44.006 — University of Seoul, Spring 2026

Lecture 4: Probability Measure on a Metric Space

June, 30, 2025

Lecturer: *Jong-June Jeon* *Scribe*: *Jong-June Jeon*

1 무한 동전 던지기 실험과 유한차원 극한분포

확률과정의 가장 기본적인 예는 동전던지기로부터 만들어지는 단순 랜덤워크이다. 이 절에서는 동전던지기 실험에서 출발하여, 적절히 rescaling된 랜덤워크가 극한에서 다변량 정규분포 구조를 갖는다는 사실을 설명한다. 이에 대한 논의는 이 랜덤워크가 랜덤한 연속함수로 표현할 수 있으며, 연속함수의 모임을 표본집합으로 가지는 확률공간을 생각하기 위함이다. 확률변수가 원래의 확률공간 $(\Omega, \mathcal{F}, \mathbb{P})$위에서 논의를 $(\mathbb{R}, \mathcal{B}(\mathbb{R}), P)$로 옮겨주었다면 지금 우리가 하고자 하는 바는 원래의 확률공간 $(\Omega, \mathcal{F}, \mathbb{P})$을 샘플경로를 표본공간으로 가지는 어떤 확률공간을 어떤방식으로 자연스럽게 생각할 수 있을까에 대한 것이다. 확률변수는 하나의 표본 ω를 하나의 실수로 변환했다면, 이번 장에서는 하나의 표본 ω가 하나의 연속함수로 변환되는 과정 위에서 확률공간을 살펴볼 것이다.

먼저 연속함수로 변환될 preimage 인 ω를 정의하기 위해 공정한 동전을 무한히 던지는 실험을 생각한다. 각 시행의 결과를 $\omega_i \in \{0, 1\},\ i = 1, 2, \dots$ 로 나타내면 하나의 실험결과(experiment outcome)는 무한열 $\omega = (\omega_1, \omega_2, \omega_3, \dots)$이다. 무한 시행을 확률변수열을 통해 정의한다:

$$X_i = \begin{cases} +1, & \omega_i = 1, \\ -1, & \omega_i = 0. \end{cases}$$

그러면 $\mathbb{P}(X_i = 1) = \mathbb{P}(X_i = -1) = \frac{1}{2}$, 또한 $\{X_i\}_{i \geq 1}$는 i.i.d.이며 $\mathbf{E}[X_i] = 0,\ \mathrm{Var}(X_i) = 1.$

다음 누적합을 통해 단순 랜덤워크를 정의한다. $S_n = X_1 + \cdots + X_n,\ n \geq 1$. 즉 S_n은 n번 동전을 던진 뒤의 위치이며, 매 시행마다 ± 1만큼 이동한다.

중심극한정리에 의해 S_n의 크기는 대략 $\sqrt{n}$ 정도로 커진다 ($S_n = O_p(\sqrt{n})$). 따라서 극한을 관찰하기 위해 다음과 같이 공간을 축소한다. $t \in [0, 1]$에 대해 $W_n\big(\frac{k}{n}\big) := \frac{1}{\sqrt{n}} S_k,\ k = 0, 1, \dots, n$ 로 정의하고, 각 구간 $\big[\frac{k}{n}, \frac{k+1}{n}\big]$에서는 선형보간하여 연속과정으로 확장한다. 즉,

$$W_n(t) := \frac{1}{\sqrt{n}} \Big(S_k + (nt - k) X_{k+1} \Big), \qquad t \in \Big[\frac{k}{n}, \frac{k+1}{n} \Big].$$

따라서 $W_n \in C[0,1]$ (여기서 $C[0,1]$은 $[0,1]$위에 정의된 연속함수의 집합)이며, 격자점에서는 원래 랜덤워크와 일치한다:

$$W_n\left(\frac{k}{n}\right) = \frac{1}{\sqrt{n}} S_k.$$

위 정의에서 보간항을 $R_n(t) := \frac{1}{\sqrt{n}}(nt-k)X_{k+1}$ 라 두면, $0 \le nt - k < 1$이고 $|X_{k+1}| = 1$이므로 $|R_n(t)| \le \frac{1}{\sqrt{n}}$. 따라서 $\sup_{t\in[0,1]} |R_n(t)| \le \frac{1}{\sqrt{n}} \to 0$. 특히 임의의 고정된 시점 $t_1, \dots, t_m$에 대해

$$\max_{1\le j\le m} |R_n(t_j)| \le \sup_{t\in[0,1]} |R_n(t)| \to 0.$$

따라서

$$\big(W_n(t_1), \dots, W_n(t_m)\big) - \Big(\frac{1}{\sqrt{n}} S_{\lfloor nt_1 \rfloor}, \dots, \frac{1}{\sqrt{n}} S_{\lfloor nt_m \rfloor}\Big) \xrightarrow{\mathbb{P}} 0.$$

Slutsky 정리에 의해 finite-dimensional distribution의 극한을 구할 때는 보간된 과정 $W_n(t)$ 대신 $\frac{1}{\sqrt{n}} S_{\lfloor nt \rfloor}$ 를 사용해도 무방하다. ($X_n \xrightarrow{d} X$ 그리고 $Y_n \xrightarrow{p} 0$ 면 $X_n + Y_n \xrightarrow{d} X$) 즉 finite-dimensional limit을 구할 때는 $W_n(t) \approx \frac{1}{\sqrt{n}} S_{\lfloor nt \rfloor}$ 로 생각할 수 있다.
이제 임의의 시점 $0 \le t_1 < \cdots < t_m \le 1$ 을 고정하고 벡터 $\big(W_n(t_1), \dots, W_n(t_m)\big)$ 의 분포 수렴을 고려한다. 각 t_j에 대해 $k_j = \lfloor nt_j \rfloor$라 하면 $W_n(t_j) \approx \frac{1}{\sqrt{n}} S_{k_j} = \frac{1}{\sqrt{n}} \sum_{i=1}^{k_j} X_i$. 두 시점 t_p, t_q에 대해 $\mathrm{Cov}(W_n(t_p), W_n(t_q)) = \frac{1}{n}\mathrm{Cov}(S_{k_p}, S_{k_q})$이다. 그런데 $S_{k_p} = \sum_{i=1}^{k_p} X_i, \qquad S_{k_q} = \sum_{j=1}^{k_q} X_j,$ 이므로 $\mathrm{Cov}(S_{k_p}, S_{k_q}) = \sum_{i=1}^{k_p}\sum_{j=1}^{k_q} \mathrm{Cov}(X_i, X_j)$. 독립성에 의해 $\mathrm{Cov}(X_i, X_j) = I(i=j)$ 이므로 $\mathrm{Cov}(S_{k_p}, S_{k_q}) = \min(k_p, k_q)$ 이며 결국

$$\mathrm{Cov}(W_n(t_p), W_n(t_q)) = \frac{1}{n}\min(k_p, k_q) \to \min(t_p, t_q).$$

Cramér–Wold 정리에 의해, 임의의 $a \in \mathbb{R}^m$에 대해 $\sum_{j=1}^m a_j W_n(t_j) \Rightarrow N\Big(0, \sum_{p,q} a_p a_q \min(t_p, t_q)\Big)$. 따라서

$$\big(W_n(t_1), \dots, W_n(t_m)\big) \Rightarrow \mathcal{N}(0, \Sigma)$$

이며, 여기서 극한 공분산 행렬은 $(\Sigma)_{pq} = \min(t_p, t_q)$로 주어진다.

2 Metric 공간과 시그마대수

앞 절에서 우리는 scaled random walk $\{W_n(t)\}_{0\le t\le 1}$의 finite-dimensional distributions가 다음과 같이 수렴함을 보였다. 임의의 시점 $0 \le t_1 < \cdots < t_m \le 1$에 대해

$$\big(W_n(t_1), \dots, W_n(t_m)\big) \Rightarrow \mathcal{N}(0, \Sigma), \qquad (\Sigma)_{pq} = \min(t_p, t_q)$$

가 성립함을 확인하였다. $\big(W_n(t_1), \ldots, W_n(t_m)\big)$은 process $\{W_n(t) : t \in [0,1]\}$ 의 고정된 시간 $(t_1, \cdots, t_m)$에서 평가된 Finite-dimensional distribution 이다. 무한열 $\omega = (\omega_1, \omega_2, \omega_3, \ldots)$ 에 대해서 $W_n(\cdot)(\omega)$는 하나의 연속경로(연속함수)를 만들고, 연속함수인 $W_n(\cdot)$는 그것을 구성할때 사용한 선형보간 구조로 인해 $t = k/n$, $k = 0, 1, \cdots, n$ 위에서 $W_n(t)$ 값으로 완전히 결정된다. 사실 우리는 무한동전 던지기 실험의 모형의 가정에 의해 $[0,1]$ 위에서 랜덤한 함수 $W_n(\cdot)$의 경로가 어떤 집합에 속할 확률을 구할 수 있다. 만약 연속된 경로의 모임 S를 생각하고 경로들의 부분집합 A을 생각한다면 랜덤한 경로가 해당 부분집합에 포함된 확률을 아래와 같이 쓸 수 있을 것이다.

$$P\{\omega : W_n(\cdot)(\omega) \in A\}$$

여기서 우리는 사건집합 A를 어떻게 잡아야 할 것인가에 대한 문제에 당면 하게 된다. 연속 함수의 모임에서 함수들의 부분집합을 잡는 널리 사용 되는 방법 중 하나가 함수 공간에 매트릭(metric)을 도입하고 그것으로부터 유도 된 열린 공을 잡아 열린 공을 포함하는 가장 작은 시그마 대수를 생각하는 것이다. 앞으로 자연스럽게 이어지는 질문은 **어떤 매트릭으로 열린 공을 정의할 것인가에 대한 것이다.** $W_n(\cdot)$에 대한 확률분포를 P_n라 놓고 우리는 P_n이 수렴하는 분포를 생각할 수 있는지에 대해 더 알아보고자 한다. 이를 위해서는 필요조건이 있는데 Prohorov 정리에 따르면 $P_n \Rightarrow P$가 되기 위해서는 $\{P_n\}$ 이 uniformly tight 해야한다. 임의의 $\varepsilon > 0$에 대하여 어떤 compact 집합 $K_\varepsilon \subset S$가 존재하여

$$\inf_n P_n(K_\varepsilon) \geq 1 - \varepsilon$$

여기서 너무 복잡한 위상구조 (열린집합들의 모임)를 만들게 되면 이러한 K_ϵ compact set 이 존재하지 않을 수도 있다. 특히 셀 수 있는 원소로 어떠한 전체집합내 원소를 근사할 수 있다는 separability 와 complete 성질은 compact set을 구성하는데 중요한 조건으로 이 성질은 metric에 의존한다. 그래서 아무 metric 이나 가져와서 열린공을 만들면 분포수렴을 정의하지 못할 수 있다. 그렇다고 너무 단순한 metric 을 가져오게 되면 함수의 근사라는 의미를 잃어버릴 수 있다. 닫힌 구간에서 정의된 연속함수(예: $C[0,1]$)의 집합을 표본공간으로 사용하는 경우 uniform metric 이 사용하면 이 위상에서 표본공간은 separable 이 알려져 있어 이 metric은 널리 사용 된다. scaled random walk 도 닫힌구간에서 정의된 연속함수이므로 이 metric을 사용한다. 아래의 두 가지 metric을 도입했을때 연속함수 공간에서

Example 2.1. (*1*) *Topology*가 상대적으로 작은 경우
예를 들어 다음 *metric*을 생각하자:

$$d(f,g) = \sum_{m=1}^{\infty} 2^{-m} \min\{1,\ |f(t_m) - g(t_m)|\},$$

여기서 $\{t_m\}_{m \geq 1}$는 $[0,1]$의 조밀한 가산 부분집합 (예: $t_m \in \mathbb{Q} \cap [0,1]$)이다. 함수의 근사에 대한 의미가

uniform metric 을 사용했을때보다 정밀하지 못하다.

Example 2.2. (*2*) *Topology*가 너무 큰 경우

$$d(f,g) = \begin{cases} 0, & f = g, \\ 1, & f \neq g. \end{cases}$$

이 *metric* 으로는 $C[0,1]$이 *non-separable* 이라 분포수렴을 정의하는데 적합하지 않다.

여기서 첫번째 우리의 결론은 $W_n(\cdot)$의 분포수렴을 논하기 전에 선결조건으로 연속함수 공간의 metric을 uniform을 사용하여 그 공간을 separable하고 complete하게 만들 수 있는 열린공을 생각하자는 것이다. 이 열린 공을 통해 우리는 사건이라는 것을 생각할 수 있다. 이어지는 질문은 다음과 같다.

- 앞서 우리가 확인한 사실은 임의의 시점 $0 \leq t_1 < \cdots < t_m \leq 1$에 대해

  $$\big(W_n(t_1), \ldots, W_n(t_m)\big) \Rightarrow \mathcal{N}(0, \Sigma), \qquad (\Sigma)_{pq} = \min(t_p, t_q)$$

  가 성립함이었다. 이러한 Finite-dimensional convergence만으로 과정 전체의 극한, 즉 경로공간 위의 확률측도를 정의할 수 있는가?

- 성질이 좋은 metric 을 도입하여 연속함수로 구성된 표본공간을 seperable 하고 complete 하게 만들었다. 이것으로 $W_n(\cdot)$ 혹은 일반적인 연속랜덤경로(연속함수)의 uniformly tightness 를 보장할 수 있는가?

3 Wiener Process

위 질문에 답하기 위해서는 먼저 극한으로 기대되는 확률과정을 정의할 필요가 있다. 유한차원 분포의 형태로부터 자연스럽게 등장하는 과정은 Wiener process이다.

Definition 3.1. Wiener process 확률과정 $\{W(t)\}_{t\in[0,1]}$를 다음 성질을 만족하는 경우 Wiener process 라 한다.

1. $W(0) = 0$ almost surely.
2. $W(t) - W(s) \sim \mathcal{N}(0, t-s)$ for $0 \leq s < t \leq 1$.
3. 서로 다른 구간의 증분은 독립이다.

4. 거의 모든 ω에 대해 $t \mapsto W(t)(\omega)$는 연속함수이다.

위 정의에 의해 Wiener process는 $W(\cdot) : \Omega \to C[0,1]$ 로 생각할 수 있으며, 이는 연속함수 공간 위의 확률변수이다. 특히 임의의 시점 $0 \le t_1 < \cdots < t_m \le 1$에 대해 $\big(W(t_1), \ldots, W(t_m)\big) \sim \mathcal{N}(0, \Sigma), \quad (\Sigma)_{pq} = \min(t_p, t_q)$, 를 만족한다. 따라서 앞서 확인한 $W_n(\cdot)$의 finite-dimensional convergence는 Wiener process의 유한차원 분포와 일치한다. 앞에서는 샘플경로를 $W(\cdot)(\omega)$로 표시했지만 이하 $W(\cdot, \omega)$로 표기하겠다.

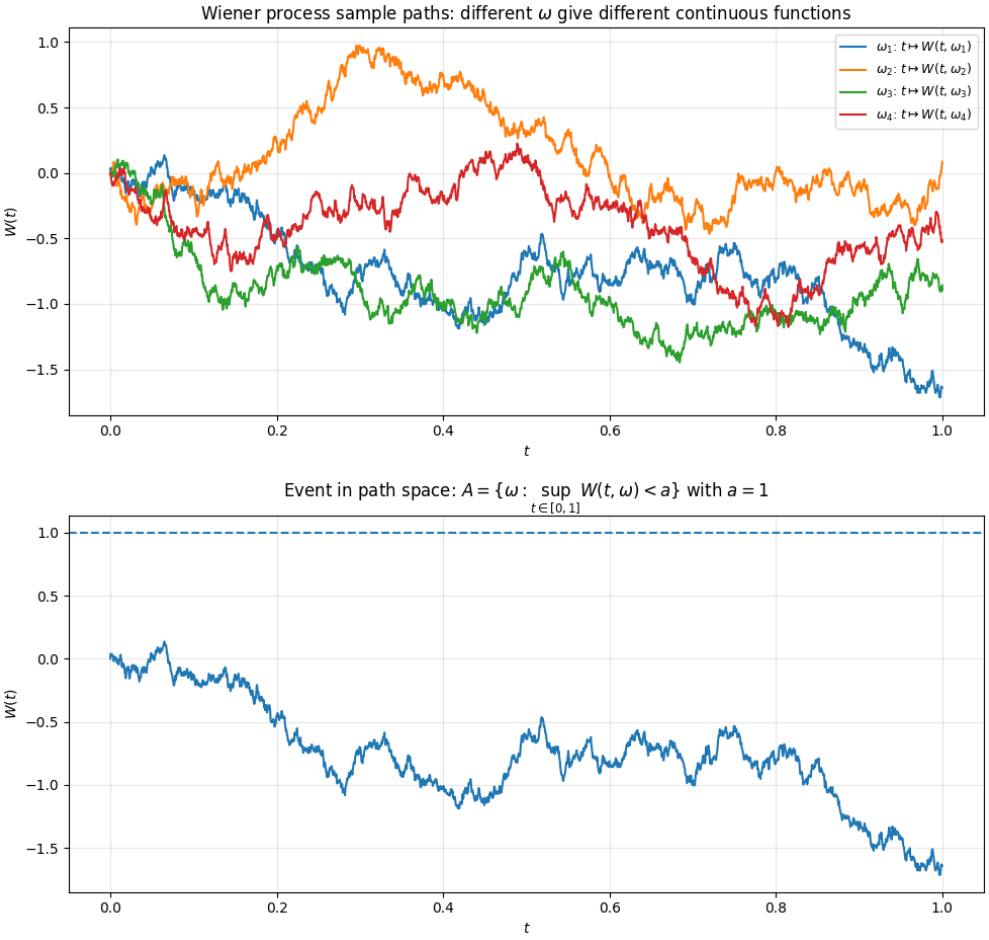


Figure 1: (위) Wiener process의 sample path, (아래) 사건(event)의 시각적 해석

Wiener process는 확률과정 $W : [0,1] \times \Omega \to \mathbb{R}, \quad (t, \omega) \mapsto W(t, \omega)$ 로 정의된다. 여기서 하나의 $\omega \in \Omega$를 고정하면 $t \mapsto W(t, \omega)$ 는 $[0,1]$ 위의 연속함수가 되므로 $W(\cdot, \omega) \in C[0,1]$ 로 생각할 수 있다. 즉 Wiener process는 $W(\cdot) : \Omega \to C[0,1]$ 인 $C[0,1]$-값 확률변수이다. 따라서 서로 다른 $\omega_1, \omega_2, \ldots$에 대해 서로 다른 sample path $W(\cdot, \omega_i)$가 나타난다.

4 Convergent Determining Class

확률측도열 $\{P_n\}$이 주어진 위상공간 S 위에서 분포수렴 $P_n \Rightarrow P$ 를 정의할 때, 이는 일반적으로 모든 Borel 집합 $A \subset S$에 대해 $P_n(A) \to P(A)$ 를 요구하는 것이 아니라, 특정한 사건들의 집합만을 통해서도 판별될 수 있다. 집합족 $\mathcal{C} \subset \mathcal{B}(S)$가 convergence-determining class라 함은, 임의의 확률측도열 $\{P_n\}$과 확률측도 P에 대해 $P_n(C) \to P(C) \quad \forall C \in \mathcal{C}$ 이면 $P_n \Rightarrow P$ 가 성립하는 것을 의미한다. 분포수렴의 정의에서 이러한 집합족이 중요한 이유는, Borel σ-대수 전체를 직접 다루는 대신 상대적으로 단순한 사건들의 모음만으로도 확률과정의 수렴을 판별할 수 있기 때문이다.

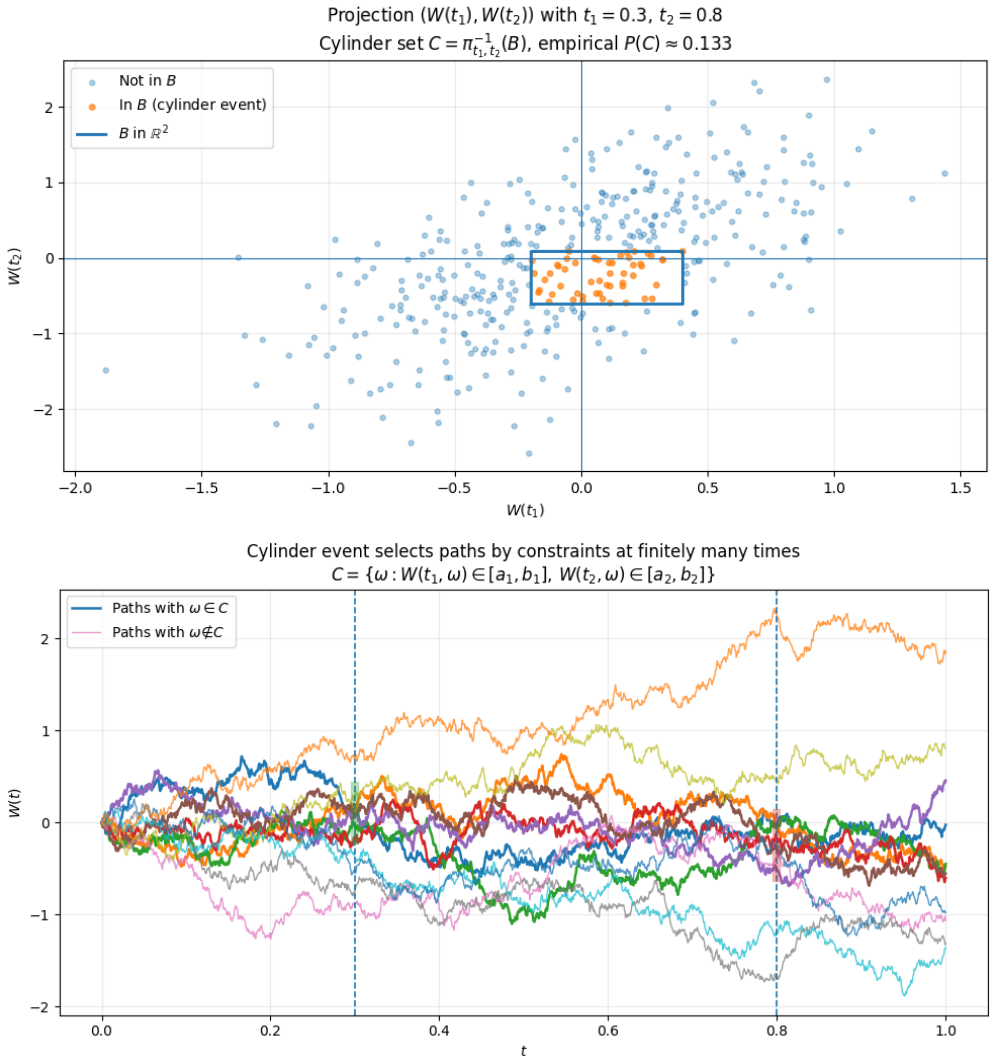


Figure 2: (위) $m = 2$ 인 경우 $B \subset \mathbb{R}^m$, (아래) $W(\cdot, \omega) \in \pi^{-1}_{t_1,\ldots,t_m}(B)$인 sample path 의 시각화

Polish space S (separable and complete space) 위에서 Portmanteau 정리에 따르면 다음 집합족들은

모두 convergence-determining class를 이룬다:

- 모든 열린집합 $G \subset S$에 대해 $\liminf_{n\to\infty} P_n(G) \geq P(G)$,
- 모든 닫힌집합 $F \subset S$에 대해 $\limsup_{n\to\infty} P_n(F) \leq P(F)$.

이제 무한 동전 던지기 실험에서 정의한 $C[0,1]$내의 연속함수들에 대해 어떤 집합을 생각하면 분포수렴을 정의할 수 있는지 살펴보겠다. 표본공간을 $S = C[0,1]$로 두고 uniform topology를 고려하자. 이 경우 열린집합은 열린공들의 가산합으로 생성되며, 대표적인 사건은 $\{f \in C[0,1] : \|f - f_0\|_\infty < \varepsilon\}$ 와 같은 형태를 가진다. 여기서는 $W_n(\cdot)$의 분포를 P_n, $W(\cdot)$의 분포를 P라고 놓자. 분포수렴 $W_n(\cdot) \xrightarrow{d} W(\cdot)$ 혹은 $P_n \Rightarrow P$를 판별하기 위해 모든 가능한 경로사건을 확인할 필요는 없고, 이러한 열린공들로 생성되는 사건들만을 통해 수렴 여부를 결정할 수 있다.
임의의 시점 $0 \leq t_1 < \cdots < t_m \leq 1$ 에 대해, 다음 사상

$$\pi_{t_1,\ldots,t_m} : C[0,1] \to \mathbb{R}^m, \qquad \pi_{t_1,\ldots,t_m}(f) = (f(t_1), \ldots, f(t_m))$$

을 *finite-dimensional projection*이라 한다. 이는 연속함수공간에서 유한 개의 시점에서의 값만을 추출하는 사상이다. 이 projection을 이용하면 $\mathbb{R}^m$의 Borel 집합 $B \in \mathcal{B}(\mathbb{R}^m)$에 대해

$$\pi^{-1}_{t_1,\ldots,t_m}(B) = \{f \in C[0,1] : (f(t_1), \ldots, f(t_m)) \in B\}$$

와 같은 사건을 정의할 수 있다. 이러한 집합을 finite-dimensional cylinder set이라 부른다. 중요한 사실은 이러한 finite-dimensional cylinder set들이 $C[0,1]$의 Borel σ-대수를 생성한다는 점이다. 즉, 두 확률측도 P, Q가

$$P\big(\pi^{-1}_{t_1,\ldots,t_m}(B)\big) = Q\big(\pi^{-1}_{t_1,\ldots,t_m}(B)\big)$$

를 모든 m, 모든 시점 $t_1, \ldots, t_m$, 모든 $B \in \mathcal{B}(\mathbb{R}^m)$에 대해 만족하면, $P = Q$ 가 성립한다. 따라서 finite-dimensional 분포들은 수렴하는 확률측도가 무엇인지 결정(determine)하기에는 충분하다. 그러나 이러한 유한차원 수렴만으로는 $C[0,1]$ 위의 확률측도 수렴을 보장하기에 충분하지 않으며, 이를 위해서는 분포열 $\{P_n\}$의 tightness가 추가로 필요하다. 즉 우리는

$$P_n(\pi^{-1}_{t_1,\ldots,t_m}(B)) \to P(\pi^{-1}_{t_1,\ldots,t_m}(B)) \text{ for all } m \in \mathbb{N},\ \pi_{t_1,\ldots,t_m} \text{ and } B \in B(\mathbb{R}^m)$$

임을 확인하였다 하더라도 $P_n \Rightarrow P$를 보장할 수 없다. 왜냐하면 finite-dimensional 사건들이 각 고정된 유한 시점에서의 분포는 수렴하더라도, 경로가 시점 사이에서 어떻게 움직이는지는 알 수 없기 때문에 이에 대한 분석이 필요하다.
이 점을 보다 구체적으로 이해하기 위해, finite-dimensional 분포는 수렴하지만 경로공간 $C[0,1]$ 위의 분포수렴은 실패하는 전형적인 상황을 생각해 보자. 각 n에 대해 연속함수 $f_n \in C[0,1]$를 다음과 같이

정의한다. f_n은 대부분의 구간에서는 0에 가깝지만, 구간 길이가 $1/n$인 매우 짧은 구간에서 높이가 1인 급격한 "spike"를 갖도록 구성한다. 이때 spike의 위치는 $[0,1]$에서 무작위로 선택된다고 하자. 이러한 함수들로부터 정의된 확률과정을 $W_n(\cdot)$라 하자.

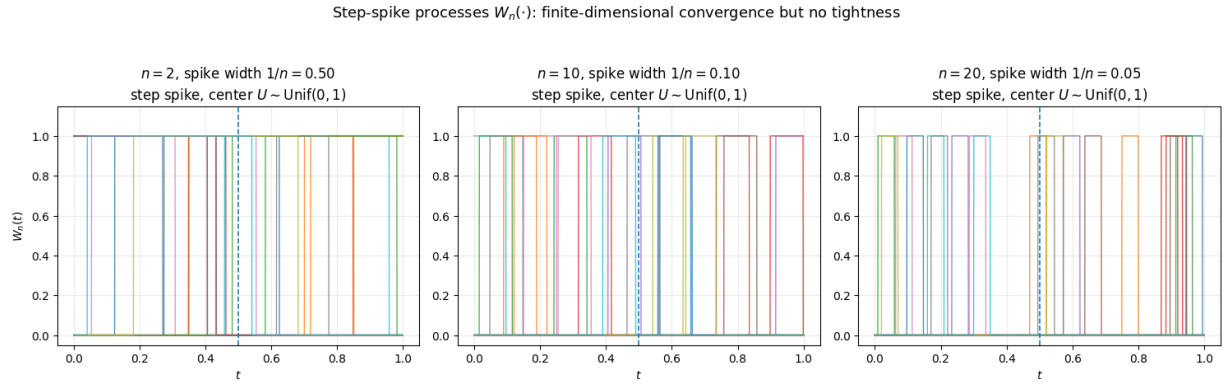


Figure 3: $n = 2, 10, 20$ 인 경우 sample path 의 시각화

이 경우 임의의 고정된 시점 $t \in [0,1]$에 대해, spike가 정확히 그 시점을 포함할 확률은 $1/n$이므로 $W_n(t) \xrightarrow{d} 0 \quad \text{as } n \to \infty$. 더 나아가 임의의 유한 시점 $0 \le t_1 < \cdots < t_m \le 1$에 대해서도 $\big(W_n(t_1), \ldots, W_n(t_m)\big) \xrightarrow{d} (0, \ldots, 0)$, 즉 finite-dimensional 분포는 모두 자명한 극한으로 수렴한다. 그러나 경로 전체의 관점에서 보면 상황은 다르다. 각 $W_n(\cdot)$는 거의 확실하게 어떤 매우 짧은 구간에서 크기 1의 급격한 진동을 가지므로, 이러한 진동은 n이 커져도 사라지지 않는다. 따라서 $\{W_n(\cdot)\}$는 uniform topology 하에서 임의의 compact 집합에 확률질량을 집중시키지 못하며, 분포열 $\{P_n\}$은 tight하지 않다. 이 예시에서 tightness가 실패하는 핵심 이유는 compactness를 totally boundedness [1] 관점에서 이해할 수 있다. $C[0,1]$는 uniform metric 에 대해 complete space이므로[2], 집합 $K \subset C[0,1]$가 compact하다는 것은 곧 K가 totally bounded하다는 것과 동치다. 즉 임의의 $\delta > 0$에 대해 K는 finitely many δ-balls로 덮일 수 있어야 한다.

그러나 spike 과정에서는 spike의 위치가 $[0,1]$에서 자유롭게 움직이며, 서로 다른 위치에 spike를 갖는 두 경로는 uniform norm에서 거리 1만큼 떨어져 있다. 따라서 $\delta < 1/2$인 경우 이러한 경로들을 덮기 위해서는 무한히 많은 열린공이 필요하게 되고, 결과적으로 $\{W_n(\cdot)\}$는 totally bounded하지 않다. 즉 확률질량을 담을 compact 집합을 구성할 수 없으므로 분포열 $\{P_n\}$은 tight하지 않다.

이 예시는 finite-dimensional 분포 수렴은 확률측도를 결정하기에는 충분하지만, 경로공간 위의 분포 수렴을 보장하기에는 본질적으로 부족하며, 이를 보완하기 위해 tightness 조건이 필요함을 암시한다. finite-dimensional cylinder set들에서의 수렴이 경로의 연속성이나 진동을 통제하지 못하기 때문에

[1] 메트릭 공간 (X,d)의 부분집합 $A \subset X$가 totally bounded하다는 것은, 임의의 $\varepsilon > 0$에 대하여 유한한 개수의 점 $x_1, \ldots, x_N \in X$가 존재하여 $A \subset \bigcup_{i=1}^{N} B(x_i, \varepsilon)$ 가 성립하는 것을 말한다. 여기서 $B(x,\varepsilon) = \{y \in X : \ d(x,y) < \varepsilon\}$는 반지름 ε인 열린공이다.

[2] 증명은 참고사항을 보아라

우리는 분포수렴을 위해 추가적인 경로함수의 통제가 필요하다는 것을 알 수 있다. 이는 경로함수의 관점에서는 얼마나 연속적인가 (modulus of continuity) 그리고 S의 위상적 관점에서는 uniformly tightness 를 만족하시키는 compact set을 잡을 수 있는지 두 가치 측면에서 분포수렴 가능성을 조사할 수 있다. 자연스럽게 첫번째 관점에서는 표본경로가 얼마나 연속적인가에 대한 문제로 귀결되어 표본경로들의 Equicontinuity 를 정의하는 문제가 도출된다. 두 번째 관점에서는 이과 같은 맥락에서 어떤 topology 혹은 metric 을 주어야 경로가 연속이라고 말할 수 있는가에 대한 문제가 된다. (이 두 관점에 대한 분석은 Arzelà–Ascoli 정리와 compactness에 대한 결과로 잘 알려져 있다.)
이번 장의 소결론은 다음과 같다. 경로공간 $C[0,1]$ 위에서의 분포수렴을 논하기 위해서는 (1) finite-dimensional 분포들이 극한과 일치하여 확률측도 자체가 무엇인지를 결정할 수 있어야 하며, (2) 동시에 분포열 $\{P_n\}$이 uniform topology 하에서 tightness를 만족하여 경로 전체의 진동과 연속성이 적절히 제어되어야 한다. Finite-dimensional cylinder set들은 확률측도를 결정(determine)하기에는 충분하지만, 경로의 연속성이나 시점 사이의 거동을 반영하지 못하므로 convergence-determining class로는 충분하지 않다. 따라서 $C[0,1]$ 위의 확률과정 수렴을 보장하기 위해서는 finite-dimensional convergence와 함께 equicontinuity를 통한 tightness 조건을 추가로 확인해야 하며, 이 두 조건이 결합될 때 비로소

$$W_n(\cdot) \Rightarrow W(\cdot) \quad \text{in } (C[0,1], \|\cdot\|_\infty)$$

와 같은 함수공간 위의 분포수렴을 얻을 수 있다. 이러한 관점은 이후 Wiener process를 극한으로 하는 functional central limit theorem의 핵심 구조를 이룬다.

5 참고사항

5.1 Wiener Process 의 성질

증명 (두 시점의 경우). $0 \le s \le t \le 1$라 하자. Wiener process의 정의에 의해

$$W(s) = W(s) - W(0) \sim \mathcal{N}(0, s), \qquad W(t) - W(s) \sim \mathcal{N}(0, t-s),$$

그리고 $W(s)$와 $W(t) - W(s)$는 서로 독립이다.

$$\begin{pmatrix} W(s) \\ W(t) \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} W(s) \\ W(t) - W(s) \end{pmatrix}$$

이고 독립이며 정규분포를 따르는 확률변수들의 선형결합은 다시 정규분포를 따르므로 $(W(s), W(t))^\top$는 이변량 정규분포를 따른다. 평균은 $(0,0)^\top$이고 공분산은

$$\text{Var}(W(s)) = s, \qquad \text{Var}(W(t)) = t,$$

$$\text{Cov}(W(s), W(t)) = \text{Cov}\big(W(s),\, W(s) + W(t) - W(s)\big) = s.$$

따라서

$$(W(s), W(t)) \sim \mathcal{N}\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} s & s \\ s & t \end{pmatrix} \right),$$

즉 공분산 행렬의 $(1,2)$-성분은 $\min(s,t)$ 가 된다.

5.2 메트릭 공간에서 Compact set에 대한 동치조건

메트릭 공간 (X,d)에서 compactness는 여러 가지 서로 다른 형태로 동치적으로 기술될 수 있으며, 이러한 동치조건들은 확률측도의 tightness나 함수공간에서의 분포수렴을 논하는 데 핵심적인 역할을 한다.

(1) 정의 부분집합 $K \subset X$가 compact하다는 것은, 임의의 열린 덮개

$$K \subset \bigcup_{\alpha \in A} G_\alpha, \qquad G_\alpha \text{ open}$$

에 대해 유한 부분덮개

$$K \subset \bigcup_{i=1}^{N} G_{\alpha_i}$$

가 존재하는 것을 의미한다. 이는 compactness의 가장 기본적인 위상적 정의이다.

(2) Sequential compactness. 메트릭 공간에서는 compactness가 수열을 통해 특징지어질 수 있다. 즉 $K \subset X$가 compact하다는 것은, K 안의 임의의 수열 $\{x_n\}$이 K 안에서 수렴 부분수열을 갖는 것과 동치이다. 이 성질은 함수공간이나 확률과정의 표본경로를 다룰 때 특히 유용하다.

(3) Complete and Totally boundedness. 메트릭 공간 (X,d)에서 다음 동치가 성립한다:

$$K \text{ compact} \quad \Longleftrightarrow \quad K \text{ is complete and totally bounded.}$$

여기서 totally bounded하다는 것은, 임의의 $\varepsilon > 0$에 대해 K를 유한 개의 ε-ball로 덮을 수 있음을

의미한다. 특히 X 자체가 complete한 공간인 경우에는

$$K \text{ compact} \quad \Longleftrightarrow \quad K \text{ is totally bounded.}$$

(4) 함수공간에서의 특수한 형태. 특히 $(C[0,1], \|\cdot\|_\infty)$와 같은 함수공간에서는 compactness가 Arzelà–Ascoli 정리에 의해 보다 구체적으로 기술된다. 즉 $A \subset C[0,1]$가 compact하기 위한 필요충분조건은 A가 uniformly bounded하고 equicontinuous한 것이다. 이 결과는 경로공간 위에서의 tightness를 표본경로의 연속성 조건으로 해석할 수 있게 해준다.

Theorem 5.1 (Arzelà–Ascoli). $(C[0,1], \|\cdot\|_\infty)$에서 부분집합 $A \subset C[0,1]$에 대하여 다음은 동치이다.

1. A의 *closure* $\overline{A}$는 *compact*하다.
2. A는 *totally bounded*하다.
3. A는 다음 두 조건을 만족한다:

   (*i*) (Uniform boundedness)
   $$\sup_{f \in A} \|f\|_\infty < \infty.$$

   (*ii*) (Equicontinuity) 임의의 $\varepsilon > 0$에 대해 어떤 $\delta > 0$가 존재하여 모든 $f \in A$와 모든 $s, t \in [0,1]$에 대하여
   $$|t - s| < \delta \quad \Rightarrow \quad |f(t) - f(s)| < \varepsilon.$$

6 자율학습 주제

1. 집합 S 위의 함수 $\rho : S \times S \to [0, \infty)$가 metric이 되기 위한 조건을 서술하라.
2. Metric 공간 (S, ρ)에서 중심 $x \in S$와 반지름 $r > 0$에 대하여 다음을 정의하라:
   - Open ball $B_r(x)$
   - Closed ball $\overline{B_r(x)}$

   각 정의를 정확한 수식으로 쓰고, 두 개념의 차이를 설명하라.

3. $C[0,1]$를 $[0,1]$ 위의 연속함수 공간이라 하자. 다음과 같은 L_2-distance를 정의한다:

$$\rho(f,g) = \left(\int_0^1 |f(t) - g(t)|^2 \, dt \right)^{1/2}.$$

함수 $f(t) = t$에 대하여 $\varepsilon > 0$일 때 open ball

$$B_\varepsilon^{(2)}(f) = \{g \in C[0,1] : \rho(g,f) < \varepsilon\}$$

의 원소가 되는 함수 g는 어떤 조건을 만족하는가? 조건을 부등식으로 명확히 쓰고, 서로 다른 예시 함수 두 개를 제시하라.

4. 위 문제에서 metric을 uniform metric

$$\rho_\infty(f,g) = \sup_{t \in [0,1]} |f(t) - g(t)|$$

으로 바꾸었을 때 open ball

$$B_\varepsilon^{(\infty)}(f) = \{g \in C[0,1] : \rho_\infty(g,f) < \varepsilon\}$$

의 원소는 어떤 조건을 만족하는가? 함수의 개형을 그려보아라. 또한 (3)번의 경우와 비교하여 두 open ball 사이의 포함관계를 논하라.

5. $\mathbb{R}^{\mathbb{N}}$ (실수 무한수열 공간)에서 다음 두 함수를 정의하라:

$$d_1(x,y) = \sum_{n=1}^{\infty} 2^{-n} \min\{1, |x_n - y_n|\},$$

$$d_2(x,y) = \sum_{n=1}^{\infty} 2^{-n} \arctan(|x_n - y_n|).$$

(a) 각 함수가 metric이 됨을 증명하라.

(b) 이 metric들이 유도하는 수렴의 의미를 설명하라.

6. 같은 집합 S 위에 서로 다른 두 metric ρ와 ρ'가 주어졌다고 하자. 어떤 수열 $\{x_n\}$이 ρ-metric 관점에서는 수렴하지만 ρ'-metric 관점에서는 수렴하지 않는다는 의미를 각각 정의로 설명하라. 또한 이러한 현상이 실제로 발생하는 구체적인 예를 구성하라.

7. Metric 공간 (S, ρ)에서 다음 정의를 써라.

    - Cauchy 수열
    - Complete metric space

8. Completeness가 주어진 metric에 의존함을 설명하라. 같은 집합 S 위에 서로 다른 두 metric을 정의하여, 한쪽은 complete이고 다른 쪽은 complete가 아닌 예를 제시하라.

9. Topology의 정의를 서술하라. Topology와 σ-algebra의 공통점과 차이점을 닫힘 성질의 관점에서 비교하라.

10. Metric ρ가 유도하는 topology의 정의를 서술하라.

    또한 수열 $\{x_n\}$이 x로 수렴한다는 의미를

    - metric의 관점
    - topology의 관점

    에서 각각 설명하고, 두 정의가 서로 동치임을 논하라.