_1)\big) \cdots \int_{B_n} p_{n-1}\big(x_{n-1}, \mu(dx_n)\big)\end{aligned}$$

note: 이때 확률측도 P는 아래의 공간 $(S_0 \times S_1 \cdots \times S_n, \mathcal{S}_0 \times \mathcal{S}_1 \cdots \times \mathcal{S}_n)$에서 정의되는 확률측도이다. 이 공간은 간단하게 $\big(S^{\{0,1,\dots,n\}}, \mathcal{S}^{\{0,1,\dots,n\}}\big)$와 같이 표현하기도 한다. 쓸대없이 말이 거창한데 P는 쉽게 말해서 아래와 같은 것을 계산할 수 있게 해주는 확률측도라 보면 된다.

$$P(X_1 = 1, X_2 = 2, X_3 = 3, X_4 = 2, X_5 = 3, \dots, X_n = 2)$$

note: 즉 결론1은 확률측도 P는 초기분포 μ와 p_n만 잘 정의하면 모순없이 정의가능하다는 것을 의미한다. 이런의미에서 보면 P는 초기분포 μ에 의해서 유도되었다고 생각할 수도 있는데 이러한 이유로 P_μ라고 쓰기도 한다.

- 아래의 정리는 확률변수열 $\{X_n\}$이 무한일 경우에도 P_μ를 잘 정의할수 있음을 보여준다.

(결론2) $(C1)$-$(C3)$의 가정하에 Kolmogorov's theorem은 확률변수열 $\{X_n\}$이 무한수열을 가질때

$$P_\mu(X_1 \in B_1, X_2 \in B_2, \dots,)$$

가 모순없이 잘 정의됨을 보여준다. 이때 P_μ는 $\big(S^{\{0,1,\dots\}}, \mathcal{S}^{\{0,1,\dots,\}}\big)$에서의 확률측도이다. 쉽게 말해서 P_μ는

$$P_\mu(X_1 = 1, X_2 = 2, X_3 = 1, \dots)$$

따위의 확률을 구할 수 있는 메져라고 생각하면 된다.

note: 아래의 기호는 외우는 것이 좋겠다.

$$\begin{aligned}P_\mu(X_0 \in B_0) &= \int_{B_0} \mu(dx_0)\\P_\mu(X_0 \in B_0, X_1 \in B_1) &= \int_{B_0} \mu(d(x_0)) \int_{B_1} p_0\big(x_0, \mu(dx_1)\big)\end{aligned}$$

(notation of P_x) $\mu = \delta_x$라고 하자. 이는 랜덤워크의 초기시작값이 항상 x라는 것을 의미한다. 그리고 기호 $P_x = P_{\delta_x}$라고 정의하자. 즉 P_x는 x에서 시작해 $\{p_n\}$의 전이확률을 가지는 랜덤워크가 가지는 분포를 의미한다.

- 일반적인 초기분포 μ에서 유도된 확률메져 P_μ는 x에서 시작할 확률 $\mu(dx)$와 x를 시작점으로 했을때 유도되는 확률메져 P_x로 표현할 수 있다.

$$P_\mu(A) = \int \mu(dx) P_x(A), \quad A \in \mathcal{S}^{\{0,1,\dots,\}}$$

(Thm 6.1.1, 결론3.) $(C1)$-$(C3)$의 조건하에 $\{X_n\}$이 마코프체인이 된다.

(Thm 6.1.2, 결론1의 변형) $(C1)$-$(C3)$의 조건중 체크하기 까다로운 것은 (1)이다. 오히려 (1)의 조건대신에 $\{X_n\}$이 마코프체인임을 가정하면 결론2와 동일한 결과를 얻을 수 있다. 즉 (1) $\{X_n\}$이 마코프체인이고 (2) transition prob $\{p_n\}$이 주어졌고 (3) initional distribution μ 가 주어졌다면 *finite dimensional distribution*이 아래와 같이 주어진다.

$$\begin{aligned}&P_\mu(X_j \in B_j, 0 \le j \le n)\\&= P_\mu(X_1 \in B_1, X_2 \in B_2, \dots, X_n \in B_n)\\&= \int_{B_0} \mu(dx_0) \int_{B_1} p_0\big(x_0, \mu(dx_1)\big) \cdots \int_{B_n} p_{n-1}\big(x_{n-1}, \mu(dx_n)\big)\end{aligned}$$

6.2. Examples

6.3. Extensions of the Markov Property

(Theorem 6.3.1. The Markov property.) 함수 $Y : \Omega_0 \to \mathbb{R}$이 (1) bounded and (2) measurable하다고 하자. 그러면 아래가 성립한다.

$$E_\mu\left(Y \circ \theta_n | \mathcal{F}_n\right) = E_{X_n} Y$$

이 정리를 쉽게 설명하면 아래와 같다.

(a) $(C1)$-$(C3)$를 가정하자. 이 가정하에서 무한확률변수열 $\{X_n\}$는 마코프체인이 된다. (Thm 6.1.1.)

(b) Y는 무한확률변수열 $X_1, X_2, \dots$으로 만든 어떠한 통계량이다. 예를들면 $Y = S_n$이라든가, $Y = \max_n X_n$과 같은 것이 있을 수 있다. 단 함수 Y는 (1) bounded and (2) measurable 해야 한다.

(c) $Y \circ \theta_n$은 무한확률변수열 $X_n, X_{n+1}, \dots$으로 만든 어떠한 통계량이다. 예를들어

$$Y = mod\{X_1, \dots, X_\tau\}$$

이었다면

$$Y \circ \theta_{100} = mod\{X_{\tau+1}, \dots, X_{\tau+100}\}$$

이 된다. (물론 mod함수가 (1) bound and (2) measurable 한지 체크해야함.)

(d) 우선 $\mathcal{F}_{100}$은 알고 있다고 가정하자. 즉 $X_1, \dots, X_{100}$의 실현값은 알고 있다. 편의상 X_{100}의 실현값을 x_{100}이라고 하자. 이제 아래식이 의미하는 바를 살펴보자.

$$E_\mu\left(Y \circ \theta_n | \mathcal{F}_n\right) = E_{X_n} Y$$

이는 $X_{101}, X_{102}, X_{103}, \dots$의 최빈값의 평균을 알기 위해서 아래와 같은 변환을 수행해 새로운 확률변수열 $Z_1, Z_2, \dots$를 만들고

$$Z_1 = X_{101}, Z_2 = X_{102}, \dots$$

$\mu = \delta_{x_{100}}$을 대입하여 이로부터 유도된 새로운 확률메져 $P_{x_{100}}$를 만들어 $Z_1, \dots, Z_\tau$의 최빈값의 평균을 계산하면 된다는 의미이다.

6.4. Recurrence and Transience

6.5. Stationary Measures

- 아래식을 만족하는 measure μ를 stationary measure라고 한다.

$$\sum_x \mu(x) p(x, y) = \mu(y)$$

note: $p(x, y)$: 노드 x에서 다음 노드 y로 이동할 확률

note: $\mu(x)$: 노드 x에 있을 확률

note: 따라서 stationary measure는 특정노드에 있을 확률을 측정하는 메져라 생각할 수 있다.

- stationary measure(=stationary distribution)가 (1) 존재하고 (2) 유일하다는 것이 조사되었다고 하자. 이제 다음 관심사는 아래식을 만족하는 staionary distribution π이다.

$$\pi p = \pi$$

(정리 6.5.6.)

(정리 6.5.6.) p가 irreducible 하다는 것과 아래는 동치이다.

(1) .

(2) stationary distribution이 존재한다.

(3) .

Asymptotic Behavior

(레마 6.6.3.) $d_x = 1$이라면 m_0보다 큰 모든 m에 대하여

$$p^m(x, x) > 0$$

를 만족시킬 수 있다.

(정리 6.6.4.) p가 (1) irreducible 하고 (2) aperiodic 하며 (3) stationary distribution π를 가진다고 하자. 그러면 아래가 성립한다.

$$p^n(x, y) \to \pi(y) \quad as\ n \to \infty.$$

note: p가 irreducible 인것만 보이면 stationary distribution π를 가진다는 것은 정리 6.5.6에 의해서 성립한다. 따라서 (1)-(2)만 조건으로 사용해도 위의 정리는 성립한다.

note: p가 에이피리오딕하다는 의미는 모든 state가 $d_x = 1$을 가진다는 것을 의미한다 .

(pf)

- $S^2 = S \times S$ 라고 하자.

- 전이확률 $\bar{p}$를 $S \times S$에서 아래와 같이 정의하자.

$$\bar{p}\big((x_1, y_1), (x_2, y_2)\big) = p(x_1, x_2) p(y_1, y_2)$$

note: 이는 각각의 coordinate가 독립적으로 움직인다는 것을 의미한다.

(step 1)

- 먼저 $\bar{p}$가 이리듀시블임을 보이자. (이는 너무 당연해서 바보같은 증명으로 보이지만 정리의 에이피리오딕조건을 사용하는 유일한 과정이다.) 우선 p가 이리듀시블하다는 조건으로부터 아래를 만족하는 적당한 K, L이 존재함을 알 수 있다.

$$p^K(x_1, x_2) > 0 \quad \text{and} \quad p^L(y_1, y_2) > 0.$$

그런데 레마 6.6.3에 의해서 M을 적당히 크게 설정한다면 아래를 만족시킬수 있음을 알 수 있다.

$$p^{L+M}(x_1, x_2) > 0 \quad \text{and} \quad p^{K+M}(y_1, y_2) > 0.$$

따라서 아래가 성립한다.

$$\bar{p}^{K+L+M}\big((x_1, y_1), (x_2, y_2)\big) > 0$$

(step 2)

- 두 코디네이츠가 독립이므로 $\bar{p}$의 stationary distribution을 아래와 같이 정의할 수 있다.

$$\bar{\pi}(a, b) = \pi(a)\pi(b).$$

- 정리 6.5.4에 의해서 $\bar{p}$의 stationary distribution이 존재한다는 것은 $\bar{p}$의 모든상태가 recurrent하다는 것을 의미한다.

- (X_n, Y_n)을 $S \times S$에서의 체인이라고 하자.

- T를 이 체인이 처음으로 대각 $\{(y, y) \in S\}$을 치는 시간이라고 하자.

- $T(x, x)$를 (x, x)를 hit하는 시간이라고하자.

- $\bar{p}$가 (1) irreducible 하고 (2) recurrent 하므로 $T(x, x) < \infty\ a.s.$ 이고 따라서 $T < \infty\ a.s.$ 이다.

(step 3)

- 우선 두개의 코디네이트 (X_n, Y_n)가 $\{T \le n\}$에서 같은 분포를 가진다는 것을 관찰하자.

- (X_n, Y_n)이 첫 교차점을 가지는 시간과 장소를 고려하여보자. 마코프 성질을 이용하면

$$P(X_n = y, T \le n) = \sum_{m=1}^{n} \sum_x P(T = m, X_m = x, X_n = y)$$

6.6. Periodicity, Tail σ-field

6.7. General State