Introduction to Probability Theory

김용환

Abstract

이 발표의 목적은 수리확률론의 기초적인 개념들을 소개하는 것이다. 독자들은 실변수함수론의 기초적인 개념들에 어느 정도 익숙하다고 가정한다.

Keywords. Conditional Expectation, Martingale

1. Conditional Expectation

$P(\Omega) = 1$인 measure가 주어진 공간 $(\Omega, \mathcal{F}, P)$를 우리는 *probability space*라고 부르고, 이 때 주어진 측도 P를 *probability measure*라 부른다. 이러한 확률측도가 주어진 공간은 앞으로 우리가 논의를 전개할 가장 기본적인 공간이다. 이제 기초확률론에서 정의한 몇 가지 개념을 르벡적분의 언어로 옮겨서 다시 정의하자.

Definition (Random Variable)**.** 앞서 정의한 Ω에 대해 $X\colon \Omega \to \mathbb{R}$가 measurable이라면 이 때 이 X를 *real-valued random variable* 이라 부른다.

Definition (Expectation)**.** 위와 같은 random variable X가 주어졌을 때, 기댓값을 $E[X] = \int X\,dP$와 같이 정의한다.

우리가 다음에 정의할 개념은 stochastic process를 다룰 때 필수적인 것이다. 다음 정리는 사실 Radon-Nikodym을 이용해서 증명할 수도 있지만, 우리는 Hilbert Space에서의 Orthogonal Projection을 이용할 것이다.

Theorem (Conditional Expectation)**.** *Integrable random variable X, 그리고 sub-σ-algebra $\mathcal{G} \subset \mathcal{F}$가 주어졌을 때, 어떤 L^1 function $E(X \mid G)\ : (\Omega, \mathcal{G}) \to \mathbb{R}$ 이 존재해, $\forall A \in \mathcal{G}$, $\int_A X\,dP = \int_A E(X \mid \mathcal{G})\,dP$가 성립한다. 또, 이러한 함수 $E(X \mid G)$는 almost everywhere에서 같은 함수값을 가진다.*

Proof. $X \in L^2(\Omega, \mathcal{F}, P)$인 경우부터 생각하자. 이 때, $L^2(\mathcal{G})$은 힐베르트 공간 $L^2(\mathcal{F})$의 닫힌 부분공간이다. 따라서, 우리는 다음의 사영 $\pi\colon L^2(\mathcal{F}) \to L^2(\mathcal{G})$을 생각할 수

있다. 이제, $E(X \mid G)$를 X를 사영시킨 $\pi(X)$으로 정의하자. 이 때, $E(X \mid G)$가 $\mathcal{G}$-measurable 함수임은 쉽게 얻어진다. 또, 임의의 $A \in \mathcal{G}$를 잡아도, $\mathbf{1}_A \in L^2(\mathcal{G})$가 성립한다. 따라서, 이 함수는 $X - E(X \mid G)$과 수직이고, 이로부터 우리는 다음의 등식을 얻는다.

$$\forall A \in \mathcal{G}, \qquad \int_A X\,dP = \langle X, \mathbf{1}_A \rangle = \langle E(X \mid G), \mathbf{1}_A \rangle = \int_A E(X \mid \mathcal{G})\,dP.$$

이제, X의 양수부분만을 생각하자. $X_n = \min(X, n)$으로 정의하자. 그러면, 전체 measure가 1이므로, $X_n \in L^2(\mathcal{G})$이고, 따라서 앞의 논의에 의해 $Y_n = E(X_n \mid G)$을 정의할 수 있고, 이 때 극한값을 $Y = \lim_{n\to\infty} Y_n$라 정의할 수 있다. 만약 $Y_n \geq Y_{n-1}$ almost everywhere라면, by the monotone convergence theorem

$$\int_A Y\,dP = \lim_{n\to\infty} \int_A Y_n\,dP = \lim_{n\to\infty} \int_A X_n\,dP = \int_A X\,dP < \infty$$

이제, 우리는 다음과 같이 $E(X \mid G)$를 정의하자:

$$E(X \mid G) = \begin{cases} Y & (Y < \infty) \\ 0 & (Y = \infty) \end{cases}.$$

$\int Y\,dP < \infty$이기 때문에 Y는 almost everywhere에서 절댓값이 유한이므로, $Y = E(X \mid G)$ almost everywhere이다. 따라서

$$\forall A \in \mathcal{G}, \qquad \int_A X\,dP = \int_A Y\,dP = \int_A E(X \mid \mathcal{G})\,dP.$$

마지막으로, 양수부분과 음수부분으로 나누어서 다음과 같이 정의하면 증명이 끝난다.

$$X = X^+ - X^-, \quad E(X \mid \mathcal{G}) = E(X^+ \mid \mathcal{G}) - E(X^- \mid \mathcal{G}).$$

이제, 앞에서 증명하지 않고 넘어간 다음 명제를 증명하자. $X_1 \geq X_2 \Rightarrow E(X_1 \mid \mathcal{G}) \geq E(X_2 \mid \mathcal{G})$. 다음의 집합을 잡자.

$$A = \{E(X_1 \mid \mathcal{G}) < E(X_2 \mid \mathcal{G})\}$$

그러면 $A \in \mathcal{G}$ 이므로,

$$\int_A (E(X_1 \mid \mathcal{G}) - E(X_2 \mid \mathcal{G}))\,dP = \int_A (X_1 - X_2)\,dP \geq 0$$

가 성립하고, 따라서 $P(A) = 0$, and $E(X_1 \mid \mathcal{G}) \geq E(X_2 \mid \mathcal{G})$ a.e.

$\square$

Proposition. *(Properties of conditional expectations)*

(i). $E(X \mid \mathcal{F}) = X$, *for* $\mathcal{F}$*-measurable* X.

(ii). $E(X \mid \mathcal{F}) = E(X)$, *for* $\mathcal{F} \perp X$.

(iii). $E(\alpha X_1 + \beta X_2 \mid \mathcal{F}) = \alpha E(X_1 \mid \mathcal{F}) + \beta E(X_2 \mid \mathcal{F})$.

(iv). $E(ZX \mid \mathcal{F}) = ZE(X \mid \mathcal{F})$, *for bounded,* $\mathcal{F}$*-measurable* Z.

(v). $E(E(X \mid \mathcal{F}) \mid \mathcal{G}) = E(X \mid \mathcal{G})$, *for* $\mathcal{G} \subset \mathcal{F}$.

(vi). $\|E(X \mid \mathcal{F})\|_p \leq \|X\|_p$, *for* $p \geq 1$.

2. Stopping Times

Definition (Filtration)**.** 증가하는 σ-algebra의 수열 $(\mathcal{F}_n)_{n\geq 0}$이 다음의 성질 $\mathcal{F}_n \subset \mathcal{F}$를 만족하면 이를 Filtration이라 정의하자.

Definition (Stopping Time)**.** 치역이 $\mathbb{N} \cup \infty$인 random variable N이 모든 $n < \infty$에 대해 $\{N = n\} \in \mathcal{F}_n$를 만족한다면 이를 Stopping Time이라 하자.

Definition. Stopping time N이 주어졌을 때, stopped process X_N를 다음과 같이 정의하자:

$$X_N(\omega) = \begin{cases} X_{N(\omega)}(\omega) & (N < \infty) \\ 0 & (N = \infty) \end{cases}$$

Definition. Stopping time N이 주어졌을 때, $\mathcal{F}_N$를 다음과 같이 정의하자.

$$\mathcal{F}_N = \{A\colon\ A \cap \{N \leq n\} \in \mathcal{F}_n\}$$

Proposition. *(Properties of Stopping Times)*

(i). *Stopping time* M, N에 대해

$$M + N,\ \min(M, N),\ \max(M, N)$$

도 *stopping time*이다.

(ii). X_n이 $\mathcal{F}_n$*-measurable for all* n이라면, X_N도 $\mathcal{F}_N$*-measurable function*이다.

(iii). Stopping time $M \leq N$에 대해, $\mathcal{F}_M \subset \mathcal{F}_N$.

(iv). The ***hitting time of A*** : $\inf\{n : S_n \in A\}$ *is a stopping time.*

3. Martingales

3.1. Basic properties

Definition (Martingale)**.** 다음의 조건을 만족하는 random variable들의 수열 X_n을 martingale이라 부르자;

(i). $E(X_n) < \infty$,

(ii). X_n is $\mathcal{F}_n$-measurable,

(iii). 모든 n에 대해, $E(X_{n+1} \mid \mathcal{F}_n) = X_n$.

등호 $=$ 대신 $\leq, \geq$ 가 세번째 조건에서 성립한다면, X_n을 각각의 경우에서 supermartingale, submartingale이라고 부른다.

Proposition. *(Properties of Martingales)*

(i). X_n이 *submartingale*이면 $E(X_n \mid \mathcal{F}_m) \geq X_m$ *for all* $n > m$.

(ii). X_n이 *submartingale*이고, ϕ가 증가하는 볼록함수라면 $\phi(X_n)$도 *submartingale* 이다.

(iii). X_n이 *supermartingale*이라면, 모든 실수 a에 대해$\min(X_n, a)$ 역시 *supermartingale*이다.

Definition (Predictable)**.** $(H_n)_{n\geq 1}$가 predictable sequence라는 것은 H_n이 모든 n에 대해 $\mathcal{F}_{n-1}$-measurable이라는 뜻이다. 이 때, 우리는 다음도 정의한다:

$$(H \cdot X)_n = \sum_{1 \leq m \leq n} H_m (X_m - X_{m-1})$$

Theorem (Unfavorable Games)**.** *Supermartingale* X_n과 *bounded predictable* H_n에 대해, $(H \cdot X)_n$ 또한 *supermartingale*이다.

Proof. $E(X_{n+1} \mid \mathcal{F}_n) \leq X_n = E(X_n \mid \mathcal{F}_n)$이므로, 우리는

$$E(X_{n+1} - X_n \mid \mathcal{F}_n) \leq 0$$

임을 얻는다. 따라서,

$$\begin{aligned} E((H \cdot X)_{n+1} \mid \mathcal{F}_n) &= (H \cdot X)_n + E(H_{n+1}(X_{n+1} - X_n) \mid \mathcal{F}_n) \\ &= (H \cdot X)_n + H_{n+1} E(X_{n+1} - X_n \mid \mathcal{F}_n) \\ &\leq (H \cdot X)_n \end{aligned}$$

□

Theorem (Doob's inequality). *(i).* $(X_n)_{n \geq 1}$가 *submartingale*이고 N이 $P(N \leq k) \leq 1$인 *stopping time*일 때, 다음의 부등식이 성립한다:

$$EX_0 \leq EX_N \leq EX_k$$

(ii). *Submartingale* $(X_n)_{n \geq 1}$과 $\lambda > 0$인 실수에 대해

$$\bar{X}_n = \max_{0 \leq m \leq n} X_m^+, \quad A = \{\bar{X}_n \geq \lambda\}$$

과 같이 정의하자. 그러면 다음의 부등식이 성립한다:

$$\lambda P(A) \leq EX_n \mathbf{1}_A \leq EX_n^+$$

Proof. 먼저, stopping time N에 대해서 $X_{N \wedge n}$가 submartingale이 됨을 보이자. $H_n = \mathbf{1}_{\{N \geq n\}}$.과 같이 정의하면, H_n은 predictable이므로,

$$(H \cdot X)_n = \sum_{1 \leq m \leq n} H_m (X_m - X_{m-1}) = X_{N \wedge n} - X_0$$

역시 submartingale이다. 따라서, 우리는 다음의 식

$$0 = E(X_{N \wedge 0} - X_0) \leq E(X_{N \wedge k} - X_0) = EX_N - EX_0$$

을 얻고, 왼쪽 부등식을 증명한다. 마찬가지의 논의를

$$K_n = \mathbf{1}_{\{N < n\}}$$

에 적용하면, $(K \cdot X)_n = X_n - X_{N \wedge n}$도 submartingale이다. 그 때,

$$0 = E(K \cdot X)_0 \leq E(K \cdot X)_k = EX_k - EX_N$$

이므로 오른쪽 부등호를 증명할 수 있다. 이제 두번째 부분을 증명하자. Stopping time

$$N = \inf\{t\colon\ X_t \geq \lambda\ \ or\ t = n\}$$

을 정의하면, 첫번째 부등식에 의해서

$$\lambda P(A) \leq EX_N \mathbf{1}_A \leq EX_n \mathbf{1}_A$$

가 성립하므로, 증명이 끝났다. □

3.2. Martingale Convergence Theorems

Theorem (Doob's upcrossing lemma)**.** *Submartingale* X_m과 임의의 두 실수 $a < b$ 가 주어졌을 때, 우리는 *crossing time* $N_0 = -1$,

$$N_{2k-1} = \{m > N_{2k-2}\colon X_m \leq a\}$$

$$N_{2k} = \{m > N_{2k-1}\colon X_m \geq b\}$$

을 정의하고, *total upcrossings until time* n을

$$U_n = sup\{k\colon N_{2k} \leq n\}$$

과 같이 정의하자. 이 때, 다음의 부등식이 성립한다;

$$(b-a)EU_n \leq E(X_n - a)^+ - E(X_0 - a)^+$$

Proof. 임의의 양의 실수 $a > 0$에 대해, $\phi(x) = (x-a)^+$는 볼록함수이다. 따라서 $Y_m = a + (X_m - a)^+$는 submartingale이고, 그것의 upcrossing number는 U_m과 같다. 이제, predictable function H_m을 다음과 같이 정의하자:

$$H_m = \begin{cases} 1 & (N_{2k-1} < m \leq N_{2k}) \\ 0 & (else) \end{cases}$$

이것은 upcrossing이 한 번 일어날 때 마다 $b-a$만큼 증가하므로, $(b-a)U_n \leq (H \cdot Y)_n$ 이다. 이제, $K_n = 1 - H_n$과 같이 정의하면, $(K \cdot Y)_n$는 supermartingale이므로, $0 = E(K \cdot Y)_0 \leq E(K \cdot Y)_n$이다. 따라서, $Y_n - Y_0 = (H \cdot Y)_n + (K \cdot Y)_n$를 이용하면 다음을 얻어, 증명이 끝난다.

$$Y_n - Y_0 = E(H \cdot Y)_0 \geq E(H \cdot Y)_n \geq (b-a)EU_n$$

□

Theorem (Martingale Convergence a.e.)**.** *$M = \sup EX_m^+ \leq \infty$인 submartingale X_m은 어떤 L^1 function X로 almost everywhere 수렴한다.*

Proof. $\liminf X_m < a < b < \limsup X_n$인 임의의 두 유리수 a, b를 잡자. 그러면 이 두 유리수에 대한 upcrossing number U_n를 정의할 수 있고, 앞의 정리에 의해

$$EU_n \leq (|a| + EX_n^+)/(b-a) \leq (M + |a|)/(b-a)$$

를 얻는다. 이제, U_n는 증가하므로, 어떤 U로 수렴한다. 그러면 앞의 부등식에 의해 $EU < \infty$이고, 따라서 $U < \infty$ almost everywhere이다. 따라서 $\liminf X_m = \limsup X_n$ a.e.이다. 이제 $X = \lim X_n$는 almost everywhere 존재한다. 우리는 증명을 마무리짓기 위해 $X > -\infty$를 보인다. 그런데,

$$EX_n^- = EX_n^+ - EX_n \leq M - EX_0$$

이므로, Fatou's Lemma에 의해

$$EX^- \leq M - EX_0 < \infty$$

이다. 따라서, $X > -\infty$ a.e이고

$$EX \leq EX^+ + EX^- < \infty.$$

가 되어 증명이 끝난다. □

Theorem (Martingale Convergence in L^p)**.** *Given a submartingale X_m and $p > 1$ such that $M = \sup \|X_n\|_p < \infty$, X_m converges a.e. and in L^p.*

Proof. 우리는 이미 앞에서 X_n가 어떤 X로 almost everywhere 수렴함은 보였다. 이제 $\sup_{0 \leq m \leq n} \|X_m\|_p \leq \frac{p}{p-1} \|X_n\|_p$가 성립한다면, $\sup |X_n| \in L^p$이므로 dominated convergence에서 $X_n \to X$ in L^p이다. 즉, 우리는 다음의 식이 성립함을 보이면 충분하다.

$$\sup_{0 \leq m \leq n} \|X_m\|_p \leq \frac{p}{p-1} \|X_n\|_p$$

이제,

$$\bar{X}_n = \sup_{0 \leq m \leq n} X_m$$

로 정의하자. 그러면 $\forall a$, Holder와 Fubini를 이용하여

$$\begin{aligned} E|\bar{X}_n \wedge a|^p &= E\int_0^{\bar{X}_n \wedge a} px^{p-1} \\ &= E\int_0^a px^{p-1}\mathbf{1}_{\bar{X}_n \geq x} \\ &= \int_0^a px^{p-1}P(\bar{X}_n \geq x) \\ &\leq EX_n\int_0^a px^{p-2}\mathbf{1}_{\bar{X}_n \geq x} \\ &= \frac{p}{p-1}EX_n \cdot (\bar{X}_n \wedge a)^{p-1} \\ &\leq \frac{p}{p-1}\|X_n\|_p\|\bar{X}_n \wedge a\|_p^{p-1} \end{aligned}$$

를 보이고, $a \to \infty$로 보내면 원하는 결론을 얻는다. □

$p > 1$일 때와 $p = 1$일 때 martingale의 수렴성은 서로 다른 성질을 가진다. $p = 1$일 때를 알아보기 위해, 우리는 다음의 새로운 개념을 도입한다.

Definition (Uniformly Integrable)**.** $\{X_i\}_{i\in I}$가 다음 조건을 만족할 때, 이를 uniformly integrable이라 부르자:

$$\lim_{M\to\infty}\{\sup_{i\in I}(\int_{\{|X_i|>M\}}|X_i|\,dP)\} = 0$$

Theorem (Martingale Convergence in L^1)**.** *Submartingale X_m에 대해서, 다음 두 성질은 동치이다;*

(i). (X_n) is uniformly integrable

(ii). (X_n) converges in L^1

Proof. (X_n)가 uniformly integrable이라면 , 우리는 충분히 큰 M을 찾아, $sup_{i\in I}E|X_i| \leq M+1 < \infty$이게 할 수 있다. 즉, 이 때 X_n은 $M+1$의 상계를 가지게 된다. 이제, X_n이 L^1 function에 의해 위로 almost everywhere bounded이고, X_n가 측도수렴한다면, L^1수렴도 한다. 즉, 우리는 $(i) \to (ii)$ 방향이 성립함을 보였다. 반대의 방향은 실변수함수론에서 convergence in measure와 관련된 기본적인 성질로, 참고문헌 [2]에서 찾을 수 있다. □

References

1. R. Durrett, *Probability: Theory and Examples*, 4th ed., Cambridge University Press, New York, NY, 2010.
2. G. Folland, *Real Analysis: Modern Techniques and their Applications*, 2nd ed., John Wiley and Sons, 2013.

서울대학교

E-mail address: kimyh4306@snu.ac.kr