Preface

확률론은 통계학을 공부하는 데 있어 굉장히 중요한 과목이다. 여기에서는 통계에 필요한 부분만 적었다.

덤으로 극단값 이론의 기초도 수록하였다.

This is a Quarto book.

To learn more about Quarto books visit https://quarto.org/docs/books.

```
1 + 1
```

```
[1] 2
```

Part I

Intro

Part II

Probability

1 Measure and Integration

1.1 Limit of sets

우리는 σ-field 에만 관심을 갖고 있지만 이를 설명하는 데 유용한 더 작은 하위 클래스들이 있다.

Definition 1.1.

- S: a collection of subsets of Ω. We say that S is a

1. π-**system**: if $A, B \in S \implies A \cap B \in S$
2. λ-**system**: if
   - $\Omega \in S$
   - $A, B \in S$ and $A \subseteq B \implies B \backslash A \in S$
   - $A_n \uparrow A$ and $A_n \in S \implies A \in S$
3. **Algebra** (field) if
   - $\emptyset, \Omega \in S$
   - $A \in S \implies A^c \in S$
   - $A, B \in S \implies A \cup B \in S$
4. **Monotone class** if
   - $A_n \in S$ and $A_n \uparrow A \implies A \in S$
   - $A_n \in S$ and $A_n \downarrow A \implies A \in S$
   - Recall that $A_n \uparrow A$ means that $A_1 \subseteq A_2 \subseteq \ldots$ and $\cup_n A_n = A$ and $A_n \downarrow A$ means that $A_n \supseteq A_2 \supseteq \ldots$ and $\cap_n A_n = A$
5. σ-**algebra** (field) if

- $\emptyset, \Omega \in S$
- $A \in S \implies A^c \in S$
- $A_n \in S \implies \cup A_n \in S$

> **Remark**
>
> - σ-algebra는 π-system, λ-system, monotone class 그리고 algebra 이다.
> - Arbitrary intersections of σ-algebra, algebra, π-system, λ-system 은 각각 σ-algebra, algebra, π-system, λ-system이다.

다음은 PROBABILITY THEORY - PART 1 MEASURE THEORETICAL FRAMEWORK 에서 가져온 몇 가지 예이다.

Ω	S(π-system)	$\mathcal{A}(S)$ (algebra generated by S)	$\sigma(S)$
$(0,1]$	$\{(a,b] : 0 < a \le b \le 1\}$	$\{\cup_{k=1}^N I_k : I_k \in S$ are pairwise disjoint$\}$	$\mathcal{B}(0,1]$
$[0,1]$	$\{(a,b] \cap [0,1] : a \le b\}$	$\{\cup_{k=1}^N R_k : R_k \in S$ are pairwise disjoint$\}$	$\mathcal{B}[0,1]$
$\mathbb{R}^d$	$\{\prod_{i=1}^d (a_i, b_i] : a_i \le b_i\}$	$\{\cup_{k=1}^N R_k : R_k \in S$ are pairwise disjoint$\}$	$\mathcal{B}_{\mathbb{R}^d}$
$\{0,1\}^{\mathbb{N}}$	Collection of all cylinder sets	Finite disjoint unions of cylinders	$\mathcal{B}(\{0,1\}^{\mathbb{N}})$

위 표를 보면, π-system이나 이 π-system으로 만든 algebra는 명시적으로 표사되어 있으나, σ-algebra를 이해하기는 쉽지 않다. 이는 Borel set은 countable number of operations on intervals로 표현하기 쉽지 않기 때문이라고 한다.

이렇게 원소를 파악하기 어려운 σ-algebra에 대해 말해주는 보조정리 뒤 개가 있다. 이들이 말하고자 하는 바는 동일하고 많은 경우에 둘을 바꿔서 사용할 수 있다고 한다.

Lemma 1.1 ($\pi - \lambda$ theorem).

- Let Ω be a set and $\mathcal{F}$ be a set of subsets of Ω.

1. $\mathcal{F}$ is a σ-algebra $\iff$ it is a π-system as well as a λ-system.
2. If S is a π-system, then $\lambda(S) = \sigma(S)$. 이때 $\lambda(S)$는 S를 포함하는 모든 λ-system들의 intersection이다.

Proof

1. 한쪽 방향은 분명하다. $\mathcal{F}$가 π-system이면서 λ-system이라고 하자. 그러면 $\Omega \in \mathcal{F}$이고 만약 $A \in \mathcal{F}$이면, $A^c = \Omega \backslash A \in \mathcal{F}$이다. 만약 $A_n \in \mathcal{F}$이면, 이것의 finite unions $B_n := \cup_{k=1}^{n} A_k = (\cap_{k=1}^{n} A_k^c)^c \in \mathcal{F}$이다. (intersection에서는 $\mathcal{F}$가 π-system이라는 것을 이용) 또한 countable union $\cup A_n$ 또한 B_n의 increasing limit이고 $\mathcal{F}$ 안에 들어가기에 λ-property를 만족

2. $\lambda(S)$ 자체도 S를 포함하는 가장 작은 λ-system이다. 이에 $\lambda(S) = \sigma(S)$를 보이려면 $\mathcal{F} := \lambda(S)$가 π-system임을 보이면 된다. 즉 $A, B \in \mathcal{F}$이면 $A \cap B \in \mathcal{F}$임을 보이면 된다. $A \in S$를 고정하고 $\mathcal{F}_A := \{B \in \mathcal{F} : B \cap A \in \mathcal{F}\}$라 하자. S가 π-system이므로 $\mathcal{F}_A \supset S$이다. 여기서 $\mathcal{F}_A$가 λ-system임을 보이고자 한다. $\Omega \in \mathcal{F}_A$이다. 만약 $B, C \in \mathcal{F}_A$이고 $B \subseteq C$이면, $\mathcal{F}$가 $C \cap A$와 $B \cap A$를 포함하는 λ-system이기 때문에 $(C \backslash B) \cap A = (C \cap A) \backslash (B \cap A) \in \mathcal{F}$이다. 따라서 $(C \backslash B) \in \mathcal{F}_A$이다. 마지막으로 만약 $B_n \in \mathcal{F}_A$이고 $B_n \uparrow B$이면, $B_n \cap A \in \mathcal{F}_A$이고 $B_n \cap A \uparrow B \cap A$이다. 따라서 $B \in \mathcal{F}_A$이다. 이것은 $\mathcal{F}_A$가 S를 포함하는 λ-system임을 의미하며 따라서 $\mathcal{F}_A \supset \mathcal{F}$이다. 다른 말로, 모든 $A \in S$와 $B \in \mathcal{F}$에 대해 $A \cap B \in \mathcal{F}$이다. 이제 $A \in \mathcal{F}$를 고정하자. 다시 $\mathcal{F}_A := \{B \in \mathcal{F} : B \cap A \in \mathcal{F}\}$라고 정의하자. 앞에서 보인대로 하면 $\mathcal{F}_A \supset S$임을 보일 수 있다. 이를 바탕으로 $\mathcal{F}_A$가 λ-system이며 모든 $A \in \mathcal{F}$에서 $\mathcal{F}_A = \mathcal{F}$임을 보일 수 있다. 즉 $\mathcal{F}$는 π-system이다.

Lemma 1.2 (Monotone class theorem).

- Let Ω be a set and let S be a collection of Ω. If S is an algebra, then the monotone class generated by S is a σ-algebra. That is, $\mathcal{M}(S) = \sigma(S)$.

Q. (Unique extension of measures) $(\mathbb{R}, \mathcal{B})$ 에서의 두 확률측도가 모든 구간에서 일치한다면, 이들은 같은가?

다음 예를 통해 거짓이라는 것을 확인할 수 있다.

Example 1.1 ($\pi - \lambda$ theorem).

- $\Omega = \{1, 2, 3, 4\}$ 이고 $S = \{\{1, 2\}, \{2, 3\}, \{3, 4\}\}$ 라고 하자. 그러면 $\sigma(S) = 2^\Omega$ (power set)이라는 것을 보일 수 있다.
- Ω 에서 두 확률측도 μ, ν 를 각각
  - $\mu_i = \frac{1}{4}, \forall i$
  - $\nu_1 = \nu_3 = \frac{1}{2}, \nu_2 = \nu_4 = 0$ 이라고 하자. 그러면 $\mu(A) = \frac{1}{2} = \nu(A)$ for all $A \in S$ 이나, $\mu \neq \nu$ on $\sigma(S)$ 이다.

다만 다음 따름정리와 같이 긍정적인 결과도 있다.

Lemma 1.3. S 가 Ω 의 부분집합의 π-system 이고 $\mathcal{F} = \sigma(S)$ 라 하자. 만약 $\mathbb{P} = \mathbb{Q}$ 가 $\mathcal{F}$ 에서의 확률측도이며 모든 $A \in S$ 에서 $\mathbb{P}(A) = \mathbb{Q}(A)$ 라 하자. 그러면 모든 $A \in \mathcal{F}$ 에서 $\mathbb{P}(A) = \mathbb{Q}(A)$ 이다.

Proof

- $\mathcal{G} = \{A \in \mathcal{F} : \mathbb{P}(A) = \mathbb{Q}(A)\}$ 라 하자. 가정에 의해 $\mathcal{G} \supseteq S$ 이다.
- 여기서 $\mathcal{G}$ 가 λ-system 임을 밝히고자 한다. 분명히 $\Omega \in \mathcal{G}$ 이;다. 만약 $A, B \in \mathcal{G}$ 이고 $A \supseteq B$ 이면, $\mathbb{P}(A \backslash B) = \mathbb{P}(A) - \mathbb{P}(B) = \mathbb{Q}(A) - \mathbb{Q}(B) = \mathbb{Q}(A \backslash B)$ 이다. 이는 즉 $A \backslash B \in \mathcal{G}$ 임을 내포한다. 마지막으로 만약 $A_n \in \mathcal{G}$ 이고 $A_n \uparrow A$ 이면, $\mathbb{P}(A) = \lim_{n\to\infty} \mathbb{P}(A_n) = \lim_{n\to\infty} \mathbb{Q}(A_n) = \mathbb{Q}(A)$ 이다. (이는 measure의 countable additivity로부터 옴)

- 따라서 $\mathcal{G} \supseteq \lambda(S)$ 이고 이는 Lemma **??** 에 의해 $\sigma(S)$ 와 같음을 보일 수 있다. 따라서 $\mathcal{F}$ 에서 $\mathbb{P} = \mathbb{Q}$ 이다.

다시 Example **??** 를 살펴보면, S 가 π-system이 아니라는 것을 알 수 있다.

Remark

- σ-algebra의 집합, 특히 Borel 집합은 항상 어떤 interval의 가산 합집합이 아니기 때문에, σ-algebra의 모든 요소에 대해 어떤 성질이 성립함을 보이기 위해, 우리는 그 성질을 가진 모든 집합의 collection을 고려하고, 이 collection이 σ-algebra임을 보여준다. 이 과정에서 그것이 λ-system이거나 monotone class (π-system과 algebra를 포함하는) 임을 보여주는 것이 더 쉬울 수 있다.

1.2 Measures

Definition 1.2 (σ-additive)**.**

- $\mathcal{A}$: a collection of subsets of Ω containing the empty set $\emptyset$
- A **set function** on $\mathcal{A}$: $\mu : \mathcal{A} \to [0, \infty]$ with $\mu(\emptyset) = 0$
- We say that μ is **countably additive**, or σ-**additive**, if for all sequences (A_n) of disjoint sets in $\mathcal{A}$ with $\cup_{n=1}^{\infty} A_n \in \mathcal{A}$

$$\mu\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} \mu(A_n).$$

Remark

- Recall that a **measurable space** is a pair $(\Omega, \mathcal{F})$, where $\mathcal{F}$ is a σ-algebra on Ω.

Definition 1.3 (Measure space)**.**

- A **measure space** is a triple $(\Omega, \mathscr{F}, \mu)$, where
  - Ω: set
  - $\mathscr{F}$: σ-algebra on Ω
  - μ: $\mathscr{F} \to [0, \infty]$ is a countably additive set function.
  - Then μ is a measure on $(\Omega, \mathscr{F})$.

이산 측도를 생각하는 것은 그냥 power set에서 생각하면 되기에 쉽다. 그러나 일반적인 상황에서는 Ω의 모든 subset을 잴 수 있는 measure를 정의할 수 없다. 그래서 "충분히 큰" collection of sets에서 정의되는 측도의 존재를 보이는 것은 쉽지 않다. 한 가지 방법으로 Ω의 smaller class of subsets에서의 측도를 만들고 σ-algebra를 generate하는 방법이 있다. 이를 위해 두 가지 문제를 해결해야 한다.

1. (Construction) 우리가 특정한 measure를 whole σ-algebra에 확장할 수 있을까? (**Caratheodory's Extension Theorem**)
2. (Uniqueness) σ-algebra에서 우리가 만들고자 하는 측도가 단 한개만 존재하는가? ($\pi - \lambda$ **systems Lemma**)

1.2.1 Lebesgue measure

Definition 1.4 (Existence and uniqueness of Lebesgue measure).

- There exists a unique Borel measure λ on $[0, 1]$ such that $\lambda(I) = |I|$ for any interval I.

Proof

- $S = \{(a, b] \cap [0, 1]\}$은 $\mathscr{B}$를 생성하는 π-system이다. 따라서 Lemma **??** 에 의해 uniqueness가 증명된다.
- Existence는 증명이 길고 복잡한데, 몇 가지 개요만을 아래에서 서술한다. 좀 더 자세한 사항은 PROBABILITY THEORY - PART 1 MEASURE THEORETICAL FRAMEWORK 를 보길 바란다. 여기서는 $\Omega = [0, 1]$로 둔다.

1. **Outer measure** λ_* 를 정의한다.

$$\lambda_*(A) = \text{inf}\{\sum |I_k| : \text{ each } I_k \text{ is an open interval and } \{I_k\} \text{ a countable cover fo}$$

참고로 outer measure의 성질은 다음과 같다.

   - 모든 subset $A \subseteq \Omega$에 대해 $0 \le \lambda_*(A) \le 1$
   - $\lambda_*(A \cup B) \le \lambda_*(A) + \lambda_*(B)$, for any $A, B \subseteq \Omega$
   - $\lambda_*(\Omega) = 1$

2. λ_* 가 측도가 되도록 하는 σ-field를 잡는다. λ_* 를 a set Ω 에서의 outer measure라고 하자. Caratheodary의 정의에 의해

$$\mathcal{F} := \{A \subseteq \Omega : \lambda_*(E) = \lambda_*(A \cap E) + \lambda_*(A^c \cap E) \text{ for any } E\}.$$

이때 $\mathcal{F}$ 가 σ-algebra이고 λ_* restricted to $\mathcal{F}$ 가 확률측도임을 보일 수 있다.

3. $A = (a, b] \subseteq [0, 1]$ 일 때 $A \in \mathcal{F}$ 임을 보여 $\mathcal{F}$ 가 모든 보렐 집합을 포함함을 보인다. 즉 $\mathcal{F}$ 가 충분히 큼을 보인다.

1.2.2 Construction of Lebesgue measure

Lebesgue 측도의 구축은 충분히 풍부한 집합 class에서 시작하여 흥미로운 측도를 구성하는 일반적인 절차로 발전할 수 있다.

1. 주어진 algebra $\mathcal{A}$ (이 경우 유한 개의 $(a, b]$ 의 합집합) 와 $\mathcal{A}$ 위의 countably additive probability measure P on $\mathcal{A}$ 에 대해, 모든 부분집합에 대해 outer measure P^* 를 $\mathcal{A}$ 의 집합들로 구성된 countable cover에 대한 최솟값을 취함으로써 구성한다. 구체적으로, 임의의 부분집합 E 에 대해

$$P^*(E) = \text{inf}\{\sum_{i=1}^{\infty} P(A_i) : E \supseteq \cup_{i=1}^{\infty} A_i, A_i \in \mathcal{A}\}.$$

2. $\mathcal{F}$ 를 위에서와 같이 정의하고, $\mathcal{F} \supset \mathcal{A}$ 가 σ-algebra이고 P^* 가 $\mathcal{A}$ 에서의 확률측도임을 보인다.

3. $P^* = P$ on $\mathcal{A}$ 임을 보인다.

1.2.3 Push-forward (image) measure

- Measure를 정의하는 방법
  - Caratheodory Extension Theorem
  - Transportation between spaces via functions.

Definition 1.5 (Push-forward measure)**.**

- $(\Omega, \mathscr{F}, \mathbb{P})$를 확률공간이라 하고, X를 $(\Omega, \mathscr{F})$에서 $(E, \mathcal{E})$로 보내는 확률변수라 하자. 그러면

$$\mathbb{Q}(A) = \mathbb{P}(X^{-1}(A)), \quad A \in \mathcal{E}$$

  는 $(E, \mathcal{E})$에서 측도를 정의한다 (**push-forward measure**).

- 이때 $\mathbb{Q} = \mathbb{P} \circ X^{-1}$이라고 쓰고 **law of the distribution** of X라고 부른다.

Remark

- 집합 E의 측도로 먼저 이를 X^{-1}을 이용해 Ω로 보내고 $\mathbb{P}$를 이용해 잰다.
- Push-forward measure는 change-of-variables formula 등 적분이론에서 많이 쓰임
- $\mathbb{Q}$가 측도인 이유는 만약 A_n이 pairwise disjoint이면 $X^{-1}(A_n)$ 또한 pairwise disjoint이기 때문이다. 그러나 B_n이 Ω에서 pairwise disjoint라고 해서, $X(B_n)$이 disjoint하지는 않다. 그래서 일반적으로 **pull-back measure**는 없다. 그러나 X가 일대일함수면 X^{-1}을 pull-back으로 생각할 수 있다.

1.3 Integration

1.3.1 Integration notations

> **Remark**
>
> 적분이론에서 굉장히 다양한 적분 notation을 쓰는데 알아두면 좋을 듯

- $(E, \mathcal{E}, \mu)$: measure space, $f : E \to \mathbb{R}$: a real-valued transformation
- 다음의 세 개의 기호는 같은 것임
  - $\int_E f(x)d\mu(x)$
  - $\int_E f d\mu$ (적분하려는 변수가 분명한 경우 생략)
  - $\int_E f(x)\mu(dx)$

1.3.2 리만-스틸체스 적분

종종 헷갈리는 표현이 기댓값을 다음과 같이 분포함수를 이용해 표현하는 경우가 있다.

$$E(X) = \int x dF(x).$$

우리가 알고 있는 정적분은 x 축을 따라가며 함수값 f(x)가 만드는 면적을 계산한다.

$$\int_a^b f(x)dx.$$

위 식을 더 확장하면 x 대신 임의의 곡선 $g(x)$ 를 적분 변수로 두고 $f(x)$ 를 단순하게 정적분 할 수도 있다.

$$\int_{x=a}^{b} f(x)dg(x).$$

여기서 $dg(x)$는 $g(x)$의 미분소(differential)로, $g(x)$의 움직임을 결정하는 x는 단조 증가하거나 감소한다. 위와 같이 리만 적분을 일반화한 정적분을 **리만-스틸체스 적분(Riemann-Stieltjes Integral)**이라 한다. 리만 적분의 정의를 이용해 리만-스틸체스의 적분을 표현할 수도 있다.

$$\int_{x=a}^{b} f(x)dg(x) = \lim_{N\to\infty} \sum_{n=0}^{N-1} f(t_n)[g(x_{n+1}) - g(x_n)].$$

여기서 x_n은 정적분을 위해 구간 $[a, b]$를 나눈 점, t_n은 닫힌 세부공간 $[x_n, x_{n+1}]$ 사이에 있는 임의점이다.

Example 1.2 (리만-스틸체스 적분을 이용한 기댓값의 계산)**.**

- X: random variable with support $R_X = [0, 1]$ and distribution function
$$F_X(x) = \begin{cases} 0, & \text{if } x < 0 \\ \frac{1}{2}x, & \text{if } 0 \le x < 1 \\ 1, & \text{if } x \ge 1. \end{cases}$$

- 이때의 기댓값은

$$\begin{aligned} E[X] &= \int_{-\infty}^{\infty} x dF_X(x) \\ &= \int_0^1 x dF_X(x) + 1 \cdot \left[F_X(1) - \lim_{x\to 1, x<1} F_X(x)\right] \\ &= \int_0^1 x \frac{d}{dx}\left(\frac{1}{2}x\right)dx + 1 \cdot \left[1 - \frac{1}{2}\right] \\ &= \left[\frac{1}{4}x^2\right]_0^1 + \frac{1}{2} = \frac{1}{4} + \frac{1}{2} = \frac{3}{4}. \end{aligned}$$

1.4 리만 적분과 르베그 적분

여기는 Confused when changing from Lebesgue Integral to Riemann Integral 에 올라왔던 내용을 살펴보기로 한다. 여기서 질문자는 리만 적분을 어떻게 르베그 적분으로 바꾸는지에 대해 관심이 있다.

다음과 같이 확률공간 $(\Omega, \mathcal{F}, P)$ 에서 정의된 음이 아닌 확률변수 X 가 지수분포를 따른다고 하자.

$$P(X < x) = 1 - e^{-\lambda x}.$$

한편, 르베그 적분으로 X 의 기댓값을 쓰면 다음과 같다.

$$E[X] = \int_{\{\omega | X(\omega) \geq 0\}} X(\omega) dP(\omega).$$

여기서 질문자는 이것을 리만 적분으로 어떻게 바꾸냐

$$E[X] = \int_0^\infty x\lambda e^{-\lambda x} dx$$

를 물어보고 있다.

답변은 이것이 적분의 문제가 아닌 변수변환의 문제라고 한다.

By definition, given $X : \Omega \to \mathbb{R}$ a random variable, $E[X] = \int_\Omega X$. X defines a measure $\tilde{m}$ in $\mathbb{R}$, called the **push-forward**, by $\tilde{m}(A) = P(X^{-1}(A))$. By definition, this measure is invariant under X, and hence

$$\int_\mathbb{R} f d\tilde{m} = \int_\Omega f \circ X dP.$$

The equality follows from the usual arguments (prove for characteristics, simple functions, then use convergence. Recall that $1_A \circ X = 1_{X^{-1}(A)}$).

Let h be the density of X. We then have, by definition of density, that $\tilde{m}(A) = P(X^{-1}(A)) = \int_A hdm$ for any $A \in \mathcal{B}(\mathbb{R})$, where m is the Lebesgue measure. By **change of variables**, we have

$$\int_\mathbb{R} f d\tilde{m} = \int_\mathbb{R} f \cdot hdm.$$

Combining these equations,

$$\int_{\mathbb{R}} f \cdot h dm = \int_{\Omega} f \circ X dP.$$

Taking $f = \text{Id}$ yields

$$\int_{\mathbb{R}} x h(x) dx = \int_{\Omega} X dP = E[X].$$

Taking $f = \text{Id} \cdot \mathbf{1}_I$, where I is some interval (for example, $(0, +\infty)$ as in your case), we have

$$\int_I x h(x) dx = \int_{X^{-1}} X dP,$$

recalling again that $\mathbf{1}_A \circ X = \mathbf{1}_{X^{-1}(A)}$. Since $P(X < 0)$ in your case is 0, this last integral is actually equal to the integral over the whole space, and hence to $E[X]$, which gives your equality.

2 Probability

2.1 Probability Triples

다음은 콜모고로프 가 정리한 수리적 기반의 확률론이다.

Q. 왜 probability triple이 필요한가? Single도 아니고 double도 아니고 왜 triple이어야 하는가?

- **Sample space** Ω (표본공간): 이것은 any non-empty set이면 된다. 예를 들어 uniform distribution일 때 $\Omega = [0, 1]$이 있다.
- $\mathcal{F}$: σ-algebra 또는 σ-field: 이것은 Ω의 subset들의 collection으로 $\emptyset$, Ω 등을 포함한다.
- **Probability** P: a mapping from $\mathcal{F}$ to $[0, 1]$ with
  - $P(\emptyset) = 0$
  - $P(\Omega) = 1$
  - P is countably additive, $P(A_1 \cup A_2 \cup \cdots) = P(A_1) + P(A_2) + \cdots$
- Recall that the elements of any field or σ-field (Definition **??** 참고) are called **random events** (or simply **events**).

2.2 Probabilities

Definition 2.1 (Probability).

- Let Ω be any set and $\mathcal{A}$ be a field of its subsets. We say that P is a **probability** on the measurable space $(\Omega, \mathcal{A})$ if P is defined for all events $A \in \mathcal{A}$ and satisfies the following axioms.

1. $P(A) \geq 0$ for each $A \in \mathcal{A}$; $P(\Omega) = 1$
2. P is **finitely additive**. That is, for any finite number of pairwise disjoint events $A_1, \ldots, A_n \in \mathcal{A}$ we have
$$P\Big(\cup_{i=1}^n A_i \Big) = \sum_{i=1}^n P(A_i).$$
3. P is continuous at $\emptyset$. That is, for any events $A_1, A_2, \ldots, \mathcal{A}$ such that $A_{n+1} \subset A_n$ and $\cap_{n=1}^\infty A_n = \emptyset$, it is true that
$$\lim_{n\to\infty} P(A_n) = 0.$$

Note that conditions 2 and 3 are equivalent to the next one 4.

4. P is σ-additive (countably additive), that is
$$P\Big(\cup_{n=1}^\infty A_n \Big) = \sum_{n=1}^\infty P(A_n)$$
for any events $A_1, A_2, \ldots \in \mathcal{A}$ which are pairwise disjoint.

Example 2.1 (A probability measure which is additive but not σ-additive)**.** Let Ω be the set of all rational numbers r of the unit interval $[0, 1]$ and $\mathcal{F}_1$ the class of the subsets of Ω of the form $[a, b]$, $(a, b]$, (a, b) or $[a, b)$ where a and b are rational numbers. Denote by $\mathcal{F}_2$ the class of all finite sums of disjoint sets of $\mathcal{F}_1$. Then $\mathcal{F}_2$ is a field. Let us define the probability measure P as follows:
$$P(A) = b - a, \quad \text{if } A \in \mathcal{F}_1,$$
$$P(B) = \sum_{i=1}^n P(A_i), \quad \text{if } B \in \mathcal{F}_2, \text{ that is, } B = \sum_{i=1}^n A_i, A_i \in \mathcal{F}_1.$$

Consider two disjoint sets of $\mathcal{F}_2$ say

$$B = \sum_{i=1}^{n} A_i \text{ and } B' = \sum_{j=1}^{m} A'_j,$$

where $A_i, A'_j \in \mathcal{F}_1$ and all A_i, A'_j are disjoint. Then $B + B' = \sum_{k=1}^{m+n} C_k$ where either $C_k = A_i$ for some $i = 1, \dots, n$, or $C_k = A'_j$ for some $j = 1, \dots, m$. Moreover,

$$\begin{aligned} P(B + B') &= P\Big(\sum_k C_k\Big) = \sum_k P(C_k) = \sum_{i,j} (P(A_i) + P(A'_j)) \\ &= P(A_i) + \sum_j P(A'_j) = P(B) + P(B'). \end{aligned}$$

and hence P is an additive measure.

Obviously every one-point set $\{r\} \in \mathcal{F}_2$ and $P(\{r\}) = 0$. Since Ω is a countable set and $\Omega = \sum_{i=1}^{\infty} \{r_i\}$, we get

$$P(\Omega) = 1 \neq 0 = \sum_{i=1}^{\infty} P(\{r_i\}).$$

This contradiction shows that P is not σ-additive.

3 Random Variables

3.1 Random Variables

Definition 3.1 (잴 수 있는 함수 (가측함수)).

- 확률공간: $(\Omega, \mathcal{F}, P)$, $f : \Omega \to \mathbb{R}$

$$B \in \mathcal{B}(\mathbb{R}) \implies f^{-1}(B) \in \mathcal{F}$$

이면 함수 f를 **잴 수 있는 함수(measurable function)**라 부름

Example 3.1 (잴 수 없는 함수의 예).

- 표본공간 $\Omega = \{1, 2, 3\}$, 사건공간 $\mathcal{F} = \{\Omega, \emptyset, \{1, 2\}, \{3\}\}$
- 이때 $f(1) = 1, f(2) = 2, f(3) = 3$인 함수 $f : (\Omega, \mathcal{F}) \to (\mathbb{R}, \mathcal{B}(\mathbb{R}))$인 함수
- 그런데 $\{1\} \in \mathcal{B}(\mathbb{R})$이지만 $f^{-1}(\{1\}) = \{1\} \notin \mathcal{F}$이므로 f는 가측이 아니며, 따라서 확률변수가 아님

Remark

- 확률변수는 확률공간 위에서 잴 수 있는 **함수**임

Definition 3.2 (Random Variables). Given a probability triple $(\Omega, \mathcal{F}, P)$, a **random variable** is a function X from Ω to $\mathbb{R}$, such that

$$\{\omega \in \Omega; X(\omega) \leq x\} \in \mathcal{F}, \quad x \in \mathbb{R}.$$

Q. Random variable을 정의하는데 왜 inverse image를 쓰는가?

Commonly a probability measure P is added to $(\Omega, \mathcal{F})$. Then sets like $\{X \in A\} := \{\omega \in \Omega | X(\omega) \in A\}$ can $= X^{-1}(A)$ be **measured** if they belong to $\mathcal{F}$. 예를 들면 $X : \Omega \to \mathbb{R}$이 확률변수일 때 $X < 1$일 확률을 구하려면 $X^{-1}(-\infty, 1)$이 가측이어야 할 것이다.

Example 3.2 (확률변수의 inverse image)**.** Proschan (2016) 예제 4.2이다.

- $(\Omega, \mathcal{F}, P) = ((0,1), \mathbb{B}_{(0,1)}, \mu_L)$에서의 확률변수

$$X(\omega) = \frac{1}{\omega(1-\omega)}$$

  를 생각하자.

- Borel set B를 $\{(6.25, \infty) \cup \{4\}\}$
- 역상: $X^{-1}(B) = \{(0, 0.2) \cup (0.8, 1) \cup \{0.5\}\} \in \mathcal{B}_{(0,1)}$

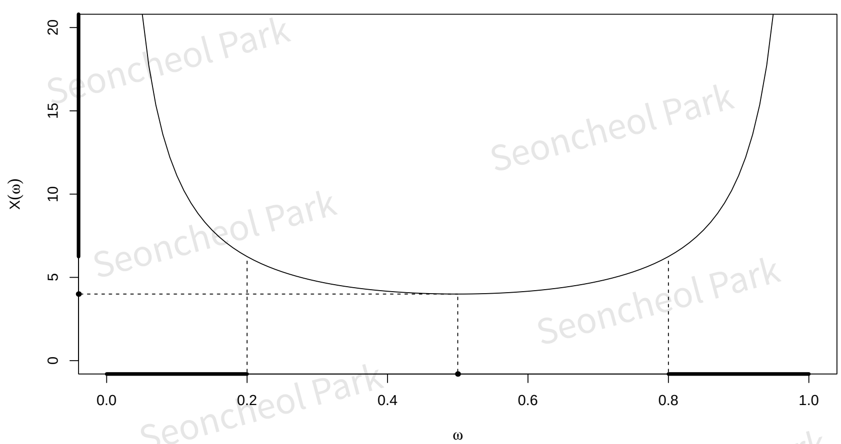

Figure 3.1: Figure: 확률변수 X의 그림.

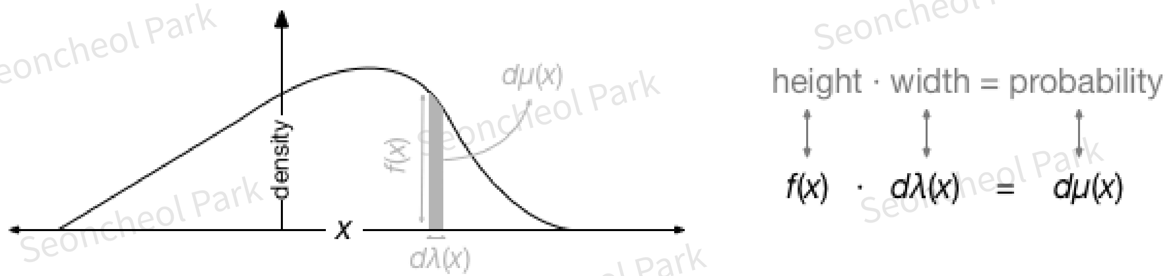

Figure 3.2: Change of measures.

3.2 Radon-nikodym derivative

확률측도는 volume element의 일반화라고 볼 수 있다.

- $\mu(x)$: probability measure, interval이나 set of points들을 인풋으로 받고 area/volume에 해당하는 확률(양수)을 아웃풋으로 주는 함수다.
- $\lambda(x)$: reference measure. We often take $\lambda(x)$ as the Lebesgue measure which is essentially just a uniform function over the sample space.

The reference measure $\lambda(x)$ is essentially just a meter-stick that allows us to express the probability measure as a simple function $f(x)$. That is, we represent the probability measure $\mu(x)$ as $f(x)$ by comparing the probability measure to some specified reference measure $\lambda(x)$. This is essentially the intuition that is given by the Radon-Nikodym derivative

$$f(x) = \frac{d\mu(x)}{d\lambda(x)}$$

or equivalently

height = area / width.

Note that we can also represent the same idea by

$$\mu(A) = \int_{A \in X} f(x) d\lambda(x),$$

where $\mu(A)$ is the sum of the probability of events in the set A which is itself a subset of the entire sample space X. Note that when $A = X$ then the integral must equal 1 by definition of probability.

라돈-니코딤 정리는 조건부 확률에 응용된다고 함.

Definition 3.3 (Integrable Random Variable)**.** Gut (2014) 의 53쪽에 따르면, $E|X| < \infty$ 인 경우 random variable X 가 integrable 하다고 부른다.

Example 3.3. Given a probability measure P and sample space Ω, it is true that

$$\int_{\Omega} dP = 1.$$

Because

$$\int_{\Omega} dP = P(\Omega) = 1.$$

More generally

$$\int_{A} dP = \int_{\Omega} 1_A dP = P(A), \quad A \in \mathcal{F}.$$

Definition 3.4 ($\mathcal{L}^p$)**.** 다음과 같은 확률공간 $(\Omega, \mathcal{F}, P)$ 를 생각하자. $p > 1$ 에 대해, 확률변수 X 가 $E|X|^p < \infty$ 이면 $X \in \mathcal{L}^p$ 라고 하며 다음과 같은 놈 $\|X_p\| = (E|X|^p)^{\frac{1}{p}}$ 를 정의할 수 있다.

3.3 Distribution

- 확률변수의 정의는 임의의 measurable subset of possible outcomes (points, bounded/unbounded intervals 등)에 measure (확률)를 부여할 수 있어야 함
- Semi-infinite interval $(-\infty, x]$ 또한 이러한 measurable subset이므로 $\mathbb{R}$ 에서 정의된 모든 확률변수는 CDF를 갖음
- PDF는 CDF의 도함수 개념이므로 CDF가 미분가능해야 전역적으로 PDF도 존재

3.4 Expectation

- **Expectation**: integral with respect to the probability measure

Definition 3.5 (Expectation).

- $(\Omega, \mathcal{A}, \mathbb{P})$: Probability space
  - Ω: set (sample space)
  - $\mathcal{A}$: σ-algebra on Ω
  - $\mathbb{P}$: Probability measure
- $X : \Omega \to \mathbb{R}$: Random variable (a measurable fct)
- **Expectation**: The concept of integral of X w.r.t. $\mathbb{P}$

$$E[X] \triangleq \int_\Omega X(\omega) d\mathbb{P}[\omega]$$

Remark

- 다음의 notation들은 모두 X의 기댓값을 의미
  - $\int_\Omega X(\omega) d\mathbb{P}[\omega]$
  - $\int_\Omega X d\mathbb{P}$ (적분하려는 변수가 분명한 경우 생략)
  - $\int_\Omega X(\omega) \mathbb{P}[d\omega]$

3.5 Cantor Random Variable

- Cantor distribution: 누적분포함수가 Cantor function인 probability distribution
- Cantor distribution은 PDF나 PMF가 존재하지 않음