PSH

수리통계학 강의록

Mathematical Statistics Lecture Note

June 21, 2019

Seoul National University

함께했던 날들과,
함께할 날들을 위해.
2018년의 끝자락, 관악에서.

Acronyms

WLOG 일반성을 잃지 않고 (Without Loss Of Generality)
TFAE 다음은 서로 동치이다. (The Followings Are Equivalent.)
ow. 그렇지 않으면 (OtherWise)
ae. 거의 어디서나 (Almost Everywhere)
LUBP Least Upper Bound Property
GLBP Greatest Lower Bound Property
MSP Monotone Sequence Property
MCT 단조수렴정리 (Monotone Convergence Theorem)
DCT Lebesgue의 지배수렴정리 (Lebesgue's Dominated Convergence Theorem)
FTC 미적분학의 기본정리 (Funcdamental Theorem of Calculus)
MVT 평균값 정리 (Mean Value Theorem)
IVT 중간값 정리 (Intermediate Value Theorem)
INFT 역함수 정리 (INverse Function Theorem)
rv. 확률변수 (Random Variable) 혹은 확률벡터 (Random Vector)
PDF 확률밀도함수 (Probability Density Function)
PMF 확률질량함수 (Probability Mass Function)
CDF 누적분포함수 (Cumulative Distribution Function)
MGF 적률생성함수 (Moment Generating Function)
CF 특성함수 (Characteristic Function)
CGF 누율생성함수 (Cumulant Generating Function)
PGF 확률생성함수 (Probability Generating Function)
iid. Independent and Identically Distributed rv.
io. Infinitely Often
as. 거의 확실하게 (Almost Surely)

Contents

Part I

Preliminary

Chapter 1

Preliminary

1.1 Analysis

1.1.1 Inequalities

다음 부등식들은 각종 정리의 증명에서 유용하게 사용된다.

Theorem 1.1 (Young's inequality) 임의의 $x, y \geq 0$와 $1/p + 1/q = 1$인 $p, q > 1$에 대해 $xy \leq x^p/p + y^q/q$이다. 이때, 등호가 성립할 필요충분조건은 $x^p = y^q$인 것이다.

Proof 임의의 $y \geq 0$를 고정하고 함수 $f : \mathbb{R} \to \mathbb{R}$를 $f : x \mapsto x^p/p + y^q/q - xy$로 두면 $f' : x \mapsto x^{p-1} - y$에서 f는 $y^{1/(1-p)} = y^{q/p}$에서 최솟값 $f(y^{q/p}) = (1/p + 1/q)y^q - y^{q/p+1} = 0$을 가진다. 여기서 마지막 등호는 $q/p + 1 = q$에서 성립하고, 증명은 이로써 충분하다. □

Theorem 1.2 (Hölder's inequality) 임의의 $x, y \in \mathbb{R}^n$와 $1/p + 1/q = 1$인 $p, q > 1$에 대해 $\sum_{i=1}^n |x_i y_i| \leq ||x||_p ||y||_q$이다. 이때, 등호가 성립할 필요충분조건은 적당한 $c \geq 0$가 존재하여 각 $i \leq n$에 대해 $|x_i|^p = c|y_i|^q$이거나 $|y_i|^q = c|x_i|^p$인 것이다.

Proof WLOG, 필요하다면 x, y를 각각 $|x|, |y|$로 대체하여 x, y의 모든 성분이 처음부터 음이 아니라 해도 된다. 한편, 만약 $||x||_p = 0$이면 $x = 0$에서 부등식이 자명하므로 $||x||_p > 0$이라 하고, 비슷한 이유로 $||y||_p > 0$이라 하자. 이제 $||x||_p = ||y||_q = 1$인 특별한 경우를 생각하면 Young의 부등식으로부터 $\sum_{i=1}^n x_i y_i \leq \sum_{i=1}^n (x_i^p/p + y_i^q/q) = ||x||_p^p/p + ||y||_q^q/q = 1/p + 1/q = 1$이다. 이제 일반적인 경우에 $\widetilde{x} = x/||x||_p, \widetilde{y} = y/||y||_q$라 하면 $||\widetilde{x}||_p = ||\widetilde{y}||_q = 1$이므로 앞선 결과로부터 $\sum_{i=1}^n \widetilde{x}_i \widetilde{y}_i \leq 1$이 되어 곧 $\sum_{i=1}^n x_i y_i \leq ||x||_p ||y||_q$임을 안다.

한편, 등호조건을 보이기 위해 $\sum_{i=1}^n |x_i y_i| = ||x||_p ||y||_q$라 하자. 만약 $||x||_p = 0$이면 $x = 0$에서 각 $i \leq n$에 대해 $|x_i|^p = 0 = 0 \cdot |y_i|^q$가 되어 더 이상 보일 것이 없으므로 $||x||_p \neq 0$

이라 하고, 비슷한 이유에서 $||y||_q \neq 0$이라 하자. 이제 $\widetilde{x} = |x|/||x||_p, \widetilde{y} = |y|/||y||_q$라 하면 $\sum_{i=1}^n \widetilde{x}_i\widetilde{y}_i = (\sum_{i=1}^n x_iy_i)/||x||_p||y||_q = 1 = ||\widetilde{x}||_p^p/p + ||\widetilde{y}||_q^q/q = \sum_{i=1}^n (\widetilde{x}_i^p/p + \widetilde{y}_i^q/q)$이므로 Young의 부등식으로부터 각 $i \le n$에 대해 $\widetilde{x}_i\widetilde{y}_i = \widetilde{x}_i^p/p + \widetilde{y}_i^q/q$이다. 이는 다시 Young의 부등식의 등호조건으로부터 각 $i \le n$에 대해 $\widetilde{x}_i^p = \widetilde{y}_i^q$임을 뜻하고, 따라서 $|x_i|^p = (||x||_p^p/||y||_q^q)|y_i|^q$이다. 역으로 적당한 $c \ge 0$가 존재하여 각 $i \le n$에 대해 $|x_i|^p = c|y_i|^q$이거나 $|y_i|^q = c|x_i|^p$인 경우에 등호가 성립함은 자명하므로 증명은 이로써 충분한다. □

Hölder의 부등식에서 조건으로 주어진 '$1/p + 1/q = 1$인 실수 $p, q > 1$' 는 다른 부등식에서도 조건으로 종종 주어지곤 하기에 특별히 이러한 $p, q \in \mathbb{R}$에 대해 **Hölder conjugate** 라는 이름이 붙어있다.

Theorem 1.3 (Minkowski's inequality) 임의의 $x, y \in \mathbb{R}^n$와 $p \ge 1$에 대해 $||x+y||_p \le ||x||_p + ||y||_p$이다. 이때, 등호가 성립한 필요충분조건은 적당한 $c \ge 0$가 존재하여 각 $i \le n$에 대해 $|x_i|^p = c|y_i|^q$이거나 $|y_i|^q = c|x_i|^p$인 것이다.

Proof 만약 $p = 1$이면 $||x+y||_1 = \sum_{i=1}^n |x_i + y_i| \le \sum_{i=1}^n |x_i| + \sum_{i=1}^n |y_i| = ||x||_1 + ||y||_1$에서 부등식이 자명하므로 $p > 1$이라 하자. 또한 $||x+y||_p = 0$인 경우에도 부등식이 자명하므로 $||x+y||_p > 0$이라 하자. 한편, Hölder의 부등식으로부터 $1/p + 1/q = 1$인 $q > 1$에 대해

$$\begin{aligned}
||x+y||_p^p &= \sum_{i=1}^n |x_i + y_i|^p \\
&\le \sum_{i=1}^n |x_i||x_i + y_i|^{p-1} + \sum_{i=1}^n |y_i||x_i + y_i|^{p-1} \\
&\le ||x||_p \left(\sum_{i=1}^n |x_i + y_i|^{(p-1)q} \right)^{1/q} + ||y||_p \left(\sum_{i=1}^n |x_i + y_i|^{(p-1)q} \right)^{1/q} \\
&= (||x||_p + ||y||_p) \left(\sum_{i=1}^n |x_i + y_i|^p \right)^{(p-1)/p} \\
&= (||x||_p + ||y||_p)||x+y||_p^{p-1}
\end{aligned}$$

이므로 $||x+y||_p \le ||x||_p + ||y||_p$임을 안다.

한편, 등호조건을 보이기 위해 $||x+y||_p = ||x||_p + ||y||_p$라 하자. 만약 $x = 0$이면 각 $i \le n$에 대해 $|x_i|^p = 0 = 0 \cdot |y_i|^p$에서 더 이상 보일 것이 없으므로 $x \neq 0$이라 하고, 비슷한 이유에서 $y \neq 0$이라 하자. 그렇다면 위의 증명과정과 Hölder의 부등식의 등호조건으로부터 적당한 $c_1, c_2 > 0$가 존재하여 $|x_i|^p = c_1|x_i + y_i|^{(p-1)q} = c_1|x_i + y_i|^p$이고 $|y_i|^p = c_2|x_i + y_i|^{(p-1)q} = c_2|x_i + y_i|^p$이므로 곧 $|x_i|^p = (c_1/c_2)c_2|x_i + y_i|^p = (c_1/c_2)|y_i|^p$이다. 역으로 적당한 $c \ge 0$가 존재하여 각 $i \le n$에 대해 $|x_i|^p = c|y_i|^q$이거나 $|y_i|^q = c|x_i|^p$인 경우에 등호가 성립함은 자명하므로 증명은 이로써 충분하다. □

1.1.2 Norms

노름은 벡터공간에서 정의되는 '크기' 의 추상화이다. 이미 해석개론 전반부에서 어느정도 논하였을 것이므로 여기서는 앞으로 사용될 몇몇 특별한 노름들을 간단히 소개하고 다변수 해석학에 필요한 작용소 노름과 행렬 노름에 대해 집중적으로 살펴본다.

Definition 1.4 벡터공간 V에서 정의된 함수 $||\cdot||:\mathsf{V}\to\mathbb{R}$가 임의의 $\lambda\in\mathbb{R}$와 임의의 $v,w\in\mathsf{V}$에 대해

i. (양의 동차성) $||\lambda v|| = |\lambda|||v||$.
ii. (삼각부등식) $||v+w|| \le ||v|| + ||w||$.
iii. (양의 정부호성) $||v|| = 0$일 필요충분조건은 $v = 0$인 것이다.

를 만족하면 이때의 함수 $||\cdot||$를 V 위의 **노름 (norm)**이라 하고, tuple $(\mathsf{V}, ||\cdot||)$를 **노름공간 (norm space)**이라 한다.

첫째로 소개할 노름은 $p-$노름으로 사실상 유클리드 공간에서의 표준으로 자리잡은 노름이다. 단순한 정의임에도 불구하고 그 기저에 깔린 훌륭한 기하학적 직관과 이가 가지는 좋은 성질들 덕분에 유클리드 공간을 넘어 수열공간이나 함수공간으로 그 정의가 확장되어 사용되지만, 무한차원 벡터공간에서는 조금 귀찮은 일들이 생기는 까닭에 여기서는 n차원 벡터공간에 한하여 $p-$노름을 도입한다. 이후 함수공간에서의 $p-$노름에 대해서는 측도론을 배우면서 L^p 공간의 개념을 통해 엄밀히 논의할 것이다. 특별히 정하지 않는 이상, $\mathbb{R}^n$에는 2−노름이 장착된 것으로 생각한다.

Definition 1.5 n차원 벡터공간 V와 임의의 $p\ge 1$에 대해 **$p-$노름 ($p-$norm)**을 $||\cdot||_p:\mathsf{V}\to\mathbb{R}_0^+$로 쓰고 $||\cdot||_p : v\mapsto(\sum_{i=1}^n |v_i|^p)^{1/p}$로 정의한다. 여기서 $v_1,\cdots,v_n$은 V의 고정된 기저에 대한 v의 좌표이다.

Proposition 1.6 n차원 벡터공간 V에 대해 $p-$노름은 노름이다. 따라서 $(\mathsf{V}, ||\cdot||_p)$는 노름공간을 이룬다.

Proof 양의 동차성과 양의 정부호성은 $p-$노름의 정의로부터 명백하고, 삼각부등식도 Minkowski의 부등식으로부터 자명하다. □

다음으로 소개할 노름은 $p-$노름만큼은 아니지만 각종 증명에서 이따금씩 등장하는 최댓값 노름이다. 아래의 정리에서 볼 수 있듯이 표기가 $p-$노름과 유사한 것은 우연의 일치가 아니며, 이러한 이유로 최댓값 노름을 $p-$노름의 $p=\infty$인 특별한 경우로 보는 경우도 있다.

Definition 1.7 n차원 벡터공간 V에 대해 **최댓값 노름 (maximum norm)**을 $||\cdot||_\infty:\mathsf{V}\to\mathbb{R}_0^+$로 쓰고 $||\cdot||_\infty : v\mapsto\max_{i=1}^n|v_i|$로 정의한다. 여기서 $v_1,\cdots,v_n$은 V의 고정된 기저에 대한 v의 좌표이다.

Proposition 1.8 n차원 벡터공간 V에 대해 maximum norm은 노름이다. 따라서 $(\mathsf{V}, ||\cdot||_\infty)$는 노름공간을 이룬다.

Proof 양의 동차성과 양의 정부호성은 maximum norm의 정의로부터 명백하므로 삼각부등식만 보이면 된다. 이를 위해 임의의 $v, w \in \mathsf{V}$를 생각하면 각 $i \le n$에 대해 $v_i + w_i \le ||v||_\infty + ||w||_\infty$이므로 $||v+w||_\infty \le ||v||_\infty + ||w||_\infty$가 성립한다. □

Theorem 1.9 n차원 벡터공간 V에 속하는 임의의 $v \in \mathsf{V}$에 대해 $p \to \infty$이면 $||v||_p \to ||v||_\infty$이다.

Proof 만약 $v = 0$이면 정리가 자명하므로 $v \neq 0$이라 하자. 한편, 간결한 논의를 위해 v의 모든 성분이 0이 아니라 하자. (다른 경우에 대해서도 이와 비슷하게 하면 된다.) 그렇다면 $||v||_\infty > 0$이고 L'Hôspital의 법칙으로부터

$$\begin{aligned}
\lim_{p\to\infty} \frac{||v||_p}{||v||_\infty} &= \lim_{p\to\infty} \left[\sum_{i=1}^n \left(\frac{|v_i|}{||v||_\infty} \right)^p \right]^{1/p} \\
&= \lim_{p\to\infty} \exp\left(\frac{1}{p} \log \sum_{i=1}^n \left(\frac{|v_i|}{||v||_\infty} \right)^p \right) \\
&= \exp\left(\lim_{p\to\infty} \frac{1}{p} \log \sum_{i=1}^n \left(\frac{|v_i|}{||v||_\infty} \right)^p \right) \\
&= \exp\left(\lim_{p\to\infty} \sum_{i=1}^n \left(\frac{|v_i|}{||v||_\infty} \right)^p \log\left(\frac{|v_i|}{||v||_\infty} \right) \Big/ \sum_{i=1}^n \left(\frac{|v_i|}{||v||_\infty} \right)^p \right) \\
&= \exp\left(\sum_{i=1}^n \left[\lim_{p\to\infty} \left(\frac{|v_i|}{||v||_\infty} \right)^p \log\left(\frac{|v_i|}{||v||_\infty} \right) \Big/ \sum_{i=1}^n \left(\frac{|v_i|}{||v||_\infty} \right)^p \right] \right)
\end{aligned}$$

인데, 각 $i \le n$에 대해 만약 $|v_i| = ||v||_\infty$이면 $\log(|v_i|/||v||_\infty) = 0$이고 그렇지 않으면 $|v_i| < |v_j|$인 $j \le n$가 적어도 하나 존재하여

$$\begin{aligned}
\lim_{p\to\infty} \left(\frac{|v_i|}{||v||_\infty} \right)^p \log\left(\frac{|v_i|}{||v||_\infty} \right) \Big/ \sum_{i=1}^n \left(\frac{|v_i|}{||v||_\infty} \right)^p &\le \lim_{p\to\infty} \left(\frac{|v_i|}{||v||_\infty} \right)^p \log\left(\frac{|v_i|}{||v||_\infty} \right) \Big/ \frac{|v_j|}{||v||_\infty} \\
&= \lim_{p\to\infty} \frac{\log(|v_i|/||v||_\infty)}{(|v_j|/|v_i|)^p} \\
&= 0
\end{aligned}$$

이므로 이상으로부터 $p \to \infty$이면 $||v||_p/||v||_\infty \to 1$임을 알고, 증명은 이로써 충분하다. □

무한차원 벡터공간으로 넘어가기 전에, 유한차원 벡터공간에 한해서만 성립하는 중요한 결과 하나를 소개한다. 이는 유한차원에서 사용할 노름의 선택을 단순한 개인의 취향 문제로 만들어버리는 강력한 결과이다.

Definition 1.10 벡터공간 V 위의 두 노름 $||\cdot||, ||\cdot||'$에 대해 $C_1 \le C_2$인 $C_1, C_2 > 0$가 존재하여 임의의 $v \in V$에 대해 $C_1||v|| \le ||v||' \le C_2||v||$이면 이때 $||\cdot||$와 $||\cdot||'$가 **equivalent**하다고 하고 $||\cdot|| \sim ||\cdot||'$으로 쓴다.

Proposition 1.11 벡터공간 V 위의 노름간의 equivalence는 동치관계이다. 즉, V 위의 노름 $||\cdot||, ||\cdot||', ||\cdot||''$에 대해 다음이 성립한다.

i. (반사성) $||\cdot|| \sim ||\cdot||$.
ii. (대칭성) $||\cdot|| \sim ||\cdot||'$이면 $||\cdot||' \sim ||\cdot||$이다.
iii. (전이성) $||\cdot|| \sim ||\cdot||'$이고 $||\cdot||' \sim ||\cdot||''$이면 $||\cdot|| \sim ||\cdot||''$이다.

Proof 반사성은 명백하므로 대칭성과 전이성만 보이자. 먼저 대칭성을 보이기 위해 $||\cdot|| \sim ||\cdot||'$라 하면 $C_1 \le C_2$인 $C_1, C_2 > 0$가 존재하여 임의의 $v \in V$에 대해 $C_1||v|| \le ||v||' \le C_2||v||$이므로 $(1/C_2)||v||' \le ||v|| \le (1/C_1)||v||'$에서 $||\cdot||' \sim ||\cdot||$이고, 곧 대칭성이 성립한다. 다음으로 전이성을 보이기 위해 $||\cdot|| \sim ||\cdot||'$이고 $||\cdot||' \sim ||\cdot||''$라 하면 $C_1 \le C_2$이고 $C_3 \le C_4$인 $C_1, C_2, C_3, C_4 > 0$가 존재하여 임의의 $v \in V$에 대해 $C_1||v|| \le ||v||' \le C_2||v||$이고 $C_3||v||' \le ||v||'' \le C_4||v||'$이므로 $C_1C_3||v|| \le C_3||v||' \le ||v||'' \le C_4||v||' \le C_2C_4||v||$에서 $||\cdot|| \sim ||\cdot||''$이고, 곧 전이성이 성립한다. □

Lemma 1.12 유한차원 노름공간 $(V, ||\cdot||_1)$ 위의 임의의 노름 $||\cdot||$은 균등연속이다.

Proof 벡터공간 V의 기저를 $\beta = \{\beta_1, \cdots, \beta_n\}$이라 하고 임의의 $\varepsilon > 0$을 택하면 $||v-w||_1 < \varepsilon / \max_{i=1}^n ||\beta_i||$인 임의의 $v, w \in V$에 대해

$$\begin{aligned}
|||v|| - ||w||| &\le ||v-w|| \\
&\le ||(v_1, v_2, \cdots, v_n) - (w_1, v_2, \cdots, v_n)|| + ||(w_1, v_2, \cdots, v_n) - (w_1, w_2, \cdots, v_n)|| \\
&\qquad + \cdots + ||(w_1, w_2, \cdots, v_n) - (w_1, w_2, \cdots, w_n)|| \\
&= \sum_{i=1}^n |v_i - w_i| ||\beta_i|| \\
&\le \sum_{i=1}^n |v_i - w_i| \max_{i=1}^n ||\beta_i|| \\
&= (\max_{i=1}^n ||\beta_i||) ||v-w||_1 \\
&< \varepsilon
\end{aligned}$$

이므로 $||\cdot||$은 균등연속이다. □

Theorem 1.13 유한차원 벡터공간 V 위의 임의의 두 노름 $||\cdot||, ||\cdot||'$에 대해 $||\cdot|| \sim ||\cdot||'$이다.

Proof 먼저 V에 1−노름을 장착하면 위의 보조정리로부터 V 위의 임의의 노름 $||\cdot||$이 연속이고, 집합 $S(1) \subseteq \mathsf{V}$이 compact하므로 $||\cdot||$은 $S(1)$ 위에서 최댓값 $M > 0$과 최솟값 $m > 0$을 가진다. 이제 0이 아닌 임의의 $v \in \mathsf{V}$에 대해 $\widetilde{v} = v/||v||_1$라 하면 $\widetilde{v} \in S(1)$이므로 $m \leq ||\widetilde{v}|| \leq M$에서 $m||v||_1 \leq ||v|| \leq M||v||_1$이고, 이는 $v = 0$인 경우에도 성립하므로 $||\cdot|| \sim ||\cdot||_1$임을 안다. 여기서 $||\cdot||$이 임의의 노름이라는 점과 명제 1.11를 생각해보면 증명은 이로써 충분하다. □

이제 작용소 노름에 대해 알아보자. 작용소 노름은 선형사상에 대해 정의되는 일종의 함수인데, 앞서 소개한 노름들과 달리 ∞를 그 값으로 가질 수 있어서 이를 진짜 노름으로 만들기 위해서는 약간의 준비가 필요하다.

Definition 1.14 노름공간 $(\mathsf{V}, ||\cdot||), (\mathsf{W}, ||\cdot||')$에 대해 **작용소 노름 (operator norm)**을 $||\cdot||_{\mathrm{op}} : \mathsf{L}(\mathsf{V},\mathsf{W}) \to \overline{\mathbb{R}}_0^+$로 쓰고 $||\cdot||_{\mathrm{op}} : T \mapsto \inf\{M \geq 0 :$ 임의의 $v \in \mathsf{V}$에 대해 $||T(v)||' \leq M||v||\}$로 정의한다.

Proposition 1.15 노름공간 $(\mathsf{V}, ||\cdot||), (\mathsf{W}, ||\cdot||')$ 사이에서 정의된 선형사상 $T, T' : \mathsf{V} \to \mathsf{W}$에 대해 다음이 성립한다.

i. 임의의 $\lambda \in \mathbb{R}$에 대해 $||\lambda T||_{\mathrm{op}} = |\lambda|||T||_{\mathrm{op}}$이다.
ii. $||T + T'||_{\mathrm{op}} \leq ||T||_{\mathrm{op}} + ||T'||_{\mathrm{op}}$.
iii. $||T||_{\mathrm{op}} = 0$일 필요충분조건은 $T = 0$인 것이다.

Proof iii. 만약 $T = 0$이면 $||T||_{\mathrm{op}} = 0$임이 자명하다. 이제 $||T||_{\mathrm{op}} = 0$인데 $T \neq 0$이라 하면 $T(v) \neq 0$인 0이 아닌 $v \in \mathsf{V}$를 적어도 하나 택할 수 있고 가정으로부터 적당한 $M < ||T(v)||'/||v||$이 존재하여 $||T(v)||' \leq M||v|| < ||T(v)||'$의 모순이 발생하므로 $T = 0$이다.

i. 만약 $\lambda = 0$이면 iii으로부터 명제가 자명하므로 $\lambda \neq 0$이라 하자. 또한, 만약 $||T||_{\mathrm{op}} = \infty$라면 임의의 $i \in \mathbb{N}$에 대해 적당한 $v_i \in \mathsf{V}$가 존재하여 $||T(v_i)||' > i||v_i||$이다. 이로부터 임의의 $j \in \mathbb{N}$에 대해 $i_j \geq j/|\lambda|$인 $\mathbb{N}$에 속하는 수열 $\{i_j\}$를 잡으면 임의의 $j \in \mathbb{N}$에 대해 $||\lambda T(v_{i_j})||' = |\lambda|||T(v_{i_j})||' > i_j|\lambda|||v_{i_j}|| \geq j||v_{i_j}||$가 되어 $||\lambda T||_{\mathrm{op}} = \infty$애서 이 경우에도 명제가 자명하므로 $||T||_{\mathrm{op}} < \infty$라 하자.

이제 임의의 $v \in \mathsf{V}$에 대해 $||T(v)||' \leq M||v||$인 $M \geq 0$을 임의로 하나 택하면 임의의 $v \in \mathsf{V}$에 대해 $||\lambda T(v)||' = |\lambda|||T(v)||' \leq |\lambda|M||v||$이므로 $||\lambda T||_{\mathrm{op}} \leq |\lambda|M$이고, 곧 $||\lambda T||_{\mathrm{op}} \leq |\lambda|||T||_{\mathrm{op}}$이다. 여기서 $\lambda \neq 0$가 임의였음을 상기한다면 $|\lambda|||T||_{\mathrm{op}} \leq ||\lambda T||_{\mathrm{op}}$도 성립하므로 이상을 종합하면 $||\lambda T||_{\mathrm{op}} = |\lambda|||T||_{\mathrm{op}}$이다.

ii. 만약 $||T||_{\mathrm{op}} = \infty$이거나 $||T'||_{\mathrm{op}} = \infty$이면 명제가 자명하므로 $||T||_{\mathrm{op}}, ||T'||_{\mathrm{op}} < \infty$라 하자. 이제 임의의 $v \in \mathsf{V}$에 대해 $||T(v)||' \leq M||v||$이고 $||T'(v)||' \leq N||v||$인 $M, N \geq 0$을 임의

로 하나씩 택하면 임의의 $v \in \mathsf{V}$에 대해 $||(T+T')(v)||' \leq ||T(v)||' + ||T'(v)||' \leq (M+N)||v||$ 이므로 $||T+T'||_{\text{op}} \leq M+N$ 이고, 곧 $||T+T'||_{\text{op}} \leq ||T||_{\text{op}} + ||T'||_{\text{op}}$ 이다. □

위의 명제를 보니 ∞의 작용소 노름을 가지는 선형사상들만 제거해주면 그 위에서 노름공간을 이룰 수 있을 듯하다.

Definition 1.16 노름공간 $(\mathsf{V}, ||\cdot||), (\mathsf{W}, ||\cdot||')$ 사이에서 정의된 선형사상 $T : \mathsf{V} \to \mathsf{W}$ 에 대해 $||T||_{\text{op}} < \infty$ 이면 이때의 선형사상 T 를 **유계 (bounded)**라 한다.

Theorem 1.17 노름공간 $(\mathsf{V}, ||\cdot||), (\mathsf{W}, ||\cdot||')$ 사이에서 정의된 선형사상 $T : \mathsf{V} \to \mathsf{W}$에 대해 T 가 유계일 필요충분조건은 이가 연속인 것이다.

Proof 먼저 충분조건임을 보이기 위해 T 가 유계라 하면 $||T||_{\text{op}} < \infty$ 이므로 적당한 $M > 0$ 이 존재하여 임의의 $v \in \mathsf{V}$에 대해 $||T(v)||' \leq M||v||$ 이다. 이제 임의의 $\varepsilon > 0$과 임의의 $v \in \mathsf{V}$를 택하면 $||v-w|| < \varepsilon/M$ 인 임의의 $w \in \mathsf{V}$에 대해 $||T(v) - T(w)||' = ||T(v-w)||' \leq M||v-w|| < \varepsilon$ 이므로 T 가 연속임을 안다. 다음으로 필요조건임을 보이기 위해 T 가 연속이라 하면 임의의 $v \in \mathsf{V}$에 대해 적당한 $\delta > 0$ 가 존재하여 $||v|| < \delta$ 이면 $||T(v)||' = ||T(v) - T(0)||' < 1$ 이다. 따라서 0이 아닌 임의의 $v \in \mathsf{V}$에 대해 $\widetilde{v} = \delta v/2||v||$ 라 하면 $||\widetilde{v}|| = \delta/2 < \delta$ 이므로 $(\delta/2||v||)||T(v)||' = ||T(\widetilde{v})||' < 1$ 에서 $||T(v)||' < (2/\delta)||v||$ 이고, 이는 $v = 0$ 인 경우에도 성립하므로 곧 $||T||_{\text{op}} \leq 2/\delta$ 에서 T 가 유계임을 안다. □

Proposition 1.18 노름공간 $(\mathsf{V}, ||\cdot||), (\mathsf{W}, ||\cdot||')$ 사이에서 정의된 모든 유계인 선형사상의 집합을 $\mathsf{BL}(\mathsf{V}, \mathsf{W}) \subseteq \mathsf{L}(\mathsf{V}, \mathsf{W})$ 라 하면 $(\mathsf{BL}(\mathsf{V}, \mathsf{W}), ||\cdot||_{\text{op}})$ 는 노름공간을 이룬다.

Proof 집합 $\mathsf{BL}(\mathsf{V}, \mathsf{W})$ 이 벡터공간임을 보일 수만 있으면 명제 1. 15로부터 명제가 자명한데, 정리 1. 17로부터 $\mathsf{BL}(\mathsf{V}, \mathsf{W})$ 의 모든 원소는 연속함수이고, 연속함수의 합과 스칼라곱은 연속함수라는 점에서 다시 정리 1. 17로부터 그 결과가 $\mathsf{BL}(\mathsf{V}, \mathsf{W})$ 에 속하여 $\mathsf{BL}(\mathsf{V}, \mathsf{W})$ 은 덧셈과 스칼라곱에 대해 닫혀있음을 안다. 이제 $\mathsf{BL}(\mathsf{V}, \mathsf{W})$ 이 벡터공간이 되기 위해 만족해야할 나머지 조건들은 쉽게 확인해볼 수 있다. □

만약 V가 유한차원이었다면 더 간단해진다.

Theorem 1.19 노름공간 $(\mathsf{V}, ||\cdot||), (\mathsf{W}, ||\cdot||')$ 사이에서 정의된 선형사상 $T : \mathsf{V} \to \mathsf{W}$에 대해 $\dim \mathsf{V} < \infty$ 이면 T 는 유계이고, 따라서 연속이다.

Proof 벡터공간 V 의 기저를 $\beta = \{\beta_1, \cdots, \beta_n\}$ 이라 하고 $M = \max_{i=1}^n ||T(\beta_i)||'$ 라 하자. 그렇다면 임의의 $v \in \mathsf{V}$에 대해

$$||T(v)||' = \left|\left|T\left(\sum_{i=1}^n v_i \beta_i\right)\right|\right|'$$

$$
\begin{aligned}
&= \left|\left| \sum_{i=1}^{n} v_i T(\beta_i) \right|\right|' \\
&\leq \sum_{i=1}^{n} |v_i| ||T(\beta_i)||' \\
&\leq M \sum_{i=1}^{n} |v_i| \\
&= M||v||_\infty
\end{aligned}
$$

인데, 정리 1. 13로부터 적당한 $C > 0$가 존재하여 $||v||_\infty \leq C||v||$이므로 $||T(v)||' \leq CM||v||$이다. 이는 곧 $||T||_{\text{op}} \leq CM < \infty$임을 뜻하므로 T는 유계이고, 정리 1. 17로부터 연속이다. □

Corollary 1.20 노름공간 $(\mathsf{V}, ||\cdot||)$, $(\mathsf{W}, ||\cdot||')$에 대해 $\dim \mathsf{V} < \infty$이면 $(\mathsf{L}(\mathsf{V}, \mathsf{W}), ||\cdot||_{\text{op}})$는 노름공간을 이룬다.

PROOF 이는 위의 정리와 명제 1. 18로부터 자명하다. □

마지막으로 작용소 노름의 중요한 성질 하나를 소개한다.

Theorem 1.21 노름공간 $(\mathsf{U}, ||\cdot||)$, $(\mathsf{V}, ||\cdot||')$, $(\mathsf{W}, ||\cdot||'')$에 대해 선형사상 $T : \mathsf{U} \to \mathsf{V}$, $T' : \mathsf{V} \to \mathsf{W}$에 대해 다음이 성립한다.

i. 임의의 $v \in \mathsf{U}$에 대해 $||T(v)||' \leq ||T||_{\text{op}}||v||$이다.
ii. $||T' \circ T||_{\text{op}} \leq ||T||_{\text{op}}||T'||_{\text{op}}$.

PROOF i. 우선 $v = 0$에 대해서는 정리가 자명하므로 $v \neq 0$이라 하자. 또한, $||T||_{\text{op}} = \infty$인 경우에도 정리가 자명하므로 $||T||_{\text{op}} < \infty$라 하자. 이제 임의의 $v \in \mathsf{V}$에 대해 $||T(v)||' \leq M||v||$인 $M \geq 0$을 임의로 하나 택하면 0이 아닌 임의의 $v \in \mathsf{V}$에 대해 $||T(v)||' \leq M||v||$에서 $||T(v)||'/||v|| \leq M$이므로 곧 $||T(v)||'/||v|| \leq ||T||_{\text{op}}$에서 증명이 끝난다.

ii. 만약 $||T||_{\text{op}} = 0$이거나 $||T'||_{\text{op}} = 0$이면 $T' \circ T = 0$이 되어 정리가 자명하므로 $||T||_{\text{op}}, ||T'||_{\text{op}} > 0$라 하자. 비슷하게, $||T||_{\text{op}} = \infty$이거나 $||T'||_{\text{op}} = \infty$인 경우에도 정리가 자명하므로 $||T||_{\text{op}}, ||T'||_{\text{op}} < \infty$라 하자. 이제 임의의 $v \in \mathsf{V}$에 대해 $||T(v)||' \leq M||v||$이고 $||T'(v)||'' \leq N||v||$인 $M, N > 0$을 임의로 하나씩 택하면 임의의 $v \in \mathsf{V}$에 대해 $||(T' \circ T)(v)||'' \leq N||T(v)||' \leq MN||v||$이므로 $||T' \circ T||_{\text{op}} \leq MN$에서 $||T' \circ T||_{\text{op}}/N \leq M$이고, 곧 $||T' \circ T||_{\text{op}}/N \leq ||T||_{\text{op}}$이다. 비슷하게, 이는 다시 $||T' \circ T||_{\text{op}}/||T||_{\text{op}} \leq N$임을 의미하여 $||T' \circ T||_{\text{op}}/||T||_{\text{op}} \leq ||T'||_{\text{op}}$이 성립하고, 곧 $||T' \circ T||_{\text{op}} \leq ||T||_{\text{op}}||T'||_{\text{op}}$가 되어 증명이 끝난다. □

마지막으로 살펴볼 노름은 행렬 노름이다. 행렬에 노름을 주는 방식은 크게 두 가지가 있는데, 우선 $\mathsf{M}_{m,n}$을 $L(\mathbb{R}^n, \mathbb{R}^m)$과 동일시하여 행렬을 선형사상으로 파악하고, 이에 방금 소개한 작용소 노름을 주는 방식을 소개한다.

Definition 1.22 행렬 $A \in \mathsf{M}_{m,n}$에 대해 이를 p−노름 (혹은 최댓값 노름) 이 장착된 $(\mathbb{R}^n, ||\cdot||_p)$와 $(\mathbb{R}^m, ||\cdot||_p)$ 사이에서 정의된 선형사상 $T_A : \mathbb{R}^n \to \mathbb{R}^m$로 생각하여 **$p$−노름 (혹은 최댓값 노름) 으로부터 유도된 작용소 노름 (operator norm induced by p-norm(resp. maximum norm))** 을 $||\cdot||_p : \mathsf{M}_{m,n} \to \mathbb{R}_0^+$로 쓰고 $||\cdot||_p : A \mapsto ||T_A||_{\text{op}}$로 정의한다.

Proposition 1.23 모든 $m \times n$ 행렬들의 집합 $\mathsf{M}_{m,n}$에 대해 p−노름으로부터 유도된 작용소 노름은 노름이다. 따라서 $(\mathsf{M}_{m,n}, ||\cdot||_p)$는 노름공간을 이룬다.

Proof 임의의 $A \in \mathsf{M}_{m,n}$에 대해 선형사상 T_A가 연속임이 자명하므로 정리 1.17로부터 이는 자명하다. □

p−노름 (혹은 최댓값 노름) 으로부터 유도된 작용소 노름은 본질적으로 작용소 노름이므로 정리 1.21가 그대로 성립한다.

Theorem 1.24 임의의 $A \in \mathsf{M}_{l,m}$와 $B \in \mathsf{M}_{m,n}$에 대해 다음이 성립한다.

i. 임의의 $x \in \mathbb{R}^m$에 대해 $||Ax||_p \leq ||A||_p ||x||_p$ (혹은 $||Ax||_\infty \leq ||A||_\infty ||x||_\infty$) 이다.
ii. $||AB||_p \leq ||A||_p ||B||_p$. (혹은 $||AB||_\infty \leq ||A||_\infty ||B||_\infty$.)

Proof 이는 p−노름 (혹은 최댓값 노름) 으로부터 유도된 작용소 노름의 정의와 정리 1.21로부터 자명하다. □

특별히, 1−노름으로부터 유도되거나 최댓값 노름으로부터 유도된 경우에는 그로부터 유도된 작용소 노름을 간단하게 구할 수 있다. 작용소 노름의 정의가 행렬의 성분과는 적어도 직접적으로는 관계가 없다는 점을 생각해보면 아래 정리는 조금 신기한 결과이다.

Theorem 1.25 임의의 $A \in \mathsf{M}_{m,n}$에 대해 다음이 성립한다.

i. $||A||_1 = \max_{j=1}^n \sum_{i=1}^m |a_{ij}|$.
ii. $||A||_\infty = \max_{i=1}^m \sum_{j=1}^n |a_{ij}|$.

Proof i. 임의의 $x \in \mathbb{R}^n$에 대해

$$\begin{aligned}
||T_A(x)||_1 &= ||Ax||_1 \\
&= \sum_{i=1}^m \left| \sum_{j=1}^n a_{ij}x_j \right| \\
&\leq \sum_{i=1}^m \sum_{j=1}^n |a_{ij}x_j| \\
&= \sum_{j=1}^n \left(\sum_{i=1}^m |a_{ij}| \right) |x_j|
\end{aligned}$$

$$
\begin{aligned}
&\le \left(\max_{j=1}^{n}\sum_{i=1}^{m}|a_{ij}|\right)\sum_{j=1}^{n}|x_j| \\
&= \left(\max_{j=1}^{n}\sum_{i=1}^{m}|a_{ij}|\right)||x||_1
\end{aligned}
$$

이므로 $||A||_1 = ||T_A||_{\text{op}} \le \max_{j=1}^{n}\sum_{i=1}^{m}|a_{ij}|$ 이다. 한편, 만약 임의의 $x \in \mathbb{R}^n$ 에 대해 $||T_A(x)||_1 \le M||x||_1$ 인 $M < \max_{j=1}^{n}\sum_{i=1}^{m}|a_{ij}|$ 이 존재한다면 $j_0 = \operatorname{argmax}_{j=1}^{n}\sum_{i=1}^{m}|a_{ij}|$ 에 대해

$$
\begin{aligned}
\max_{j=1}^{n}\sum_{i=1}^{m}|a_{ij}| &= \sum_{i=1}^{n}|a_{ij_0}| \\
&= ||A\mathbf{e}_{j_0}||_1 \\
&= ||T_A(\mathbf{e}_{j_0})||_1 \\
&\le M||\mathbf{e}_{j_0}|| \\
&= M \\
&< \max_{j=1}^{n}\sum_{i=1}^{m}|a_{ij}|
\end{aligned}
$$

에서 모순이 발생하므로 이러한 M 은 존재하지 않고, 곧 $||A||_1 = ||T_A||_{\text{op}} = \max_{j=1}^{n}\sum_{i=1}^{m}|a_{ij}|$ 이다.

ii. 만약 $A = 0$ 이면 정리가 $||A||_\infty = 0$ 이 되어 정리가 자명하므로 $A \neq 0$ 이라 하자. 그렇다면 임의의 $x \in \mathbb{R}^n$ 에 대해

$$
\begin{aligned}
||T_A(x)||_\infty &= ||Ax||_\infty \\
&= \max_{i=1}^{m}\left|\sum_{j=1}^{n}a_{ij}x_j\right| \\
&\le \max_{i=1}^{m}\sum_{j=1}^{n}|a_{ij}x_j| \\
&\le \left(\max_{i=1}^{m}\sum_{j=1}^{n}|a_{ij}|\right)\max_{j=1}^{n}|x_j| \\
&= \left(\max_{i=1}^{m}\sum_{j=1}^{n}|a_{ij}|\right)||x||_\infty
\end{aligned}
$$

이므로 $||A||_\infty = ||T_A||_{\text{op}} \le \max_{i=1}^{m}\sum_{j=1}^{n}|a_{ij}|$ 이다. 한편, 만약 임의의 $x \in \mathbb{R}^n$ 에 대해 $||T_A(x)||_\infty \le M||x||_\infty$ 인 $M < \max_{i=1}^{m}\sum_{j=1}^{n}|a_{ij}|$ 이 존재한다면 $i_0 = \operatorname{argmax}_{i=1}^{m}\sum_{j=1}^{n}|a_{ij}|$ 와 $x_0 = (\operatorname{sgn}(a_{i_0j}))$ 에 대해 x_0 의 성분 중 적어도 하나는 0 이 아니므로 $||x_0||_\infty = 1$ 이 되어

$$\begin{aligned}
\max_{i=1}^{m} \sum_{j=1}^{n} |a_{ij}| = \max_{i=1}^{m} \left| \sum_{j=1}^{n} \operatorname{sgn}(a_{i_0 j}) a_{ij} \right| \\
&= ||Ax_0||_\infty \\
&= ||T_A(x_0)||_\infty \\
&\leq M||x_0||_\infty \\
&= M \\
&< \max_{i=1}^{m} \sum_{j=1}^{n} |a_{ij}|
\end{aligned}$$

에서 모순이 발생하므로 이러한 M 은 존재하지 않고, 곧 $||A||_\infty = ||T_A||_{\mathrm{op}} = \max_{i=1}^m \sum_{j=1}^n |a_{ij}|$ 이다. 위의 식에서 첫번째 등호는 각 $i \leq m$ 에 대해 $|\sum_{j=1}^n \operatorname{sgn}(a_{i_0 j}) a_{ij}| \leq \sum_{j=1}^n |a_{ij}|$ 이고 $\max_{i=1}^m \sum_{j=1}^n |a_{ij}| = \sum_{j=1}^n |a_{i_0 j}| = \sum_{j=1}^n \operatorname{sgn}(a_{i_0 j}) a_{i_0 j}$ 이므로 성립한다. □

행렬에 노름을 주는 또다른 방식은 $\mathsf{M}_{m,n}$ 을 $\mathbb{R}^{mn}$ 과 동일시하여 행렬을 벡터로 파악하고, 이에 p−노름이나 최댓값 노름을 주는 방식이다. 사실 $\mathbb{R}^{mn}$ 에서 정의되는 어떤 노름이든 줄 수 있지만, 2−노름을 주는 경우가 일반적이며, 여기서도 이 경우만 살펴본다.

Definition 1.26 행렬 $A \in \mathsf{M}_{m,n}$ 에 대해 이를 2−노름이 장착된 $(\mathbb{R}^{m+n}, ||\cdot||)$ 에 속하는 벡터 $\mathrm{vec}A \in \mathbb{R}^{m+n}$, 로 생각하여 **Frobenius 노름 (− norm)** 을 $||\cdot||_{\mathrm{F}} : \mathsf{M}_{m,n} \to \mathbb{R}_0^+$ 로 쓰고 $||\cdot||_{\mathrm{F}} : A \mapsto ||\mathrm{vec}A||_2$ 로 정의한다.

신기하게도, Frobenius 노름은 선형사상으로서의 행렬과는 그다지 관련이 없음에도 불구하고 정리 1. 21과 유사한 성질들을 가진다.

Theorem 1.27 임의의 $A \in \mathsf{M}_{l,m}$ 와 $B \in \mathsf{M}_{m,n}$ 에 대해 다음이 성립한다.

i. 임의의 $x \in \mathbb{R}^m$ 에 대해 $||Ax|| \leq ||A||_{\mathrm{F}}||x||$ 이다.
ii. $||AB||_{\mathrm{F}} \leq ||A||_{\mathrm{F}}||B||_{\mathrm{F}}$.

Proof i. Cauchy-Schwarz의 부등식으로부터

$$\begin{aligned}
||Ax|| &= \sqrt{\sum_{i=1}^{m} \left(\sum_{j=1}^{n} a_{ij} x_j \right)^2} \\
&\leq \sqrt{\sum_{i=1}^{m} \left(\sum_{j=1}^{n} a_{ij}^2 \right) \left(\sum_{j=1}^{n} x_j^2 \right)} \\
&= ||A||_{\mathrm{F}}||x||
\end{aligned}$$

이므로 이는 자명하다.

ii. 행렬 B의 각 열을 $b_1, \cdots, b_n \in \mathbb{R}^m$이라 하면 $AB = [Ab_1, \cdots, Ab_n]$이고 곧 i로부터

$$\begin{aligned}
||AB||_{\mathrm{F}} &= \sqrt{\sum_{i=1}^{n} ||Ab_i||^2} \\
&\leq \sqrt{\sum_{i=1}^{n} ||A||_{\mathrm{F}}^2 ||b_i||^2} \\
&= ||A||_{\mathrm{F}} \sqrt{\sum_{i=1}^{n} ||b_i||^2} \\
&= ||A||_{\mathrm{F}} ||B||_{\mathrm{F}}
\end{aligned}$$

이다. □

Part II

Probability and Measure

Chapter 2

Measure Theory

Abstract 측도란, $\mathbb{R}^2$ 에서의 넓이나 $\mathbb{R}^3$ 에서의 부피를 임의의 공간으로 일반화한 개념으로, 그 용어가 뜻하는 사전적인 의미와 같이 어떤 집합을 재는 (measure) 함수이다. 언뜻 보기에 이러한 추상적인 개념으로 수리통계학을 시작하는 것이 부자연스러워 보일 수 있다. 그럼에도 불구하고, 우리가 이 책을 측도론으로 시작하는 것에는 크게 두 가지 이유가 있다. 우선 첫째로 확률이 그 자체로 측도이기 때문이다. 러시아의 대수학자 Kolmogorov가 정립한 현대의 확률론에서는 확률을 '사건' 이라는 집합에 0과 1사이의 수를 부여하는 측도로 파악한다. 따라서 측도는 확률의 본질이 되어 이를 알지 아니하고서는 확률에 대한 엄밀한 이론 전개가 어렵다. 둘째 이유는 보다 더 일반적인 것으로, 우리가 기존에 사용하던 적분론을 조금 손볼 필요가 있기 때문이다. 우리가 여태까지 사용해오던 적분은 Riemann 적분론인데, Riemann 적분론에서는 적분과 극한의 순서를 바꾸기 위해 균등수렴과 같은 상당히 강한 조건들이 필요하다. 이렇게 적분과 극한이 서로 매끄럽게 상호작용하지 못한다는 점은 Riemann 적분론의 단점으로, 이를 보완해 줄 수 있는 새로운 적분론이 필요한데, 그것이 바로 Lebesgue 적분론이다. 그리고 Lebesgue 적분론이 측도론을 기초로 하기에 우리는 측도론을 배워야 한다.

2.1 Measurable Spaces

본 장 전반에 걸쳐, 우리는 Lebesgue 측도를 구성할 것이다. Lebesgue 측도는 우리가 직관적으로 파악하는 넓이나 부피에 대응하는 $\mathbb{R}^n$ 위의 측도로 $n = 3$ 인 경우에 점이나 선에는 0의 측도를, unit cube에는 1의 측도를, 유계가 아닌 cube에는 무한대의 측도를 부여하는 이상적인 측도이다. 그러나 겉보기와는 달리, Lebesgue 측도를 구성하는 일은 생각보다 만만치 않다. 이는 놀랍게도 모든 집합에 이러한 측도를 부여하는 것이 일반적으로는 불가능하기 때문이다. 즉, $\mathbb{R}^n$ 의 부분집합이라고 해서 모두 잴 수 있는 것이 아니다. 이런 '잴

수 없는' 집합은 분명히 존재하며, 여기서는 다루지 않겠지만 이러한 집합을 직접 구성할 수도 있다. [1]따라서 우리는 Lebesgue 측도를 구성하기에 앞서 우선 측도를 부여할 $\mathbb{R}^n$ 의 부분집합들을 먼저 선정해야 한다. 그렇다면 과연 측도를 부여할 부분집합들을 어떻게 골라야 할까? 너무 많이 고르면 우리가 원하는 Lebesgue 측도의 이상적인 성질을 유지하기 힘들어지고, 그렇다고 너무 적게 고르면 실용성이 떨어지는 측도가 되고 만다. Lebesgue 측도를 구성한다고 함은, 위의 질문에 대한 아름다운 답을 찾아나가는 과정이라 볼 수 있다. 그 첫 번째 단계로 이번 절에서는 집합들의 모임인 집합족을 다루는 법을 배운다.

Definition 2.1 공집합이 아닌 집합 X 에 대해 $\mathscr{A}$ 를 X 의 부분집합의 모임이라 하고 비어있지 않다고 하자. 만약 $\mathscr{A}$ 가

i. 임의의 $A, B \in \mathscr{A}$ 에 대해 $A \cup B \in \mathscr{A}$ 이다.
ii. 임의의 $A \in \mathscr{A}$ 에 대해 $A^c \in \mathscr{A}$ 이다.

를 만족하면 이때의 $\mathscr{A}$ 를 X 위의 **대수 (algebra)** 혹은 **체 (field)**라 한다. 나아가, 만약 $\mathscr{A}$ 에 속하는 임의의 수열 $\{A_i\}$ 에 대해 $\bigcup_{i=1}^{\infty} A_i \in \mathscr{A}$ 이면 이때의 $\mathscr{A}$ 를 X 위의 **$\sigma-$대수 ($\sigma-$algebra)** 혹은 **$\sigma-$체 ($\sigma-$field)**라 한다.

Proposition 2.2 공집합이 아닌 집합 X 위의 대수 $\mathscr{A}$ 에 대해 다음이 성립한다.

i. $\emptyset, X \in \mathscr{A}$.
ii. 임의의 $A, B \in \mathscr{A}$ 에 대해 $A \cap B, A \setminus B \in \mathscr{A}$ 이다.

나아가, 만약 $\mathscr{A}$ 가 $\sigma-$대수이면 $\mathscr{A}$ 에 속하는 임의의 집합열 $\{A_i\}$ 에 대해 $\bigcap_{i=1}^{\infty} A_i \in \mathscr{A}$ 이다.

Proof ii. 이는 $A \cap B = (A^c \cup B^c)^c$ 와 $A \setminus B = A \cap B^c = (A^c \cup B)^c$ 에서 자명하다.

i. 집합족 $\mathscr{A}$ 가 비어있지 않으므로 적어도 하나의 원소 $A \in \mathscr{A}$ 를 택할 수 있다. 그렇다면 ii로부터 $\emptyset = A \setminus A \in \mathscr{A}$ 이고 $X = A \cup A^c \in \mathscr{A}$ 이다.

한편, $\mathscr{A}$ 가 $\sigma-$대수인 경우를 생각한다면, $\mathscr{A}$ 에 속하는 임의의 집합열 $\{A_i\}$ 에 대해 $\{A_i^c\}$ 도 $\mathscr{A}$ 에 속하므로 $\bigcap_{i=1}^{\infty} A_i = (\bigcup_{i=1}^{\infty} A_i^c)^c \in \mathscr{A}$ 이다. □

정의로부터, 대수는 교집합, 합집합, 차집합, 여집합에 대해 모두 닫혀있으며, 수학적 귀납법을 사용하면 이러한 연산을 유한번 적용하는 경우에도 여전히 닫혀있음을 쉽게 알 수 있다. 이에 더하여, $\sigma-$대수의 경우 이러한 연산을 가산번 반복 적용하는 경우에도 닫혀있다. 우리가 아는 집합의 연산이 대부분이 위의 네 가지 연산의 조합으로 이루어진다는 사실을 생각해보면, 대수는 굉장히 큰 모임이라고 생각할 수 있다. 실제로 대수에 속한 집합들을 연산하여 얻는 어지간한 결과들은 모두 그 대수 안에 다시 속한다.

Definition 2.3 공집합이 아닌 집합 X 위의 $\sigma-$대수 $\mathscr{A}$ 에 대해 tuple $(X, \mathscr{A})$ 를 **가측공간 (measurable space)**이라 한다. 또한, $\mathscr{A}$ 에 속하는 임의의 집합을 **($\mathscr{A}-$) 가측집합 (($\mathscr{A}-$)measurable set)**이라 한다.

갑자기 가측공간의 정의가 덜컥 주어지니 당황스럽기 그지없을 것이다. 아직 우리는 측도가 무엇인지도 확실히 모르므로 이는 당연한 반응이다. 지금으로서는, 그냥 위와 같은 정의를 만족하는 tuple을 가측공간이라 부른다는 사실만 받아들이고, 일단 용어에 익숙해지도록 하자. 이후에 Lebesgue 측도를 본격적으로 구성하게 되면 가측공간은 자연스럽게 주인공으로 떠오르게 된다. 한편, 가측공간을 정의함에 있어서 측도라는 개념이 직접적으로 필요하지 않다는 점도 기억해둘 필요가 있다. (이후 가측공간에 측도를 더해 만들어지는 측도공간과 가측공간을 혼동하면 안된다.) 가측공간의 정의에서 유일하게 필요한 개념은 $\sigma-$대수 뿐이다. 이는 $\sigma-$대수가 측도론에서 중요한 역할을 한다는 점을 시사하기도 한다.

Definition 2.4 공집합이 아닌 집합 X의 부분집합의 모임 $\mathscr{C}$에 대해 $\mathscr{C}$를 포함하는 최소의 $\sigma-$대수를 **$\mathscr{C}$가 생성하는 $\sigma-$대수 ($\sigma-$algebra generated by $\mathscr{C}$)**라 하고 $\sigma(\mathscr{C})$로 쓴다. 여기서 $\mathscr{C}$는 $\sigma(\mathscr{C})$의 **생성자 (generator)**라 한다.

Proposition 2.5 공집합이 아닌 집합 X의 부분집합의 모임 $\mathscr{C}$에 대해 이가 생성하는 $\sigma-$대수 $\sigma(\mathscr{C})$는 유일하게 존재한다. 따라서, $\sigma(\mathscr{C})$는 well-define된다.

Proof $\sigma(\mathscr{C})$가 존재하기만 한다면, 이의 유일성은 그 정의로부터 자명하다. 존재성을 보이기 위해 Σ를 $\mathscr{C}$를 포함하는 모든 $\sigma-$대수의 모임이라 하면 $\mathscr{P}(X)$가 $\mathscr{C}$를 포함하는 $\sigma-$대수임이 분명하므로 Σ는 비어있지 않다. 이제 이의 교집합인 $\bigcap\Sigma$가 $\sigma-$대수라 주장한다. 이를 보이기는 어렵지 않다. 정의 2. 1의 조건을 하나하나 따져보면 되는데, 연습삼아 두 번째 조건을 보자. 임의의 $A \in \bigcap\Sigma$에 대해 A는 Σ에 속하는 모든 $\sigma-$대수에 속하므로 곧 A^c도 Σ에 속하는 모든 $\sigma-$대수에 속하고, 이는 곧 $A^c \in \bigcap\Sigma$임을 뜻한다. 다른 조건들도 이와 같이 쉽게 따져볼 수 있다. 그렇다면 $\bigcap\Sigma$이 바로 $\sigma(\mathscr{C})$이다. □

당연히, 어떤 $\sigma-$대수를 생성하는 생성자는 유일하지 않다. 즉, 같은 $\sigma-$대수를 생성하는 서로다른 다양한 집합족이 있을 수 있다.

Definition 2.6 공집합이 아닌 집합 X에 대해 $\mathscr{P}$를 X의 부분집합의 모임이라 하고 비어있지 않다고 하자. 만약 임의의 $A, B \in \mathscr{P}$에 대해 $A \cap B \in \mathscr{P}$가 성립하면 이때의 $\mathscr{P}$를 X 위의 **$\pi-$system**이라 한다.

Definition 2.7 공집합이 아닌 집합 X에 대해 $\mathscr{L}$을 X의 부분집합의 모임이라 하자. 만약 $\mathscr{L}$이

i. $X \in \mathscr{L}$.
ii. 임의의 $A \in \mathscr{L}$에 대해 $A^c \in \mathscr{L}$이다.
iii. 집합족 $\mathscr{L}$에 속하는 임의의 서로소인 집합열 $\{A_i\}$에 대해 $\bigsqcup_{i=1}^{\infty} A_i \in \mathscr{L}$이다.

를 만족하면 이때의 $\mathscr{L}$을 X 위의 **$\lambda-$system**, **$d-$system** 혹은 **Dynkin class**라 한다.[2]

정의로부터, π−system과 λ−system은 σ−대수보다 더 단순한 구조로 이루어져 있다. π−system은 교집합에 대해서만 닫혀 있고, λ−system은 여집합과 가산번의 서로소 합집합에 대해서만 닫혀 있다. 물론 이전과 같이 수학적 귀납법을 사용하면 각각 유한번의 교집합과 여집합에 대해서도 닫혀 있음을 쉽게 알 수 있지만 그래도 σ−대수의 구조에 비하면 π−system과 λ−system의 구조는 한참 단순하다. 그러나 이 둘과 σ−대수는 Dynkin의 π−λ 정리에 의해 아름답게 연결된다.

Definition 2.8 공집합이 아닌 집합 X의 부분집합의 모임 $\mathscr{C}$에 대해 $\mathscr{C}$를 포함하는 최소의 λ−system을 $\mathscr{C}$ **가 생성하는 λ−system(− generated by** $\mathscr{C}$)라 하고 $\lambda(\mathscr{C})$로 쓴다. 여기서 $\mathscr{C}$는 $\lambda(\mathscr{C})$의 **생성자 (generator)**라 한다.

Proposition 2.9 공집합이 아닌 집합 X의 부분집합의 모임 $\mathscr{C}$에 대해 이가 생성하는 λ−system $\lambda(\mathscr{C})$는 유일하게 존재한다. 따라서, $\lambda(\mathscr{C})$는 well-define된다.

Proof $\lambda(\mathscr{C})$가 존재하기만 한다면, 이의 유일성은 그 정의로부터 자명하다. 존재성을 보이기 위해 Λ를 $\mathscr{C}$를 포함하는 모든 λ−system의 모임이라 하면 $\mathscr{P}(X)$가 $\mathscr{C}$를 포함하는 λ−system임이 분명하므로 Λ는 비어있지 않다. 이제 이의 교집합인 $\bigcap\Lambda$가 λ−system이라 주장한다. 이를 보이기는 어렵지 않다. 정의 2.7의 조건을 하나하나 따져보면 되는데, 연습삼아 첫 번째 조건을 보자. 임의의 $A\in\bigcap\Lambda$에 대해 A는 Λ에 속하는 모든 λ−system에 속하므로 곧 A^c도 Λ에 속하는 모든 λ−system에 속하고, 이는 곧 $A^c\in\bigcap\Lambda$임을 뜻한다. 다른 조건들도 이와 같이 쉽게 따져볼 수 있다. 그렇다면 $\bigcap\Lambda$이 바로 $\lambda(\mathscr{C})$이다. □

Lemma 2.10 공집합이 아닌 집합 X의 부분집합의 모임 $\mathscr{A}$가 π−system인 동시에 λ−system이라면 이는 σ−대수이다.

Proof 집합족 $\mathscr{A}$가 가산 합집합에 대해 닫혀있음을 보이면 충분하다. 즉, $\mathscr{A}$에 속하는 임의의 집합열 $\{A_i\}$에 대해 $\bigcup_{i=1}^{\infty}A_i\in\mathscr{A}$임을 보이면 된다. 이를 위해 집합열 $\{B_i\}$를 $B_i := A_i\setminus\bigcup_{j=1}^{i-1}A_j = A_i\cap\bigcap_{j=1}^{i-1}A_j^c$로 정의하면 $\mathscr{A}$가 유한번의 교집합과 여집합에 대해 닫혀 있으므로 이는 명백히 $\mathscr{A}$에 속하는 서로소인 집합열이다. 또한, $\bigcup_{i=1}^{\infty}A_i=\bigsqcup_{i=1}^{\infty}B_i$이므로 $\bigcup_{i=1}^{\infty}A_i\in\mathscr{A}$가 되어 원하는 결론을 얻는다. □

Theorem 2.11 (Dynkin's π-λ theorem) 공집합이 아닌 집합 X 위의 π−system $\mathscr{P}$에 대해 $\sigma(\mathscr{P})=\lambda(\mathscr{P})$이다.

Proof $\lambda(\mathscr{P})\subseteq\sigma(\mathscr{P})$임이 분명하므로 역의 포함관계만 보이면 되고, 다시 위의 보조정리로부터 $\lambda(\mathscr{P})$가 π−system임을 보이는 것으로 충분하다. 이를 위해 임의의 $A\in\lambda(\mathscr{P})$를 택하여 집합 $\lambda_A=\{B\subseteq X : A\cap B\in\lambda(\mathscr{P})\}$를 생각하고 $\lambda(\mathscr{P})\subseteq\lambda_A$임을 보이자. 우선 $A\in\lambda_A$이므로 $X\in\lambda_A$임은 분명하고, 임의의 $B\in\lambda_A$에 대해 $A\cap B^c=A\setminus(A\cap B)$이므로

$B^c \in \lambda_A$ 이다. 마지막으로 λ_A 에 속하는 임의의 서로소인 집합열 $\{B_i\}$ 에 대해 $\{A \cap B_i\}$ 가 $\lambda(\mathscr{P})$ 에 속하는 서로소인 집합열이므로 $\bigsqcup_{i=1}^{\infty} B_i \in \lambda_A$ 가 성립하여 λ_A 가 $\lambda-$system임을 안다. 이상의 결론이 임의의 $A \in \lambda(\mathscr{P})$ 에 대해 성립함을 상기한다면, 임의의 $B \in \mathscr{P}$ 에 대해 λ_B 도 $\lambda-$system인데, λ_B 의 정의에서 $\mathscr{P} \subseteq \lambda_B$ 임이 분명하므로 곧 $\lambda(\mathscr{P}) \subseteq \lambda_B$ 이고, 따라서 $A \in \lambda_B$ 이다. 그렇다면 다시 λ_B 의 정의에서 $B \in \lambda_A$ 이며, 곧 $\mathscr{P} \subseteq \lambda_A$ 가 되고, 증명은 이로써 충분하다. □

앞서 대수가 꽤나 큰 집합족이라는 사실을 설명한 바 있는데, 이는 마치 양날의 검과 같다. 우선 그 크기가 충분히 커서 우리가 필요로 하는 대부분의 집합들을 포함시킬 수 있으며, 어지간한 연산에 대해 닫혀있다는 점은 분명 대수의 장점이다. 그러나 대수의 이런 거대한 크기 때문에 때로는 이를 직접 다루기 힘든 경우도 있다. 이러한 어려움을 극복하기 위해 중간다리의 역할을 하는 구조가 필요한데, 바로 semi-algebra이다.

Definition 2.12 공집합이 아닌 집합 X 에 대해 $\mathscr{A}$ 를 X 의 부분집합의 모임이라 하자. 만약 $\mathscr{A}$ 가

i. $X \in \mathscr{A}$.
ii. 임의의 $A, B \in \mathscr{A}$ 에 대해 $A \cap B \in \mathscr{A}$ 이다.
iii. 임의의 $A \in \mathscr{A}$ 에 대해 A^c 를 $\mathscr{A}$ 에 속하는 서로소인 집합들의 유한 합집합으로 쓸 수 있다.

를 만족하면 이때의 $\mathscr{A}$ 를 X 위의 **semi-algebra**라 한다.

Proposition 2.13 공집합이 아닌 집합 X 위의 semi-algebra $\mathscr{A}$ 에 대해 다음이 성립한다.

i. $\emptyset \in \mathscr{A}$.
ii. 임의의 $A, B \in \mathscr{A}$ 에 대해 $A \setminus B$를 $\mathscr{A}$ 에 속하는 서로소인 집합들의 유한 합집합으로 쓸 수 있다.

Proof i. $X \in \mathscr{A}$ 이므로 $\emptyset = X^c$ 는 적당한 서로소인 집합들 $B_1, \cdots, B_k \in \mathscr{A}$ 에 대해 이들의 서로소 합집합으로 쓸 수 있는데, 이는 곧 $B_1 = \cdots = B_k = \emptyset$ 임을 뜻하고, 따라서 $\emptyset \in \mathscr{A}$ 이다.

ii. 적당한 서로소인 집합들 $C_1, \cdots, C_k \in \mathscr{A}$ 에 대해 $B^c = \bigsqcup_{i=1}^{k} C_i$ 이고 여기서 각 $i \leq k$ 에 대해 $A \cap C_i \in \mathscr{A}$ 이며 이들이 서로소인데, $A \setminus B = A \cap B^c = A \cap \bigsqcup_{i=1}^{k} C_i = \bigsqcup_{i=1}^{k} (A \cap C_i)$ 가 되어 증명이 끝난다. □

대수와 semi-algebra의 가장 큰 차이점은 semi-algebra의 경우 더 이상 여집합에 대해 닫혀있지 않다는 점이다. 위의 정의에서 볼 수 있듯이 여집합에 대해 닫혀있어야 한다는 대수의 조건이 조건 iii과 같이 약화되었다. 그러나 이번에도, semi-algebra와 대수를 연결해주는 정리가 존재한다.

Definition 2.14 공집합이 아닌 집합 X의 부분집합의 모임 $\mathscr{C}$에 대해 $\mathscr{C}$를 포함하는 최소의 대수를 **$\mathscr{C}$가 생성하는 대수 (algebra generated by $\mathscr{C}$)**라 하고 $A(\mathscr{C})$로 쓴다. 여기서 $\mathscr{C}$는 $A(\mathscr{C})$의 **생성자 (generator)**라 한다.

Proposition 2.15 공집합이 아닌 집합 X의 부분집합의 모임 $\mathscr{C}$에 대해 이가 생성하는 대수 $A(\mathscr{C})$는 유일하게 존재한다. 따라서, $A(\mathscr{C})$는 well-define된다.

Proof $A(\mathscr{C})$가 존재하기만 한다면, 이의 유일성은 그 정의로부터 자명하다. 존재성을 보이기 위해 Γ를 $\mathscr{C}$를 포함하는 모든 대수의 모임이라 하면 $\mathscr{P}(X)$가 $\mathscr{C}$를 포함하는 대수임이 분명하므로 Γ는 비어있지 않다. 이제 이의 교집합인 $\bigcap\Gamma$가 대수라 주장한다. 이를 보이기는 어렵지 않다. 정의 2. 1의 조건을 하나하나 따져보면 되는데, 연습삼아 두 번째 조건을 보자. 임의의 $A\in\bigcap\Gamma$에 대해 A는 Γ에 속하는 모든 대수에 속하므로 곧 A^c도 Γ에 속하는 모든 대수에 속하고, 이는 곧 $A^c\in\bigcap\Gamma$임을 뜻한다. 다른 조건들도 이와 같이 쉽게 따져볼 수 있다. 그렇다면 $\bigcap\Gamma$가 바로 $A(\mathscr{C})$이다. □

Theorem 2.16 공집합이 아닌 집합 X 위의 semi-algebra $\mathscr{A}$에 대해 $A(\mathscr{A})$는 $\mathscr{A}$에 속하는 서로소인 집합들의 유한 합집합으로 만들 수 있는 모든 집합들의 집합이다.

Proof 집합 $\mathscr{C}=\{\bigsqcup_{i=1}^k A_i\subseteq X : A_1,\cdots,A_k\in\mathscr{A}\}$를 생각하면 $\mathscr{C}\subseteq A(\mathscr{A})$임은 분명하므로 역의 포함관계를 보이는 것으로 충분하다. 이를 위해, $\mathscr{C}$가 대수임을 보이자. 만약 이를 보일 수 있으면 $\mathscr{A}\subseteq\mathscr{C}$임이 분명하므로 증명이 끝난다. 이하의 증명에서 A,B는 $\mathscr{C}$의 임의의 두 원소로 적당한 서로소인 집합 $A_1,\cdots,A_k\in\mathscr{A}$와 $B_1,\cdots,B_l\in\mathscr{A}$이 존재하여 $A=\bigsqcup_{i=1}^k A_i, B=\bigsqcup_{j=1}^l B_j$로 쓸 수 있다고 하자. 먼저 $A_1\cap B_1,\cdots,A_k\cap B_l\in\mathscr{A}$이 모두 서로소이고 $A\cap B=\bigsqcup_{i=1}^k\bigsqcup_{j=1}^l(A_i\cap B_j)$라는 점에서 $\mathscr{C}$가 교집합에 대해 닫혀있음을 안다. 다음으로, $A^c=(\bigsqcup_{i=1}^k A_i)^c=\bigcap_{i=1}^k A_i^c$이고 각 $i\le k$에 대해 A_i^c는 $\mathscr{A}$에 속하는 서로소인 집합들의 유한 합집합으로 쓸 수 있으므로 명백히 $\mathscr{C}$에 속하여 곧 $A^c\in\mathscr{C}$이고 $\mathscr{C}$가 여집합에 대해 닫혀있음도 안다. 또한, 만약 A,B가 서로소라면 $A_1\cdots,A_k,B_1,\cdots,B_l$이 모두 서로소가 되어 $A\sqcup B=\bigsqcup_{i=1}^k A_i\sqcup\bigsqcup_{j=1}^l B_j$에서 $\mathscr{C}$는 서로소 합집합에 대해서도 닫혀있다. 이제 마지막으로 A,B가 서로소가 아니라 하더라도 $A\cup B=(A\setminus B)\sqcup(B\setminus A)\sqcup(A\cap B)$에서 $A\setminus B=A\cap B^c, B\setminus A=B^c\cap A, A\cap B$가 모두 $\mathscr{C}$에 속하므로 $A\cup B\in\mathscr{C}$이고 곧 $\mathscr{C}$가 합집합에 대해 닫혀있음을 안다. 이로써 $\mathscr{C}$가 대수의 조건을 모두 만족함을 안다. □

2.2 Measures and Premeasures

지금까지 우리는 다양한 구조의 집합족을 다루는 방법들과 서로다른 구조를 연결하는 몇몇 유용한 정리들을 익혔다. 이번 절에서는 측도가 무엇인지 알아보고 이전 절에서 덩그러니

던져진 가측공간에 측도를 부여하는 과정을 알아보도록 한다. 그리고 이로써, 우리의 첫 번째 Lebesgue 측도의 구성이 본격적으로 시작된다.

Definition 2.17 가측공간 $(X, \mathscr{A})$에 대해 함수 $\mu : \mathscr{A} \to \overline{\mathbb{R}}_0^+$ 가

i. $\mu(\emptyset) = 0$.
ii. (σ−가법성) σ−대수 $\mathscr{A}$ 에 속하는 임의의 서로소인 집합열 $\{A_i\}$ 에 대해 $\mu(\bigsqcup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} \mu(A_i)$ 이다.

를 만족하면 이때의 μ 를 $\mathscr{A}$ 위의 **측도 (measure)**라 하고, triple $(X, \mathscr{A}, \mu)$ 를 **측도공간 (measure space)**이라 한다.

보통 측도는 그리스 문자 $\mu, \nu, \cdots$ 로 표기한다. 한편, 위의 정의를 유심히 보면 '측정'의 본질을 측도라는 개념에 어떻게 수학적으로 담아내었는지 살펴볼 수 있다. 우선, 측정의 결과는 반드시 0 이상의 실수이거나 무한대이다. 둘째로, 비어있는 집합을 측정한 결과는 반드시 0이다. 마지막으로 서로 겹치지 않는 가산개의 집합을 측정한 결과는 각각의 집합을 측정한 후에 이를 모두 더한 것과 일치한다. (이때, 더하는 항이 모두 양수이므로 덧셈의 순서는 문제가 되지 않는다.) 특히 마지막 성질을 흔히 **σ−가법성 (σ−additivity)** 혹은 **가산가법성 (countable additivity)**이라 부르며 어떤 함수가 측도인지를 결정하는 데 있어 가장 강한 조건으로 작용한다. 한편, 서로소인 유한개의 집합들에 공집합을 dummy term 으로 추가하여 이를 서로소인 집합열로 만들고 이에 위의 정의의 두 조건을 적용하면 유한개의 서로소인 집합들에 대해서도 σ−가법성과 비슷한 성질이 성립함을 알 수 있는데, 이를 **(유한) 가법성 ((finite) additivity)**이라 한다. 방금 설명한 것과 같이 측도는 σ−가법성과 가법성을 모두 가진다.

이제 측도의 정의를 알았으니 본격적으로 Lebesgue 측도의 구성을 시작해보자. Lebesgue 측도의 첫 번째 구성은 이른바 '확장정리' 들을 이용한다. 이전 절에서 언급한 바와 같이 σ−대수가 너무 거대한 까닭에 그 위에 곧바로 측도를 구성하기는 어려우니, 일단 그 구조가 간단한 $\mathbb{R}^n$ 위의 semi-algebra 하나를 적당히 택하여 이 위에 궁극적으로 측도가 될 함수를 정의하고, 이를 몇 단계를 거쳐 확장함으로써 최종적으로 Lebesgue 측도를 구성해내는 것이다. 여기서 처음에 semi-algebra 위에 정의되는 함수는 측도가 아님에 유의하자. (측도가 되려면 정의역이 σ−대수이어야 한다.) 이렇게 확장을 거치기 전의 '궁극적으로 측도가 될' 함수를 premeasure라 한다.

Definition 2.18 공집합이 아닌 집합 X 위의 semi-algebra $\mathscr{A}$ 에 대해 함수 $\rho : \mathscr{A} \to \overline{\mathbb{R}}_0^+$ 가

i. $\rho(\emptyset) = 0$.
ii. (σ−가법성) Semi-algebra $\mathscr{A}$ 에 속하는 임의의 서로소인 집합열 $\{A_i\}$ 에 대해 $\bigsqcup_{i=1}^{\infty} A_i \in \mathscr{A}$ 이면 $\mu(\bigsqcup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} \mu(A_i)$ 이다.

를 만족하면 이때의 ρ 를 $\mathscr{A}$ 위의 **premeasure**라 한다.

보통 premeasure는 그리스 문자 $\rho, \cdots$ 로 표기한다. 위의 정의에서 알 수 있듯이 premeasure는 $\sigma-$대수가 아닌 semi-algebra 위에서 보다 일반적으로 정의된다는 점을 제외하면 측도와 거의 같은 개념이며 당연히 측도는 premeasure이다. 다만, 더 이상 정의역이 $\sigma-$대수가 아니므로 $\sigma-$가법성의 경우 서로소 합집합의 결과가 정의역에 다시 포함된다는 추가적인 조건이 붙어야 한다. (일반적으로 정의역이 $\sigma-$대수가 아니어서 이러한 조건이 필수적인 경우, $\sigma-$가법성은 이러한 조건을 암묵적으로 포함하는 것으로 본다. 이는 가법성의 경우에도 마찬가지이다.) 이번 절의 남은 부분은 모두 premeasure와 측도의 성질에 관한 내용들로, 측도는 premeasure이므로 premeasure의 성질은 모두 측도에 대해서도 동일하게 성립한다는 점에 유의하기 바란다.

Lemma 2.19 공집합이 아닌 집합 X 위의 semi-algebra $\mathscr{A}$ 에 대해 함수 $f : \mathscr{A} \to \overline{\mathbb{R}}_0^+$ 가 가법성을 갖는다고 하면 임의의 집합 $A, B_1, \cdots, B_k \in \mathscr{A}$ 에 대해 다음이 성립한다.

i. 만약 $B_1, \cdots, B_k$ 가 서로소이고 $\bigsqcup_{i=1}^k B_i \subseteq A$ 이면 $\sum_{i=1}^k f(B_i) \le f(A)$ 이다.
ii. 만약 $A \subseteq \bigcup_{i=1}^k B_i$ 이면 $f(A) \le \sum_{i=1}^k f(B_i)$ 이다.

나아가, 만약 f 가 $\sigma-$가법성을 가지면 ii보다 일반적인 명제가 성립한다. 즉, 임의의 $A \in \mathscr{A}$ 와 $\mathscr{A}$ 에 속하는 임의의 집합열 $\{B_i\}$ 에 대해 다음이 성립한다.

ii° 만약 $A \subseteq \bigcup_{i=1}^\infty B_i$ 이면 $f(A) \le \sum_{i=1}^\infty f(B_i)$ 이다.

Proof i. 명제 2.13의 ii로부터 적당한 서로소인 $C_1, \cdots, C_l \in \mathscr{A}$ 이 존재하여 $A \setminus \bigsqcup_{i=1}^k B_i = A \setminus B_1 \setminus \cdots \setminus B_k = \bigsqcup_{j=1}^l C_j$ 이고, 따라서

$$\begin{aligned} f(A) &= f\left(\bigsqcup_{i=1}^k B_i \sqcup \left(A \setminus \bigsqcup_{i=1}^k B_i\right)\right) \\ &= f\left(\bigsqcup_{i=1}^k B_i \sqcup \bigsqcup_{j=1}^l C_j\right) \\ &= \sum_{i=1}^k f(B_i) + \sum_{j=1}^l f(C_j) \\ &\ge \sum_{i=1}^k f(B_i) \end{aligned}$$

이다.

ii. 각 $i \le k$에 대해 집합 A_i 를 $A_i := A \cap (B_i \setminus \bigcup_{j=1}^{i-1} B_j) = (A \cap B_i) \setminus B_1 \setminus \cdots \setminus B_{i-1}$ 로 두면 $A_1, \cdots, A_k$ 는 서로소이다. 또한 각 $i \le k$에 대해 $A_i \subseteq B_i$ 이며 명제 2.13의 ii로부터 적당한 서로소인 $C_{i1}, \cdots, C_{il_i} \in \mathscr{A}$ 가 존재하여 $A_i = \bigsqcup_{j=1}^{l_i} C_{ij}$ 이므로 방금 보인 i로부터

$$\begin{aligned} f(A) &= f\left(\bigsqcup_{i=1}^{k}\bigsqcup_{j=1}^{l_i} C_{ij}\right) \\ &= \sum_{i=1}^{k}\sum_{j=1}^{l_i} f(C_{ij}) \\ &\le \sum_{i=1}^{k} f(B_i) \end{aligned}$$

이다.

한편, f가 $\sigma-$가법성을 가지는 경우에 ii° 에 대해서는 집합열 $\{A_i\}$를 $A_i := A\cap(B_i \setminus \bigcup_{j=1}^{i-1} B_j) = (A\cap B_i)\setminus B_1 \setminus \cdots \setminus B_{i-1}$로 두고 방금 보인 ii의 증명과 비슷하게 하면 된다. □

Proposition 2.20 공집합이 아닌 집합 X 위의 semi-algebra $\mathscr{A}$에 대해 ρ를 $\mathscr{A}$ 위의 premeasure라 하면 다음이 성립한다.

i. (단조성) 임의의 $A, B \in \mathscr{A}$에 대해 $A \subseteq B$이면 $\rho(A) \le \rho(B)$이다. 나아가 $B\setminus A \in \mathscr{A}$이면 $\rho(B) = \rho(A) + \rho(B\setminus A)$이다.
ii. ($\sigma-$반가법성) Semi-algebra $\mathscr{A}$에 속하는 임의의 집합열 $\{A_i\}$에 대해 $\bigcup_{i=1}^{\infty} A_i \in \mathscr{A}$이면 $\rho(\bigcup_{i=1}^{\infty} A_i) \le \sum_{i=1}^{\infty}\rho(A_i)$이다.

Proof i. 이는 위의 보조정리의 i로부터 자명하다. 한편, 만약 $B\setminus A \in \mathscr{A}$이면 당연히 $\rho(B) = \rho(A) + \rho(B\setminus A)$이다.

ii. 이는 위의 보조정리의 ii° 로부터 자명하다. □

따라서 premeasure는 $\sigma-$가법성과 $\sigma-$반가법성, 가법성, 그리고 반가법성을 모두 가진다고 정리할 수 있다.

Theorem 2.21 (Inclusion-exclusion principle) 공집합이 아닌 집합 X 위의 대수 $\mathscr{A}$에 대해 ρ를 $\mathscr{A}$ 위의 premeasure라 하자. 그렇다면 임의의 $A_1, \cdots, A_l \in \mathscr{A}$에 대해 $\bigcup_{i=1}^{l} A_i \in \mathscr{A}$이고 $\rho(\bigcup_{i=1}^{l} A_i) < \infty$이면

$$\rho\left(\bigcup_{i=1}^{l} A_i\right) = \sum_{i=1}^{l}(-1)^{i-1}\sum_{1\le j_1<\cdots<j_i\le l}\rho\left(\bigcap_{k=1}^{i} A_{j_k}\right)$$

이다.

Proof 우선, 임의의 $j_1 < \cdots < j_i \le l$에 대해 $\rho(\bigcap_{k=1}^{i} A_{j_k}) \le \rho(\bigcup_{i=1}^{l} A_i) < \infty$임은 분명하다. 그렇다면 $l = 1$인 경우는 자명하고, $l = 2$인 경우에는

$$\rho(A_1\cup A_2) = \rho(A_1\setminus A_2) + \rho(A_2\setminus A_1) + \rho(A_1\cap A_2)$$

$$\rho(A_1) = \rho(A_1 \setminus A_2) + \rho(A_1 \cap A_2)$$
$$\rho(A_2) = \rho(A_2 \setminus A_1) + \rho(A_1 \cap A_2)$$

와 $\rho(A_1 \setminus A_2), \rho(A_2 \setminus A_1) \leq \rho(A_1 \cup A_2) < \infty$에서 $\rho(A_1 \cup A_2) = \rho(A_1) + \rho(A_2) - \rho(A_1 \cap A_2)$이다. 이제 $l \geq 3$인 경우에 대해서는 수학적 귀납법을 사용하면 된다. 귀납가정으로서 $l \in \mathbb{N}$에 대해 위의 명제가 성립한다고 하면 $\rho(\bigcup_{i=1}^{l} A_i), \rho(\bigcup_{i=1}^{l}(A_i \cap A_{l+1})) \leq \rho(\bigcup_{i=1}^{l+1} A_i) < \infty$이므로

$$\begin{aligned}
\rho\left(\bigcup_{i=1}^{l+1} A_i\right) &= \rho\left(\bigcup_{i=1}^{l} A_i\right) + \rho(A_{l+1}) - \rho\left(\bigcup_{i=1}^{l}(A_i \cap A_{l+1})\right) \\
&= \sum_{i=1}^{l}(-1)^{i-1} \sum_{1 \leq j_1 < \cdots < j_i \leq l} \rho\left(\bigcap_{k=1}^{i} A_{j_k}\right) + \rho(A_{l+1}) \\
&\qquad - \sum_{i=1}^{l}(-1)^{i-1} \sum_{1 \leq j_1 < \cdots < j_i \leq l} \rho\left(\bigcap_{k=1}^{i} A_{j_k} \cap A_{l+1}\right) \\
&= \sum_{i=1}^{l+1}(-1)^{i-1} \sum_{1 \leq j_1 < \cdots < j_i \leq l+1} \rho\left(\bigcap_{k=1}^{i} A_{j_k}\right)
\end{aligned}$$

에서 $l+1$에 대해서도 위의 명제가 성립함을 알 수 있다. □

다음은 집합열의 극한에 관련된 성질들이다.

Definition 2.22 집합열 $\{A_i\}$에 대해 이의 **하극한 (limit infimum)**과 **상극한 (limit supremum)**을 각각 $\liminf_{i\to\infty} A_i, \limsup_{i\to\infty} A_i$로 쓰고

$$\liminf_{i\to\infty} A_i = \bigcup_{j=1}^{\infty}\bigcap_{i=j}^{\infty} A_i, \qquad \limsup_{i\to\infty} A_i = \bigcap_{j=1}^{\infty}\bigcup_{i=j}^{\infty} A_i$$

로 정의한다. 나아가, 만약 $\liminf_{i\to\infty} A_i = \limsup_{i\to\infty} A_i$이면 이때 $\{A_i\}$가 **수렴 (converge)**한다고 하며, 이때의 공통의 값을 $\lim_{i\to\infty} A_i$로 쓴다.

Proposition 2.23 집합열 $\{A_i\}$에 대해 다음이 성립한다.

i. $\liminf_{i\to\infty} A_i \subseteq \limsup_{i\to\infty} A_i$.
ii. 만약 $\{A_i\}$가 증가하면 (혹은 감소하면) $A_i \uparrow \bigcup_{i=1}^{\infty} A_i$ (혹은 $A_i \downarrow \bigcap_{i=1}^{\infty} A_i$) 이다.
iii. 가측공간 $(X, \mathscr{A})$가 $\{A_i\}$를 포함하면 $\liminf_{i\to\infty} A_i, \limsup_{i\to\infty} A_i \in \mathscr{A}$이다.

Proof i. 임의의 $x \in \liminf_{i\to\infty} A_i = \bigcup_{j=1}^{\infty}\bigcap_{i=j}^{\infty} A_i$에 대해 적당한 $i_0 \in \mathbb{N}$이 존재하여 $i \geq i_0$이면 $x \in A_i$이므로 임의의 $j \in \mathbb{N}$에 대해 $i = \max\{i_0, j\}$로 두면 $x \in A_i$가 되어 $x \in \bigcap_{j=1}^{\infty}\bigcup_{i=j}^{\infty} A_i = \limsup_{i\to\infty} A_i$이고, 곧 $\liminf_{i\to\infty} A_i \subseteq \limsup_{i\to\infty} A_i$임을 안다.

ii. 간결한 논의를 위해 $\{A_i\}$가 증가하는 경우만 생각하자. (반대의 경우도 비슷하게 하면 된다.) 집합열 $\{A_i\}$가 증가하므로 $\liminf_{i\to\infty} A_i = \bigcup_{i=1}^{\infty} A_i$이고 임의의 $j \in \mathbb{N}$에 대해 $\bigcup_{i=j}^{\infty} A_i = \bigcup_{i=1}^{\infty} A_i$이므로 $\limsup_{i\to\infty} A_i = \bigcap_{j=1}^{\infty}\bigcup_{i=j}^{\infty} A_i = \bigcup_{i=1}^{\infty} A_i$이 되어 정의로부터 $A_i \uparrow \bigcup_{i=1}^{\infty} A_i$이다.

iii. 이는 정의로부터 자명하다. □

Theorem 2.24 측도공간 $(X, \mathscr{A}, \mu)$에 속하는 임의의 집합열 $\{A_i\}$에 대해 다음이 성립한다.

i. 만약 $A_i \uparrow A$이면 $\mu(A_i) \uparrow \mu(A)$이다.
ii. 만약 $A_i \downarrow A$이고 $\mu(A_1) < \infty$이면 $\mu(A_i) \downarrow \mu(A)$이다.

Proof i. 만약 어떤 $i \in \mathbb{N}$에 대해 $\mu(A_i) = \infty$이면 $\mu(A_i) \le \mu(A) = \infty$가 되어 명제가 자명하므로 모든 $i \in \mathbb{N}$에 대해 $\mu(A_i) < \infty$라 하자. 이제 집합열 $\{B_i\}$를 $B_i := A_i \setminus A_{i-1}$ $(B_1 := A_1)$로 정의하면 이는 $\mathscr{A}$에 속하는 서로소인 집합열이며 $A = \bigcup_{i=1}^{\infty} A_i = \bigsqcup_{i=1}^{\infty} B_i$이므로

$$\begin{aligned}
\mu(A) &= \mu\left(\bigsqcup_{i=1}^{\infty} B_i\right) \\
&= \sum_{i=1}^{\infty} \mu(B_i) \\
&= \mu(A_1) + \sum_{i=2}^{\infty} \mu(A_i \setminus A_{i-1}) \\
&= \lim_{j\to\infty} \left\{\mu(A_1) + \sum_{i=1}^{j} [\mu(A_i) - \mu(A_{i-1})]\right\} \\
&= \lim_{j\to\infty} \mu(A_j).
\end{aligned}$$

이다. 이제 수열 $\{\mu(A_i)\}$가 증가함이 분명하므로 명제가 성립한다.

ii. 집합열 $\{B_i\}$를 $B_i := A_i \setminus A_{i+1}$로 정의하면 이는 $\mathscr{A}$에 속하는 서로소인 집합열이며 $A_1 = \bigsqcup_{i=1}^{\infty} B_i \sqcup A$이다. 한편, 가정으로부터 모든 $i \in \mathbb{N}$에 대해 $\mu(A_i) \le \mu(A_1) < \infty$이고 $\mu(A) \le \mu(A_1) < \infty$이므로

$$\begin{aligned}
\mu(A_1) &= \mu\left(\bigsqcup_{i=1}^{\infty} B_i \sqcup A\right) \\
&= \sum_{i=1}^{\infty} \mu(B_i) + \mu(A) \\
&= \sum_{i=1}^{\infty} \mu(A_i \setminus A_{i+1}) + \mu(A) \\
&= \lim_{j\to\infty} \sum_{i=1}^{j} [\mu(A_i) - \mu(A_{i+1})] + \mu(A)
\end{aligned}$$

$$= \mu(A_1) - \lim_{j\to\infty} \mu(A_j) + \mu(A)$$

에서 $\mu(A_i) \to \mu(A)$ 임을 안다. 이제 수열 $\{\mu(A_i)\}$ 가 감소함이 분명하므로 명제가 성립한다. □

Definition 2.25 공집합이 아닌 집합 X 위의 semi-algebra $\mathscr{A}$ 에 대해 ρ 를 $\mathscr{A}$ 위의 premeasure 라 하자. 만약 $\rho(X) < \infty$ 이면 ρ 가 **유한 (finite)** 하다고 한다. 만약 $\mathscr{A}$ 에 속하는 적당한 집합열 $\{A_i\}$ 가 존재하여 각 $i \in \mathbb{N}$ 에 대해 $\rho(A_i) < \infty$ 이고 $\bigcup_{i=1}^{\infty} A_i = X$ 이면 ρ 가 **σ−유한 (σ−finite)** 하다고 한다. 특별히, $\mathscr{A}$ 의 생성자 $\mathscr{C}$ 에 대해 방금의 조건을 만족하는 집합열이 $\mathscr{C}$ 에 속하면 ρ 가 $\mathscr{C}$ **위에서 σ−유한 (σ−finite on $\mathscr{C}$)** 하다고 한다.

Theorem 2.26 유한 측도공간 $(X, \mathscr{A}, \mu)$ 와 $\mathscr{A}$ 에 속하는 임의의 집합열 $\{A_i\}$ 에 대해

$$\mu(\liminf_{i\to\infty} A_i) \le \liminf_{i\to\infty} \mu(A_i) \le \limsup_{i\to\infty} \mu(A_i) \le \mu(\limsup_{i\to\infty} A_i)$$

가 성립한다. 특별히, $A_i \to A$ 이면 $\mu(A_i) \to \mu(A)$ 이다.

Proof 위의 식에서 $\liminf_{i\to\infty} \mu(A_i) \le \limsup_{i\to\infty} \mu(A_i)$ 부분은 자명하므로 첫 번째와 세 번째 부등식만 보이면 되는데, 간결한 논의를 위해 여기서는 첫 번째 부등식만 보이자. (세 번째 부등식도 비슷하게 하면 된다.) 먼저 $B = \liminf_{i\to\infty} A_i$ 라 하고 집합열 $\{B_i\}$ 를 $B_i := \bigcap_{j=i}^{\infty} A_j$ 로 두면 이는 $\mathscr{A}$ 에 속하는 증가하는 집합열로서 $B_i \uparrow B$ 이다. 따라서 정리 2.24의 i로부터 $\mu(B) = \lim_{i\to\infty} \mu(B_i) \le \sup_{i\in\mathbb{N}} \mu(\bigcap_{j=i}^{\infty} A_j) \le \sup_{i\in\mathbb{N}} \inf_{j\ge i} \mu(A_j) = \liminf_{i\to\infty} \mu(A_i)$ 를 얻는다. 한편, $A_i \to A$ 인 경우에는 방금 보인 부등식으로부터 $\mu(A_i) \to \mu(A)$ 임이 자명하다. □

나중에 사용할 정리를 하나 증명하는 것으로 이번 절을 끝마친다.

Definition 2.27 측도공간 $(X, \mathscr{A}, \mu)$ 와 $A \in \mathscr{A}$ 에 대해 $\mu_A : B \mapsto \mu(A \cap B)$ 로 정의된 함수 $\mu_A : \mathscr{A} \to \overline{\mathbb{R}}_0^+$ 를 μ 의 A 로의 **제한 (restriction)** 이라 한다.

Proposition 2.28 측도공간 $(X, \mathscr{A}, \mu)$ 와 집합 $A \in \mathscr{A}$ 에 대해 μ_A 는 $\mathscr{A}$ 위의 측도이다.

Proof 거의 자명하다. 우선 $\mu_A(\emptyset) = \mu(\emptyset) = 0$ 임은 분명하고 $\mathscr{A}$ 에 속하는 임의의 서로소인 집합열 $\{B_i\}$ 에 대해 $\mu_A(\bigsqcup_{i=1}^{\infty} B_i) = \mu(\bigsqcup_{i=1}^{\infty}(A \cap B_i)) = \sum_{i=1}^{\infty} \mu(A \cap B_i) = \sum_{i=1}^{\infty} \mu_A(B_i)$ 가 성립하므로 μ_A 는 $\mathscr{A}$ 위의 측도가 된다. □

Theorem 2.29 가측공간 $(X, \mathscr{A})$ 와 $\mathscr{A}$ 의 생성자이자 π−system인 $\mathscr{P}$ 에 대해 μ, ν 를 $\mathscr{A}$ 위의 측도로서 $\mathscr{P}$ 위에서 σ−유한하다고 하자. 만약 μ, ν 가 $\mathscr{P}$ 위에서 일치하면 $\mu = \nu$ 이다.

Proof 가정으로부터 $\mathscr{P}$에 속하는 집합열 $\{A_i\}$가 존재하여 $\mu(A_i) < \infty$이고 $\bigcup_{i=1}^{\infty} A_i = X$이다. 여기서 집합열 $\{B_i\}$를 $B_i := A_i \setminus \bigcup_{j=1}^{i-1} A_j$로 두면 이는 $\mathscr{A}$에 속하는 서로소인 집합열로 $\bigsqcup_{i=1}^{\infty} B_i = X$이며 포함배제의 원리로부터 모든 $i \in \mathbb{N}$에 대해

$$\begin{aligned}
\mu(B_i) &= \mu\left(A_i \setminus \bigcup_{j=1}^{i-1} A_j\right) \\
&= \mu(A_i) - \mu\left(\bigcup_{j=1}^{i-1} (A_i \cap A_j)\right) \\
&= \mu(A_i) - \sum_{j=1}^{i-1} (-1)^{j-1} \sum_{1 \le k_1 < \cdots < k_j \le i-1} \mu\left(\bigcap_{l=1}^{j} (A_i \cap A_{k_l})\right) \\
&= \nu(A_i) - \sum_{j=1}^{i-1} (-1)^{j-1} \sum_{1 \le k_1 < \cdots < k_j \le i-1} \nu\left(\bigcap_{l=1}^{j} (A_i \cap A_{k_l})\right) \\
&= \nu(A_i) - \nu\left(\bigcup_{j=1}^{i-1} (A_i \cap A_j)\right) \\
&= \nu\left(A_i \setminus \bigcup_{j=1}^{i-1} A_j\right) \\
&= \nu(B_i)
\end{aligned}$$

가 성립한다. 이를 이용하여 우선 μ, ν가 유한한 경우에 대해 위의 정리를 증명하자. 일단 $\mu(X) = \mu(\bigsqcup_{i=1}^{\infty} B_i) = \sum_{i=1}^{\infty} \mu(B_i) = \sum_{i=1}^{\infty} \nu(B_i) = \nu(\bigsqcup_{i=1}^{\infty} B_i) = \nu(X)$가 성립하므로 집합족 $\mathscr{L} = \{A \in \mathscr{A} : \mu(A) = \nu(A)\}$를 생각하면 이가 $\mathscr{P}$를 포함하는 λ−system임이 거의 자명하다. 그렇다면 Dynkin의 π−λ 정리로부터 $\mathscr{A} = \sigma(\mathscr{P}) = \lambda(\mathscr{P}) \subseteq \mathscr{L}$이 되어 증명이 끝난다.

이제 일반적인 경우를 증명하자. 이를 위해 각 $i \in \mathbb{N}$에 대해 측도 μ_{B_i}, ν_{B_i}를 생각하고 앞서 포함배제의 원리를 사용한 것과 비슷하게 하면 μ_{B_i}, ν_{B_i}가 $\mathscr{A}$ 위의 유한 측도이며 $\mathscr{P}$ 위에서 일치함을 보일 수 있다. 따라서 앞선 결론으로부터 각 $i \in \mathbb{N}$에 대해 $\mu_{B_i} = \nu_{B_i}$임을 안다. 그렇다면 임의의 $A \in \mathscr{A}$에 대해 $\mu(A) = \mu(\bigsqcup_{i=1}^{\infty} (A \cap B_i)) = \sum_{i=1}^{\infty} \mu(A \cap B_i) = \sum_{i=1}^{\infty} \mu_{B_i}(A) = \sum_{i=1}^{\infty} \nu_{B_i}(A) = \nu(\bigsqcup_{i=1}^{\infty} (A \cap B_i)) = \nu(A)$이므로 $\mu = \nu$임을 안다. □

위의 정리는 σ−유한인 측도는 그 측도가 정의된 σ−대수의 π−system인 생성자 위에서의 값만으로 특정됨을 함의한다. 일반적으로 σ−대수의 생성자는 그가 생성하는 σ−대수보다 그 크기가 작고, 대부분의 경우 그 원소를 명시적으로 표현할 수 있기에 위의 정리는 상당히 유용하고 강력한 정리이다. 한편, 이가 사실상 Dynkin의 π−λ 정리의 따름정리 격이라는 점도 눈여겨 볼 만하다.

2.3 Extension Theorems

이번 절에서는 Lebesgue 측도의 구성의 핵심이 되는 확장정리들에 대해 자세히 알아본다. 앞서 잠시 언급했듯이 확장정리들은 semi-algebra와 같은 약한 구조에서 정의된 premeasure를 대수나 $\sigma-$대수와 같은 보다 강한 구조를 가진 곳에서의 premeasure나 측도로 확장시킬 수 있는 방법을 제공한다. 여기에서는 크게 두 가지의 확장정리를 소개할텐데, 첫 번째 확장정리에 비해 두 번째 확장정리는 조금 깊은 이해를 필요로 한다. 마음의 준비를 하도록 하자.

Theorem 2.30 공집합이 아닌 집합 X 위의 semi-algebra $\mathscr{A}$ 에 대해 ρ 를 $\mathscr{A}$ 위의 premeasure라 하면 이의 $A(\mathscr{A})$ 에서의 premeasure로의 확장이 유일하게 존재한다.

Proof 정리 2. 16로부터 임의의 $A \in A(\mathscr{A})$ 에 대해 적당한 서로소인 집합 $A_1, \cdots, A_k \in \mathscr{A}$ 가 존재하여 $A = \bigsqcup_{i=1}^k A_i$ 이다. 이제 함수 $\widetilde{\rho} : A(\mathscr{A}) \to \overline{\mathbb{R}}_0^+$ 를 위와 같이 표현되는 A 에 대해 $\widetilde{\rho} : A \mapsto \sum_{i=1}^k \rho(A_i)$ 인 함수로 정의하자. 그렇다면 이는 명백히 ρ 의 $A(\mathscr{A})$ 로의 확장이다. 한편, 이의 well-definedness를 담보하기 위해 위의 A 를 또다른 서로소인 $B_1, \cdots, B_l \in \mathscr{A}$ 에 대해 $A = \bigsqcup_{j=1}^l B_j$ 로도 쓸 수 있다고 하자. 그렇다면 $A_1 \cap B_1, \cdots, A_k \cap B_l \in \mathscr{A}$ 이 모두 서로소이므로 $\sum_{i=1}^k \rho(A_i) = \sum_{i=1}^k \rho(\bigsqcup_{j=1}^l (A_i \cap B_j)) = \sum_{i=1}^k \sum_{j=1}^l \rho(A_i \cap B_j) = \sum_{j=1}^l \rho(\bigsqcup_{i=1}^k (A_i \cap B_j)) = \sum_{j=1}^l \rho(B_j)$ 가 되어 $\widetilde{\rho}$ 가 well-define됨을 안다.

이제 $\widetilde{\rho}$ 가 premeasure임을 보이자. 정의로부터 $\widetilde{\rho}(\emptyset) = \rho(\emptyset) = 0$ 임은 분명하므로 이의 $\sigma-$가법성을 보이는 것으로 충분하다. 이를 위해, $A(\mathscr{A})$ 에 속하는 임의의 서로소인 집합열 $\{A_i\}$ 에 대해 $A = \bigsqcup_{i=1}^\infty A_i \in A(\mathscr{A})$ 라 하자. 그렇다면 다시 정리 2. 16로부터 적당한 서로소인 집합 $B_1, \cdots, B_l \in \mathscr{A}$ 가 존재하여 $A = \bigsqcup_{j=1}^l B_j$ 이고 각 $i \in \mathbb{N}$ 에 대해서도 적당한 서로소인 집합 $C_{i1}, \cdots, C_{im_i} \in \mathscr{A}$ 가 존재하여 $A_i = \bigsqcup_{k=1}^{m_i} C_{ik}$ 이다. 이로부터

$$\begin{aligned}
\widetilde{\rho}(A) &= \widetilde{\rho}\left(\bigsqcup_{j=1}^l B_j\right) \\
&= \sum_{j=1}^l \rho(B_j) \\
&= \sum_{j=1}^l \rho\left(\bigsqcup_{i=1}^\infty \bigsqcup_{k=1}^{m_i} (B_j \cap C_{ik})\right) \\
&= \sum_{j=1}^l \sum_{i=1}^\infty \sum_{k=1}^{m_i} \rho(B_j \cap C_{ik}) \\
&= \sum_{i=1}^\infty \widetilde{\rho}\left(\bigsqcup_{k=1}^{m_i} \bigsqcup_{j=1}^l (B_j \cap C_{ik})\right)
\end{aligned}$$

$$
\begin{aligned}
&= \sum_{i=1}^{\infty} \widetilde{\rho}\left(\bigsqcup_{k=1}^{m_i} C_{ik}\right) \\
&= \sum_{i=1}^{\infty} \widetilde{\rho}(A_i)
\end{aligned}
$$

가 되어 $\widetilde{\rho}$가 $\sigma-$가법성을 가짐을 안다.

이제 이러한 확장의 유일성만 보이면 되는데, ρ의 $A(\mathscr{A})$에서의 premeasure로의 어떠한 확장이든 그 확장은 $\sigma-$가법성을 가지고, 곧 가법성을 가지므로 앞서 정의한 $\widetilde{\rho}$와 일치할 수 밖에 없다. 따라서 유일성은 자명하다. □

위의 확장정리는 semi-algebra 위에서의 premeasure를 그 semi-algebra가 생성하는 대수 위의 premeasure로 확장시키는 방법을 제공한다. 그리고 이를 $\sigma-$대수 위의 측도로 끌어올리는 확장정리가 바로 두 번째로 소개할 확장정리이다. 다만, 두 번째 확장정리는 첫 번째보다 조금 더 복잡해서, 대수 위의 premeasure를 $\sigma-$대수 위의 측도로 곧바로 확장시키는 것이 아니라, 우선 이를 전체 공간으로 확장시킨 다음 다시 적당한 $\sigma-$대수로 축소하여 측도로의 확장을 얻는다. 당연히 이때 전체 공간으로 확장시킨 함수는 측도가 요구하는 강한 조건을 만족시키지 못하므로 측도가 아니며, 이를 외측도라 부른다.

Definition 2.31 공집합이 아닌 집합 X에 대해 함수 $\mu^* : \mathscr{P}(X) \to \overline{\mathbb{R}}_0^+$가

i. $\mu^*(\emptyset) = 0$.
ii. (단조성) 임의의 $A, B \subseteq X$에 대해 $A \subseteq B$이면 $\mu^*(A) \leq \mu^*(B)$이다.
iii. ($\sigma-$반가법성) 멱집합 $\mathscr{P}(X)$에 속하는 임의의 집합열 $\{A_i\}$에 대해 $\mu^*(\bigcup_{i=1}^{\infty} A_i) \leq \sum_{i=1}^{\infty} \mu^*(A_i)$이다.

를 만족하면 이때의 μ^*를 $\mathscr{P}(X)$ 위의 **외측도 (outer measure)**라 한다.

보통 외측도는 그리스 문자에 asterisk $*$을 위첨자로 붙여 $\mu^*, \nu^*, \cdots$ 로 표기한다. Premeasure가 $\sigma-$대수보다 약한 구조에서 정의된 일반적인 측도라면, 외측도는 측도의 강력한 조건인 $\sigma-$가법성을 과감히 포기하고 이를 단조성과 $\sigma-$반가법성으로 대체하여 얻는 전체 공간에서의 일반적인 측도로 볼 수 있다. 위의 정의의 세 조건을 만족하는 함수라면 모두 외측도라 할 수 있지만, 우리는 외측도를 확장정리의 중간 과정으로만 이용할 계획이므로 지금부터는 premeasure로부터 유도되는 외측도에 집중하도록 하자.

Definition 2.32 공집합이 아닌 집합 X 위의 대수 $\mathscr{A}$에 대해 ρ를 $\mathscr{A}$ 위의 premeasure라 하면 **ρ로부터 유도되는 외측도 (outer measure induced by ρ)**를 ρ^*로 쓰고

$$
\rho^* : A \mapsto \inf\left\{ \sum_{i=1}^{\infty} \rho(A_i) \in \overline{\mathbb{R}}_0^+ : \{A_i\}\text{는 } \mathscr{A}\text{에 속하는 } A\text{의 가산 덮개이다.} \right\}
$$

로 정의한다.

Proposition 2.33 공집합이 아닌 집합 X 위의 대수 $\mathscr{A}$ 에 대해 ρ 를 $\mathscr{A}$ 위의 premeasure라 하자. 그렇다면 이로부터 유도되는 외측도 ρ^* 에 대해 다음이 성립한다.

i. 만약 ρ 가 유한하면 임의의 $A \subseteq X$ 에 대해 $\rho^*(A) < \infty$ 이다.
ii. 함수 ρ^* 와 ρ 는 $\mathscr{A}$ 위에서 일치한다. 따라서 ρ^* 는 ρ 의 확장이다.
iii. 함수 ρ^* 는 외측도이다.

Proof i. 이는 $\{X\}$ 가 임의의 $A \subseteq X$ 의 $\mathscr{A}$ 에 속하는 가산 덮개라는 점에서 자명하다.

ii. 임의의 $A \in \mathscr{A}$ 를 택하자. 우선 $\{A\}$ 가 그 자체로 $\mathscr{A}$ 에 속하는 A 의 가산 덮개이므로 $\rho^*(A) \le \rho(A)$ 이다. 역으로, $\mathscr{A}$ 에 속하는 A 의 임의의 가산 덮개 $\{A_i\}$ 에 대해 WLOG, 필요하다면 이에 공집합을 추가하여 이가 집합열을 이룬다고 하면 $\sigma-$반가법성으로부터 $\rho(A) \le \sum_{i=1}^{\infty} \rho(A_i)$ 이므로 $\rho(A) \le \rho^*(A)$ 이다.

iii. ii로부터 $\rho^*(\emptyset) = \rho(\emptyset) = 0$ 임은 분명하다. 또한 임의의 두 $A, B \subseteq X$ 에 대해 $A \subseteq B$ 라면 B 의 모든 가산 덮개는 곧 A 의 가산 덮개이기도 하므로 $\rho^*(A) \le \rho^*(B)$ 에서 ρ^* 는 단조성도 가진다. 곧 ρ^* 의 $\sigma-$반가법성을 보이면 충분하므로 이를 위해 $\mathscr{P}(X)$ 에 속하는 임의의 집합열 $\{A_i\}$ 를 택하자. 만약 $\rho^*(A_{i_0}) = \infty$ 인 $i_0 \in \mathbb{N}$ 가 존재하면 $\rho^*(\bigcup_{i=1}^{\infty} A_i) \ge \rho^*(A_{i_0}) = \infty$ 에서 $\sigma-$반가법성이 자명하므로 모든 $i \in \mathbb{N}$ 에 대해 $\rho^*(A_i) < \infty$ 라 가정하자. 이제 임의의 $\varepsilon > 0$ 을 택하면 각 $i \in \mathbb{N}$ 에 대해 $\mathscr{A}$ 에 속하는 A_i 의 적당한 가산 덮개 $\{B_{ij}\}_j$ 가 존재하여 WLOG, 필요하다면 이에 공집합을 추가하여 이가 집합열을 이룬다고 하면 $\sum_{j=1}^{\infty} \rho(B_{ij}) < \rho^*(A_i) + \varepsilon/2^i$ 이다. 한편, 이러한 덮개를 모두 합한 $\{B_{ij}\}_{ij}$ 가 $\mathscr{A}$ 에 속하는 $\bigcup_{i=1}^{\infty} A_i$ 의 가산 덮개이므로 $\rho^*(\bigcup_{i=1}^{\infty} A_i) \le \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} \rho(B_{ij}) \le \sum_{i=1}^{\infty} [\rho^*(A_i) + \varepsilon/2^i] = \sum_{i=1}^{\infty} \rho^*(A_i) + \varepsilon$ 에서 $\rho^*(\bigcup_{i=1}^{\infty} A_i) \le \sum_{i=1}^{\infty} \rho^*(A_i)$ 가 되어 ρ^* 가 $\sigma-$반가법성을 가짐을 안다. □

이제부터 다소 혼란스러운 부분이 시작된다. 앞서 1절에서 이미 가측집합을 $\sigma-$대수의 원소로 정의한 바 있는데, 우리는 여기서 가측집합을 다시 정의한다.

Definition 2.34 공집합이 아닌 집합 X 의 멱집합 $\mathscr{P}(X)$ 위의 외측도 μ^* 에 대해 만약 $A \subseteq X$ 가 임의의 $S \subseteq X$ 에 대해 $\mu^*(S \cap A) + \mu^*(S \cap A^c) = \mu^*(S)$ 를 만족하면 이때의 A 를 **(μ^*-) 가측집합 ((μ^*-)measurable set)**이라 한다.

외측도의 $\sigma-$반가법성을 이용하면 가측일 조건을 다음과 같이 쓸 수도 있다.

Proposition 2.35 공집합이 아닌 집합 X 의 멱집합 $\mathscr{P}(X)$ 위의 외측도 μ^* 에 대해 집합 $A \subseteq X$ 가 가측일 필요충분조건은 임의의 $S \subseteq X$ 에 대해 $\mu^*(S \cap A) + \mu^*(S \cap A^c) \le \mu^*(S)$ 가 성립하는 것이다.

Proof 외측도 μ^* 의 $\sigma-$반가법성으로부터 $\mu^*(S) = \mu^*((S \cap A) \sqcup (S \setminus A)) \le \mu^*(S \cap A) + \mu^*(S \cap A^c)$ 가 항상 성립하므로 이는 자명하다. □

더 이상 진행하기 전에, 가측집합에 대한 정의로 인한 혼란을 조금 해결하고 가는 것이 좋을 것 같다. 혼란의 원인은 앞서 제시된 두 개의 정의가 서로 전혀 동등하게 보이지 않고, 실제로도 동등하지 않다는 점이다. 하지만, 다행히 다음 정리에 의하면 두 번째 가측집합의 정의는 첫 번째 가측집합의 정의에 부합하여 위의 두 정의는, 비록 서로 같지는 않지만, 양립 가능하다. (사실 첫 번째 정의가 두 번째 정의보다 더 일반적인 정의인 격이다.) 이러한 혼란에도 불구하고, 전체 공간으로 확장된 외측도를 다시 적당한 $\sigma-$대수로 축소하여 측도를 얻기 위해서 두 번째 가측집합의 정의는 반드시 필요하다.

Lemma 2.36 공집합이 아닌 집합 X의 멱집합 $\mathscr{P}(X)$ 위의 외측도 μ^*에 대해 $\mathscr{M}$을 모든 가측집합들의 모임이라 하면 μ^*는 $\mathscr{M}$ 위에서 가법성을 가진다.

Proof 임의의 서로소인 $A, B \in \mathscr{M}$에 대해 $\mu^*(A \sqcup B) = \mu^*((A \sqcup B) \cap A) + \mu^*((A \sqcup B) \cap A^c) = \mu^*(A) + \mu^*(B)$가 성립하고 여기에 수학적 귀납법을 적용하면 이 보조정리는 쉽게 보일 수 있다. □

Theorem 2.37 공집합이 아닌 집합 X의 멱집합 $\mathscr{P}(X)$ 위의 외측도 μ^*에 대해 $\mathscr{M}$을 모든 가측집합들의 모임이라 하면 이는 $\sigma-$대수이다.

Proof 먼저 $\mathscr{M}$이 대수임을 보이자. 우선 $\emptyset$이 가측이므로 $\mathscr{M}$은 비어있지 않다. 또한 임의의 $A \in \mathscr{M}$에 대해 $A^c \in \mathscr{M}$이므로 $\mathscr{M}$은 여집합에 대해 닫혀있다. 다음으로, 임의의 $A, B \in \mathscr{M}$를 택하면 임의의 $S \subseteq X$에 대해 $\mu^*(S) = \mu^*(S \cap A) + \mu^*(S \cap A^c) = \mu^*(S \cap A \cap B) + \mu^*(S \cap A \cap B^c) + \mu^*(S \cap A^c)$이고 $(S \cap A \cap B^c) \cup (S \cap A^c) = S \cap (A \cap B)^c$이므로 $\mu^*(S) \geq \mu^*(S \cap (A \cap B)) + \mu^*(S \cap (A \cap B)^c)$가 되어 명제 2. 35로부터 $A \cap B \in \mathscr{M}$이고, 곧 $\mathscr{M}$이 교집합에 대해서도 닫혀있음을 안다. 그렇다면 임의의 $A, B \in \mathscr{M}$에 대해 $A \cup B = (A^c \cap B^c)^c \in \mathscr{M}$이므로 $\mathscr{M}$은 합집합에 대해 닫혀있고, 이로써 $\mathscr{M}$은 대수임을 안다.

이제 $\mathscr{M}$이 가산 합집합에 대해 닫혀있다는 것만 보이면 된다. 이를 위해 $\mathscr{M}$에 속하는 임의의 집합열 $\{A_i\}$에 대해 WLOG, 필요하다면 각 항을 $B_i := A_i \setminus \bigcup_{j=1}^{i-1} A_j$로 바꾸어 $\{A_i\}$가 처음부터 서로소인 집합열이라고 해도 된다. 그렇다면 방금 $\mathscr{M}$이 대수임을 보였으므로 각 $k \in \mathbb{N}$에 대해 $\bigsqcup_{i=1}^{k} A_i \in \mathscr{M}$이고 위의 보조정리로부터 각 $k \in \mathbb{N}$와 임의의 $S \subseteq X$에 대해

$$\begin{aligned}
\mu^*(S) &= \mu^*\left(S \cap \bigsqcup_{i=1}^{k} A_i\right) + \mu^*\left(S \cap \left(\bigsqcup_{i=1}^{k} A_i\right)^c\right) \\
&= \mu^*\left(\bigsqcup_{i=1}^{k} (S \cap A_i)\right) + \mu^*\left(S \cap \left(\bigsqcup_{i=1}^{k} A_i\right)^c\right) \\
&\geq \sum_{i=1}^{k} \mu^*(S \cap A_i) + \mu^*\left(S \cap \left(\bigsqcup_{i=1}^{\infty} A_i\right)^c\right)
\end{aligned}$$

가 성립한다. 이로부터 임의의 $S \subseteq X$에 대해

$$\begin{aligned}\mu^*(S) &\geq \sum_{i=1}^{\infty}\mu^*(S\cap A_i)+\mu^*\left(S\cap\left(\bigsqcup_{i=1}^{\infty}A_i\right)^c\right)\\ &\geq \mu^*\left(\bigsqcup_{i=1}^{\infty}(S\cap A_i)\right)+\mu^*\left(S\cap\left(\bigsqcup_{i=1}^{\infty}A_i\right)^c\right)\\ &= \mu^*\left(S\cap\bigsqcup_{i=1}^{\infty}A_i\right)+\mu^*\left(S\cap\left(\bigsqcup_{i=1}^{\infty}A_i\right)^c\right)\end{aligned}$$

가 성립하여 명제 2.35으로부터 $\bigsqcup_{i=1}^{\infty}A_i\in\mathscr{M}$ 이고, 곧 증명이 끝난다. □

자, 정의 2.34에서의 가측집합을 생각해보자. 어떤 집합 A가 μ^*−가측이라 함은 곧 모든 μ^*−가측집합들의 모임 $\mathscr{M}$에 대해 $A\in\mathscr{M}$ 임을 뜻하고, 따라서 정의 2.3에 따르더라도 A는 $\mathscr{M}$−가측집합이 되어 앞서 말한 바와 같이 비록 두 정의가 서로 동등하지는 않지만, 이 두 정의는 서로 양립가능하다. 이제 두 번째 확장정리인 Carathéodory의 확장정리를 보자.

Lemma 2.38 공집합이 아닌 집합 X 위의 대수 $\mathscr{A}$에 대해 함수 $f:\mathscr{A}\to\overline{\mathbb{R}}_0^+$를 $f(\emptyset)=0$인 가법성을 갖는 함수라 하면 f가 σ−가법성을 가질 필요충분조건은 f가 σ−반가법성을 갖는 것이다.

Proof 만약 f가 σ−가법성을 가지면 이는 premeasure이므로 명제 2.20의 ii로부터 f가 σ−반가법성을 가지게 되어 충분조건임은 자명하다. 필요조건임을 보이기 위해 $\mathscr{A}$에 속하는 임의의 서로소인 집합열 $\{A_i\}$를 생각하여 $\bigsqcup_{i=1}^{\infty}A_i\in\mathscr{A}$라 하면 가정으로부터 $f(\bigsqcup_{i=1}^{\infty}A_i)\leq\sum_{i=1}^{\infty}f(A_i)$가 성립하고 보조정리 2.19의 i로부터 임의의 $k\in\mathbb{N}$에 대해 $\sum_{i=1}^{k}f(A_i)=f(\bigsqcup_{i=1}^{k}A_i)\leq f(\bigsqcup_{i=1}^{\infty}A_i)$가 되어 $f(\bigsqcup_{i=1}^{\infty}A_i)\geq\sum_{i=1}^{\infty}f(A_i)$도 성립한다. 이로부터 $f(\bigsqcup_{i=1}^{\infty}A_i)=\sum_{i=1}^{\infty}f(A_i)$가 되어 f가 σ−가법성을 가짐을 안다. □

Theorem 2.39 (Carathéodory's extension theorem) 공집합이 아닌 집합 X 위의 대수 $\mathscr{A}$에 대해 ρ를 $\mathscr{A}$ 위의 premeasure라 하고 $\mathscr{M}$을 모든 ρ^*−가측집합들의 모임이라 하자. 그렇다면 $\mu:=\rho^*|_{\mathscr{M}}$는 ρ의 $\mathscr{M}$에서의 측도로의 확장이고, 따라서 $(X,\mathscr{M},\mu)$는 측도공간이 된다.

Proof 우선 $\mu(\emptyset)=0$이고 μ가 정의로부터 σ−반가법성을 가지며 보조정리 2.36로부터 가법성도 가지므로 위의 보조정리에 의해 이는 σ−가법성을 가지고, 곧 μ는 $\mathscr{M}$ 위의 측도가 된다. 이제 μ가 ρ의 확장임을 보이기만 하면 되는데, 만약 $\mathscr{M}$이 $\mathscr{A}$를 포함한다는 것을 보인다면 명제 2.33의 ii로부터 μ가 ρ의 확장임이 분명하다. 이를 위해 임의의 $A\in\mathscr{A}$와 임의의 $S\subseteq X$ 그리고 임의의 $\varepsilon>0$을 택하면 $\mathscr{A}$에 속하는 적당한 가산 덮개 $\{S_i\}$가 존재하여 WLOG, 필요하다면 공집합을 추가하여 이가 집합열을 이룬다고 하면 $\sum_{i=1}^{\infty}\rho(S_i)\leq\rho^*(S)+\varepsilon$이다. 여기서 $\{S_i\cap A\},\{S_i\cap A^c\}$가 각각 $\mathscr{A}$에 속하는 $S\cap A, S\cap A^c$의 가산 덮개이므로 $\rho^*(S\cap A)+\rho^*(S\cap A^c)\leq\sum_{i=1}^{\infty}[\rho(S_i\cap A)+\rho(S_i\cap A^c)]=\sum_{i=1}^{\infty}\rho(S_i)\leq\rho^*(S)+\varepsilon$이 되어 곧 $\rho^*(S\cap A)+\rho^*(S\cap A^c)\leq\rho^*(S)$이다. 따라서 명제 2.35로부터 $A\in\mathscr{M}$이고, 곧 $\mathscr{A}\subseteq\mathscr{M}$이므로 증명이 끝난다. □

이와 같이 어떤 premeasure ρ 에 Carathéodory의 확장정리를 적용함으로써 얻는 측도를 **ρ 의 Carathéodory 확장 (− extension of ρ)** 이라 한다. 다만, 일반적으로 어떤 premeasure ρ 를 모든 ρ^*−가측집합들의 모임 $\mathscr{M}$ 으로 확장하는 방법은 유일하지 않다. 즉, ρ 의 $\mathscr{M}$ 에서의 측도로의 확장은 ρ 의 Carathéodory 확장 말고도 얼마든지 있을 수 있다.[3]하지만 만약 ρ 가 σ−유한하다면 그러한 확장은 Carathéodory 확장으로 유일함을 보일 수 있다.

Theorem 2.40 공집합이 아닌 집합 X 위의 대수 $\mathscr{A}$ 에 대해 ρ 를 $\mathscr{A}$ 위의 σ−유한 premeasure 라 하고 $\mathscr{M}$ 을 모든 ρ^*−가측집합들의 모임이라 하자. 또한 μ 를 ρ 의 Carathéodory 확장이라 하자. 그렇다면 μ 도 σ−유한하고, 임의의 측도공간 $(X, \mathscr{N}, \nu)$ 에 대해 만약 ν 가 ρ 의 확장이라면 μ 와 ν 는 $\mathscr{M} \cap \mathscr{N}$ 에서 일치한다.

Proof 우선 μ 가 σ−유한함은 거의 자명하다. 가정으로부터 ρ 가 σ−유한하므로 $\mathscr{A}$ 에 속하는 적당한 집합열 $\{A_i\}$ 가 존재하여 모든 $i \in \mathbb{N}$ 에 대해 $\mu(A_i) = \rho(A_i) < \infty$ 이고 $\bigcup_{i=1}^{\infty} A_i = X$ 이며, 따라서 μ 도 σ−유한하다.

다음으로, $\mathscr{M} \cap \mathscr{N}$ 에서 μ 와 ν 가 일치함을 보이자. 이를 위해 임의의 $A \in \mathscr{M} \cap \mathscr{N}$ 를 택하면 $\mathscr{A}$ 에 속하는 A 의 임의의 가산 덮개 $\{A_i\}$ 에 대해 WLOG, 필요하다면 공집합을 추가하여 이가 집합열을 이룬다고 하면 $\nu(A) \leq \nu(\bigcup_{i=1}^{\infty} A_i) \leq \sum_{i=1}^{\infty} \nu(A_i) = \sum_{i=1}^{\infty} \rho(A_i)$ 이므로 $\nu(A) \leq \rho^*(A) = \mu(A)$ 이다.

이제 임의의 $A \in \mathscr{M} \cap \mathscr{N}$ 를 고정하고 $\mu(A) < \infty$ 인 경우를 생각해보자. 이 경우에는 $\mathscr{A}$ 에 속하는 적당한 A 의 가산 덮개 $\{A_i\}$ 가 존재하여 $\sum_{i=1}^{\infty} \rho(A_i) < \infty$ 이고, WLOG, 필요하다면 각 항을 $B_i := A_i \setminus \bigcup_{j=1}^{i-1} A_j$ 로 교체하여 $\{A_i\}$ 가 처음부터 서로소인 집합열이라 해도 된다. 여기서 $B = \bigsqcup_{i=1}^{\infty} A_i$ 라 하면 $B \setminus A \in \mathscr{M} \cap \mathscr{N}$ 이고, 곧 앞서 결론으로부터 $\nu(A) \leq \mu(A), \nu(B \setminus A) \leq \mu(B \setminus A)$ 이다. 또한, $\mu(B) = \sum_{i=1}^{\infty} \mu(A_i) = \sum_{i=1}^{\infty} \rho(A_i) = \sum_{i=1}^{\infty} \nu(A_i) = \nu(B) < \infty$ 에서 $\nu(A) + \nu(B \setminus A) = \nu(B) = \mu(B) = \mu(A) + \mu(B \setminus A)$ 이다. 지금까지 얻은 결론들을 정리하면

i. $\nu(A) \leq \mu(A)$.
ii. $\nu(B \setminus A) \leq \mu(B \setminus A)$.
iii. $\nu(B) = \mu(B) < \infty$.
iv. $\nu(A) + \nu(B \setminus A) = \mu(A) + \mu(B \setminus A)$.

인데, 이를 종합하면 $\mu(A) = \nu(A)$ 임을 쉽게 알 수 있다.

마지막으로, 이번에는 $\mu(A) = \infty$ 인 경우를 생각해보자. 앞서 보인 바와 같이 μ 가 σ−유한하므로 $\mathscr{A}$ 에 속하는 적당한 집합열 $\{B_i\}$ 가 존재하여 모든 $i \in \mathbb{N}$ 에 대해 $\mu(B_i) < \infty$ 이고 $\bigcup_{i=1}^{\infty} B_i = X$ 이며, WLOG, $\mu(A) < \infty$ 인 경우에서와 비슷하게 하여 $\{B_i\}$ 가 처음부터 서로소인 집합열이라 해도 된다. 그렇다면 모든 $i \in \mathbb{N}$ 에 대해 $\mu(A \cap B_i) \leq \mu(B_i) < \infty$ 이므로 앞선 결론으로부터 $\mu(A) = \mu(\bigsqcup_{i=1}^{\infty}(A \cap B_i)) = \sum_{i=1}^{\infty} \mu(A \cap B_i) = \sum_{i=1}^{\infty} \nu(A \cap B_i) = \nu(\bigsqcup_{i=1}^{\infty}(A \cap B_i)) = \nu(A)$ 가 되어 증명이 끝난다. □

나중에 사용할 정리 하나를 증명하는 것으로 이번 절을 마무리하자. 다음 정리는 측도의 근사에 대한 정리로, semi-algebra $\mathscr{A}$가 생성하는 σ−대수 위에서 정의된 σ−유한한 측도 μ는 $\mathscr{A}$에서의 정보로 충분히 근사될 수 있음을 함의한다. 이는 μ도 결국에는 $\mathscr{A}$에서의 적당한 premeasure에 조금씩 정보를 더해나가며 확장시켜 얻은 것이라는 점을 생각하면 자연스러운 결과이다.

Theorem 2.41 공집합이 아닌 집합 X 위의 semi-algebra $\mathscr{A}$에 대해 μ를 $\sigma(\mathscr{A})$ 위의 σ−유한한 측도라 하자. 그렇다면 임의의 $A \in \sigma(\mathscr{A})$와 임의의 $\varepsilon > 0$에 대해 다음이 성립한다.

i. Semi-algebra $\mathscr{A}$에 속하는 적당한 서로소인 집합열 $\{A_i\}$가 존재하여 $A \subseteq \bigsqcup_{i=1}^\infty A_i$이고 $\mu(\bigsqcup_{i=1}^\infty A_i \setminus A) < \varepsilon$이다.
ii. 만약 $\mu(A) < \infty$이면 적당한 서로소인 $A_1, \cdots, A_k \in \mathscr{A}$가 존재하여 $\mu(A \triangle \bigsqcup_{i=1}^k A_i) < \varepsilon$이다.

Proof i. 먼저 $\mu(A) < \infty$인 경우부터 생각하자. Carathéodory 확장의 유일성으로부터 $(\mu|_{A(\mathscr{A})})^*|_{\sigma(\mathscr{A})} = \mu$이므로 $A(\mathscr{A})$에 속하는 적당한 A의 가산 덮개 $\{A_i\}$가 존재하여, WLOG, 필요하다면 공집합을 추가하여 이가 집합열을 이룬다고 하면 $\mu(\bigcup_{i=1}^\infty A_i) \leq \sum_{i=1}^\infty \mu(A_i) < \mu(A) + \varepsilon$이다. 이제 집합열 $\{B_i\}$를 $B_i := A_i \setminus \bigcup_{j=1}^{i-1} A_j$로 두면 이는 $A(\mathscr{A})$에 속하는 서로소인 집합열이고 $\bigcup_{i=1}^\infty A_i = \bigsqcup_{i=1}^\infty B_i$이므로 WLOG, 필요하다면 $\{A_i\}$를 $\{B_i\}$로 바꾸어 처음부터 $\{A_i\}$가 서로소인 집합열이라 해도 된다. 한편, 각 $i \in \mathbb{N}$에 대해 $A_i \in A(\mathscr{A})$이므로 정리 2.16로부터 서로소인 $A_{i1}, \cdots, A_{ik_i} \in \mathscr{A}$가 존재하여 $A_i = \bigsqcup_{j=1}^{k_i} A_{ij}$이고, 곧 $\mu(\bigsqcup_{i=1}^\infty \bigsqcup_{j=1}^{k_i} A_{ij}) = \sum_{i=1}^\infty \mu(\bigsqcup_{j=1}^{k_i} A_{ij}) = \sum_{i=1}^\infty \mu(A_i) < \mu(A) + \varepsilon$에서 $\mu(\bigsqcup_{i=1}^\infty \bigsqcup_{j=1}^{k_i} A_{ij} \setminus A) = \mu(\bigsqcup_{i=1}^\infty \bigsqcup_{j=1}^{k_i} A_{ij}) - \mu(A) < \varepsilon$이 되어 정리가 성립한다.

다음으로 $\mu(A) = \infty$인 경우를 생각하면, μ가 σ−유한하므로 $\mathscr{A}$에 속하는 적당한 집합열 $\{B_i\}$가 존재하여 각 $i \in \mathbb{N}$에 대해 $\mu(B_i) < \infty$이고 $\bigcup_{i=1}^\infty B_i = X$이다. 그렇다면 앞선 결과로부터 임의의 $\varepsilon > 0$과 각 $i \in \mathbb{N}$에 대해 $\mathscr{A}$에 속하는 적당한 서로소인 집합열 $\{A_{ij}\}_j$가 존재하여 $A \cap B_i \subseteq \bigsqcup_{j=1}^\infty A_{ij}$이고 $\mu(\bigsqcup_{j=1}^\infty A_{ij} \setminus (A \cap B_i)) < \varepsilon/2^i$이다. 이상으로부터 $A = \bigcup_{i=1}^\infty (A \cap B_i) \subseteq \bigcup_{i=1}^\infty \bigsqcup_{j=1}^\infty A_{ij}$이고 $\mu(\bigcup_{i=1}^\infty \bigsqcup_{j=1}^\infty A_{ij} \setminus A) \leq \mu(\bigcup_{i=1}^\infty [\bigsqcup_{j=1}^\infty A_{ij} \setminus (A \cap B_i)]) \leq \sum_{i=1}^\infty \mu(\bigsqcup_{j=1}^\infty A_{ij} \setminus (A \cap B_i)) < \sum_{i=1}^\infty \varepsilon/2^i = \varepsilon$인데, 이전과 같은 방법으로 WLOG, $\{A_{ij}\}_{ij}$가 서로소인 집합열이라 가정할 수 있으므로 곧 모든 경우에 대해 정리가 성립한다.

ii. i의 $\mu(A) < \infty$인 경우에 대한 증명에서와 같이 $A(\mathscr{A})$에 속하는 $\{A_i\}$를 택하여 $A \subseteq \bigsqcup_{i=1}^\infty A_i$이고 $\mu(\bigsqcup_{i=1}^\infty A_i) < \mu(A) + \varepsilon/2$이도록 할 수 있다. 따라서 집합열 $\{B_i\}$를 $B_i := \bigsqcup_{j=1}^i A_j$로 두면 $B_i \uparrow \bigsqcup_{j=1}^\infty A_j =: B$에서 정리 2.24로부터 $\mu(B_i) \uparrow \mu(B)$이고, 곧 충분히 큰 $i_0 \in \mathbb{N}$를 택하면 $\mu(B \setminus B_{i_0}) < \varepsilon/2$가 되어 $\mu(A \triangle \bigsqcup_{i=1}^{i_0} A_i) = \mu(A \setminus \bigsqcup_{i=1}^{i_0} A_i) + \mu(\bigsqcup_{i=1}^{i_0} A_i \setminus A) \leq \mu(B \setminus B_{i_0}) + \mu(\bigsqcup_{i=1}^\infty A_i \setminus A) < \varepsilon$이다. 이제 i의 증명에서와 같이 정리 2.16를 사용하면 원하는 결론을 얻는다. □

2.4 Construction of The Lebesgue Measure I

지금까지 Lebesgue 측도를 구성하기 위한 준비를 부단히 해 왔으니, 이제 때가 되었다. 이번 절에서는 전에 언급한 바와 같이 확장정리를 사용하여 Lebesgue 측도를 구성해보도록 하자.

Definition 2.42 구간 $I \subseteq \mathbb{R}$가 적당한 $x, y \in \mathbb{R}$에 대해 $(x, y], (-\infty, y], (x, \infty), (-\infty, \infty)$ 중 하나의 형태로 나타내어지면 이때의 I를 **반열린구간 (semi-open interval)**이라 한다. 나아가, 만약 집합 $B \in \mathbb{R}^n$가 반열린구간들의 Cartesian 곱으로 표현되면 이때의 B를 **semi-open box**라 한다.

이제부터는 $\mathbb{R}^n$의 모든 semi-open box의 모임을 $\mathscr{S}_n$로 쓰도록 하겠다.

Proposition 2.43 집합족 $\mathscr{S}_n$은 semi-algebra이다.

Proof 간결한 논의를 위해 $n = 1$인 경우만 생각하도록 하자. (일반적인 경우에도 이와 비슷하게 보일 수 있다.) 우선 $\mathscr{S}_1$이 비어있지 않음은 분명하다. 이제 임의의 $B, B' \in \mathscr{S}_1$에 대해 정의로부터 각각이 $(x, y], (-\infty, y], (x, \infty), (-\infty, \infty)$ 중 하나의 형태로 나타내어지므로, 총 16가지의 가능성이 있다. 이 16가지 경우 각각에 대해 잠시 생각해보면, 어느 경우이든 $B \cap B'$이 다시 반열린구간임을 쉽게 알 수 있고, 따라서 $B \cap B' \in \mathscr{S}_1$이다. 한편, $\mathscr{S}_1$이 semi-algebra가 되기 위한 다른 조건들도 비슷하게 따져볼 수 있다. □

Proposition 2.44 함수 $\rho_n : \mathscr{S}_n \to \overline{\mathbb{R}}_0^+$를

$$\rho_n : B \mapsto \begin{cases} \displaystyle\prod_{i=1}^{n}(y_i - x_i) & \text{집합 } B\text{가 적당한 } x, y \in \mathbb{R}\text{에 대해 } B = \prod_{i=1}^{n}(x_i, y_i]\text{로 표현되는 경우} \\ \infty & \text{ow. 즉, 집합 } B\text{가 유계가 아닌 경우} \end{cases}$$

로 정의하면 이는 $\mathscr{S}_n$ 위의 premeasure이다.

Proof 우선 $\rho_n(\emptyset) = 0$임은 분명하므로, ρ_n이 $\sigma-$가법성을 가짐을 보이는 것으로 충분하다. 이를 위해 $\mathscr{S}_n$에 속하는 임의의 서로소인 집합열 $\{B_i\}$를 생각하여 $B := \bigsqcup_{i=1}^{\infty} B_i \in \mathscr{S}_n$이라 하자. 여기서 만약 B_{i_0}가 유계가 아닌 $i_0 \in \mathbb{N}$가 존재한다면 이 경우에 $\sigma-$가법성은 자명하므로 각 $i \in \mathbb{N}$에 대해 B_i가 유계라 가정하고, 따라서 적당한 $x_i, y_i \in \mathbb{R}^n$에 대해 $B_i = \prod_{j=1}^{n}(x_i^j, y_i^j]$로 쓸 수 있다. 이제 경우를 나누어 $n = 1$인 경우를 먼저 생각하자.

$n = 1$인 경우에 B는 적당한 $x, y \in \mathbb{R}$에 대해 $(x, y], (-\infty, y], (x, \infty), (-\infty, \infty)$ 중 하나의 형태로 나타낼 수 있다. 첫 번째의 경우, 반드시 어떤 $i_0 \in \mathbb{N}$가 존재하여 $y_{i_0} = y$이므로 B_{i_0}부터 모든 $i \in \mathbb{N}$에 대해 $x_i = y_{i+1}$이고 $x_i \downarrow x$가 되도록 집합열 $\{B_i\}$를 재배열할 수 있다. 그렇다면 $\sum_{i=1}^{\infty} \rho_1(B_i) = \lim_{k\to\infty} \sum_{i=1}^{k}(y_i - x_i) = y - \lim_{k\to\infty} x_k = y - x = \rho_1(B)$

가 성립한다. 두 번째의 경우도 첫 번째와 동일하게 재배열하여 같은 결론을 얻을 수 있다. (단, 이때는 $x_i \to -\infty$가 되도록 한다.) 세 번째 경우, 우선 (x,∞)를 $I = (x, x_1]$과 $J = (x_1,\infty)$의 두 개의 구간으로 나누고, J에 속하는 $\{B_i\}$의 항들만 추려내어 부분열 $\{B_{i_j}\}_j$를 구성한다. 그렇다면, B_1은 분명 그 부분열에 포함되어 있을 것이므로 B_1부터 모든 $j \in \mathbb{N}$에 대해 $y_{i_j} = x_{i_{j+1}}$이고 $y_{i_j} \to \infty$가 되도록 집합열 $\{B_{i_j}\}_j$를 재배열할 수 있다. 그렇다면 $\sum_{j=1}^{\infty} \rho_1(B_{i_j}) = \lim_{k\to\infty} \sum_{j=1}^{k} (y_{i_j} - x_{i_j}) = \lim_{k\to\infty} y_{i_k} - x_1 = \infty$이고, 따라서 $\sum_{i=1}^{\infty} \rho_1(B_i) \geq \sum_{j=1}^{\infty} \rho_1(B_{i_j}) = \infty = \rho_1(B)$가 성립한다. 마지막 경우도 세 번째 경우와 동일하게 재배열하여 같은 결론을 얻을 수 있다. (단, 이때는 B를 $I = (-\infty, x_1]$과 $J = (x_1,\infty)$의 두 개의 구간으로 나눈다.) 이상으로부터 어떠한 경우에도 ρ_1이 σ−가법성을 가짐을 안다.

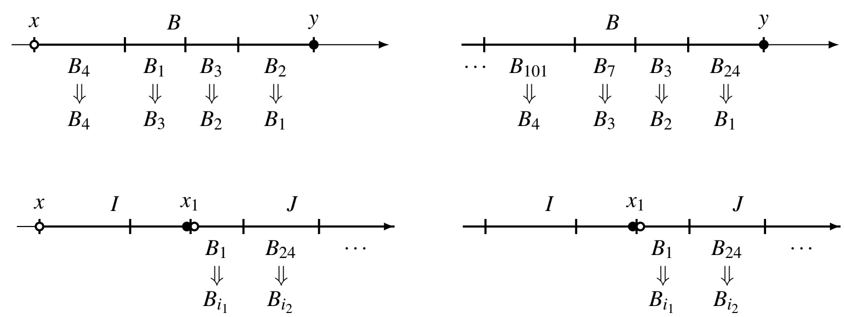


Figure 2.1 명제 2.44의 증명에서 $n = 1$인 경우 B의 형태로 가능한 네 가지 경우에 대한 $\{B_i\}$의 재배열 과정. 각각 $B = (x, y]$ (왼쪽 위), $B = (-\infty, y]$ (오른쪽 위), $B = (x,\infty)$ (왼쪽 아래), $B = (-\infty,\infty)$ (오른쪽 아래)인 경우를 나타낸다.

이제 일반적인 $n \in \mathbb{N}$의 경우를 생각하면 적당한 반열린구간 $I_1, \cdots, I_n \subseteq \mathbb{R}$에 대해 $B = \prod_{j=1}^{n} I_j$로 쓸 수 있다. 우선 $\{B_i\}$가 B의 'regular tilling'을 이루는 경우를 먼저 생각해 보자. 여기서 regular tilling을 이룬다고 함은 $\{B_i\}$가 B를 딱 맞아떨어지게 채워넣는 경우를 이르는 것으로, 보다 엄밀하게, 각 $j \leq n$에 대해 I_j를 가산개의 반열린구간 $I_{j1}, I_{j2}, \cdots$로 분할할 수 있어서 이러한 분할들이 이루는 semi-open box들이 모두 정확히 하나의 B_i와 일치하는 경우를 이른다.

이 경우, 만약 B가 유계라면 곧 모든 $I_1, \cdots, I_n$도 유계이고, 앞선 결과로부터

$$\begin{aligned}\rho_n(B) &= \prod_{j=1}^{n} \rho_1(I_j) \\ &= \prod_{j=1}^{n} \sum_{k=1}^{l_j} \rho_1(I_{jk})\end{aligned}$$

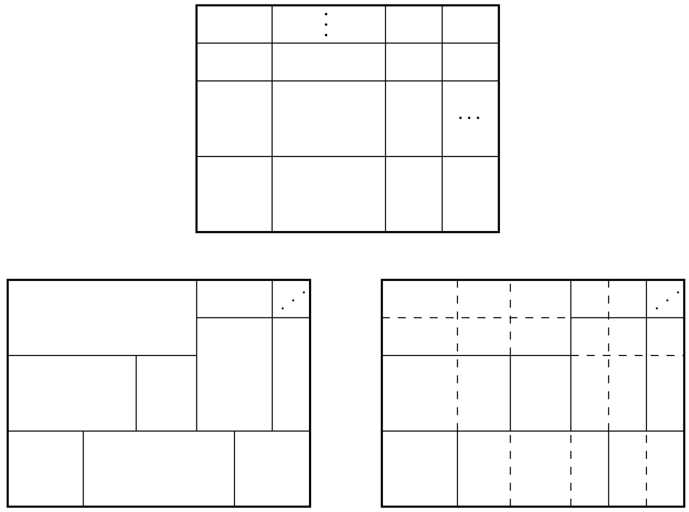


Figure 2.2 명제 2. 44의 증명에서 집합열 $\{B_i\}$가 B의 regular tilling을 이루는 경우. (위) 집합열 $\{B_i\}$가 regular tilling을 이루지 않는 경우에 이의 각 면을 연장하여 regular tilling을 이루도록 하는 과정. (아래)

$$
\begin{aligned}
&= \sum_{k_1=1}^{l_1} \cdots \sum_{k_n=1}^{l_n} \prod_{j=1}^{n} \rho_1(I_{jk_j}) \\
&= \sum_{k_1=1}^{l_1} \cdots \sum_{k_n=1}^{l_n} \rho_n\left(\prod_{j=1}^{n} I_{jk_j}\right) \\
&= \sum_{i=1}^{\infty} \rho_n(B_i)
\end{aligned}
$$

임을 안다. (여기서 $l_1, \cdots, l_n$은 유한할 수도 있고, ∞일 수도 있다.) 만약 B가 유계가 아니라면 WLOG, I_1이 유계가 아니라 해도 되고, 이번에도 앞선 결과로부터

$$
\begin{aligned}
\sum_{k=1}^{l_1} \rho_n\left(I_{1k} \times \prod_{j=2}^{n} I_{j1}\right) &= \left[\sum_{k=1}^{l_1} \rho_1(I_{1k})\right] \prod_{j=2}^{n} \rho_1(I_{j1}) \\
&= \rho_1(I_1) \prod_{j=2}^{n} \rho_1(I_{j1}) \\
&= \infty
\end{aligned}
$$

가 되어 $\sum_{i=1}^{\infty} \rho_n(B_i) \geq \sum_{k=1}^{l_1} \rho_n(I_{1k} \times \prod_{j=2}^{n} I_{j1}) = \infty = \rho(B)$를 얻는다. 다음으로, $\{B_i\}$가 B의 regular tilling을 이루지 않는 경우를 생각해보자. 이 경우에는 각 B_i의 면을 연장하여 B를 가산개의 semi-open box $C_1, C_2, \cdots$로 분할할 수 있으며 WLOG, 필요하다면 공집

합을 추가하여 이가 집합열 $\{C_j\}$ 을 이룬다고 해도 된다. (공집합은 $\emptyset = \prod_{j=1}^{n}(0,0]$ 에서 명백히 semi-open box이다.) 그렇다면 모든 $j \in \mathbb{N}$ 에 대해 $C_j \subseteq B_i$ 인 $i \in \mathbb{N}$ 가 유일하게 존재하며, $\{C_j\}$ 는 모든 B_i 와 B 의 regular tilling을 이룬다. 따라서 $\rho_n(B) = \sum_{j=1}^{\infty} \rho_n(C_j) = \sum_{i=1}^{\infty} \sum_{j:C_j \subseteq B_i} \rho_n(C_j) = \sum_{i=1}^{\infty} \rho_n(B_i)$ 이고, 곧 어떠한 경우에도 ρ_n 이 $\sigma-$가법성을 가짐을 알 수 있으므로 증명이 끝난다. □

이와 같이 정의된 $\mathscr{S}_n$ 위의 premeasure ρ_n 이 우리의 출발점이다. 여기에 정리 2. 30를 사용하면 $\mathscr{S}_n$ 위의 premeasure ρ_n 을 $A(\mathscr{S}_n)$ 에서의 premeasure로 유일하게 확장할 수 있고, 이제부터는 표기를 남용하여 이러한 유일한 확장을 ρ_n 이라 쓰도록 하겠다. 이러한 확장 ρ_n 이 $\sigma-$유한함은 쉽게 확인할 수 있다.

Proposition 2.45 대수 $A(\mathscr{S}_n)$ 위의 premeasure ρ_n 은 $\sigma-$유한하다.

Proof $\bigsqcup_{x \in \mathbb{Z}^n} \prod_{i=1}^{n}(x_i, x_i+1] = \mathbb{R}^n$ 이고 모든 $x \in \mathbb{Z}^n$ 에 대해 $\rho_n(\prod_{i=1}^{n}(x_i, x_i+1]) = 1$ 이므로 이는 자명하다. □

마지막 단계로, premeasure ρ_n 으로부터 유도되는 $\mathscr{P}(\mathbb{R}^n)$ 위의 외측도 ρ_n^* 을 생각하고 $\mathscr{M}_n$ 을 모든 ρ_n^*-가측집합의 모임이라 하자. 그렇다면 Carathéodory의 확장정리와 정리 2. 40로부터 $A(\mathscr{S}_n)$ 위의 premeasure ρ_n 을 $\mathscr{M}_n$ 위의 측도로 유일하게 확장할 수 있고, 이가 우리가 그토록 원하던 Lebesgue 측도이다.

Definition 2.46 위에서와 같이 확장된 $A(\mathscr{S}_n)$ 위의 premeasure ρ_n 에 대해 이의 Carathéodory 확장을 **Lebesgue 측도 (– measure)** 라 하고 λ_n 으로 쓴다. 또한, 이때 λ_n 의 정의역인 모든 ρ_n^*-가측집합들의 모임을 $\mathscr{M}_n$ 으로 쓰고, triple $(\mathbb{R}^n, \mathscr{M}_n, \lambda_n)$ 을 **Lebesgue 측도공간 (– measure space)** 이라 한다.

보통 특별히 정하지 않는 이상 $\mathbb{R}^n$ 에는 Lebesgue 측도 λ_n 이 주어진 것으로 보고, 이는 이 책에서도 마찬가지이다.

이로써 우리의 첫 번째 Lebesgue 측도의 구성이 일단락되었다. 그러나 아직 몇몇 답하기 어려운 질문들이 남아있다. 그 중 하나로, $\sigma-$대수 $\mathscr{M}_n$ 을 한 번 살펴보자. 과연 이는 어떤 집합들로 이루어져 있을까? Lebesgue 측도의 구성 과정에서 $A(\mathscr{S}_n)$ 의 구조까지는 쉽게 상상이 가지만 Carathéodory의 확장정리의 다소 현란한 확장 방식 덕에 $\mathscr{M}_n$ 의 구조는 쉽게 손에 잡히지 않는다. 우리가 앞으로 기본적인 측도공간으로 삼을 Lebesgue 측도공간의 구조를 모른다는 것은 그다지 유쾌한 일이 아니며, 이는 곧 Lebesgue 측도의 구성은 마쳤지만 $\mathbb{R}^n$ 의 부분집합 중에서 Lebesgue 측도로써 잴 수 있는 집합과 그렇지 않은 집합을 구별할 수 있는 기준은 여전히 모호하다는 뜻이다.

사실, $\mathscr{M}_n$ 은 생각보다 굉장히 큰 $\sigma-$대수이다. 우리가 상식적인 선에서 생각해 낼 수 있는 거의 대부분의 집합, 예컨대 (표준 위상에 대해) 열린집합, 닫힌집합, 가산 집합, box

나 ball, 혹은 $\mathbb{N}^n, \mathbb{Z}^n, \mathbb{Q}^n$ 따위의 집합들은 $\mathscr{M}_n$ 의 극히 일부분을 차지할 뿐이다. 집합족 $\mathscr{M}_n$ 의 대부분은 정말 상상조차 하기 힘든 괴상한 집합들로 이루어져 있어 $\mathscr{M}_n$ 은 그다지 경제적인 σ−대수는 못된다.[4]어쩌면 1절에서 지적한 것과 같이 측도를 부여할 집합으로 너무 많은 것들을 택한 것일 수도 있다. 그럼에도 불구하고 (이후에 자세히 보겠지만) 이 거대한 $\mathscr{M}_n$ 위에서 Lebesgue 측도는 놀라우리만치 바람직한 성질들을 가지기에 단순히 $\mathscr{M}_n$ 이 너무 크다는 이유로 힘들게 구성해낸 Lebesgue 측도공간을 포기하기는 뭔가 아쉽다. 이 상황을 타개할 현명한 해결책은 바로 $\mathscr{M}_n$ 중에서 이상한 집합들을 모두 걷어내고 '실용적' 인 집합들만 골라 $\mathscr{M}_n$ 보다 작은 σ−대수를 적당히 만든 다음 Lebesgue 측도를 그 σ−대수로 제한시키는 것이다. 이렇게 하면 이 새로운 측도공간은 Lebesgue 측도공간의 바람직한 성질들을 그대로 상속받으면서 훨씬 더 가볍고 경제적인 σ−대수로 구성되어 그 구조도 선명히 파악할 수 있을 것이다. 언뜻 보기에 이 과정 또한 Lebesgue 측도의 구성 못지않게 꽤나 복잡할 것 같지만, 사실 $\mathbb{R}^n$ 의 열린집합을 모으는 것만으로 충분하다.

Definition 2.47 집합족 $\mathscr{U}$ 를 $\mathbb{R}^n$ 의 모든 열린집합의 모임이라 하자. 이때 $\mathscr{U}$ 가 생성하는 σ−대수를 **Borel σ−대수 (− σ−algebra)**라 하고 $\mathscr{B}_n$ 으로 쓴다. 나아가, $\mathscr{B}_n$ 에 속하는 임의의 집합을 **Borel 집합 (− set)**이라 한다.

Proposition 2.48 집합 $\mathbb{R}^n$ 의 열린집합이나 닫힌집합에 교집합, 합집합, 여집합, 차집합 혹은 이들을 가산번 조합하여 정의한 연산을 취하여 얻는 집합은 Borel이다.

PROOF 임의의 닫힌집합 $F \subseteq \mathbb{R}^n$ 에 대해 그 정의로부터 F^c 가 열린집합이므로 F 는 Borel 이다. 이제 $\mathscr{B}_n$ 이 σ−대수로서 교집합, 합집합, 여집합, 차집합 그리고 이들을 가산번 조합한 연산에 대해 모두 닫혀있음을 생각하면 이 명제는 자명하다. □

단순히 $\mathbb{R}^n$ 의 열린집합을 모아 이들로 σ−대수를 생성한 것 뿐인데 우리가 상식적으로 떠올릴 수 있는 거의 대부분이 집합이 Borel σ−대수에 속한다는 사실을 알 수 있다. 따라서 $\mathscr{B}_n \subseteq \mathscr{M}_n$ 만 성립한다면 모든 문제가 해결된다. 이는 $\mathscr{B}_n$ 의 생성자를 조금 더 조사해보면 바로 알 수 있다.

Theorem 2.49 집합족 $\mathscr{C}$ 를 $\mathbb{R}^n$ 의 모든 유계인 열린 box의 모임이라 하면 이는 $\mathscr{B}_n$ 의 생성 자이다.

PROOF 집합족 $\mathscr{U}$ 를 $\mathbb{R}^n$ 의 모든 열린집합의 모임이라 하면 $\mathscr{C} \subseteq \mathscr{U}$ 이므로 $\sigma(\mathscr{C}) \subseteq \mathscr{B}_n$ 임은 분명하다. 이제 $\mathbb{R}^n$ 의 모든 열린집합은 유계인 열린 box의 가산 합집합으로 표현된다는 점[5]을 생각해보면 $\mathscr{U} \subseteq \sigma(\mathscr{C})$ 이고, 곧 $\mathscr{B}_n \subseteq \sigma(\mathscr{C})$ 가 되어 증명이 끝난다. □

Theorem 2.50 집합족 $\mathscr{S}_n$ 은 $\mathscr{B}_n$ 의 생성자이다.

Proof 간결한 논의를 위해 $n = 1$인 경우만 생각하자. (일반적인 경우에도 이와 비슷하게 보일 수 있다.) 또한 $\mathscr{C}$를 $\mathbb{R}$의 모든 유계인 열린 box의 모임이라 하자. 그렇다면 임의의 $B \in \mathscr{C}$에 대해 적당한 $x, y \in \mathbb{R}$가 존재하여 $B = (x, y)$이고, $(x, y) = \bigcup_{i=1}^{\infty}(x, y - 1/i]$에서 $\mathscr{C} \subseteq \sigma(\mathscr{S}_1)$이므로 곧 $\sigma(\mathscr{C}) \subseteq \sigma(\mathscr{S}_1)$이다. 이제 역의 포함관계를 보이기 위해 임의의 $I \in \mathscr{S}_1$를 택하면 정의로부터 I는 적당한 $x, y \in \mathbb{R}$에 대해 $(x, y], (-\infty, y], (x, \infty), (-\infty, \infty)$ 중 하나의 형태로 나타낼 수 있다. 이번에도 간결한 논의를 위해 첫 번째 경우만 생각해보자. (나머지 경우들도 비슷하게 따져보면 동일한 결론을 얻는다.) $I = (x, y]$라 하면 $I = (x, \infty) \setminus (y, \infty) = \bigcup_{i=1}^{\infty}(x, x + i) \setminus \bigcup_{i=1}^{\infty}(y, y + i)$이므로 $I \in \sigma(\mathscr{C})$이고, 곧 $\mathscr{S}_1 \subseteq \sigma(\mathscr{C})$가 되어 $\sigma(\mathscr{S}_1) \subseteq \sigma(\mathscr{C})$임을 안다. 이제 정리 2.49로부터 $\sigma(\mathscr{S}_1) = \mathscr{B}_1$이 되어 정리가 증명된다. □

여기서 $\mathscr{M}_n$이 $\mathscr{S}_n$을 포함한다는 사실을 떠올린다면, $\mathscr{B}_n \subseteq \mathscr{M}_n$임은 당연하다.[6]이제 우리는 다음을 정의할 수 있다.

Definition 2.51 Lebesgue 측도 λ_n에 대해 이의 $\mathscr{B}_n$으로의 제한 $\lambda_n|_{\mathscr{B}_n}$을 **Borel 측도 (– measure)**라 하고 μ_n으로 쓴다. 나아가, triple $(\mathbb{R}^n, \mathscr{B}_n, \mu_n)$을 **Borel 측도공간 (– measure space)**이라 한다.

나중에 사용하기 위해 $\mathscr{B}_n$의 생성자를 하나만 더 소개한다.

Theorem 2.52 집합족 $\mathscr{C}$를 어떤 $x \in \mathbb{R}^n$에 대해 $\prod_{i=1}^{n}(-\infty, x_i]$로 표현되는 $\mathbb{R}^n$의 모든 부분집합의 모임이라 하면 이는 $\mathscr{B}_n$의 생성자이다.

Proof 간결한 논의를 위해 $n = 1$인 경우만 생각하자. (일반적인 경우에도 이와 비슷하게 보일 수 있다.) 우선, $\mathscr{C} \subseteq \mathscr{S}_1$이므로 $\sigma(\mathscr{C}) \subseteq \sigma(\mathscr{S}_1)$임은 분명하다. 이제 역의 포함관계를 보이기 위해 임의의 $I \in \mathscr{S}_1$를 택하면 정의로부터 I는 적당한 $x, y \in \mathbb{R}$에 대해 $(x, y], (-\infty, y], (x, \infty), (-\infty, \infty)$ 중 하나의 형태로 나타낼 수 있다. 이번에도 간결한 논의를 위해 첫 번째 경우만 생각해보자. (나머지 경우들도 비슷하게 따져보면 동일한 결론을 얻는다.) $I = (x, y]$라 하면 $I = (-\infty, y] \setminus (-\infty, x]$이므로 $I \in \sigma(\mathscr{C})$이고, 곧 $\mathscr{S}_1 \subseteq \sigma(\mathscr{C})$가 되어 $\sigma(\mathscr{S}_1) \subseteq \sigma(\mathscr{C})$임을 안다. 이제 정리 2.50로부터 $\sigma(\mathscr{C}) = \mathscr{B}_1$이 되어 정리가 증명된다. □

Lebesgue 측도의 중요한 성질 몇 가지를 알아보자.

Theorem 2.53 임의의 유계인 $A \in \mathscr{M}_n$에 대해 $\lambda_n(A) < \infty$이다.

Proof 집합 A가 유계이므로 적당한 $M > 0$에 대해 $A \subseteq (-M, M]^n$이고, 곧 $\lambda_n(A) \leq \lambda_n((-M, M]^n) = (2M)^n < \infty$에서 이 정리는 자명하다. □

Theorem 2.54 Lebesgue 측도 λ_n는 regular하다. 즉, 다음 두 가지 성질이 성립한다.

i. 임의의 $A \in \mathscr{M}_n$ 에 대해 $\lambda_n(A) = \inf\{\lambda_n(U) \in \overline{\mathbb{R}}_0^+ : A \subseteq U \subseteq \mathbb{R}^n, U$는 열린집합$\}$ 이다.
ii. 임의의 $A \in \mathscr{M}_n$ 에 대해 $\lambda_n(A) = \sup\{\lambda_n(K) \in \overline{\mathbb{R}}_0^+ : K \subseteq A, K$는 compact 집합$\}$ 이다.

Proof i. 만약 $\lambda_n(A) = \infty$ 이면 어떠한 열린집합 $U \supseteq A$ 에 대해서도 $\lambda_n(U) \geq \lambda_n(A) = \infty$ 가 되어 정리가 자명하므로 $\lambda_n(A) < \infty$ 라 가정하자. 한편, 임의의 열린집합 $U \supseteq A$ 에 대해 $\lambda_n(U) \geq \lambda_n(A)$ 임이 분명하므로 임의의 $\varepsilon > 0$ 을 택하고 적당한 열린집합 $U \supseteq A$ 가 존재하여 $\lambda_n(U) < \lambda_n(A) + \varepsilon$ 임을 보이면 충분하다. 이를 위해, $\lambda_n(A)$ 의 정의가 $\lambda_n(A) = \inf\{\sum_{i=1}^{\infty} \rho_n(A_i) \in \overline{\mathbb{R}}_0^+ : \{A_i\}$는 $A(\mathscr{S}_n)$ 에 속하는 A 의 가산 덮개이다.$\}$ 라는 점을 상기한다면 $A(\mathscr{S}_n)$ 에 속하는 적당한 A 의 가산 덮개 $\{A_i\}$ 가 존재하여 WLOG, 필요하다면 공집합을 추가하여 이가 집합열을 이룬다고 하고 $\sum_{i=1}^{\infty} \rho_n(A_i) < \lambda_n(A) + \varepsilon$ 이다. 한편, 정리 2. 16 로부터 각 $i \in \mathbb{N}$ 에 대해 적당한 서로소인 $B_{i1}, \cdots, B_{ik_i} \in \mathscr{S}_n$ 이 존재하여 $A_i = \bigsqcup_{j=1}^{k_i} B_{ij}$ 이고, 따라서 WLOG, 필요하다면 A_i 를 $B_{i1}, \cdots, B_{ik_i}$ 로 바꾸어 $\{A_i\}$ 가 처음부터 $\mathscr{S}_n$ 에 속한다고 해도 된다. 또한, $\lambda_n(A) < \infty$ 이므로 각 A_i 도 유계이고 WLOG, 필요하다면 각 A_i 를 조금씩 크게 하여 $A \subseteq \bigcup_{i=1}^{\infty} A_i^{\circ} := U$ 라 해도 된다.[7] 그렇다면 U 가 열려있으므로 이는 가측이고, 곧 $\lambda_n(U) \leq \lambda_n(\bigcup_{i=1}^{\infty} A_i) \leq \sum_{i=1}^{\infty} \rho_n(A_i) \leq \lambda_n(A) + \varepsilon$ 이 되어 증명이 끝난다.

ii. 먼저 A 가 유계인 경우를 생각하면 적당한 $r > 0$ 에 대해 $A \subseteq B(r)$ 이다. 한편, 임의의 compact한 $K \subseteq A$ 에 대해 $\lambda_n(K) \leq \lambda_n(A)$ 이므로 임의의 $\varepsilon > 0$ 을 택하고 적당한 compact 집합 $K \subseteq A$ 가 존재하여 $\lambda_n(K) > \lambda_n(A) - \varepsilon$ 임을 보이면 충분하다. 이를 위해 앞서 보인 i을 이용하면 적당한 열린집합 $U \supseteq \overline{B(r)} \setminus A$ 가 존재하여 $\lambda_n(U) < \lambda(\overline{B(r)} \setminus A) + \varepsilon$ 이다. 여기서 $K = \overline{B(r)} \setminus U$ 라 하면 이는 compact하고 $\overline{B(r)} \subseteq U \sqcup K$ 에서 $\lambda_n(U) + \lambda_n(K) = \lambda_n(U \sqcup K) \geq \lambda_n(\overline{B(r)})$ 인데, 이는 곧 $\lambda_n(K) \geq \lambda_n(\overline{B(r)}) - \lambda_n(U) > \lambda_n(\overline{B(r)}) - \lambda_n(\overline{B(r)} \setminus A) - \varepsilon = \lambda_n(A) - \varepsilon$ 임을 뜻하므로 이 경우에 대해서는 증명이 끝난다. 이제 일반적인 A 에 대해 유계인 집합열 $\{A \cap B(i)\}_i$ 를 생각하면 $A \cap B(i) \uparrow A$ 이고 따라서 앞선 결론으로부터

$$\begin{aligned}
\lambda_n(A) &= \lim_{i \to \infty} \lambda_n(A \cap B(i)) \\
&= \sup_{r > 0} \lambda_n(A \cap B(r)) \\
&= \sup_{r > 0} \sup\{\lambda_n(K) \in \overline{\mathbb{R}}_0^+ : K \subseteq A \cap B(i), K\text{는 compact 집합}\} \\
&= \sup\{\lambda_n(K) \in \overline{\mathbb{R}}_0^+ : K \subseteq A, K\text{는 compact 집합}\}
\end{aligned}$$

이 성립하여 일반적인 경우에 대해서도 증명이 끝난다. □

Theorem 2.55 Lebesgue 측도 λ_n 는 이동 불변성을 갖는다. 즉, 임의의 $A \in \mathscr{M}_n$ 와 임의의 $x \in \mathbb{R}^n$ 에 대해 $A + x \in \mathscr{M}_n$ 이고 $\lambda_n(A + x) = \lambda_n(A)$ 이다.

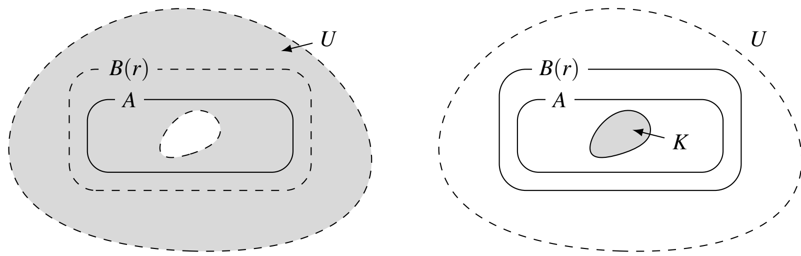


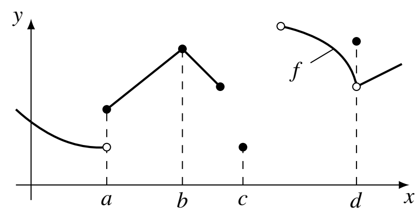
Figure 2.3 정리 2.54의 ii의 증명에서 집합 U (왼쪽)와 집합 K (오른쪽).

Proof 만약 $A \in A(\mathscr{S}_n)$ 이라면 이 정리는 자명하다. 이제 임의의 $B \subseteq \mathbb{R}^n$ 에 대해 $\{B_i\}$ 를 $A(\mathscr{S}_n)$ 에 속하는 B 의 가산 덮개라 하면 $\{B_i + x\}$ 는 $A(\mathscr{S}_n)$ 에 속하는 $B+x$ 의 가산 덮개가 되고, 이의 역도 성립하므로 $\rho_n^*(B) = \rho_n^*(B+x)$ 임을 안다. 이러한 결론이 임의의 $B \subseteq \mathbb{R}^n$ 에 대해 성립한다는 사실을 상기한다면, $A+x$가 가측이라는 사실만 보이는 것으로 충분함을 알 수 있다. 이를 위해 임의의 $S \subseteq \mathbb{R}^n$ 를 생각하면 A가 가측이라는 사실과 앞선 결과로부터 $\rho_n^*(S \cap (A+x)) + \rho_n^*(S \cap (A+x)^c) = \rho_n^*((S-x) \cap A) + \rho_n^*((S-x) \cap A^c) = \rho_n^*(S-x) = \rho_n^*(S)$ 가 되어 $A+x$가 가측임을 알고, 곧 증명이 끝난다. □

마지막으로 $\mathscr{B}_n$ 에 관한 놀라운 결과 하나를 소개하는 것으로 이번 절을 마치도록 하자. 이러한 결과는 $\mathscr{B}_n$ 이 $\mathscr{M}_n$ 보다 작은 σ−대수로 그 구조가 비교적 명확하다는 점에 크게 기인한다. ($\mathscr{B}_n$ 의 생성자는 서너개 알고 있는 반면, $\mathscr{M}_n$ 의 생성자나 구조는 여전히 오리무중이다.)

Definition 2.56 집합 $A \subseteq \mathbb{R}^n$ 에서 정의된 함수 $f : A \to \mathbb{R}$와 한 점 $x \in A$를 생각하자. 만약 임의의 $\varepsilon > 0$에 대해 적당한 $\delta > 0$가 존재하여 $||x-y|| < \delta$ 이고 $x \leq y$ 인 모든 $y \in A$ 에 대해 $|f(x) - f(y)| < \varepsilon$ 이 성립하면 f가 **x에서 오른쪽 연속 (right-continuous at x)** 혹은 **continuous from above at** x라 한다. 나아가, 만약 f가 모든 $x \in A$에서 오른쪽 연속이면 이때의 f를 **오른쪽 연속 (right-continuous)** 혹은 **continuous from above**라 한다. 이와 비슷하게 **x에서 왼쪽 연속 (left-continuous at x)**과 **왼쪽 연속 (left-continuous)**도 정의된다.

Figure 2.4 함수 f는 a에서는 오른쪽 연속이지만 연속이 아니고, b에서는 연속이면서 오른쪽 연속이며, c에서는 이가 고립점이므로 연속이면서 오른쪽 연속이고, d에서는 오른쪽 연속도, 연속도 아니다.



Definition 2.57 집합 $A \subseteq \mathbb{R}^n$ 에서 정의된 함수 $f : A \to \mathbb{R}$ 와 유계인 semi-open box $B \subseteq A$ 에 대해 B 가 적당한 $x, y \in \mathbb{R}^n$ 에 대해 $B = \prod_{i=1}^n (x_i, y_i]$ 와 같이 표현된다고 하자. 이때,

$$\sum_{i_1, \cdots, i_n \in \{0,1\}} (-1)^{i_1 + \cdots + i_n} f(x_1^{1-i_1} y_1^{i_1}, \cdots, x_n^{1-i_n} y_n^{i_n})$$

를 f 의 B **에서의 difference**라 하고 $\Delta_B f$ 로 쓴다.

Theorem 2.58 유한 측도공간 $(\mathbb{R}^n, \mathscr{B}_n, \mu)$ 에 대해 함수 $F_\mu : \mathbb{R}^n \to \mathbb{R}$ 를 $F_\mu : x \mapsto \mu(\prod_{i=1}^n (-\infty, x_i])$ 로 정의하면 다음이 성립한다.

i. 함수 F_μ 는 $\mu(\mathbb{R}^n)$ 에 의해 위로 유계이다.
ii. 임의의 유계인 semi-open box $B \subseteq \mathbb{R}^n$ 에 대해 $\Delta_B F_\mu = \mu(B) \geq 0$ 이다.
iii. 함수 F_μ 는 각 변수에 대해 증가한다.
iv. 함수 F_μ 는 오른쪽 연속이다.
v. 함수 F_μ 는 각 변수에 대해 오른쪽 연속이다.
vi. 각 $i \leq n$ 에 대해 $\lim_{x_i \to -\infty} F_\mu(x) = 0$ 이다. [8]
vii. $\lim_{x_1, \cdots, x_n \to \infty} F_\mu(x) = \mu(\mathbb{R}^n)$. [9]

Proof i. 임의의 $x \in \mathbb{R}^n$ 에 대해 $F_\mu(x) = \mu(\prod_{i=1}^n (-\infty, x_i]) \leq \mu(\mathbb{R}^n) < \infty$ 이므로 이는 자명하다.

ii. 간결한 논의를 위해 $n = 2$ 인 경우만 생각하도록 하자. ($n = 1$ 인 경우를 포함한 일반적인 경우도 이와 비슷하게 보일 수 있다.) 그렇다면 적당한 $x, y \in \mathbb{R}^2$ 에 대해 $B = (x_1, y_1] \times (x_2, y_2]$ 로 쓸 수 있으므로

$$\begin{aligned}
\Delta_B F_\mu &= F_\mu(y_1, y_2) - F_\mu(x_1, y_2) - F_\mu(y_1, x_2) + F_\mu(x_1, x_2) \\
&= \mu((-\infty, y_1] \times (-\infty, y_2]) - \mu((-\infty, x_1] \times (-\infty, y_2]) \\
&\qquad\qquad - \mu((-\infty, y_1] \times (-\infty, x_2] + \mu((-\infty, x_1] \times (-\infty, x_2]) \\
&= \mu((-\infty, y_1] \times (-\infty, y_2] \setminus (-\infty, x_1] \times (-\infty, y_2]) \\
&\qquad\qquad - \mu((-\infty, y_1] \times (-\infty, x_2] \setminus (-\infty, x_1] \times (-\infty, x_2]) \\
&= \mu((x_1, y_1] \times (-\infty, y_2]) - \mu((x_1, y_1] \times (-\infty, x_2]) \\
&= \mu((x_1, y_1] \times (-\infty, y_2] \setminus (x_1, y_1] \times (-\infty, x_2]) \\
&= \mu((x_1, y_1] \times (x_2, y_2]) \\
&= \mu(B)
\end{aligned}$$

가 성립한다.

iii. 간결한 논의를 위해 F_μ가 첫 번째 변수에 대해 증가한다는 것만 보이자. (다른 변수에 대해서도 이와 비슷하게 보일 수 있다.) 이제 임의의 $x \in \mathbb{R}^n$와 임의의 $h \geq 0$에 대해 $B = \prod_{i=1}^n (-\infty, x_i]$, $B' = (-\infty, x_1 + h] \times \prod_{i=2}^n (-\infty, x_i]$라 하면 $F_\mu(x_1 + h, \cdots) - F_\mu(x_1, \cdots) = \mu(B') - \mu(B) = \mu(B' \setminus B) \geq 0$에서 증명이 끝난다.

iv. 임의의 $x \in \mathbb{R}^n$를 택하여 $B = \prod_{i=1}^n (-\infty, x_i]$라 하고, 집합열 $\{B_j\}$를 $B_j := \prod_{i=1}^n (-\infty, x_i + 1/j]$로 두면 이는 B로 수렴하는 $\mathscr{S}_n$에 속하는 감소하는 집합열이므로 정리 2.24의 ii에서 $\mu(B_j) \downarrow \mu(B) = F_\mu(x)$이고, 곧 임의의 $\varepsilon > 0$을 택하면 적당한 $j_0 \in \mathbb{N}$가 존재하여 $\mu(B_{j_0}) - F_\mu(x) < \varepsilon$이다. 이제 $\delta = 1/j_0$라 하면 $||x - y|| < \delta$이고 $x \leq y$인 모든 $y \in \mathbb{R}^n$에 대해 $B \subseteq \prod_{i=1}^n (-\infty, y_i] \subseteq B_{j_0}$에서 $F_\mu(x) = \mu(B) \leq F_\mu(y) = \mu(\prod_{i=1}^n (-\infty, y_i]) \leq \mu(B_{j_0}) < F_\mu(x) + \varepsilon$이므로 곧 $|F_\mu(x) - F_\mu(y)| < \varepsilon$이 되어 증명이 끝난다.

v. 이는 iv로부터 자명하다.

vi. 간결한 논의를 위해 $i = 1$인 경우에 대해서만 보이도록 하자. (다른 경우에도 이와 비슷하게 보일 수 있다.) 이제 임의의 $x_2, \cdots, x_n \in \mathbb{R}$을 택하고 집합열 $\{B_j\}$를 $B_j := (-\infty, -j] \times \prod_{i=2}^n (-\infty, x_i]$로 두면 이는 $\emptyset$으로 수렴하는 $\mathscr{S}_n$에 속하는 감소하는 집합열이므로 정리 2.24의 ii에서 $\mu(B_j) \downarrow \mu(\emptyset) = 0$이고, 곧 임의의 $\varepsilon > 0$을 택하면 적당한 $j_0 \in \mathbb{N}$가 존재하여 $\mu(B_{j_0}) < \varepsilon$이다. 이제 $x_1 \leq -j_0$인 임의의 $x_1 \in \mathbb{R}$에 대해 $F_\mu(x_1, x_2, \cdots) \leq F_\mu(-j_0, x_2, \cdots) = \mu(B_{j_0}) < \varepsilon$이 되어 증명이 끝난다.

vii. 집합열 $\{B_j\}$를 $B_j := (-\infty, j]^n$으로 두면 이는 $\mathbb{R}^n$으로 수렴하는 $\mathscr{S}_n$에 속하는 증가하는 집합열이므로 정리 2.24의 i에서 $\mu(B_j) \uparrow \mu(\mathbb{R}^n)$이고, 곧 임의의 $\varepsilon > 0$을 택하면 적당한 $j_0 \in \mathbb{N}$가 존재하여 $\mu(\mathbb{R}^n) - \mu(B_{j_0}) < \varepsilon$이다. 이제 $x \geq j_0\mathbf{1}$인 임의의 $x \in \mathbb{R}^n$에 대해 $B_{j_0} \subseteq \prod_{i=1}^n (-\infty, x_i]$인데, 이와 i로부터 $\mu(\mathbb{R}^n) \geq F_\mu(x) = \mu(\prod_{i=1}^n (-\infty, x_i]) \geq F_\mu(j_0\mathbf{1}) = \mu(B_{j_0}) > \mu(\mathbb{R}^n) - \varepsilon$이므로 곧 $|F_\mu(x) - \mu(\mathbb{R}^n)| < \varepsilon$이 되어 증명이 끝난다. □

Theorem 2.59 함수 $F : \mathbb{R}^n \to \mathbb{R}$에 대해 이가

i. 함수 F는 오른쪽 연속이고 각 변수에 대해 증가한다.
ii. 유계인 semi-open box $B \subseteq \mathbb{R}^n$에 대해 $\Delta_B F \geq 0$이다.
iii. 각 $i \leq n$에 대해 $\lim_{x_i \to -\infty} F(x) = 0$이다.

를 만족하면 $F(x) = \mu(\prod_{i=1}^n (-\infty, x_i])$인 $\mathscr{B}_n$ 위의 측도 μ가 유일하게 존재한다.

Proof 집합족 $\mathscr{C}$를 어떤 $x \in \mathbb{R}^n$에 대해 $\prod_{i=1}^n (-\infty, x_i]$로 표현되는 $\mathbb{R}^n$의 모든 부분집합의 모임이라 하자. 일단 정리의 조건을 만족하는 측도가 존재하기만 한다면 정리 2.52로부터 $\mathscr{C}$가 $\mathscr{B}_n$의 생성자이고 이는 π-system이므로 정리 2.29에 의해 그 측도의 유일성은 자명하다. 따라서 존재성을 보이는 것으로 증명은 충분하다. 이를 위해 우리는 Lebesgue 측도를 구성할 때 썼던 방법을 비슷하게 사용하려고 한다. 우선 $\mathscr{S}_n$ 위의 적당한 premeasure ρ를 정의하고, 이에 정리 2.30와 Carathéodory의 확장정리를 연달아 사용하여 ρ를 측도로

확장한 다음, 이를 다시 $\mathscr{B}_n$ 으로 제한하는 것이다. 또한, 간결한 논의를 위해 $n=2$ 인 경우만 생각하도록 하자. ($n=1$ 인 경우를 포함한 일반적인 경우에도 이와 비슷하게 보일 수 있다.)

먼저 함수 $\rho : \mathscr{S}_2 \to \overline{\mathbb{R}}_0^+$ 를

$$\rho : B \mapsto \begin{cases} \Delta_B F & \text{집합 } B\text{가 유계인 경우} \\ \infty & \text{ow.} \end{cases}$$

로 정의하면 $\rho(\emptyset)=0$ 임은 분명하다. 이제 이가 $\sigma-$가법성을 가진다는 것만 보이면 ρ 가 premeasure임을 알 수 있는데, 이를 위해서는 보조정리 2.38로부터 ρ 가 가법성과 $\sigma-$반가법성을 가짐을 보이는 것으로 충분하다. 먼저 가법성을 보이기 위해 임의의 서로소인 $B_1,\cdots,B_l \in \mathscr{S}_2$ 를 생각하여 $B := \bigsqcup_{i=1}^{l} B_i \in \mathscr{S}_2$ 이라 하면 적당한 반열린구간 $I_1, I_2 \subseteq \mathbb{R}$ 에 대해 $B = I_1 \times I_2$ 로 쓸 수 있다. 만약 $B_1,\cdots,B_l$ 중에서 유계가 아닌 집합이 있다면 당연히 B 도 유계가 아니게 되어 이 경우에는 가법성이 자명하므로 $B_1,\cdots,B_l$ 이 모두 유계라 하자. 이제 경우를 나누어 $B_1,\cdots,B_l$ 이 B 의 regular tilling을 이루는 경우를 먼저 생각해보자. (Regular tilling을 이룬다는 것의 의미는 명제 2.44의 증명을 참조하기 바란다.) 이 경우에는 $\sum_{i=1}^{l} \rho(B_i) = \sum_{i=1}^{l} \Delta_{B_i} F = \Delta_B F = \rho(B)$ 가 되어 ρ 가 가법성을 가진다. 이제 $B_1,\cdots,B_l$ 이 regular tilling을 이루지 않는 경우를 생각해보자. 이 경우에는 $B_1,\cdots,B_l$ 들의 각 면을 연장하여 B 를 유한개의 semi-open box $C_1,\cdots,C_{l'}$ 로 분할할 수 있으며, 모든 $k \le l'$ 에 대해 $C_k \subseteq B_i$ 인 $i \le l$ 가 유일하게 존재하여 $C_1,\cdots,C_{l'}$ 은 모든 $B_1,\cdots,B_l$ 과 B 의 regular tilling을 이룬다. 따라서 앞선 결과로부터 $\rho(B) = \Delta_B F = \sum_{k=1}^{l'} \Delta_{C_k} F = \sum_{i=1}^{l} \sum_{k:C_k \subseteq B_i} \Delta_{C_k} F = \sum_{i=1}^{l} \Delta_{B_i} F = \sum_{i=1}^{l} \rho(B_i)$ 가 되어 어느 경우에나 ρ 는 가법성을 가짐을 안다.

다음으로 ρ 가 $\sigma-$반가법성을 가짐을 보이기 위해 $\mathscr{S}_2$ 에 속하는 임의의 집합열 $\{B_i\}$ 를 생각하여 $B := \bigcup_{i=1}^{\infty} B_i \in \mathscr{S}_2$ 이라 하자. 한편, 만약 B_{i_0} 가 유계가 아닌 $i_0 \in \mathbb{N}$ 가 존재한다면 당연히 B 도 유계가 아니게 되어 이 경우에는 $\sigma-$반가법성이 자명하므로 모든 $i \in \mathbb{N}$ 에 대해 B_i 가 유계라고 하자. 비슷하게, B 가 유계가 아닌 경우에도 $\sigma-$반가법성이 자명하므로 B 도 유계라 하자. 그렇다면 각 $i \in \mathbb{N}$ 에 대해 적당한 $x_i, y_i \in \mathbb{R}^2$ 가 존재하여 $B_i = (x_i^1, y_i^1] \times (x_i^2, y_i^2]$ 로 쓸 수 있고, 비슷하게 적당한 $x, y \in \mathbb{R}^2$ 가 존재하여 $B = (x^1, y^1] \times (x^2, y^2]$ 로 쓸 수 있다. 한편, F 가 오른쪽 연속이므로 임의의 $\varepsilon > 0$ 을 택하면 각 $i \in \mathbb{N}$ 에 대해 적당한 $\delta_i > 0$ 가 존재하여

$$\begin{aligned} |F(y_i^1+\delta_i, y_i^2+\delta_i) - F(y_i^1, y_i^2)| &< \varepsilon/3\cdot 2^{i+1}, \\ |F(x_i^1, y_i^2+\delta_i) - F(x_i^1, y_i^2)| &< \varepsilon/3\cdot 2^{i+1}, \\ |F(y_i^1+\delta_i, x_i^2) - F(y_i^1, x_i^2)| &< \varepsilon/3\cdot 2^{i+1} \end{aligned}$$

이고, 곧 $B_i' := (x_i^1, y_i^1+\delta_i] \times (x_i^2, y_i^2+\delta_i] \supseteq B_i$ 로 두면

$$\begin{aligned}
|\rho(B_i') - \rho(B_i)| &= |\Delta_{B_i'} F - \Delta_{B_i} F| \\
&= |[F(y_i^1+\delta_i, y_i^2+\delta_i) - F(x_i^1, y_i^2+\delta_i) - F(y_i^1+\delta_i, x_i^2) + F(x_i^1, x_i^2)] \\
&\qquad - [F(y_i^1, y_i^2) - F(x_i^1, y_i^2) - F(y_i^1, x_i^2) + F(x_i^1, x_i^2)]| \\
&\le |F(y_i^1+\delta_i, y_i^2+\delta_i) - F(y_i^1, y_i^2)| + |F(x_i^1, y_i^2+\delta_i) - F(x_i^1, y_i^2)| \\
&\qquad + |F(y_i^1+\delta_i, x_i^2) - F(y_i^1, x_i^2)| \\
&< \varepsilon/2^{i+1}.
\end{aligned}$$

이 성립한다. 비슷하게,

$$\begin{aligned}
|F(x^1+\delta, x^2+\delta) - F(x^1, x^2)| &< \varepsilon/6, \\
|F(x^1+\delta, y^2) - F(x^1, y^2)| &< \varepsilon/6, \\
|F(y^1, x^2+\delta) - F(y^1, x^2)| &< \varepsilon/6
\end{aligned}$$

인 $\delta > 0$ 도 존재하여 $B' := (x^1+\delta, y^1] \times (x^2+\delta, y^2] \subseteq B$ 로 두면 $|\rho(B') - \rho(B)| < \varepsilon/2$ 이 성립한다. 여기서 $B' \subseteq \overline{B'} \subseteq B \subseteq \bigcup_{i=1}^{\infty} (B_i')^\circ$ 이므로 $\{(B_i')^\circ\}$ 는 compact한 $\overline{B'}$ 의 열린 덮개가 되어 이의 유한 부분덮개가 존재하고, WLOG, 필요하다면 relabel하여 그 유한 부분덮개를 $\{(B_1')^\circ, \cdots, (B_k')^\circ\}$ 라 할 수 있다. 이제 $B' \subseteq \bigcup_{i=1}^{k} (B_i')^\circ \subseteq \bigcup_{i=1}^{k} B_i'$ 에 보조정리 2. 19의 ii를 적용하면 $\rho(B) < \rho(B') + \varepsilon/2 \le \sum_{i=1}^{k} \rho(B_i') + \varepsilon/2 < \sum_{i=1}^{\infty} [\rho(B_i) + \varepsilon/2^{i+1}] + \varepsilon/2 = \sum_{i=1}^{\infty} \rho(B_i) + \varepsilon$ 이 되어 ρ 가 σ−반가법성을 가짐을 안다.

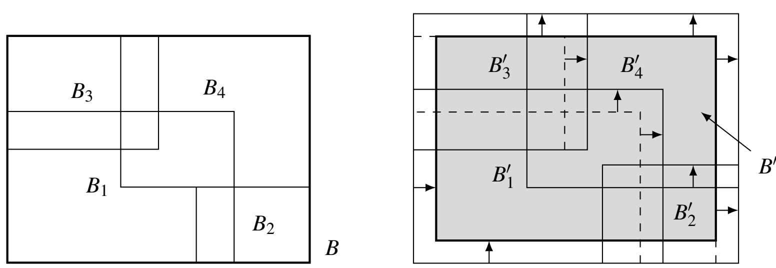


Figure 2.5 정리 2. 59의 증명에서의 집합 B, B_i (왼쪽)와 집합 B', B_i' (오른쪽).

이제 ρ 가 $\mathscr{S}_2$ 위의 premeasure임을 알았으므로 앞서 설명한 바와 같이 이를 $\mathscr{B}_2$ 위의 측도로 확장하는 과정은 자연스럽게 이루어진다. 남은 일은 과연 이러한 확장을 통해 얻은 측도 μ 가 정리의 조건을 만족하는지를 확인하는 일인데, 이는 어렵지 않다. 우선 임의의 $x \in \mathbb{N}$ 에 대해 집합열 $\{B_i\}$ 를 $B_i := (x^1 - i, x^1] \times (x^2 - i, x^2]$ 로 두면 이는 $B_i \uparrow (-\infty, x^1] \times (-\infty, x^2]$ 인 $\mathscr{S}_2$ 에 속하는 증가하는 집합열임이 분명하고 μ 의 구성 과정

으로부터 $\mu(B_i) = \Delta_{B_i} F = F(x^1, x^2) - F(x^1 - i, x^2) - F(x^1, x^2 - i) + F(x^1 - i, x^2 - i)$ 이다. 그런데 가정으로부터 $F(x^1 - i, x^2), F(x^1, x^2 - i) \to 0$ 이고 F 의 함숫값은 항상 0보다 크거나 같으므로 (만약 어떤 $x \in \mathbb{R}^2$ 에 대해 $F(x) < 0$ 이라면 F 가 첫 번째 변수에 대해 감소한다는 사실로부터 $0 = \lim_{i\to\infty} F(x^1 - i, x^2) \le F(x) < 0$ 의 모순이 발생하므로 모든 $x \in \mathbb{R}^2$ 에 대해 $F(x) \ge 0$ 이다.) $F(x^1 - i, x^2 - i) \le F(x^1 - i, x^2) \to 0$ 에서 $F(x^1 - i, x^2 - i) \to 0$ 이 되어 $\mu((-\infty, x^1] \times (-\infty, x^2]) = \lim_{i\to\infty} \mu(B_i) = F(x)$ 임을 안다. □

이상의 결과는 가측공간 $(\mathbb{R}^n, \mathscr{B}_n)$ 위의 유한 측도 μ 에 대한 정보는 정리 2. 58에서와 같이 정의된 함수 F_μ 에 모두 담겨있음을 뜻한다. 따라서 F_μ 만 알고 있다면 원래의 측도 μ 를 완벽히 복원해낼 수 있다. 이 놀라운 결과는 한참 후에 CDF를 소개하며 다시 등장할 것이다.

2.5 Completion of Measures

측도가 0인 집합을 생각해보자. 상식적으로, 이 집합의 모든 부분집합은 0의 측도를 가질 것 같지만, 아쉽게도 이가 항상 성립하는 것은 아니다. 어떤 부분집합은 가측이 아닐수도 있기 때문이다. 이렇게 측도가 0인 집합의 부분집합이 가측이 아닐 수도 있다는 사실은 측도론의 전개에 있어 상당히 신경쓰이는 부분이다. 이번 절에서는 Lebesgue 측도의 구성이라는 우리의 목표에서 잠시 옆으로 빠져 이런 문제를 해결하는 방법에 대해 알아본다. 그리고 놀랍게도, 이를 공부함으로써 우리에게 숙제로 남겨진 $\mathscr{M}_n$ 의 구조에 대한 질문에 어느 정도 만족스러운 답을 할 수 있게 될 것이다.

Definition 2.60 측도공간 $(X, \mathscr{A}, \mu)$ 와 부분집합 $N \subseteq X$ 에 대해 만약 $\mu(A) = 0$ 이고 $N \subseteq A$ 인 $A \in \mathscr{A}$ 가 존재하면 이때의 N 을 **$(\mu-)$ 영집합 ($(\mu-)$null set)**이라 한다. 또한, 만약 모든 영집합이 $\mathscr{A}-$가측이면 이때의 측도공간 $(X, \mathscr{A}, \mu)$ 를 **완비측도공간 (complete measure space)**이라 한다.

어떤 측도공간을 완비측도공간으로 만드는 체계적인 방법 즉, 완비화 과정을 찾는 것이 이번 절에서의 우리의 목표이다. 우선 영집합에 대한 몇몇 기본적인 성질들을 보자.

Proposition 2.61 측도공간 $(X, \mathscr{A}, \mu)$ 에 대해 다음이 성립한다.

i. 영집합 N 의 임의의 부분집합 $M \subseteq N$ 도 영집합이다.
ii. 영집합으로 구성된 집합열 $\{N_i\}$ 에 대해 $\bigcup_{i=1}^{\infty} N_i$ 도 영집합이다.

Proof i. 정의로부터 적당한 $A \in \mathscr{A}$ 가 존재하여 $M \subseteq N \subseteq A$ 이고 $\mu(A) = 0$ 이므로 이 명제는 자명하다.

ii. 이번에도 정의로부터 각 $i \in \mathbb{N}$에 대해 적당한 $A_i \in \mathscr{A}$가 존재하여 $N_i \subseteq A_i$이고 $\mu(A_i) = 0$이므로 곧 $\bigcup_{i=1}^{\infty} N_i \subseteq \bigcup_{i=1}^{\infty} A_i$이고 $\mu(\bigcup_{i=1}^{\infty} A_i) \leq \sum_{i=1}^{\infty} \mu(A_i) = 0$이 되어 $\bigcup_{i=1}^{\infty} N_i$도 영집합이다. □

영집합이라는 용어에서 알 수 있듯이 영집합은 '거의' 공집합과 같다. 물론, 아예 아무런 원소를 포함하지 않는 빈 집합인 공집합과 원소를 포함할 수는 있지만 이들이 마치 여기저기 흩어진 점과 같아 의미있는 측도를 이루지 못하는 영집합이 엄밀하게는 서로 다르지만, 이러한 차이점은 많은 경우에 흐려져 영집합을 마치 공집합과 같이 사용하는 경우가 측도론에서는 흔하다. 특히 어떤 성질을 만족하지 않는 경우가 영집합을 이루는 경우, 마치 이런 예외가 아예 없는 것과 같이 생각하곤 한다. 이는 Lebesgue 적분론을 배우기 시작하면 충분히 느낄 수 있으니, 일단은 앞서 설명한 것과 같이 완비화 과정을 정립하는 것에 집중하자.

Proposition 2.62 측도공간 $(X, \mathscr{A}, \mu)$와 모든 영집합들의 모임 $\mathscr{N}$에 대해 $\sigma(\mathscr{A} \cup \mathscr{N}) = \{A \cup N \subseteq X : A \in \mathscr{A}, N \in \mathscr{N}\}$이다.

Proof 표기의 편의를 위해 $\overline{\mathscr{A}} = \{A \cup N \subseteq X : A \in \mathscr{A}, N \in \mathscr{N}\}$라 하자. 그렇다면 $\emptyset \in \mathscr{A} \cap \mathscr{N}$에서 $\mathscr{A} \cup \mathscr{N} \subseteq \overline{\mathscr{A}}$임은 분명하다. 또한, $\mathscr{A} \cup \mathscr{N}$을 포함하는 모든 σ−대수가 $\overline{\mathscr{A}}$를 포함해야 함도 분명하므로 $\overline{\mathscr{A}}$가 σ−대수임을 보이는 것으로 증명은 충분하다. 우선 $\mathscr{A}$가 비어있지 않으므로 $\overline{\mathscr{A}}$도 비어있지 않다. 한편, 임의의 $B \in \overline{\mathscr{A}}$에 대해 적당한 $A \in \mathscr{A}, N \in \mathscr{N}$이 존재하여 $B = A \cup N$인데, WLOG, 필요하다면 N을 $N \setminus A$로 바꾸어 A와 N이 처음부터 서로소라고 해도 된다. 또한, 정의로부터 적당한 $C \in \mathscr{A}$가 존재하여 $\mu(C) = 0$이고 $N \subseteq C$인데, $\mu(C \setminus A) \leq \mu(C) = 0$이고 $N \subseteq C \setminus A$이므로 WLOG, 필요하다면 C를 $C \setminus A$로 바꾸어 A와 C가 처음부터 서로소라고 해도 된다. 그렇다면 $B^c = (A \sqcup N)^c = (A \sqcup C)^c \cup (C \setminus N)$에서 $B^c \in \overline{\mathscr{A}}$가 되어 $\overline{\mathscr{A}}$가 여집합에 대해 닫혀있음을 안다. 마지막으로 $\overline{\mathscr{A}}$에 속하는 임의의 집합열 $\{B_i\}$에 대해 이번에도 정의로부터 각 $i \in \mathbb{N}$에 대해 적당한 $A_i \in \mathscr{A}, N_i \in \mathscr{N}$가 존재하여 $B_i = A_i \cup N_i$로 쓸 수 있다. 그렇다면 $\bigcup_{i=1}^{\infty} B_i = \bigcup_{i=1}^{\infty} A_i \cup \bigcup_{i=1}^{\infty} N_i$에서 명제 2. 61의 ii로부터 $\bigcup_{i=1}^{\infty} N_i$도 영집합이므로 $\bigcup_{i=1}^{\infty} B_i \in \overline{\mathscr{A}}$가 되어 $\overline{\mathscr{A}}$가 가산 합집합에 대해서도 닫혀있음을 알고, 따라서 이는 σ−대수이다. □

Proposition 2.63 측도공간 $(X, \mathscr{A}, \mu)$에 대해 $A, B \in \mathscr{A}$이고 $N, M \subseteq X$이 영집합이라 하자. 만약 $A \cup N = B \cup M$이면 $\mu(A) = \mu(B)$이다.

Proof 집합 M이 영집합이므로 적당한 $C \in \mathscr{A}$가 존재하여 $\mu(C) = 0$이고 $M \subseteq C$이다. 이는 곧 $A \subseteq A \cup N = B \cup M \subseteq B \cup C$에서 $\mu(A) \leq \mu(B \cup C) \leq \mu(B) + \mu(C) = \mu(B)$임을 의미하는데, 이와 비슷하게 $\mu(A) \geq \mu(B)$도 보일 수 있으므로 증명이 끝난다. □

Theorem 2.64 측도공간 $(X,\mathscr{A},\mu)$와 모든 $\mu-$영집합들의 모임 $\mathscr{N}$에 대해 $\overline{\mathscr{A}}=\{A\cup N\subseteq X : A\in\mathscr{A}, N\in\mathscr{N}\}$라 하고 함수 $\overline{\mu}:\overline{\mathscr{A}}\to\overline{\mathbb{R}}_0^+$를 $\overline{\mu}:A\cup N\mapsto\mu(A)$로 정의하자. (여기서 $A\in\mathscr{A}, N\in\mathscr{N}$이다.) 그렇다면 $\overline{\mu}$는 μ의 $\overline{\mathscr{A}}$에서의 측도로의 유일한 확장이고 $(X,\overline{\mathscr{A}},\overline{\mu})$는 완비측도공간이 된다.

Proof 명제 2.62, 2.63로부터 $\overline{\mathscr{A}}$는 $\mathscr{A}$를 포함하는 $\sigma-$대수이며 $\overline{\mu}$는 well-define된다. 이제 $\overline{\mu}$가 $\overline{\mathscr{A}}$ 위의 측도임을 보이자. 이를 위해 $\overline{\mathscr{A}}$에 속하는 임의의 서로소인 집합열 $\{B_i\}$를 생각하면 각 $i\in\mathbb{N}$에 대해 적당한 $A_i\in\mathscr{A}, N_i\in\mathscr{N}$가 존재하여 $B_i=A_i\cup N_i$이고 $\{A_i\}$와 $\{N_i\}$가 명백히 서로소이므로, $\overline{\mu}(\bigsqcup_{i=1}^{\infty}B_i)=\overline{\mu}(\bigsqcup_{i=1}^{\infty}A_i\cup\bigsqcup_{i=1}^{\infty}N_i)=\mu(\bigsqcup_{i=1}^{\infty}A_i)=\sum_{i=1}^{\infty}\mu(A_i)=\sum_{i=1}^{\infty}\overline{\mu}(B_i)$이다. 한편, $\overline{\mu}(\emptyset)=0$임은 분명하므로 $\overline{\mu}$는 명백히 $\overline{\mathscr{A}}$ 위의 측도이다. 그렇다면 $\emptyset$이 $\mu-$영집합이라는 사실로부터 $\overline{\mu}$가 μ의 확장임은 자명하다.

다음으로, 측도공간 $(X,\overline{\mathscr{A}},\overline{\mu})$의 완비성을 보이자. 이를 위해 임의의 $B\in\overline{\mathscr{A}}$를 택하여 $\overline{\mu}(B)=0$이라 하면, 적당한 $A\in\mathscr{A}, N\in\mathscr{N}$이 존재하여 $B=A\cup N$이고 $\mu(A)=\overline{\mu}(B)=0$이다. 이는 $A\in\mathscr{N}$임을 뜻하므로 임의의 $M\subseteq B$에 대해 $M\in\mathscr{N}\subseteq\overline{\mathscr{A}}$이고, 곧 B의 임의의 부분집합이 $\overline{\mathscr{A}}-$가측이 되어 $(X,\overline{\mathscr{A}},\overline{\mu})$가 완비측도공간임을 안다.

마지막으로 유일성을 보이기 위해 ν를 μ의 $\overline{\mathscr{A}}$에서의 측도로의 확장이라 하고 $B\in\overline{\mathscr{A}}$를 임의로 하나 택하자. 그렇다면 적당한 $A\in\mathscr{A}$와 $N\in\mathscr{N}$이 존재하여 $B=A\cup N$이고, 다시 적당한 $C\in\mathscr{A}$가 존재하여 $N\subseteq C$이고 $\mu(C)=0$이다. 이로부터 $\nu(B)=\nu(A\cup N)\geq\nu(A)=\mu(A)=\overline{\mu}(B)$이고 $\nu(B)=\nu(A\cup N)\leq\nu(A)+\nu(N)\leq\nu(A)+\nu(C)=\mu(A)=\overline{\mu}(B)$에서 $\nu=\overline{\mu}$가 되어 원하던 유일성을 얻는다. □

이상으로부터 다음과 같은 well-define된 완비화를 얻을 수 있다.

Definition 2.65 측도공간 $(X,\mathscr{A},\mu)$와 모든 $\mu-$영집합들의 모임 $\mathscr{N}$에 대해 $\overline{\mathscr{A}}=\{A\cup N\subseteq X : A\in\mathscr{A}, N\in\mathscr{N}\}$라 하고 함수 $\overline{\mu}:\overline{\mathscr{A}}\to\overline{\mathbb{R}}_0^+$를 $\overline{\mu}:A\cup N\mapsto\mu(A)$로 정의하자. (여기서 $A\in\mathscr{A}, N\in\mathscr{N}$이다.) 이때의 완비측도공간 $(X,\overline{\mathscr{A}},\overline{\mu})$를 측도공간 $(X,\mathscr{A},\mu)$의 **완비화 (completion)**라 한다.

위의 정의에서와 같이 일반적으로 측도공간의 완비화를 통해 얻는 $\sigma-$대수와 측도는 각각 원래의 $\sigma-$대수와 측도에 overline을 그어 표기한다. 나중을 위해 정리 하나를 소개한다.

Theorem 2.66 측도공간 $(X,\mathscr{A},\mu)$와 부분집합 $A\subseteq X$에 대해 $A\in\overline{\mathscr{A}}$일 필요충분조건은 적당한 $B,C\in\mathscr{A}$가 존재하여 $B\subseteq A\subseteq C$이고 $\mu(C\setminus B)=0$인 것이다.

Proof 충분조건임을 보이기 위해 $A\in\overline{\mathscr{A}}$라 하면 적당한 $B\in\mathscr{A}$와 영집합 $N\subseteq X$이 존재하여 $A=B\cup N$이며, 다시 적당한 $D\in\mathscr{A}$가 존재하여 $N\subseteq D$이고 $\mu(D)=0$이다. 이제 $C=B\cup D$라 하면 $B\subseteq A\subseteq C$이고 $\mu(C\setminus B)\leq\mu(D)=0$이 되어 충분조건임이 보여진다. 다음으로, 필요조건임을 보이기 위해 적당한 $B,C\in\mathscr{A}$가 존재하여 $B\subseteq A\subseteq C$이고 $\mu(C\setminus B)=0$이라

하자. 그렇다면 $A = (A \setminus B) \sqcup B$ 이고 $A \setminus B \subseteq C \setminus B$ 에서 $A \setminus B$ 가 영집합이므로 $A \in \overline{\mathscr{A}}$ 가 되어 필요조건임도 보여진다. □

이번 절의 남은 부분에서는 앞서 예고했던 바와 같이 이러한 완비화가 $\mathscr{M}_n$ 의 구조에 대한 질문으로 어떻게 연결되는지 살펴보려고 한다. 우선 Carathéodory의 확장정리로 잠시 돌아가자.

Theorem 2.67 공집합이 아닌 집합 X 위의 대수 $\mathscr{A}$ 에 대해 ρ 를 $\mathscr{A}$ 위의 premeasure라 하고, $\mathscr{M}$ 을 모든 ρ^*−가측집합들의 모임이라 하자. 또한 μ 를 ρ 의 Carathéodory 확장이라 하면 $(X, \mathscr{M}, \mu)$ 는 완비측도공간이다.

Proof 집합 $N \subseteq X$ 이 영집합이라 하면 정의로부터 적당한 $A \subseteq \mathscr{M}$ 가 존재하여 $\mu(A) = 0$ 이고 $N \subseteq A$ 이므로 임의의 $S \subseteq X$ 에 대해 $\rho^*(S \cap N) + \rho^*(S \cap N^c) \leq \rho^*(A) + \rho^*(S \cap N^c) = \mu(A) + \rho^*(S \cap N^c) = \rho^*(S \cap N^c) \leq \rho^*(S)$ 가 성립하여 $N \in \mathscr{M}$ 이고, 곧 $(X, \mathscr{M}, \mu)$ 는 완비측도공간임을 안다. □

따라서 Carathéodory의 확장정리를 통해 얻는 모든 측도공간은 완비성을 가지고, 특히 Lebesgue 측도공간은 완비측도공간이다. 그렇다면 조심스럽게, 어쩌면 Lebesgue 측도공간이 Borel 측도공간의 완비화가 아닐까하는 추측을 해 볼 수도 있다. 이 추측의 결론은 다음과 같다.

Theorem 2.68 Lebesgue 측도공간 $(\mathbb{R}^n, \mathscr{M}_n, \lambda_n)$ 은 Borel 측도공간 $(\mathbb{R}^n, \mathscr{B}_n, \mu_n)$ 의 완비화이다. 즉, $(\mathbb{R}^n, \mathscr{M}_n, \lambda_n) = (\mathbb{R}^n, \overline{\mathscr{B}_n}, \overline{\mu_n})$.

Proof 정리 2.64로부터 완비화를 통해 얻는 σ−대수로의 측도의 확장은 유일하므로 $\overline{\mathscr{B}_n} = \mathscr{M}_n$ 을 보이는 것으로 충분하다. 우선 $\overline{\mathscr{B}_n} \subseteq \mathscr{M}_n$ 부터 보이자. 임의의 집합 $A \in \overline{\mathscr{B}_n}$ 를 택하면 적당한 $B \in \mathscr{B}_n$ 와 μ_n−영집합 $N \subseteq \mathbb{R}^n$ 이 존재하여 $A = B \cup N$ 이고, 다시 적당한 $C \in \mathscr{B}_n$ 가 존재하여 $N \subseteq C$ 이고 $\mu_n(C) = 0$ 이다. 이제 Lebesgue 측도공간은 정리 2.67로부터 완비측도공간이므로 $N \in \mathscr{M}_n$ 이 되어 $A = B \cup N \in \mathscr{M}_n$ 에서 곧 $\overline{\mathscr{B}_n} \subseteq \mathscr{M}_n$ 이다.

역의 포함 관계를 보이는 것은 조금 더 복잡하다. 임의의 $A \in \mathscr{M}_n$ 를 택하고 먼저 $\lambda_n(A) < \infty$ 인 경우를 생각하자. 이 경우, 정리 2.54로부터 각 $i \in \mathbb{N}$ 에 대해 compact한 $K_i \in \mathscr{B}_n$ 와 열린 $U_i \in \mathscr{B}_n$ 가 존재하여 $K_i \subseteq A \subseteq U_i$ 이고 $\lambda_n(A) - 1/i < \lambda_n(K_i) \leq \lambda_n(A) \leq \lambda_n(U_i) < \lambda_n(A) + 1/i$ 이므로 이들로써 집합열 $\{K_i\}, \{U_i\}$ 를 구성할 수 있으며 WLOG, 필요하다면 각각의 항을 $F_i := \bigcup_{j=1}^{i} K_j, V_i := \bigcap_{j=1}^{i} U_j$ 로 바꾸어 $\{K_i\}, \{U_i\}$ 각각을 증가하고 감소하는 집합열이라 해도 된다. 그렇다면 그 구성으로부터 $\{K_i\}, \{U_i\}$ 는 각각 $K := \bigcup_{i=1}^{\infty} K_i, U := \bigcap_{i=1}^{\infty} U_i$ 로 수렴하고, 곧 $K \subseteq A \subseteq U$ 이며 $\lambda_n(K) = \lambda_n(A) = \lambda_n(U)$ 이다. 따라서 $\mu_n(U \setminus K) = \lambda_n(U) - \lambda_n(K) = 0$ 이 되어 정리 2.66로부터 $A \in \overline{\mathscr{B}_n}$ 임을 안다. 한편, $\lambda_n(A) = \infty$ 인 경우에는 각 $i \in \mathbb{N}$ 에 대해 $A \cap B(i)$ 가 유계이므로 정리 2.53로부터 이는 유한한 Lebesgue 측도를 가지고, 곧

앞선 결과로부터 $A\cap B(i)\in\overline{\mathscr{B}_n}$ 이 되어 적당한 $B_i\in\mathscr{B}_n$ 과 μ_n–영집합 $N_i\subseteq\mathbb{R}^n$ 가 존재하여 $A\cap B(i)=B_i\cup N_i$ 이다. 이로부터 $A=\bigcup_{i=1}^{\infty}[A\cap B(i)]=\bigcup_{i=1}^{\infty}B_i\cup\bigcup_{i=1}^{\infty}N_i$ 인데 명제 2.61의 ii에서 $\bigcup_{i=1}^{\infty}N_i$ 도 μ_n–영집합이므로 $A\in\overline{\mathscr{B}_n}$ 임을 안다. 곧 어느 경우에나 $A\in\overline{\mathscr{B}_n}$ 이므로 $\overline{\mathscr{B}_n}\supseteq\mathscr{M}_n$ 이 되어 증명이 끝난다. □

실로 놀라운 결과가 아닐 수 없다! 우리는 $\mathscr{M}_n$ 의 구조를 전혀 모르는 상태에서 단지 이가 너무 크다는 생각에 이 중에서 쓸만하다고 생각되는 극히 일부의 집합만 뽑아내어 $\mathscr{B}_n$ 을 구성했는데, $\mathscr{B}_n$ 은 정말로 $\mathscr{M}_n$ 에서 중요한 원소들로만 이루어진 알짜배기 σ–대수였던 것이다.

2.6 Measurable Functions

이제 Lebesgue 측도를 다른 방법으로 다시 구성하는 것으로 본 장의 후반부를 시작한다. Lebesgue 측도의 두 번째 구성 또한 기본적으로는 확장을 이용한 방법으로, 이번에는 마치 n 개의 $\mathbb{R}$ 에 Cartesian 곱을 취하여 $\mathbb{R}^n$ 을 얻는 것과 비슷하게 n 개의 Lebesgue 측도공간 $(\mathbb{R},\mathscr{M}_1,\lambda_1)$ 에 적당한 곱연산을 취하여 $(\mathbb{R}^n,\mathscr{M}_n,\lambda_n)$ 을 얻어보려고 한다. 다만, 이번에도 $\mathscr{M}_n$ 이 너무 거대한 탓에 곱연산을 취한다고 해서 곧바로 $\mathscr{M}_n$ 을 얻지는 못하는데, 다행히 이는 완비화를 통해 해결할 수 있다. 이러한 방법은 Carathéodory의 확장정리를 비롯한 확장정리들을 사용한 방법보다 훨씬 더 직관적인 방법인데, 그럼에도 불구하고 이 방법을 두 번째로 소개하는 까닭은 이때 사용할 '적당한 곱연산' 을 정의하기가 꽤나 어렵기 때문이다. 마치 배보다 배꼽이 더 큰 격이다. 특히, 이 곱연산을 정의하려면 Lebesgue 적분론의 전개가 반드시 필요하므로 일단은 Lebesgue 적분론에 집중하도록 하고, 이 과정이 끝나면 다시 원래의 목표로 돌아와 Lebesgue 측도의 두 번째 구성을 마무리하도록 하겠다.

앞서 초록에서 지적한 바와 같이 Lebesgue 적분론은 우리가 측도론을 배우는 주된 이유 중 하나이다. 물론, Lebesgue 적분론에서 극한과 적분의 상호작용이 Riemann 적분론에서보다 자연스럽다는 것도 Lebesgue 적분론의 장점 중 하나이지만, Lebesgue 적분론의 가장 큰 수학적 의의는 측도가 정의되는 임의의 공간에서 적분론을 전개할 수 있도록 적분의 개념을 확장했다는 데 있다. 우리가 $\mathbb{R}^2$ 나 $\mathbb{R}^3$ 에 한정되었던 넓이나 부피의 개념을 일반적인 공간으로 일반화하여 측도의 개념을 정립하였으니 어찌보면 이는 당연한 수순이다.

Lebesgue 적분론을 정립하기 위한 준비로서, 과연 어떤 함수를 적분의 대상으로 삼아야 하는지에 대해 생각해보자. 당연히, 더 이상 함수의 정의역이 $\mathbb{R}^n$ 의 부분집합일 필요가 없다. 정의역에서 적당한 측도가 정의되기만 하면 충분하다. 한편, Riemann 적분론에서 properly 적분가능한 함수는 반드시 유계여야 했는데, 이 또한 왠지 불필요하게 과도한 제한이라 생각된다. 결론부터 말하자면, Lebesgue 적분론에서 적분의 대상이 되는 함수는

가측함수들이다. (가측함수라고 해서 모두 적분가능하다는 뜻이 아니다. 마치 Riemann 적분론에서 유계함수라고 해서 모두 적분가능하지는 않은 것처럼 Lebesgue 적분론에서도 적분불가능한 가측함수가 얼마든지 있다.) 이번 절에서는 이러한 가측함수에 대해 알아보도록 하겠다.

Definition 2.69 가측공간 $(X, \mathscr{A}), (Y, \mathscr{B})$와 함수 $f : X \to Y$를 생각하자. 만약 임의의 $A \in \mathscr{B}$에 대해 $f^{-1}(A) \in \mathscr{A}$이면 이때의 f를 **($\mathscr{A}/\mathscr{B}$−) 가측함수 (($\mathscr{A}/\mathscr{B}$−) measurable function)**라 한다. 특히, 만약 $\mathscr{A}$가 Borel σ−대수이면 이때의 f를 **Borel 함수 (− function)**라 한다.

Theorem 2.70 가측공간 $(X, \mathscr{A}), (Y, \mathscr{B})$와 함수 $f : X \to Y$에 대해 $\mathscr{C}$를 $\mathscr{B}$의 생성자라 하면 f가 가측일 필요충분조건은 임의의 $A \in \mathscr{C}$에 대해 $f^{-1}(A) \in \mathscr{A}$인 것이다.

Proof 충분조건임은 정의로부터 자명하므로 필요조건임만 보이면 된다. 이를 위해 $\mathscr{D} = \{A \subseteq Y : f^{-1}(A) \in \mathscr{A}\}$를 생각하면 가정으로부터 $\mathscr{C} \subseteq \mathscr{D}$에서 $\mathscr{D}$는 비어있지 않다. 또한, 임의의 $A \in \mathscr{D}$에 대해 $f^{-1}(A^c) = [f^{-1}(A)]^{-1} \in \mathscr{A}$에서 $A^c \in \mathscr{D}$가 되어 $\mathscr{D}$는 여집합에 대해 닫혀있다. 마지막으로 $\mathscr{D}$에 속하는 임의의 집합열 $\{A_i\}$에 대해 $f^{-1}(\bigcup_{i=1}^{\infty} A_i) = \bigcup_{i=1}^{\infty} f^{-1}(A_i) \in \mathscr{A}$에서 $\bigcup_{i=1}^{\infty} A_i \in \mathscr{D}$이므로 $\mathscr{D}$는 가산 합집합에 대해서도 닫혀있어서 곧 이는 σ−대수이고 $\mathscr{B} \subseteq \mathscr{D}$가 되어 증명이 끝난다. □

적분론에서는 이론 전개상의 편의를 위해 실수체계에 $\pm\infty$를 추가하는 것이 편하다. 이렇게 실수에 $\pm\infty$를 추가한 수 체계를 **확장된 실수 (extended real number)**라 하며 $\overline{\mathbb{R}}$로 쓴다. 우리가 $\mathbb{R}$에서 성립하는 것으로 알고있는 거의 대부분의 성질들, 예컨대 LUBP, GLBP, MSP 따위의 성질들은 $\overline{\mathbb{R}}$에서도 그대로 (사실은 더 잘) 성립한다. 다만, $\overline{\mathbb{R}}$에서는 $\mathbb{R}$에서의 표준 거리함수가 더 이상 제 역할을 하지 못하므로 (당장 $d(-\infty, \infty)$를 뭐라고 해야 할지부터가 대략 난감하다.) $\overline{\mathbb{R}}$에서의 열림과 닫힘을 정의하기 위해서는 새로운 아이디어가 필요하다. 여기에서는 위상수학의 내용을 피하기 위해 결론만 받아들이도록 하자. 우리는 집합 $U \subseteq \overline{\mathbb{R}}$가 적당한 $x, y \in \mathbb{R}$에 대해 $(x, y), (x, \infty], [-\infty, y)$ 중 하나로 표현되는 구간들의 가산 합집합으로 표현되면 이때의 U를 열린집합이라 정의한다. 이제 닫힌집합의 정의는 $\mathbb{R}$에서와 마찬가지이다. 한편, $\mathbb{R}$에서의 열린집합이 유계인 열린구간의 가산 합집합으로 표현된다는 점을 상기하면 (각주 5 참조.), 이는 $\overline{\mathbb{R}}$에서도 여전히 열려있음을 알 수 있고, 곧 $\mathscr{U}, \mathscr{U}'$를 각각 $\mathbb{R}$과 $\overline{\mathbb{R}}$에서의 모든 열린집합의 모임이라 하면 $\mathscr{U} \subseteq \mathscr{U}'$이 성립한다.[10]

이제 $\mathscr{B}_1$을 정의한 것과 비슷하게 $\overline{\mathbb{R}}$ 위의 σ−대수 $\mathscr{B}_1'$을 $\mathscr{U}'$이 생성하는 σ−대수로 정의하자.[11] 그렇다면 $\mathscr{U} \subseteq \mathscr{U}'$에서 $\mathscr{B}_1 \subseteq \mathscr{B}_1'$이 성립한다. 사실, 대부분의 경우 WLOG, 실함수의 공역을 $\overline{\mathbb{R}}$라 해도 되므로 $\mathscr{B}_1$과 $\mathscr{B}_1'$을 엄밀하게 구분해야 할 필요가 그렇게 많지 않고, 따라서 표기를 남용하여 $\mathscr{B}_1'$을 그냥 $\mathscr{B}_1$로 쓴다. 나아가 정리 2.49, 2.50, 2.52의 증명을 비슷하게 반복하면 $\overline{\mathbb{R}}$ 위의 $\mathscr{B}_1$에 대해 다음 결과를 얻는다.

Theorem 2.52° 집합족 $\mathscr{C}$ 를 어떤 $x \in \mathbb{R}$ 에 대해 $[-\infty, x]$ 로 표현되는 모든 구간의 모임이라 하면 이는 $\mathscr{B}_1$ 의 생성자이다.

한편, 앞서 가측함수를 굉장히 일반적으로 정의했지만, 적분론의 전개를 목표로 하는 우리는 함수의 공역이 $\overline{\mathbb{R}}$ 이거나 $\mathbb{R}^n$ 인 두 가지 경우에 특히 관심이 있다. 그리고 예외적으로, $\overline{\mathbb{R}}$ 이나 $\mathbb{R}^n$ 이 함수의 공역을 쓰인 경우, 여기에는 Lebesgue 측도가 아닌 Borel 측도가 주어진 것으로 가정한다. 따라서 다음 보조정리가 자주 쓰인다.

Corollary 2.71 가측공간 $(X, \mathscr{A})$ 에서 정의된 함수 $f : X \to \mathbb{R}^n$ (혹은 $f : X \to \overline{\mathbb{R}}$) 에 대해 f 가 가측일 필요충분조건은 임의의 $x \in \mathbb{R}^n$ (혹은 $x \in \mathbb{R}$) 에 대해 $f^{-1}(\prod_{i=1}^n (-\infty, x_i]) \in \mathscr{A}$ (혹은 $f^{-1}([-\infty, x]) \in \mathscr{A}$) 인 것이다.

Proof 이는 정리 2. 52°, 2. 70로부터 자명하다. □

정의만 가지고는 어떤 함수가 가측인지의 여부를 직접 판단하는 것은 쉬운 일이 아니므로 이에 도움이 될 만한 정리들을 다수 소개한다.

Proposition 2.72 가측공간 $(X, \mathscr{A})$ 에 대해 집합 $A \subseteq X$ 가 가측일 필요충분조건은 $\mathbf{1}_A$ 가 가측인 것이다.

Proof 정의로부터 $\mathbf{1}_A$ 가 가측인 것과 임의의 $x \in \mathbb{R}$ 에 대해 집합

$$\mathbf{1}_A^{-1}([-\infty, x]) = \begin{cases} X & x \geq 1 \text{ 인 경우} \\ A^c & 0 \leq x < 1 \text{ 인 경우} \\ \emptyset & \text{ow.} \end{cases}$$

이 가측인 것은 동치이고, 이는 다시 A 가 가측이라는 것과 동치이므로 이 명제는 자명하다. □

Theorem 2.73 가측공간 $(X, \mathscr{A})$ 에서 정의된 가측함수 $f : X \to \mathbb{R}^m$ 와 Borel 함수 $g : \mathbb{R}^m \to \mathbb{R}^n$ (혹은 $g : \mathbb{R}^m \to \overline{\mathbb{R}}$) 에 대해 합성 $g \circ f$ 는 가측이다.

Proof 임의의 $x \in \mathbb{R}^n$ 를 택하면 가정으로부터 $g^{-1}(\prod_{i=1}^n (-\infty, x_i])$ 가 Borel이고, 따라서 $(g \circ f)^{-1}(\prod_{i=1}^n (-\infty, x_i]) = f^{-1}(g^{-1}(\prod_{i=1}^n (-\infty, x_i]))$ 가 가측이 되어 $g \circ f$ 가 가측임을 안다.

한편, $g : \mathbb{R}^m \to \overline{\mathbb{R}}$ 의 경우에도 이와 비슷하게 보일 수 있다. □

위의 정리에서 함수 g 가 Borel이라는 조건은 필수적이다. 만약 g 가 단지 가측함수에 그친다면, f, g 는 모두 가측이지만 그 합성이 가측이 아닌 반례가 존재한다.[12]

Theorem 2.74 모든 연속함수 $f : \mathbb{R}^m \to \mathbb{R}^n$ 는 Borel이다.

Proof 임의의 $x \in \mathbb{R}^n$ 에 대해 $\prod_{i=1}^n (-\infty, x_i]$ 가 닫혀있으므로 그 역상 $f^{-1}(\prod_{i=1}^n (-\infty, x_i])$ 도 닫혀있고, 곧 Borel이므로 f 는 Borel이다. □

Theorem 2.75 가측공간 $(X,\mathscr{A})$에서 정의된 함수 $f:X\to\mathbb{R}^n$에 대해 f가 가측일 필요충분조건은 f의 각 성분 $f_1,\cdots,f_n:X\to\mathbb{R}$이 가측인 것이다.

Proof 충분조건임을 보이기 위해 임의의 $x\in\mathbb{R}$를 택하면 각 $i\le n$에 대해 $\mathbb{R}^{i-1}\times(-\infty,x]\times\mathbb{R}^{n-i}$이 닫혀있으므로 곧 Borel이 되어 가정으로부터 $f_i^{-1}((-\infty,x])=f^{-1}(\mathbb{R}^{i-1}\times(-\infty,x]\times\mathbb{R}^{n-i})$가 가측이고, 따라서 f_i가 가측이 되어 충분조건임이 보여진다. 반대로, 필요조건임을 보이기 위해 임의의 $x\in\mathbb{R}^n$를 택하면 가정으로부터 모든 $i\le n$에 대해 $f_i^{-1}((-\infty,x_i])$가 가측이므로 $f^{-1}(\prod_{i=1}^n(-\infty,x_i])=\bigcap_{i=1}^n f_i^{-1}((-\infty,x_i])$가 가측이고, 따라서 f가 가측이 되어 필요조건임이 보여진다. □

Theorem 2.76 가측공간 $(X,\mathscr{A})$에서 정의된 가측함수 $f,g:X\to\mathbb{R}^n$(혹은 $f,g:X\to\overline{\mathbb{R}}$)와 가측함수 $h:X\to\mathbb{R}$(혹은 $h:X\to\overline{\mathbb{R}}$)에 대해 $f\pm g, hf, f/h$는 모두 가측이다. 단, 연산의 과정에서 0으로 나누게 되거나 $\infty-\infty,\infty/\infty,0\cdot\infty$와 같은 부정형이 등장하여 그 결과가 well-define되지 않는 경우 이를 $c\in\mathbb{R}$와 같은 dummy value로 두는 관례를 전제하자.

Proof 각 $i\le n$에 대해 정리 2.75로부터 f_i,g_i는 가측이고, 함수 $+,-,\cdot:\mathbb{R}^2\to\mathbb{R}$가 모두 연속함수이므로 정리 2.74로부터 이들은 Borel이다. 그렇다면 정리 2.73로부터 $f_i\pm g_i, hf_i$가 모두 가측이며, 다시 정리 2.75에서 $f\pm g, hf$가 가측임을 안다. 마지막으로, f/h가 가측임을 보이기 위해서는 $1/h$가 가측임을 보이는 것으로 충분한데, $h^{-1}(0)=\emptyset$인 경우에는

$$(1/h)^{-1}((-\infty,x])=\begin{cases}h^{-1}((-\infty,0))\cup h^{-1}([1/x,\infty)) & x>0\text{인 경우}\\ h^{-1}((-\infty,0)) & x=0\text{인 경우}\\ h^{-1}([1/x,0)) & \text{ow.}\end{cases}$$

이므로 $1/h$는 가측이다. 이제 일반적인 h에 대해서도 $A=h^{-1}(0)$에 대해 이가 닫힌집합이므로 곧 Borel이고 $1/h=(1/h)\mathbf{1}_{X\setminus A}+c\mathbf{1}_A$에서 $1/h$가 가측이 되어 증명이 끝난다. 한편, $f,g,h:X\to\overline{\mathbb{R}}$인 경우에도 비슷하게 하면 된다. □

Theorem 2.77 가측공간 $(X,\mathscr{A})$에서 정의된 가측함수 $f_i:X\to\overline{\mathbb{R}}$의 열 $\{f_i\}$에 대해 다음이 성립한다.

i. 함수 $\sup_{i\in\mathbb{N}}f_i, \inf_{i\in\mathbb{N}}f_i, \limsup_{i\to\infty}f_i, \liminf_{i\to\infty}f_i$는 모두 가측이다.
ii. 만약 모든 $x\in X$에 대해 $\lim_{i\to\infty}f_i(x)$가 존재하면 이는 가측이다.
iii. 집합 $\{x\in X:\{f_i(x)\}\text{가 수렴}\}$이 가측이다.
iv. 임의의 가측인 $f:X\to\overline{\mathbb{R}}$에 대해 집합 $\{x\in X:f_i(x)\to f(x)\}$가 가측이다.

한편, 가측함수 $f_i:X\to\mathbb{R}^n$의 열 $\{f_i\}$에 대해서도 i을 제외한 나머지 성질들이 성립한다. 다만, iv의 경우 가측함수 $f:X\to\mathbb{R}^n$에 대해 성립한다

Proof i. 임의의 $x \in \mathbb{R}$에 대해 $\{y \in X : \sup_{i\in\mathbb{N}} f_i(y) \le x\} = \bigcap_{i=1}^{\infty} f_i^{-1}([-\infty, x])$가 가측이므로 곧 $\sup_{i\in\mathbb{N}} f_i$가 가측이고, 비슷하게 $\inf_{i\in\mathbb{N}} f_i$도 가측임을 보일 수 있다. 한편, $\limsup_{i\to\infty} f_i = \inf_{j\in\mathbb{N}} \sup_{i\ge j} f_i$이고 $\liminf_{i\to\infty} f_i = \sup_{j\in\mathbb{N}} \inf_{i\ge j} f_i$이므로 앞선 결과로부터 이들 또한 가측이다.

ii. 이는 $\lim_{i\to\infty} f_i = \limsup_{i\to\infty} f_i = \liminf_{i\to\infty} f_i$에서 자명하다.

iii. 정리에서 주어진 집합이 곧 $\{x \in X : \limsup_{i\to\infty} f_i(x) = \liminf_{i\to\infty} f_i(x)\}$이므로 자명하다.

iv. 정리에서 주어진 집합이 곧 $\{x \in X : \limsup_{i\to\infty} f_i(x) = \liminf_{i\to\infty} f_i(x) = f(x)\}$이므로 자명하다.

한편, 가측함수 $f_i : X \to \mathbb{R}^n$의 열 $\{f_i\}$와 가측함수 $f : X \to \mathbb{R}^n$의 경우에는 앞선 결과를 성분별로 적용하면 된다. □

이전 절에서 공집합과 영집합 간의 차이가 흐려져 마치 공집합과 영집합을 같은 것처럼 생각하는 경우가 많다고 하였는데, 이를 조금 더 명확히 하기 위해 '거의 어디서나' 라는 개념을 소개한다. 이 뭔가 굉장히 수학스럽지 않은 개념은 측도론의 발전이 낳은 사고방식 중 하나이다.

Definition 2.78 측도공간 $(X, \mathscr{A}, \mu)$에서 적당한 영집합 $N \subseteq X$이 존재하여 어떤 성질이 모든 $x \in X \setminus N$에 대해 성립한다면 그 성질은 X의 **(μ−) 거의 어디서나 ((μ−)almost everywhere)** 혹은 **(μ−) 거의 대부분의 $x \in X$에 대해 ((μ−)almost every $x \in X$)** 성립한다고 한다.

어떤 성질 P가 거의 어디서나 성립한다는 것을 때로는 P (ae.)라 줄여 간단히 표기할 때도 있다. 한편, 거의 어디서나 성립하는 성질 중 가운데 특히 자주 쓰이는 것이 바로 **거의 어디서나 같다 (almost everywhere equality)**는 개념이다.

Proposition 2.79 거의 어디서나 같다는 관계는 동치관계이다. 즉, 측도공간 $(X, \mathscr{A}, \mu)$와 집합 Y, 함수 $f, g, h : X \to Y$에 대해 다음이 성립한다.

i. (반사성) $f = f$ (ae.)
ii. (대칭성) $f = g$ (ae.)이면 $g = f$ (ae.)이다.
iii. (전이성) $f = g$ (ae.)이고 $g = h$ (ae.)이면 $f = h$ (ae.)이다.

Proof 반사성과 대칭성은 자명하므로 전이성만 보이면 되는데, $\{x \in X : f(x) \ne h(x)\} \subseteq \{x \in X : f(x) \ne g(x)\} \cup \{x \in X : g(x) \ne h(x)\}$에서 우변의 두 집합이 모두 영집합이므로 좌변의 집합도 영집합이 되어 전이성이 성립함을 알고, 증명이 끝난다. □

Theorem 2.80 완비측도공간 $(X, \mathscr{A}, \mu)$, 가측공간 $(Y, \mathscr{B})$와 함수 $f, g : X \to Y$에 대해 만약 f가 가측이고 $f = g$ (ae.)이면 g도 가측이다.

Proof 임의의 $A \in \mathscr{B}$ 를 택하고 $N = f^{-1}(A) \triangle g^{-1}(A)$ 라 하면 $N \subseteq \{x \in X : f(x) \neq g(x)\}$ 이므로 이는 가측이고, 따라서 $g^{-1}(A) = f^{-1}(A) \triangle N$ 도 가측이 되어 g 가 가측임을 안다. □

Theorem 2.81 완비측도공간 $(X, \mathscr{A}, \mu)$ 에서 정의된 가측함수 $f_i : X \to \mathbb{R}^n$ (혹은 $f_i : X \to \overline{\mathbb{R}}$) 의 열 $\{f_i\}$, 함수 $f : X \to \mathbb{R}^n$ (혹은 $f : X \to \overline{\mathbb{R}}$) 에 대해 만약 $f_i \to f$ (ae.) 이면 f 도 가측이다.

Proof 집합 $N = \{x \in X : f_i(x) \not\to f(x)\}$ 를 생각하여 함수 $g_i : X \to \mathbb{R}^n$ (혹은 $g_i : X \to \overline{\mathbb{R}}$) 의 열 $\{g_i\}$ 를 $g_i = f_i \mathbf{1}_{X \setminus N}$ 로 두고 비슷하게 함수 $g : X \to \mathbb{R}^n$ (혹은 $g : X \to \overline{\mathbb{R}}$) 를 $g = f \mathbf{1}_{X \setminus N}$ 로 두면 모든 $i \in \mathbb{N}$ 에 대해 $f_i = g_i$ (ae.) 이고 $f = g$ (ae.) 이며 $g_i \to g$ 이다. 이제 정리 2.80로부터 $\{g_i\}$ 는 가측함수열이며, 정리 2.77의 ii로부터 g 는 가측이고, 다시 정리 2.80로부터 f 도 가측이다. □

정리 2.80, 2.81에서 측도공간의 완비성은 필수적인 조건이다. 완비성을 가지는 Lebesgue 측도공간과는 달리 우리가 주로 다루게 될 측도공간인 확률공간은 완비성을 가질 필요가 없으므로 이와 같이 완비성을 조건으로 하는 정리는 사용하기 전에 완비성 조건을 반드시 확인해줘야 한다. (확률공간은 다음 장에서 소개될 것이다.) 다음으로 소개할 내용은 가측함수의 근사에 관한 내용이다.

Definition 2.82 가측공간 $(X, \mathscr{A})$ 에서 정의된 가측함수 $f : X \to \mathbb{R}^n$ 에 대해 이의 치역 $f(X)$ 가 유한하면 이때의 f 를 **단순함수 (simple function)** 라 한다.

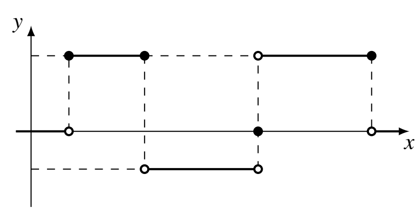


Figure 2.6 단순함수의 예시.

임의의 단순함수 $f : X \to \mathbb{R}^n$ 에 대해 이가 갖는 유한개의 서로다른 함숫값을 $a_1, \cdots, a_k \in \mathbb{R}^n$ 라 하고 각 $i \leq k$ 에 대해 $A_i = f^{-1}(a_i)$ 라 하면 $A_1, \cdots, A_k$ 는 가측이고 서로소이며 $\bigsqcup_{i=1}^k A_i = X$ 이다. 따라서 $f = \sum_{i=1}^k a_i \mathbf{1}_{A_i}$ 와 같이 지시함수를 사용하여 이를 나타낼 수 있는데, 이러한 형태를 f 의 **표준형 (standard expression)** 이라 한다.

Theorem 2.83 가측공간 $(X, \mathscr{A})$ 에서 정의된 가측함수 $f : X \to \overline{\mathbb{R}}$ 에 대해 적당한 단순함수 $f_i : X \to \mathbb{R}$ 의 열 $\{f_i\}$ 가 존재하여 다음을 만족한다.

i. $f(x) \geq 0$ 인 $x \in X$ 에 대해서는 $0 \leq f_i(x) \uparrow f(x)$ 이고 $f(x) < 0$ 인 $x \in X$ 에 대해서는 $0 \geq f_i(x) \downarrow f(x)$ 이다.

ii. 임의의 $x \in X$와 모든 $i \in \mathbb{N}$에 대해 $|f(x) - f_i(x)| < 2^{-i}$이거나 $|f_i(x)| = i$이다.

한편, 가측함수 $f : X \to \mathbb{R}^n$에 대해서도 적당한 단순함수 $f_i : X \to \mathbb{R}^n$의 열 $\{f_i\}$가 존재하여 $f_i \to f$이다.

Proof 함수열 $\{f_i\}$를

$$f_i : x \mapsto \begin{cases} -i & f(x) \le -i \text{인 경우} \\ -(j-1)2^{-i} & \text{적당한 } j \le i2^i \text{에 대해 } -j2^{-i} < f(x) \le -(j-1)2^{-i} \text{인 경우} \\ (j-1)2^{-i} & \text{적당한 } j \le i2^i \text{에 대해 } (j-1)2^{-i} \le f(x) < j2^{-i} \text{인 경우} \\ i & f(x) \ge i \text{인 경우} \end{cases}$$

로 정의하면 이는 명백히 단순함수이고 정리의 조건 i, ii를 모두 만족함을 쉽게 보일 수 있다.

한편, 가측함수 $f : X \to \mathbb{R}^n$의 경우에는 각 성분이 단순함수인 함수도 단순함수라는 사실로부터 앞선 결과를 성분별로 적용하기만 하면 된다. □

비록 가측함수는 꽤나 다양한 함수를 포괄하는 개념이지만 모든 가측함수는 단순함수의 열로써 근사가 가능하다는 공통된 성질을 가지고, 이는 이후 Lebesgue 적분론의 전개에 있어 매우 중요한 역할을 하게 될 것이다. 나중에 사용할 가측함수에 관한 몇 가지 개념과 성질을 소개하는 것으로 이번 절을 마치도록 하겠다.

Definition 2.84 공집합이 아닌 집합 X, 가측공간 $(Y, \mathscr{B})$에 대해 $\mathscr{F}$를 함수 $f : X \to Y$의 모임이라 할 때, 모든 $f \in \mathscr{F}$가 $\mathscr{A}/\mathscr{B}$−가측이도록 하는 가장 작은 σ−대수 $\mathscr{A}$를 **$\mathscr{F}$가 생성하는 σ−대수 (σ−algebra generated by $\mathscr{F}$)**라 하고 $\sigma(\mathscr{F})$로 쓴다.

Proposition 2.85 공집합이 아닌 집합 X, 가측공간 $(Y, \mathscr{B})$에 대해 $\mathscr{F}$를 함수 $f : X \to Y$의 모임이라 하면 이가 생성하는 σ−대수 $\sigma(\mathscr{F})$는 유일하게 존재한다. 따라서, $\sigma(\mathscr{F})$는 well-define된다.

Proof $\sigma(\mathscr{F})$가 존재하기만 한다면, 이의 유일성은 그 정의로부터 자명하다. 존재성의 경우 Σ를 모든 $f \in \mathscr{F}$가 $\mathscr{A}/\mathscr{B}$−가측이도록 하는 모든 σ−대수 $\mathscr{A}$의 모임이라 두고 명제 2. 5의 증명을 그대로 따라가면 된다. □

특히, $\mathscr{F}$가 한원소 집합인 경우에는 $\sigma(\mathscr{F})$의 원소들과 $\sigma(\mathscr{F})/\mathscr{B}_n$−가측인 함수들을 명시적으로 표현할 수 있다.

Theorem 2.86 공집합이 아닌 집합 X, 가측공간 $(Y, \mathscr{B})$와 함수 $f : X \to Y$에 대해 $\sigma(f) = \{f^{-1}(A) \subseteq X : A \in \mathscr{B}\}$이다.

Proof 집합족 $\mathscr{C} = \{f^{-1}(A) \subseteq X : A \in \mathscr{B}\}$ 에 대해 이가 σ−대수라면 f 가 $\mathscr{C}/\mathscr{B}$−가측임이 분명하므로 $\sigma(f) \subseteq \mathscr{C}$ 인 한편, 정의로부터 역의 포함관계는 자명하므로 증명이 끝난다. 따라서 $\mathscr{C}$ 가 σ−대수임을 보이는 것으로 충분하다. 우선 $\emptyset = f^{-1}(\emptyset) \in \mathscr{C}$ 에서 이는 비어있지 않다. 또한, 임의의 $B \in \mathscr{C}$ 에 대해 적당한 $A \in \mathscr{B}$ 가 존재하여 $B = f^{-1}(A)$ 이므로 $B^c = [f^{-1}(A)]^c = f^{-1}(A^c) \in \mathscr{C}$ 가 되어 $\mathscr{C}$ 는 여집합에 대해 닫혀있다. 마지막으로 $\mathscr{C}$ 에 속하는 임의의 집합열 $\{B_i\}$ 를 생각하면 각 $i \in \mathbb{N}$ 에 대해 적당한 $A_i \in \mathscr{B}$ 가 존재하여 $B_i = f^{-1}(A_i)$ 이므로 $\bigcup_{i=1}^{\infty} B_i = \bigcup_{i=1}^{\infty} f^{-1}(A_i) = f^{-1}(\bigcup_{i=1}^{\infty} A_i) \in \mathscr{C}$ 가 되어 $\mathscr{C}$ 는 가산 합집합에 대해서도 닫혀 있고, 곧 이는 σ−대수이다. □

Theorem 2.87 공집합이 아닌 집합 X 에서 정의된 $f : X \to \mathbb{R}^m$ 에 대해 함수 $g : X \to \mathbb{R}^n$ (혹은 $g : X \to \overline{\mathbb{R}}$) 가 $\sigma(f)/\mathscr{B}_n$−가측 (혹은 $\sigma(f)/\mathscr{B}_1$−가측) 일 필요충분조건은 적당한 Borel 함수 $h : \mathbb{R}^m \to \mathbb{R}^n$ (혹은 $h : \mathbb{R}^m \to \overline{\mathbb{R}}$) 가 존재하여 $g = h \circ f$ 인 것이다.

Proof 먼저 충분조건임을 보이기 위해 g 가 단순함수인 경우를 생각해보자. 단순함수 g 의 표준형을 $\sum_{i=1}^{k} a_i \mathbf{1}_{A_i}$ 라 하면 각 $i \le k$ 에 대해 $A_i \in \sigma(f)$ 이므로 정리 2. 86로부터 적당한 $B_i \in \mathscr{B}_m$ 가 존재하여 $A_i = f^{-1}(B_i)$ 이다. 이제 $h = \sum_{i=1}^{k} a_i \mathbf{1}_{B_i}$ 라 하면 이는 명백히 Borel 함수이다. 또한, 각 $i \in \mathbb{N}$ 와 임의의 $x \in A_i$ 에 대해 $f(x) \in B_i$ 이고 $j \neq i$ 에 대해서는 $f(x) \in B_j$ 일 수 없으므로 $g = h \circ f$ 이다. (만약 $f(x) \in B_j$ 이면 $A_1, \cdots, A_k$ 가 서로소라는 표준형의 전제에 모순된다.)

이제 일반적인 $\sigma(f)/\mathscr{B}_n$−가측함수 g 를 생각하면 정리 2. 83로부터 적당한 단순함수 $g_i : X \to \mathbb{R}^n$ 의 열 $\{g_i\}$ 가 존재하여 $g_i \to g$ 이고 각 $i \in \mathbb{N}$ 에 대해 g_i 는 $\sigma(f)/\mathscr{B}_n$−가측이므로 앞선 결과로부터 적당한 Borel 함수 $h_i : \mathbb{R}^m \to \mathbb{R}^n$ 가 존재하여 $g_i = h_i \circ f$ 이다. 이제 $A = \{x \in \mathbb{R}^m : \{h_i(x)\}_i\text{가 수렴}\}$ 라 두고 함수 h 를

$$h : x \mapsto \begin{cases} \lim_{i\to\infty} h_i(x) & x \in A \text{ 인 경우} \\ 0 & \text{ow.} \end{cases}$$

로 정의하면 이는 Borel이고, 임의의 $x \in X$ 에 대해 $\lim_{i\to\infty} h_i(f(x)) = \lim_{i\to\infty} g_i(x) = g(x)$ 에서 $f(x) \in A$ 이므로 $g = h \circ f$ 임을 안다.

반대로, 필요조건임을 보이는 것은 쉽다. 임의의 $x \in \mathbb{R}^n$ 에 대해 $h^{-1}(\prod_{i=1}^{n}(-\infty, x_i])$ 가 Borel이므로 정리 2. 86로부터 $g^{-1}(\prod_{i=1}^{n}(-\infty, x_i]) = f^{-1}(h^{-1}(\prod_{i=1}^{n}(-\infty, x_i])) \in \sigma(f)$ 가 되어 g 는 $\sigma(f)/\mathscr{B}_n$−가측이다.

한편, 함수 $g : X \to \overline{\mathbb{R}}$ 의 경우에도 이와 비슷하게 보일 수 있다. □

다음은 측도공간의 완비화와 관련이 있다. 결론부터 말하자면, 측도공간의 완비화는 함수의 가측성을 보존하며, 완비화된 공간에서 가측인 함수는 원래 공간에서의 가측함수에 의해 근사될 수 있다는 것이 다음 정리의 함의이다.

Theorem 2.88 측도공간 $(X, \mathscr{A}, \mu)$에서 정의된 함수 $f : X \to \mathbb{R}^n$(혹은 $f : X \to \overline{\mathbb{R}}$)에 대해 다음이 성립한다.

i. 만약 f가 $\mathscr{A}/\mathscr{B}_n$−가측(혹은 $\mathscr{A}/\mathscr{B}_1$−가측)이면 이는 $\overline{\mathscr{A}}/\mathscr{B}_n$−가측(혹은 $\overline{\mathscr{A}}/\mathscr{B}_1$−가측)이다.

ii. 만약 f가 $\overline{\mathscr{A}}/\mathscr{B}_n$−가측(혹은 $\overline{\mathscr{A}}/\mathscr{B}_1$−가측)이면 적당한 $\mathscr{A}/\mathscr{B}_n$−가측함수 $g : X \to \mathbb{R}^n$(혹은 $\mathscr{A}/\mathscr{B}_1$−가측함수 $g : X \to \overline{\mathbb{R}}$)가 존재하여 $f = g$ (ae.)이다.

PROOF i. 이는 $\mathscr{A} \subseteq \overline{\mathscr{A}}$에서 자명하다.

ii. 먼저 f가 단순함수인 경우를 생각해보자. 이제 f의 표준형을 $f = \sum_{i=1}^k a_i \mathbf{1}_{A_i}$라 하면 가정으로부터 각 $i \le k$에 대해 적당한 $B_i \in \mathscr{A}$와 영집합 $N_i \subseteq X$가 존재하여 $A_i = B_i \cup N_i$이다. 그렇다면 함수 g를 $g = \sum_{i=1}^k a_i \mathbf{1}_{B_i}$라 하면 이는 명백히 $\mathscr{A}/\mathscr{B}_n$−가측이고, $\{x \in X : f(x) \neq g(x)\} = \bigcup_{i=1}^k (A_i \setminus B_i) \subseteq \bigcup_{i=1}^k N_i$에서 $\bigcup_{i=1}^k N_i$도 영집합이므로 $f = g$ (ae.)이다.

한편, 일반적인 $\overline{\mathscr{A}}/\mathscr{B}_n$−가측함수 f를 생각하면 정리 2. 83로부터 적당한 단순함수 $f_i : X \to \mathbb{R}^n$의 열 $\{f_i\}$가 존재하여 $f_i \to f$이고, 앞선 결과로부터 각 $i \in \mathbb{N}$에 대해 적당한 $\mathscr{A}/\mathscr{B}_n$−가측함수 $g_i : X \to \mathbb{R}^n$가 존재하여 $f_i = g_i$ (ae.)이다. 이제 각 $i \in \mathbb{N}$에 대해 $N_i = \{x \in X : f_i(x) \neq g_i(x)\}$라 하면 이는 영집합이므로 $N = \bigcup_{i=1}^\infty N_i$도 영집합이어서 적당한 $C \in \mathscr{A}$에 대해 $N \subseteq C$이고 $\mu(C) = 0$이다. 이제 임의의 $x \in X \setminus C$에 대해 $g_i(x) \to f(x)$이므로 함수 $g : X \to \mathbb{R}^n$를

$$g : x \mapsto \begin{cases} \lim\limits_{i\to\infty} g_i(x) & x \in X \setminus C\text{인 경우} \\ 0 & \text{ow.} \end{cases}$$

로 두면 이는 $\mathscr{A}/\mathscr{B}_n$−가측이고, $f = g$ (ae.)임을 안다.

한편, 함수 $f : X \to \overline{\mathbb{R}}$의 경우에도 이와 비슷하게 보일 수 있다. □

위의 정리의 ii에서 임의의 μ−영집합은 $\overline{\mu}$−영집합이고 그 역도 성립하므로 일부러 어떤 측도를 기준으로 거의 어디서나인지를 나타내지 않았다.

2.7 Integration

이번 절의 전반부에서 우리는 Lebesgue 적분을 총 세 단계에 걸쳐 정의한다. 먼저 첫 번째 단계에서는 음이 아닌 단순함수들에 대해 Lebesgue 적분을 정의한다. 이어지는 두 번째 단계에서는 정리 2. 83를 사용하여 Lebesgue 적분의 정의를 음이 아닌 가측함수로 확장한다. 마지막으로 세 번째 단계에서는 함수를 양의 부분과 음의 부분으로 나눔으로써 Lebesgue 적분의 정의를 일반적인 가측함수로 확장하여 Lebesgue 적분의 가장 일반적인 정의를

얻는다. 이렇게 Lebesgue 적분의 정의를 끝마친 이후에는 앞서 언급한 바와 같이 Lebesgue 적분론에서 적분과 극한의 순서를 바꾸는 데 사용되는 유용한 정리들을 살펴보도록 하겠다.

Definition 2.89 측도공간 $(X, \mathscr{A}, \mu)$에서 정의된 음이 아닌 단순함수 $f : X \to \mathbb{R}_0^+$에 대해 이의 표준형을 $f = \sum_{i=1}^k a_i \mathbf{1}_{A_i}$라 하면 이때 μ에 대한 f의 **(Lebesgue) 적분 (− integral)**을 $\int_X f(x)\,d\mu(x)$ 혹은 간단히 $\int_X f\,d\mu$로 쓰고 $\int_X f\,d\mu = \sum_{i=1}^k a_i \mu(A_i)$로 정의한다.

Proposition 2.90 측도공간 $(X, \mathscr{A}, \mu)$에 대해 $A_1, \cdots, A_k \in \mathscr{A}$를 $X = \bigsqcup_{i=1}^k A_i$인 서로소인 집합이라 하자. 이제 $a_1, \cdots, a_k \geq 0$에 대해 함수 $f : X \to \mathbb{R}_0^+$를 $f = \sum_{i=1}^k a_i \mathbf{1}_{A_i}$라 하면 $\int_X f\,d\mu = \sum_{i=1}^k a_i \mu(A_i)$이다.

Proof 양수 $a_1, \cdots, a_k$ 중에서 서로 같은 값들을 하나만 남기고 제거하여 이를 $b_1, \cdots, b_l \in \mathbb{R}$라 하고 각 $j \leq l$에 대해 $B_j = f^{-1}(b_j) = \bigsqcup_{i:a_i=b_j} A_i$라 하면 이는 가측이고 $f = \sum_{j=1}^l b_j \mathbf{1}_{B_j}$이다. 그렇다면 이는 곧 단순함수 f의 표준형이므로 Lebesgue 적분의 정의로부터 $\int_X f\,d\mu = \sum_{j=1}^l b_j \mu(B_j) = \sum_{j=1}^l b_j \mu(\bigsqcup_{i:a_i=b_j} A_i) = \sum_{j=1}^l b_j \sum_{i:a_i=b_j} \mu(A_i) = \sum_{i=1}^k a_i \mu(A_i)$이다. □

위의 명제에서는 $a_1, \cdots, a_k$가 서로 다르다는 조건이 없으므로 이는 음의 아닌 단순함수에 대한 적분의 정의보다 조금 더 일반적이다. 다음으로 Lebesgue 적분의 기초적인 성질들을 보자.

Proposition 2.91 측도공간 $(X, \mathscr{A}, \mu)$에서 정의된 음이 아닌 단순함수 $f, g : X \to \mathbb{R}_0^+$에 대해 다음이 성립한다.

i. 임의의 $c \geq 0$에 대해 $\int_X cf\,d\mu = c\int_X f\,d\mu$이다.
ii. $\int_X f + g\,d\mu = \int_X f\,d\mu + \int_X g\,d\mu$.
iii. 만약 $f \leq g$이면 $\int_X f\,d\mu \leq \int_X g\,d\mu$이다.

Proof 단순함수 f, g의 표준형을 각각 $f = \sum_{i=1}^k a_i \mathbf{1}_{A_i}$, $g = \sum_{j=1}^l b_j \mathbf{1}_{B_j}$라 하자.

i. 함수 $cf = \sum_{i=1}^k ca_i \mathbf{1}_{A_i}$가 음이 아닌 단순함수임이 분명하므로 명제 2. 90로부터 $\int_X cf\,d\mu = \sum_{i=1}^k ca_i \mu(A_i) = c\sum_{i=1}^k a_i \mu(A_i) = c\int_X f\,d\mu$이다.

ii. 함수 $f+g = \sum_{i=1}^k \sum_{j=1}^l (a_i + b_j) \mathbf{1}_{A_i \cap B_j}$가 음이 아닌 단순함수임이 분명하고 $\bigsqcup_{i=1}^k \bigsqcup_{j=1}^l (A_i \cap B_j) = X$이므로 명제 2. 90로부터 $\int_X f + g\,d\mu = \sum_{i=1}^k \sum_{j=1}^l (a_i + b_j)\mu(A_i \cap B_j)$이다. 한편, $f = \sum_{i=1}^k \sum_{j=1}^l a_i \mathbf{1}_{A_i \cap B_j}$, $g = \sum_{i=1}^k \sum_{j=1}^l b_j \mathbf{1}_{A_i \cap B_j}$로 쓸 수 있으므로 다시 명제 2. 90로부터 $\int_X f\,d\mu = \sum_{i=1}^k \sum_{j=1}^l a_i \mu(A_i \cap B_j)$이고 $\int_X g\,d\mu = \sum_{i=1}^k \sum_{j=1}^l b_j \mu(A_i \cap B_j)$가 되어 명제가 성립한다.

iii. 함수 $g - f$가 음이 아닌 단순함수이므로 ii로부터 $\int_X g\,d\mu = \int_X g - f\,d\mu + \int_X f\,d\mu \geq \int_X f\,d\mu$이다. □

Corollary 2.92 측도공간 $(X, \mathscr{A}, \mu)$와 $A_1, \cdots, A_k \in \mathscr{A}$, $a_1, \cdots, a_k \geq 0$에 대해 함수 $f : X \to \mathbb{R}_0^+$를 $f = \sum_{i=1}^k a_i \mathbf{1}_{A_i}$라 하면 $\int_X f\,d\mu = \sum_{i=1}^k a_i \mu(A_i)$이다.

Proof 이는 $\int_X f\,d\mu = \sum_{i=1}^k a_i \int_X \mathbf{1}_{A_i}\,d\mu = \sum_{i=1}^k a_i\mu(A_i)$ 에서 자명하다. □

위의 따름정리에서는 $A_1, \cdots, A_k$ 가 서로소일 필요도, X 의 덮개일 필요도 없다. 따라서 이는 명제 2.90의 결과에서 한 발짝 더 나아간 결과이다. 이제 두 번째 단계로 넘어가자.

Definition 2.93 측도공간 $(X, \mathscr{A}, \mu)$ 에서 정의된 음이 아닌 가측함수 $f : X \to \overline{\mathbb{R}}_0^+$ 에 대해 μ 에 대한 f 의 **(Lebesgue) 적분 (− integral)**을 $\int_X f(x)\,d\mu(x)$ 혹은 간단히 $\int_X f\,d\mu$ 로 쓰고

$$\int_X f\,d\mu = \sup\left\{ \int_X g\,d\mu \in \overline{\mathbb{R}} : \text{함수 } g : X \to \mathbb{R}_0^+ \text{ 는 } g \le f \text{ 인 음이 아닌 단순함수} \right\}$$

로 정의한다.

정리 2.83로부터 임의의 음이 아닌 가측함수는 이로 점별수렴하는 음이 아닌 함수열을 가지므로 위의 정의는 well-define되어 있다. 하지만, 이런 정의를 직접 사용하는 것은 그다지 편리하지 않으므로 이를 대신할 정리를 소개한다.

Lemma 2.94 측도공간 $(X, \mathscr{A}, \mu)$ 와 $A \in \mathscr{A}$ 에 대해 음이 아닌 단순함수 $f_i : X \to \mathbb{R}_0^+$ 의 열 $\{f_i\}$ 가 증가하고 어떤 $M \ge 0$ 과 모든 $x \in A$ 에 대해 $\lim_{i\to\infty} f_i(x) \ge M$ 을 만족한다고 하면 $\lim_{i\to\infty} \int_X f_i \mathbf{1}_A\,d\mu \ge M\mu(A)$ 이다.

Proof 만약 $M = 0$ 이거나 $\mu(A) = 0$ 이면 보조정리가 자명하므로 $M, \mu(A) > 0$ 라 하자. 또한, $\{\int_X f_i \mathbf{1}_A\,d\mu\}_i$ 가 증가하므로 만약 이가 위로 유계가 아니라면 이번에도 보조정리가 자명하여 $\{\int_X f_i \mathbf{1}_A\,d\mu\}$ 가 위로 유계라 하자. 그렇다면 MSP로부터 이는 수렴한다. 이제 임의의 양수 $a < M$ 를 택하고 집합열 $\{A_i\}$ 를 $A_i = f_i^{-1}([a, \infty)) \cap A$ 로 두면 이는 가측이고 $A_i \uparrow A$ 이다. 따라서 정리 2.24의 i로부터 $\mu(A_i) \uparrow \mu(A)$ 이고, 곧 임의의 $b < \mu(A)$ 를 택하면 적당한 $i_0 \in \mathbb{N}$ 가 존재하여 임의의 $i \ge i_0$ 에 대해 $\mu(A_i) > b$ 이다. 이상의 결과로부터 임의의 $i \ge i_0$ 에 대해 $\int_X f_i \mathbf{1}_A\,d\mu \ge \int_X f_i \mathbf{1}_{A_i}\,d\mu \ge \int_X a\mathbf{1}_{A_i}\,d\mu = a\mu(A_i) > ab$ 가 성립하여 $\lim_{i\to\infty} \int_X f_i \mathbf{1}_A\,d\mu \ge M\mu(A)$ 임을 안다. □

Lemma 2.95 측도공간 $(X, \mathscr{A}, \mu)$ 에서 정의된 음이 아닌 단순함수 $f : X \to \mathbb{R}_0^+$ 에 대해 음이 아닌 단순함수 $g_i : X \to \mathbb{R}_0^+$ 의 열 $\{g_i\}$ 가 증가하고 $\lim_{i\to\infty} g_i \ge f$ 를 만족한다면 $\lim_{i\to\infty} \int_X g_i\,d\mu \ge \int_X f\,d\mu$ 이다.

Proof 단순함수 f 의 표준형을 $f = \sum_{j=1}^k a_j \mathbf{1}_{A_j}$ 라 하면 임의의 $j \le k$ 와 임의의 $x \in A_j$ 에 대해 $\lim_{i\to\infty} g_i(x) \ge a_j$ 이므로 위의 보조정리로부터 $\lim_{i\to\infty} \int_X g_i \mathbf{1}_{A_j}\,d\mu \ge a_j\mu(A_j)$ 이다. 한편, 각 $i \in \mathbb{N}$ 에 대해 $\int_X g_i\,d\mu = \int_X g_i \sum_{j=1}^k \mathbf{1}_{A_j}\,d\mu = \sum_{j=1}^k \int_X g_i \mathbf{1}_{A_j}\,d\mu$ 이므로 $\lim_{i\to\infty} \int_X g_i\,d\mu = \sum_{j=1}^k \lim_{i\to\infty} \int_X g_i \mathbf{1}_{A_j}\,d\mu \ge \sum_{j=1}^k a_j\mu(A_j) = \int_X f\,d\mu$ 를 얻는다. □

Theorem 2.96 측도공간 $(X, \mathscr{A}, \mu)$에서 정의된 음이 아닌 가측함수 $f : X \to \overline{\mathbb{R}}_0^+$에 대해 $f_i \uparrow f$인 임의의 음이 아닌 단순함수 $f_i : X \to \mathbb{R}_0^+$의 열 $\{f_i\}$를 생각하면 $\int_X f_i \, d\mu \uparrow \int_X f \, d\mu$이다.

Proof 수열 $\{\int_X f_i \, d\mu\}_i$가 증가하므로 이는 수렴하거나 ∞로 발산하는데, 후자의 경우 $\int_X f \, d\mu = \infty$가 되어 정리가 자명하므로 $\{\int_X f_i \, d\mu\}$가 수렴한다고 가정하자. 그렇다면 $\{\int_X f_i \, d\mu\}$는 $\int_X f \, d\mu$에 의해 위로 유계이므로 $\lim_{i\to\infty} \int_X f_i \, d\mu \le \int_X f \, d\mu$이다. 한편, $g \le f$인 임의의 음이 아닌 단순함수 $g : X \to \mathbb{R}_0^+$에 대해 $f_i \uparrow f \ge g$이므로 위의 보조정리로부터 $\lim_{i\to\infty} \int_X f_i \, d\mu \ge \int_X g \, d\mu$이고, 곧 $\lim_{i\to\infty} \int_X f_i \, d\mu \ge \int_X f \, d\mu$가 되어 이를 앞선 결과와 합하면 $\int_X f_i \, d\mu \uparrow \int_X f \, d\mu$를 얻는다. □

이제 마지막 단계이다.

Definition 2.97 측도공간 $(X, \mathscr{A}, \mu)$에서 정의된 음이 아닌 가측함수 $f : X \to \overline{\mathbb{R}}_0^+$에 대해 $\int_X f \, d\mu < \infty$이면 이때의 f를 **(μ−Lebesgue) 적분가능 (− integrable)**하다고 한다. 나아가, 일반적인 가측함수 $f : X \to \overline{\mathbb{R}}$에 대해 $f_\pm$이 모두 적분가능하면 이때의 f를 **(μ−Lebesgue) 적분가능 (− integrable)**하다고 한다. 그리고 이때 μ에 대한 f의 **(Lebesgue) 적분 (− integral)**을 $\int_X f(x) \, d\mu(x)$ 혹은 간단히 $\int_X f \, d\mu$로 쓰고 $\int_X f \, d\mu = \int_X f_+ \, d\mu - \int_X f_- \, d\mu$로 정의한다. 여기서 $f_\pm$은 각각 f의 양의 부분과 음의 부분을 의미하는 것으로 각각 $f_\pm = (|f| \pm f)/2$로 정의된다.

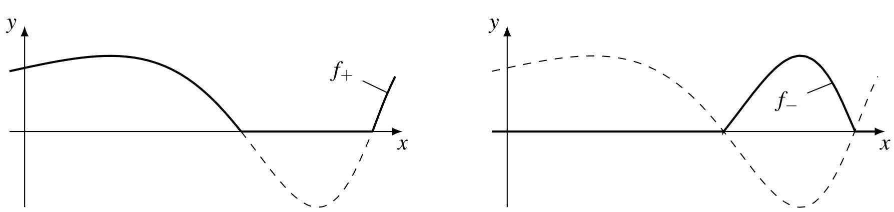


Figure 2.7 함수 f (파선)에 대한 f_+ (왼쪽)와 f_- (오른쪽)의 그래프.

Definition 2.98 측도공간 $(X, \mathscr{A}, \mu)$에서 정의된 가측함수 $f : X \to \overline{\mathbb{R}}$와 $A \in \mathscr{A}$에 대해 함수 $f\mathbf{1}_A$가 적분가능하면 이때의 f를 A 위에서 **(μ−Lebesgue) 적분가능 (integrable)**하다고 한다. 그리고 이때 μ에 대한 f의 **A에서의 (Lebesgue) 적분 (− integral over A)**을 $\int_A f(x) \, d\mu(x)$ 혹은 간단히 $\int_A f \, d\mu$로 쓰고 $\int_A f \, d\mu = \int_X f\mathbf{1}_A \, d\mu$로 정의한다.

Lebesgue 적분의 정의가 끝났으니 이제 이의 성질들을 알아보도록 하자. 표기의 편의와 간결한 논의를 위해 대부분의 명제나 정리들에서 적분 영역이 전체 공간 X로 주어져 있지만

임의의 가측집합 $A \subseteq X$ 에 대해 f 대신 $f\mathbf{1}_A$ 를 생각함으로써 이들을 임의의 가측집합을 적분영역으로 갖는 경우로 자명하게 일반화할 수 있다.

Theorem 2.99 측도공간 $(X, \mathscr{A}, \mu)$ 에서 정의된 음이 아닌 가측함수 $f, g : X \to \overline{\mathbb{R}}_0^+$ 와 $A, B \in \mathscr{A}$ 에 대해 다음이 성립한다.

i. 임의의 $c \geq 0$에 대해 $\int_X cf\,d\mu = c\int_X f\,d\mu$ 이다.
ii. $\int_X f+g\,d\mu = \int_X f\,d\mu + \int_X g\,d\mu$.
iii. 만약 $f \leq g$이면 $\int_X f\,d\mu \leq \int_X g\,d\mu$ 이다.
iv. 만약 $A \subseteq B$이면 $\int_A f\,d\mu \leq \int_B f\,d\mu$ 이다.

한편, 적분가능한 함수 $f, g : X \to \overline{\mathbb{R}}$ 에 대해서도 iv를 제외한 나머지 성질들이 성립한다. 다만, i의 경우 더 이상 $c \geq 0$의 조건이 요구되지 않으며, cf 와 $f+g$ 의 적분가능이 보장된다.

i°. 임의의 $c \in \overline{\mathbb{R}}$ 에 대해 cf 가 적분가능하고 $\int_X cf\,d\mu = c\int_X f\,d\mu$ 이다.
ii°. 함수 $f+g$가 적분가능하고 $\int_X f+g\,d\mu = \int_X f\,d\mu + \int_X g\,d\mu$ 이다. [13]

나아가 이 경우, 음이 아닌 가측함수 $f : X \to \overline{\mathbb{R}}_0^+$ 에 대해서는 자명한 다음 성질도 성립한다.

v. $|\int_X f\,d\mu| \leq \int_X |f|\,d\mu$.

Proof 정리 2. 83로부터 음이 아닌 단순함수 $f_i, g_i : X \to \mathbb{R}_0^+$ 의 열 $\{f_i\}, \{g_i\}$ 가 존재하여 $f_i \uparrow f, g_i \uparrow g$ 이고, 그렇다면 정리 2. 96로부터 $\int_X f_i\,d\mu \uparrow \int_X f\,d\mu, \int_x g_i\,d\mu \uparrow \int_X g\,d\mu$ 이다.

i. 함수열 $\{cf_i\}$ 음이 아닌 단순함수열로서 $cf_i \uparrow cf$ 임이 분명하므로 정리 2. 96로부터 $\int_X cf\,d\mu = \lim_{i\to\infty}\int_X cf_i\,d\mu = c\lim_{i\to\infty}\int_X f_i\,d\mu = c\int_X f\,d\mu$ 이다.

ii. 함수열 $\{f_i+g_i\}$ 가 음이 아닌 단순함수열로서 $f_i+g_i \uparrow f+g$ 임이 분명하므로 정리 2. 96로부터 $\int_X f+g\,d\mu = \lim_{i\to\infty}\int_X f_i+g_i\,d\mu = \lim_{i\to\infty}\int_X f_i\,d\mu + \lim_{i\to\infty}\int_X g_i\,d\mu = \int_X f\,d\mu + \int_X g\,d\mu$ 이다.

iii. 상수 함수 $h = 0$은 $g - f \geq h$인 음이 아닌 단순함수로 볼 수 있으므로 정의로부터 $\int_X g-f\,d\mu \geq 0$ 이고, 곧 ii로부터 $\int_X g\,d\mu = \int_X g-f\,d\mu + \int_X f\,d\mu \geq \int_X f\,d\mu$ 이다.

iv. 이는 $f\mathbf{1}_A \leq f\mathbf{1}_B$ 와 iii으로부터 자명하다.

한편, 적분가능한 함수 $f, g : X \to \overline{\mathbb{R}}$ 의 경우에는 이를 각각 양과 음의 부분으로 나누어 앞선 결과를 적용하면 된다. 특히 i° 의 경우 $c \geq 0$인 경우와 $c < 0$인 경우를 나누어 생각하면 된다.

v. 이는 $-|f| \leq f \leq |f|$ 와 iii으로부터 $-\int_X |f|\,d\mu \leq \int_X f\,d\mu \leq \int_X |f|\,d\mu$ 이므로 자명하다. □

Theorem 2.100 측도공간 $(X, \mathscr{A}, \mu)$ 에서 정의된 가측함수 $f : X \to \overline{\mathbb{R}}$ 와 $A, B \in \mathscr{A}$ 에 대해 다음이 성립한다.

i. 함수 f가 적분가능할 필요충분조건은 $|f|$가 적분가능한 것이다.
ii. 만약 f가 적분가능하면 이는 거의 어디서나 유한하다.
iii. 만약 $A \subseteq B$이고 f가 B 위에서 적분가능하면 이는 A 위에서도 적분가능하다.

Proof i. 정의로부터 $|f|$가 적분가능하다는 것은 $\int_X |f|\, d\mu = \int_X f_+\, d\mu + \int_X f_-\, d\mu < \infty$인 것과 동치이고, 이는 다시 f_+, f_-가 적분가능하다는 것과 동치이므로 곧 정의로부터 f가 적분가능하다는 것과 동치이다.

ii. 우선 $\{\pm\infty\}$가 닫혀있으므로 Borel이고, 따라서 $N = f^{-1}(\pm\infty)$이 가측이다. 그런데 만약 $\mu(N) > 0$이라면 $\int_X |f|\, d\mu \geq \int_N |f|\, d\mu = \infty$에서 f가 적분가능하다는 가정에 모순되므로 N은 영집합이고, 곧 명제가 성립한다.

iii. i로부터 $\int_A |f|\, d\mu \leq \int_B |f|\, d\mu < \infty$가 되어 이는 자명하다. □

Theorem 2.101 측도공간 $(X, \mathscr{A}, \mu)$에서 정의된 가측함수 $f : X \to \overline{\mathbb{R}}$에 대해 다음이 성립한다.

i. 만약 $f = 0$ (ae.)이면 이는 적분가능하고 $\int_X f\, d\mu = 0$이다.
ii. 만약 f가 적분가능하고 거의 어디서나 음이 아니며 $\int_X f\, d\mu = 0$이면 $f = 0$ (ae.)이다.

Proof i. 먼저 f가 음이 아닌 경우를 생각하고 $f \geq g$인 임의의 음이 아닌 단순함수 $g : X \to \mathbb{R}_0^+$를 택하여 이의 표준형을 $g = \sum_{i=1}^k a_i \mathbf{1}_{A_i}$라 하자. 그렇다면 $g = 0$ (ae.)이므로 WLOG, $a_1 = 0$이라 한다면 각 $1 < i \leq k$에 대해 $\mu(A_i) = 0$이 되어 $\int_X g\, d\mu = \sum_{i=1}^k a_i \mu(A_i) = 0$에서 $\int_X f\, d\mu = 0$이다. 이제 일반적인 가측함수 f에 대해서 $f_+ = f_- = 0$ (ae.)이므로 앞선 결과로부터 $\int_x f\, d\mu = \int_X f_+\, d\mu - \int_X f_-\, d\mu = 0$이다.

ii. 먼저 f가 음이 아닌 경우를 생각하고 모순을 유도하기 위해 $f = 0$ (ae.)가 아니라 가정하자. 이제 집합열 $\{A_i\}$를 $A_i := f^{-1}((1/i, \infty])$로 두면 이는 가측이고 곧 $\bigcup_{i=1}^\infty A_i = f^{-1}(\bigcup_{i=1}^\infty (1/i, \infty]) = f^{-1}((0, \infty])$도 가측이 되어 가정으로부터 양의 측도를 가진다. 나아가 $A_i \uparrow f^{-1}((0, \infty])$이므로 정리 2. 24의 i로부터 $\mu(A_i) \uparrow \mu(f^{-1}((0, \infty])) > 0$이 되어 $\mu(A_{i_0}) > 0$인 적당한 $i_0 \in \mathbb{N}$가 존재한다. 그러나, $g = \mathbf{1}_{A_{i_0}}/i_0$라 하면 이가 $f \geq g$인 음이 아닌 단순함수이므로 $\int_X f\, d\mu \geq \int_X g\, d\mu = \mu(A_{i_0})/i_0 > 0$의 모순이 발생한다. 따라서 $f = 0$ (ae.)이다. 이제 일반적인 가측함수 f에 대해 가정으로부터 $f_- = 0$ (ae.)이므로 i에서 $\int_X f_+\, d\mu = \int_X f_-\, d\mu = 0$이고, 곧 앞선 결론으로부터 $f_+ = 0$ (ae.)가 성립하여 $f = 0$ (ae.)임을 안다. □

Corollary 2.102 측도공간 $(X, \mathscr{A}, \mu)$에서 정의된 음이 아닌 가측함수 $f, g : X \to \overline{\mathbb{R}}_0^+$에 대해 다음이 성립한다.

i. 만약 $N \in \mathscr{A}$이 영집합이면 f는 N 위에서 적분가능하며 $\int_N f\, d\mu = 0$이다.
ii. 만약 $f = g$ (ae.)이면 $\int_X f\, d\mu = \int_X g\, d\mu$이다.

iii. 만약 $\{x \in X : f(x) > 0\}$ 이 양의 측도를 가지면 $\int_X f\,d\mu > 0$ 이다.

한편, 적분가능한 함수 $f : X \to \overline{\mathbb{R}}$ 와 일반적인 가측함수 $g : X \to \overline{\mathbb{R}}$ 에 대해서도 iii을 제외한 나머지 성질들이 성립한다. 특히 ii의 경우 g 의 적분가능성까지 보장된다.

ii° 만약 $f = g$ (ae.) 이면 g 도 적분가능하고 $\int_X f\,d\mu = \int_X g\,d\mu$ 이다.

Proof i. 이는 $f\mathbf{1}_N = 0$ (ae.) 에서 정리 2. 101의 i로부터 자명하다.

ii. 가정으로부터 적당한 영집합 $N \in \mathscr{A}$ 가 존재하여 $X \setminus N$ 에서 $f = g$ 이고, i로부터 $\int_X f\,d\mu = \int_{X\setminus N} f\,d\mu + \int_N f\,d\mu = \int_{X\setminus N} g\,d\mu = \int_{X\setminus N} g\,d\mu + \int_N g\,d\mu = \int_X g\,d\mu$ 이다.

iii. 만약 $\int_X f\,d\mu = 0$ 이면 정리 2. 101의 ii로부터 $f = 0$ (ae.) 의 모순이 발생하므로 명제가 자명하다.

한편, 적분가능한 함수 $f : X \to \overline{\mathbb{R}}$ 와 일반적인 가측함수 $g : X \to \overline{\mathbb{R}}$ 의 경우에는 이를 각각 양과 음의 부분으로 나누어 앞선 결과를 적용하면 된다. □

이제 이번 절의 후반부로서 Lebesgue 적분과 극한이 서로 어떻게 상호작용하는지에 대해 알아보도록 하자. 이러한 상호작용을 규정하는 정리들을 흔히 '수렴정리' 라 부른다. 아쉽게도, 어느 경우에나 항상 사용할 수 있는 일반적인 수렴정리는 존재하지 않는다. 곧 각각의 수렴정리는 나름의 조건을 요구하는데, 수렴정리를 적용하기 전에 이러한 조건을 반드시 확인해야 한다.[14] 첫 번째로 살펴볼 수렴정리는 함수열이 음이 아니며 증가할 것을 요구한다.

Theorem 2.103 (Monotone convergence theorem) 측도공간 $(X, \mathscr{A}, \mu)$ 에서 정의된 음이 아닌 가측함수 $f_i : X \to \overline{\mathbb{R}}_0^+$ 의 열 $\{f_i\}$ 가 증가하며 적당한 음이 아닌 가측함수 $f : X \to \overline{\mathbb{R}}_0^+$ 에 대해 $f_i \uparrow f$ (ae.) 라 하면 $\int_X f_i\,d\mu \uparrow \int_X f\,d\mu$ 이다.

Proof 먼저 $f_i \uparrow f$ 인 경우를 생각하면 수열 $\{\int_X f_i\,d\mu\}_i$ 가 증가하고 $\int_X f\,d\mu$ 에 의해 위로 유계이므로 $\lim_{i\to\infty} \int_X f_i\,d\mu \le \int_X f\,d\mu$ 이다. 따라서 역의 대소관계만 보인다면 이 경우에 대해서는 증명이 끝난다. 이를 위해 각 $i \in \mathbb{N}$ 에 대해 음이 아닌 단순함수 $g_{ij} : X \to \mathbb{R}_0^+$ 의 열 $\{g_{ij}\}_j$ 를 정리 2. 83의 증명에서의 함수열과 같이 잡고 함수열 $\{h_i\}$ 를 $h_i := g_{ii}$ 로 두자. 이제 $h_i \uparrow f$ 임을 보이기 위해 $\{h_i\}$ 가 증가한다는 것과 f 로 수렴한다는 것을 주장한다.

함수열 $\{h_i\}$ 가 증가한다는 점은 경우를 나누어 하나하나 확인해봄으로써 어렵지 않게 알 수 있다. 우선 임의의 $x \in X$ 와 임의의 $i \in \mathbb{N}$ 를 택하면 우리가 생각해야 할 경우는 다음의 네 가지이다.

i. $f_i(x) \ge i$ 이고 $f_{i+1}(x) \ge i+1$ 인 경우.
ii. $i \le f_i(x) \le f_{i+1}(x) < i+1$ 인 경우.
iii. $f_i(x) < i$ 이고 $f_{i+1}(x) \ge i+1$ 인 경우.
iv. $f_i(x) \le f_{i+1}(x) < i$ 인 경우.

연습삼아 처음 두 개의 경우에 대해서만 확인해보면, i의 경우에는 $h_{i+1}(x) = i+1 \geq i = h_i(x)$에서 자명하다. 한편, ii의 경우에는 적당한 자연수 $j \leq (i+1)2^{i+1}$가 존재하여 $(j-1)2^{-i-1} \leq f(x) < j2^{-i-1}$인데, 이는 $i2^{i+1} < j \leq (i+1)2^{i+1}$임을 뜻하므로 곧 $h_{i+1}(x) = (j-1)2^{-i-1} \geq i = h_i(x)$에서 이 또한 자명하다. 남은 두 경우에 대해서도 이와 같이 쉽게 따져볼 수 있으므로 $\{h_i\}$는 증가함을 안다.

다음으로 $h_i \uparrow f$임을 확인해보자. 일단, 임의의 $x \in X$를 택하면 임의의 $i \in \mathbb{N}$에 대해 $|f_i(x) - h_i(x)| < 2^{-i}$이거나 $h_i(x) = i$이다. 여기서 만약 무한히 많은 $i \in \mathbb{N}$에 대해 $h_i(x) = i$라면 $f_i(x) \to \infty$에서 $f(x) = \infty$인 한편, $h_i(x) \to \infty$이므로 $h_i \uparrow f$이다. 만약 그렇지 않다면 충분히 큰 $i \in \mathbb{N}$에 대해 $|f_i(x) - h_i(x)| < 2^{-i}$이고, 곧 $h_i(x) \uparrow f(x)$임이 분명하다.

이제 $h_i \uparrow f$임을 알았으므로 정리 2. 96로부터 $\int_X h_i \, d\mu \uparrow \int_X f \, d\mu$이고 $\{h_i\}$의 구성으로부터 모든 $i \in \mathbb{N}$에 대해 $h_i \leq f_i$이므로 $\int_X f \, d\mu \leq \lim_{i\to\infty} f_i \, d\mu$에서 증명이 끝난다.

한편, $f_i \uparrow f$ (ae.)인 경우에 대해서는 $N = \{x \in X : f_i(x) \not\to f(x)\}$을 생각하면 이가 영집합이고 정리 2. 77의 iv로부터 가측이다. 이제 함수 $g_i : X \to \overline{\mathbb{R}}_0^+$의 열 $\{g_i\}$와 함수 $g : X \to \overline{\mathbb{R}}_0^+$를 각각 $g_i = f_i \mathbf{1}_{X \setminus N}$, $g = f \mathbf{1}_{X \setminus N}$과 같이 정의하면 $\{g_i\}$는 음이 아닌 가측함수의 열로, 증가하고 $g_i \uparrow g$이다. 그렇다면 앞서 얻은 결과로부터 $\int_X g_i \, d\mu \uparrow \int_X g \, d\mu$임을 바로 알 수 있고, 곧 모든 $i \in \mathbb{N}$에 대해 $f_i = g_i$ (ae.)이며 $f = g$ (ae.)이므로 따름정리 2. 102의 ii로부터 $\int_X f_i \, d\mu \uparrow \int_X f \, d\mu$를 얻는다. □

두 번째 수렴정리는 함수열이 음이 아닐 것만을 요구하는 대신, 등호가 아닌 부등호의 관계만 보장한다.

Theorem 2.104 (Fatou's lemma) 측도공간 $(X, \mathscr{A}, \mu)$에서 정의된 음이 아닌 가측함수 $f_i : X \to \overline{\mathbb{R}}_0^+$의 열 $\{f_i\}$에 대해 $\liminf_{i\to\infty} \int_X f_i \, d\mu \geq \int_X \liminf_{i\to\infty} f_i \, d\mu$이다.

Proof 함수열 $\{g_i\}$를 $g_i := \inf_{j \geq i} f_j$로 두면 이는 증가하는 함수열로 정리 2. 77의 i로부터 각 g_i가 가측이므로 MCT에서 $\lim_{i\to\infty} \int_X g_i \, d\mu = \int_X \lim_{i\to\infty} g_i \, d\mu = \int_X \sup_{i\in\mathbb{N}} \inf_{j\geq i} f_j \, d\mu = \int_X \liminf_{i\to\infty} f_i \, d\mu$이다. 그렇다면 모든 $i \in \mathbb{N}$에 대해 $g_i \leq f_i$이므로 $\liminf_{i\to\infty} \int_X f_i \, d\mu \geq \liminf_{i\to\infty} \int_X g_i \, d\mu = \lim_{i\to\infty} \int_X g_i \, d\mu = \int_X \liminf_{i\to\infty} f_i \, d\mu$이다. □

마지막 수렴정리는 그 이름에서 알 수 있듯이 '지배'라는 독특한 조건을 요구한다. 즉, 함수열이 어떤 적분가능한 함수에 의해 지배되어야 한다.

Theorem 2.105 (Lebesgue's dominated convergence theorem) 측도공간 $(X, \mathscr{A}, \mu)$에서 정의된 가측함수 $f_i : X \to \overline{\mathbb{R}}$의 열 $\{f_i\}$와 가측함수 $f : X \to \overline{\mathbb{R}}$에 대해 $f_i \to f$ (ae.)라 하자. 또한, 어떤 적분가능한 함수 $g : X \to \overline{\mathbb{R}}_0^+$가 존재하여 모든 $i \in \mathbb{N}$에 대해 $|f_i| \leq g$이면 각 f_i와 f는 적분가능하고 $\int_X f_i \, d\mu \to \int_X f \, d\mu$이다.

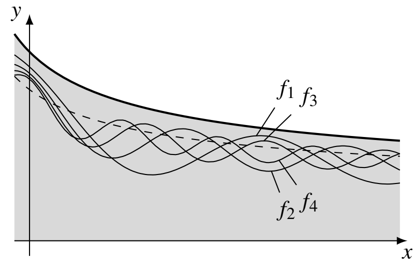


Figure 2.8 함수열 $\{f_i\}$가 g (굵은선)에 의해 지배당하고 있다. 그림에서 $\{f_i\}$는 결국 f (파선)로 점별수렴한다.

Proof 가정으로부터 각 f_i에 대해 $\int_X |f_i|\,d\mu \le \int_X g\,d\mu < \infty$이므로 이는 적분가능하고, 비슷한 이유로 f도 적분가능하다. 이제 $f_i \to f$인 경우를 생각해보자. 이 경우에 함수열 $\{g - f_i\}, \{g + f_i\}$를 생각하면 이는 음이 아닌 가측함수열로서 Fatou의 보조정리로부터 $\liminf_{i\to\infty} \int_X g - f_i\,d\mu \ge \int_X g - f\,d\mu = \int_X g\,d\mu - \int_X f\,d\mu$이고, $\liminf_{i\to\infty} \int_X g - f_i\,d\mu = \int_X g\,d\mu - \limsup_{i\to\infty} \int_X f_i\,d\mu$에서 $\limsup_{i\to\infty} \int_X f_i\,\mu \le \int_X f\,d\mu$를 얻는다. 한편, $\{g + f_i\}$에 대해서도 이와 비슷하게 하면 $\liminf_{i\to\infty} \int_X f_i\,d\mu \ge \int_X f\,d\mu$를 얻어 곧 $\int_X f_i\,d\mu \to \int_X f\,d\mu$가 된다.

한편, $f_i \to f$ (ae.)인 경우에 대해서는 $N = \{x \in X : f_i(x) \not\to f(x)\}$을 생각하면 이가 영집합이고 정리 2.77의 iv로부터 가측이다. 이제 함수 $h_i : X \to \overline{\mathbb{R}}$의 열 $\{h_i\}$와 함수 $h : X \to \overline{\mathbb{R}}$를 각각 $h_i = f_i \mathbf{1}_{X\setminus N}, h = f\mathbf{1}_{X\setminus N}$과 같이 정의하면 $\{h_i\}$는 음이 아닌 가측함수의 열로, 증가하고 모든 $i \in \mathbb{N}$에 대해 $|h_i| \le g$이며 $h_i \uparrow h$이다. 그렇다면 앞서 얻은 결과로부터 $\int_X h_i\,d\mu \to \int_X h\,d\mu$임을 바로 알 수 있고, 곧 모든 $i \in \mathbb{N}$에 대해 $f_i = h_i$ (ae.)이며 $f = h$ (ae.)이므로 따름정리 2.102의 ii로부터 $\int_X f_i\,d\mu \to \int_X f\,d\mu$를 얻는다. □

이러한 수렴정리들로부터 얻어지는 다음 결과들도 유용하다.

Corollary 2.106 (Bounded convergence theorem) 유한 측도공간 $(X, \mathscr{A}, \mu)$에서 정의된 가측함수 $f_i : X \to \overline{\mathbb{R}}$의 열 $\{f_i\}$에 대해 적당한 가측함수 $f : X \to \overline{\mathbb{R}}$가 존재하여 $f_i \uparrow f$ (ae.)라 하자. 또한, $\{f_i\}$가 균등하게 유계라 하자. 즉, 어떤 $M > 0$이 존재하여 모든 $i \in \mathbb{N}$에 대해 $|f_i| \le M$이라 하자. 그렇다면 각 f_i와 f는 적분가능하고 $\int_X f_i\,d\mu \to \int_X f\,d\mu$이다.

Proof 상수 함수 $g = M$가 명백히 적분가능하므로 DCT로부터 이는 자명하다. □

Corollary 2.107 측도공간 $(X, \mathscr{A}, \mu)$에서 정의된 가측함수 $f_i : X \to \overline{\mathbb{R}}$의 열 $\{f_i\}$에 대해 다음이 성립한다.

i. 만약 각 f_i가 음이 아니면 $\int_X \sum_{i=1}^{\infty} f_i\,d\mu = \sum_{i=1}^{\infty} \int_X f_i\,d\mu$이다.
ii. 만약 $\sum_{i=1}^{\infty} f_i$가 거의 어디서나 수렴하거나 $\pm\infty$로 발산하며,[15]적당한 적분가능한 $g : X \to \overline{\mathbb{R}}_0^+$가 존재하여 모든 $k \in \mathbb{N}$에 대해 $|\sum_{i=1}^{k} f_i| \le g$라 하면 각 f_i와 $\sum_{i=1}^{\infty} f_i$는 적분가능하고 $\int_X \sum_{i=1}^{\infty} f_i\,d\mu = \sum_{i=1}^{\infty} \int_X f_i\,d\mu$이다.

PROOF 각각 MCT와 DCT로부터 자명하다. 특히 ii의 경우 모든 $k \in \mathbb{N}$에 대해 $\sum_{i=1}^{k} f_i$가 적분가능하므로 곧 임의의 $i \in \mathbb{N}$에 대해 $f_i = \sum_{j=1}^{i} f_j - \sum_{j=1}^{i-1} f_j$가 적분가능함을 안다. □

Corollary 2.108 측도공간 $(X, \mathscr{A}, \mu)$와 열린집합 $U \subseteq \mathbb{R}^n$ 대해 함수 $f : X \times U \to \mathbb{R}$가 다음을 조건을 만족한다고 하자.

i. 임의의 $y \in U$에 대해 함수 $x \mapsto f(x, y)$가 적분가능하다.
ii. 거의 대부분의 $x \in X$에 대해 함수 $y \mapsto f(x, y)$가 연속이다.
iii. 적당한 적분가능한 함수 $g : X \to \overline{\mathbb{R}}_0^+$가 존재하여 $X \times U$에서 $|f| \leq g$이다.

그렇다면 함수 $y \mapsto \int_X f(x, y)\, d\mu(x)$는 연속이다.

PROOF DCT로부터 거의 자명하다. 임의의 $y \in U$를 고정하고 U에 속하는 임의의 수열 $\{y_i\}$를 생각하여 $y_i \to y$라 하면 주어진 조건으로부터 거의 대부분의 $x \in X$에 대해 $f(x, y_i) \to f(x, y)$이고 임의의 $i \in \mathbb{N}$와 임의의 $x \in X$에 대해 $|f(x, y_i)| \leq g(x)$이므로, 곧 DCT에서 $\lim_{i\to\infty} \int_X f(x, y_i)\, d\mu(x) = \int_X f(x, y)\, d\mu(x)$가 되어 증명이 끝난다. □

Corollary 2.109 (Leibniz's rule) 측도공간 $(X, \mathscr{A}, \mu)$와 열린집합 $U \subseteq \mathbb{R}$ 대해 함수 $f : X \times U \to \mathbb{R}$가 다음 조건을 만족한다고 하자.

i. 임의의 $y \in U$에 대해 함수 $x \mapsto f(x, y)$가 적분가능하다.
ii. 거의 대부분의 $x \in X$에 대해 함수 $y \mapsto f(x, y)$가 미분가능하다.
iii. 임의의 $y \in U$에 대해 함수 $x \mapsto (\partial/\partial y) f(x, y)$가 가측이다.[16]
iv. 적당한 적분가능한 함수 $g : X \to \overline{\mathbb{R}}_0^+$가 존재하여 $X \times U$에서 $|\partial f/\partial y| \leq g$이다.

그렇다면 함수 $y \mapsto \int_X f(x, y)\, d\mu(x)$는 미분가능하고 임의의 $y \in U$에 대해 함수 $x \mapsto (\partial f/\partial y)(x, y)$는 적분가능하며 다음이 성립한다.

$$\frac{d}{dy} \int_X f(x, y)\, d\mu(x) = \int_X \frac{\partial}{\partial y} f(x, y)\, d\mu(x)$$

PROOF 편의를 위해 함수 $f_x : U \to \mathbb{R}$를 $f_x : y \mapsto f(x, y)$로 정의하자. 이제 임의의 $y \in U$를 고정하고 적당한 열린 ball B에 대해 함수 $k : X \times B \to \mathbb{R}$를

$$(x, h) \mapsto \begin{cases} \dfrac{f(x, y+h) - f(x, y)}{h} & h \neq 0 \text{ 이고 } f_x\text{가 미분가능한 경우} \\ \dfrac{\partial}{\partial y} f(x, y) & h = 0 \text{ 이고 } f_x\text{가 미분가능한 경우} \\ 0 & \text{ow.} \end{cases}$$

로 정의하면 이는 well-define되고, 만약 이가 따름정리 2. 108의 조건을 모두 만족함을 보일 수 있으면 곧바로

$$\lim_{h\to 0}\int_X \frac{f(x,y+h)-f(x,y)}{h}\,d\mu(x) = \lim_{h\to 0}\int_X k(x,h)\,d\mu(x)$$
$$= \int_X \frac{\partial}{\partial y} f(x,y)\,d\mu(x)$$

가 되어 증명이 끝난다. 그런데 주어진 조건으로부터 따름정리 2. 108의 조건 i, ii가 만족됨은 자명하므로 iii이 만족됨을 보이는 것만 보이자. 임의의 $(x,h)\in X\times V$에 대해 만약 f_x가 미분가능하지 않다면 $k(x,h)=0$에서 $|k(x,h)|\le g(x)$임이 분명하므로 f_x가 미분가능하다고 하자. 그렇다면 $h=0$인 경우에는 가정으로부터 $|k(x,h)|=|(\partial f/\partial y)(x,y)|\le g(x)$이고, 그렇지 않은 경우에는 MVT에서 y와 $y+h$ 사이의 적당한 $z\in U$가 존재하여 $|k(x,h)| = |(\partial f/\partial y)(x,z)|\le g(x)$가 되어 iii이 만족된다. 증명은 이로써 충분하다. □

측도를 이용해 적분을 정의한 것과 반대로 적분을 이용해 새로운 측도를 유도할 수도 있다. 이러한 기법은 확률론에서 널리 쓰이므로 여기서 잠시 살펴보는 것이 좋을 것 같다.

Theorem 2.110 측도공간 $(X,\mathscr{A},\mu)$에서 정의된 음이 아닌 가측함수 $f:X\to\overline{\mathbb{R}}_0^+$에 대해 함수 $\mu_f:\mathscr{A}\to\overline{\mathbb{R}}_0^+$를 $\mu_f:A\mapsto\int_A f\,d\mu$로 두면 μ_f는 $\mathscr{A}$ 위의 측도이다.

PROOF 당장 $\mu_f(\emptyset)=0$임은 분명하므로 μ_f가 σ−가법성을 가짐을 보이는 것으로 충분하다. 이를 위해 $\mathscr{A}$에 속하는 임의의 집합열 $\{A_i\}$를 택하면 MCT에서 $\mu_f(\bigsqcup_{i=1}^\infty A_i)=\int_{\bigsqcup_{i=1}^\infty A_i} f\,d\mu = \int_X \sum_{i=1}^\infty f\mathbf{1}_{A_i}\,d\mu = \sum_{i=1}^\infty\int_{A_i} f\,d\mu = \sum_{i=1}^\infty \mu_f(A_i)$가 되어 μ_f가 σ−가법성을 가짐을 알고, 곧 증명이 끝난다. □

Definition 2.111 가측공간 $(X,\mathscr{A},\mu)$에서 정의된 음이 아닌 가측함수 $f:X\to\overline{\mathbb{R}}_0^+$에 대해 f**로부터 유도되는 측도 (measure induced by** f**)**를 μ_f로 쓰고 $\mathscr{A}$ 위의 측도 $\mu_f:A\mapsto\int_A f\,d\mu$로 정의한다.

그렇다면 과연 이의 역도 성립할까에 대한 궁금증이 자연스럽게 피어난다. 어떤 고정된 측도 μ에 대해 임의의 측도 ν를 적당한 음이 아닌 가측함수 f를 사용하여 $\mu_f=\nu$와 같이 쓸 수 있을까? 만약 그렇다면 이러한 f는 유일할까? 만약 그 존재성이 일반적으로 보장되는 것이 아니라면, 어떤 조건에서 존재성을 말할 수 있을까? 이런 질문들에 대한 답은 자연스럽게 도함수의 일반화에 대한 아이디어로 이어지고, 이는 확률론이 꽃필 수 있는 견고한 기반이 된다. 다만, 그 자체로 꽤나 많은 논의가 필요하므로 일단 이와 관련된 논의는 잠시 미뤄두고, 여기에서는 적분에 관한 이야기를 마저 하도록 하자.

Theorem 2.112 측도공간 $(X,\mathscr{A},\mu)$에서 정의된 음이 아닌 가측함수 $f:X\to\overline{\mathbb{R}}_0^+$와 $\mathscr{A}$에 속하는 집합열 $\{A_i\}$에 대해 다음이 성립한다.

i. 만약 $A_i\uparrow A$이면 $\int_{A_i} f\,d\mu\uparrow\int_A f\,d\mu$이다.
ii. 만약 $A_i\downarrow A$이고 f가 A_1 위에서 적분가능하면 $\int_{A_i} f\,d\mu\downarrow\int_A f\,d\mu$이다.

한편, 적분가능한 함수 $f : X \to \overline{\mathbb{R}}$에 대해서도 위의 성질들이 성립한다. 특히 ii의 경우 f의 A_1 위에서의 적분가능성이 자명해지므로 이 조건을 생략할 수 있다.

ii° 만약 $A_i \downarrow A$ 이면 $\int_{A_i} f\,d\mu \downarrow \int_A f\,d\mu$ 이다.

PROOF 이는 μ_f 가 $\mathscr{A}$ 위에서의 측도를 이루므로 자명하다. 한편, 적분가능한 함수 $f : X \to \overline{\mathbb{R}}$에 대해서는 이를 각각 양과 음의 부분으로 나누어 각각이 $\mathscr{A}$ 위에서의 측도를 유도한다는 사실을 이용하면 된다. □

Theorem 2.113 측도공간 $(X, \mathscr{A}, \mu)$에서 정의된 음이 아닌 가측함수 $f, g : X \to \overline{\mathbb{R}}_0^+$ 에 대해 $\int_X g\,d\mu_f = \int_X fg\,d\mu$ 이다. 한편, μ_f−적분가능한 (혹은 fg가 μ−적분가능한 가측함수) $g : X \to \overline{\mathbb{R}}$에 대해서도 동일한 결과가 성립하며 이때 fg는 μ−적분가능하다 (혹은 g는 μ_f−적분가능하다).

PROOF 먼저 g가 음이 아닌 단순함수인 경우를 생각하여 이의 표준형을 $g = \sum_{i=1}^k a_i \mathbf{1}_{A_i}$ 라 하자. 그렇다면 $\int_X g\,d\mu_f = \sum_{i=1}^k a_i \mu_f(A_i) = \sum_{i=1}^k a_i \int_{A_i} f\,d\mu = \int_X \sum_{i=1}^n a_i f \mathbf{1}_{A_i}\,d\mu = \int_X fg\,d\mu$ 임이 분명하다. 이제 g를 음이 아닌 가측함수라 하면 정리 2. 83로부터 음이 아닌 단순함수 $g_i : X \to \mathbb{R}_0^+$ 의 열 $\{g_i\}$가 존재하여 $g_i \uparrow g$ 이므로 MCT에서 $\int_X g d\mu_f = \lim_{i\to\infty} \int_X g_i d\mu_f = \lim_{i\to\infty} \int_X f g_i d\mu = \int_X fg d\mu$ 이다.

한편, μ_f−적분가능한 $g : X \to \overline{\mathbb{R}}$에 대해서, 이를 양과 음의 부분으로 나누어 앞선 결과를 적용하면 fg가 μ−적분가능하며 $\int_X g\,d\mu_f = \int_X fg\,d\mu$ 임을 쉽게 알 수 있고, fg가 μ−적분가능한 가측함수 $g : X \to \overline{\mathbb{R}}$에 대해서도 비슷하게 하면 된다. □

Theorem 2.114 측도공간 $(X, \mathscr{A}, \mu)$에서 정의된 음이 아닌 가측함수 $f, g : X \to \overline{\mathbb{R}}_0^+$ 에 대해 만약 μ 가 σ−유한하며 임의의 $A \in \mathscr{A}$ 에 대해 $\int_A f\,d\mu = \int_A g\,d\mu$ 이면 $f = g$ (ae.) 이다. 한편, 적분가능한 $f, g : X \to \overline{\mathbb{R}}$에 대해서도 동일한 결과가 성립하며, 이 경우에는 μ 가 σ−유한하다는 조건을 생략할 수 있다.

PROOF 가정으로부터 $\mathscr{A}$ 에 속하는 적당한 집합열 $\{A_i\}$가 존재하여 $\mu(A_i) < \infty$ 이고 $\bigcup_{i=1}^\infty A_i = X$ 이고, WLOG, 필요하다면 각 항을 $C_i := \bigcup_{j=1}^i A_j$ 로 바꾸어 $\{A_i\}$가 처음부터 증가하는 집합열이라 해도 된다. 이제 집합열 $\{B_i\}$를 $B_i := \{x \in X : f(x) < g(x), f(x) < i\}$ 로 두면 이는 $\mathscr{A}$ 에 속하는 집합열이며 모든 $i \in \mathbb{N}$에 대해 $\int_{A_i \cap B_i} f\,d\mu = \int_{A_i \cap B_i} g\,d\mu \le \int_{A_i \cap B_i} i\,d\mu = i\mu(A_i \cap B_i) < \infty$ 이다. 이로부터 각 $i \in \mathbb{N}$에 대해 $\int_{A_i \cap B_i} g - f\,d\mu = 0$ 이고, 이는 보조정리 2. 102의 iii으로부터 $A_i \cap B_i$가 영집합임을 의미하므로 $\{x \in X : f(x) < g(x)\} = \bigcup_{i=1}^\infty (A_i \cup B_i)$ 도 영집합이 되어 $f \ge g$ (ae.) 이다. 한편, 비슷하게 하면 $f \le g$ (ae.) 임도 보일 수 있으므로 곧 $f = g$ (ae.) 임을 안다.

한편, 적분가능한 $f, g : X \to \overline{\mathbb{R}}$에 대해서도 비슷하게 하면 된다. 다만, 이 경우에는 f, g가 적분가능하여 $\int_{B_i} f\,d\mu = \int_{B_i} g\,d\mu < \infty$ 이므로 μ 가 σ−유한하다는 조건이 필요하지 않다. □

측도공간의 완비화와 관련된 정리를 끝으로 길었던 Lebesgue 적분론의 전개를 마무리하도록 하겠다.

Theorem 2.115 측도공간 $(X, \mathscr{A}, \mu)$에서 정의된 음이 아닌 $\mathscr{A}/\mathscr{B}_1$−가측함수 $f : X \to \overline{\mathbb{R}}_0^+$에 대해 $\int_X f\,d\mu = \int_X f\,d\overline{\mu}$이다. 한편, μ−적분가능한(혹은 $\overline{\mu}$−적분가능한 $\mathscr{A}/\mathscr{B}_1$−가측함수) $f : X \to \overline{\mathbb{R}}$에 대해서도 동일한 결과가 성립하며 이때 f는 $\overline{\mu}$−적분가능하다(혹은 μ−적분가능하다).

Proof 먼저 f가 음이 아닌 단순함수인 경우를 생각하여 이의 표준형을 $f = \sum_{i=1}^k a_i \mathbf{1}_{A_i}$라 하자. 그렇다면 $\int_X f\,d\mu = \sum_{i=1}^k a_i \mu(A_i) = \sum_{i=1}^k a_i \overline{\mu}(A_i) = \int_X f\,d\overline{\mu}$임이 분명하다. 이제 f를 음이 아닌 가측함수라 하면 정리 2. 83로부터 음이 아닌 단순함수 $f_i : X \to \mathbb{R}_0^+$의 열 $\{f_i\}$가 존재하여 $f_i \uparrow f$이므로 MCT에서 $\int_X f\,d\mu = \lim_{i\to\infty} \int_X f_i\,d\mu = \lim_{i\to\infty} \int_X f_i\,d\overline{\mu} = \int_X f\,d\overline{\mu}$이다.

한편, μ−적분가능한 $f : X \to \overline{\mathbb{R}}$에 대해서, 이를 양과 음의 부분으로 나누어 앞선 결과를 적용하면 f가 $\overline{\mu}$−적분가능하며 $\int_X f\,d\mu = \int_X f\,d\overline{\mu}$임을 쉽게 알 수 있고, $\overline{\mu}$−적분가능한 $\mathscr{A}/\mathscr{B}_1$−가측함수 $f : X \to \overline{\mathbb{R}}$에 대해서도 비슷하게 하면 된다. □

2.8 Product Measures and Contruction of The Lebesgue Measure II

이전 절에서 Lebesgue 적분론의 전개가 마무리 되었으므로 이제 우리의 원래 목적으로 되돌아와 Lebesgue 측도의 두 번째 구성을 본격적으로 시작해보려고 한다. 앞서 언급했듯이 이를 위해서는 우선 적당한 곱연산을 정의해야 한다.

Definition 2.116 가측공간 $(X, \mathscr{A}), (Y, \mathscr{B})$에 대해 적당한 $A \in \mathscr{A}, B \in \mathscr{B}$가 존재하여 $C = A \times B$로 쓸 수 있는 집합 $C \subseteq X \times Y$를 **measurable rectangle**이라 하고, 모든 measurable rectangle의 모임을 $\mathscr{A} \times \mathscr{B}$로 쓴다. 이때 $\sigma(\mathscr{A} \times \mathscr{B})$를 $\mathscr{A}, \mathscr{B}$의 **product σ−algebra**라 하며 $\mathscr{A} \otimes \mathscr{B}$로 쓴다.

위의 정의로부터 σ−대수의 $\times$곱은 결합법칙을 만족함이 분명하다. 따라서 가측공간 $(X_1, \mathscr{A}_1), \cdots, (X_i, \mathscr{A}_k)$에 대해 $\prod_{i=1}^k \mathscr{A}_i := \mathscr{A}_1 \times \cdots \times \mathscr{A}_k$가 well-define된다. 한편, σ−대수의 $\otimes$곱이 결합법칙을 만족하는가는 조금 덜 자명하다.

Theorem 2.117 가측공간 $(X, \mathscr{A}), (Y, \mathscr{B}), (Z, \mathscr{C})$에 대해 $(\mathscr{A} \otimes \mathscr{B}) \otimes \mathscr{C} = \mathscr{A} \otimes (\mathscr{B} \otimes \mathscr{C}) = \sigma(\mathscr{A} \times \mathscr{B} \times \mathscr{C})$이다.

Proof 표기의 편의를 위해 $\mathscr{D} = \sigma(\mathscr{A} \times \mathscr{B} \times \mathscr{C})$라 하자. 만약 $(\mathscr{A} \otimes \mathscr{B}) \otimes \mathscr{C} = \mathscr{A} \otimes (\mathscr{B} \otimes \mathscr{C}) = \mathscr{D}$임을 보일 수 있으면 증명이 끝난다. 이를 위해 임의의 $C \in \mathscr{C}$를 고정하

고 $\mathscr{E} = \{A \in \mathscr{A} \otimes \mathscr{B} : A \times C \in \mathscr{D}\}$를 생각하면 $X \times Y \in \mathscr{E}$임이 분명하다. 또한, 임의의 $A \in \mathscr{E}$에 대해 $A^c \times C = [(X \times Y) \times C] \setminus (A \times C) \in \mathscr{D}$에서 $\mathscr{E}$는 여집합에 대해 닫혀있다. 한편, $\mathscr{E}$에 속하는 임의의 서로소인 집합열 $\{A_i\}$에 대해 $(\bigsqcup_{i=1}^{\infty} A_i) \times C = \bigsqcup_{i=1}^{\infty}(A_i \times C) \in \mathscr{D}$이므로 이상으로부터 $\mathscr{E}$이 $\lambda-$system임을 알 수 있다. 한편, 임의의 $A, B \in \mathscr{E}$에 대해 $(A \cap B) \times C = (A \times C) \cap (B \times C) \in \mathscr{E}$에서 $\mathscr{E}$는 교집합에 대해서도 닫혀있어 곧 $\pi-$system이기도 하다. 그렇다면 보조정리 2. 10로부터 $\mathscr{E}$는 $\sigma-$대수이고, 정의로부터 $\mathscr{A} \times \mathscr{B} \subseteq \mathscr{E}$임이 분명하므로 $\mathscr{A} \otimes \mathscr{B} \subseteq \mathscr{E}$이다. 이제 이상의 결과가 임의의 $C \in \mathscr{C}$에 대해 성립함을 상기한다면 $(\mathscr{A} \otimes \mathscr{B}) \times \mathscr{C} \subseteq \mathscr{D}$가 되어 곧 $(\mathscr{A} \otimes \mathscr{B}) \otimes \mathscr{C} \subseteq \mathscr{D}$이다. 한편, $\mathscr{A} \times \mathscr{B} \times \mathscr{C} \subseteq (\mathscr{A} \otimes \mathscr{B}) \otimes \mathscr{C}$가 분명하므로 $\mathscr{D} \subseteq (\mathscr{A} \otimes \mathscr{B}) \otimes \mathscr{C}$에서 역의 포함관계도 성립하고, 곧 $(\mathscr{A} \otimes \mathscr{B}) \otimes \mathscr{C} = \mathscr{D}$이다. 비슷한 방법으로 $\mathscr{A} \otimes (\mathscr{B} \otimes \mathscr{C}) = \mathscr{D}$도 보일 수 있다. □

이로부터 $\sigma-$대수의 $\otimes$곱도 결합법칙을 만족하므로 $\bigotimes_{i=1}^{k} \mathscr{A}_i = \mathscr{A}_1 \otimes \cdots \otimes \mathscr{A}_k$가 well-define된다. 만약 $\mathscr{A}_1 = \cdots = \mathscr{A}_k =: \mathscr{A}$이면 이들의 $\otimes$곱을 $\mathscr{A}^{\otimes k}$로 간단히 쓰기도 한다.

Definition 2.118 집합 X, Y와 $A \subseteq X \times Y$, 점 $x_0 \in X, y_0 \in Y$에 대해 A의 **section** A_{x_0}, A^{y_0}를 각각 $A_{x_0} = \{y \in Y : (x_0, y) \in A\}, A^{y_0} = \{x \in X : (x, y_0) \in A\}$로 정의한다.

Definition 2.119 집합 X, Y, Z와 함수 $f : X \times Y \to Z$, 점 $x_0 \in X, y_0 \in Y$에 대해 f의 **section** $f_{x_0} : Y \to Z, f^{y_0} : X \to Z$를 각각 $f_{x_0} : y \mapsto f(x_0, y), f^{y_0} : x \mapsto f(x, y_0)$로 정의한다.

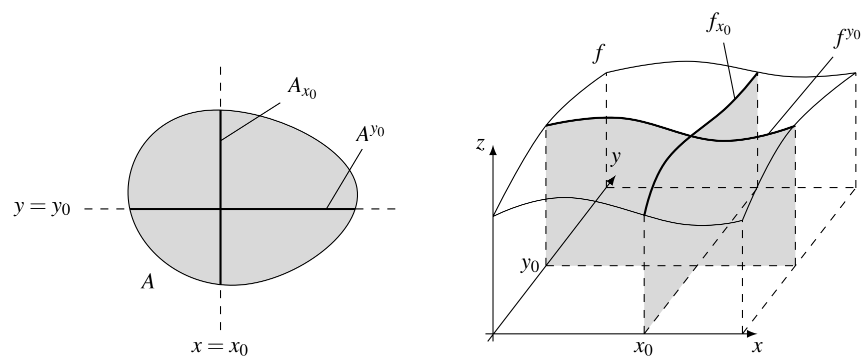


Figure 2.9 집합 $A \subseteq \mathbb{R}^2$(왼쪽)와 함수 $f : \mathbb{R}^2 \to \mathbb{R}$(오른쪽)의 section들.

다음 정리에 따르면 product $\sigma-$algebra에 속하는 measurable rectangle을 포함한 모든 집합은 항상 가측인 section을 가진다. 이로써 우리는 product $\sigma-$algebra의 구조를 조금 더 잘 이해할 수 있다.

Theorem 2.120 가측공간 $(X, \mathscr{A}), (Y, \mathscr{B}), (Z, \mathscr{C})$에 대해 다음이 성립한다.

i. 임의의 $A \in \mathscr{A} \otimes \mathscr{B}$와 임의의 $x \in X, y \in Y$에 대해 $A_x \in \mathscr{B}, A^y \in \mathscr{A}$이다.
ii. $\mathscr{A} \otimes \mathscr{B}/\mathscr{C}-$가측함수 $f : X \times Y \to Z$와 임의의 $x \in X, y \in Y$에 대해 f_x, f^y는 각각 $\mathscr{B}/\mathscr{C}-$가측이고 $\mathscr{A}/\mathscr{C}-$가측이다.

Proof i. 임의의 $x \in X$를 고정하고 $\mathscr{D} = \{A \in \mathscr{A} \otimes \mathscr{B} : A_x \in \mathscr{B}\}$를 생각하면 임의의 $A \times B \in \mathscr{A} \times \mathscr{B}$에 대해

$$(A \times B)_x = \begin{cases} B & x \in A \text{인 경우} \\ \emptyset & \text{ow.} \end{cases}$$

이므로 $\mathscr{A} \times \mathscr{B} \subseteq \mathscr{D}$이고, 따라서 $\mathscr{D}$가 $\sigma-$대수라는 것만 보이면 $\mathscr{A} \otimes \mathscr{B} \subseteq \mathscr{D}$가 되어 증명이 끝난다. 이를 위해 임의의 $A \in \mathscr{D}$를 생각하면 $(A^c)_x = (A_x)^c \in \mathscr{B}$이므로 $A^c \in \mathscr{D}$에서 $\mathscr{D}$는 여집합에 대해 닫혀있다. 또한 $\mathscr{D}$에 속하는 임의의 집합열 $\{A_i\}$에 대해 $(\bigcup_{i=1}^{\infty} A_i)_x = \bigcup_{i=1}^{\infty} (A_i)_x \in \mathscr{B}$이므로 $\bigcup_{i=1}^{\infty} A_i \in \mathscr{D}$에서 $\mathscr{D}$는 가산 합집합에 대해서도 닫혀있다. 이제 이가 공집합이 아님은 분명하므로 곧 $\sigma-$대수이고, 증명이 끝난다. 한편, A^y에 대해서도 이와 비슷하게 하면 된다.

ii. 임의의 $A \in \mathscr{C}$에 대해 $(f_x)^{-1}(A) = [f^{-1}(A)]_x$이고 $(f^y)^{-1}(A) = [f^{-1}(A)]^y$이므로 이는 i로부터 자명하다. □

이제 product $\sigma-$algebra에 측도를 부여하는 방법에 대해 생각해보자. 만약 곱해지는 두 가측공간에 적당한 측도가 정의되어 있었다면 이를 사용하면서도 곱셈이라는 구조를 잘 반영할 수 있도록 product $\sigma-$algebra에 측도를 부여하는 것이 자연스러울 것이다. 이런 관점에서, 다음 정리는 product $\sigma-$algebra에 자연스러운 측도를 정의할 구체적인 방법을 제공하는 동시에, 이러한 측도의 유일성까지 보장한다.

Lemma 2.121 $\sigma-$유한 측도공간 $(X, \mathscr{A}, \mu), (Y, \mathscr{B}, \nu)$와 임의의 $A \in \mathscr{A} \otimes \mathscr{B}$에 대해 함수 $\varphi_A : X \to \overline{\mathbb{R}}_0^+$를 $\varphi_A : x \mapsto \nu(A_x)$로 정의하면 이는 $\mathscr{A}/\mathscr{B}_1-$가측이다. 비슷하게, 함수 $\psi_A : Y \to \overline{\mathbb{R}}_0^+$를 $\psi_A : y \mapsto \mu(A^y)$로 정의하면 이는 $\mathscr{B}/\mathscr{B}_1-$가측이다.

Proof 먼저 ν가 유한한 경우를 생각하여 $\mathscr{L} = \{A \in \mathscr{A} \otimes \mathscr{B} : \varphi_A\text{가 } \mathscr{A}/\mathscr{B}_1\text{−가측}\}$을 생각하면 임의의 $A \times B \in \mathscr{A} \times \mathscr{B}$에 대해 $\varphi_{A\times B}(x) = \nu((A \times B)_x) = \nu(B)\mathbf{1}_A(x)$이므로 이가 단순함수가 되어 $A \times B \in \mathscr{L}$이고, 곧 $\mathscr{A} \times \mathscr{B} \subseteq \mathscr{L}$이다. 이제 $\mathscr{A} \times \mathscr{B}$가 $\pi-$system임은 분명하므로 만약 우리가 $\mathscr{L}$이 $\lambda-$system임을 보일 수 있으면 Dynkin의 $\pi-\lambda$ 정리로부터 $\mathscr{L} = \mathscr{A} \otimes \mathscr{B}$가 되어 증명이 끝난다. 이를 위해 임의의 $A \in \mathscr{L}$를 택하면 $\varphi_{A^c}(x) = \nu((A^c)_x) = \nu(Y \setminus A_x) = \nu(Y) - \nu(A_x) = \nu(Y) - \varphi_A(x)$에서 이가 $\mathscr{A}/\mathscr{B}_1-$가측이므로 $A^c \in \mathscr{L}$이고, 곧 $\mathscr{L}$은 여집합에 대해 닫혀있다. 또한, $\mathscr{L}$에 속하는 임의의 서로소인 집합열 $\{A_i\}$에 대해 $\varphi_{\bigsqcup_{i=1}^{\infty} A_i}(x) = \nu((\bigsqcup_{i=1}^{\infty} A_i)_x) = \nu(\bigsqcup_{i=1}^{\infty} (A_i)_x) = \sum_{i=1}^{\infty} \nu((A_i)_x) = \sum_{i=1}^{\infty} \varphi_{A_i}(x)$에서 이가

$\mathscr{A}/\mathscr{B}_1$−가측이므로 $\bigsqcup_{i=1}^{\infty} A_i \in \mathscr{L}$ 이고, $X \times Y \in \mathscr{L}$ 임은 분명하므로 곧 $\mathscr{L}$ 이 λ−system이 되어 증명이 끝난다.

한편, ν 가 σ−유한한 경우에는 정의로부터 $\mathscr{B}$ 에 속하는 적당한 집합열 $\{B_i\}$ 가 존재하여 각 $i \in \mathbb{N}$ 에 대해 $\nu(B_i) < \infty$ 이고 $\bigcup_{i=1}^{\infty} B_i = Y$ 이다. WLOG, 필요하다면 각 항을 $C_i := B_i \setminus \bigcup_{j=1}^{i-1} A_j$ 로 바꾸어 $\{B_i\}$ 가 처음부터 서로소였다고 해도 된다. 이제 각 $i \in \mathbb{N}$ 에 대해 측도 ν_{B_i} 가 유한하므로 곧 임의의 $A \in \mathscr{A} \otimes \mathscr{B}$ 에 대해 함수 $\varphi_A^i : X \to \overline{\mathbb{R}}_0^+$ 를 $\varphi_A^i : x \mapsto \nu_{B_i}(A_x)$ 로 두면 앞선 결과로부터 이가 $\mathscr{A}/\mathscr{B}_1$−가측이고, 곧 $\varphi_A = \sum_{i=1}^{\infty} \varphi_A^i$ 에서 φ_A 가 $\mathscr{A}/\mathscr{B}_1$−가측이다. 남은 ψ_A 에 대해서도 비슷하게 하면 된다. □

Theorem 2.122 σ−유한 측도공간 $(X, \mathscr{A}, \mu), (Y, \mathscr{B}, \nu)$ 에 대해 적당한 $\mathscr{A} \otimes \mathscr{B}$ 위의 측도 $\mu \otimes \nu : \mathscr{A} \otimes \mathscr{B} \to \overline{\mathbb{R}}_0^+$ 가 존재하여 임의의 $A \times B \in \mathscr{A} \times \mathscr{B}$ 에 대해 $(\mu \otimes \nu)(A \times B) = \mu(A)\nu(B)$ 이다. 나아가, 이러한 측도는 유일하고, σ−유한이며 $(\mu \otimes \nu)(A) = \int_X \nu(A_x)\, d\mu(x) = \int_Y \mu(A^y)\, d\nu(y)$ 이다.

Proof 가정으로부터 각각 $\mathscr{A}, \mathscr{B}$ 에 속하는 적당한 집합열 $\{A_i\}, \{B_j\}$ 가 존재하여 각 $i, j \in \mathbb{N}$ 에 대해 $\mu(A_i), \nu(B_j) < \infty$ 이고 $\bigcup_{i=1}^{\infty} A_i = X, \bigcup_{j=1}^{\infty} B_j = Y$ 이다. 이제, 집합열 $\{A_i \times B_j\}_{ij}$ 는 $\mathscr{A} \otimes \mathscr{B}$ 에 속하는 집합열로 $\bigcup_{i=1}^{\infty} \bigcup_{j=1}^{\infty} (A_i \times B_j) = X \times Y$ 이며 만약 정리의 조건을 만족하는 측도가 존재하기만 한다면 $(\mu \otimes \nu)(A_i \times B_j) = \mu(A_i)\nu(B_i) < \infty$ 가 되어 그 측도는 항상 σ−유한하다. 또한, 집합족 $\mathscr{A} \times \mathscr{B}$ 가 π−system임이 분명하므로 정리 2. 29에서 우리가 원하는 측도는 존재하기만 한다면 유일하다. 따라서 증명은 정리의 조건을 만족하는 측도가 존재함을 보이는 것으로 충분하다.

이를 위해 함수 $\xi : \mathscr{A} \otimes \mathscr{B} \to \overline{\mathbb{R}}_0^+$ 를 $\xi : A \mapsto \int_X \nu(A_x)\, d\mu(x)$ 로 두고 이가 정리의 조건을 만족하는 측도임을 보이자. 우선 위의 보조정리로부터 이는 well-define되며 $\xi(\emptyset) = \int_X \nu(\emptyset)\, d\mu = 0$ 임은 분명하고, $\mathscr{A} \otimes \mathscr{B}$ 에 속하는 임의의 서로소인 집합열 $\{A_i\}$ 에 대해 MCT로부터

$$\begin{aligned}
\xi\left(\bigsqcup_{i=1}^{\infty} A_i\right) &= \int_X \nu\left(\left(\bigsqcup_{i=1}^{\infty} A_i\right)_x\right) d\mu \\
&= \int_X \nu\left(\bigsqcup_{i=1}^{\infty} (A_i)_x\right) d\mu \\
&= \int_X \sum_{i=1}^{\infty} \nu((A_i)_x) d\mu \\
&= \sum_{i=1}^{\infty} \int_X \nu((A_i)_x) d\mu \\
&= \sum_{i=1}^{\infty} \xi(A_i)
\end{aligned}$$

이므로 ξ가 $\mathscr{A}\otimes\mathscr{B}$ 위의 측도임을 안다. 한편, 임의의 $A\times B\in\mathscr{A}\times\mathscr{B}$에 대해 $\xi(A\times B)=\int_X \nu((A\times B)_x)\,d\mu(x)=\int_X \nu(B)\mathbf{1}_A(x)\,d\mu(x)=\mu(A)\nu(B)$에서 ξ는 정리의 조건도 만족하여 증명이 끝난다. 한편, 비슷한 방법으로 $(\mu\otimes\nu)(A)=\int_Y \mu(A^y)\,d\nu(y)$임도 보일 수 있다. □

Definition 2.123 σ−유한 측도공간 $(X,\mathscr{A},\mu),(Y,\mathscr{B},\nu)$에 대해 μ와 ν의 **곱측도 (product measure)**를 $\mu\otimes\nu$로 쓰고 모든 $A\times B\in\mathscr{A}\times\mathscr{B}$에 $\mu(A)\nu(B)$의 측도를 부여하는 $\mathscr{A}\otimes\mathscr{B}$ 위의 유일한 측도로 정의한다. 나아가, 이때의 측도공간 $(X\times Y,\mathscr{A}\otimes\mathscr{B},\mu\otimes\nu)$를 $(X,\mathscr{A},\mu)$와 $(Y,\mathscr{B},\nu)$의 **product measure space**라 한다.

위의 정리에서 곱해지는 두 측도공간이 σ−유한해야 한다는 조건을 필수적이다. 만약 이들의 σ−유한성이 보장되지 않으면 Carathéodory의 확장정리로부터 존재성은 여전히 보장되지만 유일성은 더 이상 성립하지 않는다. [17]하지만, 우리가 주로 다룰 Lebesgue 측도공간, Borel 측도공간, 나중에 소개될 확률공간이 모두 σ−유한하므로 σ−유한의 가정이 깨어지는 경우에 대해 이 이상으로 논하지는 않겠다. 한편, 이렇게 정의된 측도 간의 곱이 과연 결합법칙을 만족시키는지에 대한 의문이 생겨난다.

Theorem 2.124 σ−유한 측도공간 $(X,\mathscr{A},\mu),(Y,\mathscr{B},\nu),(Z,\mathscr{C},\xi)$에 대해 $(\mu\otimes\nu)\otimes\xi=\mu\otimes(\nu\otimes\xi)$이다.

Proof 정리 2.117로부터 $\mathscr{A}\times\mathscr{B}\times\mathscr{C}$가 $\mathscr{A}\otimes\mathscr{B}\otimes\mathscr{C}$의 생성자이고, 이가 명백히 π−system이므로 정리 2.29로부터 $(\mu\otimes\nu)\otimes\xi$와 $\mu\otimes(\nu\otimes\xi)$가 $\mathscr{A}\times\mathscr{B}\times\mathscr{C}$에서 일치함을 보이는 것으로 충분한데, 이는 곱측도의 정의로부터 자명하다. □

따라서 측도의 곱이 결합법칙을 만족하므로 σ−유한 측도공간 $(X_1,\mathscr{A}_1,\mu_1),\cdots,(X_k,\mathscr{A}_k,\mu_k)$에 대해 $\bigotimes_{i=1}^k\mu_i=\mu_1\otimes\cdots\otimes\mu_k$가 well-define된다. 만약 $\mu_1=\cdots=\mu_k=:\mu$이면 이들의 곱을 $\mu^{\otimes k}$로 간단히 쓰기도 한다.

이로써 Lebesgue 측도의 두 번째 구성을 위해 준비가 끝났다.

Theorem 2.125 $l+m=n$인 $l,m,n\in\mathbb{N}$에 대해 다음이 성립한다.

i. Borel 측도공간 $(\mathbb{R}^n,\mathscr{B}_n,\mu_n)$은 $(\mathbb{R}^l,\mathscr{B}_l,\mu_l)$과 $(\mathbb{R}^m,\mathscr{B}_m,\mu_m)$의 product measure space이다. 즉, $(\mathbb{R}^n,\mathscr{B}_n,\mu_n)=(\mathbb{R}^l\times\mathbb{R}^m,\mathscr{B}_l\otimes\mathscr{B}_m,\mu_l\otimes\mu_m)$.
ii. Lebesgue 측도공간 $(\mathbb{R}^n,\mathscr{M}_n,\lambda_n)$은 $(\mathbb{R}^l,\mathscr{M}_l,\lambda_l)$과 $(\mathbb{R}^m,\mathscr{M}_m,\lambda_m)$의 product measure space의 완비화이다. 즉, $(\mathbb{R}^n,\mathscr{M}_n,\lambda_n)=(\mathbb{R}^l\times\mathbb{R}^m,\overline{\mathscr{M}_l\otimes\mathscr{M}_m},\overline{\lambda_l\otimes\lambda_m})$.

Proof i과 ii를 총 4단계에 걸쳐 한 번에 증명하도록 하자.

Step 1. $\mathscr{B}_n=\mathscr{B}_l\otimes\mathscr{B}_m$.

임의의 열린 $U\subseteq\mathbb{R}^n$에 대해 이는 $\mathbb{R}^l$과 $\mathbb{R}^m$의 열린집합의 Cartesian 곱의 가산 합집합으로 쓸 수 있으므로[18] $\mathscr{U}$를 $\mathbb{R}^n$의 모든 열린집합의 모임이라 하면 $\mathscr{U}\subseteq\mathscr{B}_l\otimes\mathscr{B}_m$

에서 $\mathscr{B}_n \subseteq \mathscr{B}_l \otimes \mathscr{B}_m$ 이다. 이제 $\pi_1 : \mathbb{R}^n \to \mathbb{R}^l, \pi_2 : \mathbb{R}^n \to \mathbb{R}^m$ 를 각각 처음 l 개의 좌표와 마지막 m 개의 좌표로의 사영이라 한다면 이들이 연속함수이므로 Borel이고, 곧 임의의 $A \in \mathscr{B}_l, B \in \mathscr{B}_m$ 에 대해 $A \times \mathbb{R}^m = \pi_1^{-1}(A), \mathbb{R}^l \times B = \pi_2^{-1}(B)$ 가 모두 Borel이 되어 $A \times B = (A \times \mathbb{R}^m) \cap (\mathbb{R}^l \times B) \in \mathscr{B}_n$ 이므로 $\mathscr{B}_l \times \mathscr{B}_m \subseteq \mathscr{B}_n$ 에서 $\mathscr{B}_l \otimes \mathscr{B}_m \subseteq \mathscr{B}_n$ 이고, 이로써 $\mathscr{B}_n = \mathscr{B}_l \otimes \mathscr{B}_m$ 임을 안다.

Step 2. 임의의 $A \in \mathscr{B}_n$ 에 대해 이는 $\mathscr{M}_l \otimes \mathscr{M}_m$–가측이고 $\lambda_n(A) = (\lambda_l \otimes \lambda_m)(A)$ 이다.

1단계로부터 $\mathscr{B}_n = \mathscr{B}_l \otimes \mathscr{B}_m \subseteq \mathscr{M}_l \otimes \mathscr{M}_m$ 이고, 따라서 A 가 $\mathscr{M}_l \otimes \mathscr{M}_m$–가측임은 분명하다. 이제 남은 부분을 보이기 위해 먼저 $A \in \mathscr{S}_n$ 인 경우를 생각해보면, 적당한 반열린구간 $I_1, \cdots, I_n \subseteq \mathbb{R}$ 이 존재하여 $A = \prod_{i=1}^n I_i$ 이다. 이로부터 $\lambda_n(A) = \prod_{i=1}^n \lambda_1(I_i) = [\prod_{i=1}^l \lambda_1(I_i)][\prod_{i=l+1}^n \lambda_1(I_i)] = \lambda_l(\prod_{i=1}^l I_i)\lambda_m(\prod_{i=l+1}^n I_i) = (\lambda_l \otimes \lambda_m)(A)$ 가 성립한다. 다음으로, A 가 열린집합인 경우를 생각해보면 $\mathscr{S}_n$ 에 속하는 적당한 서로소인 집합열 $\{B_j\}$ 에 대해 $A = \bigsqcup_{j=1}^\infty B_j$ 로 쓸 수 있으므로[19]이 경우에도 $\lambda_n(A) = \lambda_n(\bigsqcup_{j=1}^\infty B_j) = \sum_{j=1}^\infty \lambda_n(B_j) = \sum_{j=1}^\infty (\lambda_l \otimes \lambda_m)(B_j) = (\lambda_l \otimes \lambda_m)(\bigsqcup_{j=1}^\infty B_j) = (\lambda_l \otimes \lambda_m)(A)$ 가 성립한다. 이번에는 A 가 compact한 경우를 생각해보면 이는 유계이므로 적당한 유계인 열린집합 $U \subseteq \mathbb{R}^n$ 가 존재하여 $A \subseteq U$ 이고 $\lambda_n(U) < \infty$ 이다. 따라서 $V = U \setminus A$ 라 하면 이 또한 열려있고 유계이므로 $\lambda_n(V) < \infty$ 이어서 이전의 결과로부터 $\lambda_n(A) = \lambda_n(U \setminus V) = \lambda_n(U) - \lambda_n(V) = (\lambda_l \otimes \lambda_m)(U) - (\lambda_l \otimes \lambda_m)(V) = (\lambda_l \otimes \lambda_m)(U \setminus V) = (\lambda_l \otimes \lambda_m)(A)$ 를 얻는다. 마지막으로, 일반적인 $A \in \mathscr{B}_n$ 에 대해서는 앞선 결과들과 정리 2.54로부터

$$
\begin{aligned}
\lambda_n(A) &= \inf\{\lambda_n(U) \in \overline{\mathbb{R}}_0^+ : A \subseteq U \subseteq \mathbb{R}^n, U\text{는 열린집합}\} \\
&= \inf\{(\lambda_l \otimes \lambda_m)(U) \in \overline{\mathbb{R}}_0^+ : A \subseteq U \subseteq \mathbb{R}^n, U\text{는 열린집합}\} \\
&\geq (\lambda_l \otimes \lambda_m)(A)
\end{aligned}
$$

이고

$$
\begin{aligned}
\lambda_n(A) &= \sup\{\lambda_n(K) \in \overline{\mathbb{R}}_0^+ : K \subseteq A, K\text{는 compact 집합}\} \\
&= \sup\{(\lambda_l \otimes \lambda_m)(K) \in \overline{\mathbb{R}}_0^+ : K \subseteq A, K\text{는 compact 집합}\} \\
&\leq (\lambda_l \otimes \lambda_m)(A)
\end{aligned}
$$

이므로 $\lambda_n(A) = (\lambda_l \otimes \lambda_m)(A)$ 임을 안다.

이상의 결론으로부터 임의의 $A \times B \in \mathscr{B}_l \times \mathscr{B}_m \subseteq \mathscr{B}_n$ 에 대해 $\mu_n(A \times B) = \mu_l(A)\mu_m(B)$ 이므로 정리 2.29로부터 $\mu_n = \mu_l \otimes \mu_m$ 이고, 곧 i의 증명이 끝난다.

Step 3. $\mathscr{M}_l \otimes \mathscr{M}_m \subseteq \mathscr{M}_n$.

임의의 $A \in \mathscr{M}_l = \overline{\mathscr{B}_l}$ 를 택하면 정리 2.66로부터 적당한 $B, C \in \mathscr{B}_l$ 가 존재하여 $B \subseteq A \subseteq C$ 이고 $\mu_l(C \setminus B) = 0$ 이다. 이제 $\pi : \mathbb{R}^n \to \mathbb{R}^l$ 를 처음 l 개의 좌표로의 사영이라 하면 이가 연

속이므로 Borel이고 곧 $B\times\mathbb{R}^m=\pi^{-1}(B), C\times\mathbb{R}^m=\pi^{-1}(C)$ 가 Borel이다. 여기에 앞선 단계의 결과들을 적용하면 $B\times\mathbb{R}^m\subseteq A\times\mathbb{R}^m\subseteq C\times\mathbb{R}^m$ 이고 $\lambda_n((C\times\mathbb{R}^m)\setminus(B\times\mathbb{R}^m))=\lambda_n((C\setminus B)\times\mathbb{R}^m)=\lambda_l(C\setminus B)\lambda_m(\mathbb{R}^m)=0$ 이 되어 정리 2.66에서 $A\times\mathbb{R}^m\in\mathscr{M}_n$ 이다. 비슷하게 임의의 $B\in\mathscr{M}_m$ 에 대해 $\mathbb{R}^l\times B\in\mathscr{M}_n$ 임을 보일 수 있으므로 임의의 $A\times B\in\mathscr{M}_l\times\mathscr{M}_m$ 에 대해 $A\times B=(A\times\mathbb{R}^m)\cap(\mathbb{R}^l\times B)\in\mathscr{M}_n$ 이 되어 $\mathscr{M}_l\times\mathscr{M}_m\subseteq\mathscr{M}_n$ 에서 $\mathscr{M}_l\otimes\mathscr{M}_m\subseteq\mathscr{M}_n$ 이다.

Step 4. $\mathscr{M}_n=\overline{\mathscr{M}_l\otimes\mathscr{M}_m}$ 이고 $\lambda_n=\overline{\lambda_l\otimes\lambda_m}$ 이다.

먼저 $\mathscr{M}_n\subseteq\overline{\mathscr{M}_l\otimes\mathscr{M}_m}$ 이고 $\lambda_n=\overline{\lambda_l\otimes\lambda_m}|_{\mathscr{M}_n}$ 임을 보이기 위해 임의의 $A\in\mathscr{M}_n=\overline{\mathscr{B}_n}$ 를 택하면 정리 2.66로부터 적당한 $B,C\in\mathscr{B}_n\subseteq\mathscr{M}_l\otimes\mathscr{M}_m$ 가 존재하여 $B\subseteq A\subseteq C$ 이고 $(\lambda_l\otimes\lambda_m)(C\setminus B)=\lambda_n(C\setminus B)=0$ 이다. 그런데 이는 다시 정리 2.66로부터 $A\in\overline{\mathscr{M}_l\otimes\mathscr{M}_m}$ 임을 뜻하여 곧 $\mathscr{M}_n\subseteq\overline{\mathscr{M}_l\otimes\mathscr{M}_m}$ 이다. 나아가 $A=B\sqcup(A\setminus B)$ 이고 $A\setminus B\subseteq C\setminus B$ 가 $\lambda_l\otimes\lambda_m$–영집합이므로 $\overline{\lambda_l\otimes\lambda_m}(A)=(\lambda_l\otimes\lambda_m)(B)=\lambda_n(B)=\lambda_n(B)+\lambda_n(A\setminus B)=\lambda_n(B\sqcup(A\setminus B))=\lambda_n(A)$ 에서 $\overline{\lambda_l\otimes\lambda_m}|_{\mathscr{M}_n}=\lambda_n$ 임을 안다. 이제 $\mathscr{M}_n\supseteq\overline{\mathscr{M}_l\otimes\mathscr{M}_m}$ 만 보이면 증명이 끝난다. 이를 위해 임의의 $A\in\overline{\mathscr{M}_l\otimes\mathscr{M}_m}$ 를 택하면 정리 2.66와 앞선 결과로부터 적당한 $B,C\in\mathscr{M}_l\otimes\mathscr{M}_m\subseteq\mathscr{M}_n$ 가 존재하여 $B\subseteq A\subseteq C$ 이고 $\lambda_n(C\setminus B)=(\lambda_l\otimes\lambda_m)(C\setminus B)=0$ 이다. 그렇다면 $A\setminus B\subseteq C\setminus B$ 는 λ_n–영집합이고 정리 2.68로부터 $\mathscr{M}_n$ 이 완비성을 가지므로 $A\setminus B\in\mathscr{M}_n$ 이 되어 $A=B\sqcup(A\setminus B)\in\mathscr{M}_n$ 이다. 그렇다면 $\mathscr{M}_n\supseteq\overline{\mathscr{M}_l\otimes\mathscr{M}_m}$ 에서 ii의 증명도 끝난다. □

이로부터 우리는 n차원 Lebesgue 측도공간을 재귀적으로 구성할 수 있다. 먼저 첫 번째 Lebesgue 측도의 구성을 따라 $(\mathbb{R},\mathscr{M}_1,\lambda_1)$ 을 구성하고, 이로써 2차원 Lebesgue 측도공간을 $(\mathbb{R}^2,\mathscr{M}_2,\lambda_2):=(\mathbb{R}\times\mathbb{R},\overline{\mathscr{M}_1\otimes\mathscr{M}_1},\overline{\lambda_1\otimes\lambda_1})$ 와 같이 정의한다. 다음으로 3차원 Lebesgue 측도공간을 $(\mathbb{R}^3,\mathscr{M}_3,\lambda_3):=(\mathbb{R}^2\times\mathbb{R},\overline{\mathscr{M}_2\otimes\mathscr{M}_1},\overline{\lambda_2\otimes\lambda_1})$ 로 정의하고, 이를 충분히 반복하면 원하는 모든 $n\in\mathbb{N}$ 에 대해 n차원 Lebesgue 측도공간을 구성할 수 있다. 한편, n차원 Borel 측도공간은 Lebesgue 측도공간의 구성과 달리 완비화가 필요하지 않으므로 $(\mathbb{R}^n,\mathscr{B}_1^{\otimes n},\mu_1^{\otimes n})$ 으로 더 깔끔하게 정의할 수 있다.

본 장에서의 우리의 이야기는 여기서 끝을 맺는다. 이어지는 절에서는 이후의 논의에서 필요하다고 생각되는 흥미로운 주제들에 대해 알아보도록 하겠다.

2.9 Riemann vs. Lebesgue

앞서 살펴본 바와 같이 Lebesgue 적분은 Riemann 적분이 극한과 매끄럽게 상호작용하지 못한다는 단점을 잘 보완해준다. 그런데, 우리는 여기서 과연 Lebesgue 적분이 정말 Riemann 적분의 일반화인지에 대한 질문을 조용히 숨겨두었다. 만약 우리가 그토록 공을 들여 전개한 Lebesgue 적분론이 Riemann 적분론과 다른 별개의 이론이라면 Riemann

적분론에서 성립하는 FTC와 같은 수많은 이론들을 Lebesgue 적분론에서는 사용할 수가 없을 것이다. 이건 빈대 잡자고 초가삼간 다 태우는 격이 아닌가! 당장 FTC가 없으면 아주 단순한 경우가 아니고서야 적분값을 계산할 수조차 없다. 이번 절에서는 우리가 회피했던 이 질문을 다루어보고자 한다. 한편, 다양한 문헌에서 서로 동등하지만 조금씩 다른 접근법을 통해 Riemann 적분을 정의하므로 일관된 논의를 위해 이번 절에서 사용할 Riemann 적분의 정의를 중간중간 옮겨 두었다. 이하의 정의는 [1] 에서의 Riemann 적분의 정의를 측도론의 맥락에 맞도록 조금 수정한 것으로 Darboux의 접근방식이라 불린다.

Definition 2.126 유계인 함수 $f:\mathbb{R}^n\to\mathbb{R}$와 유계인 집합 $A\in\mathscr{M}_n$에 대해 $A\subseteq B$인 임의의 유계인 닫힌 box $B\subseteq\mathbb{R}^n$를 택하여 $\mathscr{P}$를 B의 분할이라 하자. 이때, $\mathscr{P}$에 대한 f의 **하합 (lower sum)** 을 $L(f,\mathscr{P})$로 쓰고 $L(f,\mathscr{P})=\sum_{P\in\mathscr{P}}\inf_{x\in P}f\mathbf{1}_A\mathrm{vol}(P)$로 정의한다. (여기서 $\mathrm{vol}(\cdot)$은 $\mathbb{R}^n$의 유계인 box의 '부피' 로서 각 모서리의 길이의 곱으로 정의된다.) 비슷하게, $\mathscr{P}$에 대한 f의 **상합 (upper sum)** 을 $U(f,\mathscr{P})$로 쓰고 $U(f,\mathscr{P})=\sum_{P\in\mathscr{P}}\sup_{x\in P}f\mathbf{1}_A\mathrm{vol}(P)$로 정의한다. 이제 f의 **A에서의 하적분 (lower integral over A)** 을 $\underline{\int_A}f(x)\,dx$ 혹은 간단히 $\underline{\int_A}f$로 쓰고, $\underline{\int_A}f=\sup_{\mathscr{P}\text{ partition}}L(f,\mathscr{P})$로 정의한다. 비슷하게, f의 **A에서의 상적분 (upper integral over A)** 을 $\overline{\int_A}f(x)\,dx$ 혹은 간단히 $\overline{\int_A}f$로 쓰고 $\overline{\int_A}f=\inf_{\mathscr{P}\text{ partition}}U(f,\mathscr{P})$로 정의하여 만약 $\underline{\int_A}f=\overline{\int_A}f$이면 이때 f가 **A 위에서 (Riemann) 적분가능 (− integrable over A)** 하다고 하고, 그 공통의 값을 $\int_A f(x)\,dx$ 혹은 간단히 $\int_A f$로 쓴다.

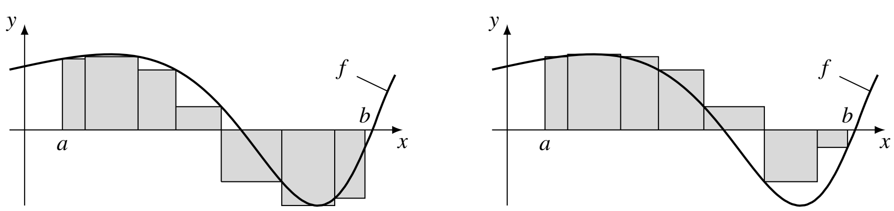


Figure 2.10 구간 $[a,b]$에서의 분할 $\mathscr{P}$에 대한 함수 f의 하합 $L(f,\mathscr{P})$ (오른쪽)와 상합 $U(f,\mathscr{P})$ (왼쪽).

Lemma 2.127 임의의 $i\le n$와 임의의 $x_0\in\mathbb{R}$에 대해 $\mathbb{R}^n$에 속하는 초평면 $\pi:x_i=x_0$는 영집합이다.

Proof 우선 π가 닫힌집합이므로 Borel이고, 따라서 이는 가측이다. 한편, Lebesgue 측도가 이동 불변성을 가지므로 $x_i=0$인 경우에 대해서만 증명하는 것으로 충분하고, 간결한 논의를 위해 $i=1$인 경우에 대해서만 보이도록 하자. (다른 경우에도 이와 비슷하게 보일 수 있다.) 이제 임의의 $\varepsilon>0$을 택하고 집합열 $\{A_i\}$를 $A_i=(-\varepsilon/2^{i+n}i^{n-1},\varepsilon/2^{i+n}i^{n-1}]\times(-i,i]^{n-1}$로 두면 이는 $\mathscr{S}_n$에 속하는 집합열로서 각 $i\in\mathbb{N}$에 대해 $\lambda_n(A_i)=\varepsilon/2^i$이고 $\pi\subseteq\bigcup_{i=1}^\infty A_i$이다. 이로부터 $\lambda_n(\pi)\le\lambda_n(\bigcup_{i=1}^\infty A_i)\le\sum_{i=1}^\infty\lambda_n(A_i)=\varepsilon$이 되어 보조정리가 성립한다. □

Theorem 2.128 유계인 함수 $f : \mathbb{R}^n \to \mathbb{R}$와 유계인 집합 $A \in \mathscr{M}_n$에 대해 만약 f가 A 위에서 Riemann 적분가능하다면 $f\mathbf{1}_A$는 Lebesgue 적분가능하며, 이때의 두 적분은 서로 일치한다. 즉, $\int_A f = \int_{\mathbb{R}^n} f\mathbf{1}_A \, d\lambda_n$이다.

Proof 먼저 A가 닫힌 box이고 f가 음이 아닌 특별한 경우를 생각하여 A의 임의의 분할 $\mathscr{P}$에 대해 단순함수 $\varphi_{\mathscr{P}}, \psi_{\mathscr{P}} : \mathbb{R}^n \to \mathbb{R}_0^+$를 각각 $\varphi_{\mathscr{P}} = \sum_{P \in \mathscr{P}} \inf_{x \in P} f(x)\mathbf{1}_{\widetilde{P}}$, $\psi_{\mathscr{P}} = \sum_{P \in \mathscr{P}} \sup_{x \in P} f(x)\mathbf{1}_{\widetilde{P}}$로 두자. (여기서 $\widetilde{P}$는 box P의 남서쪽 면을 빼서 얻는 semi-open box로서 적당한 $x, y \in \mathbb{R}^n$에 대해 $P = \prod_{i=1}^n [x_i, y_i]$라면 $\widetilde{P} = \prod_{i=1}^n (x_i, y_i]$이다.) 그렇다면

$$\begin{aligned}
\int_{\mathbb{R}^n} \varphi_{\mathscr{P}} \, d\lambda_n &= \sum_{P \in \mathscr{P}} \inf_{x \in P} f(x) \int_{\mathbb{R}^n} \mathbf{1}_{\widetilde{P}} \, d\lambda_n \\
&= \sum_{P \in \mathscr{P}} \inf_{x \in P} f(x) \lambda_n(\widetilde{P}) \\
&= \sum_{P \in \mathscr{P}} \inf_{x \in P} f(x) \mathrm{vol}(P) \\
&= L(f, \mathscr{P})
\end{aligned}$$

이고 비슷하게 $\int_{\mathbb{R}^n} \psi_{\mathscr{P}} \, d\lambda_n = U(f, \mathscr{P})$이다. 이제 f가 A위에서 Riemann 적분가능하므로 임의의 $i \in \mathbb{N}$에 대해 적당한 A의 분할 $\mathscr{P}_i$가 존재하여 $U(f, \mathscr{P}_i) - L(f, \mathscr{P}_i) < 1/i$이고, 이들을 모아 분할의 열 $\{\mathscr{P}_i\}$를 구성할 수 있고, WLOG, 필요하다면 각 항을 $\mathscr{Q}_i := \bigcup_{j=1}^i \mathscr{P}_j$로 바꾸어 $\{\mathscr{P}_i\}$가 처음부터 증가했다고 해도 된다. 그렇다면 $\{\varphi_{\mathscr{P}_i}\}, \{\psi_{\mathscr{P}_i}\}$는 각각 증가하고 감소하는 함수열로 $\{\psi_{\mathscr{P}_i} - \varphi_{\mathscr{P}_i}\}$가 감소하는 음이 아닌 함수열이 되어 이가 모든 점에서 수렴하고, 곧 Fatou의 보조정리로부터

$$\begin{aligned}
\int_{\mathbb{R}^n} \lim_{i \to \infty} (\psi_{\mathscr{P}_i} - \varphi_{\mathscr{P}_i}) \, d\lambda_n &\le \liminf_{i \to \infty} \int_{\mathbb{R}^n} \psi_{\mathscr{P}_i} - \varphi_{\mathscr{P}_i} \, d\lambda_n \\
&= \lim_{i \to \infty} \left(\int_{\mathbb{R}^n} \psi_{\mathscr{P}_i} \, d\lambda_n - \int_{\mathbb{R}^n} \varphi_{\mathscr{P}_i} \, d\lambda_n \right) \\
&= \lim_{i \to \infty} [U(f, \mathscr{P}_i) - L(f, \mathscr{P}_i)] \\
&= 0
\end{aligned}$$

이다. 이는 정리 2.101의 ii로부터 $\psi_{\mathscr{P}_i} - \varphi_{\mathscr{P}_i} \downarrow 0$ (ae.) 임을 뜻하고, 특히 위의 보조정리로부터 모든 $i \in \mathbb{N}$에 대해 $\varphi_{\mathscr{P}_i} \le f\mathbf{1}_A \le \psi_{\mathscr{P}_i}$ (ae.)가 성립하므로[20] $\varphi_{\mathscr{P}_i} \uparrow f\mathbf{1}_A$ (ae.)가 되어 정리 2.81로부터 $f\mathbf{1}_A$가 가측임을 안다. 이제 MCT로부터 $\int_{\mathbb{R}^n} f\mathbf{1}_A \, d\lambda_n = \lim_{i \to \infty} \int_{\mathbb{R}^n} \varphi_{\mathscr{P}_i} \, d\lambda_n = \lim_{i \to \infty} U(f, \mathscr{P}_i) = \int_A f < \infty$이므로 $f\mathbf{1}_A$는 Lebesgue 적분가능하고 두 적분이 일치한다.

한편, 일반적인 경우에 대해서는 $A \subseteq B \in \mathbb{R}^n$인 임의의 닫힌 box B를 잡으면 이전의 결과로부터

$$\begin{aligned}\int_{\mathbb{R}^n} f\mathbf{1}_A\,d\lambda_n &= \int_{\mathbb{R}^n}(f\mathbf{1}_A)_+\,d\lambda_n - \int_{\mathbb{R}^n}(f\mathbf{1}_A)_-\,d\lambda_n\\ &= \int_{\mathbb{R}^n} f_+\mathbf{1}_A\mathbf{1}_B\,d\lambda_n - \int_{\mathbb{R}^n} f_-\mathbf{1}_A\mathbf{1}_B\,d\lambda_n\\ &= \int_B f_+\mathbf{1}_A - \int_B f_-\mathbf{1}_A\\ &= \int_A f_+ - \int_A f_-\\ &= \int_A f\end{aligned}$$

이고, 곧 일반적인 경우에도 $f\mathbf{1}_A$ 는 Lebesgue 적분가능하며 두 적분이 일치한다. □

Definition 2.129 유계인 음이 아닌 함수 $f:\mathbb{R}^n \to \mathbb{R}_0^+$ 와 집합 $A \in \mathscr{M}_n$ 에 대해 $f\mathbf{1}_A$ 가 적당한 $x>0$에 대해 $[-x,x]^n$ 으로 표현되는 모든 닫힌 box 위에서 Riemann 적분가능하다고 하자. 만약 $\lim_{x\to\infty}\int_{[-x,x]^n} f\mathbf{1}_A < \infty$ 이면 이때 f 가 A **위에서 (Riemann) 적분가능 (− integrable over** A)하다고 하고, 그 극한을 $\int_A f(x)\,dx$ 혹은 간단히 $\int_A f$ 로 쓴다. 나아가, 함수 $f:\mathbb{R}^n\to\mathbb{R}$ 에 대해서는 $f_\pm$ 이 모두 A 위에서 Riemann 적분가능하다면 이때 f 가 A **위에서 (Riemann) 적분가능 (− integrable over** A)하다고 하고, $\int_A f_+ - \int_A f_-$ 의 값을 $\int_A f(x)\,dx$ 혹은 간단히 $\int_A f$ 로 쓴다.

Theorem 2.130 유계인 함수 $f:\mathbb{R}^n\to\mathbb{R}$ 와 집합 $A\in\mathscr{M}_n$ 에 대해 만약 f 가 A 위에서 Riemann 적분가능하다면 $f\mathbf{1}_A$ 는 Lebesgue 적분가능하며, 이때의 두 적분은 서로 일치한다. 즉, $\int_A f = \int_{\mathbb{R}^n} f\mathbf{1}_A\,d\mu$ 이다.

Proof 함수 $f_i:\mathbb{R}^n\to\mathbb{R}$ 의 열 $\{f_i\}$ 를 $f_i = f\mathbf{1}_{A\cap[-i,i]^n}$ 으로 정의하면 이는 유계함수열이고 $f_i \to f\mathbf{1}_A$ 인데 각 $i\in\mathbb{N}$ 에 대해 f_i 가 $[-i,i]^n$ 위에서 Riemann 적분가능하므로 정리 2. 128 로부터 가측이고, 곧 정리 2. 81로부터 $f\mathbf{1}_A$ 도 가측이다. 이제 f 가 음이 아닌 특별한 경우를 생각하면 WLOG, $\{f_i\}$ 가 증가한다고 할 수 있다. 그렇다면 MCT와 정리 2. 128로부터 $\int_{\mathbb{R}^n} f\mathbf{1}_A\,d\lambda_n = \lim_{i\to\infty}\int_{\mathbb{R}^n} f_i\,d\lambda_n = \lim_{i\to\infty}\int_{[-i,i]^n} f\mathbf{1}_A = \int_A f < \infty$ 가 되어 $f\mathbf{1}_A$ 는 Lebesgue 적분가능하고 두 적분이 일치한다. 한편, 일반적인 f 에 대해서는

$$\begin{aligned}\int_{\mathbb{R}^n} f\mathbf{1}_A\,d\lambda_n &= \int_{\mathbb{R}^n}(f\mathbf{1}_A)_+\,d\lambda_n - \int_{\mathbb{R}^n}(f\mathbf{1}_A)_-\,d\lambda_n\\ &= \int_{\mathbb{R}^n} f_+\mathbf{1}_A\,d\lambda_n - \int_{\mathbb{R}^n} f_-\mathbf{1}_A\,d\lambda_n\\ &= \int_A f_+ - \int_A f_-\\ &= \int_A f\end{aligned}$$

이므로 이때에도 같은 결론이 성립한다. □

Definition 2.131 음이 아닌 함수 $f:\mathbb{R}^n \to \mathbb{R}_0^+$ 와 집합 $A \in \mathscr{M}_n$, 충분히 큰 $M \in \mathbb{R}$에 대해 $f_{(M)} := \min\{f, M\}$ 이 A 위에서 Riemann 적분가능하다고 하자. 만약 $\lim_{M\to\infty}\int_A f_{(M)} < \infty$ 이면 이때 f가 A **위에서 (Riemann) 적분가능 (− integrable over** A)하다고 하고, 그 극한을 $\int_A f(x)\,dx$ 혹은 간단히 $\int_A f$ 로 쓴다. 나아가, 유계인 함수 $f:\mathbb{R}^n \to \mathbb{R}$에 대해서는 $f_\pm$ 이 모두 A 위에서 Riemann 적분가능하다면 이때 f가 A **위에서 (Riemann) 적분가능 (− integrable over** A)하다고 하고, $\int_A f_+ - \int_A f_-$ 의 값을 $\int_A f(x)\,dx$ 혹은 간단히 $\int_A f$ 로 쓴다.

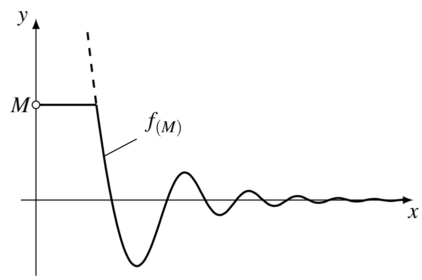


Figure 2.11 함수 f (파선)에 대한 $f_{(M)}$ 의 그래프. 여기서 f는 $x\downarrow 0$이면 ∞로 발산한다.

Theorem 2.132 함수 $f:\mathbb{R}^n \to \mathbb{R}$와 집합 $A \in \mathscr{M}_n$에 대해 만약 f가 A 위에서 Riemann 적분가능하다면 $f\mathbf{1}_A$ 는 Lebesgue 적분가능하며, 이때의 두 적분은 서로 일치한다. 즉, $\int_A f = \int_{\mathbb{R}^n} f\mathbf{1}_A\,d\mu$ 이다.

PROOF 함수 $f_i:\mathbb{R}^n \to \mathbb{R}$의 열 $\{f_i\}$를 $f_i = f_{(i)}\mathbf{1}_A$ 로 정의하면 이는 증가하는 유계함수열이고 $f_i \uparrow f\mathbf{1}_A$ 인데 충분히 큰 $i \in \mathbb{N}$에 대해 f_i가 A 위에서 Riemann 적분가능하므로 정리 2. 130로부터 가측이고, 곧 정리 2. 81로부터 $f\mathbf{1}_A$ 도 가측이다. 여기서 WLOG, f_i가 모든 $i \in \mathbb{N}$에 대해 Riemann 적분가능하다고 하자. 이제 f가 음이 아닌 특별한 경우를 생각하면 MCT와 정리 2. 130로부터 $\int_{\mathbb{R}^n} f\mathbf{1}_A\,d\lambda_n = \lim_{i\to\infty}\int_{\mathbb{R}^n} f_i\,d\lambda_n = \lim_{i\to\infty}\int_A f_{(i)} = \int_A f < \infty$가 되어 $f\mathbf{1}_A$ 는 Lebesgue 적분가능하고 두 적분이 일치한다. 한편, 일반적인 f에 대해서는

$$\begin{aligned}\int_{\mathbb{R}^n} f\mathbf{1}_A\,d\lambda_n &= \int_{\mathbb{R}^n}(f\mathbf{1}_A)_+\,d\lambda_n - \int_{\mathbb{R}^n}(f\mathbf{1}_A)_-\,d\lambda_n \\ &= \int_{\mathbb{R}^n} f_+\mathbf{1}_A\,d\lambda_n - \int_{\mathbb{R}^n} f_-\mathbf{1}_A\,d\lambda_n \\ &= \int_A f_+ - \int_A f_- \\ &= \int_A f\end{aligned}$$

이므로 이때에도 같은 결론이 성립한다. □

따라서 다행히도 Lebesgue 적분은 Riemann 적분을 포괄하는 개념으로, 이론 전개를 할 때는 Lebesgue 적분으로 생각하여 이론을 전개하고, 실제로 계산을 해야 할 때에는 이를

Riemann 적분으로 생각하여 FTC 등을 사용해 계산해도 대부분의 경우 문제가 없다. 다만, 여기서도 조심해야 할 부분이 있는데, 바로 improper한 Riemann 적분은 이가 절대수렴하는 경우에만 Lebesgue 적분과 그 값이 일치한다는 것이다. Riemann 적분이 조건수렴에 그치는 $\int_0^\infty (\sin x)/x\,dx$와 같은 경우 이의 Lebesgue 적분가능성은 일반적으로 보장되지 않는다. 물론, 이 책에서는 조건수렴에 그치는 Riemann 적분을 다룰 일이 (아마) 없을 것이므로 안심해도 좋다.

2.10 L^p Space and Radon-Nikodým Theorem

지금까지 우리는 측도라는 개념으로 넓이나 부피 따위를 일반화하였고, 이를 이용하여 Lebesgue 적분의 개념을 정립함으로써 $\mathbb{R}^n$에서 뿐만 아니라 측도가 주어진 임의의 공간에서 적분을 할 수 있는 이론적 토대를 마련하였다. 이번 절에서는 이런 추상화의 흐름에 발맞추어 적분의 역연산인 미분의 개념을 임의의 공간으로 확장하려고 한다. 한편, 이를 위해서는 L^p 공간이라는 경유지가 필요한데, L^p 공간은 그 자체로도 꽤나 흥미로운 탐구의 대상이어서 파고들자면 끝도 없으므로 여기서는 우리에게 필요한 정도만 선택적으로 그 내용을 소개하도록 하겠다. 이에 대해 조금 더 관심이 있는 독자들은 [2]와 같은 실해석학 교재를 참조하기 바란다.

Definition 2.133 측도공간 $(X, \mathscr{A}, \mu)$에서 정의된 가측함수 $f: X \to \overline{\mathbb{R}}$와 $p \geq 1$에 대해 $(\int_X |f|^p\,d\mu)^{1/p}$를 f의 **p−노름 (p−norm)**이라 하고 $||f||_{p,\mu}$ 혹은 간단히 $||f||_p$로 쓴다.

Proposition 2.134 측도공간 $(X, \mathscr{A}, \mu)$에서 정의된 가측함수 $f: X \to \overline{\mathbb{R}}$와 $p \geq 1$에 대해 다음이 성립한다.

i. 임의의 $c \in \mathbb{R}$에 대해 $||cf||_p = |c|||f||_p$이다.
ii. $||f||_p = 0$일 필요충분조건은 $f = 0$ (ae.) 이다.

Proof i. 이는 $||cf||_p = (\int_X |cf|^p\,d\mu)^{1/p} = |c|(\int_X |f|^p\,d\mu)^{1/p} = |c|||f||_p$에서 자명하다.
ii. $\int_X |f|^p\,d\mu = ||f||_p^p = 0$이므로 정리 2. 101의 ii로부터 $|f| = 0$ (ae.) 이고, 곧 $f = 0$ (ae.) 이다. 한편, 역은 자명하다. □

Theorem 2.135 (Hölder's inequality) 측도공간 $(X, \mathscr{A}, \mu)$에서 정의된 가측함수 $f, g: X \to \overline{\mathbb{R}}$와 $1/p + 1/q = 1$인 $p, q > 1$에 대해 $\int_X |fg|\,d\mu \leq ||f||_p||g||_q$이다. 이때, $||f||_p, ||g||_q < \infty$라면 등호가 성립할 필요충분조건은 적당한 $c \geq 0$가 존재하여 $|f|^p = c|g|^q$ (ae.) 이거나 $|g|^q = c|f|^p$ (ae.) 인 것이다.

PROOF WLOG, 필요하다면 f,g를 각각 $|f|,|g|$로 대체하여 f,g가 처음부터 음이 아니라 해도 된다. 한편, 만약 $||f||_p=0$이면 명제 2.134의 ii로부터 $f=0$ (ae.)이고, 곧 $\int_X|fg|\,d\mu=0$이 되어 부등식이 자명하므로 $||f||_p>0$이라 하고, 비슷한 이유로 $||g||_p>0$이라 하자. 또한 $||f||_p=\infty$이거나 $||g||_p=\infty$인 경우에도 부등식이 자명하므로 $||f||_p,||g||_p<\infty$라 하자. 이제 $||f||_p=||g||_q=1$인 특별한 경우를 생각하면 Young의 부등식으로부터 $\int_X fg\,d\mu\le\int_X f^p/p+g^q/q\,d\mu=||f||_p^p/p+||g||_q^q/q=1/p+1/q=1$이다. 이제 일반적인 경우에 $\widetilde{f}=f/||f||_p,\widetilde{g}=g/||g||_q$라 하면 명제 2.134의 i로부터 $||\widetilde{f}||_p=||\widetilde{g}||_q=1$이므로 앞선 결과로부터 $\int_X|\widetilde{f}\widetilde{g}|\,d\mu\le1$이 되어 곧 $\int_X|fg|\,d\mu\le||f||_p||g||_q$임을 안다.

한편, 등호조건을 보이기 위해 $||f||_p,||g||_q<\infty$이고 $\int_X|fg|\,d\mu=||f||_p||g||_q$라 하자. 만약 $||f||_p=0$이면 $f=0$ (ae.)에서 $|f|^p=0=0\cdot|g|^q$ (ae.)가 되어 더 이상 보일 것이 없으므로 $||f||_p\neq0$이라 하고, 비슷한 이유에서 $||g||_q\neq0$이라 하자. 이제 $\widetilde{f}=|f|/||f||_p,\widetilde{g}=|g|/||g||_q$라 하면 $\int_X\widetilde{f}\widetilde{g}\,d\mu=(\int_X|fg|\,d\mu)/||f||_p||g||_q=1=||\widetilde{f}||_p^p/p+||\widetilde{g}||_q^q/q=\int_X\widetilde{f}^p/p+\widetilde{g}^q/q\,d\mu$이므로 Young의 부등식과 정리 2.101의 ii로부터 $\widetilde{f}\widetilde{g}=\widetilde{f}^p/p+\widetilde{g}^q/q$ (ae.)이다. 이는 다시 Young의 부등식의 등호조건으로부터 $\widetilde{f}^p=\widetilde{g}^q$ (ae.)를 뜻하고, 따라서 $|f|^p=(||f||_p^p/||g||_q^q)|g|^q$ (ae.)이다. 역으로 적당한 $c\ge0$가 존재하여 $|f|^p=c|g|^q$ (ae.) 이거나 $|g|^q=c|f|^p$ (ae.)인 경우에 등호가 성립함은 자명하므로 증명은 이로써 충분한다. □

Theorem 2.136 (Minkowski's inequality) 측도공간 $(X,\mathscr{A},\mu)$에서 정의된 가측함수 $f,g:X\to\overline{\mathbb{R}}$와 $p\ge1$에 대해 $||f+g||_p\le||f||_p+||g||_p$이다. 이때, $||f||_p,||g||_p<\infty$이면 등호가 성립한 필요충분조건은 적당한 $c\ge0$가 존재하여 $|f|^p=c|g|^p$ (ae.)이거나 $|g|^p=c|f|^p$ (ae.)인 것이다.

PROOF 만약 $p=1$이면 $||f+g||_1=\int_X|f+g|\,d\mu\le\int_X|f|\,d\mu+\int_X|g|\,d\mu=||f||_1+||g||_1$에서 부등식이 자명하므로 $p>1$이라 하자. 또한, 만약 $||f||_p=\infty$이거나 $||g||_p=\infty$이거나 $||f+g||_p=0$인 경우에도 부등식이 자명하므로 $||f||_p,||g||_p<\infty$이고 $||f+g||_p>0$이라 하자. 그렇다면 $|f|,|g|\le(|f|^p+|g|^p)^{1/p}$이므로 $|f+g|^p\le(|f|+|g|)^p\le2^p(|f|^p+|g|^p)$에서 $\int_X|f+g|^p\,d\mu\le2^p\int_X|f|^p+|g|^p\,d\mu=2^p(||f||_p^p+||g||_p^p)<\infty$이고, 곧 $||f+g||_p<\infty$이다. 한편, Hölder의 부등식으로부터 $1/p+1/q=1$인 $q>1$에 대해

$$\begin{aligned}
||f+g||_p^p&=\int_X|f+g|^p\,d\mu\\
&\le\int_X|f||f+g|^{p-1}\,d\mu+\int_X|g||f+g|^{p-1}\,d\mu\\
&\le||f||_p|||f+g|^{p-1}||_q+||g||_p|||f+g|^{p-1}||_q\\
&=(||f||_p+||g||_p)\left(\int_X|f+g|^{(p-1)q}\,d\mu\right)^{1/q}\\
&=(||f||_p+||g||_p)\left(\int_X|f+g|^p\,d\mu\right)^{(p-1)/p}
\end{aligned}$$

$$= (||f||_p + ||g||_p)||f+g||_p^{p-1}$$

이므로 $||f+g||_p \le ||f||_p + ||g||_p$ 임을 안다.

한편, 등호조건을 보이기 위해 $||f||_p, ||g||_p < \infty$ 이고 $||f+g||_p = ||f||_p + ||g||_p$ 라 하자. 만약 $f = 0$ (ae.) 이면 $|f|^p = 0 = 0 \cdot |g|^p$ (ae.) 에서 더 이상 보일 것이 없으므로 $f = 0$ (ae.) 이 아니라 하고, 비슷한 이유에서 $g = 0$ (ae.) 도 아니라 하자. 그렇다면 위의 증명과정과 Hölder의 부등식의 등호조건으로부터 적당한 $c_1, c_2 > 0$ 가 존재하여 $|f|^p = c_1|f+g|^{(p-1)q} = c_1|f+g|^p$ (ae.) 이고 $|g|^p = c_2|f+g|^{(p-1)q} = c_2|f+g|^p$ (ae.) 이므로 곧 $|f|^p = (c_1/c_2)c_2|f+g|^p = (c_1/c_2)|g|^p$ (ae.) 이다. 역으로 적당한 $c \ge 0$ 가 존재하여 $|f|^p = c|g|^p$ (ae.) 이거나 $|g|^p = c|f|^p$ (ae.) 인 경우에 등호가 성립함은 자명하므로 증명은 이로써 충분하다. □

당장은 p−노름은 노름이 아니라는 점에 주의하자. 명제 2. 134의 i과 Minkowski의 부등식으로부터 p−노름이 양의 동차성과 삼각부등식은 만족하지만, 아직 이가 양의 정부호성을 만족하는지는 알지 못한다. 그리고 조금만 생각해보면, p−노름이 양의 정부호성을 갖길 기대하는 건 부질없는 것임을 깨달을 수 있다. 적분값이 0이라고 해서 피적분함수가 반드시 0일 필요는 없기 때문이다. 그럼에도 이에 p−노름이라는 이름을 괜히 붙인 것은 아닐테니, p−노름을 정말 노름으로 만들기 위해서는 '도약' 이 필요하다.

Definition 2.137 측도공간 $(X, \mathscr{A}, \mu)$ 에서 정의된 가측함수 $f : X \to \overline{\mathbb{R}}$ 와 $p \ge 1$ 에 대해 만약 $||f||_p < \infty$ 이면 이를 **$L^p(\mu)$−함수 ($L^p(\mu)$−function)** 혹은 간단히 **L^p−함수 (L^p−function)** 라 한다. 나아가, X 에서 정의된 모든 L^p−함수의 모임을 $\mathscr{L}^p(\mu)$ 혹은 간단히 $\mathscr{L}^p$ 로 쓴다.

Definition 2.138 측도공간 $(X, \mathscr{A}, \mu)$ 와 $p \ge 1$ 에 대해 $\mathscr{L}^p(\mu)$ 위의 동치관계 $\overset{\mu}{\sim}$ 를 $f \overset{\mu}{\sim} g :\Leftrightarrow f = g$ (ae.) 로 정의하고 이의 몫공간 $\mathscr{L}^p(\mu)/\overset{\mu}{\sim}$ 를 **Lebesgue 공간 (− space)**, **$L^p(\mu)$−공간 ($L^p(\mu)$−space)** 혹은 간단히 **L^p−공간 (L^p−space)** 이라 하고 $L^p(\mu)$ 혹은 간단히 L^p 로 쓴다.

언듯 보아서는 뭔가 이상한 짓을 한 것 같지만 이러한 L^p 공간의 구성은 굉장히 직관적인 구성이다. 앞서 보았듯이 p−노름이 가지는 문제점은 양의 정부호성이 성립하지 않는다는 것이었는데, 명제 2. 134의 ii에 의하면 $||f||_p = 0$ 이면 f 가 거의 어디서나 0인 것까지는 보장이 되었다. 우리는 이에 착안하여 거의 어디서나 같은 것을 정말 같은 것으로 볼 수 있는 공간을 구성한 것이다. 이제 우리의 도약은 L^p 공간에 적당한 노름공간의 구조를 부여함으로써 마무리된다.

Proposition 2.139 측도공간 $(X, \mathscr{A}, \mu)$ 와 $p \ge 1$ 에 대해 함수 $+ : (L^p)^2 \to L^p, \cdot : \mathbb{R} \times L^p \to L^p$ 를 각각 $+ : ([f], [g]) \mapsto [f+g], \cdot : (\lambda, [f]) \mapsto [\lambda f]$ 로 정의하면 $(L^p, +, \cdot, [0], 1)$ 은 벡터공간이다.

Proof 정리에서 정의된 덧셈과 스칼라 곱이 well-define된다는 것만 보이면 $(L^p, +, \cdot, [0], 1)$ 이 벡터공간이 되기 위한 조건은 쉽게 확인할 수 있으므로 여기서는 well-definedness만 보이도록 하자. 이를 위해 임의의 $f, g \in \mathscr{L}^p$ 와 임의의 $\lambda \in \mathbb{R}$ 를 생각하면 Minkowski의 부등식으로부터 $||f+g||_p \leq ||f||_p + ||g||_p < \infty$ 이고 명제 2. 134의 i로부터 $||\lambda f||_p = |\lambda|||f||_p < \infty$ 이므로 $f+g, \lambda f \in \mathscr{L}^p$ 이다. 나아가, 주어진 덧셈과 스칼라 곱의 결과가 representative의 선택에 의존하지 않음이 자명하므로 이는 well-define된다. □

Proposition 2.140 측도공간 $(X, \mathscr{A}, \mu)$ 와 $p \geq 1$ 에 대해 함수 $||\cdot||_p : L^p \to \mathbb{R}_0^+$ 를 $||\cdot||_p : [f] \mapsto ||f||_p$ 로 정의하면 $(L^p, ||\cdot||_p)$ 는 노름공간이다.

Proof 우선 따름정리 2. 102의 ii로부터 $||\cdot||_p$ 는 representative의 선택에 의존하지 않으므로 well-define된다. 나아가 명제 2. 134의 i과 Minkowski의 부등식으로부터 양의 동차성과 삼각부등식이 성립하고, 명제 2. 134의 ii와 L^p 공간의 구성으로부터 이가 양의 정부호성도 가지므로 $(L^p, ||\cdot||_p)$ 은 노름공간이다. □

보통 특별히 정하지 않는 이상 L^p 공간에는 $p-$노름이 주어진 것으로 생각하고, 이는 이 책에서도 마찬가지이다. 또한, 일반적으로 노름공간 $(L^p, ||\cdot||_p)$ 의 원소 $[f]$ 를 표기를 남용하여 그냥 f 로 쓴다. 이렇게 집합인 동치류를 마치 함수처럼 표기하는 것에 처음에는 거부감을 느낄 수 있다. 하지만 L^p 에서 진행되는 논의는 대부분 적분에 관한 논의이고, 곧 이러한 맥락에서는 함수의 특정 점에서의 값보다는 그 함수의 적분값이 중요한데, 거의 어디서나 같은 함수는 적어도 적분에서는 구별할 필요가 없으므로 이러한 표기법은 굉장히 편리하다. 또한, L^p 공간의 기저에 깔려있는 거의 어디서나 같은 것은 정말 같다는 그 철학을 생각한다면 오히려 이러한 표기가 자연스러운 표기법이라고 생각할 수도 있다.

다만, 이러한 표기의 남용으로 인하여 L^p 에 속하는 동치류로서 $[f]$ 를 뜻하는 f 와 $[f]$ 의 임의의 representative로서 $\mathscr{L}^p$ 에 속하는 구체적인 함수를 뜻하는 f 가 적어도 표기로써는 구별되지 않으므로 이를 논의의 맥락으로써 잘 구별해야 한다. 대부분의 경우, 이러한 구별은 그다지 어렵지 않다. 한 예시로 아래의 따름정리 2. 143의 'L^p 에 속하는 함수열 $\{f_i\}$ 가 적당한 $f \in L^p$ 에 대해 $f_i \to f$ in sense of L^p 이면 $\{f_i\}$ 의 적당한 부분함수열이 존재하여 거의 어디서나 f 로 수렴한다.' 는 표현을 보자. 먼저 'L^p 에 속하는 함수열 $\{f_i\}$' 에서의 f_i 는 L^p 에 속하는 동치류로 $[f_i]$ 를 이르는 것이고, 이를 표기를 남용하여 f_i 라 쓰고 그에 걸맞게 '함수열' 이라 표현하였다. 이어지는 '적당한 $f \in L^p$ 에 대해 $f_i \to f$ in sense of L^p' 또한 동치류의 열 $\{[f_i]\}$ 가 $p-$노름으로써 동치류 $[f] \in L^p$ 로 수렴한다는 뜻으로, 'in sense of L^p' 라는 표현에서 그 의미가 분명하다. 한편, 끝부분의 '$\{f_i\}$ 의 적당한 부분함수열이 존재하여 거의 어디서나 f 로 수렴한다' 는 $[f_i]$ 의 임의의 representative를 택하여 구성한 $\mathscr{L}^p$ 에 속하는 (진짜) 함수열 $\{f_i\}$ 에 대해 적당한 부분함수열 $\{f_{i_j}\}$ 가 존재하여 $[f]$ 의 임의의 representative f 로 거의 어디서나 점별수렴함을 뜻하는 것으로, 이러한 거의 어디서나 점별수렴한다는 결론이 $[f_i]$ 나 $[f]$ 의 representative의 선택에 의존하지 아니하므로 그 의미의

전달이 분명하다. 이와 같이 맥락에 따라 그 의미를 파악하며 읽는 것은, 처음에는 다소 시간이 걸리겠지만, 머지않아 익숙해 질 것이다.

우리는 $p-$노름이 정말 노름이 되도록 하기 위해 L^p 공간을 도입했지만, L^p 공간은 우리가 생각하는 것보다 훨씬 더 아름다운 공간이다. 지금부터는 L^p 공간이 함수공간으로서 가지는 성질을 알아보자.

Definition 2.141 노름공간 $(\mathsf{V}, ||\cdot||)$에 대해 $||\cdot||$이 완비성을 가지면, 즉 임의의 Cauchy 수열이 수렴하면, 이때의 $(\mathsf{V}, ||\cdot||)$를 **Banach 공간 (− space)**이라 한다.

Theorem 2.142 측도공간 $(X, \mathscr{A}, \mu)$와 $p \geq 1$에 대해 $(L^p, ||\cdot||_p)$는 Banach 공간이다.

Proof L^p 공간에 속하는 임의의 Cauchy 수열 $\{f_i\}$를 택하면 임의의 $j \in \mathbb{N}$에 대해 적당한 $i_j \in \mathbb{N}$가 존재하여 임의의 $i, i' \geq i_j$에 대해 $||f_i - f_{i'}||_p < 2^{-j}$이다. WLOG, 필요하다면 i_j를 더 크게 잡음으로써 $\{i_j\}_j$를 증가수열이라 해도 되고, 이로부터 부분열 $\{f_{i_j}\}_j$를 생각하면 이는 임의의 $j \in \mathbb{N}$에 대해 $||f_{i_j} - f_{i_{j+1}}||_p < 2^{-j}$인 함수열이다. 이제 함수 $g_j : X \to \overline{\mathbb{R}}$의 열 $\{g_j\}$를 $g_j = \sum_{k=1}^{j} |f_{i_{k+1}} - f_{i_k}|$로 정의하고 함수 $g : X \to \overline{\mathbb{R}}$를 $g = \sum_{k=1}^{\infty} |f_{i_{k+1}} - f_{i_k}| = \lim_{j\to\infty} g_j$로 두자. 그렇다면 임의의 $j \in \mathbb{N}$에 대해 Minkowski의 부등식으로부터 $||g_j||_p = ||\sum_{k=1}^{j} |f_{i_{k+1}} - f_{i_k}|||_p \leq \sum_{k=1}^{j} ||f_{i_{k+1}} - f_{i_k}||_p < \sum_{k=1}^{\infty} 2^{-k} = 1$이다. 나아가 임의의 $j \in \mathbb{N}$에 대해 $g_j^p \leq g_{j+1}^p$이고 $g_j^p \to g^p$이므로 MCT로부터 $||g||_p = (\int_X |g|^p \, d\mu)^{1/p} = \lim_{j\to\infty} (\int_X |g_j|^p \, d\mu)^{1/p} = \lim_{j\to\infty} ||g_j||_p \leq 1$에서 $g \in L^p$이다. 이는 $|g|^p$가 적분가능함을 뜻하므로 정리 2. 100의 ii로부터 g는 거의 어디서나 유한하고, 곧 적당한 영집합 $N \subseteq X$에 대해 $X \setminus N$에서 $\sum_{j=1}^{\infty} (f_{i_{j+1}} - f_{i_j})$는 절대수렴한다. 이제 함수 $f : X \to \overline{\mathbb{R}}$를 $f = [f_{i_1} + \sum_{j=1}^{\infty} (f_{i_{j+1}} - f_{i_j})] \mathbf{1}_{X\setminus N}$이라 하면 $f_{i_j} \mathbf{1}_{X\setminus N} \to f$이므로 정리 2. 77의 ii로부터 f는 가측이다. 한편, 임의의 $\varepsilon > 0$을 택하면 적당한 $i_0 \in \mathbb{N}$가 존재하여 임의의 $i, i' \geq i_0$에 대해 $||f_i - f_{i'}||_p < \varepsilon/2$이므로 Fatou의 보조정리로부터 $i \geq i_0$이면

$$\begin{aligned}
\int_X |f_i - f|^p \, d\mu &= \int_X \lim_{j\to\infty} |f_i - f_{i_j} \mathbf{1}_{X\setminus N}|^p \, d\mu \\
&\leq \liminf_{j\to\infty} \int_X |f_i - f_{i_j} \mathbf{1}_{X\setminus N}|^p \, d\mu \\
&= \liminf_{j\to\infty} \int_X |f_i - f_{i_j}|^p \, d\mu \\
&= \liminf_{j\to\infty} ||f_i - f_{i_j}||_p^p \\
&\leq \left(\frac{\varepsilon}{2}\right)^p
\end{aligned}$$

가 되어 곧 임의의 $i \geq i_0$에 대해 $||f_i - f||_p < \varepsilon$이다. 이는 $f \in L^p$임을 뜻하기도 하므로 곧 $f_i \to f$ in sense of L^p가 되어 $(L^p, ||\cdot||_p)$가 Banach 공간임을 안다. □

Corollary 2.143 측도공간 $(X, \mathscr{A}, \mu)$와 $p \geq 1$에 대해 L^p에 속하는 함수열 $\{f_i\}$가 적당한 $f \in L^p$에 대해 $f_i \to f$ in sense of L^p이면 $\{f_i\}$의 적당한 부분함수열이 존재하여 거의 어디서나 f로 수렴한다.

Proof 함수열 $\{f_i\}$가 수렴하므로 이는 Cauchy이고, 따라서 위의 정리의 증명에서와 같이 $\{f_{i_j}\}$를 잡으면 이는 거의 어디서나 f로 수렴한다. □

요컨대, L^p 공간은 적어도 그 구조에서는 실수체에 버금갈 정도로 좋은 공간이고, 이러한 이유로 전술하였다시피 L^p 공간 자체에 대해 많은 연구가 이루어져 있다. 한편, $p = 2$인 경우에는 여기서 한 발 더 나아갈 수 있다.

Definition 2.144 내적공간 $(\mathsf{V}, \langle\cdot,\cdot\rangle)$에 대해 $\langle\cdot,\cdot\rangle$로부터 유도되는 노름이 완비성을 가지면 이때의 $(\mathsf{V}, \langle\cdot,\cdot\rangle)$를 **Hilbert 공간 (− space)**이라 한다.

Theorem 2.145 측도공간 $(X, \mathscr{A}, \mu)$에 대해 함수 $\langle\cdot,\cdot\rangle_\mu : (L^2)^2 \to \mathbb{R}$를 $\langle\cdot,\cdot\rangle_\mu : (f, g) \mapsto \int_X fg\,d\mu$로 정의하면 $(L^2, \langle\cdot,\cdot\rangle_\mu)$는 Hilbert 공간이다.

Proof 우선 임의의 $f, g \in L^2$에 대해 Hölder의 부등식으로부터 $\int_X |fg|\,d\mu \leq ||f||_2||g||_2 < \infty$이므로 fg는 적분가능하고, 곧 $\langle\cdot,\cdot\rangle_\mu$는 well-define된다. 이제 $\langle\cdot,\cdot\rangle_\mu$가 내적임은 거의 자명하고, 이가 유도하는 노름이 2−노름이므로 정리 2.142로부터 $(L^2, \langle\cdot,\cdot\rangle_\mu)$는 Hilbert 공간이다. □

보통 특별히 정하지 않는 이상 L^2 공간에는 위의 정리에서와 같은 내적이 주어진 것으로 생각하고, 혼동의 우려가 없다면 위 내적의 아래첨자 μ를 생략하여 그냥 $\langle\cdot,\cdot\rangle$로 쓴다. 한편, 위의 정리로부터 $(L^2, \langle\cdot,\cdot\rangle)$이 Hilbert 공간을 이루므로, 우리는 잠시 Hilbert 공간이라는 추상의 세계로 건너가 이가 가지는 성질을 알아보고자 한다. 이번 절을 시작하며 L^p 공간은 도함수를 일반화하기 위한 경유지라고 했는데 어쩐지 점점 더 우리의 목적으로부터 멀어지는 듯하다. 그러나 대부분의 추상화가 그러하듯, 추상의 세계에서 우리는 본질을 더 잘 바라볼 수 있는 법이다. 우리가 Hilbert 공간에서 특히 관심을 가질 부분은 이에서 정의된 선형사상의 characterization이다.

Theorem 2.146 Hilbert 공간 $(\mathsf{V}, ||\cdot||)$의 공집합이 아닌 부분집합 $K \subseteq \mathsf{V}$가 닫혀있고 볼록하다면 임의의 $v \in K$에 대해 $||v^*|| \leq ||v||$인 $v^* \in K$가 유일하게 존재한다.

Proof 먼저 존재성을 보이기 위해 $\delta = \inf_{v \in K} ||v||$라 하고 $||v_i|| \to \delta$인 K에 속하는 수열 $\{v_i\}$를 생각하자. 이제 임의의 $\varepsilon > 0$을 택하면 적당한 $i_0 \in \mathbb{N}$가 존재하여 임의의 $i \geq i_0$에 대해 $|||v_i||^2 - \delta^2| < \varepsilon/4$ 즉, $||v_i||^2 < \delta^2 + \varepsilon/4$이다. 한편, 임의의 $i, j \geq i_0$에 대해 K가 볼록집합이므로 $(v_i + v_j)/2 \in K$이고, 따라서 $||v_i + v_j|| \geq 2\delta$가 되어 이상으로부터

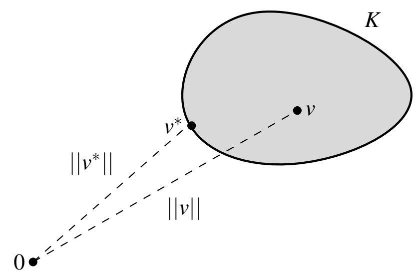


Figure 2.12 정리 2.146에서의 집합 K와 점 v^*.

$||v_i - v_j||^2 = 2||v_i||^2 + 2||v_j||^2 - ||v_i + v_j||^2 < 4\delta^2 + \varepsilon - 4\delta^2 = \varepsilon$ 이므로 $\{v_i\}$는 Cauchy이다. 그런데 가정으로부터 $\{v_i\}$는 K에서 수렴하고 그 수렴값을 $v^* \in K$라 하면 $||v_i|| \to ||v^*|| = \delta$ 이므로 임의의 $v \in K$에 대해 $||v|| \geq ||v^*||$가 되어 존재성이 보여진다.

유일성을 보이는 것은 보다 간단하다. 만약 $||v^*|| = ||\widetilde{v}|| = \delta$인 v^*와 다른 $\widetilde{v} \in K$가 존재한다면 K가 볼록집합이므로 $(v^* + \widetilde{v})/2 \in K$이고, 곧 $||v^* + \widetilde{v}|| \geq 2\delta$이다. 그런데 이는 $||v^* - \widetilde{v}||^2 = 2||v^*||^2 + 2||\widetilde{v}||^2 - ||v^* + \widetilde{v}||^2 \leq 4\delta^2 - 4\delta^2 = 0$에서 $v^* = \widetilde{v}$의 모순을 함의하므로 유일성이 보여진다. □

마침내, 짧았던 L^p 공간과 추상세계에서의 여정에 마침표를 찍을 정리를 소개한다. Riesz의 표현정리라 불리는 다음 정리의 결론은 Hilbert 공간에서의 모든 유계인 선형사상은 그 공간의 벡터 하나로 완벽히 특정된다는 것이다. 즉, 선형사상이 유계이면 그 선형사상에 대한 정보는 점 하나에 모두 담겨있다!

Theorem 2.147 (Riesz representation theorem) Hilbert 공간 $(\mathsf{V}, ||\cdot||)$에서 정의된 선형사상 $T : \mathsf{V} \to \mathbb{R}$가 유계라 하면 임의의 $v \in \mathsf{V}$에 대해 $T(v) = \langle v, v^* \rangle$인 $v^* \in \mathsf{V}$가 유일하게 존재한다. 나아가, 이러한 v^*에 대해 $||v^*|| = ||T||_{\text{op}}$가 성립한다.

Proof 먼저 존재성을 보이자. 만약 $T = 0$이면 $v^* = 0$이 정리의 조건을 만족하고 곧 정리가 자명하므로 $T \neq 0$이라 하자. 이제 집합 $K = T^{-1}(1)$을 생각하면 가정으로부터 $T(v) \neq 0$인 $v \in \mathsf{V}$를 적어도 하나 택할 수 있고, 이러한 v에 대해 $T(v/||T(v)||) = 1$에서 $v/||T(v)|| \in K$이므로 K는 공집합이 아니다. 또한, T가 유계이므로 이는 연속이고, 따라서 K는 닫혀있으며, 임의의 $v, w \in K$와 임의의 $t \in [0, 1]$에 대해 $T(tv + (1-t)w) = tT(v) + (1-t)T(w) = 1$이므로 $tv + (1-t)w \in K$가 되어 K는 볼록집합이다. 그렇다면 정리 2.146로부터 임의의 $v \in K$에 대해 $||\widetilde{v}|| \leq ||v||$인 $\widetilde{v} \in K$가 유일하게 존재한다.

존재성에 대한 증명을 마무리하기 위해서는 임의의 $v \in \ker T$에 대해 $\langle v, \widetilde{v} \rangle = 0$임을 보이는 것으로 충분하다. 만약 이를 보일 수 있으면 임의의 $v \in \mathsf{V}$에 대해 $T(v - T(v)\widetilde{v}) = T(v) - T(v) = 0$에서 $v - T(v)\widetilde{v} \in \ker T$이므로 곧 $0 = \langle v - T(v)\widetilde{v}, \widetilde{v} \rangle = \langle v, \widetilde{v} \rangle - T(v)||\widetilde{v}||^2$에서 $v^* = \widetilde{v}/||\widetilde{v}||^2$라 하면 $T(v) = \langle v, \widetilde{v} \rangle/||\widetilde{v}||^2 = \langle v, v^* \rangle$가 되어 존재성이 명백하기 때문이

다. 이제 임의의 $v \in \ker T$ 를 택하면 임의의 $t \in \mathbb{R}$에 대해 $T(\widetilde{v}+tv) = T(\widetilde{v}) + tT(v) = 1$ 이므로 $\widetilde{v}+tv \in K$에서 $||\widetilde{v}||^2 \leq ||\widetilde{v}+tv||^2 = ||\widetilde{v}||^2 + 2t\langle v, \widetilde{v}\rangle + t^2||v||^2$이다. 이는 곧 함수 $t \mapsto ||\widetilde{v}||^2 + 2t\langle v, \widetilde{v}\rangle + t^2||v||^2$가 $t = 0$에서 최댓값을 가짐을 뜻하므로 그 도함수가 0을 근으로 가진다는 사실로부터 $\langle v, \widetilde{v}\rangle = 0$이다.

유일성을 보이는 것은 보다 간단하다. 만약 임의의 $v \in \mathsf{V}$에 대해 $T(v) = \langle v, v^*\rangle = \langle v, \widetilde{v}\rangle$인 v^*가 아닌 $\widetilde{v} \in \mathsf{V}$가 존재한다면 $\langle v, v^* - \widetilde{v}\rangle = 0$인데, 여기서 $v = v^* - \widetilde{v}$를 택하면 $||v^* - \widetilde{v}||^2 = \langle v^* - \widetilde{v}, v^* - \widetilde{v}\rangle = 0$이 되어 $v^* = \widetilde{v}$의 모순이 발생하므로 유일성이 보여진다.

마지막으로, $||v^*|| = ||T||_{\text{op}}$임을 보이자. 만약 $v^* = 0$이면 명백히 $T = 0$이므로 $||v^*|| = ||T||_{\text{op}} = 0$에서 정리가 자명하므로 $v^* \neq 0$이라 하자. 그렇다면 $||v^*|| = \langle v^*, v^*\rangle / ||v^*|| = T(v^*)/||v^*|| \leq |T||_{\text{op}}$이다. 한편, Cauchy-Schwarz 부등식으로부터 임의의 $v \in \mathsf{V}$에 대해 $|T(v)| = |\langle v, v^*\rangle| \leq ||v|| ||v^*||$이므로 $||v^*|| \geq |T(v)|/||v||$에서 $||T||_{\text{op}} \leq ||v^*||$이 되어 이상으로부터 $||v^*|| = ||T||_{\text{op}}$임을 안다. □

Corollary 2.148 측도공간 $(X, \mathscr{A}, \mu)$와 선형사상 $T : L^2 \to \mathbb{R}$에 대해 T가 유계라면 임의의 $g \in L^2$에 대해 $T(g) = \int_X fg\,d\mu$인 $f \in L^2$가 존재하고, 거의 어디서나 같은 함수를 하나로 본다면 유일하다. 나아가 $||T||_{\text{op}} = ||f||_2$이다.

Proof 정리 2. 145로부터 $(L^2, \langle\cdot,\cdot\rangle)$이 Hilbert 공간으므로 Riesz의 표현정리로부터 이는 거의 자명하다. □

이제 다시 측도공간으로 돌아오자. 다음의 정리나 개념들은 훨씬 전에 소개할 수도 있었지만 논의의 전개가 산만해지는 것을 피하기 위해 일부러 조금 미루어 두었던 것들이다. 먼저 σ−대수의 제한을 소개한다.

Definition 2.149 가측공간 $(X, \mathscr{A})$와 공집합이 아닌 $Y \in \mathscr{A}$에 대해 집합족 $\{A \in \mathscr{A} : A \subseteq Y\}$를 $\mathscr{A}$의 Y로의 **제한 (restriction)**이라 하고 $\mathscr{A}|_Y$로 쓴다.

Proposition 2.150 가측공간 $(X, \mathscr{A})$와 집합 $Y \in \mathscr{A}$에 대해 $\mathscr{A}|_Y$는 Y 위의 σ−대수이다.

Proof 우선 $Y \in \mathscr{A}|_Y$이므로 이는 비어있지 않고 임의의 $A \in \mathscr{A}|_Y$에 대해 $Y \setminus A \in \mathscr{A}|_Y$에서 $\mathscr{A}|_Y$는 (Y를 전체 공간으로 보았을 때의) 여집합에 대해서 닫혀있다. 한편, 이가 가산 합집합에 대해 닫혀있음은 분명하므로 이상으로부터 $\mathscr{A}|_Y$는 Y 위의 σ−대수이다. □

Theorem 2.151 측도공간 $(X, \mathscr{A}, \mu)$와 집합 $Y \in \mathscr{A}$에 대해 $\mathscr{B} = \mathscr{A}|_Y$라 하면 다음이 성립한다.

i. 제한 $\mu|_{\mathscr{B}}$는 $\mathscr{B}$ 위의 측도이다. 따라서 $(X, \mathscr{B}, \mu|_{\mathscr{B}})$는 측도공간을 이룬다.
ii. 가측공간 $(Z, \mathscr{C})$와 $\mathscr{A}/\mathscr{C}$−가측함수 $f : X \to Z$에 대해 $f|_Y$는 $\mathscr{B}/\mathscr{C}$−가측이다.

iii. 음이 아닌 $\mathscr{A}/\mathscr{B}_1$−가측함수 $f: X \to \overline{\mathbb{R}}_0^+$ 에 대해 $\int_Y f\,d\mu = \int_Y f|_Y\,d\mu|_{\mathscr{B}}$ 이다. 한편, Y 에서 μ−적분가능한 (혹은 $f|_Y$ 가 $\mu|_{\mathscr{B}}$−적분가능한 $\mathscr{A}/\mathscr{B}_1$−가측함수) $f: X \to \overline{\mathbb{R}}$ 에 대해서도 동일한 결과가 성립하며, 이때 $f|_Y$ 는 $\mu|_{\mathscr{B}}$−적분가능하다 (혹은 f 는 Y 에서 μ−적분가능하다).

Proof i. 이는 명제 2. 150로부터 $\mathscr{B}$가 Y 위의 σ−대수라는 점에서 자명하다.

ii. 이 또한 거의 자명하다. 임의의 $A \in \mathscr{C}$ 에 대해 $(f|_Y)^{-1}(A) = f^{-1}(A) \cap Y \in \mathscr{B}$ 이므로 $f|_Y$ 가 $\mathscr{B}/\mathscr{C}$−가측이다.

iii. 먼저 f가 음이 아닌 단순함수인 경우를 생각하여 이의 표준형을 $f = \sum_{i=1}^k a_i \mathbf{1}_{A_i}$ 라 하면 $f|_Y = \sum_{i=1}^k a_i \mathbf{1}_{A_i \cap Y}$ 이고, 이로부터 $\int_Y f|_Y\,d\mu|_{\mathscr{B}} = \sum_{i=1}^k a_i \mu|_{\mathscr{B}}(A_i \cap Y) = \sum_{i=1}^k a_i \mu(A_i \cap Y) = \int_X f \mathbf{1}_Y\,d\mu = \int_Y f\,d\mu$ 이다. 이제 f 를 음이 아닌 가측함수라 하면 정리 2. 83로부터 음이 아닌 단순함수 $f_i: X \to \mathbb{R}_0^+$ 의 열 $\{f_i\}$ 가 존재하여 $f_i \uparrow f$ 이므로 MCT에서 $\int_Y f|_Y\,d\mu|_{\mathscr{B}} = \lim_{i\to\infty} \int_Y f_i|_Y\,d\mu|_{\mathscr{B}} = \lim_{i\to\infty} \int_Y f_i\,d\mu = \int_Y f\,d\mu$ 이다.

한편, Y 에서 μ−적분가능한 $f: X \to \overline{\mathbb{R}}$ 에 대해서, 이를 양과 음의 부분으로 나누어 앞선 결과를 적용하면 $f|_Y$ 가 $\mu|_{\mathscr{B}}$−적분가능하며 $\int_Y f|_Y\,d\mu|_{\mathscr{B}} = \int_Y f\,d\mu$ 임을 쉽게 알 수 있고, $f|_Y$ 가 $\mu|_{\mathscr{B}}$−적분가능한 $\mathscr{A}/\mathscr{B}_1$−가측함수 $f: X \to \overline{\mathbb{R}}$ 에 대해서도 비슷하게 하면 된다. □

위의 정리는, 비록 그 표현은 복잡해 보이지만, 본질적으로는 피적분함수와 적분에 사용되는 측도의 적분영역 밖의 정보는 그 적분에 있어 아무런 필요가 없다는 지극히 상식적인 아이디어를 확인한 것에 불과하다.

다음으로, 측도의 선형결합에 대한 내용으로 넘어가자. 이 또한 매우 직관적인 결과들이다.

Theorem 2.152 가측공간 $(X, \mathscr{A})$ 위의 측도 μ, ν 에 대해 다음이 성립한다.

i. 임의의 $c \geq 0$에 대해 함수 $c\mu$ 도 측도이다.
ii. 함수 $\mu + \nu$ 도 측도이다.

Proof 자명하다. i의 $c\mu$ 와 ii의 $\mu + \nu$ 가 측도의 조건들을 만족함을 쉽게 확인할 수 있다. □

Theorem 2.153 가측공간 $(X, \mathscr{A})$ 위의 측도 μ, ν 와 음이 아닌 가측함수 $f: X \to \overline{\mathbb{R}}_0^+$ 에 대해 다음의 성립한다.

i. 임의의 $c \geq 0$에 대해 $\int_X f\,d(c\mu) = c\int_X f\,d\mu$ 이다.
ii. $\int_X f\,d(\mu+\nu) = \int_X f\,d\mu + \int_X f\,d\nu$ 이다.

한편, μ−적분가능한 $f: X \to \overline{\mathbb{R}}$ 에 대해서도 i의 성질이 성립하고, 이때 f 는 $c\mu$−적분가능하다. 비슷하게, μ−적분가능한 동시에 ν−적분가능한 (혹은 $(\mu+\nu)$−적분가능한) $f: X \to \overline{\mathbb{R}}$ 에 대해서도 ii의 성질이 성립하고, 이때 f 는 $(\mu+\nu)$−적분가능하다 (혹은 μ−적분가능한 동시에 ν−적분가능하다).

PROOF 편의를 위해 i만 증명하도록 한다. (ii는 i과 비슷하게 하면 된다.) 먼저 f가 음이 아닌 단순함수인 경우를 생각하여 이의 표준형을 $f = \sum_{i=1}^k a_i \mathbf{1}_{A_i}$ 라 하면 $\int_X f\,d(c\mu) = \sum_{i=1}^k a_i(c\mu)(A_i) = c\sum_{i=1}^k a_i\mu(A_i) = c\int_X f\,d\mu$ 가 성립한다. 이제 f를 음이 아닌 가측함수라 하면 정리 2.83로부터 음이 아닌 단순함수 $f_i : X \to \mathbb{R}_0^+$ 의 열 $\{f_i\}$가 존재하여 $f_i \uparrow f$ 이므로 MCT에서 $\int_X f\,d(c\mu) = \lim_{i\to\infty}\int_X f_i\,d(c\mu) = c\lim_{i\to\infty}\int_X f_i\,d\mu = c\int_X f\,d\mu$ 임을 안다.

한편, $\mu-$적분가능한 $f : X \to \overline{\mathbb{R}}$에 대해서, 이를 양과 음의 부분으로 나누어 앞선 결과를 적용하면 f가 $c\mu-$적분가능하며 $\int_X f\,d(c\mu) = c\int_X f\,d\mu$ 임을 쉽게 알 수 있다. □

이로써 도함수를 일반화하기 위한 준비가 모두 끝났다. 아마 아직은 감이 오지 않겠지만, 우리가 L^p 공간과 Hilbert 공간을 오가며 얻은 결과들과 뜬금없이 소개한 $\sigma-$대수의 제한과 같은 내용들이 마법과 같이 하나로 합쳐져 우리가 원하는 도함수의 일반화로 이어질 것이다.

Definition 2.154 가측공간 $(X, \mathscr{A})$ 위의 측도 μ, ν를 생각하자. 만약 $\nu(N) = 0$인 모든 $N \in \mathscr{A}$에 대해 $\mu(N) = 0$이면 이때 μ는 ν에 대해 **절대연속 (absolutely continuous)**이라 하고 $\mu \ll \nu$로 쓴다.

Lemma 2.155 유한 측도공간 $(X, \mathscr{A}, \mu)$에 대해 $L^2 \subseteq L^1$이다.

PROOF 임의의 $f \in L^2$에 대해 Cauchy-Schwarz 부등식으로부터 $\int_X |f|\,d\mu = \langle |f|, 1\rangle \le ||f||_2||1||_2 = \sqrt{(\int_X f^2\,d\mu)(\int_X d\mu)} = \sqrt{\mu(X)}||f||_2 < \infty$이고, 곧 $f \in L^1$이므로 $L^2 \subseteq L^1$ 임을 안다. □

Theorem 2.156 (Radon-Nikodým) 가측공간 $(X, \mathscr{A})$ 위의 측도 μ와 $\sigma-$유한 측도 ν에 대해 TFAE.

i. 측도 μ는 $\sigma-$유한하고 $\mu \ll \nu$이다.
ii. 음이 아닌 가측함수 $f : X \to \mathbb{R}_0^+$가 존재하여 $\mu = \nu_f$이다.

특히, 만약 i이 성립한다면 ii에서의 가측함수 f는 $\nu-$거의 어디서나 같은 함수를 하나로 본다면 유일하다. 나아가, 만약 μ가 유한하다면 f는 $\nu-$적분가능하다.

PROOF ii $\Rightarrow$ i. 우선 임의의 $N \in \mathscr{A}$에 대해 $\nu(N) = 0$이라 하면 보조정리 2.102의 i로부터 $\mu(N) = \nu_f(N) = 0$이 되어 $\mu \ll \nu$ 임은 분명하므로 μ가 $\sigma-$유한함만 보이면 된다. 가정으로부터 ν가 $\sigma-$유한하므로 $\mathscr{A}$에 속하는 적당한 $\{A_i\}$가 존재하여 각 $i \in \mathbb{N}$에 대해 $\nu(A_i) < \infty$이고 $\bigcup_{i=1}^\infty A_i = X$이며, WLOG, 필요하다면 각 항을 $C_i := \bigcup_{j=1}^i A_j$로 바꾸어 $\{A_i\}$가 처음부터 증가하는 집합열이라 해도 된다. 이제 집합열 $\{B_i\}$를 $B_i = f^{-1}((-\infty, i]) \cap A_i$라 하면 이는 명백히 증가하는 집합열로 각 $i \in \mathbb{N}$에 대해 $\mu(B_i) = \int_{B_i} f\,d\nu \le \int_{A_i} i\,d\nu = i\nu(A_i) < \infty$이고 $\bigcup_{i=1}^\infty B_i = X$ 임이 분명하므로 μ도 $\sigma-$유한하다.

i $\Rightarrow$ ii. 우선 ν, μ가 모두 유한한 특별한 경우를 생각하자. 이제 함수 $T : L^2(\nu+\mu) \to \mathbb{R}$를 $T : g \mapsto \int_X g\,d\mu$로 두면 ν, μ가 유한하므로 $\nu + \mu$도 유한하여 위의 보조정리와 정리

2.153의 ii로부터 $L^2(\nu+\mu) \subseteq L^1(\nu+\mu) \subseteq L^1(\mu)$이고, 곧 T는 well-define된다. 또한, T가 선형사상임이 분명하고 임의의 $g \in L^2(\nu+\mu)$에 대해 Cauchy-Schwarz 부등식으로부터

$$\begin{aligned}
|T(g)| &= \left| \int_X g\, d\mu \right| \\
&\leq \int_X |g|\, d\mu \\
&\leq \int_X |g|\, d\nu + \int_X |g|\, d\mu \\
&= \int_X |g|\, d(\nu+\mu) \\
&= \langle |g|, 1 \rangle_{\nu+\mu} \\
&\leq ||g||_{2,\nu+\mu} ||1||_{2,\nu+\mu} \\
&= \sqrt{\left[\int_X g^2\, d(\nu+\mu)\right]\left[\int_X d(\nu+\mu)\right]} \\
&= \sqrt{(\nu+\mu)(X)}||g||_{2,\nu+\mu}
\end{aligned}$$

이므로 T는 유계인 선형사상이 되어 Riesz의 표현정리로부터 임의의 $g \in L^2(\nu+\mu)$에 대해 $\int_X g\, d\mu = T(g) = \int_X gh\, d(\nu+\mu)$인 $h \in L^2(\nu+\mu)$가 존재한다. 나아가 임의의 $g \in L^2(\nu+\mu)$에 대해 $\int_X g(1-h)\, d(\nu+\mu) = \int_X g\, d(\nu+\mu) - \int_X gh\, d(\nu+\mu) = \int_X g\, d(\nu+\mu) - \int_X g\, d\mu = \int_X g\, d\nu$도 성립한다.

더 진행하기 전에 $0 \leq h < 1$ $((\nu+\mu)\text{–ae.})$임을 확인하고 가자. 이를 위해 집합 $A = h^{-1}((-\infty, 0)), B = h^{-1}([1,\infty))$를 생각하면 $\mu(A) = \int_X \mathbf{1}_A\, d\mu = \int_X h\mathbf{1}_A\, d(\nu+\mu) \leq 0$이므로 $\int_X h\mathbf{1}_A\, d(\nu+\mu) = 0$이고 곧 정리 2.101의 ii로부터 $h\mathbf{1}_A = 0$ $((\nu+\mu)\text{–ae.})$가 되어 $(\nu+\mu)(A) = (\nu+\mu)((h\mathbf{1}_A)^{-1}((-\infty,0))) = 0$이다. 비슷하게, $\nu(B) = \int_X \mathbf{1}_B\, d\nu = \int_X (1-h)\mathbf{1}_B\, d(\nu+\mu) \leq 0$이므로 $\nu(B) = 0$이고, $\mu \ll \nu$라는 가정으로부터 $\mu(B) = 0$이 되어 $(\nu+\mu)(B) = 0$이다. 이상으로부터 A, B가 모두 $(\nu+\mu)$–영집합이 되어 $0 \leq h < 1$ $((\nu+\mu)\text{–ae.})$이다.

WLOG, 필요하다면 $(\nu+\mu)$–영집합에서의 h의 값을 0으로 바꾸어 그냥 $0 \leq h < 1$이라 해도 된다. 한편, 임의의 $A \in \mathscr{A}$에 대해 $\nu(A) = \int_X \mathbf{1}_A\, d\nu = \int_X (1-h)\mathbf{1}_A\, d(\nu+\mu) = \int_A 1-h\, d(\nu+\mu) = (\nu+\mu)_{1-h}(A)$이므로 임의의 음이 아닌 가측함수 $g : X \to \overline{\mathbb{R}}_0^+$에 대해 $\int_X g\, d\nu = \int_X g\, d(\nu+\mu)_{1-h} = \int_X g(1-h)\, d(\nu+\mu)$이다. 따라서 함수 $f : X \to \overline{\mathbb{R}}_0^+$를 $f = h/(1-h)$로 두면 이는 가측이고 임의의 $A \in \mathscr{A}$에 대해 $\mu(A) = \int_X \mathbf{1}_A\, d\mu = \int_X h\mathbf{1}_A\, d(\nu+\mu) = \int_X f(1-h)\mathbf{1}_A\, d(\nu+\mu) = \int_X f\mathbf{1}_A\, d\nu = \int_A f\, d\nu$가 되어 $\mu = \nu_f$에서 정리가 성립한다.

이제 σ–유한한 ν, μ를 생각하면 정의로부터 $\mathscr{A}$에 속하는 적당한 $\{A_i\}, \{B_i\}$가 존재하여 $\bigcup_{i=1}^\infty A_i = \bigcup_{i=1}^\infty B_i = X$이고 각 $i \in \mathbb{N}$에 대해 $\nu(A_i), \mu(B_i) < \infty$이다. 또한, 앞서 ii가 i을

함의함을 증명할 때와 비슷하게 하여 WLOG, $\{A_i\}, \{B_i\}$가 증가한다고 해도 된다. 이제 집합열 $\{Y_i\}$를 $Y_i = A_i \cap B_i$로 두면 이는 증가하는 집합열로서 $\bigcup_{i=1}^{\infty} Y_i = \bigcup_{i=1}^{\infty} A_i \cap \bigcup_{i=1}^{\infty} B_i = X$이고 각 $i \in \mathbb{N}$에 대해 $\nu(Y_i) \leq \nu(A_i) < \infty$이며 $\mu(Y_i) \leq \mu(B_i) < \infty$이다. 또한, 표기의 편의를 위해 각 $i \in \mathbb{N}$에 대해 $\mathscr{B}_i = \mathscr{A}|_{Y_i}, \nu_i = \nu|_{\mathscr{B}_i}, \mu_i = \mu|_{\mathscr{B}_i}$로 두면 ν_i, μ_i는 유한하며 $\mathscr{B}_{i+1}|_{Y_i} = \mathscr{B}_i, \nu_{i+1}|_{\mathscr{B}_i} = \nu_i, \mu_{i+1}|_{\mathscr{B}_i} = \mu_i$이다. 그렇다면 각 $i \in \mathbb{N}$에 대해 $\mu_i \ll \nu_i$임이 분명하므로 앞선 결과로부터 적당한 음이 아닌 가측함수 $f_i : Y_i \to \mathbb{R}_0^+$가 존재하여 $\mu_i = (\nu_i)_{f_i}$이고, 각 $i \in \mathbb{N}$와 임의의 $A \in \mathscr{B}_i$에 대해 정리 2.151로부터 $\int_A f_i \, d\nu_i = (\nu_i)_{f_i}(A) = \mu_i(A) = \mu_{i+1}(A) = (\nu_{i+1})_{f_{i+1}}(A) = \int_A f_{i+1} \, d\nu_{i+1} = \int_A f_{i+1}|_{Y_i} \, d\nu_i$이다. 이는 곧 정리 2.114로부터 각 $i \in \mathbb{N}$에 대해 $f_i = f_{i+1}|_{Y_i}$ (ν_i–ae.)임을 함의하므로 WLOG, 필요하다면 ν_i–영집합에서의 f_{i+1}의 값을 조금씩 바꾸어 그냥 $f_i = f_{i+1}|_{Y_i}$라 해도 되고, 이로부터 함수 $f : X \to \mathbb{R}_0^+$를 임의의 $i \in \mathbb{N}$에 대해 $f|_{Y_i} = f_i$가 되도록 정의할 수 있다. 그렇다면 임의의 $x \in \mathbb{R}$에 대해 $f^{-1}((-\infty, x]) = \bigcup_{i=1}^{\infty} [f^{-1}((-\infty, x]) \cap Y_i] = \bigcup_{i=1}^{\infty} (f_i)^{-1}((-\infty, x])$이므로 f는 가측이다. 이제 임의의 $C \in \mathscr{A}$에 대해 집합열 $\{C_i\}$를 $C_i = C \cap Y_i$로 두면 $C_i \uparrow C$이므로 정리 2.24의 i과 정리 2.112의 i로부터

$$
\begin{aligned}
\mu(C) &= \lim_{i\to\infty} \mu(C_i) \\
&= \lim_{i\to\infty} \mu_i(C_i) \\
&= \lim_{i\to\infty} (\nu_i)_{f_i}(C_i) \\
&= \lim_{i\to\infty} \int_{C_i} f_i \, d\nu_i \\
&= \lim_{i\to\infty} \int_{C_i} f \, d\nu \\
&= \int_C f \, d\nu \\
&= \nu_f(C)
\end{aligned}
$$

가 되어 $\mu = \nu_f$이다.

한편, 정리 2.114로부터 이러한 f가 $\nu-$거의 어디서나 같은 함수를 하나로 본다면 유일함이 분명하고, 만약 μ가 유한하다면 $\int_X f \, d\nu = \nu_f(X) = \mu(X) < \infty$에서 f가 ν–적분가능하므로 증명이 끝난다. □

Definition 2.157 가측공간 $(X, \mathscr{A})$ 위의 σ–유한 측도 μ, ν에 대해 $\mu \ll \nu$이면 이때 $\mu = \nu_f$를 만족하는 함수 $f : X \to \mathbb{R}_0^+$를 μ의 ν에 대한 **Radon-Nikodým 도함수 (− derivative)**라 하고 $d\mu/d\nu$로 쓴다.

치환적분법을 떠오르게 하는 다음 따름정리를 보면 방금 우리가 얻은 정의를 왜 도함수의 일반화로 볼 수 있는지가 더욱 분명해 질 것이다.

Corollary 2.158 가측공간 $(X, \mathscr{A})$ 위의 σ−유한 측도 ν, μ에 대해 $\mu \ll \nu$라 하자. 그렇다면 임의의 음이 아닌 가측함수 $f : X \to \overline{\mathbb{R}}_0^+$에 대해 $\int_X f\,d\mu = \int_X f(d\mu/d\nu)\,d\nu$이다. 한편, μ−적분가능한 (혹은 $f(d\mu/d\nu)$가 ν−적분가능한 가측함수) $f : X \to \overline{\mathbb{R}}$에 대해서도 동일한 결과가 성립하며 이때 $f(d\mu/d\nu)$는 ν−적분가능하다 (혹은 f는 μ−적분가능하다).

Proof 이는 정리 2.113로부터 자명하다. □

2.11 Fubini's Theorem

이번 절에서는 그 제목에서와 같이 다중적분을 반복적분으로써 계산할 수 있도록 해 주는 Fubini의 정리를 증명한다. 사실 이전 절에서 보인 바와 같이 Lebesgue 적분은 Riemann 적분을 포괄하는 개념이므로 실제로 그 적분값을 구해야 할 필요가 있을 때에는 이를 Riemann 적분으로 생각하여 계산하면 되고, 미적분학이나 해석학에서 Riemann 적분에서의 Fubini의 정리를 이미 배웠으므로 이번 절에서의 논의가 없더라도 계산에 있어 Fubini의 정리를 사용할 수 있다. 그럼에도 불구하고, 하나의 절을 할애하여 이에 관한 논의를 진행하는 것에는 나름의 이유가 있다.

먼저, 종전에 우리가 배운 Riemann 적분에서의 Fubini의 정리는 피적분함수에 대해, 예컨대 이가 미분가능해야 한다거나, 혹은 조각적 연속이어야 한다거나 하는 다소 강한 조건을 요구한다. 또한, 어디까지나 Lebesgue 적분이 Riemann 적분을 포괄하는 것이지 이 두 개념이 서로 동등한 것이 아니므로 Lebesgue 적분을 언제나 Riemann 적분으로 생각하여 계산할 수 있는 것은 아니다. 예컨대 Riemann 적분은 불가능하지만 Lebesgue 적분은 가능한 대표적인 함수인 $\mathbf{1}_{\mathbb{Q}\cap[0,1]}$의 적분을 계산해야 하는 경우, 이는 Lebesgue 적분으로만 계산할 수 있다.[21]비록 이러한 경우가 이 책에서는 (아마) 없겠지만, 그래도 이론적인 측면에서 조금 불안한 것이 사실이다. 이에 이번 절에서 Fubini의 정리를 Lebesgue 적분론의 틀 안에서 최대한 일반적으로 다시 증명함으로써 이런 이론적인 불안정함을 해소하고자 한다.

우선 가장 기본적인 형태로 product measure space에서 정의된 음이 아닌 가측함수에 대한 Fubini의 정리를 보자.

Lemma 2.159 σ−유한 측도공간 $(X, \mathscr{A}, \mu), (Y, \mathscr{B}, \nu)$와 음이 아닌 $(\mathscr{A} \otimes \mathscr{B})/\mathscr{B}_1$−가측함수 $f : X \times Y \to \overline{\mathbb{R}}_0^+$에 대해 함수 $\varphi : X \to \overline{\mathbb{R}}_0^+, \psi : Y \to \overline{\mathbb{R}}_0^+$를 각각 $\varphi : x \mapsto \int_Y f_x\,d\nu, \psi : y \mapsto \int_X f^y\,d\mu$라 하면 이는 well-define되며 각각 $\mathscr{A}/\mathscr{B}_1$−가측이고 $\mathscr{B}/\mathscr{B}_1$−가측이다.

Proof 우선 정리 2.120로부터 모든 $x \in X$에 대해 f_x가 $\mathscr{B}/\mathscr{B}_1$−가측이므로 함수 φ는 well-define된다. 이제 f가 음이 아닌 단순함수인 특별한 경우를 생각하여 이의 표준형을 $f = \sum_{i=1}^k a_i \mathbf{1}_{A_i}$라 하자. 그렇다면 $\varphi(x) = \int_Y (\sum_{i=1}^k a_i \mathbf{1}_{A_i})_x\,d\nu = \sum_{i=1}^k a_i \int_Y \mathbf{1}_{(A_i)_x}\,d\nu =$

$\sum_{i=1}^{k} a_i \nu((A_i)_x)$ 이므로 보조정리 2. 121로부터 φ가 $\mathscr{A}/\mathscr{B}_1$−가측이다. 한편, 음이 아닌 가측함수 f의 경우에는 정리 2. 83로부터 음이 아닌 단순함수 $f_i : X\times Y \to \mathbb{R}_0^+$의 열 $\{f_i\}$가 존재하여 $f_i \uparrow f$이고, 각 $i \in \mathbb{N}$에 대해 함수 $\varphi_i : X \to \overline{\mathbb{R}}_0^+$를 $\varphi_i : x \mapsto \int_Y (f_i)_x \, d\nu$라 하면 앞선 결과로부터 이는 $\mathscr{A}/\mathscr{B}_1$−가측이다. 그렇다면 MCT로부터 $\lim_{i\to\infty} \varphi_i(x) = \int_Y \lim_{i\to\infty}(f_i)_x \, d\nu = \int_Y f_x \, d\nu = \varphi(x)$이므로 정리 2. 77의 ii로부터 φ는 $\mathscr{A}/\mathscr{B}_1$−가측이다. 이제 ψ에 대해서도 이와 비슷하게 하면 된다. □

Theorem 2.160 (Fubini) σ−유한 측도공간 $(X, \mathscr{A}, \mu), (Y, \mathscr{B}, \nu)$와 $(\mathscr{A}\otimes\mathscr{B})/\mathscr{B}_1$−가측함수 $f : X\times Y \to \overline{\mathbb{R}}_0^+$에 대해

$$\int_{X\times Y} f(x,y)\, d(\mu\otimes\nu)(x,y) = \int_X \left[\int_Y f(x,y)\, d\nu(y)\right] d\mu(x) = \int_Y \left[\int_X f(x,y)\, d\mu(x)\right] d\nu(y)$$

가 성립한다.

Proof 먼저 위의 보조정리로부터 이중적분들이 well-define됨을 안다. 이제 함수 $\varphi : X \to \overline{\mathbb{R}}_0^+$를 위의 보조정리에서와 같이 두고 f가 음이 아닌 단순함수인 특별한 경우를 생각하여 이의 표준형을 $f = \sum_{i=1}^{k} a_i \mathbf{1}_{A_i}$라 하면 위의 보조정리의 증명에서 $\varphi(x) = \sum_{i=1}^{k} a_i \nu((A_i)_x)$이므로 정리 2. 122로부터 $\int_X \varphi\, d\mu = \sum_{i=1}^{k} a_i \int_X \nu((A_i)_x)\, d\mu(x) = \sum_{i=1}^{k} a_i (\mu\otimes\nu)(A_i) = \int_{X\times Y} f\, d(\mu\otimes\nu)$이다. 한편, 음이 아닌 가측함수 f의 경우에는 정리 2. 83로부터 음이 아닌 단순함수 $f_i : X\times Y \to \mathbb{R}_0^+$의 열 $\{f_i\}$가 존재하여 $f_i \uparrow f$이고 함수열 $\{\varphi_i\}$를 위의 보조정리의 증명에서와 같이 두면 $\varphi_i \uparrow \varphi$이므로 MCT와 앞선 결론으로부터 $\int_X \varphi\, d\mu = \lim_{i\to\infty} \int_X \varphi_i\, d\mu = \lim_{i\to\infty} \int_{X\times Y} f_i\, d(\mu\otimes\nu) = \int_{X\times Y} f\, d(\mu\otimes\nu)$이다. 이제, 두 번째 등식에 대해서도 이와 비슷하게 하면 된다. □

다음으로, product measure space에서 정의된 적분가능한 함수에 대한 Fubini의 정리를 볼텐데, 이 경우에는 피적분함수의 section이 항상 적분가능하지 않으므로 조금 더 복잡하다.

Lemma 2.161 σ−유한 측도공간 $(X, \mathscr{A}, \mu), (Y, \mathscr{B}, \nu)$와 $(\mu\otimes\nu)$−적분가능한 함수 $f : X\times Y \to \overline{\mathbb{R}}$에 대해 다음이 성립한다.

i. Section $f_x : Y \to \overline{\mathbb{R}}$는 μ−거의 대부분의 $x \in X$에 대해 ν−적분가능하며 함수 $\varphi : X \to \mathbb{R}$를

$$\varphi : x \mapsto \begin{cases} \displaystyle\int_Y f_x\, d\nu & f_x\text{가 } \nu\text{−적분가능한 경우} \\ 0 & \text{ow.} \end{cases}$$

라 하면 이는 μ−적분가능하다.

ii. Section $f^y : X \to \overline{\mathbb{R}}$는 ν−거의 대부분의 $y \in Y$에 대해 μ−적분가능하며 함수 $\psi : Y \to \mathbb{R}$를

$$\psi : x \mapsto \begin{cases} \int_X f^y \, d\mu & f^y \text{가 } \mu\text{−적분가능한 경우} \\ 0 & \text{ow.} \end{cases}$$

라 하면 이는 ν−적분가능하다.

PROOF 간결한 논의를 위해 i만 보이도록 하자. (ii는 i과 비슷하게 하면 된다.) 함수 $f_\pm$가 $(\mathscr{A} \otimes \mathscr{B})/\mathscr{B}_1$−가측이므로 함수 $\widetilde{\varphi}^\pm : X \to \overline{\mathbb{R}}_0^+$를 $\widetilde{\varphi}^\pm : x \mapsto \int_Y (f_\pm)_x \, d\nu$로 두면 (여기서 위첨자로 쓰인 $\pm$은 정의 2.97에서 아래첨자로 사용된 $\pm$와는 달리 함수의 양과 음의 부분을 의미하는 것이 아니다.) 이는 보조정리 2.159로부터 $\mathscr{A}/\mathscr{B}_1$−가측이고, 곧 Fubini의 정리로부터 $\int_X \widetilde{\varphi}^\pm \, d\mu = \int_{X \times Y} f_\pm \, d(\mu \otimes \nu) \le \int_{X \times Y} |f| \, d(\mu \otimes \nu) < \infty$가 되어 $\widetilde{\varphi}^\pm$이 μ−적분가능함을 안다. 그렇다면 $\widetilde{\varphi}^\pm$이 μ−거의 어디서나 유한하므로 명백히 $\mathscr{A}$−가측인 집합 $N = (\widetilde{\varphi}^\pm)^{-1}(\infty) = \{x \in X : \int_Y |f_x| \, d\nu = \infty\}$에 대해 $\mu(N) = 0$이고, 곧 f_x는 μ−거의 대부분의 $x \in X$에 대해 ν−적분가능하다. 이제 함수 $\varphi^\pm : X \to \mathbb{R}_0^+$를 $\varphi^\pm = \widetilde{\varphi}^\pm \mathbf{1}_{X \setminus N}$로 두면 $\varphi^\pm = \widetilde{\varphi}^\pm$ (μ−ae.)이므로 따름정리 2.102의 ii로부터 $\varphi^\pm$은 μ−적분가능하고, 임의의 $x \in X$에 대해 $(f_x)_\pm = (f_\pm)_x$가 성립하여 $\varphi = \varphi^+ - \varphi^-$이므로 증명이 끝난다. □

Theorem 2.162 (Fubini) σ−유한 측도공간 $(X, \mathscr{A}, \mu), (Y, \mathscr{B}, \nu)$와 $(\mu \otimes \nu)$−적분가능한 함수 $f : X \times Y \to \overline{\mathbb{R}}$에 대해

$$\int_{X \times Y} f(x, y) \, d(\mu \otimes \nu)(x, y) = \int_X \left[\int_Y f(x, y) \, d\nu(y) \right] d\mu(x) = \int_Y \left[\int_X f(x, y) \, d\mu(x) \right] d\nu(y)$$

가 성립한다. 단, 만약 고정된 $x \in X$에 대해 $f(x, y)$가 ν−적분가능하지 않다면 이때의 $\int_Y f(x, y) \, d\nu(y)$의 값은 0으로 간주하고, 이는 고정된 $y \in Y$에 대해 $f(x, y)$가 μ−적분가능하지 않은 경우에도 마찬가지이다.

PROOF 먼저 위의 보조정리로부터 이중적분들이 well-define됨을 안다. 이제 함수 $\varphi : X \to \mathbb{R}$와 $\varphi^\pm : X \to \mathbb{R}_0^+$를 각각 위의 보조정리와 그 증명에서와 같이 두면 $\varphi = \varphi^+ - \varphi^-$에서 $\int_X \varphi \, d\mu = \int_X \varphi^+ \, d\mu - \int_X \varphi^- \, d\mu = \int_{X \times Y} f_+ \, d(\mu \otimes \nu) - \int_{X \times Y} f_- \, d(\mu \otimes \nu) = \int_{X \times Y} f \, d(\mu \otimes \nu)$임을 쉽게 알 수 있다. 이제 두 번째 등식에 대해서도 이와 비슷하게 하면 된다. □

비록 위의 정리에서는 적분불가능한 section에 대해 그 적분값을 0이라는 dummy value로 두었지만, 보조정리 2.161로부터 피적분함수의 거의 대부분의 section이 적분가능하므로 전체 적분의 값은 dummy value의 선택과는 무관하다. 그리고 이로부터 위의 반복적분 표기는 나름 make sense한다. 이러한 사실은 뒤따르는 나머지 Fubini의 정리에 대해서도 마찬가지이다.

이제 중간 결과로서 Borel 측도공간에서의 Fubini의 정리를 얻는다.

Corollary 2.163 $l + m = n$인 $l, m, n \in \mathbb{N}$과 음이 아닌 Borel 함수 $f : \mathbb{R}^n \to \overline{\mathbb{R}}_0^+$에 대해

$$\int_{\mathbb{R}^n} f(x,y)\,d\mu_n(x,y) = \int_{\mathbb{R}^l}\left[\int_{\mathbb{R}^m} f(x,y)\,d\mu_m(y)\right] d\mu_l(x) = \int_{\mathbb{R}^m}\left[\int_{\mathbb{R}^l} f(x,y)\,d\mu_l(x)\right] d\mu_m(y)$$

가 성립한다. 한편, 적분가능한 Borel 함수 $f:\mathbb{R}^n\to\overline{\mathbb{R}}$에 대해서도 동일한 결과가 성립한다. 단, 이 경우에는 만약 고정된 $x\in\mathbb{R}^l$에 대해 $f(x,y)$가 적분가능하지 않다면 이때의 $\int_{\mathbb{R}^m} f(x,y)\,d\mu_m(y)$의 값은 0으로 간주하고, 이는 고정된 $y\in\mathbb{R}^m$에 대해 $f(x,y)$가 적분가능하지 않은 경우에도 마찬가지이다.

Proof 이는 Fubini의 정리와 정리 2. 125의 i로부터 자명하다. □

한편, Lebesgue 측도공간은 Borel 측도공간과는 달리 product measure space의 완비화로써 구성되므로 조금 더 복잡한 Fubini의 정리가 필요하다. 논의의 전체적인 전개 방향은 이전과 비슷하다.

Lemma 2.164 σ−유한 완비측도공간 $(X,\mathscr{A},\mu)$, $(Y,\mathscr{B},\nu)$와 음이 아닌 $\overline{\mathscr{A}\otimes\mathscr{B}}/\mathscr{B}_1$−가측 함수 $f:X\times Y\to\overline{\mathbb{R}}_0^+$에 대해 다음이 성립한다.

i. Section $f_x:Y\to\overline{\mathbb{R}}$는 μ−거의 대부분의 $x\in X$에 대해 $\mathscr{B}/\mathscr{B}_1$−가측이며 함수 $\varphi:X\to\overline{\mathbb{R}}_0^+$를

$$\varphi: x\mapsto\begin{cases}\displaystyle\int_Y f_x\,d\nu & f_x\text{가 } \mathscr{B}/\mathscr{B}_1\text{−가측인 경우}\\ 0 & \text{ow.}\end{cases}$$

라 하면 이는 $\mathscr{A}/\mathscr{B}_1$−가측이다.

ii. Section $f^y:X\to\overline{\mathbb{R}}$는 ν−거의 대부분의 $y\in Y$에 대해 $\mathscr{A}/\mathscr{B}_1$−가측이며 함수 $\psi:Y\to\overline{\mathbb{R}}_0^+$를

$$\psi: x\mapsto\begin{cases}\displaystyle\int_X f^y\,d\mu & f^y\text{가 } \mathscr{A}/\mathscr{B}_1\text{−가측인 경우}\\ 0 & \text{ow.}\end{cases}$$

라 하면 이는 $\mathscr{B}/\mathscr{B}_1$−가측이다.

Proof 간결한 논의를 위해 i만 보이도록 하자. (ii는 i과 비슷하게 하면 된다.) 정리 2. 88로부터 적당한 $(\mathscr{A}\otimes\mathscr{B})/\mathscr{B}_1$−가측함수 $g:X\times Y\to\overline{\mathbb{R}}$가 존재하여 $f=g$ $((\mu\otimes\nu)$−ae.) 이고, WLOG, 필요하다면 $(\mu\otimes\nu)$−영집합에서의 g의 값을 0으로 바꾸어 g를 음이 아니라 해도 된다. 그렇다면 집합 $M=\{(x,y)\in X\times Y: f(x,y)\neq g(x,y)\}$에 대해 $(\mu\otimes\nu)(M)=0$이므로 정리 2. 122로부터 $\int_X \nu(M_x)\,d\mu(x)=(\mu\otimes\nu)(M)=0$이고, 정리 2. 101로부터 명백히 $\mathscr{A}$−가측인 집합 $N=\{x\in X:\nu(M_x)\neq 0\}$에 대해 $\mu(N)=0$이다. 이상을 종합하면 임의의 $x\in X\setminus N$에 대해 $[(X\times Y)\setminus M]_x=Y\setminus M_x$ 위에서 $f_x=g_x$이므로 곧 μ−거의 대부분의 $x\in X$

에 대해 $f_x = g_x$ (ν−ae.) 이고, 정리 2. 80와 보조정리 2. 159로부터 f_x는 μ−거의 대부분의 $x \in X$에 대해 $\mathscr{B}/\mathscr{B}_1$−가측이다. 나아가, 함수 $\widetilde{\varphi} : X \to \overline{\mathbb{R}}_0^+$를 $\widetilde{\varphi} : x \mapsto \int_Y g_x \, d\nu$로 두면 다시 보조정리 2. 159로부터 이는 $\mathscr{A}/\mathscr{B}_1$−가측이 되고, 따름정리 2. 102의 ii에서 $\varphi = \widetilde{\varphi}$ (μ−ae.) 이므로 정리 2. 80로부터 φ는 $\mathscr{A}/\mathscr{B}_1$−가측이다. □

Theorem 2.165 (Fubini) σ−유한 완비측도공간 $(X, \mathscr{A}, \mu), (Y, \mathscr{B}, \nu)$와 음이 아닌 $\overline{\mathscr{A} \otimes \mathscr{B}}/\mathscr{B}_1$−가측함수 $f : X \times Y \to \overline{\mathbb{R}}_0^+$에 대해

$$\int_{X \times Y} f(x, y) \, d\overline{\mu \otimes \nu}(x, y) = \int_X \left[\int_Y f(x, y) \, d\nu(y) \right] d\mu(x) = \int_Y \left[\int_X f(x, y) \, d\mu(x) \right] d\nu(y)$$

가 성립한다. 단, 만약 고정된 $x \in X$에 대해 $f(x, y)$가 $\mathscr{B}/\mathscr{B}_1$−가측이 아니라면 이때의 $\int_Y f(x, y) \, d\nu(y)$의 값은 0으로 간주하고, 이는 고정된 $y \in Y$에 대해 $f(x, y)$가 $\mathscr{A}/\mathscr{B}_1$−가측이 아닌 경우에도 마찬가지이다.

f

Proof 먼저 위의 보조정리로부터 이중적분들이 well-define됨을 안다. 이제 함수 $\widetilde{\varphi} : X \to \overline{\mathbb{R}}_0^+$와 $g : X \times Y \to \overline{\mathbb{R}}_0^+$를 위의 보조정리의 증명에서와 같이 두면 $\varphi = \widetilde{\varphi}$ (μ−ae.) 에서 Fubini의 정리와 따름정리 2. 102의 ii, 정리 2. 115로부터 $\int_X \varphi \, d\mu = \int_X \widetilde{\varphi} \, d\mu = \int_{X \times Y} g \, d(\mu \otimes \nu) = \int_{X \times Y} g \, d\overline{\mu \otimes \nu} = \int_{X \times Y} f \, d\overline{\mu \otimes \nu}$이다. 이제 두 번째 등식에 대해서도 이와 비슷하게 하면 된다. □

Lemma 2.166 σ−유한 완비측도공간 $(X, \mathscr{A}, \mu), (Y, \mathscr{B}, \nu)$와 $\overline{\mu \otimes \nu}$−적분가능한 함수 $f : X \times Y \to \overline{\mathbb{R}}$에 대해 다음이 성립한다.

i. Section $f_x : Y \to \overline{\mathbb{R}}$는 μ−거의 대부분의 $x \in X$에 대해 ν−적분가능하며 함수 $\varphi : X \to \mathbb{R}$를

$$\varphi : x \mapsto \begin{cases} \displaystyle\int_Y f_x \, d\nu & f_x\text{가 } \nu\text{−적분가능한 경우} \\ 0 & \text{ow.} \end{cases}$$

라 하면 이는 μ−적분가능하다.

ii. Section $f^y : X \to \overline{\mathbb{R}}$는 ν−거의 대부분의 $y \in Y$에 대해 μ−적분가능하며 함수 $\psi : Y \to \mathbb{R}$를

$$\psi : x \mapsto \begin{cases} \displaystyle\int_X f^y \, d\mu & f^y\text{가 } \mu\text{−적분가능한 경우} \\ 0 & \text{ow.} \end{cases}$$

라 하면 이는 ν−적분가능하다.

Proof 증명은 보조정리 2. 161와 거의 비슷하다. 이번에도 간결한 논의를 위해 i만 보이도록 하자. (ii는 i과 비슷하게 하면 된다.) 함수 $f_\pm$가 $\overline{\mathscr{A}\otimes\mathscr{B}}/\mathscr{B}_1$−가측이므로 함수 $\widetilde{\varphi}^\pm : X \to \overline{\mathbb{R}}_0^+$를

$$\widetilde{\varphi}^\pm : x \mapsto \begin{cases} \int_Y (f_\pm)_x \, d\nu & (f_\pm)_x \text{가 } \mathscr{B}/\mathscr{B}_1\text{−가측인 경우} \\ 0 & \text{ow.} \end{cases}$$

로 두면(여기서 위첨자로 쓰인 $\pm$은 정의 2. 97에서 아래첨자로 사용된 $\pm$와는 달리 함수의 양과 음의 부분을 의미하는 것이 아니다.) 보조정리 2. 164로부터 $(f_\pm)_x$ 중 하나라도 $\mathscr{B}/\mathscr{B}_1$−가측이 아닌 $x \in X$의 집합 M은 μ−영집합이고, $\widetilde{\varphi}^\pm$는 $\mathscr{A}/\mathscr{B}_1$−가측이 되어 곧 Fubini의 정리로부터 $\int_X \widetilde{\varphi}^\pm \, d\mu = \int_{X\times Y} f_\pm \, d\overline{\mu\otimes\nu} \leq \int_{X\times Y} |f| \, d\overline{\mu\otimes\nu} < \infty$에서 $\widetilde{\varphi}^\pm$이 μ−적분가능함을 안다. 그렇다면 $\widetilde{\varphi}^\pm$이 μ−거의 어디서나 유한하므로 명백히 $\mathscr{A}$−가측인 집합 $N = (\widetilde{\varphi}^\pm)^{-1}(\infty) = \{x \in X \setminus M : \int_Y |f_x| \, d\nu = \infty\}$에 대해 $\mu(N) = 0$이고, 곧 f_x는 μ−거의 대부분의 $x \in X$에 대해 ν−적분가능하다. 이제 함수 $\varphi^\pm : X \to \mathbb{R}_0^+$를 $\varphi^\pm = \widetilde{\varphi}^\pm \mathbf{1}_{X\setminus(M\cup N)}$으로 두면 $\varphi^\pm = \widetilde{\varphi}^\pm$ (μ−ae.)이므로 따름정리 2. 102의 ii로부터 $\varphi^\pm$은 μ−적분가능하고, 임의의 $x \in X$에 대해 $(f_x)_\pm = (f_\pm)_x$가 성립하여 $\varphi = \varphi^+ - \varphi^-$이므로 증명이 끝난다. □

Theorem 2.167 (Fubini) σ−유한 완비측도공간 $(X, \mathscr{A}, \mu), (Y, \mathscr{B}, \nu)$와 $\overline{\mu\otimes\nu}$−적분가능한 함수 $f : X \times Y \to \overline{\mathbb{R}}$에 대해

$$\int_{X\times Y} f(x,y) \, d\overline{\mu\otimes\nu}(x,y) = \int_X \left[\int_Y f(x,y) \, d\nu(y)\right] d\mu(x) = \int_Y \left[\int_X f(x,y) \, d\mu(x)\right] d\nu(y)$$

가 성립한다. 단, 만약 고정된 $x \in X$에 대해 $f(x,y)$가 ν−적분가능하지 않다면 이때의 $\int_Y f(x,y) \, d\nu(y)$의 값은 0으로 간주하고, 이는 고정된 $y \in Y$에 대해 $f(x,y)$가 μ−적분가능하지 않은 경우에도 마찬가지이다.

Proof 증명은 앞선 보인 Fubini의 정리와 거의 비슷하다. 먼저 위의 보조정리로부터 이중적분들이 well-define됨을 안다. 이제 함수 $\varphi : X \to \mathbb{R}$와 $\varphi^\pm : X \to \mathbb{R}_0^+$를 각각 위의 보조정리와 그 증명에서와 같이 두면 $\varphi = \varphi^+ - \varphi^-$에서 $\int_X \varphi \, d\mu = \int_X \varphi^+ \, d\mu - \int_X \varphi^- \, d\mu = \int_{X\times Y} f_+ \, d\overline{\mu\otimes\nu} - \int_{X\times Y} f_- \, d\overline{\mu\otimes\nu} = \int_{X\times Y} f \, d\overline{\mu\otimes\nu}$임을 쉽게 알 수 있다. 이제 두 번째 등식에 대해서도 이와 비슷하게 하면 된다. □

마침내, Lebesgue 측도공간에서 적용될 수 있는 일반적인 Fubini의 정리를 얻을 수 있다.

Corollary 2.168 $l + m = n$인 $l, m, n \in \mathbb{N}$과 음이 아닌 가측함수 $f : \mathbb{R}^n \to \overline{\mathbb{R}}_0^+$에 대해

$$\int_{\mathbb{R}^n} f(x,y) \, d\lambda_n(x,y) = \int_{\mathbb{R}^l} \left[\int_{\mathbb{R}^m} f(x,y) \, d\lambda_m(y)\right] d\lambda_l(x) = \int_{\mathbb{R}^m} \left[\int_{\mathbb{R}^l} f(x,y) \, d\lambda_l(x)\right] d\lambda_m(y)$$

가 성립한다. 한편, 적분가능한 $f:\mathbb{R}^n \to \overline{\mathbb{R}}$에 대해서도 동일한 결과가 성립한다. 단, 어느 경우에나 만약 고정된 $x \in \mathbb{R}^l$에 대해 $f(x,y)$가 가측이 아니거나 적분가능하지 않다면 이때의 $\int_{\mathbb{R}^m} f(x,y)\,d\lambda_m(y)$의 값은 0으로 간주하고, 이는 고정된 $y \in \mathbb{R}^m$에 대해 $f(x,y)$가 가측이 아니거나 적분가능하지 않은 경우에도 마찬가지이다.

Proof 이는 Fubini의 정리와 정리 2.68, 2.125의 ii로부터 자명하다. □

이렇게 얻은 일반적인 측도에 대한 Fubini의 정리를 사용하여 앞서 증명한 Minkowski의 부등식을 일반화하는 것으로 이번 절을 마무리한다. 이 결과는 이후 합성곱에 대해 공부할 때 유용하게 사용될 것이다.

Lemma 2.169 측도공간 $(X,\mathscr{A},\mu)$와 $1/p+1/q=1$인 $p,q>1$, 음이 아닌 가측함수 $f: X \to \overline{\mathbb{R}}_0^+$를 생각하자. 만약 적당한 $C \geq 0$가 존재하여 음이 아닌 임의의 $g \in L^q$에 대해 $\int_X fg\,d\mu \leq C||g||_q$를 만족하면 다음이 성립한다.

i. 만약 $||f||_p < \infty$이면 $||f||_p \leq C$이다.
ii. 측도 μ가 σ−유한하면 $||f||_p \leq C$이다.

Proof i. 만약 $||f||_p = 0$이면 정리가 자명하므로 $||f||_p > 0$이라 하자. 가정과 정리 2.100의 ii로부터 $N = f^{-1}(\infty)$는 영집합이므로 함수 $h: X \to \mathbb{R}_0^+$를 $h = f\mathbf{1}_{X\setminus N}$으로 두면 $f = h$ (ae.)이고 $||h||_p = ||f||_p < \infty$이며 임의의 음이 아닌 $g \in L^q$에 대해 $\int_X hg\,d\mu = \int_X fg\,d\mu \leq C||g||_q$이다. 이제 함수 $g: X \to \mathbb{R}_0^+$를 $g = h^{p-1}$로 두면 $g^q = h^{(p-1)q} = h^p = hg$에서 $||g||_q = (\int_X h^p\,d\mu)^{1-1/p} = ||h||_p^{p-1} < \infty$가 되어 $g \in L^q$이고 곧 $||h||_p^p = \int_X h^p\,d\mu = \int_X hg\,d\mu \leq C||g||_q = C||h||_p^{p-1}$이다. 그런데 앞서 $||h||_p = ||f||_p > 0$이라 가정하였으므로 이는 $||f||_p = ||h||_p \leq C$임을 함의하여 증명이 끝난다.

ii. 모순을 유도하기 위해 $||f||_p > C$라 하면 정의로부터 적당한 음이 아닌 단순함수 $h: X \to \mathbb{R}_0^+$가 존재하여 $h \leq f$이고 $||h||_p > C$이다. 여기서 $||h||_p < \infty$인 경우에는 $g \geq 0$인 임의의 $g \in L^q$에 대해 $\int_X hg\,d\mu \leq \int_X fg\,d\mu \leq C||g||_q$이므로 i로부터 $||h||_p \leq C$의 모순이 발생하여 증명이 끝나므로 $||h||_p = \infty$인 경우에 대해서만 조금 더 생각해보자. 이 경우에는 h의 표준형을 $h = \sum_{i=1}^l a_i \mathbf{1}_{A_i}$라 하면 적당한 $i_0 \leq k$에 대해 $a_{i_0} > 0$이고 $\mu(A_{i_0}) = \infty$이다. 한편, 가정으로부터 $\mathscr{A}$에 속하는 적당한 집합열 $\{B_j\}$가 존재하여 각 $j \in \mathbb{N}$에 대해 $\mu(B_j) < \infty$이고 $\bigcup_{j=1}^{\infty} B_j = X$이므로 집합열 $\{C_k\}$를 $C_k := A_{i_0} \cap \bigcup_{j=1}^{k} B_j$로 두면 이는 증가하는 집합열로서 A_{i_0}으로 수렴함이 분명하므로 정리 2.24의 i로부터 충분히 큰 $k_0 \in \mathbb{N}$에 대해 $(C/a_{i_0})^p < \mu(C_{k_0}) \leq \mu(\bigcup_{j=1}^{k_0} B_j) \leq \sum_{j=1}^{k_0} \mu(B_j) < \infty$이다. 따라서 단순함수 $\widetilde{h}: X \to \mathbb{R}_0^+$를 $\widetilde{h} = a_{i_0}\mathbf{1}_{C_{k_0}}$로 두면 $||\widetilde{h}||_p = a_{i_0}[\mu(C_{k_0})]^{1/p}$에서 $C < ||\widetilde{h}||_p < \infty$이고 $\widetilde{h} \leq h \leq f$이므로 앞서 $||h||_p < \infty$인 경우에 모순을 유도한 것과 비슷하게 $||\widetilde{h}||_p \leq C$의 모순을 유도할 수 있고, 증명은 이로써 충분하다. □

Theorem 2.170 (Minkowski's integral inequality) $\sigma-$유한 측도공간 $(X,\mathscr{A},\mu)$, $(Y,\mathscr{B},\nu)$에 대해 함수 $f:X\times Y\to\overline{\mathbb{R}}$가 $(\mathscr{A}\otimes\mathscr{B})/\mathscr{B}_1-$가측이라 하자. 그렇다면 $p\geq 1$에 대해

$$\left\{\int_X\left[\int_Y|f(x,y)|\,d\nu(y)\right]^p d\mu(x)\right\}^{1/p}\leq\int_Y\left[\int_X|f(x,y)|^p\,d\mu(x)\right]^{1/p}d\nu(y)$$

이다. 즉, $||\int_Y|f_x|\,d\nu||_p\leq\int_Y||f^y||_p\,d\nu$이다.

Proof 우선 Fubini의 정리로부터 위 부등식의 적분은 모두 well-define된다. 한편, $p=1$이면 부등식이 자명하여 더 이상 보일 것이 없으므로 $p>1$이라 하자. 또한, $\int_Y||f^y||_p\,d\nu=\infty$인 경우에도 부등식이 자명하므로 $C:=\int_Y||f^y||_p\,d\nu<\infty$라 하자. 이제 $1/p+1/q=1$인 $q>1$에 대해 음이 아닌 임의의 $g\in L^q(\mu)$를 택하고 함수 $\varphi:X\to\overline{\mathbb{R}}_0^+$를 $\varphi:x\mapsto\int_Y|f_x|\,d\nu$로 두면 이는 Fubini의 정리에서 $\mathscr{A}/\mathscr{B}_1-$가측이고 φg는 $(\mathscr{A}\otimes\mathscr{B})/\mathscr{B}_1-$가측이므로 다시 Fubini의 정리와 Hölder의 부등식으로부터 $\int_X\varphi g\,d\mu=\int_X[\int_Y|f(x,y)|g(x)\,d\nu(y)]\,d\mu(x)=\int_Y[\int_X|f(x,y)|g(x)\,d\mu(x)]\,d\nu(y)\leq\int_Y||f^y||_p||g||_q\,d\nu(y)=C||g||_q$이다. 그렇다면 위의 보조정리로부터 $||\varphi||_p\leq C$가 되어 증명이 끝난다. □

2.12 Transformation formula

이번 절에서는 변수변환 공식이라는 단 하나의 정리를 증명하도록 한다. 이렇게 하나의 정리를 증명하는데 하나의 절을 오롯이 할애하는 것에는 이전 절이 Fubini의 정리의 증명에 할애된 것과 같은 맥락의 이유와 더불어, 그 증명이 굉장히 길기 때문이다. (아마 이 책을 통틀어 가장 긴 증명이 아닐까 싶다.) 대부분의 독자들이 변수변환 공식이 어떤 공식인지, 언제 사용하는지 등에 대해 잘 알고 있을 것이고, 긴 증명과정과는 달리 이에 별다른 준비가 필요하지는 않으므로 이만 각설하고, 바로 증명을 시작하자. 증명은 2개의 보조정리의 도움을 받아 총 10단계로 이루어져 있다.

Lemma 2.171 집합 $N\subseteq\mathbb{R}^n$이 영집합일 필요충분조건은 임의의 $\varepsilon>0$에 대해 적당한 열린 ball의 열 $\{B(x_i,r_i)\}$가 존재하여 $N\subseteq\bigcup_{i=1}^\infty B(x_i,r_i)$이고 $\sum_{i=1}^\infty r_i^n<\varepsilon$인 것이다.

Proof 먼저 충분조건임을 보이기 위해 임의의 $\varepsilon>0$을 택하면 Lebesgue 측도가 regular 하므로 적당한 열린집합 $U\subseteq\mathbb{R}^n$가 존재하여 $N\subseteq U$이고 $\lambda_n(U)<(2/\sqrt{n})^n\varepsilon$이다. 또한, $\mathscr{S}_n$에 속하는 적당한 집합열 $\{B_i\}$가 존재하여 $U=\bigsqcup_{i=1}^\infty B_i$이고 각 $i\in\mathbb{N}$에 대해 B_i의 모든 모서리의 길이가 l_i로 같다.[22]그렇다면 $\sum_{i=1}^\infty l_i^n=\sum_{i=1}^\infty\lambda_n(B_i)=\lambda_n(\bigsqcup_{i=1}^\infty B_i)=\lambda_n(U)<(2/\sqrt{n})^n\varepsilon$이 성립하므로 각 $i\in\mathbb{N}$에 대해 x_i를 B_i의 중심이라 하여 $\{B(x_i,\sqrt{n}l_i/2)\}$로 정의된 열린 ball의 열을 생각하면 $\sum_{i=1}^\infty(\sqrt{n}l_i/2)^n<\varepsilon$이다. 이제 각 $i\in\mathbb{N}$에 대해 $B_i\subseteq B(x_i,\sqrt{n}l_i/2)$임은 분명하므로 $r_i=\sqrt{n}l_i/2$라 하면 충분조건임이 보여진다.

한편, 필요조건임은 거의 자명하다. 가정으로부터 임의의 $\varepsilon > 0$을 택하면 적당한 열린 ball의 열 $\{B(x_i, r_i)\}$가 존재하여 $\overline{N} \subseteq \bigcup_{i=1}^{\infty} \overline{B(x_i, r_i)}$이고 $\sum_{i=1}^{\infty} r_i^n < \varepsilon/4^n$인데, 각 $i \in \mathbb{N}$에 대해 $\overline{B(x_i, r_i)} \subseteq \prod_{i=1}^{n}(-2r_i, 2r_i] + x_i$이므로 $\lambda_n(\overline{B(x_i, r_i)}) \leq (4r_i)^n$에서 $\lambda_n(\overline{N}) \leq \lambda_n(\bigcup_{i=1}^{\infty} \overline{B(x_i, r_i)}) \leq \sum_{i=1}^{\infty} \lambda_n(\overline{B(x_i, r_i)}) \leq \sum_{i=1}^{\infty}(4r_i)^n < \varepsilon$이다. 이는 곧 $\lambda_n(\overline{N}) = 0$임을 뜻하므로 필요조건임도 보여진다. □

Lemma 2.172 영집합 $N \subseteq \mathbb{R}^n$과 Lipschitz 연속인 $f : \mathbb{R}^n \to \mathbb{R}^n$에 대해 $f(N)$도 영집합이다.

Proof 함수 f의 Lipschitz 상수를 $L > 0$이라 하자. 위의 보조정리로부터 임의의 $\varepsilon > 0$을 택하면 적당한 열린 ball의 열 $\{B(x_i, r_i)\}$가 존재하여 $N \subseteq \bigcup_{i=1}^{\infty} B(x_i, r_i)$이고 $\sum_{i=1}^{\infty} r_i^n < \varepsilon/L^n$이다. 한편, 각 $i \in \mathbb{N}$와 임의의 $x \in B(x_i, r_i)$에 대해 $||f(x) - f(x_i)|| < L||x - x_i|| < Lr_i$에서 $f(B(x_i, r_i)) \subseteq B(f(x_i), Lr_i)$이다. 이로부터 $f(N) \subseteq f(\bigcup_{i=1}^{\infty} B(x_i, r_i)) = \bigcup_{i=1}^{\infty} f(B(x_i, r_i)) \subseteq \bigcup_{i=1}^{\infty} B(f(x_i), Lr_i)$이고 $\sum_{i=1}^{\infty}(Lr_i)^n < \varepsilon$이므로 다시 위의 보조정리로부터 $f(N)$이 영집합임을 안다. □

Theorem 2.173 (Transformation formula) 열린집합 $U, V \subseteq \mathbb{R}^n$ 사이에서 정의된 $\mathscr{C}^1$급 미분동형사상 $\Phi : U \to V$와 음이 아닌 가측함수 $f : \mathbb{R}^n \to \overline{\mathbb{R}}_0^+$에 대해 합성 $f \circ \Phi : U \to \overline{\mathbb{R}}_0^+$는 가측이고 $\int_V f\, d\lambda_n = \int_U (f \circ \Phi)|\det \mathbf{D}\Phi|\, d\lambda_n$이다. 한편, V에서 적분가능한 (혹은 $(f \circ \Phi)|\det \mathbf{D}\Phi|$가 U에서 적분가능한) $f : \mathbb{R}^n \to \overline{\mathbb{R}}$에 대해서도 동일한 결과가 성립하며, 이때 $(f \circ \Phi)|\det \mathbf{D}\Phi|$는 U에서 적분가능하다 (혹은 f는 V에서 적분가능하다).[23]

Proof 증명은 총 10단계로 구성되어 있다. 증명이 상당히 짜증날 것이므로 인내심을 갖길 바란다. (대략 5~6단계를 전후하여 고비가 온다.) 이제부터 임의의 음이 아닌 가측함수 f에 대해 위의 정리를 만족하는 미분동형사상 Φ를 good이라 하자.

Step 1. 미분동형사상 $\Phi : U \to V$가 good일 필요충분조건은 임의의 가측인 $A \subseteq V$에 대해 $\Phi^{-1}(A)$도 가측이고 $\lambda_n(A) = \int_{\Phi^{-1}(A)} |\det \mathbf{D}\Phi|\, d\lambda_n$인 것이다.

먼저 충분조건임을 보이기 위해 임의의 가측집합 $A \subseteq V$를 고정하면 $\mathbf{1}_A$가 가측이고 가정으로부터 Φ가 good이므로 $\mathbf{1}_{\Phi^{-1}(A)} = \mathbf{1}_A \circ \Phi$가 가측이 되어 곧 $\Phi^{-1}(A)$가 가측이고, $\lambda_n(A) = \int_V \mathbf{1}_A\, d\lambda_n = \int_U (\mathbf{1}_A \circ \Phi)|\det \mathbf{D}\Phi|\, d\lambda_n = \int_{\Phi^{-1}(A)} |\det \mathbf{D}\Phi|\, d\lambda_n$이다. 이제 필요조건임을 보이기 위해 먼저 f가 음이 아닌 단순함수인 경우를 생각하여 단순함수 $f\mathbf{1}_V$의 표준형을 $f\mathbf{1}_V = \sum_{i=1}^{k} a_i \mathbf{1}_{A_i}$라 하면 $f \circ \Phi = (f\mathbf{1}_V) \circ \Phi = \sum_{i=1}^{k} a_i \mathbf{1}_{\Phi^{-1}(A_i)}$가 가측이고

$$\begin{aligned}\int_V f\, d\lambda_n &= \sum_{i=1}^{k} a_i \lambda_n(A_i) \\ &= \sum_{i=1}^{k} a_i \int_{\Phi^{-1}(A_i)} |\det \mathbf{D}\Phi|\, d\lambda_n \\ &= \int_U \sum_{i=1}^{k} a_i \mathbf{1}_{\Phi^{-1}(A_i)} |\det \mathbf{D}\Phi|\, d\lambda_n\end{aligned}$$

$$= \int_U \sum_{i=1}^{k} a_i(\mathbf{1}_{A_i} \circ \Phi)|\det \mathbf{D}\Phi| d\lambda_n$$
$$= \int_U (f \circ \Phi)|\det \mathbf{D}\Phi| d\lambda_n$$

이 성립한다. 다음으로, 일반적인 음이 아닌 가측함수 f를 생각하면 정리 2.83로부터 음이 아닌 단순함수 $f_i : \mathbb{R}^n \to \overline{\mathbb{R}}_0^+$ 의 열 $\{f_i\}$에 대해 $f_i \uparrow f$ 인데, 각 $i \in \mathbb{N}$에 대해 앞선 결과로부터 $f_i \circ \Phi$가 가측이므로 정리 2.77의 ii로부터 $f \circ \Phi$도 가측이다. 또한 MCT로부터 $\int_V f\, d\lambda_n = \lim_{i\to\infty} \int_V f_i\, d\lambda_n = \lim_{i\to\infty} \int_U (f_i \circ \Phi)|\det \mathbf{D}\Phi|\, d\lambda_n = \int_U (f \circ \Phi)|\det \mathbf{D}\Phi|\, d\lambda_n$ 도 성립하므로 Φ가 good임을 안다.

Step 2. 열린집합 $U, V, W \subseteq \mathbb{R}^n$ 에 대해 미분동형사상 $\Phi : U \to V, \Psi : V \to W$ 가 good이라면 합성 $\Psi \circ \Phi : U \to W$ 도 good이다.

임의의 음이 아닌 가측함수 f에 대해 가정으로부터 $f \circ (\Psi \circ \Phi)$가 가측임은 자명하다. 또한 $\int_W f\, d\lambda_n = \int_V (f \circ \Psi)|\det \mathbf{D}\Psi|\, d\lambda_n$ 인데, 여기서 $g := (f \circ \Psi)|\det \mathbf{D}\Psi|$도 음이 아닌 가측함수이므로 다시 가정으로부터 $\int_V g\, d\lambda_n = \int_U (g \circ \Phi)|\det \mathbf{D}\Phi|\, d\lambda_n$ 이 되어 이상을 종합하면

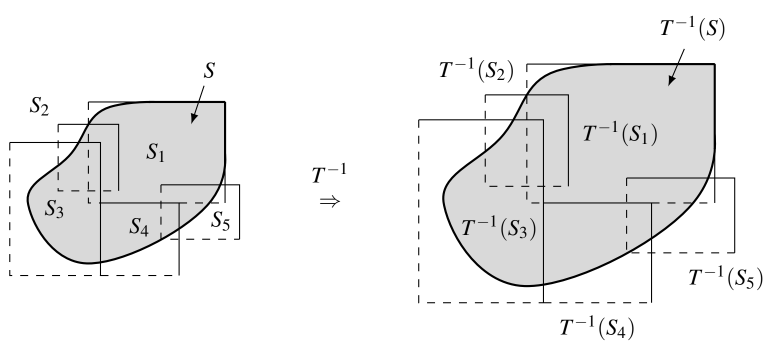
$$\int_W f\, d\lambda_n = \int_U (f \circ \Psi \circ \Phi)|\det \mathbf{D}\Psi \circ \Phi||\det \mathbf{D}\Phi| d\lambda_n$$
$$= \int_U [f \circ (\Psi \circ \Phi)]|\det(\mathbf{D}\Psi \circ \Phi)\mathbf{D}\Phi| d\lambda_n$$
$$= \int_U [f \circ (\Psi \circ \Phi)]|\det \mathbf{D}(\Psi \circ \Phi)| d\lambda_n$$

에서 $\Psi \circ \Phi$가 good임을 안다.

Step 3. 모든 평행이동 $T : \mathbb{R}^n \to \mathbb{R}^n$ 는 good이다.

평행이동인 T 에 대해 $|\det \mathbf{D}T| = 1$ 이고 Lebesgue 측도가 이동 불변성을 가지므로 임의의 가측인 $A \subseteq \mathbb{R}^n$ 에 대해 $T^{-1}(A)$ 가 가측이며 $\lambda_n(A) = \lambda_n(T^{-1}(A)) = \int_{T^{-1}(A)} |\det \mathbf{D}T|\, d\lambda_n$ 이므로 1단계에서 T 는 good임을 안다.

Step 4. 모든 homothety $T : \mathbb{R}^n \to \mathbb{R}^n$ 는 good이다.



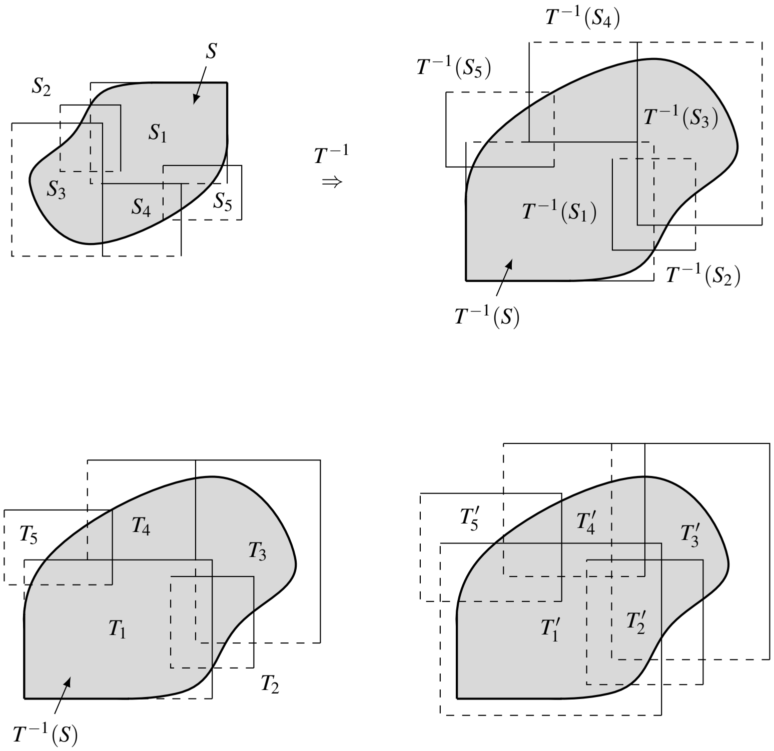


Figure 2.13 변수변환 공식의 증명의 4단계에서의 집합 S와 임의의 가산 덮개 $\{S_i\}$가 homothety T에 의해 변환되는 모습. 각각 $c>0$인 경우(위)와 $c<0$인 경우(중간). 한편, $c<0$인 경우에는 $\{T_i\}$가 $T^{-1}(S)$의 가산 덮개가 되지 못할 수도 있으므로(왼쪽 아래) 각 T_i를 조금씩 늘려 $\{T'_i\}$를 구성한다(오른쪽 아래).

Homothety인 T는 적당한 0이 아닌 $c\in\mathbb{R}$에 대해 $T: x\mapsto cx$로 쓸 수 있다. 이번에도 단계 1로써 T가 good임을 보이도록 하자. 먼저 임의의 semi-open box $B\subseteq\mathbb{R}^n$를 생각하면 T가 연속이므로 곧 가측이고, 따라서 $T^{-1}(B)$도 가측이다. 이제 $\lambda_n(T^{-1}(B))=\rho_n(B)/|c|^n$임을 보이고자 하는데, 만약 B가 유계가 아니라면 $T^{-1}(B)$도 유계가 아니어서 이가 자명하므로 B가 유계라 하자. 그렇다면 적당한 $x,y\in\mathbb{R}^n$가 존재하여 $B=\prod_{i=1}^n(x_i,y_i]$이고, 곧 보조정리 2.127로부터 $\lambda_n(T^{-1}(B))=\lambda_n(B/c)=\prod_{i=1}^n(y_i-x_i)/|c|=\rho_n(B)/|c|^n$이다.[24] 나아가, 정리 2.16로부터 임의의 $A\in A(\mathscr{S}_n)$에 대해서도 $T^{-1}(A)$가 가측이고 $\lambda_n(T^{-1}(A))=\rho_n(A)/|c|^n$임이 자명하다.

마지막으로, 임의의 $S\subseteq\mathbb{R}^n$를 택하고 이의 $A(\mathscr{S}_n)$에 속하는 임의의 가산 덮개 $\{S_i\}$를 생각하여, WLOG, 필요하다면 공집합을 추가하여 이가 집합열을 이룬다고 하자. 만약 $c>0$이라면 $\{T^{-1}(S_i)\}$가 $A(\mathscr{S}_n)$에 속하는 $T^{-1}(S)$의 가산 덮개임이 분명하므로 앞선 결과로부터 $\rho_n^*(T^{-1}(S))=\rho_n^*(S)/|c|^n$임이 분명하다. 한편, $c<0$인 경우에는 조금 복잡하다.

먼저 각 $i \in \mathbb{N}$에 대해 $T^{-1}(S_i)$를 구성하는 서로소인 box들에서 원래 속한 면은 빼고, 원래 빠진 면은 더하여 얻는 집합을 T_i라 하면 이는 다시 semi-open box들의 서로소 합집합이 되고, 앞선 결과와 보조정리 2.127로부터 $\rho_n(T_i) = \lambda_n(T^{-1}(S_i)) = \rho_n(S_i)/|c|^n$이다. (각주 24 참조.) 그러나 원래 속한 면을 빼는 과정에서 $\{T_i\}$가 더 이상 $T^{-1}(S)$의 가산 덮개가 되지 못할 수 있다. 이 문제를 해결하기 위해 임의의 $\varepsilon > 0$을 택하여 각 $i \in \mathbb{N}$에 대해 T_i를 이루는 semi-open box들을 조금씩 늘려 얻은 T_i'가 $\rho_n(T_i') = \rho_n(T_i) + \varepsilon/2^i$가 되도록 하면 이때의 $\{T_i'\}$는 다시 $T^{-1}(S)$의 가산 덮개가 되면서 $\sum_{i=1}^{\infty} \rho_n(T_i') = \rho_n(S_i)/|c|^n + \varepsilon$을 만족한다. 이로부터 $\rho_n^*(T^{-1}(S)) \le \rho_n(S_i)/|c|^n + \varepsilon$이고, 곧 $\rho_n^*(T^{-1}(S)) \le \rho_n^*(S)/|c|^n$이다. 그런데 여기서 T는 임의의 homothety이고, T^{-1}도 $T^{-1} : x \mapsto x/c$인 homothety이므로 $\rho_n^*(T(S))/|c|^n \le \rho_n^*(S)$에서 $\rho_n^*(S)/|c|^n \le \rho_n^*(T^{-1}(S))$가 되어 이상으로부터 c의 부호에 무관하게 $\rho_n^*(T^{-1}(S)) = \rho_n^*(S)/|c|^n$임을 안다. 이러한 결론이 임의의 $S \subseteq \mathbb{R}^n$에 대해 성립한다는 점을 상기하면, 임의의 가측인 $A \subseteq \mathbb{R}^n$와 임의의 $S \subseteq \mathbb{R}^n$에 대해

$$\begin{aligned}
\rho_n^*(S \cap T^{-1}(A)) + \rho_n^*(S \cap [T^{-1}(A)]^c) &= \rho_n^*(S \cap T^{-1}(A)) + \rho_n^*(S \cap T^{-1}(A^c)) \\
&= \frac{\rho_n^*(T(S) \cap A) + \rho_n^*(T(S) \cap A^c)}{|c|^n} \\
&= \frac{\rho_n^*(T(S))}{|c|^n} \\
&= \rho_n^*(S)
\end{aligned}$$

이므로 $T^{-1}(A)$도 가측이며, 앞선 결론으로부터 $\lambda_n(T^{-1}(A)) = \lambda_n(A)/|c|^n$이다. 이제 $\lambda_n(A) = |c|^n \lambda_n(T^{-1}(A)) = \int_{T^{-1}(A)} |\det \mathbf{D}T| \, d\lambda_n$이므로 단계 1로부터 T가 good임을 안다.

Step 5. 모든 가역인 선형사상 $T : \mathbb{R}^n \to \mathbb{R}^n$는 good이다.

모든 가역인 선형사상은 (표준기저에 대해) 기본행렬로 나타내어지는 유한개의 선형사상의 합성으로 쓸 수 있으므로 2단계로부터 T가 다음 3가지와 같은 경우에 대해서만 생각하면 된다.

i. $T : (\cdots, x_i, \cdots) \mapsto (\cdots, cx_i, \cdots)$ (단, $c \neq 0$이다.)
ii. $T : (\cdots, x_i, \cdots) \mapsto (\cdots, x_i + cx_j, \cdots)$
iii. $T : (\cdots, x_i, \cdots, x_j, \cdots) \mapsto (\cdots, x_j, \cdots, x_i, \cdots)$

간결한 논의를 위해 $i = 1$, $j = 2$인 경우에 대해서만 보이도록 하고, 이하 $f : \mathbb{R}^n \to \overline{\mathbb{R}}_0^+$를 임의의 가측함수라 하자. (다른 i, j에 대해서도 이와 비슷하게 하면 된다.) 우선 $f \circ T$가 가측임을 보이려고 하는데, 만약 f가 Borel이면 정리 2.73로부터 $f \circ T$가 가측임이 분명하다. 한편, 정리 2.68와 2.88로부터 적당한 Borel 함수 $g : \mathbb{R}^n \to \overline{\mathbb{R}}$가 존재하여 $f = g$ (ae.)이다. 이제 $N = \{x \in \mathbb{R}^n : f(x) \neq g(x)\}$라 하면 이가 영집합이고 T^{-1}가 Lipschitz 연속이므로 위의 보조정리로부터 $T^{-1}(N) = \{x \in \mathbb{R}^n : (f \circ T)(x) \neq (g \circ T)(x)\}$가 영집합이

되어 $f \circ T = g \circ T$ (ae.)이다. 그렇다면 Lebesgue 측도공간이 완비성을 가진다는 사실과 정리 2.80로부터 $f \circ T$ 도 가측이다.

다음으로, 위의 세 가지 경우에 대해 $\int_{\mathbb{R}^n} f\,d\lambda_n = \int_{\mathbb{R}^n} (f \circ T)|\det \mathbf{D}T|\,d\lambda_n$ 임을 보이자. 먼저 첫 번째 경우를 보면 Fubini의 정리와 4단계의 결과로부터

$$\begin{aligned}
\int_{\mathbb{R}^n} f\,d\lambda_n &= \int_{\mathbb{R}} \cdots \int_{\mathbb{R}} f(x_1, \cdots, x_n)\,d\lambda_1(x_1) \cdots d\lambda_1(x_n) \\
&= \int_{\mathbb{R}} \cdots \int_{\mathbb{R}} f(cx_1, \cdots, x_n)|c|\,d\lambda_1(x_1) \cdots d\lambda_1(x_n) \\
&= \int_{\mathbb{R}} \cdots \int_{\mathbb{R}} (f \circ T)(x_1, \cdots, x_n)|c|\,d\lambda_1(x_1) \cdots d\lambda_1(x_n) \\
&= \int_{\mathbb{R}^n} (f \circ T)|\det \mathbf{D}T|\,d\lambda_n
\end{aligned}$$

이다. 두 번째 경우에도 Fubini의 정리와 3단계의 결과로부터

$$\begin{aligned}
\int_{\mathbb{R}^n} f\,d\lambda_n &= \int_{\mathbb{R}} \cdots \int_{\mathbb{R}} f(x_1, \cdots, x_n)\,d\lambda_1(x_1) \cdots d\lambda_1(x_n) \\
&= \int_{\mathbb{R}} \cdots \int_{\mathbb{R}} f(x_1 + cx_2, \cdots, x_n)\,d\lambda_1(x_1) \cdots d\lambda_1(x_n) \\
&= \int_{\mathbb{R}} \cdots \int_{\mathbb{R}} (f \circ T)(x_1, \cdots, x_n)\,d\lambda_1(x_1) \cdots d\lambda_1(x_n) \\
&= \int_{\mathbb{R}^n} (f \circ T)|\det \mathbf{D}T|\,d\lambda_n
\end{aligned}$$

이며 마지막 경우에도 Fubini의 정리로부터

$$\begin{aligned}
\int_{\mathbb{R}^n} f\,d\lambda_n &= \int_{\mathbb{R}} \cdots \int_{\mathbb{R}} f(x_1, x_2, \cdots, x_n)\,d\lambda_1(x_2)\,d\lambda_1(x_1) \cdots d\lambda_1(x_n) \\
&= \int_{\mathbb{R}} \cdots \int_{\mathbb{R}} f(x_2, x_1, \cdots, x_n)\,d\lambda_1(x_1)\,d\lambda_1(x_2) \cdots d\lambda_1(x_n) \\
&= \int_{\mathbb{R}} \cdots \int_{\mathbb{R}} (f \circ T)(x_1, \cdots, x_n)\,d\lambda_1(x_1)\,d\lambda_1(x_2) \cdots d\lambda_1(x_n) \\
&= \int_{\mathbb{R}^n} (f \circ T)|\det \mathbf{D}T|\,d\lambda_n
\end{aligned}$$

이다. 이상으로부터 어느 경우에나 T 가 good이 되어 임의의 가역인 선형사상이 good임을 안다.

2단계, 3단계, 5단계의 결과를 종합하면 임의의 Affine 변환이 good임을 안다. 이제 증명의 후반부를 위해 임의의 $\mathscr{C}^1$ 급 미분동형사상 $\Phi : U \to V$ 와 임의의 compact한 $K \subseteq U$ 를 고정하고, $\mathbb{R}^n$ 에 최대값 노름 $||\cdot||_\infty$ 를 장착하자.

Step 6. 임의의 $\varepsilon > 0$에 대해 적당한 $\delta > 0$가 존재하여 임의의 $x, y \in K$가 $||x-y||_\infty < \delta$와 $\prod_{i=1}^n [x_i, y_i] \subseteq K$를 만족하면 $||\Phi(x) - \Phi(y) - \mathbf{D}\Phi(x)(y-x)||_\infty < \varepsilon ||x-y||_\infty$ 이다.

편의를 위해 $\prod_{i=1}^{n}[x_i, y_i] \subseteq K$ 를 만족하는 임의의 $x, y \in K$ 를 고정하자. 이제 함수 $\gamma : [0,1] \to \mathbb{R}^n$ 를 $\gamma : t \mapsto \Phi((1-t)x+ty)$ 로 두면 이는 $\mathscr{C}^1$ 급이고, FTC로부터 $\Phi(y)-\Phi(x) = \int_0^1 \nabla\gamma = \int_0^1 \mathbf{D}\Phi((1-t)x+ty)(y-x)\,dt$ 이다. 이로부터 함수 $\Gamma : [0,1] \to \mathsf{L}(\mathbb{R}^n, \mathbb{R}^n)$ 를 $\Gamma : t \mapsto \mathbf{D}\Phi((1-t)x+ty) - \mathbf{D}\Phi(x)$ 로 두면 $\Phi(y)-\Phi(x) = \int_0^1 \mathbf{D}\Phi((1-t)x+ty)(y-x)\,dt = \int_0^1 [\mathbf{D}\Phi(x)+\Gamma(t)](y-x)\,dt = \mathbf{D}\Phi(x)(y-x) + \int_0^1 \Gamma(t)(y-x)\,dt$ 이다. 한편, $\mathbf{D}\Phi : \mathbb{R}^n \to \mathsf{L}(\mathbb{R}^n, \mathbb{R}^n)$ 가 K 위에서 균등연속이므로 임의의 $\varepsilon > 0$ 을 택하면 적당한 $\delta > 0$ 가 존재하여 $||x-y||_\infty < \delta$ 이면 $||\mathbf{D}\Phi(x) - \mathbf{D}\Phi(y)||_{\text{op}} < \varepsilon$ 이다. 이제 $||x-y||_\infty < \delta$ 라 하고 임의의 $t \in [0,1]$ 에 대해 $z = (1-t)x+ty$ 라 하면 $||x-z||_\infty \le ||x-y||_\infty < \delta$ 에서 $||\Gamma(t)||_{\text{op}} = ||\mathbf{D}\Phi(x) - \mathbf{D}\Phi(z)||_{\text{op}} < \varepsilon$ 이 되어 이상으로부터 각 $i \le n$ 에 대해 $\int_0^1 \Gamma_i(t)(y-x)\,dt \le \int_0^1 ||\Gamma(t)(y-x)||_\infty\,dt \le \int_0^1 ||\Gamma(t)||_{\text{op}}||x-y||_\infty\,dt < \varepsilon||x-y||_\infty$ 이므로 곧 $||\Phi(x)-\Phi(y)-\mathbf{D}\Phi(x)(y-x)||_\infty = ||\int_0^1 \Gamma(t)(y-x)\,dt||_\infty < \varepsilon||x-y||_\infty$ 가 성립한다.

이제 임의의 $x \in U$ 에 대해 함수 $\Psi_x : \mathbb{R}^n \to \mathbb{R}^n$ 를 $\Psi_x : y \mapsto \Phi(x) + \mathbf{D}\Phi(x)(x-y)$ 로 두면 이는 Affine 변환이며 곧 good이다. 나아가 6단계의 결과로부터 임의의 $\varepsilon > 0$ 에 대해 적당한 $\delta > 0$ 가 존재하여 임의의 $x, y \in K$ 가 $||x-y||_\infty < \delta$ 와 $\prod_{i=1}^{n}[x_i, y_i] \subseteq K$ 를 만족하면 $||\Phi(y) - \Psi_x(y)||_\infty < \varepsilon||x-y||_\infty$ 이다.

Step 7. 임의의 $0 < \varepsilon < 1$ 에 대해 적당한 $\delta > 0$ 가 존재하여 임의의 열린 ball $B(x, r) \subseteq K$ 이 $r < \delta$ 를 만족하면 $\Psi_x(B(x, (1-\varepsilon)r)) \subseteq \Phi(B(x, r)) \subseteq \Psi_x(B(x, (1+\varepsilon)r))$ 이다.

미분동형사상 Φ 가 $\mathscr{C}^1$ 급이고 집합 K 가 compact하여 $(\mathbf{D}\Phi(\cdot))^{-1} : K \to \mathsf{L}(\mathbb{R}^n, \mathbb{R}^n)$ 가 유계이므로 $M := \sup_{x\in K} ||(\mathbf{D}\Phi(x))^{-1}||_{\text{op}} < \infty$ 이다. 그렇다면 임의의 $x \in K$ 와 임의의 $y, z \in \mathbb{R}^n$ 에 대해 $\Psi_x^{-1}(y) = (\mathbf{D}\Phi(x))^{-1}(\Phi(x) - y) + x$ 이고, 이는 z 에 대해서도 마찬가지이므로 $||\Psi_x^{-1}(y) - \Psi_x^{-1}(z)||_\infty = ||(\mathbf{D}\Phi(x))^{-1}(y-z)||_\infty \le ||(\mathbf{D}\Phi(x))^{-1}||_{\text{op}}||y-z||_\infty \le M||y-z||_\infty$ 이다. 이제 임의의 $x \in K^\circ$ 를 고정하고 임의의 $0 < \varepsilon < 1$ 을 택하면 적당한 $\delta > 0$ 가 존재하여 임의의 $y \in K$ 가 $||x-y||_\infty < \delta$ 와 $\prod_{i=1}^{n}[x_i, y_i] \subseteq K$ 를 만족하면 $||\Phi(y) - \Psi_x(y)||_\infty < [\varepsilon/(M+1)]||x-y||_\infty$ 이다. 이러한 δ 에 대해 임의의 $r < \delta$ 도 고정하되, WLOG, 필요하다면 r 을 더 작게 하여 $B(x, r) \subseteq K$ 라 하면 임의의 $y \in B(x, r)$ 에 대해 $||x-y||_\infty < r < \delta$ 이고 $\prod_{i=1}^{n}[x_i, y_i] \subseteq B(x, r) \subseteq K$ 가 되어 $||\Phi(y) - \Psi_x(y)||_\infty < [\varepsilon/(M+1)]||x-y||_\infty < \varepsilon r/(M+1)$ 이다. 이상을 종합하면 임의의 $y \in B(x, r)$ 에 대해 $||(\Psi_x^{-1} \circ \Phi)(y) - y||_\infty = ||(\Psi_x^{-1} \circ \Phi)(y) - (\Psi_x^{-1} \circ \Psi_x)(y)||_\infty \le M||\Phi(y) - \Psi_x(y)||_\infty < \varepsilon r$ 이 성립하여 $||(\Psi_x^{-1} \circ \Phi)(y) - x||_\infty \le ||(\Psi_x^{-1} \circ \Phi)(y) - y||_\infty + ||x-y||_\infty < (1+\varepsilon)r$ 이므로 $\Phi(y) \in \Psi_x(B(x, (1+\varepsilon)r))$ 이고, 곧 $\Phi(B(x, r)) \subseteq \Psi_x(B(x, (1+\varepsilon)r))$ 임을 안다.

다른 포함관계를 보이기 위해 $x \in K$ 와 $r > 0$, $0 < \varepsilon < 1$ 을 위에서와 같이 두고 임의의 $y \in S(x, r)$ 를 생각하면 $||x-y||_\infty = r$ 이므로 $||(\Psi_x^{-1} \circ \Phi)(y) - x||_\infty \ge ||x-y||_\infty - ||(\Psi_x^{-1} \circ \Phi)(y) - y||_\infty > (1-\varepsilon)r$ 에서 $\Phi(y) \notin \Psi_x(B(x, (1-\varepsilon)r))$ 이다. 이는 곧 $\Psi_x(B(x, (1-\varepsilon)r))$ 과 $\Phi(S(x, r))$ 이 서로소임을 뜻하므로 집합 $A = \Psi_x(B(x, (1-\varepsilon)r)) \cap \Phi(B(x, r))$, $B = \Psi_x(B(x, (1-\varepsilon)r)) \setminus \Phi(B(x, r))$ 를 생각하면 $A \sqcup B = \Psi_x(B(x, (1-\varepsilon)r))$ 이며 A 는 열린집합임이 자명하고 $B =$

$\Psi_x(B(x,\,(1-\varepsilon)r))\setminus[\Phi(B(x,\,r))\cup\Phi(S(x,\,r))]=\Psi_x(B(x,\,(1-\varepsilon)r))\setminus\Phi(\overline{B(x,\,r)})$ 에서 B 도 열린집합이다.[25] 그러나 $\Psi_x(B(x,\,(1-\varepsilon)r))$ 이 연결집합이므로 A,B 중 하나는 공집합인데, $\Psi_x(x)=\Phi(x)\in A$ 이므로 $B=\emptyset$ 이 되어 $A=\Psi_x(B(x,\,(1-\varepsilon)r))$ 이고, 곧 $\Psi_x(B(x,\,(1-\varepsilon)r))\subseteq\Phi(B(x,\,r))$ 이 되어 7단계의 증명이 끝난다.

Step 8. 가측집합 $A\subseteq U$ 에 대해 $\Phi(A)$ 도 가측이다.

함수 Φ 가 미분동형사상이므로 Φ^{-1} 가 연속이고, 곧 가측이므로 임의의 Borel 집합 $B\subseteq U$ 에 대해 $\Phi(B)=(\Phi^{-1})^{-1}(B)$ 가 가측이다. 한편, 정리 2. 66, 2. 68로부터 임의의 가측집합 $A\subseteq U$ 에 대해 적당한 Borel 집합 $B,C\subseteq\mathbb{R}^n$ 가 존재하여 $B\subseteq A\subseteq C$ 이며 $C\setminus B$ 가 영집합이고, $C\setminus A\subseteq C\setminus B$ 이므로 $N:=C\setminus A$ 도 영집합이다. 나아가 WLOG, 필요하다면 B,C 를 각각 $B\cap U,C\cap U$ 로 바꾸어 $B,C\subseteq U$ 라 해도 된다. 그렇다면 $\Phi(A)=\Phi(C)\setminus\Phi(N)$ 에서 $\Phi(C)$ 는 앞선 결과로부터 가측이고, $\Phi(N)$ 은 위의 보조정리로부터 영집합이 되어 Lebesgue 측도공간이 완비성을 가진다는 점에서 곧 가측이므로 $\Phi(A)$ 는 가측이다.

Step 9. 열린집합 $W\subseteq U$ 에 대해 $\lambda_n(\Phi(W))=\int_W|\det\mathbf{D}\Phi|\,d\lambda_n$ 이다.

먼저 $W\subseteq K$ 인 특별한 경우를 생각하자. 임의의 $\varepsilon>0$ 을 택하고 $k\geq 1$ 를 충분히 크게 잡으면 $\log|\det\mathbf{D}\Phi(\cdot)|:K\to\mathbb{R}$ 가 균등연속이므로 적당한 $\delta>0$ 가 존재하여 임의의 $x,y\in K$ 가 $\|x-y\|_\infty<\delta$ 를 만족하면 $|\log|\det\mathbf{D}\Phi(x)/\det\mathbf{D}\Phi(y)||=|\log|\det\mathbf{D}\Phi(x)|-\log|\det\mathbf{D}\Phi(y)||<\varepsilon/k$ 이고 곧 $1-\varepsilon\leq 1-\varepsilon/k\leq e^{-\varepsilon/k}<|\det\mathbf{D}\Phi(x)/\det\mathbf{D}\Phi(y)|<e^{\varepsilon/k}\leq 1+\varepsilon$ 이다. 또한, 7단계의 결과로부터 WLOG, 필요하다면 $\delta>0$ 를 더 작게 하여 임의의 열린 ball $B(x,r)\subseteq K$ 이 $r<\delta$ 를 만족하면 $\Psi_x(B(x,\,(1-\varepsilon)r))\subseteq\Phi(B(x,\,r))\subseteq\Psi_x(B(x,\,(1+\varepsilon)r))$ 이도록 할 수 있다. 나아가, $\mathscr{P}(W)$ 에 속하는 적당한 서로소인 열린 ball의 열 $\{B(x_i,r_i)\}$ 가 존재하여 각 $i\in\mathbb{N}$ 에 대해 $r_i<\delta$ 이고 $\lambda_n(W\setminus\bigsqcup_{i=1}^\infty B(x_i,r_i))=0$ 이며,[26] 위의 보조정리로부터 $\lambda_n(\Phi(W)\setminus\bigsqcup_{i=1}^\infty\Phi(B(x_i,r_i)))=\lambda_n(\Phi(W\setminus\bigsqcup_{i=1}^\infty B(x_i,r_i)))=0$ 이다. 이상을 종합하면 Affine 변환이 good이라는 점과 따름정리 2. 102, 정리 2. 112로부터 각 $i\in\mathbb{N}$ 에 대해

$$\begin{aligned}\lambda_n(\Psi_{x_i}(B(x_i,\,(1+\varepsilon)r_i)))&=\int_{B(x_i,\,(1+\varepsilon)r_i)}|\det\mathbf{D}\Psi_{x_i}|\,d\lambda_n\\&=|\det\mathbf{D}\Phi(x_i)|\lambda_n(B(x_i,\,(1+\varepsilon)r_i))\\&=|\det\mathbf{D}\Phi(x_i)|(1+\varepsilon)^n\lambda_n(B(x_i,\,r_i))\\&\leq(1+\varepsilon)^{n+1}\int_{B(x_i,r_i)}|\det\mathbf{D}\Phi|\,d\lambda_n\end{aligned}$$

이고, 이로부터

$$\begin{aligned}\lambda_n(\Phi(W))&=\lambda_n\left(\bigsqcup_{i=1}^\infty\Phi(B(x_i,\,r_i))\right)\\&=\sum_{i=1}^\infty\lambda_n(\Phi(B(x_i,\,r_i)))\end{aligned}$$

$$\begin{aligned}
&\le \sum_{i=1}^{\infty} \lambda_n(\Psi_{x_i}(B(x_i,(1+\varepsilon)r_i))) \\
&\le (1+\varepsilon)^{n+1} \sum_{i=1}^{\infty} \int_{B(x_i,r_i)} |\det \mathbf{D}\Phi|\, d\lambda_n \\
&= (1+\varepsilon)^{n+1} \int_{\bigsqcup_{i=1}^{\infty} B(x_i,r_i)} |\det \mathbf{D}\Phi|\, d\lambda_n \\
&= (1+\varepsilon)^{n+1} \int_W |\det \mathbf{D}\Phi|\, d\lambda_n
\end{aligned}$$

이다. 이제 $(1-\varepsilon)^{n+1} \int_W |\det \mathbf{D}\Phi|\, d\lambda_n \le \lambda_n(\Phi(W))$ 임을 이와 비슷하게 보일 수 있으므로 곧 $\lambda_n(\Phi(W)) = \int_W |\det \mathbf{D}\Phi|\, d\lambda_n$ 임을 안다.

이제 일반적인 $W \subseteq K$에 대해 $\mathscr{P}(U)$에 속하는 적당한 서로소인 열린 ball의 열 $\{B_i\}$가 존재하여 각 $i \in \mathbb{N}$에 대해 $\overline{B_i} \subseteq U$ 이고 $\lambda_n(U \setminus \bigsqcup_{i=1}^{\infty} B_i) = 0$ 이다. [27]따라서 위의 보조정리로부터 $\lambda_n(\Phi(W) \setminus \bigsqcup_{i=1}^{\infty} \Phi(W \cap B_i)) = \lambda_n(\Phi(W \setminus \bigsqcup_{i=1}^{\infty} B_i)) \le \lambda_n(\Phi(U \setminus \bigsqcup_{i=1}^{\infty} B_i)) = 0$ 이고, 각 $i \in \mathbb{N}$에 대해 $W \cap B_i \subseteq \overline{B_i} \subseteq U$ 이므로 앞선 결과에서 $K \subseteq U$가 임의의 compact한 집합이었음을 상기한다면

$$\begin{aligned}
\lambda_n(\Phi(W)) &= \lambda_n\left(\bigsqcup_{i=1}^{\infty} \Phi(W \cap B_i) \right) \\
&= \sum_{i=1}^{\infty} \lambda_n(\Phi(W \cap B_i)) \\
&= \sum_{i=1}^{\infty} \int_{W \cap B_i} |\det \mathbf{D}\Phi|\, d\lambda_n \\
&= \int_{\bigsqcup_{i=1}^{\infty} (W \cup B_i)} |\det \mathbf{D}\Phi|\, d\lambda_n \\
&= \int_W |\det \mathbf{D}\Phi|\, d\lambda_n
\end{aligned}$$

이다.

Step 10. 닫힌집합 $F \subseteq U$에 대해 $\lambda_n(\Phi(F)) = \int_F |\det \mathbf{D}\Phi|\, d\lambda_n$ 이다.

만약 $\Phi(F)$가 유계라면 적당한 열린집합 $W \subseteq U$가 존재하여 $\Phi(W)$가 유계이며 $\Phi(F) \subseteq \Phi(W)$ 이고, [28]9단계의 결과로부터 $\lambda_n(\Phi(W \setminus F)) = \int_{W \setminus F} |\det \mathbf{D}\Phi|\, d\lambda_n < \infty$ 이고 $\lambda_n(\Phi(W)) = \int_W |\det \mathbf{D}\Phi|\, d\lambda_n < \infty$ 이므로

$$\begin{aligned}
\lambda_n(\Phi(F)) &= \lambda_n(\Phi(W)) - \lambda_n(\Phi(W \setminus F)) \\
&= \int_W |\det \mathbf{D}\Phi|\, d\lambda_n - \int_{W \setminus F} |\det \mathbf{D}\Phi|\, d\lambda_n \\
&= \int_F |\det \mathbf{D}\Phi|\, d\lambda_n
\end{aligned}$$

이다. 이제 $\Phi(F)$가 유계가 아닌 경우에는 집합열 $\{F_i\}$를 $F_i := F \cap \Phi^{-1}(B(i))$로 두면 각 $i \in \mathbb{N}$에 대해 $\Phi(F_i) = \Phi(F) \cap B(i)$에서 $\Phi(F_i)$는 유계이고, $F_i \uparrow F$임이 자명하다. 한편, $\Phi(F_i) \uparrow \Phi(F)$임도 분명하므로 앞선 결과와 정리 2.24의 i, 2.112의 i로부터 $\lambda_n(\Phi(F)) = \lim_{i\to\infty} \lambda_n(\Phi(F_i)) = \lim_{i\to\infty} \int_{F_i} |\det \mathbf{D}\Phi|\, d\lambda_n = \int_F |\det \mathbf{D}\Phi|\, d\lambda_n$이다.

증명을 끝내기 위해, 임의의 가측집합 $A \subseteq U$를 생각하면 Lebesgue 측도가 regular하므로

$$\begin{aligned}
\lambda_n(\Phi(A)) &= \inf\{\lambda_n(W) \in \overline{\mathbb{R}}_0^+ : \Phi(A) \subseteq W \subseteq \mathbb{R}^n, W\text{는 열린집합}\} \\
&= \inf\{\lambda_n(\Phi(W)) \in \overline{\mathbb{R}}_0^+ : A \subseteq W \subseteq U, W\text{는 열린집합}\} \\
&= \inf\left\{ \int_W |\det \mathbf{D}\Phi|\, d\lambda_n \in \overline{\mathbb{R}}_0^+ : A \subseteq W \subseteq U, W\text{는 열린집합}\right\} \\
&\geq \int_A |\det \mathbf{D}\Phi|\, d\lambda_n
\end{aligned}$$

이고 비슷하게

$$\begin{aligned}
\lambda_n(\Phi(A)) &= \sup\{\lambda_n(K) \in \overline{\mathbb{R}}_0^+ : K \subseteq \Phi(A), K\text{는 compact 집합}\} \\
&= \sup\{\lambda_n(\Phi(K)) \in \overline{\mathbb{R}}_0^+ : K \subseteq A, K\text{는 compact 집합}\} \\
&= \sup\left\{ \int_K |\det \mathbf{D}\Phi|\, d\lambda_n \in \overline{\mathbb{R}}_0^+ : K \subseteq A, K\text{는 compact 집합}\right\} \\
&\leq \int_A |\det \mathbf{D}\Phi|\, d\lambda_n
\end{aligned}$$

이 되어 곧 $\lambda_n(\Phi(A)) = \int_A |\det \mathbf{D}\Phi|\, d\lambda_n$이다. 이제 1단계와 8단계로부터 임의의 $\mathscr{C}^1$급 미분동형사상이 good임을 안다.

한편, V에서 적분가능한 $f : \mathbb{R}^n \to \overline{\mathbb{R}}$의 경우에는 이를 각각 양과 음의 부분으로 나누어 앞선 결과를 적용하면 $(f \circ \Phi)|\det \mathbf{D}\Phi|$는 U에서 적분가능하며 $\int_V f\, d\lambda_n = \int_U (f \circ \Phi)|\det \mathbf{D}\Phi|\, d\lambda_n$임을 쉽게 알 수 있고, $(f \circ \Phi)|\det \mathbf{D}\Phi|$가 U에서 적분가능한 $f : \mathbb{R}^n \to \overline{\mathbb{R}}$에 대해서도 비슷하게 하면 된다. □

여기서 한 발 더 나아가 Sard의 정리와 INFT로부터 변수변환 공식에서의 Φ의 조건을 $\mathscr{C}^1$급 전단사함수로 완화할 수 있다.

Theorem 2.174 (Sard) 열린집합 $U \subseteq \mathbb{R}^n$에서 정의된 $\mathscr{C}^1$급 단사함수 $f : U \to \mathbb{R}^n$의 모든 임계점의 집합 C에 대해 $f(C)$는 영집합이다.

Proof 먼저 $U = (0,1)^n$인 특별한 경우를 생각하자. 가정으로부터 f가 $\mathscr{C}^1$급이므로 곧 Lipschitz 연속이며, 이때의 Lipschitz 상수를 $L > 0$이라 하고 임의의 $x \in U$에 대해 함수 $T_x : \mathbb{R}^n \to \mathbb{R}^n$를 $T_x : y \mapsto f(x) + \mathbf{D}f(x)(y - x)$로 두자. 증명은 크게 2단계로 이루어진다.

Step 1. 임의의 $x, y \in U$에 대해 $||f(y) - T_x(y)|| = o(||x - y||)$이다.

서로다른 $x, y \in U$를 임의로 택하면 MVT에서 x와 y를 잇는 선분 위에 적당한 $z_1, \cdots, z_n \in U$이 존재하여 각 $i \leq n$에 대해 $f_i(y) - T_x^i(y) = f_i(y) - f_i(x) - \mathbf{D} f_i(x)(y-x) = [\mathbf{D} f_i(z_i) - \mathbf{D} f_i(x)](y-x)$이다. 이는 각 $i \leq n$에 대해 $|f_i(y) - T_x^i(y)| = |[\mathbf{D} f_i(z_i) - \mathbf{D} f_i(x)](y-x)| \leq ||\mathbf{D} f_i(z_i) - \mathbf{D} f_i(x)||_{\mathrm{op}}||x-y||$ 임을 뜻하므로 $||f(y) - T_x(y)||/||x-y|| \leq \sum_{i=1}^n |f_i(y) - T_x^i(y)|/||x-y|| \leq \sum_{i=1}^n ||\mathbf{D} f_i(z_i) - \mathbf{D} f_i(x)||_{\mathrm{op}}$이고, 곧 f가 $\mathscr{C}^1$ 급이라는 사실로부터 $||f(y) - T_x(y)|| = o(||x-y||)$ 임을 안다.

Step 2. 임의의 $\varepsilon > 0$에 대해 적당한 $R > 0$이 존재하여 임의의 $0 < r < R$에 대해 다음이 성립한다 : 임의의 $x \in C$에 대해 적당한 box $B \subseteq \mathbb{R}^n$가 존재하여 $\lambda_n(B) = \varepsilon r^n$이고 $f(B(x,r)) \subseteq B$이다.

1단계의 결과로부터 적당한 $R > 0$이 존재하여 $0 < r < R$이면 $||x-y|| < r$인 임의의 $x, y \in U$에 대해 $||f(y) - T_x(y)|| \leq (\varepsilon/L^{n-1})||x-y|| < \varepsilon r/L^{n-1}$이며 $||f(x) - f(y)|| < L||x-y|| < Lr$이다. 이제 임의의 $x \in C$를 택하면 적당한 $n-1$차원 부분공간 $\mathsf{V} < \mathbb{R}^n$가 존재하여 $\mathrm{im}\, T_x \leq \mathsf{V}$이고, 곧 $||x-y|| < r$인 임의의 $y \in U$에 대해 $d(f(y), \mathsf{V}) \leq ||f(y) - T_x(y)|| < \varepsilon r/L^{n-1}$이다. 따라서 $f(x)$를 중심으로 하고 각 모서리의 길이가 Lr로 같은 $n-1$차원 box $B' \subseteq \mathsf{V}$와 이를 V에 수직인 방향으로 $\pm\varepsilon r/2L^{n-1}$씩 평행이동시켜 얻는 n차원 box $B \subseteq \mathbb{R}^n$를 생각하면 임의의 $y \in B(x,r) \cap U$에 대해 $f(y) \in B$임이 자명하여 곧 $f(B(x,r)) \subseteq B$이고 이때 $\lambda_n(B) = \varepsilon r^n$이다. (이는 적당한 Affine 변환에 변수변환 공식을 적용한 결과로, 직관적으로도 자명하다. 이제부터는 이 정도의 변수변환 공식의 가벼운 응용은 특별한 언급 없이 사용하도록 한다.)

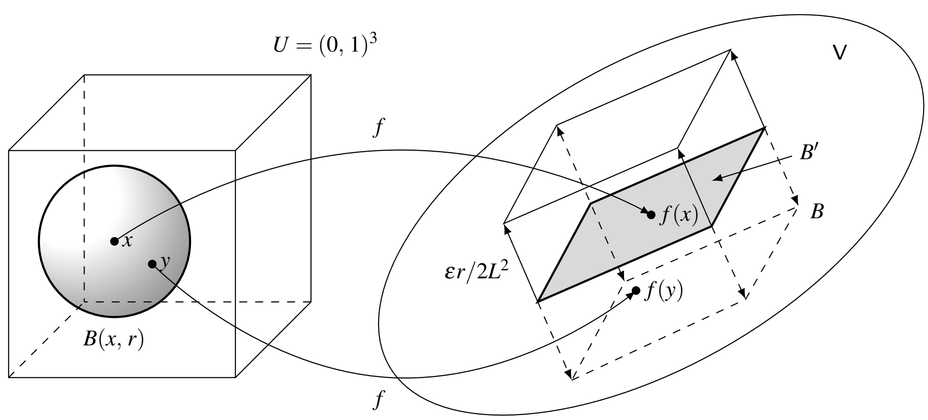


Figure 2.14 Sard의 정리의 증명에서의 집합 B'와 B.

증명을 끝내기 위해 임의의 $\varepsilon > 0$을 택하면 2단계의 결과로부터 적당한 $0 < r < 1$이 존재하여 다음이 성립한다 : 임의의 $x \in C$에 대해 적당한 box $B \subseteq \mathbb{R}^n$가 존재하여 $\lambda_n(B) = \varepsilon r^n/(\sqrt{n}+1)^n$이고 $f(B(x,r)) \subseteq B$이다. 이제 $N = \lceil \sqrt{n}/r \rceil$과 임의의 $j_1, \cdots, j_n \leq N$

에 대해 $B_{j_1\cdots j_n} = \prod_{i=1}^{n}((j_i-1)/N, j_i/N)$라 하고, 이 중에 C의 원소를 적어도 하나 포함하는 것만을 택하여 $B_1, \cdots, B_k$ $(k \le N^n)$라 하자. 그렇다면 각 $j \le k$에 대해 $x_j \in B_j \cap C$를 택할 수 있으므로 적당한 box $\widetilde{B}_j \subseteq \mathbb{R}^n$가 존재하여 $\lambda_n(\widetilde{B}_j) = \varepsilon r^n/(\sqrt{n}+1)^n$이고 $B_j \subseteq B(x_j, \sqrt{n}/N) \subseteq B(x_j, r)$에서 $f(B_j) \subseteq f(B(x_j, r)) \subseteq \widetilde{B}_j$이다. 이제 $f(C) \subseteq \bigcup_{j=1}^{k} f(B_j) \cup \bigcup_{j_1,\cdots,j_n=1}^{N} f(\partial B_{j_1\cdots j_n}) \subseteq \bigcup_{j=1}^{k} \widetilde{B}_j \cup \bigcup_{j_1,\cdots,j_n=1}^{N} f(\partial B_{j_1\cdots j_n})$인데, 보조정리 2.127, 2.172에서 $\bigcup_{j_1,\cdots,j_n=1}^{N} f(\partial B_{j_1\cdots j_n})$가 영집합이므로

$$\begin{aligned}
\rho_n^*(f(C)) &\le \lambda_n\left(\bigcup_{j=1}^{k} \widetilde{B}_j \cup \bigcup_{j_1,\cdots,j_n=1}^{N} f(\partial B_{j_1\cdots j_n})\right) \\
&\le \lambda_n\left(\bigcup_{j=1}^{k} \widetilde{B}_j\right) \\
&\le \sum_{j=1}^{k} \lambda_n(\widetilde{B}_j) \\
&= \sum_{j=1}^{k} \frac{\varepsilon r^n}{(\sqrt{n}+1)^n} \\
&\le N^n \frac{\varepsilon r^n}{(\sqrt{n}+1)^n} \\
&< \left(\frac{\sqrt{n}}{r}+1\right)^n \frac{\varepsilon r^n}{(\sqrt{n}+1)^n} \\
&= \left(\frac{\sqrt{n}+r}{\sqrt{n}+1}\right)^n \varepsilon \\
&< \varepsilon
\end{aligned}$$

이다. 이는 곧 임의의 $S \subseteq \mathbb{R}^n$에 대해 $\rho_n^*(S \cap f(C)) + \rho_n^*(S \cap [f(C)]^c) \le \rho_n^*(f(C)) + \rho_n^*(S \cap [f(C)]^c) = \rho_n^*(S \cap [f(C)]^c) \le \rho_n^*(S)$임을 뜻하여 명제 2.35로부터 $f(C)$가 영집합임을 안다.

이제 f를 적당히 scaling하고 평행이동함으로써 이상의 결론이 U가 임의의 유계인 열린 box인 경우에도 성립함을 알 수 있다. 한편, 일반적인 열린집합 $U \subseteq \mathbb{R}^n$를 생각하면 이는 적당한 유계인 열린 box $B_1, B_2, \cdots$에 대해 $U = \bigcup_{i=1}^{k} B_i$로 표현되므로 (각주 5 참조.) $f(C) = f(\bigcup_{i=1}^{k}(C \cap B_i)) = \bigcup_{i=1}^{k} f(C \cap B_i) = \bigcup_{i=1}^{k} f|_{B_i}(C)$에서 $f(C)$가 영집합임을 안다. (여기서 k는 유한할 수도, ∞일 수도 있다.) $\square$

Corollary 2.175 (Transformation formula) 열린집합 $U, V \subseteq \mathbb{R}^n$ 사이에서 정의된 $\mathscr{C}^1$ 급 전단사함수 $\Phi : U \to V$와 음이 아닌 가측함수 $f : \mathbb{R}^n \to \overline{\mathbb{R}}_0^+$에 대해 $(f \circ \Phi)|\det \mathbf{D}\Phi| : U \to \overline{\mathbb{R}}_0^+$는 가측이고 $\int_V f\, d\lambda_n = \int_U (f \circ \Phi)|\det \mathbf{D}\Phi|\, d\lambda_n$이다. 한편, V에서 적분가능한 (혹은 $(f \circ \Phi)|\det \mathbf{D}\Phi|$가 U에서 적분가능한) $f : \mathbb{R}^n \to \overline{\mathbb{R}}$에 대해서도 동일한 결과가 성립하며, 이때 $(f \circ \Phi)|\det \mathbf{D}\Phi|$는 U에서 적분가능하다 (혹은 f는 V에서 적분가능하다).

PROOF 함수 Φ의 모든 임계점의 집합 $C = (\det \mathbf{D} f)^{-1}(0)$를 생각하면 $\det \mathbf{D} f : U \to \mathbb{R}$가 연속이므로 이는 U에 대해 닫혀있다. 이로부터 $W = U \setminus C$는 열려있으며 INFT에서 $\Phi(W)$도 열려있고 $\Phi|_W : W \to \Phi(W)$는 $\mathscr{C}^1$급 미분동형사상이다. 그렇다면 Sard의 정리로부터 $\Phi(C)$가 영집합이고, 앞서 증명한 변수변환 공식으로부터 $(f \circ \Phi)|\det \mathbf{D}\Phi| = (f \circ \Phi|_W)\mathbf{1}_W$는 가측이며 $\int_V f\,d\lambda_n = \int_{V \setminus \Phi(C)} f\,d\lambda_n = \int_{\Phi(W)} f\,d\lambda_n = \int_W (f \circ \Phi|_W)|\det \mathbf{D}\Phi|_W|\,d\lambda_n = \int_U (f \circ \Phi)|\det \mathbf{D}\Phi|\,d\lambda_n$이다.

한편, V에서 적분가능한 $f : \mathbb{R}^n \to \overline{\mathbb{R}}$의 경우에는 이를 각각 양과 음의 부분으로 나누어 앞선 결과를 적용하면 $(f \circ \Phi)|\det \mathbf{D}\Phi|$는 U에서 적분가능하며 $\int_V f\,d\lambda_n = \int_U (f \circ \Phi)|\det \mathbf{D}\Phi|\,d\lambda_n$임을 쉽게 알 수 있고, $(f \circ \Phi)|\det \mathbf{D}\Phi|$가 U에서 적분가능한 $f : \mathbb{R}^n \to \overline{\mathbb{R}}$에 대해서도 비슷하게 하면 된다. □

이제 변수변환 공식으로부터 쉽게 얻어지는 따름정리들을 간단히 소개하는 것으로 이번 절을 마무리하자. 첫 번째로 소개할 따름정리는 미적분학에서 배우는 직교좌표계에서 극좌표계, 원통좌표계, 구면좌표계로의 변수변환 공식이다.

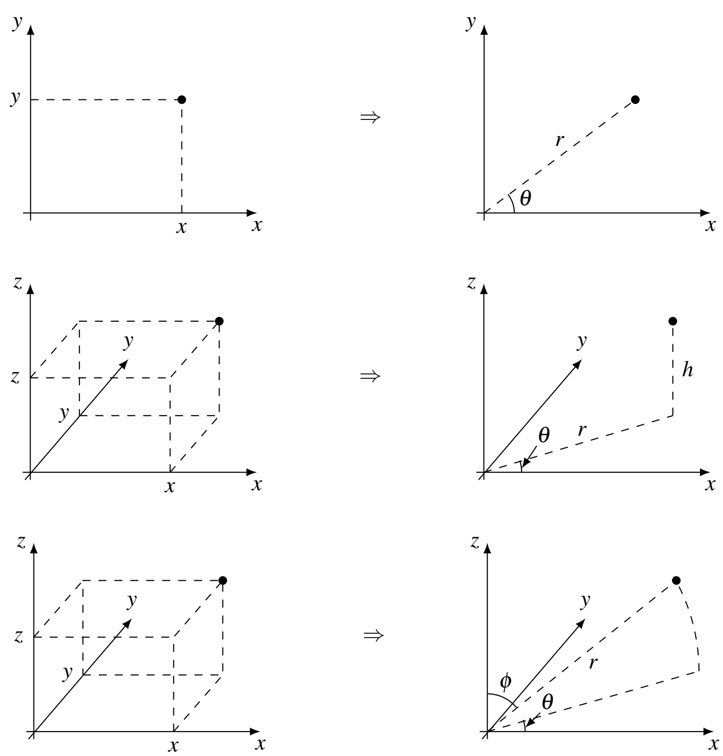


Figure 2.15 기본적인 직교좌표계에서 극좌표계 (위), 원통좌표계 (중간), 구면좌표계 (아래) 로의 변환.

Corollary 2.176 음이 아닌 가측함수 $f : \mathbb{R}^2 \to \overline{\mathbb{R}}_0^+$, $g : \mathbb{R}^3 \to \overline{\mathbb{R}}_0^+$ 에 대해 다음이 성립한다.

i. (극좌표계 변환)

$$\int_{\mathbb{R}^2} f(x, y)\, d\lambda_2(x, y) = \int_{\mathbb{R}_0^+ \times [0, 2\pi)} f(r\cos\theta, r\sin\theta) r\, d\lambda_2(r, \theta)$$

ii. (원통좌표계 변환)

$$\int_{\mathbb{R}^3} g(x, y, z)\, d\lambda_3(x, y, z) = \int_{\mathbb{R}_0^+ \times [0, 2\pi) \times \mathbb{R}} g(r\cos\theta, r\sin\theta, h) r\, d\lambda_3(r, \theta, h)$$

iii. (구면좌표계 변환)

$$\begin{aligned}&\int_{\mathbb{R}^3} g(x, y, z)\, d\lambda_3(x, y, z) \\ &\quad = \int_{\mathbb{R}_0^+ \times [0, \pi] \times [0, 2\pi)} g(r\sin\theta\cos\phi, r\sin\theta\sin\phi, r\cos\theta) r^2 \sin\theta\, d\lambda_3(r, \theta, \phi)\end{aligned}$$

한편, 적분가능한 $f : \mathbb{R}^2 \to \overline{\mathbb{R}}$, $g : \mathbb{R}^3 \to \overline{\mathbb{R}}$에 대해서도 동일한 결과가 성립한다.

Proof 함수

$$\begin{aligned}&\Phi_1 : \mathbb{R}^+ \times (0, 2\pi) \to \mathbb{R}^2 \setminus (\mathbb{R}_0^+ \times \{0\}) \\ &\Phi_2 : \mathbb{R}^+ \times (0, 2\pi) \times \mathbb{R} \to \mathbb{R}^3 \setminus (\mathbb{R}_0^+ \times \{0\} \times \mathbb{R}) \\ &\Phi_3 : \mathbb{R}^+ \times (0, \pi) \times (0, 2\pi) \to \mathbb{R}^3 \setminus (\mathbb{R}_0^+ \times \{0\} \times \mathbb{R})\end{aligned}$$

을 각각

$$\begin{aligned}&\Phi_1 : (r, \theta) \mapsto (r\cos\theta, r\sin\theta) \\ &\Phi_2 : (r, \theta, h) \mapsto (r\cos\theta, r\sin\theta, h) \\ &\Phi_3 : (r, \theta, \phi) \mapsto (r\sin\theta\cos\phi, r\sin\theta\sin\phi, r\cos\theta)\end{aligned}$$

로 정의하면 이들이 $\mathscr{C}^1$ 급 미분동형사상임을 쉽게 확인할 수 있고,

$$\det \mathbf{D}\Phi_1(r, \theta) = \det \begin{bmatrix} \cos\theta & -r\sin\theta \\ \sin\theta & r\cos\theta \end{bmatrix} = r$$

$$\det \mathbf{D}\Phi_2(r, \theta, h) = \det \begin{bmatrix} \cos\theta & -r\sin\theta & 0 \\ \sin\theta & r\cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix} = r$$

$$\det \mathbf{D}\Phi_3(r,\theta,\phi) = \det \begin{bmatrix} \sin\theta\cos\phi & r\cos\theta\cos\phi & -r\sin\theta\sin\phi \\ \sin\theta\sin\phi & r\cos\theta\sin\phi & r\sin\theta\cos\phi \\ \cos\theta & -r\sin\theta & 0 \end{bmatrix} = r^2\sin\theta$$

이므로 변수변환 공식과 보조정리 2. 127로부터 정리가 자명하다. □

다음은 변수변환 공식의 Borel 측도공간 버전이다.

Corollary 2.177 열린집합 $U, V \subseteq \mathbb{R}^n$ 사이에서 정의된 $\mathscr{C}^1$ 급 전단사함수 $\Phi : U \to V$ 와 음이 아닌 Borel 함수 $f : \mathbb{R}^n \to \overline{\mathbb{R}}_0^+$ 에 대해 합성 $f \circ \Phi : U \to \overline{\mathbb{R}}_0^+$ 는 Borel이고 $\int_V f\, d\mu_n = \int_U (f\circ\Phi)|\det \mathbf{D}\Phi|\, d\mu_n$ 이다. 한편, V 에서 μ_n–적분가능한 (혹은 $(f\circ\Phi)|\det \mathbf{D}\Phi|$ 가 U 에서 μ_n–적분가능한) $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ 에 대해서도 동일한 결과가 성립하며, 이때 $(f\circ\Phi)|\det \mathbf{D}\Phi|$ 는 U 에서 μ_n–적분가능하다 (혹은 f 는 V 에서 μ_n–적분가능하다).

Proof 이는 변수변환 공식과 정리 2. 73, 2. 115로부터 자명하다. □

세 번째는 보조정리 2. 127의 일반화로, 직관적으로도 꽤 자명하다.

Corollary 2.178 모든 부분공간 $\mathsf{V} < \mathbb{R}^n$ 는 영집합이다.

Proof 영집합의 정의로부터 $\dim \mathsf{V} = n-1$ 인 경우만 생각하면 되고, 이때에 V 는 $\mathbb{R}^n$ 에서의 초평면이므로 만약 이가 $\mathbb{R}^n$ 의 좌표축 중 어느 하나와 직교한다면 보조정리 2. 127로부터 정리가 자명하다. 만약 V 와 직교하는 좌표축이 존재하지 않는다면, 적당한 가역인 선형사상 $T : \mathbb{R}^n \to \mathbb{R}^n$ 가 존재하여 $T(\mathsf{V})$ 는 $\mathbb{R}^n$ 의 좌표축 중 어느 하나와 직교하는 초평면이고 $|\det T| = 1$ 이므로[29] 변수변환 공식과 앞선 결과로부터 $\lambda_n(T(\mathsf{V})) = \int_{\mathbb{R}^n} \mathbf{1}_{T(\mathsf{V})}\, d\lambda_n = \int_{\mathbb{R}^n} (\mathbf{1}_{T(\mathsf{V})} \circ T)|\det T|\, d\lambda_n = \int_{\mathbb{R}^n} \mathbf{1}_{\mathsf{V}}\, d\lambda_n = \lambda_n(\mathsf{V})$ 에서 V 가 영집합임을 안다. □

마지막 따름정리는 선형사상의 행렬식과 측도로써 정의되는 넓이의 개념을 아름답게 연결해준다.

Corollary 2.179 가측집합 $A \subseteq \mathbb{R}^n$ 와 선형사상 $T : \mathbb{R}^n \to \mathbb{R}^n$ 에 대해 $T(A)$ 도 가측이고, $\lambda_n(T(A)) = |\det T|\lambda_n(A)$ 이다.

Proof 만약 T 가 가역이라면 변수변환 공식으로부터 $\lambda_n(A) = \int_{T(A)} |\det T^{-1}|\, d\lambda_n = \lambda_n(T(A))/|\det T|$ 가 되어 정리가 자명하다. 한편, T 가 가역이 아니라면 $\operatorname{im} T < \mathbb{R}^n$ 이므로 따름정리 2. 178로부터 $\lambda_n(T(A)) \le \lambda_n(\operatorname{im} T) = 0 = |\det T|\lambda_n(A)$ 이다. □

2.13 Fundamental Theorem of Calculus

이번 절에서는 FTC로 돌아간다. FTC는 적분과 미분을 이어주는 다리 역할을 하는 정리로서 미적분학의 근간을 이루는 중요한 정리이다. 먼저, 우리가 1학년 미적분학 시간에 배운 기본적인 FTC를 보자.

Theorem (Fundamental theorem of calculus) 다음이 성립한다.

i. 적분가능한 연속함수 $f:\mathbb{R}\to\mathbb{R}$에 대해 함수 $F:\mathbb{R}\to\mathbb{R}$를 $F:x\mapsto\int_{-\infty}^{x} f\,d\lambda_1$로 두면 이는 f의 역도함수이다. 즉, $F'=f$이다.
ii. 함수 $F:\mathbb{R}\to\mathbb{R}$가 $\mathscr{C}^1$급이고 그 도함수가 적분가능하면 $f:=F'$에 대해 $F(x)=\int_{-\infty}^{x} f\,d\lambda_1$ 이다.

쉽게 생각하면, i은 적분한 뒤 미분하면 원래의 함수를 얻는다는 의미이고, ii는 반대로 미분한 뒤 적분해도 원래의 함수를 얻는다는 의미로, 이 둘은 서로 필요충분과 비슷한 관계이다. 우리는 하도 많이 쓰는 까닭에 자명하게까지 느껴지는 이 FTC에 한 가지 간단하지만, 중요한 질문을 던진다. 만약 여기서 f가 연속이 아니면 어떻게 될까? 지금부터 우리는 이에 대해 답을 찾아나선다.

먼저 i을 보면, 함수 F를 정의할 수는 있어야 하므로 f가 적분가능하다는 조건은 필수적이다. 그러나 f가 적분가능하다는 조건만으로는 F의 미분가능성을 보장할 수가 없다. 실제로 적분가능하지만 F가 특정 점에서 미분불가능한 f를 쉽게 찾을 수 있다.[30]하지만 이렇게 허무하게 끝날 것이었으면 애초에 시작하지도 않았을 것이다. 놀랍게도, 비록 F가 모든 점에서 미분가능하지는 않지만, 이가 거의 대부분의 점에서 '미분 비슷한 무언가' 를 할 수 있음을 보일 수 있다. 다만, 이를 증명하기 위해 준비해야 할 것들이 꽤나 많다. 먼저 비교적 간단한 적분가능한 함수의 근사부터 시작하자.

Theorem 2.180 적분가능한 $f:\mathbb{R}^n\to\overline{\mathbb{R}}$와 임의의 $\varepsilon>0$에 대해 다음이 성립한다.

i. 단순함수 $g:\mathbb{R}^n\to\mathbb{R}$가 존재하여 이는 적당한 $a_1,\cdots,a_k\in\mathbb{R}$와 유계인 $B_1,\cdots,B_k\in\mathscr{I}_n$에 대해 $g=\sum_{i=1}^{k}a_i\mathbf{1}_{B_i}$로 쓸 수 있으며 $\int_{\mathbb{R}^n}|f-g|\,d\lambda_n<\varepsilon$이다.
ii. 적분가능한 연속함수 $g:\mathbb{R}^n\to\mathbb{R}$가 존재하여 $\int_{\mathbb{R}^n}|f-g|\,d\lambda_n<\varepsilon$이다.

Proof i. 먼저 f가 음이 아닌 특별한 경우를 생각하면 정리 2.83로부터 적당한 음이 아닌 단순함수 $f_i:\mathbb{R}^n\to\mathbb{R}_0^+$의 열 $\{f_i\}$가 존재하여 $f-f_i\downarrow 0$이고 각 $i\in\mathbb{N}$에 대해 $f-f_i\le f$이다. 그렇다면 DCT로부터 $\int_{\mathbb{R}^n}|f-f_i|\,d\lambda_n\to 0$이고, 곧 임의의 $\varepsilon>0$을 택하면 충분히 큰 $i_0\in\mathbb{N}$에 대해 $\int_{\mathbb{R}^n}|f-f_{i_0}|\,d\lambda_n<\varepsilon/2$이다.

표기의 편의를 위해 $\widetilde{f}:=f_{i_0}$로 두자. 함수 $\widetilde{f}$의 표준형을 $\sum_{j=1}^{l}a_j\mathbf{1}_{A_j}$라 하고, WLOG, 모든 a_j가 0이 아니라 하면 정리 2.66, 2.68로부터 각 $j\le l$에 대해 적당한 $B_j,C_j\in\mathscr{B}_n$가

존재하여 $B_j \subseteq A_j \subseteq C_j$ 이고 $C_j \setminus B_j$ 가 영집합이므로 $A_j \setminus B_j$ 도 영집합이다. 나아가, 집합족 $\mathscr{S}_n$ 이 $\mathscr{B}_n$ 의 생성자이고 각 $j \le l$ 에 대해 $\mu_n(B_j) < \infty$ 임이 자명하므로 (만약 어떤 $j_0 \le l$ 에 대해 $\mu_n(B_{j_0}) = \infty$ 이면 $\lambda_n(A_{j_0}) = \infty$ 이고, 곧 $\int_{\mathbb{R}^n} f\,d\lambda_n \ge \int_{\mathbb{R}^n} \widetilde{f}\,d\lambda_n = \sum_{j=1}^{l} a_j \lambda_n(A_j) \ge a_{j_0}\lambda(A_{j_0}) = \infty$ 가 되어 f 가 적분가능하다는 가정에 모순된다.) 정리 2.41의 ii로부터 서로소인 $B_{j1}, \cdots, B_{jm_j} \in \mathscr{S}_n$ 이 존재하여 $\mu_n(B_j \triangle \bigsqcup_{k=1}^{m_j} B_{jk}) < \varepsilon/2a_j l$ 이다. 이제 단순함수 $g = \sum_{j=1}^{l} a_j \mathbf{1}_{\bigsqcup_{k=1}^{m_j} B_{jk}} = \sum_{j=1}^{l}\sum_{k=1}^{m_j} a_j \mathbf{1}_{B_{jk}}, h = \sum_{j=1}^{l} a_j \mathbf{1}_{B_j}$ 를 생각하면 이상으로부터

$$
\begin{aligned}
\int_{\mathbb{R}^n} |g-h|\,d\lambda_n &= \int_{\mathbb{R}^n} \left| \sum_{j=1}^{l} a_j (\mathbf{1}_{\bigsqcup_{k=1}^{m_j} B_{jk}} - \mathbf{1}_{B_j}) \right| d\lambda_n \\
&\le \int_{\mathbb{R}^n} \sum_{j=1}^{l} a_j \mathbf{1}_{B_j \triangle \bigsqcup_{k=1}^{m_j} B_{jk}}\,d\lambda_n \\
&= \sum_{j=1}^{l} a_j \lambda_n \left(B_j \triangle \bigsqcup_{k=1}^{m_j} B_{jk} \right) \\
&< \sum_{j=1}^{l} \frac{\varepsilon}{2l} \\
&= \frac{\varepsilon}{2}
\end{aligned}
$$

이고 $\widetilde{f} = h$ (ae.) 이므로 따름정리 2.102의 ii로부터 $\int_{\mathbb{R}^n} |f-g|\,d\lambda_n \le \int_{\mathbb{R}^n} |f-\widetilde{f}|\,d\lambda_n + \int_{\mathbb{R}^n} |\widetilde{f}-g|\,d\lambda_n < \varepsilon/2 + \int_{\mathbb{R}^n} |g-h|\,d\lambda_n < \varepsilon$ 에서 정리가 성립한다.

한편, 적분가능한 $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ 에 대해서는 이를 양과 음의 부분으로 나누어 앞선 결과를 적용하면 된다.

ii. 함수 f 가 적분가능하므로 정리 2.100의 ii로부터 거의 어디서나 유한하고, WLOG, 필요하다면 영집합에서의 함숫값을 0으로 바꾸어 f 가 항상 유한하다고 해도 된다. 그렇다면 i로부터 임의의 $\varepsilon > 0$ 을 택하면 단순함수 $h : \mathbb{R}^n \to \mathbb{R}$ 가 존재하여 이는 적당한 $a_1, \cdots, a_k \in \mathbb{R}$ 와 유계인 $B_1, \cdots, B_k \in \mathscr{S}_n$ 에 대해 $h = \sum_{i=1}^{k} a_i \mathbf{1}_{B_i}$ 로 쓸 수 있으며 $\int_{\mathbb{R}^n} |f-h|\,d\lambda_n < \varepsilon/2$ 이다. 이제 각 $i \le k$ 와 임의의 $j \in \mathbb{N}$ 에 대해 B_i 의 모서리를 조금씩 줄여 (단, 이때 $2/j$ 이상으로는 줄이지 않는 것으로 하자.) 만든 semi-open box를 B_i' 라 하고 $B_i \setminus B_i'$ 에서 h 의 함숫값인 a_i 를 0과 선형으로 보간하여 얻는 함수를 $g_j : \mathbb{R}^n \to \mathbb{R}$ 라 하면 $|g_j - h| \to 0$ (ae.) 이고 (각 $i \le k$ 에 대해 $j \to \infty$ 이면 $B_i' \uparrow B_i$ 이므로 함수열 $\{h_j\}$ 는 $\bigcup_{i=1}^{k} \partial B_i$ 를 제외한 모든 곳에서 g 로 수렴한다. 그런데 주석 26에서와 같은 이유로 $\lambda_n(\bigcup_{i=1}^{k} \partial B_i) = 0$ 이므로 결국 $|g - h_j| \to 0$ (ae.) 이다.) $|g_j - h| \le |g_j| + |h| \le 2|h|$ 이므로 DCT에서 $\int_{\mathbb{R}^n} |g_j - h|\,d\lambda_n \to 0$ 이다. 이는 곧 충분히 큰 $j_0 \in \mathbb{N}$ 를 택하면 $\int_{\mathbb{R}^n} |g_{j_0} - h|\,d\lambda_n < \varepsilon/2$ 임을 뜻하므로 $g := g_{j_0}$ 라 하면 $\int_{\mathbb{R}^n} |f-g|\,d\lambda_n \le \int_{\mathbb{R}^n} |f-h|\,d\lambda_n + \int_{\mathbb{R}^n} |g-h|\,d\lambda_n < \varepsilon$ 이다. 이때 g 가 연속임이 분명하고 $\int_X |g|\,d\lambda_n \le \int_X |f-g|\,d\lambda_n + \int_X |f|\,d\lambda_n < \infty$ 에서 이가 적분가능하므로 정리가 성립한다. □

다음으로 필요한 것은 $\sigma-$대수와 측도의 pushfowarding이다. 이는 pushfoward라는 단어의 사전적 뜻 그대로 고정된 $\sigma-$대수나 측도에 대해 어떤 함수를 앞으로 '밀어넣어' 새로운 $\sigma-$대수나 측도를 얻는 방법이다. 엄밀한 정의를 보면 그 느낌을 더 잘 알 수 있을 것이다.

Definition 2.181 측도공간 $(X, \mathscr{A}, \mu)$와 집합 Y 사이에서 정의된 함수 $\varphi : X \to Y$를 생각하자. 이때 φ에 대한 $\mathscr{A}$의 **pushfoward $\sigma-$대수 (− $\sigma-$algebra)**를 $\varphi_*\mathscr{A}$로 쓰고 $\varphi_*\mathscr{A} = \{A \subseteq Y : \varphi^{-1}(A) \in \mathscr{A}\}$로 정의한다. 비슷하게, φ에 대한 μ의 **pushfoward 측도 (− measure)**를 $\varphi_*\mu : \varphi_*\mathscr{A} \to \overline{\mathbb{R}}_0^+$로 쓰고 $\varphi_*\mu : A \mapsto \mu(\varphi^{-1}(A))$로 정의한다.

물론, 이때의 정의가 well-define되는지 확인해줘야 한다.

Proposition 2.182 측도공간 $(X, \mathscr{A}, \mu)$와 집합 Y 사이에서 정의된 함수 $\varphi : X \to Y$에 대해 $(Y, \varphi_*\mathscr{A}, \varphi_*\mu)$는 측도공간을 이룬다. 따라서 pushfoward $\sigma-$대수와 측도는 well-define 된다.

Proof 먼저 $\varphi_*\mathscr{A}$가 $\sigma-$대수임을 보이자. 이를 위해서는 정의 2.1의 조건을 하나하나 따져보면 된다. 그 중 몇 개만 연습삼아 살펴보면, $\emptyset \in \mathscr{A}$에서 $\emptyset = \varphi^{-1}(\emptyset) \in \varphi_*\mathscr{A}$에서 $\varphi_*\mathscr{A}$가 비어있지 않음이 자명하다. 또한 임의의 $A \in \varphi_*\mathscr{A}$에 대해 $\varphi^{-1}(A) \in \mathscr{A}$이므로 $\varphi^{-1}(A^c) = [\varphi^{-1}(A)]^c \in \mathscr{A}$에서 $A^c \in \varphi_*\mathscr{A}$이다. 이제 남은 조건들도 이와 같이 쉽게 따져볼 수 있다.

다음으로, $\varphi_*\mu$가 측도임을 보이자. 우선 $\varphi_*\mu(\emptyset) = \mu(\varphi^{-1}(\emptyset)) = \mu(\emptyset) = 0$임은 자명하고, $\varphi_*\mathscr{A}$에 속하는 임의의 서로소인 집합열 $\{A_i\}$에 대해 $\varphi_*\mu(\bigsqcup_{i=1}^{\infty} A_i) = \mu(\varphi^{-1}(\bigsqcup_{i=1}^{\infty} A_i)) = \mu(\bigsqcup_{i=1}^{\infty} \varphi^{-1}(A_i)) = \sum_{i=1}^{\infty} \mu(\varphi^{-1}(A_i)) = \sum_{i=1}^{\infty} \varphi_*\mu(A_i)$이므로 $\varphi_*\mu$가 측도임을 알고, 증명이 끝난다. □

위의 정의에서 알 수 있듯이 pushfowarding은 φ에게 아무런 조건을 요구하지 않는다. 연속일 필요도, 전단사여야 할 필요도 없고, 심지어 가측일 필요도 없다. 앞서 배운 음이 아닌 가측함수의 적분으로 새로운 측도를 유도하는 것과 비교하면 pushfowarding은 보다 일반적인 상황에서도 적용할 수 있는 기법인 셈이다. 이러한 pushfowarding은 나중에 확률변수를 배우며 다시 등장할 것이다. 여기서는 pushfowarding에 대한 측도론적인 성질들을 조금 더 살펴보자.

Theorem 2.183 가측공간 $(X, \mathscr{A})$와 집합 Y 사이에서 정의된 함수 $\varphi : X \to Y$와 Y와 가측공간 $(Z, \mathscr{B})$ 사이에서 정의된 함수 $f : Y \to Z$에 대해 f가 $\varphi_*\mathscr{A}/\mathscr{B}-$가측일 필요충분조건은 $f \circ \varphi$가 $\mathscr{A}/\mathscr{B}-$가측인 것이다.

Proof 충분조건임을 보이기 위해 임의의 $A \in \mathscr{B}$를 생각하면 가정으로부터 $f^{-1}(A) \in \varphi_*\mathscr{A}$이므로 $(f \circ \varphi)^{-1}(A) = \varphi^{-1}(f^{-1}(A)) \in \mathscr{A}$에서 $f \circ \varphi$가 $\mathscr{A}/\mathscr{B}-$가측이 되어 충분조건임이 보여진다. 이제 필요조건임도 이와 비슷하게 보일 수 있다. □

Theorem 2.184 (Transformation formula) 측도공간 $(X, \mathscr{A}, \mu)$와 집합 Y 사이에서 정의된 함수 $\varphi : X \to Y$에 대해 음이 아닌 함수 $f : Y \to \overline{\mathbb{R}}_0^+$가 $\varphi_*\mathscr{A}/\mathscr{B}_1$−가측이라 하자. 그렇다면 합성 $f \circ \varphi : X \to \overline{\mathbb{R}}_0^+$는 가측이고 $\int_Y f\, d\varphi_*\mu = \int_X f \circ \varphi\, d\mu$이다. 한편, $\varphi_*\mu$−적분가능한 (혹은 $f \circ \varphi$가 μ−적분가능한 가측함수) $f : Y \to \overline{\mathbb{R}}$에 대해서도 동일한 결과가 성립하며 이때 $f \circ \varphi$는 μ−적분가능하다(혹은 f는 $\varphi_*\mu$−적분가능하다).

Proof 먼저 f가 음이 아닌 단순함수인 경우를 생각하여 이의 표준형을 $f = \sum_{i=1}^k a_i \mathbf{1}_{A_i}$라 하자. 그렇다면 $f \circ \varphi = \sum_{i=1}^k a_i \mathbf{1}_{\varphi^{-1}(A_i)}$에서 이가 가측이고 $\int_Y f\, d\varphi_*\mu = \sum_{i=1}^k a_i \varphi_*\mu(A_i) = \sum_{i=1}^k a_i \mu(\varphi^{-1}(A_i)) = \int_X \sum_{i=1}^k a_i \mathbf{1}_{\varphi^{-1}(A_i)}\, d\mu = \int_X f \circ \varphi\, d\mu$임이 분명하다. 이제 f를 음이 아닌 $\varphi_*\mathscr{A}/\mathscr{B}_1$−가측함수라 하면 정리 2. 83로부터 음이 아닌 단순함수 $f_i : Y \to \mathbb{R}_0^+$의 열 $\{f_i\}$가 존재하여 $f_i \uparrow f$이다. 여기서 각 $i \in \mathbb{N}$에 대해 $f_i \circ \varphi$가 가측이므로 정리 2. 77의 ii에서 $f \circ \varphi$도 가측이고 MCT에서 $\int_Y f\, d\varphi_*\mu = \lim_{i\to\infty} \int_Y f_i\, d\varphi_*\mu = \lim_{i\to\infty} \int_Y f_i \circ \varphi\, d\mu = \int_Y f \circ \varphi\, d\mu$이다.

한편, $\varphi_*\mu$−적분가능한 $f : Y \to \overline{\mathbb{R}}$에 대해서도 이를 양과 음의 부분으로 나누어 앞선 결과를 적용하면 $f \circ \varphi$가 μ−적분가능하며 $\int_Y f\, d\varphi_*\mu = \int_X f \circ \varphi\, d\mu$임을 쉽게 알 수 있고, $f \circ \varphi$가 μ−적분가능한 가측함수 $f : Y \to \overline{\mathbb{R}}$에 대해서도 비슷하게 하면 된다. □

마지막 성질은 앞서 살펴본 변수변환 공식과 묘하게 닮은 부분이 있기에 두 정리 모두 변수변환 공식이라 불린다. 물론, 이러한 유사함은 기분탓만은 아니다. 더 파보면 분명 무언가 더 있겠지만 우리의 갈 길이 먼 까닭에 pushfowarding에 대해서는 이쯤으로 정리하자. [31]

이번 절에서 pushfowarding이 필요한 이유는 바로 다음의 유명한 부등식 때문이다. 보통 확률론에서 쓰이는 이 부등식은 측도론에서도 유용하게 쓰이는 경우가 종종 있다. 나중에 확률론을 배우며 다른 부등식들도 만나볼 수 있을 것이다.

Lemma 2.185 측도공간 $(X, \mathscr{A}, \mu)$와 집합 Y 사이에서 정의된 함수 $\varphi : X \to Y$와 집합 $A \in \varphi_*\mathscr{A}$에 대해 $\mu(\varphi^{-1}(A)) = \int_X \mathbf{1}_A \circ \varphi\, d\mu$이다.

Proof 이는 $\int_X \mathbf{1}_A \circ \varphi\, d\mu = \int_X \mathbf{1}_A\, d\varphi_*\mu = \varphi_*\mu(A) = \mu(\varphi^{-1}(A))$로부터 자명하다. □

Theorem 2.186 (Markov's inequality) 측도공간 $(X, \mathscr{A}, \mu)$에서 정의된 가측함수 $f : X \to \overline{\mathbb{R}}$와 임의의 $M, p > 0$에 대해

$$\mu\{|f(x)| \geq M\} \leq \frac{1}{M^p} \int_X |f|^p\, d\mu$$

가 성립한다. [32]

Proof 부등식 $\mathbf{1}_{(-M,M)^c} \circ f = \mathbf{1}_{(-1,1)^c} \circ (f/M) \leq (|f|/M)^p \mathbf{1}_{(-1,1)^c} \circ (f/M) \leq (|f|/M)^p$와 위의 보조정리로부터 $\mu\{|f(x)| \geq M\} = \int_X \mathbf{1}_{(-M,M)^c} \circ f\, d\mu \leq \int_X (|f|/M)^p\, d\mu$가 성립한다. □

이로써 준비는 끝났다. 이제 본격적으로 FTC의 i을 확장해보자. 먼저 '미분 비슷한 무언가' 를 엄밀히 도입하기 위한 framework이 필요한데, 곧 증명할 Hardy-Littlewood의 정리가 그 역할을 해 줄 것이다.

Definition 2.187 적분가능한 함수 $f:\mathbb{R}^n\to\overline{\mathbb{R}}$에 대해 이의 **(Hardy-Littlewood) maximal function**을 $Mf:\mathbb{R}^n\to\overline{\mathbb{R}}_0^+$로 쓰고

$$Mf:x\mapsto\sup_{r>0}\frac{1}{\lambda_n(B(x,r))}\int_{B(x,r)}|f|\,d\lambda_n$$

로 정의한다.

Proposition 2.188 적분가능한 함수 $f:\mathbb{R}^n\to\overline{\mathbb{R}}$에 대해 이의 maximal function Mf는 가측이다.

Proof 만약 $(Mf)(x)=0$인 $x\in\mathbb{R}^n$가 존재하면 이는 곧 임의의 $i\in\mathbb{N}$에 대해 $\int_{B(x,i)}|f|\,d\lambda_n=0$임을 뜻하여 정리 2.112의 i로부터 $\int_{\mathbb{R}^n}|f|\,d\lambda_n=\lim_{i\to\infty}\int_{B(x,i)}|f|\,d\lambda_n=0$이고, 곧 정리 2.101의 ii로부터 $f=0$ (ae.)이다. 그런데 이 경우에는 $Mf=0$이 되어 이가 명백히 가측이므로 $Mf>0$이라 가정하자.

이제 임의의 $x>0$에 대해 $(Mf)^{-1}((x,\infty])$가 가측이면 $(Mf)^{-1}([-\infty,x])=[(Mf)^{-1}((x,\infty])]^c$에서 $(Mf)^{-1}([-\infty,x])$도 가측이므로 $(Mf)^{-1}((x,\infty])$가 가측임을 보이는 것으로 충분하다. 이를 위해 임의의 $x>0$를 고정하고 $(Mf)^{-1}((x,\infty])$가 열려있음을 보이자. 여기서 임의의 $y\in(Mf)^{-1}((x,\infty])$도 고정하면 적당한 $r>0$이 존재하여

$$\frac{1}{\lambda_n(B(y,r))}\int_{B(y,r)}|f|\,d\lambda_n>x$$

이다. 한편, $r_i\downarrow r$인 임의의 $\{r_i\}$에 대해 정리 2.24의 ii로부터 $\lambda_n(B(y,r_i))\downarrow\lambda_n(B(y,r))$이므로 $\lim_{s\downarrow r}\lambda_n(B(y,s))=\lambda_n(B(y,r))$이고, 곧

$$\lim_{s\downarrow r}\frac{1}{\lambda_n(B(y,s))}\int_{B(y,r)}|f|\,d\lambda_n=\frac{1}{\lambda_n(B(y,r))}\int_{B(y,r)}|f|\,d\lambda_n>x$$

에서 적당한 $s>r$가 존재하여

$$\frac{1}{\lambda_n(B(y,s))}\int_{B(y,r)}|f|\,d\lambda_n>x$$

이다. 또한, 임의의 $z\in B(y,s-r)$에 대해 $w\in B(y,r)$이면 $||z-w||\le||z-y||+||y-w||<s$이므로 $B(y,r)\subseteq B(z,s)$이고, 곧 Lebesgue 측도의 이동 불변성으로부터

$$x<\frac{1}{\lambda_n(B(y,s))}\int_{B(y,r)}|f|\,d\lambda_n$$

$$\begin{aligned}
&\le \frac{1}{\lambda_n(B(y,s))}\int_{B(z,s)} |f|\,d\lambda_n \\
&= \frac{1}{\lambda_n(B(z,s))}\int_{B(z,s)} |f|\,d\lambda_n \\
&= (Mf)(z)
\end{aligned}$$

가 되어 $z \in (Mf)^{-1}((x,\infty])$, 즉 $B(y,s-r) \subseteq (Mf)^{-1}((x,\infty])$에서 증명이 끝난다. □

Lemma 2.189 (Vitali covering lemma) 집합 $\mathbb{R}^n$에 속하는 열린 ball의 모임 $\mathscr{C}$에 대해 $\bigcup\mathscr{C}$가 유계라면 $\mathscr{C}$에서 가산개의 서로소인 열린 ball $B(x_1,r_1), B(x_2,r_2), \cdots$를 적당히 택하여 $\bigcup\mathscr{C} \subseteq \bigcup_{i=1}^{k} B(x_i,5r_i)$이도록 할 수 있다. (여기서 k는 유한할 수도, ∞일 수도 있다.)

Proof 먼저 $d_1 = \sup\{r \in \mathbb{R}_0^+ : B(x,r) \in \mathscr{C}\} < \infty$에 대해 $r > d_1/2$인 열린 ball $B(x,r) \in \mathscr{C}$을 (가능하다면) 택하여 이를 $B(x_1,r_1)$이라 하자. 비슷하게 $d_2 = \sup\{r \in \mathbb{R}_0^+ : B(x,r) \in \mathscr{C}, B(x,r) \cap B(x_1,r_1) = \emptyset\} < \infty$에 대해 $r > d_2/2$인 열린 ball $B(x,r) \in \mathscr{C}$을 (가능하다면) 택하여 이를 $B(x_2,r_2)$라 하고, 이를 가능할 때까지 반복하여 가산개의 열린 ball $B(x_1,r_1), B(x_2,r_2), \cdots$를 얻는다. 이제 이들이 모두 서로소임은 그 구성으로부터 자명하므로 $\bigcup\mathscr{C} \subseteq \bigcup_i^k B(x_i,5r_i)$만 보이면 된다. (여기서 k는 유한할 수도, ∞일 수도 있는데, 이제부터는 $k=\infty$라 생각하고 논의를 진행하도록 한다. k가 유한한 경우에도 비슷하게 하면 된다.)

만약 어떤 $B(x,r) \in \mathscr{C}$가 존재하여 모든 $i \in \mathbb{N}$에 대해 $B(x,r) \cap B(x_i,r_i) = \emptyset$이라면 위의 구성으로부터 각 $i \in \mathbb{N}$에 대해 $r \le d_i$이고, 곧 $r_i > d_i/2 \ge r/2$에서 $\lambda_n(\bigsqcup_i^\infty B(x_i,r_i)) = \sum_{i=1}^\infty \lambda_n(B(x_i,r_i)) \ge \sum_{i=1}^\infty \lambda_n(B(x_i,r/2)) = \infty$가 되어 $\bigcup\mathscr{C}$가 유계라는 가정에 모순되므로 임의의 $B(x,r) \in \mathscr{C}$에 대해 $B(x,r) \cap B(x_i,r_i) \ne \emptyset$인 $i \in \mathbb{N}$가 적어도 하나 존재한다. 이제 임의의 $B(x,r) \in \mathscr{C}$를 고정하고 $B(x,r) \cap B(x_i,r_i) \ne \emptyset$인 $i \in \mathbb{N}$ 중 가장 작은 것을 택하여 $i_0 \in \mathbb{N}$라 하면 $B(x,r) \cap \bigsqcup_{i=1}^{i_0-1} B(x_i,r_i) = \emptyset$이고 곧 위의 구성으로부터 $r \le d_{i_0} < 2r_{i_0}$이다. 또한 i_0의 정의로부터 $B(x,r) \cap B(x_{i_0},r_{i_0}) \ne \emptyset$이므로 임의의 $y \in B(x,r)$에 대해 $||x_{i_0} - y|| \le ||x_{i_0} - x|| + ||x - y|| < 2r + r_{i_0} < 5r_{i_0}$가 되어 $y \in B(x_{i_0},5r_{i_0})$, 즉 $B(x,r) \subseteq B(x_{i_0},5r_{i_0})$에서 $\bigcup\mathscr{C} \subseteq \bigcup_i^\infty B(x_i,5r_i)$임을 안다. □

Lemma 2.190 적분가능한 함수 $f : \mathbb{R}^n \to \overline{\mathbb{R}}$와 임의의 $M > 0$에 대해 $\lambda_n\{(Mf)(x) > M\} \le 5^n||f||_1/M$이다.

Proof 임의의 $M > 0$과 임의의 $i \in \mathbb{N}$를 고정하면 임의의 $x \in (Mf)^{-1}((M,\infty]) \cap B(i) =: A_i$에 대해 적당한 $r_x > 0$가 존재하여

$$\frac{1}{\lambda_n(B(x,r_x))}\int_{B(x,r_x)} |f|\,d\lambda_n > M$$

에서 $\lambda_n(B(x, r_x)) < \int_{B(x,r_x)} |f|/M\, d\lambda_n \le ||f||_1/M$ 이므로 $x \in A_i$ 의 선택과는 무관하게 적당한 $R > 0$ 에 대해 $r_x \le R$ 이고, 곧 $\bigcup_{x\in A_i} B(x, r_x) \subseteq \bigcup_{x\in A_i} B(x, R) \subseteq B(R+i)$ 에서 $\bigcup_{x\in A_i} B(x, r_x)$ 는 유계이다. 그렇다면 Vitali covering lemma로부터 $\{B(x, r_x)\}_{x\in A_i}$ 에 속하는 서로소인 열린 ball의 열 $\{B(x_j, r_{x_j})\}$ 가 존재하여 $\bigcup_{x\in A_i} B(x, r_x) \subseteq \bigcup_{j=1}^{\infty} B(x_j, 5r_{x_j})$ 이다. 이로부터

$$\begin{aligned}
\lambda_n(A_i) &\le \lambda_n\left(\bigcup_{x\in A_i} B(x, r_x)\right) \\
&\le \lambda_n\left(\bigcup_{j=1}^{\infty} B(x_j, 5r_{x_j})\right) \\
&\le \sum_{j=1}^{\infty} \lambda_n(B(x_j, 5r_{x_j})) \\
&= 5^n \sum_{j=1}^{\infty} \lambda_n(B(x_j, r_{x_j})) \\
&\le \frac{5^n}{M} \sum_{j=1}^{\infty} \int_{B(x_j, r_{x_j})} |f|\, d\lambda_n \\
&= \frac{5^n}{M} \int_{\bigsqcup_{j=1}^{\infty} B(x_j, r_{x_j})} |f|\, d\lambda_n \\
&\le \frac{5^n}{M} ||f||_1
\end{aligned}$$

이므로 정리 2. 24의 i에서 $\lambda_n\{(Mf)(x) > M\} = \lim_{i\to\infty} \lambda_n(A_i) \le 5^n ||f||_1/M$ 이다. □

Theorem 2.191 (Hardy-Littlewood) 적분가능한 함수 $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ 와 거의 대부분의 $x \in \mathbb{R}^n$ 에 대해 $r \downarrow 0$ 이면

$$\frac{1}{\lambda_n(B(x, r))} \int_{B(x,r)} |f - f(x)|\, d\lambda_n \to 0$$

이다.

Proof 임의의 적분가능한 $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ 에 대해 함수 $f^* : \mathbb{R}^n \to \overline{\mathbb{R}}_0^+$ 를

$$f^* : x \mapsto \limsup_{r\downarrow 0} \frac{1}{\lambda_n(B(x, r))} \int_{B(x,r)} |f - f(x)|\, d\lambda_n$$

로 정의하고 이의 성질을 살펴보는 것으로 증명을 시작한다. 함수 $f, g : \mathbb{R}^n \to \overline{\mathbb{R}}$ 를 임의의 적분가능한 함수라 하면

$$\limsup_{r\downarrow 0} \frac{1}{\lambda_n(B(x, r))} \int_{B(x,r)} |(f+g) - (f+g)(x)|\, d\lambda_n$$

$$\leq \limsup_{r\downarrow 0} \frac{1}{\lambda_n(B(x,r))}\left[\int_{B(x,r)} |f-f(x)|\,d\lambda_n + \int_{B(x,r)} |g-g(x)|\,d\lambda_n\right]$$
$$\leq \limsup_{r\downarrow 0} \frac{1}{\lambda_n(B(x,r))}\int_{B(x,r)} |f-f(x)|\,d\lambda_n + \limsup_{r\downarrow 0} \frac{1}{\lambda_n(B(x,r))}\int_{B(x,r)} |g-g(x)|\,d\lambda_n$$

에서 $(f+g)^* \leq f^* + g^*$ 이다. 나아가, 만약 g가 $x_0 \in \mathbb{R}^n$ 에서 연속이라면 임의의 $\varepsilon > 0$에 대해 적당한 $\delta > 0$가 존재하여 $||x-x_0|| < \delta$ 이면 $|g(x)-g(x_0)| < \varepsilon$ 이므로 $r < \delta$ 이면

$$\frac{1}{\lambda_n(B(x_0,r))}\int_{B(x_0,r)} |g-g(x_0)|\,d\lambda_n \leq \varepsilon$$

에서 $g^*(x_0) = 0$ 이다. 따라서 g가 연속이면 $g^* = (-g)^* = 0$ 이 되어 $(f-g)^* \leq f^* + (-g)^* = f^* \leq (f-g)^* + g^* = (f-g)^*$ 에서 $(f-g)^* = f^*$ 이다. 한편,

$$\begin{aligned}\frac{1}{\lambda_n(B(x,r))}\int_{B(x,r)} |f-f(x)|\,d\lambda_n &\leq \frac{1}{\lambda_n(B(x,r))}\int_{B(x,r)} |f|+|f(x)|\,d\lambda_n \\ &= \frac{1}{\lambda_n(B(x,r))}\int_{B(x,r)} |f|\,d\lambda_n + |f(x)| \\ &\leq (Mf)(x) + |f(x)|\end{aligned}$$

에서 $f^* \leq Mf + |f|$ 도 성립한다.

이상의 성질들을 정리하면 다음과 같다.

i. 만약 $x_0 \in \mathbb{R}^n$ 에서 g가 연속이면 $g^*(x_0) = 0$ 이다.
ii. 만약 g가 연속이면 $(f-g)^* = f^*$ 이다.
iii. $f^* \leq Mf + |f|$.

이제 임의의 적분가능한 $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ 를 고정하고 $f^* = 0$ (ae.) 임을 보이자. 이를 위해 임의의 $\varepsilon > 0$과 임의의 $M > 0$을 택하면 정리 2.180의 ii로부터 적당한 적분가능한 연속함수 $g : \mathbb{R}^n \to \mathbb{R}$가 존재하여 $||f-g||_1 < M\varepsilon/2(5^n+1)$ 이다. 한편, $(f-g)^* \leq M(f-g) + |f-g|$ 에서 $(f-g)^* > M$ 이면 $M(f-g) > M/2$ 이거나 $|f-g| > M/2$ 이므로 $[(f-g)^*]^{-1}((M,\infty]) \subseteq [M(f-g)]^{-1}((M/2,\infty]) \cup (|f-g|)^{-1}((M/2,\infty])$ 이고, 위의 보조정리와 Markov의 부등식으로부터

$$\begin{aligned}&\lambda_n\left\{(M(f-g))(x) \geq \frac{M}{2} \text{ or } |(f-g)(x)| \geq \frac{M}{2}\right\} \\ &\qquad\leq \lambda_n\left\{(M(f-g))(x) \geq \frac{M}{2}\right\} + \lambda_n\left\{|f-g|(x) \geq \frac{M}{2}\right\} \\ &\qquad\leq \frac{2\cdot 5^n}{M}||f-g||_1 + \frac{2}{M}||f-g||_1 \\ &\qquad< \varepsilon\end{aligned}$$

이다. 이는 곧 $[M(f-g)]^{-1}((M/2,\infty]) \cup (|f-g|)^{-1}((M/2,\infty])$가 영집합임을 함의하므로 $[(f-g)^*]^{-1}((M,\infty])$도 영집합이고, $(f-g)^* = f^*$에서 $(f^*)^{-1}((M,\infty])$도 영집합이다. 이제 M이 임의의 양수임을 상기한다면 $\{x \in \mathbb{R}^n : f^*(x) \neq 0\} = \bigcup_{i=1}^{\infty}(f^*)^{-1}((1/i,\infty])$에서 $f^* = 0$ (ae.) 이고, 증명은 이로써 충분하다. $\square$

Definition 2.192 적분가능한 함수 $f : \mathbb{R}^n \to \overline{\mathbb{R}}$와 $x \in \mathbb{R}^n$에 대해 $r \downarrow 0$일 때

$$\frac{1}{\lambda_n(B(x,r))} \int_{B(x,r)} |f - f(x)|\, d\lambda_n \to 0$$

이면 이때의 $x \in \mathbb{R}^n$를 f의 **Lebesgue 점 (− point)**라 한다.

Corollary 2.193 적분가능한 함수 $f : \mathbb{R}^n \to \overline{\mathbb{R}}$에 대해 이가 $x \in \mathbb{R}^n$에서 연속이라 하면 x는 Lebesgue 점이다.

Proof 이는 위의 정리의 증명에서 밝힌 f^*의 성질 i로부터 자명하다. $\square$

Hardy-Littlewood의 정리는 적분가능한 함수에 대해 모든 Lebesgue 점을 모은 집합은 영집합임을 함의한다. 뭔가 우리가 원하는 결과가 얻어지는 듯한 느낌이 든다. 그러나 통상적인 미분과 Hardy-Littlewood의 정리를 자연스럽게 연결하기에는 아직 조금 부족하다.

Definition 2.194 집합족 $\mathscr{P}(\mathbb{R}^n)$에 속하는 가측인 집합열 $\{A_i\}$와 점 $x \in \mathbb{R}^n$에 대해 만약 적당한 상수 $C > 0$와 0으로 수렴하는 양의 수열 $\{r_i\}$가 존재하여 각 $i \in \mathbb{N}$에 대해 $A_i \subseteq B(x, r_i)$이고 $\lambda_n(B(x,r_i)) \leq C\lambda_n(A_i)$이면 이때의 집합열 $\{A_i\}$가 x로 **regular하게 수렴한다 (converge regularly)**고 한다.

Theorem 2.195 적분가능한 함수 $f : \mathbb{R}^n \to \overline{\mathbb{R}}$와 이의 Lebesgue 점 $x \in \mathbb{R}^n$, 그리고 x로 regular하게 수렴하는 $\mathscr{P}(\mathbb{R}^n)$에 속하는 가측인 집합열 $\{A_i\}$에 대해

$$\frac{1}{\lambda_n(A_i)} \int_{A_i} |f - f(x)|\, d\lambda_n \to 0$$

이다.

Proof 가정으로부터 적당한 $C > 0$와 0으로 수렴하는 양의 수열 $\{r_i\}$가 존재하여 $A_i \subseteq B(x,r_i)$이고 $\lambda_n(B(x,r_i)) \leq C\lambda_n(A_i)$이므로 임의의 $i \in \mathbb{N}$에 대해

$$\frac{1}{\lambda_n(A_i)} \int_{A_i} |f - f(x)|\, d\lambda_n \leq \frac{C}{\lambda_n(B(x,r_i))} \int_{B(x,r_i)} |f - f(x)|\, d\lambda_n$$

이다. 그런데 Hardy-Littlewood의 정리로부터 위 식의 우변이 $i \to \infty$이면 0으로 수렴하므로 증명이 끝난다. $\square$

Corollary 2.196 적분가능한 함수 $f:\mathbb{R}^n \to \overline{\mathbb{R}}$와 이의 Lebesgue 점 $x \in \mathbb{R}^n$, 그리고 x로 regular하게 수렴하는 $\mathscr{P}(\mathbb{R}^n)$에 속하는 가측인 집합열 $\{A_i\}$에 대해

$$\frac{1}{\lambda_n(A_i)}\int_{A_i} f\,d\lambda_n \to f(x)$$

이다.

Proof 임의의 $i \in \mathbb{N}$에 대해 $|\int_{A_i} f\,d\lambda_n - f(x)\lambda_n(A_i)| = |\int_{A_i} f - f(x)\,d\lambda_n| \le \int_{A_i} |f - f(x)|\,d\lambda_n$ 이므로 따름정리는 자명하다. □

마침내 다음 정리를 얻는다. 이는 일반적인 n차원에서 적분가능한 함수에 대한 정리이므로 우리가 원했던 결과보다 더 강력한 결과이다.

Definition 2.197 열린집합 $U \subseteq \mathbb{R}^n$에서 정의된 함수 $f:U \to \mathbb{R}$와 한 점 $x \in U$에 대해 극한 $\lim_{h\downarrow 0}\Delta_{\prod_{i=1}^n [x_i-h,x_i+h]} f/(2h)^n$이 수렴하면 f는 x에서 **symmetric differentiable**하다고 하고, 이때의 극한값을 $f^{(s)}(x)$로 쓴다. 나아가, 만약 f가 모든 $x \in U$에서 symmetric differentiable하면, 이때의 f를 **symmetric differentiable**하다고 한다.

Corollary 2.198 적분가능한 함수 $f:\mathbb{R}^n \to \overline{\mathbb{R}}$에 대해 함수 $F:\mathbb{R}^n \to \mathbb{R}$를

$$F : x \mapsto \int_{\prod_{i=1}^n (-\infty,x_i]} f\,d\lambda_n$$

로 두면 $F^{(s)} = f$ (ae.)이다. 나아가, $x_0 \in \mathbb{R}^n$의 근방에서 $\partial^n F/\partial x_1 \cdots \partial x_n$이 존재하고 이가 x_0에서 연속이면 $F^{(s)}(x_0) = (\partial^n F/\partial x_1 \cdots \partial x_n)(x_0)$이다. 이때, 편미분의 순서는 중요하지 않다.

Proof 간결한 논의를 위해 $n=2$인 경우만 생각하자. ($n=1$인 경우를 포함한 일반적인 경우에도 이와 비슷하게 보일 수 있다.) 이제 임의의 Lebesgue 점 $x \in \mathbb{R}^2$와 0으로 수렴하는 임의의 수열 $\{h_i\}$를 택하여 $B_x(h_i) = [x_1-h_i, x_1+h_i] \times [x_2-h_i, x_2+h_i]$라 하면 $\{B_x(h_i)\}$가 x로 regular하게 수렴함을 쉽게 보일 수 있다. 그렇다면 따름정리 2.196로부터

$$\begin{aligned}
\frac{\Delta_{B_x(h_i)}F}{(2h_i)^2} &= \frac{1}{(2h_i)^2}\Bigg[\left(\int_{(-\infty,x_1+h_i]\times(-\infty,x_2+h_i]} f\,d\lambda_2 - \int_{(-\infty,x_1-h_i]\times(-\infty,x_2+h_i]} f\,d\lambda_2\right) \\
&\qquad - \left(\int_{(-\infty,x_1+h_i]\times(-\infty,x_2-h_i]} f\,d\lambda_2 - \int_{(-\infty,x_1-h_i]\times(-\infty,x_2-h_i]} f\,d\lambda_2\right)\Bigg] \\
&= \frac{1}{(2h_i)^2}\int_{[x_1-h_i,x_1+h_i]\times[x_2-h_i,x_2+h_i]} f\,d\lambda_2 \\
&= \frac{1}{\lambda_2(B_x(h_i))}\int_{B_x(h_i)} f\,d\lambda_2 \\
&\to f(x)
\end{aligned}$$

이므로 $F^{(s)}(x) = f(x)$ 이고, Hardy-Littlewood의 정리로부터 $\mathbb{R}^n$ 의 거의 대부분의 점들이 이러한 Lebesgue 점이므로 $F^{(s)} = f$ (ae.) 임을 안다.

한편, $x_0 \in \mathbb{R}^n$ 의 근방에서 $\partial^n F/\partial x_1 \cdots \partial x_n$ 이 존재하고 이가 x_0 에서 연속이라 하자. 또한, 편의를 위해 편미분이 $\partial x_1 \partial x_2$ 의 순서로 주어진 경우만 생각하자. (다른 경우에 대해서도 이와 비슷하게 보일 수 있다.) 그렇다면 따름정리 2.193에서 x_0 는 Lebesgue 점이므로 앞선 결과로부터 F 가 x_0 에서 symmetric differentiable하고, MVT와 L'Hôspital의 법칙으로부터

$$
\begin{aligned}
&\lim_{h\downarrow 0}\frac{\Delta_{[x_0^1-h,x_0^1+h]\times[x_0^2-h,x_0^2+h]}F}{(2h)^2}\\
&\quad=\lim_{h\downarrow 0}\frac{1}{2h}\left[\frac{F(x_0^1+h,x_0^2+h)-F(x_0^1+h,x_0^2-h)}{2h}-\frac{F(x_0^1-h,x_0^2+h)-F(x_0^1-h,x_0^2-h)}{2h}\right]\\
&\quad=\lim_{h\downarrow 0}\frac{1}{2h}\left[\frac{\partial F(x_0^1+h,y_h)}{\partial x_2}-\frac{\partial F(x_0^1-h,z_h)}{\partial x_2}\right]\\
&\quad=\lim_{h\downarrow 0}\frac{1}{2}\left[\frac{\partial^2 F(x_0^1+h,y_h)}{\partial x_1\partial x_2}-\frac{\partial^2 F(x_0^1-h,z_h)}{\partial x_1\partial x_2}\right]\\
&\quad=\frac{\partial^2 F}{\partial x_1\partial x_2}(x_0)
\end{aligned}
$$

가 되어 증명이 끝난다. 여기서 y_h, z_h 는 모두 $x_0^2 - h$ 와 $x_0^2 + h$ 를 잇는 선분 위에 존재하는 점들로 MVT에서 그 존재가 보장된다. □

위의 정리에서 $n = 1$ 인 경우를 기존의 FTC의 i에 대입하면 다음을 얻는다.

Theorem (Fundamental theorem of calculus) i. 적분가능한 함수 $f : \mathbb{R} \to \mathbb{R}$ 에 대해 함수 $F : \mathbb{R} \to \mathbb{R}$ 를 $F : x \mapsto \int_{-\infty}^{x} f\, d\lambda_1$ 로 두면 $F^{(s)} = f$ (ae.) 이다.

여기서의 극한은 꽤나 익숙하다. 고등학교 문제집에서 한 번쯤은 보았을 이 극한은 정말 '미분 비슷한 무언가' 이지만 항상 미분과 그 값이 같은 것도 아니고, 이 극한의 존재성이 미분가능성을 함의하는 것도 아니다. 하지만, 이는 f 의 연속성을 포기한 대가이므로 이 정도에 만족하도록 하자.

이제 기존의 FTC의 ii를 살펴보는 것으로 이번 절의 후반부를 시작한다. 앞서 i을 일반화 시킨 것과 비슷하게, F 가 $\mathscr{C}^1$ 급이라는 조건을 F 가 거의 어디서나 symmetric differentiable 하다는 조건으로 바꾸어보자. 아쉽게도 아직은 한참 부족하다. 생각해보면, 이는 굉장히 순진한 시도였다. 만약 F 가 정말 $F(x) = \int_{-\infty}^{x} f\, d\lambda_1$ 을 만족한다면 이는 정리 2.58의 iv와 $F(x) = \int_{-\infty}^{x} f_+\, d\lambda_1 - \int_{-\infty}^{x} f_-\, d\lambda_1 = (\lambda_1)_{f_+}((-\infty, x]) - (\lambda_1)_{f_-}((-\infty, x])$ 에서 반드시 오른쪽 연속인데, 우리는 F 가 거의 어디서나 symmetric differentiable해야 할 것을 요구했을 뿐, 이의 연속성에 대해서는 생각하지 않았기 때문이다.

그렇다면 조금 더 머리를 써서 $F^{(s)} = f$ (ae.) 일 뿐만 아니라 F가 연속일 것까지 요구해 보자. 그러나 아직도 부족하다. 즉, $F^{(s)} = f$ (ae.) 이고 F가 연속이지만 $F(x) \neq \int_{-\infty}^{x} f\, d\lambda_1$ 인 반례가 존재한다.[33] 어쩌면 f의 연속성을 포기한 대가가 너무나도 커서 이를 대체할 좋은 조건이 없는 것은 아닐까? 다행히, 우리의 질문이 이렇게 재미없게 결론나지는 않는다. 지금 우리에게 필요한 적당한 조건은 바로 F의 절대연속성이다.

Definition 2.199 함수 $f : \mathbb{R}^n \to \mathbb{R}$를 생각하자. 임의의 $\varepsilon > 0$에 대해 적당한 $\delta > 0$가 존재하여 임의의 유계이고 서로소인 $B_1, \cdots, B_k \in \mathscr{S}_n$에 대해 $\sum_{i=1}^{k} \mu_n(B_i) < \delta$이면 $\sum_{i=1}^{k} |\Delta_{B_i} f| < \varepsilon$을 만족하면 이때의 f를 **절대연속 (absolutely continuous)**이라 한다.

이전 절에서 Radon-Nikodým 도함수를 도입하며 측도의 절대연속을 정의했는데, 그 용어가 그대로 쓰이는 것은 단순한 우연의 일치가 아니다. 측도의 절대연속성과 함수의 절대연속성은 서로 긴밀히 연결되어 있다.

Theorem 2.200 (Borel-Cantelli) 측도공간 $(\mathbb{R}^n, \mathscr{B}_n, \mu)$에 속하는 임의의 집합열 $\{A_i\}$에 대해 $\sum_{i=1}^{\infty} \mu(A_i) < \infty$이면 $\mu(\limsup_{i\to\infty} A_i) = 0$이다.

Proof 임의의 $\varepsilon > 0$을 택하면 충분히 큰 $i_0 \in \mathbb{N}$에 대해 $\sum_{i=i_0}^{\infty} \mu(A_i) < \varepsilon$인데, 정의로부터 $\limsup_{i\to\infty} A_i \subseteq \bigcup_{i=i_0}^{\infty} A_i$이므로 $\mu(\limsup_{i\to\infty} A_i) \le \mu(\bigcup_{i=i_0}^{\infty} A_i) \le \sum_{i=i_0}^{\infty} \mu(A_i) < \varepsilon$에서 $\mu(\limsup_{i\to\infty} A_i) = 0$임을 안다. □

Theorem 2.201 $\sigma-$유한 측도공간 $(\mathbb{R}^n, \mathscr{B}_n, \mu)$에 대해 함수 $F_\mu : \mathbb{R}^n \to \mathbb{R}_0^+$를 $F_\mu : x \mapsto \mu(\prod_{i=1}^{n} (-\infty, x_i])$로 두자. 그렇다면 F_μ가 절대연속일 필요충분조건은 $\mu \ll \mu_n$인 것이다.

Proof 충분조건임을 보이기 위해서는 $\mu_n(N) = 0$인 임의의 $N \in \mathscr{B}_n$을 고정하고 $\mu(N) = 0$임을 보이면 충분하다. 이제 임의의 $\varepsilon > 0$을 택하면 정리 2.58의 ii와 가정으로부터 임의의 $\delta > 0$가 존재하여 임의의 유계이고 서로소인 $B_1, \cdots, B_k \in \mathscr{S}_n$에 대해 $\sum_{i=1}^{k} \mu_n(B_i) < \delta$이면 $\sum_{i=1}^{k} \mu(B_i) = \sum_{i=1}^{k} \Delta_{B_i} F_\mu < \varepsilon$이다. 한편, 정리 2.41의 i과 2.50로부터 $\mathscr{S}_n$에 속하는 적당한 서로소인 집합열 $\{B_i\}$가 존재하여 $N \subseteq \bigsqcup_{i=1}^{\infty} B_i$이고 $\sum_{i=1}^{\infty} \mu_n(B_i) = \mu_n(\bigsqcup_{i=1}^{\infty} B_i) - \mu_n(N) = \mu_n(\bigsqcup_{i=1}^{\infty} B_i \setminus N) < \delta$이다. 이상으로부터 임의의 $k \in \mathbb{N}$에 대해 $\sum_{i=1}^{k} \mu_n(B_i) < \delta$이므로 $\sum_{i=1}^{k} \mu(B_i) < \varepsilon$이고, 곧 $\sum_{i=1}^{\infty} \mu(B_i) \le \varepsilon$이 되어 $\mu(N) \le \mu(\bigsqcup_{i=1}^{\infty} B_i) = \sum_{i=1}^{\infty} \mu(B_i) \le \varepsilon$에서 $\mu(N) = 0$임을 안다.

다음으로, 필요조건임을 보이기 위해 F_μ가 절대연속이 아니면 $\mu \ll \mu_n$이 아님을 보이자. 정리 2.58의 ii와 가정으로부터 적당한 $\varepsilon > 0$이 존재하여 임의의 $i \in \mathbb{N}$에 대해 유계이고 서로소인 $B_{i1}, \cdots, B_{ik_i} \in \mathscr{S}_n$가 존재하여 $\mu_n(\bigsqcup_{j=1}^{k_i} B_{ij}) = \sum_{j=1}^{k_i} \mu_n(B_{ij}) < 1/i^2$이고 $\mu(\bigsqcup_{j=1}^{k_i} B_{ij}) = \sum_{j=1}^{k_i} \mu(B_{ij}) = \sum_{j=1}^{k_i} \Delta_{B_{ij}} F_\mu \ge \varepsilon$이다. 이로부터 집합열 $\{B_i\}$를 $B_i := \bigsqcup_{j=1}^{k_i} B_{ij}$로 정의하면 $\sum_{i=1}^{\infty} \mu_n(B_i) = \sum_{i=1}^{\infty} \mu_n(\bigsqcup_{j=1}^{k_i} B_{ij}) < \sum_{i=1}^{\infty} 1/i^2 < \infty$이므로 정리 2.26와 Borel-Cantelli의 정리로부터 $\mu_n(\limsup_{i\to\infty} B_i) = 0$이지만 $\mu(\limsup_{i\to\infty} B_i) \ge \limsup_{i\to\infty} \mu(B_i) = \limsup_{i\to\infty} \mu(\bigsqcup_{j=1}^{k_i} B_{ij}) \ge \varepsilon$이 되어 $\mu \ll \mu_n$이 아님을 안다. □

위의 정리는 놀라운 발견이라기보다 이러한 관계가 성립하도록 함수의 절대연속의 정의가 디자인되었다고 보는 것이 맞을 것 같다. (균등연속이나 Lipschitz 연속과 같은 다른 함수의 연속성의 정의에 비해 절대연속의 정의는 다소 복잡하고 기하학적 직관도 부족하다.) 아무튼 이로부터 우리가 원하던 결론을 얻을 수 있다.

Theorem 2.202 함수 $F:\mathbb{R}^n \to \mathbb{R}$가 다음을 만족한다고 하자.

i. 함수 F는 거의 어디서나 symmetric differentiable하다.
ii. 함수 F는 절대연속이고 오른쪽 연속이며 각 변수에 대해 증가한다.
iii. 유계인 semi-open box $B \subseteq \mathbb{R}^n$에 대해 $\Delta_B F \geq 0$이다.
iv. 각 $i \leq n$에 대해 $\lim_{x_i \to -\infty} F(x) = 0$이다.
v. $\lim_{x_1,\cdots,x_n \to \infty} F(x) < \infty$.

그렇다면 $f := F^{(s)}$에 대해 $F = \int_{\prod_{i=1}^n (-\infty, x_i]} f \, d\lambda_n$이다. (단, 여기서 F가 $x \in \mathbb{R}^n$에서 symmetric differentiable하지 않은 경우 $f(x) = 0$으로 생각한다.)

Proof 먼저 정리 2. 59로부터 $\mathscr{B}_n$ 위의 적당한 측도 μ가 존재하여 $F(x) = \mu(\prod_{i=1}^n (-\infty, x_i])$이고, 조건 v와 정리 2. 58의 vii로부터 이때의 μ는 유한하다. 그렇다면 조건 ii와 정리 2. 201로부터 $\mu \ll \mu_n$이 되어 Radon-Nikodým 도함수 $g := d\mu/d\mu_n$가 존재하고 정리 2. 115로부터 $F(x) = \mu(\prod_{i=1}^n (-\infty, x_i]) = \int_{\prod_{i=1}^n (-\infty, x_i]} g \, d\mu_n = \int_{\prod_{i=1}^n (-\infty, x_i]} g \, d\lambda_n$이다. 그런데 우리가 앞서 일반화한 FTC i로부터 $h \downarrow 0$이면 $\lim_{h \downarrow 0} \Delta_{B_x(h)} F/(2h)^n \to g(x)$ (ae.)이므로 $f = g$ (ae.)가 되어 정리 2. 102의 ii로부터 $F = \int_{\prod_{i=1}^n (-\infty, x_i]} f \, d\lambda_n$이다. □

위의 정리에서 $n = 1$인 경우를 기존의 FTC에 대입하면 다음을 얻는다.

Proposition 2.203 절대연속함수 $f:\mathbb{R} \to \mathbb{R}$는 균등연속이다.

Proof 거의 자명하다. 함수 f가 절대연속이므로 임의의 $\varepsilon > 0$에 대해 적당한 $\delta > 0$가 존재하여 임의의 유계인 반열린구간 $I = (x, y) \subseteq \mathbb{R}$에 대해 $|x - y| = \mu_1(I) < \delta$이면 $|f(x) - f(y)| = |\Delta_I f| < \varepsilon$이고, 이로부터 f가 균등연속임이 분명하다. □

Theorem (Fundamental theorem of calculus) ii. 함수 $F:\mathbb{R} \to \mathbb{R}$가 다음을 만족한다고 하자.

a. 함수 F는 거의 어디서나 symmetric differentiable하다.
b. 함수 F는 절대연속인 증가함수이다.
c. $\lim_{x \to -\infty} F(x) = 0$.
d. $\lim_{x \to \infty} F(x) < \infty$.

그렇다면 $f := F^{(s)}$ 에 대해 $F = \int_{-\infty}^{x} f\,d\lambda_1$ 이다. (단, 여기서 F가 $x \in \mathbb{R}$에서 symmetric differentiable하지 않은 경우 $f(x) = 0$으로 생각한다.)

정리 2. 202에서의 조건들이 $n = 1$인 경우에는 많이 간단해진 것을 볼 수 있다. 그렇지만 뭔가 만족스럽지 못한 것도 사실이다. 일반적인 $n \in \mathbb{N}$에 대해 절대연속이라는 개념이 그 이름이 무색하게 오른쪽 연속성조차 함의하지 못하고, FTC의 ii의 확장이라고는 했지만 F에는 여전히 많은 조건들이 붙어있다. 특히 F가 증가함수인 경우로 한정된 것은 누가 보더라도 불필요하게 강하게 주어진 감이 있다. 하지만 나중에 필요한 정리나 개념은 모두 소개했고, 보다 나은 결과를 얻기 위해서는 지금까지의 논의와는 다른 별개의 논의가 필요한데, 이는 하나의 절에 담아내기에는 벅찬 내용이다. 이에 우리는 아쉬움을 뒤로 하고 여기서 멈추어 다음의 FTC를 얻은 것에 만족하기로 한다.[34]

Theorem 2.204 (Fundamental theorem of calculus) 다음이 성립한다.

i. 적분가능한 함수 $f : \mathbb{R} \to \mathbb{R}$에 대해 함수 $F : \mathbb{R} \to \mathbb{R}$를 $F : x \mapsto \int_{-\infty}^{x} f\,d\lambda_1$로 두면 $F^{(s)} = f$ (ae.)이다.
ii. 함수 $F : \mathbb{R} \to \mathbb{R}$가 다음을 만족한다고 하자.
   a. 함수 F는 거의 어디서나 symmetric differentiable하다.
   b. 함수 F는 절대연속인 증가함수이다.
   c. $\lim_{x\to-\infty} F(x) = 0$.
   d. $\lim_{x\to\infty} F(x) < \infty$.

   그렇다면 $f := F^{(s)}$에 대해 $F = \int_{-\infty}^{x} f\,d\lambda_1$이다. (단, 여기서 F가 $x \in \mathbb{R}$에서 symmetric differentiable하지 않은 경우 $f(x) = 0$으로 생각한다.)

2.14 Convolution

이번 절에서는 적분으로 정의되는 유명한 연산인 합성곱에 대한 엄밀한 이론 전개를 해보기로 한다. 이미 대부분의 독자들이 알고 있겠지만 합성곱은 두 함수를 '곱하는' 특유의 방식으로 신호처리에서 두 신호를 합성하는 방법의 일종으로 널리 쓰인다. 우리도 이와 비슷하게 이후 확률론을 배우면서 두 확률변수를 더하는 과정에서 합성곱을 사용하게 될 것이다. 함수의 합성곱과 관련하여 필요한 이론의 전개가 이루어진 후에는 이와 유사하게 측도의 합성곱을 정의하고, 함수의 합성곱과 측도의 합성곱이 어떻게 연결되어 있는지를 살펴보도록 하자.

Proposition 2.205 가측함수 $f : \mathbb{R}^n \to \overline{\mathbb{R}}$와 임의의 $x \in \mathbb{R}^n$에 대해 함수 $t \mapsto f(x-t)$는 가측이다.

Proof 함수 $f_x : \mathbb{R}^n \to \mathbb{R}$와 미분동형사상 $\Phi_x : \mathbb{R}^n \to \mathbb{R}^n$를 각각 $f_x : t \mapsto f(x-t), \Phi_x : t \mapsto x-t$로 두면 변수변환 공식으로부터 $f_x = f \circ \Phi_x$가 가측이므로 이 명제는 자명하다. □

Definition 2.206 가측함수 $f, g : \mathbb{R}^n \to \mathbb{R}$에 대해 집합

$$E(f,g) = \{x \in \mathbb{R}^n : \text{함수 } t \mapsto f(x-t)g(t) \text{가 적분가능하지 않다.}\}$$

를 생각하자. 이제 f와 g의 **합성곱 (convolution)**을 $f * g : \mathbb{R}^n \to \mathbb{R}$로 쓰고

$$f * g : x \mapsto \begin{cases} \int_{\mathbb{R}^n} f(x-t)g(t)\, d\lambda_n(t) & x \notin E(f,g) \text{인 경우} \\ 0 & \text{ow.} \end{cases}$$

로 정의한다.

명제 2.205에서 함수 $t \mapsto f(x-t)g(t)$가 가측이므로 합성곱은 well-define되며 이로써 $f * g$를 아주 일반적으로 정의할 수 있다. 이제 합성곱의 기본적인 성질들을 알아보자.

Theorem 2.207 가측함수 $f, \widetilde{f}, g, \widetilde{g} : \mathbb{R}^n \to \overline{\mathbb{R}}$에 대해 $f = \widetilde{f}$ (ae.)이고 $g = \widetilde{g}$ (ae.)이면 $E(f,g) = E(\widetilde{f}, \widetilde{g})$이고 $f * g = \widetilde{f} * \widetilde{g}$이다.

Proof 임의의 $x \in \mathbb{R}^n$에 대하여 함수 $f_x, \widetilde{f}_x : \mathbb{R}^n \to \mathbb{R}$와 미분동형사상 $\Phi_x : \mathbb{R}^n \to \mathbb{R}^n$를 각각 $f_x : t \mapsto f(x-t), \widetilde{f}_x : t \mapsto \widetilde{f}(x-t), \Phi_x : t \mapsto x-t$로 두면 명제 2.205로부터 f_x와 $\widetilde{f}_x$가 가측이다. 이제 집합 $M = \{x \in \mathbb{R}^n : f(x) \neq \widetilde{f}(x)\}, N = \{x \in \mathbb{R}^n : g(x) \neq \widetilde{g}(x)\}$이 영집합이므로 임의의 $x \in \mathbb{R}^n$를 택하여 집합 $A_x = \{t \in \mathbb{R}^n : (f_x g)(t) \neq (\widetilde{f}_x \widetilde{g})(t)\}$를 생각하면 $A_x \subseteq \{t \in \mathbb{R}^n : x - t \in M\} \cup N = \Phi_x^{-1}(M) \cup N$이다. 그런데 Φ_x^{-1}가 미분동형사상이고 곧 Lipschitz 연속이므로 보조정리 2.172로부터 $\Phi_x^{-1}(M)$이 영집합이 되어 A_x도 영집합이고, 이로부터 $f_x g = \widetilde{f}_x \widetilde{g}$ (ae.)임을 안다. 이는 임의의 $x \in \mathbb{R}^n$에 대해 $f_x g$와 $\widetilde{f}_x \widetilde{g}$의 적분가능성이 일치함을 의미하므로 $E(f,g) = E(\widetilde{f}, \widetilde{g})$가 성립하고, 이어서 $f * g = \widetilde{f} * \widetilde{g}$도 성립한다. □

Theorem 2.208 가측함수 $f, g : \mathbb{R}^n \to \overline{\mathbb{R}}$에 대해 집합 $E(f,g)$는 Borel이고 함수 $f * g$도 Borel이다.

Proof 정리 2.68, 2.88의 ii로부터 적당한 Borel 함수 $\widetilde{f}, \widetilde{g} : \mathbb{R}^n \to \overline{\mathbb{R}}$가 존재하여 $f = \widetilde{f}$ (ae.)이고 $g = \widetilde{g}$ (ae.)이다. 이제 함수 $F : \mathbb{R}^{2n} \to \overline{\mathbb{R}}$를 $F : (x,t) \mapsto \widetilde{f}(x-t)\widetilde{g}(t)$로 두면 명제 2.205의 증명에서와 비슷하게 하여 이가 Borel임을 쉽게 보일 수 있다. 이로부터 함수 $G : \mathbb{R}^n \to \overline{\mathbb{R}}$를 $G : x \mapsto \int_{\mathbb{R}^n} |F(x,t)|\, d\mu(t)$로 정의하면 Fubini의 정리로부터 이는 Borel이고, 정리 2.207로부터 $E(f,g) = E(\widetilde{f}, \widetilde{g}) = G^{-1}(\infty) =: E$도 Borel이다. 증명을 끝내기 위해 함수 $\widetilde{F}^{\pm} : \mathbb{R}^{2n} \to \overline{\mathbb{R}}$를

$$\widetilde{F}^{\pm} : (x,t) \mapsto \begin{cases} F_{\pm}(x,t) & x \notin E \\ 0 & \text{ow.} \end{cases}$$

로 두면 $\widetilde{F}^{\pm}$는 Borel이고, 다시 정리 2.207로부터 임의의 $x \in \mathbb{R}^n$에 대해 $(f*g)(x) = (\widetilde{f}*\widetilde{g})(x) = \int_{\mathbb{R}^n} \widetilde{F}^{+}(x,t)\,d\mu(t) - \int_{\mathbb{R}^n} \widetilde{F}^{-}(x,t)\,d\mu(t)$이므로 Fubini의 정리로부터 $f*g$가 Borel임을 안다. □

이건 사실 조금 신기한 결과이다. 위의 정리에 따르면 비록 f, g 각각은 가측인 함수에 그쳐 다루기 힘든 '이상한' 함수일 수 있지만, $f*g$는 Borel이 되어 그 합성곱은 반드시 '예쁜' 함수로 얻어진다. 이어서 살펴볼 것은 합성곱의 교환법칙과 분배법칙이다.

Theorem 2.209 가측함수 $f, g : \mathbb{R}^n \to \overline{\mathbb{R}}$와 $c \neq 0$에 대해 다음이 성립한다.

i. $E(f,g) = E(cf,g) = E(f,cg)$이고 $c(f*g) = (cf)*g = f*(cg)$이다.
ii. $E(f,g) = E(g,f)$이고 $f*g = g*f$이다.

PROOF i. 우선 $E(f,g) = E(cf,g) = E(f,cg) =: E$임은 정의로부터 자명하고, 임의의 $x \in \mathbb{R}^n \setminus E$에 대해 $[c(f*g)](x) = c\int_{\mathbb{R}^n} f(x-t)g(t)\,d\lambda_n(t) = \int_{\mathbb{R}^n} cf(x-t)g(t)\,d\lambda_n(t) = [(cf)*g](x)$이므로 $c(f*g) = (cf)*g$이다. 이제 이와 비슷하게 $c(f*g) = f*(cg)$임도 보일 수 있다.

ii. 임의의 $x \in \mathbb{R}^n$에 대하여 함수 $f_x, g_x : \mathbb{R}^n \to \overline{\mathbb{R}}$와 미분동형사상 $\Phi_x : \mathbb{R}^n \to \mathbb{R}^n$를 각각 $f_x : t \mapsto f(x-t)$, $g_x : t \mapsto g(x-t)$, $\Phi_x : t \mapsto x-t$로 정의하면 명제 2.205로부터 f_x와 g_x가 가측이다. 이제 임의의 $x \in \mathbb{R}^n$에 대해 $g_x f = (f_x g) \circ \Phi_x$이므로 변수변환 공식으로부터 $f_x g$와 $g_x f$의 적분가능성이 일치하여 $E(f,g) = E(g,f) =: E$이다. 나아가 임의의 $x \in \mathbb{R}^n \setminus E$에 대해 $(f*g)(x) = \int_{\mathbb{R}^n} f_x g\,d\lambda_n = \int_{\mathbb{R}^n} [(f_x g) \circ \Phi_x] |\det \mathbf{D}\Phi_x|\,d\lambda_n = \int_{\mathbb{R}^n} g_x f\,d\lambda_n = (g*f)(x)$이므로 $f*g = g*f$가 성립한다. □

Theorem 2.210 가측함수 $f, g, h : \mathbb{R}^n \to \overline{\mathbb{R}}$에 대해 $E(f,g+h) \subseteq E(f,g) \cup E(f,h)$이고 $\mathbb{R}^n \setminus [E(f,g) \cup E(f,h)]$에서 $f*(g+h) = f*g + f*h$이다. 단, 이때 $g+h$를 계산하는 과정에서 $\infty - \infty$와 같은 부정형은 나오지 않는다고 하자.

PROOF 임의의 $x \in \mathbb{R}^n$에 대하여 함수 $f_x : \mathbb{R}^n \to \mathbb{R}$를 $f_x : t \mapsto f(x-t)$로 정의하면 명제 2.205로부터 f_x가 가측이다. 이제 $|f_x(g+h)| \leq |f_x g| + |f_x h|$이므로 정리 2.100의 i로부터 $E(f,g+h) \subseteq E(f,g) \cup E(f,h) =: E$임을 안다. 한편, 임의의 $x \in \mathbb{R}^n \setminus E$에 대해

$$\begin{aligned}[f*(g+h)](x) &= \int_{\mathbb{R}^n} f(x-t)(g+h)(t)\,d\lambda_n(t) \\ &= \int_{\mathbb{R}^n} f(x-t)g(t)\,d\lambda_n(t) + \int_{\mathbb{R}^n} f(x-t)h(t)\,d\lambda_n(t)\end{aligned}$$

$$= (f*g)(x) + (f*h)(x)$$

이므로 $\mathbb{R}^n \setminus E$ 에서 $f*(g+h) = f*g + f*h$가 성립한다. □

다음으로, 합성곱의 결합법칙을 보자. 방금 살펴본 합성곱의 교환법칙과 분배법칙과는 달리, 결합법칙은 조금 복잡하다.

Theorem 2.211 가측함수 $f, g : \mathbb{R}^n \to \overline{\mathbb{R}}$에 대해 $|f*g| \le |f|*|g|$이다.

Proof 임의의 $x \in \mathbb{R}^n$에 대해 만약 $x \in E(f, g) = E(|f|, |g|)$이면 $(f*g)(x) = (|f|*|g|)(x) = 0$이 되어 정리가 자명하므로 $x \notin E(f, g)$라 하자. 그렇다면 $|(f*g)(x)| = |\int_{\mathbb{R}^n} f(x-t)g(t)\,d\lambda_n(t)| \le \int_{\mathbb{R}^n} |f(x-t)g(t)|\,d\lambda_n(t) = (|f|*|g|)(x)$가 되어 증명이 끝난다. □

Lemma 2.212 Borel 함수 $f, g, h : \mathbb{R}^n \to \overline{\mathbb{R}}$에 대해 $E(f, g), E(g, h)$가 영집합이면 $E(|f|, |g|*|h|) = E(|f|*|g|, |h|)$이다.

Proof 임의의 $x \in \mathbb{R}^n$를 택하여 함수 $F_x : \mathbb{R}^{2n} \to \overline{\mathbb{R}}$를 각각 $F_x : (s, t) \mapsto f(x-t)g(t-s)h(s)$로 두면 명제 2. 205의 증명에서와 비슷하게 하여 이가 Borel임을 쉽게 보일 수 있다. 그렇다면 $\int_{\mathbb{R}^n} |F_x(s, t)|\,d\lambda_n(s) = |f(x-t)| \int_{\mathbb{R}^n} |g(t-s)h(s)|\,d\lambda_n(s) = |f(x-t)|(|g|*|h|)(t)$가 임의의 $t \in \mathbb{R}^n \setminus E(|g|, |h|)$에 대해 성립하는데, 가정으로부터 $E := E(|g|, |h|) = E(g, h)$가 영집합이므로 Fubini의 정리에서

$$\begin{aligned}
||F_x||_1 &= \int_{\mathbb{R}^n} \left[\int_{\mathbb{R}^n} |F_x(s, t)|\,d\lambda_n(s) \right] d\lambda_n(t) \\
&= \int_{\mathbb{R}^n} |f(x-t)| \left[\int_{\mathbb{R}^n} |g(t-s)h(s)|\,d\lambda_n(s) \right] d\lambda_n(t) \\
&= \int_{\mathbb{R}^n \setminus E} |f(x-t)| \left[\int_{\mathbb{R}^n} |g(t-s)h(s)|\,d\lambda_n(s) \right] d\lambda_n(t) \\
&= \int_{\mathbb{R}^n \setminus E} |f(x-t)|(|g|*|h|)(t)\,d\lambda_n(t) \\
&= \int_{\mathbb{R}^n} |f(x-t)|(|g|*|h|)(t)\,d\lambda_n(t)
\end{aligned}$$

가 되어 $x \in E(|f|, |g|*|h|)$와 $||F_x||_1 = \infty$가 서로 동치이다. 이와 비슷하게 $x \in E(|f|*|g|, |h|)$와 $||F_x||_1 = \infty$가 서로 동치임도 보일 수 있으므로 $E(|f|, |g|*|h|) = E(|f|*|g|, |h|)$임을 안다. □

Theorem 2.213 가측함수 $f, g, h : \mathbb{R}^n \to \overline{\mathbb{R}}$에 대해 $E(f, g), E(g, h)$가 영집합이면 $E := E(|f|, |g|*|h|) = E(|f|*|g|, |h|)$이고 $\mathbb{R}^n \setminus E$에서 $f*(g*h) = (f*g)*h$이다.

Proof 정리 2. 68, 2. 88의 ii로부터 적당한 Borel 함수 $\widetilde{f}, \widetilde{g}, \widetilde{h} : \mathbb{R}^n \to \overline{\mathbb{R}}$가 존재하여 $f = \widetilde{f}$ (ae.), $g = \widetilde{g}$ (ae.), $h = \widetilde{h}$ (ae.)이고 정리 2. 207와 위의 보조정리로부터 $E(|f|, |g|*|h|) =$

$E(|\widetilde{f}|,|\widetilde{g}|*|\widetilde{h}|)=E(|\widetilde{f}|*|\widetilde{g}|,|\widetilde{h}|)=E(|f|*|g|,|h|)$ 이다. 이제 임의의 $x\in\mathbb{R}^n\setminus E$ 를 택하여 함수 $F_x:\mathbb{R}^{2n}\to\overline{\mathbb{R}}$ 를 $F_x:(s,t)\mapsto\widetilde{f}(x-t)\widetilde{g}(t-s)\widetilde{h}(s)$ 로 두면 명제 2.205의 증명에서와 비슷하게 하여 이가 Borel임을 쉽게 보일 수 있다. 여기서 만약 $x\in E(\widetilde{f},\widetilde{g}*\widetilde{h})$ 이면 정리 2.211로부터 $\infty=\int_{\mathbb{R}^n}|\widetilde{f}(x-t)||(\widetilde{g}*\widetilde{h})(t)|\,d\lambda_n(t)\le\int_{\mathbb{R}^n}|\widetilde{f}(x-t)|(|\widetilde{g}|*|\widetilde{h}|)(t)\,d\lambda_n(t)<\infty$ 의 모순이 발생하므로 $x\notin E(\widetilde{f},\widetilde{g}*\widetilde{h})$ 이고, 곧 Fubini의 정리, 정리 2.207와 $E(g,h)=E(\widetilde{g},\widetilde{h})$ 가 영집합이라는 점으로부터

$$\begin{aligned}
\int_{\mathbb{R}^{2n}}F_x\,d\lambda_{2n}&=\int_{\mathbb{R}^n}\left[\int_{\mathbb{R}^n}F_x(s,t)\,d\lambda_n(s)\right]d\lambda_n(t)\\
&=\int_{\mathbb{R}^n}\widetilde{f}(x-t)\left[\int_{\mathbb{R}^n}\widetilde{g}(t-s)\widetilde{h}(s)\,d\lambda_n(s)\right]d\lambda_n(t)\\
&=\int_{\mathbb{R}^n\setminus E(\widetilde{g},\widetilde{h})}\widetilde{f}(x-t)\left[\int_{\mathbb{R}^n}\widetilde{g}(t-s)\widetilde{h}(s)\,d\lambda_n(s)\right]d\lambda_n(t)\\
&=\int_{\mathbb{R}^n\setminus E(\widetilde{g},\widetilde{h})}\widetilde{f}(x-t)(\widetilde{g}*\widetilde{h})(t)\,d\lambda_n(t)\\
&=\int_{\mathbb{R}^n}\widetilde{f}(x-t)(\widetilde{g}*\widetilde{h})(t)\,d\lambda_n(t)\\
&=[\widetilde{f}*(\widetilde{g}*\widetilde{h})](x)\\
&=[f*(g*h)](x)
\end{aligned}$$

이다. 한편, 비슷한 방법으로 $\int_{\mathbb{R}^{2n}}F_x\,d\lambda_{2n}=[(f*g)*h](x)$ 임을 보일 수 있으므로 $\mathbb{R}^n\setminus E$ 에서 $f*(g*h)=(f*g)*h$ 임을 안다. □

비록 앞서 얻은 분배법칙과 결합법칙은 훌륭한 결과이지만, 각각이 성립하기 위한 조건들 때문에 실용성이 떨어지는 것이 사실이다. 이러한 조건들을 생략하기 위해서는 합성곱을 취하는 함수들의 종류를 조금 제한할 필요가 있다. 즉, 지금까지는 임의의 두 가측함수에 합성곱을 취하는 것을 생각했다면, 이를 조금 제한하여 합성곱을 취해 보자는 것이다. 물론, 이때 너무 제한적으로 그 종류를 제한하면 또다시 실용성을 잃고 말겠지만, 다행히 다음 정리는 적당히 적분가능한 함수로 제한하는 것만으로 우리의 목적을 달성할 수 있음을 말해준다.

Theorem 2.214 (Young's convolution inequality) 임의의 $1/p+1/q=1+1/r$ 인 $p,q,r\ge1$ 과 임의의 $f\in L^p(\lambda_n),g\in L^q(\lambda_n)$ 에 대해 $E(f,g)$ 는 영집합이고 $||f*g||_r\le||f||_p||q||_q$ 이다. 따라서 $f*g\in L^r$ 이다.

Proof 함수 $h:\mathbb{R}^n\to\overline{\mathbb{R}}_0^+$ 를 $h:x\mapsto\int_{\mathbb{R}^n}|f(x-t)g(t)|\,d\lambda_n(t)$ 로 두면 이는 명제 2.205로부터 well-define되고 $E(f,g)=E(|f|,|g|)=h^{-1}(\infty)$ 이다. 한편, 정리 2.211로부터 $|f*g|\le h$ 이므로 $||h||_r\le||f||_p||g||_q$ 만 보일 수 있으면 $||f*g||_r\le||f||_p||g||_q$ 는 자명해지고 $h\in L^r$ 이 되어 정리 2.100의 ii로부터 h가 거의 어디서나 유한하므로 증명이 끝난다. 이를 위해

임의의 $x \in \mathbb{R}^n$를 택하여 함수 $f_x : \mathbb{R}^n \to \overline{\mathbb{R}}$를 $f_x : t \mapsto f(x-t)$로 두고 $\lambda = p(1-1/q)$라 하면 명제 2.205로부터 f_x는 가측이고 $0 \le \lambda < 1$이다. 먼저 $\lambda > 0$인 경우를 생각하여 $s = p/\lambda$라 하면 $1/q + 1/s = 1$이므로 Hölder의 부등식으로부터 $h(x) = \int_{\mathbb{R}^n} |f_x|^{\lambda} |f_x|^{1-\lambda} |g| \, d\lambda_n \le |||f_x|^{1-\lambda}|g|||_q |||f_x|^{\lambda}||_s$이고, 곧

$$\begin{aligned}
[h(x)]^q &\le \left[\int_{\mathbb{R}^n} |f_x|^{(1-\lambda)q} |g|^q \, d\lambda_n \right] \left(\int_{\mathbb{R}^n} |f_x|^{\lambda s} \, d\lambda_n \right)^{q/s} \\
&= \left[\int_{\mathbb{R}^n} |f_x|^{(1-\lambda)q} |g|^q \, d\lambda_n \right] \left(\int_{\mathbb{R}^n} |f_x|^{p} \, d\lambda_n \right)^{\lambda q/p} \\
&= \left[\int_{\mathbb{R}^n} |f_x|^{(1-\lambda)q} |g|^q \, d\lambda_n \right] ||f_x||_p^{\lambda q}
\end{aligned}$$

이며 이는 $\lambda = 0$인 경우에도 성립함을 쉽게 알 수 있다. 그렇다면 $||f_x||_p = ||f||_p$임은 자명하므로 Minkowski의 적분부등식으로부터 $s' = r/q$에 대해

$$\begin{aligned}
||h||_r^q &= \left(\int_{\mathbb{R}^n} h^r \, d\lambda_n \right)^{q/r} \\
&= \left(\int_{\mathbb{R}^n} h^{qs'} \, d\lambda_n \right)^{1/s'} \\
&\le ||f||_p^{\lambda q} \left\{ \int_{\mathbb{R}^n} \left[\int_{\mathbb{R}^n} |f(x-t)|^{(1-\lambda)q} |g(t)|^q \, d\lambda_n(t) \right]^{s'} d\lambda_n(x) \right\}^{1/s'} \\
&\le ||f||_p^{\lambda q} \int_{\mathbb{R}^n} \left[\int_{\mathbb{R}^n} |f(x-t)|^{(1-\lambda)qs'} |g(t)|^{qs'} \, d\lambda_n(x) \right]^{1/s'} d\lambda_n(t) \\
&= ||f||_p^{\lambda q} \int_{\mathbb{R}^n} \left[\int_{\mathbb{R}^n} |f(x)|^{(1-\lambda)qs'} \, d\lambda_n(x) \right]^{1/s'} |g(t)|^q \, d\lambda_n(t) \\
&= ||f||_p^{\lambda q} ||f||_p^{(1-\lambda)q} ||g||_q^q \\
&= (||f||_p ||g||_q)^q
\end{aligned}$$

에서 $||h||_r \le ||f||_p ||g||_q$임을 안다. 여기서 마지막에서 두 번째 등호는 $(1-\lambda)qs' = (1-\lambda)r = p$에서 성립한다. □

Corollary 2.215 적분가능한 $f, g : \mathbb{R}^n \to \overline{\mathbb{R}}$에 대해 $E(f, g)$는 영집합이고 $f * g$도 적분가능하다.

Proof 이는 Young의 합성곱 부등식에서 $p = q = r = 1$인 특별한 경우이다. □

따라서 f, g가 모두 적분가능한 경우에는 분배법칙과 결합법칙의 조건들을 생략하여 다음과 같이 간단히 쓸 수 있다. 이와 더불어 L^1 공간은 합성곱 $*$에 대해 닫혀있다는 이상의 결론은 기억해둘만한 가치가 있다.

Corollary 2.216 적분가능한 $f, g, h : \mathbb{R}^n \to \overline{\mathbb{R}}$와 $c \in \mathbb{R}$에 대해 다음이 성립한다.

i. (교환법칙) $c(f*g) = (cf)*g = f*(cg)$.
ii. (교환법칙) $f*g = g*f$.
iii. (결합법칙) $f*(g*h) = (f*g)*h$ (ae.).
iv. (분배법칙) $f*(g+h) = f*g + f*h$ (ae.).

Proof 이는 정리 2. 209, 2. 210, 2. 213에 따름정리 2. 215를 적용한 결과로 자명하다. □

이제 측도의 합성곱을 정의하는 것으로 이번 절의 후반부를 시작한다.

Proposition 2.217 Borel 함수 $f : \mathbb{R}^n \to \overline{\mathbb{R}}$에 대해 함수 $F : \mathbb{R}^{2n} \to \overline{\mathbb{R}}$를 $F : (x, y) \mapsto f(x+y)$로 정의하면 이는 Borel이다.

Proof 함수 $g : \mathbb{R}^{2n} \to \mathbb{R}^n$를 $g : (x, y) \mapsto x+y$로 두면 이가 연속이므로 곧 Borel이고, 정리 2. 73에서 $F = f \circ g$는 Borel이다. □

Definition 2.218 Borel $\sigma-$대수 $\mathscr{B}_n$ 위의 측도 μ, ν에 대해 μ와 ν의 **합성곱 (convolution)**을 $\mu * \nu : \mathscr{B}_n \to \overline{\mathbb{R}}_0^+$로 쓰고 $\mu * \nu : A \mapsto \int_{\mathbb{R}^{2n}} \mathbf{1}_A(x+y)\, d(\mu \otimes \nu)(x, y)$로 정의한다.

Proposition 2.219 Borel $\sigma-$대수 $\mathscr{B}_n$ 위의 측도 μ, ν에 대해 함수 $\mu * \nu$는 측도이다.

Proof 우선 $(\mu * \nu)(\emptyset) = 0$임은 분명하고, $\mathscr{B}_n$에 속하는 임의의 서로소인 집합열 $\{A_i\}$에 대해 MCT로부터

$$\begin{aligned}(\mu * \nu)\left(\bigsqcup_{i=1}^{\infty} A_i\right) &= \int_{\mathbb{R}^{2n}} \mathbf{1}_{\bigsqcup_{i=1}^{\infty} A_i}(x+y)\, d(\mu \otimes \nu)(x, y) \\ &= \int_{\mathbb{R}^{2n}} \sum_{i=1}^{\infty} \mathbf{1}_{A_i}(x+y)\, d(\mu \otimes \nu)(x, y) \\ &= \sum_{i=1}^{\infty} \int_{\mathbb{R}^{2n}} \mathbf{1}_{A_i}(x+y)\, d(\mu \otimes \nu)(x, y) \\ &= \sum_{i=1}^{\infty} (\mu * \nu)(A_i)\end{aligned}$$

이므로 $\mu * \nu$는 측도이다. □

명제 2. 217로부터 $\mathbf{1}_A(x+y)$가 Borel이므로 측도의 합성곱은 well-define된다. 다음 순서는 측도의 합성곱의 기본적인 성질들을 살피는 것이다. 다행히 측도의 합성곱의 성질들은 함수의 합성곱의 성질들보다 쉽게 얻어진다.

Theorem 2.220 Borel $\sigma-$대수 $\mathscr{B}_n$ 위의 측도 μ, ν, ξ에 대해 $\xi = \mu * \nu$일 필요충분조건은 임의의 음이 아닌 Borel 함수 $f : \mathbb{R}^n \to \overline{\mathbb{R}}_0^+$에 대해 $\int_{\mathbb{R}^n} f\, d\xi = \int_{\mathbb{R}^{2n}} f(x+y)\, d(\mu \otimes \nu)(x, y)$인 것이다.

Proof 충분조건임을 보이기 위해 먼저 f가 단순함수인 경우를 생각하여 이의 표준형을 $f = \sum_{i=1}^k a_i \mathbf{1}_{A_i}$라 하면 $\int_{\mathbb{R}^n} f\, d\xi = \sum_{i=1}^k a_i \xi(A_i) = \sum_{i=1}^k a_i \int_{\mathbb{R}^{2n}} \mathbf{1}_{A_i}(x+y)\, d(\mu \otimes \nu)(x, y) = \int_{\mathbb{R}^{2n}} f(x+y)\, d(\mu \otimes \nu)(x, y)$이다. 이제 일반적인 음이 아닌 Borel 함수 f를 생각하면 정리 2.83로부터 적당한 단순함수 $f_i : \mathbb{R}^n \to \mathbb{R}_0^+$의 열 $\{f_i\}$가 존재하여 $f_i \uparrow f$이므로 MCT로부터 $\int_{\mathbb{R}^n} f\, d\xi = \lim_{i\to\infty} \int_{\mathbb{R}^n} f_i\, d\xi = \lim_{i\to\infty} \int_{\mathbb{R}^{2n}} f_i(x+y)\, d(\mu \otimes \nu)(x, y) = \int_{\mathbb{R}^{2n}} f(x+y)\, d(\mu \otimes \nu)(x, y)$가 되어 충분조건임이 보여진다.

반대로, 필요조건임을 보이는 것은 쉽다. 임의의 $A \in \mathscr{B}_n$에 대해 $\mathbf{1}_A$이 Borel 함수이므로 $\xi(A) = \int_{\mathbb{R}^n} \mathbf{1}_A\, d\xi = \int_{\mathbb{R}^{2n}} \mathbf{1}_A(x+y)\, d(\mu \otimes \nu)(x, y) = (\mu * \nu)(A)$에서 필요조건임이 분명하다. □

Theorem 2.221 Borel $\sigma-$대수 $\mathscr{B}_n$ 위의 측도 μ, ν, ξ와 $c \geq 0$에 대해 다음이 성립한다.

i. (교환법칙) $c(\mu * \nu) = (c\mu) * \nu = \mu * (c\nu)$.
ii. (교환법칙) $\mu * \nu = \nu * \mu$.
iii. (결합법칙) $\mu * (\nu * \xi) = (\mu * \nu) * \xi$.
iv. (분배법칙) $\mu * (\nu + \xi) = \mu * \nu + \mu * \xi$.

Proof 임의의 $A \in \mathscr{B}_n$를 택하자.

i. Fubini의 정리로부터

$$\begin{aligned}
[c(\mu * \nu)](A) &= c \int_{\mathbb{R}^{2n}} \mathbf{1}_A(x+y)\, d(\mu \otimes \nu)(x, y) \\
&= c \int_{\mathbb{R}^n} \left[\int_{\mathbb{R}^n} \mathbf{1}_A(x+y)\, d\mu(x) \right] d\nu(y) \\
&= \int_{\mathbb{R}^n} \left[\int_{\mathbb{R}^n} \mathbf{1}_A(x+y)\, d(c\mu)(x) \right] d\nu(y) \\
&= \int_{\mathbb{R}^{2n}} \mathbf{1}_A(x+y)\, d[(c\mu) \otimes \nu](x, y) \\
&= [(c\mu) * \nu](A)
\end{aligned}$$

에서 $c(\mu * \nu) = (c\mu) * \nu$이고 이와 비슷하게 $c(\mu * \nu) = \mu * (c\nu)$도 보일 수 있다.

ii. Fubini의 정리로부터

$$\begin{aligned}
(\mu * \nu)(A) &= \int_{\mathbb{R}^{2n}} \mathbf{1}_A(x+y)\, d(\mu \otimes \nu)(x, y) \\
&= \int_{\mathbb{R}^n} \left[\int_{\mathbb{R}^n} \mathbf{1}_A(x+y)\, d\mu(x) \right] d\nu(y)
\end{aligned}$$

$$
\begin{aligned}
&= \int_{\mathbb{R}^n} \left[\int_{\mathbb{R}^n} \mathbf{1}_A(x+y)\, d\mu(y) \right] d\nu(x) \\
&= \int_{\mathbb{R}^{2n}} \mathbf{1}_A(x+y)\, d(\nu \otimes \mu)(x,y) \\
&= (\nu * \mu)(A)
\end{aligned}
$$

에서 $\mu * \nu = \nu * \mu$ 이다.

iii. Fubini의 정리와 정리 2.220로부터

$$
\begin{aligned}
[\mu * (\nu * \xi)](A) &= \int_{\mathbb{R}^{2n}} \mathbf{1}_A(x+y)\, d[\mu \otimes (\nu * \xi)](x,y) \\
&= \int_{\mathbb{R}^n} \left[\int_{\mathbb{R}^n} \mathbf{1}_A(x+y)\, d\mu(x) \right] d(\nu * \xi)(y) \\
&= \int_{\mathbb{R}^{2n}} \left[\int_{\mathbb{R}^n} \mathbf{1}_A(x+y+z)\, d\mu(x) \right] d(\nu \otimes \xi)(y,z) \\
&= \int_{\mathbb{R}^n} \left\{ \int_{\mathbb{R}^n} \left[\int_{\mathbb{R}^n} \mathbf{1}_A(x+y+z)\, d\mu(x) \right] d\nu(y) \right\} d\xi(z) \\
&= \int_{\mathbb{R}^n} \left[\int_{\mathbb{R}^{2n}} \mathbf{1}_A(x+y+z)\, d(\mu \otimes \nu)(x,y) \right] d\xi(z) \\
&= \int_{\mathbb{R}^n} \left[\int_{\mathbb{R}^n} \mathbf{1}_A(x+z)\, d(\mu * \nu)(x) \right] d\xi(z) \\
&= \int_{\mathbb{R}^{2n}} \mathbf{1}_A(x+z)\, d[(\mu * \nu) \otimes \xi](x,z) \\
&= [(\mu * \nu) * \xi](A)
\end{aligned}
$$

에서 $\mu * (\nu * \xi) = (\mu * \nu) * \xi$ 이다.

iv. Fubini의 정리로부터

$$
\begin{aligned}
[\mu * (\nu + \xi)](A) &= \int_{\mathbb{R}^{2n}} \mathbf{1}_A(x+y)\, d[\mu \otimes (\nu + \xi)](x,y) \\
&= \int_{\mathbb{R}^n} \left[\int_{\mathbb{R}^n} \mathbf{1}_A(x+y)\, d\mu(x) \right] d(\nu + \xi)(y) \\
&= \int_{\mathbb{R}^n} \left[\int_{\mathbb{R}^n} \mathbf{1}_A(x+y)\, d\mu(x) \right] d\nu(y) + \int_{\mathbb{R}^n} \left[\int_{\mathbb{R}^n} \mathbf{1}_A(x+y)\, d\mu(x) \right] d\xi(y) \\
&= \int_{\mathbb{R}^{2n}} \mathbf{1}_A(x+y)\, d(\mu \otimes \nu)(x,y) + \int_{\mathbb{R}^{2n}} \mathbf{1}_A(x+y)\, d(\mu \otimes \xi)(x,y) \\
&= (\mu * \nu)(A) + (\mu * \xi)(A)
\end{aligned}
$$

에서 $\mu * (\nu + \xi) = \mu * \nu + \mu * \xi$ 이다. □

한편, 함수의 합성곱과 측도의 합성곱은 다음 정리에 의해 서로 연결되어 있다.

Theorem 2.222 음이 아닌 적분가능한 Borel 함수 $f, g : \mathbb{R}^n \to \overline{\mathbb{R}}_0^+$ 에 대해 $(\mu_n)_f * (\mu_n)_g = (\mu_n)_{f*g}$ 이다.

Proof 함수 $F : \mathbb{R}^{2n} \to \overline{\mathbb{R}}_0^+$ 를 $F : (x, y) \mapsto f(x-y)$ 로 두면 명제 2. 217의 증명에서와 비슷하게 하여 이가 Borel임을 쉽게 보일 수 있다. 이제 임의의 $A \in \mathscr{B}_n$ 를 택하면 Fubini의 정리로부터

$$
\begin{aligned}
[(\mu_n)_f * (\mu_n)_g](A) &= \int_{\mathbb{R}^{2n}} \mathbf{1}_A(x+y)\, d[(\mu_n)_f \otimes (\mu_n)_g](x, y) \\
&= \int_{\mathbb{R}^n} \left[\int_{\mathbb{R}^n} \mathbf{1}_A(x+y)\, d(\mu_n)_f(x) \right] d(\mu_n)_g(y) \\
&= \int_{\mathbb{R}^n} \left[\int_{\mathbb{R}^n} f(x)g(y)\mathbf{1}_A(x+y)\, d\mu_n(x) \right] d\mu_n(y) \\
&= \int_{\mathbb{R}^n} \left[\int_{\mathbb{R}^n} f(x-y)g(y)\mathbf{1}_A(x)\, d\mu_n(x) \right] d\mu_n(y) \\
&= \int_{\mathbb{R}^n} \left[\int_{\mathbb{R}^n} f(x-y)g(y)\mathbf{1}_A(x)\, d\mu_n(y) \right] d\mu_n(x)
\end{aligned}
$$

인데, 정리 2. 208와 따름정리 2. 215로부터 $E(f, g)$ 가 영집합이고 $f * g$ 도 음이 아닌 적분가능한 Borel 함수이므로 $[(\mu_n)_f * (\mu_n)_g](A) = \int_{\mathbb{R}^n} \mathbf{1}_A(f*g)\, d\mu_n = \int_{\mathbb{R}^n} \mathbf{1}_A\, d(\mu_n)_{f*g} = (\mu_n)_{f*g}(A)$ 이다. □

나중에 사용할 정리 하나를 증명하는 것으로 이번 절을 끝마치도록 하자.

Theorem 2.223 Borel $\sigma-$대수 $\mathscr{B}_n$ 위의 측도 μ, ν 와 임의의 $A \in \mathscr{B}_n$ 에 대해 함수 $x \mapsto \mu(A-x)$ 는 Borel이고 $(\mu * \nu)(A) = \int_{\mathbb{R}^n} \mu(A-x)\, d\nu(x)$ 이다.

Proof Fubini의 정리로부터

$$
\begin{aligned}
(\mu * \nu)(A) &= \int_{\mathbb{R}^{2n}} \mathbf{1}_A(x+y)\, d(\mu \otimes \nu)(y, x) \\
&= \int_{\mathbb{R}^n} \left[\int_{\mathbb{R}^n} \mathbf{1}_A(x+y)\, d\mu(y) \right] d\nu(x) \\
&= \int_{\mathbb{R}^n} \left[\int_{\mathbb{R}^n} \mathbf{1}_{A-x}(y)\, d\mu(y) \right] d\nu(x) \\
&= \int_{\mathbb{R}^n} \mu(A-x)\, d\nu(x)
\end{aligned}
$$

이고 이때 함수 $x \mapsto \mu(A-x)$ 가 Borel임이 분명하다. □

2.15 Epilog

본 장의 마지막을 장식할 이번 절에서는 별개의 절로 구성할 정도의 내용은 아니지만, 그래도 놓치기는 아까운 주제들에 대해 가볍게 알아보도록 하겠다. 다루게 될 주제는 측도의 분해, 셈측도, 그리고 벡터함수와 복소함수의 적분으로 크게 3가지이다.

2.15.1 Lebesgue's decomposition theorem

우리는 주어진 임의의 측도를 보다 단순하거나 특정한 구조를 가지는 측도들의 합으로 분해할 수 있도록 해 주는 정리들을 통틀어 흔히 '분해정리' 라 부른다. 이런 분해정리들은 다분히 이론적인 도구로, 마치 인수분해를 통해 정수의 성질을 더 깊이 탐구할 수 있는 것처럼 분해정리들로 측도를 적당히 분해함으로써 측도론의 깊이있는 이론을 전개할 수 있다. 이렇게 측도론에서 큰 의의를 가지는 정리들이지만 분해정리가 본격적으로 필요할 정도의 추상적인 측도론은 지나치다는 생각에 아쉽게도 이 책에서는 분해정리가 빛을 보지 못한 것이 사실이다. 그럼에도 불구하고, 이를 에필로그에서라도 뒤늦게 소개하는 까닭은 이가 확률론에서 그 자체로 나름 함의하는 바가 있고, 그 함의가 어쩌면 고등학교 때부터 궁금했을 수도 있는 '왜 이산확률변수와 연속확률변수에만 관심을 가지는가? ' 라는 질문에 대한 훌륭한 답변이 되기 때문이다. 그 답변은 확률의 현실적인 응용에서는 이산확률변수와 연속확률변수로 충분하다는 식의 공학도의 마인드가 물씬 풍기는 타협적인 답변보다 훨씬 지적으로 우아하고 만족스러울 것이다.

일단 이번 절은 에필로그인 만큼 그 답변을 듣는 것은 다음 장으로 잠시 미루어두고, 여기서는 분해정리를 증명하고, 내친 김에 이를 Radon-Nikodým 도함수에 간단히 응용하는 정도로 만족하도록 하자. 이제 singular라는 개념과 함께 앞서 미쳐 살펴보지 못했던 측도의 절대연속에 관한 성질을 살펴보는 것으로 논의를 시작한다.

Definition 2.224 가측공간 $(X, \mathscr{A})$ 위의 측도 μ, ν에 대해 적당한 $A \in \mathscr{A}$가 존재하여 $\mu(A) = \nu(A^c) = 0$이면 이때 μ와 ν는 서로 **singular**하다고 하고 $\mu \perp \nu$로 쓴다.

측도 μ와 ν가 서로 singular하다는 것은, μ와 ν가 측도를 부여하는 데 있어 이 둘은 서로 공통된 정보를 사용하지 않는다는, 따라서 서로 독립적이라는 정도의 의미이다. 즉, 정의로부터 적당한 $A \in \mathscr{A}$가 존재하여 $\mu(A) = \nu(A^c) = 0$인데, 임의의 $B \in \mathscr{A}$를 생각하면 $\mu(B) = \mu(B \cap A) + \mu(B \cap A^c) = \mu(B \cap A^c)$이고 $\nu(B) = \nu(B \cap A) + \nu(B \cap A^c) = \nu(B \cap A)$이므로 μ는 $B \cap A^c$의 정보만으로 B에 측도를 부여하고, ν는 $B \cap A$의 정보만으로 B에 측도를 부여하여 이 둘은 B의 측도를 부여하는 데 있어 서로 공통된 정보를 사용하지 않는다.

한편, 앞서 소개한 절대연속은 이와 정반대의 의미이다. 예컨대 $\mu \ll \nu$라 하고 적당한 $A \in \mathscr{A}$에 대해 $\nu(A^c) = 0$이라 하면 ν는 측도를 부여하는 데 있어 A에 속한 정보만을 사용한다고 생각할 수 있다. 그런데 정의로부터 $\mu(A^c) = 0$이므로 μ 또한 A에 속한 정보만을 사용하고, A 밖의 정보는 사용하지 못한다. 즉, μ는 측도를 부여하는 데 있어 철저하게 ν가 사용하는 정보만 사용할 수 있고, 곧 μ는 ν에 종속적이다.

Proposition 2.225 가측공간 $(X, \mathscr{A})$ 위의 측도 μ, ν, ξ에 대해 다음이 성립한다.

i. 만약 $\mu \perp \xi$이고 $\nu \perp \xi$이면 $(\mu+\nu) \perp \xi$이다.
ii. 만약 $\mu \ll \xi$이고 $\nu \ll \xi$이면 $(\mu+\nu) \ll \xi$이다.
iii. 만약 $\mu \ll \xi$이고 $\nu \perp \xi$이면 $\mu \perp \nu$이다.
iv. 만약 $\mu \ll \xi$이고 $\mu \perp \xi$이면 $\mu = 0$이다.

Proof i. 가정으로부터 적당한 $A, B \in \mathscr{A}$가 존재하여 $\mu(A) = \nu(B) = \xi(A^c) = \xi(B^c) = 0$이므로 $\mu(A \cap B) \leq \mu(A) = 0, \nu(A \cap B) \leq \nu(B) = 0, \xi((A \cap B)^c) = \xi(A^c \cup B^c) \leq \xi(A^c) + \xi(B^c) = 0$이고, 이로부터 $(\mu+\nu) \perp \xi$이다.

ii. 가정으로부터 임의의 $N \in \mathscr{A}$에 대해 $\xi(N) = 0$이면 $(\mu+\nu)(N) = \mu(N) + \nu(N) = 0$이므로 $(\mu+\nu) \ll \xi$임이 자명하다.

iii. 가정으로부터 적당한 $A \in \mathscr{A}$가 존재하여 $\nu(A) = \xi(A^c) = 0$인데, 다시 가정으로부터 $\mu(A^c) = 0$이 되어 $\mu \perp \nu$이다.

iv. 가정으로부터 적당한 $A \in \mathscr{A}$가 존재하여 $\mu(A) = \xi(A^c) = 0$인데, 다시 가정으로부터 $\mu(A^c) = 0$이므로 $\mu(X) = \mu(A \sqcup A^c) = \mu(A) + \mu(A^c) = 0$에서 $\mu = 0$이다. □

많은 분해정리 중에서 우리가 특히 주목할 것은 Lebesgue의 분해정리이다.

Theorem 2.226 (Lebesgue's decomposition theorem) 가측공간 $(X, \mathscr{A})$ 위의 σ−유한 측도 μ, ν를 생각하면 ν에 대해 절대연속인 $\mathscr{A}$ 위의 측도 μ_{ac}와 ν와 singular한 $\mathscr{A}$ 위의 측도 μ_{s}가 유일하게 존재하여 $\mu = \mu_{\mathrm{ac}} + \mu_{\mathrm{s}}$이다.

Proof 먼저 존재성을 보이자. 이를 위해 측도 $\xi = \mu + \nu$를 생각하면 이는 $\mathscr{A}$ 위의 σ−유한 측도이고 $\mu, \nu \ll \xi$임이 자명하므로 Radon-Nikodým 도함수 $f := d\mu/d\xi, g := d\nu/d\xi$가 존재한다. 이제 집합 $A = \{x \in X : g(x) > 0\}$를 생각하여 $\mu_{\mathrm{ac}} := \mu_A, \mu_{\mathrm{s}} := \mu_{A^c}$로 두면 이들은 모두 $\mathscr{A}$ 위의 측도이며 $\mu = \mu_{\mathrm{ac}} + \mu_{\mathrm{s}}$이다. 따라서 $\mu_{\mathrm{ac}} \ll \nu$와 $\mu_{\mathrm{s}} \perp \nu$만 보이면 존재성이 보여지는데, $\mu_{\mathrm{s}}(A) = \mu(A \cap A^c) = 0$이고 $\nu(A^c) = \int_{A^c} g\, d\xi = 0$이므로 $\mu_{\mathrm{s}} \perp \nu$임은 쉽게 알 수 있다. 다음으로, $\nu(N) = 0$인 임의의 $N \in \mathscr{A}$을 고정하면 $\int_X g\mathbf{1}_N\, d\xi = \int_N g\, d\xi = \nu(N) = 0$이므로 정리 2. 101의 ii로부터 $g\mathbf{1}_N = 0$ (ξ−ae.)이다. 이로부터 $g\mathbf{1}_{A\cap N} = 0$ (ξ−ae.)인데 $A \cap N$ 위에서 $g > 0$이므로 $\xi(A \cap N) = 0$이 되어 정리 2. 102의 i로부터 $\mu_{\mathrm{ac}}(N) = \mu(A \cap N) = \int_{A\cap N} f\, d\xi = 0$이고, 곧 $\mu_{\mathrm{ac}} \ll \nu$임을 안다.

다음으로, 유일성을 보이기 위해 $\mu_{\rm ac}, \mu'_{\rm ac} \ll \nu$ 이고 $\mu_{\rm s}, \mu'_{\rm s} \perp \nu$ 인 $\mathscr{A}$ 위의 측도 $\mu_{\rm ac}, \mu'_{\rm ac}$, $\mu_{\rm s}, \mu'_{\rm s}$ 가 존재하여 $\mu = \mu_{\rm ac} + \mu_{\rm s} = \mu'_{\rm ac} + \mu'_{\rm s}$ 라 하면 적당한 $A, A' \in \mathscr{A}$ 가 존재하여 $\mu_{\rm s}(A) = \mu'_{\rm s}(A') = \nu(A^c) = \nu(A'^c) = 0$ 이다. 이제 임의의 $B \in \mathscr{A}$ 를 고정하면 $\mu_{\rm s}(B \cap A \cap A') \le \mu_{\rm s}(A) = 0$, $\mu'_{\rm s}(B \cap A \cap A') \le \mu'_{\rm s}(A') = 0$ 이고, 이로부터 $\mu_{\rm ac}(B \cap A \cap A') = \mu(B \cap A \cap A') = \mu'_{\rm ac}(B \cap A \cap A')$ 임을 안다. 나아가 $\nu(B \cap (A \cap A')^c) \le \nu((A \cap A')^c) = \nu(A^c \cup A'^c) \le \nu(A) + \nu(A'^c) = 0$ 이므로 $\mu_{\rm ac}(B \cap (A \cap A')^c) = \mu'_{\rm ac}(B \cap (A \cap A')^c) = 0$ 이 되어 $\mu_{\rm s}(B \cap (A \cap A')^c) = \mu(B \cap (A \cap A')^c) = \mu'_{\rm s}(B \cap (A \cap A')^c)$ 임을 안다. 이상을 종합하면

$$
\begin{aligned}
\mu_{\rm ac}(B) &= \mu_{\rm ac}(B \cap A \cap A') + \mu_{\rm ac}(B \cap (A \cap A')^c) \\
&= \mu_{\rm ac}(B \cap A \cap A') \\
&= \mu'_{\rm ac}(B \cap A \cap A') \\
&= \mu'_{\rm ac}(B \cap A \cap A') + \mu'_{\rm ac}(B \cap (A \cap A')^c) \\
&= \mu'_{\rm ac}(B)
\end{aligned}
$$

에서 $\mu_{\rm ac} = \mu'_{\rm ac}$ 가 보여지고 비슷하게 $\mu_{\rm s} = \mu'_{\rm s}$ 도 보여지므로 유일성이 증명된다. □

약간의 노력으로 Lebesgue 분해정리를 조금 더 refine할 수 있다.

Definition 2.227 측도공간 $(X, \mathscr{A}, \mu)$ 와 집합 $A \in \mathscr{A}$ 에 대해 $\mu(A^c) = 0$ 이면 이때의 집합 A 를 측도 μ 의 **지지집합 (support)** 이라 하고 $\operatorname{supp}\mu = A$ 로 쓴다. 나아가, μ 가 가산인 지지집합을 가지면 이때의 측도 μ 를 **이산측도 (discrete measure)** 라 한다.

Definition 2.228 가측공간 $(X, \mathscr{A})$ 위의 측도 μ, ν 를 생각하고 X 의 모든 한원소 집합이 $\mathscr{A}$ 에 속한다고 하자. 만약 $\mu \perp \nu$ 이고 X 의 모든 한원소 집합이 μ–영집합이면 이때 μ 는 ν 에 대해 **특이연속 (singular continuous)** 이라 한다.

일반적인 측도는 전체 공간 X 의 측도가 X 전반에 걸쳐 연속적으로 퍼져있다면 이산측도는 가산개의 점에만 전체 공간의 측도가 띄엄띄엄 집중되어있는 느낌이다. 흔히 이렇게 측도가 집중된 듯한 점들을 point mass라 부르곤 한다. 한편, μ 가 ν 에 대해 특이연속이라는 것은 μ 와 ν 가 서로 독립적일 뿐만 아니라 μ 가 point mass를 가지지 않음을 의미한다.

Lemma 2.229 σ–유한 측도공간 $(X, \mathscr{A}, \mu)$ 에 대해 X 의 모든 한원소 집합이 $\mathscr{A}$ 에 속하다면 $(X, \mathscr{A}, \mu)$ 는 가산개의 point mass를 가진다. 즉, 집합 $\{x \in X : \mu\{x\} > 0\}$ 은 가산이다.

Proof 주어진 집합을 A 라 하고 모순을 유도하기 위해 A 가 비가산이라 하자. 그렇다면 가정으로부터 $\mathscr{A}$ 에 속하는 적당한 집합열 $\{B_i\}$ 가 존재하여 각 $i \in \mathbb{N}$ 에 대해 $\mu(B_i) < \infty$ 이고 $\bigcup_{i=1}^{\infty} B_i = X$ 이므로 $A \cap B_{i_0}$ 가 비가산인 $i_0 \in \mathbb{N}$ 가 적어도 하나 존재하여 집합열 $\{C_i\}$ 를 $C_i := \{x \in B_{i_0} : \mu\{x\} > 1/i\}$ 로 두면 이는 증가하는 집합열이 되어 $C_i \uparrow A \cap B_{i_0}$ 이다. 그

렇다면 다시 적당한 $i_1 \in \mathbb{N}$ 이 존재하여 C_{i_1} 이 비가산이다. 이제 여기서 서로다른 가산개의 원소를 뽑아 수열 $\{x_i\}$ 를 구성하면 $\mu\{x_1, x_2, \cdots\} = \sum_{i=1}^{\infty} \mu\{x_i\} > \sum_{i=1}^{\infty} 1/i_1 = \infty$ 인 동시에 $\mu\{x_1, x_2, \cdots\} \leq \mu(C_{i_1}) \leq \mu(B_{i_0}) < \infty$ 가 되어 모순이 발생한다. 증명은 이로써 충분하다. □

Theorem 2.230 (Lebesgue's decomposition theorem) 기측공간 $(X, \mathscr{A})$ 에 대해 $\mathscr{A}$ 위의 $\sigma-$유한 측도 μ, ν 를 생각하고 X 의 모든 한원소 집합이 $\mathscr{A}$ 에 속한다고 하자. 그렇다면 ν 에 대해 절대연속인 $\mathscr{A}$ 위의 측도 μ_{ac}, ν 에 대해 특이연속인 $\mathscr{A}$ 위의 측도 μ_{sc}, $\mathscr{A}$ 위의 이산측도 μ_{pp} 가 유일하게 존재하여 $\mu = \mu_{\mathrm{ac}} + \mu_{\mathrm{sc}} + \mu_{\mathrm{pp}}$ 이다.

Proof 먼저 존재성을 보이자. 앞서 보인 Lebesgue의 분해정리로부터 $\mu = \mu_{\mathrm{ac}} + \mu_{\mathrm{s}}$ 이고 $\mu_{\mathrm{ac}} \ll \nu, \mu_{\mathrm{s}} \perp \nu$ 인 $\mathscr{A}$ 위의 측도 $\mu_{\mathrm{ac}}, \mu_{\mathrm{s}}$ 가 유일하게 존재한다. 여기서 μ 가 $\sigma-$유한하므로 μ_{s} 도 $\sigma-$유한하여 집합 $A = \{x \in X : \mu_{\mathrm{s}}\{x\} > 0\}$ 를 생각하면 위의 보조정리로부터 이는 가산이다. 이제 측도 $\mu_{\mathrm{pp}} = (\mu_{\mathrm{s}})_A, \mu_{\mathrm{sc}} = (\mu_{\mathrm{s}})_{A^c}$ 를 생각하면 이는 $\mathscr{A}$ 위의 측도이고 $\mu = \mu_{\mathrm{ac}} + \mu_{\mathrm{sc}} + \mu_{\mathrm{pp}}$ 가 되어 μ_{sc} 가 ν 에 대해 특이연속이고 μ_{pp} 가 이산측도라는 점만 보이면 존재성이 보여진다. 우선 $\mu_{\mathrm{s}} \perp \nu$ 에서 적당한 $B \in \mathscr{A}$ 가 존재하여 $\mu_{\mathrm{s}}(B) = \nu(B^c) = 0$ 이므로 $\mu_{\mathrm{sc}}(B) = \mu_{\mathrm{s}}(B \cap A^c) \leq \mu_{\mathrm{s}}(B) = 0$ 에서 $\mu_{\mathrm{sc}} \perp \nu$ 임을 안다. 나아가 임의의 $x \in X$ 에 대해 만약 $x \in A$ 라면 $\mu_{\mathrm{sc}}\{x\} = \mu_{\mathrm{s}}(\{x\} \cap A^c) = 0$ 이고 반대로 $x \notin A$ 라도 $\mu_{\mathrm{sc}}\{x\} = \mu_{\mathrm{s}}\{x\} = 0$ 이므로 곧 어느 경우에나 $\mu_{\mathrm{sc}}\{x\} = 0$ 이 되어 μ_{sc} 는 ν 에 대해 특이연속이다. 한편, $\mu_{\mathrm{pp}}(A^c) = \mu_{\mathrm{s}}(A^c \cap A) = 0$ 에서 가산인 A 가 μ_{pp} 의 지지집합이 되어 이가 이산측도임도 쉽게 알 수 있다.

다음으로, 유일성을 보이기 위해서는 ν 에 대해 특이연속인 $\mathscr{A}$ 위의 측도 $\mu_{\mathrm{sc}}, \mu'_{\mathrm{sc}}$ 와 $\mathscr{A}$ 위의 이산측도 $\mu_{\mathrm{pp}}, \mu'_{\mathrm{pp}}$ 에 대해 $\mu_{\mathrm{s}} = \mu_{\mathrm{sc}} + \mu_{\mathrm{pp}} = \mu'_{\mathrm{sc}} + \mu'_{\mathrm{pp}}$ 이면 $\mu_{\mathrm{sc}} = \mu'_{\mathrm{sc}}, \mu_{\mathrm{pp}} = \mu'_{\mathrm{pp}}$ 임을 보이면 충분하다. 여기서 μ_{pp} 와 μ'_{pp} 의 가산 지지집합을 각각 $A, A' \in \mathscr{A}$ 라 하면 $\mu_{\mathrm{pp}}((A \cup A')^c) = \mu_{\mathrm{pp}}(A^c \cap A'^c) \leq \mu_{\mathrm{pp}}(A^c) = 0$ 이고 비슷하게 $\mu'_{\mathrm{pp}}((A \cup A')^c) = 0$ 이므로 WLOG, 필요하다면 A 와 A' 를 $A \cup A'$ 로 바꾸어 μ_{pp} 와 μ'_{pp} 가 공통의 가산 지지집합 $A \in \mathscr{A}$ 를 가진다고 해도 된다. 이제 임의의 $B \in \mathscr{A}$ 를 고정하면 $\mu_{\mathrm{pp}}(B \cap A^c) \leq \mu_{\mathrm{pp}}(A^c) = 0, \mu'_{\mathrm{pp}}(B \cap A^c) \leq \mu'_{\mathrm{pp}}(A^c) = 0$ 이다. 나아가, A 의 가산개의 원소를 나열하여 $x_1, x_2, \cdots$ 라 하면 $\mu_{\mathrm{sc}}(B \cap A) \leq \mu_{\mathrm{sc}}(A) = \sum_{i=1}^{k} \mu_{\mathrm{sc}}\{x_i\} = 0$ 이고 (여기서 k 는 유한할 수도 있고, ∞ 일 수도 있다.) 비슷하게 $\mu'_{\mathrm{sc}}(B \cap A) = 0$ 이므로 곧 $\mu_{\mathrm{pp}}(B \cap A) = \mu_{\mathrm{s}}(B \cap A) = \mu'_{\mathrm{pp}}(B \cap A)$ 가 되어 이상으로부터

$$\begin{aligned}
\mu_{\mathrm{pp}}(B) &= \mu_{\mathrm{pp}}(B \cap A) + \mu_{\mathrm{pp}}(B \cap A^c) \\
&= \mu_{\mathrm{pp}}(B \cap A) \\
&= \mu'_{\mathrm{pp}}(B \cap A) \\
&= \mu'_{\mathrm{pp}}(B \cap A) + \mu'_{\mathrm{pp}}(B \cap A^c) \\
&= \mu'_{\mathrm{pp}}(B)
\end{aligned}$$

에서 $\mu_{\text{pp}} = \mu'_{\text{pp}}$ 임이 보여진다. 한편, 가정으로부터 적당한 $C, C' \in \mathscr{A}$ 가 존재하여 $\mu_{\text{sc}}(C) = \mu'_{\text{sc}}(C') = \nu(C^c) = \nu(C'^c) = 0$ 인데, 앞선 결과로부터 $\mu_{\text{sc}}(A \cup C) \leq \mu_{\text{sc}}(A) + \mu_{\text{sc}}(C) = 0$ 이고, 비슷하게 $\mu'_{\text{sc}}(A \cup C') = 0$ 이므로 WLOG, 필요하다면 C, C' 를 각각 $A \cup C, A \cup C'$ 로 바꾸어 $A \subseteq C, C'$ 라 해도 된다. 그렇다면 $\mu_{\text{sc}}(B \cap C \cap C') \leq \mu_{\text{sc}}(C) = 0$, $\mu'_{\text{sc}}(B \cap C \cap C') \leq \mu'_{\text{sc}}(C') = 0$ 이고 $\mu'_{\text{pp}}(B \cap (C \cap C')^c) = \mu_{\text{pp}}(B \cap (C \cap C')^c) \leq \mu_{\text{pp}}(A^c) = 0$ 이므로 $\mu_{\text{sc}}(B \cap (C \cap C')^c) = \mu_{\text{s}}(B \cap (C \cap C')^c) = \mu'_{\text{sc}}(B \cap (C \cap C')^c)$ 이다. 이상을 종합하면

$$\begin{aligned}
\mu_{\text{sc}}(B) &= \mu_{\text{sc}}(B \cap C \cap C') + \mu_{\text{sc}}(B \cap (C \cap C')^c) \\
&= \mu_{\text{sc}}(B \cap C \cap C') \\
&= \mu'_{\text{sc}}(B \cap C \cap C') \\
&= \mu'_{\text{sc}}(B \cap C \cap C') + \mu'_{\text{sc}}(B \cap (C \cap C')^c) \\
&= \mu'_{\text{sc}}(B)
\end{aligned}$$

에서 $\mu_{\text{sc}} = \mu'_{\text{sc}}$ 도 보여지므로 유일성이 증명된다. $\square$

따라서 Lebesgue의 분해정리는 임의의 측도 μ 를 기준 측도 ν 에 종속적인 측도 하나, ν 와 서로 독립이고 point mass 없는 측도 하나, point mass로만 이루어진 측도 하나로 분해시켜주며, 이들을 각각 **절대연속성분 (absolutely continuous component)**, **특이연속성분 (singular continuous component)**, **순수 점 성분 (pure point component)**이라 한다. 이제 예고했던 바와 같이 Lebesgue의 분해정리를 간단히 응용하여 Radon-Nikodým 도함수를 일반화시켜보자.

Definition 2.231 가측공간 $(X, \mathscr{A})$ 위의 σ−유한 측도 μ, ν 에 대해 μ 가 $\mu_{\text{ac}} \ll \nu$ 이고 $\mu_{\text{s}} \perp \nu$ 인 $\mathscr{A}$ 위의 측도 $\mu_{\text{ac}}, \mu_{\text{s}}$ 로써 $\mu = \mu_{\text{ac}} + \mu_{\text{s}}$ 와 같이 Lebesgue 분해된다고 하자. 이때 μ 의 ν 에 대한 **Radon-Nikodým 도함수 (− derivative)**를 $d\mu_{\text{ac}}/d\nu$ 로 정의하고 $d\mu/d\nu$ 로 쓴다.

위의 정의에서 만약 $\mu \ll \nu$ 이면 Lebesgue 분해의 유일성으로부터 $\mu_{\text{ac}} = \mu$, $\mu_{\text{s}} = 0$ 이므로 위의 정의에서의 Radon-Nikodým 도함수가 정의 2. 157에서의 Radon-Nikodým 도함수와 같아진다. 그러나, 이렇게 일반화된 Radon-Nikodým 도함수에 대해 따름정리 2. 158는 다소 불완전한 형태로만 성립한다.

Corollary 2.232 가측공간 $(X, \mathscr{A})$ 위의 σ−유한 측도 μ, ν 와 임의의 음이 아닌 가측함수 $f : X \to \overline{\mathbb{R}}_0^+$ 에 대해 $\int_X f\,d\mu \geq \int_X f(d\mu/d\nu)\,d\nu$ 이다.

Proof 측도 μ 가 $\mu_{\text{ac}} \ll \nu$ 이고 $\mu_{\text{s}} \perp \nu$ 인 $\mathscr{A}$ 위의 측도 $\mu_{\text{ac}}, \mu_{\text{s}}$ 로써 $\mu = \mu_{\text{ac}} + \mu_{\text{s}}$ 와 같이 Lebesgue 분해된다고 하자. 그렇다면 $\int_X f\,d\mu = \int_X f\,d(\mu_{\text{ac}} + \mu_{\text{s}}) = \int_X f\,d\mu_{\text{ac}} + \int_X f\,d\mu_{\text{s}} \geq \int_X f\,d\mu_{\text{ac}} = \int_X f(d\mu/d\nu)\,d\nu$ 이다. $\square$

2.15.2 Counting measure

여태껏 우리는 Lebesgue 측도와 Borel 측도에 지대한 관심을 기울여왔는데, 확률론에서는 이만큼이나 중요한 역할을 하는 '셈측도' 라는 측도가 하나 더 있다. 이렇게나 중요한 셈측도를 Lebesgue 측도와 달리 본 장의 끝에서야 소개하는 것은 이가 매우 단순한 측도이기 때문이다. 별도의 구성 과정도 필요없고, 관련된 성질들도 우리가 이미 알고 있는 것을 측도론의 언어로 번역한 것에 불과하다.

Definition 2.233 가측공간 $(X,\mathscr{A})$에 대해 다음과 같이 정의되는 함수 $\#:\mathscr{A}\to\overline{\mathbb{R}}_0^+$를 $\mathscr{A}$ 위의 **셈측도** (counting measure)라 한다.

$$\#:A\mapsto\begin{cases}|A| & A\text{가 유한한 경우}\\ \infty & \text{ow.}\end{cases}$$

Proposition 2.234 가측공간 $(X,\mathscr{A})$에 대해 셈측도 $\#$는 측도이다. 따라서 $(X,\mathscr{A},\#)$는 측도공간을 이룬다.

Proof 우선 $\#(\emptyset)=0$임은 분명하다. 이제 $\mathscr{A}$에 속하는 임의의 서로소인 집합열 $\{A_i\}$를 생각하자. 만약 $\{A_i\}$가 공집합이 아닌 집합을 무한히 많이 포함하거나 적어도 하나의 무한집합을 포함한다면 $\bigsqcup_{i=1}^{\infty}A_i$가 무한집합이므로 $\#(\bigsqcup_{i=1}^{\infty}A_i)=|\bigsqcup_{i=1}^{\infty}A_i|=\infty=\sum_{i=1}^{\infty}|A_i|=\sum_{i=1}^{\infty}\#(A_i)$임이 자명하다. 반대로 $\{A_i\}$가 오직 유한개의 유한집합으로만 이루어져있다면 적당한 $i_0\in\mathbb{N}$에 대해 $i>i_0$이면 A_i가 공집합이고 $i\le i_0$이면 A_i는 유한집합이므로 $\bigsqcup_{i=1}^{\infty}A_i=\bigsqcup_{i=1}^{i_0}A_i$가 유한집합이 되어 $\#(\bigsqcup_{i=1}^{\infty}A_i)=\#(\bigsqcup_{i=1}^{i_0}A_i)=|\bigsqcup_{i=1}^{i_0}A_i|=\sum_{i=1}^{i_0}|A_i|=\sum_{i=1}^{i_0}\#(A_i)=\sum_{i=1}^{\infty}\#(A_i)$이다. 이상으로부터 $\#$가 측도임을 안다. □

보통 특별히 정하지 않는 이상 집합 X에서의 셈측도는 $\mathscr{P}(X)$ 위에서 정의된 것으로 보고, 만약 집합 X도 주어지지 않았다면 $(\mathbb{N},\mathscr{P}(\mathbb{N}))$에서 정의된 것으로 본다.

Proposition 2.235 측도공간 $(X,\mathscr{A},\#)$에서 영집합은 $\emptyset$ 뿐이다.

Proof 임의의 $N\in\mathscr{A}$에 대해 셈측도의 정의로부터 $\#(N)=0$은 $N=\emptyset$를 뜻하므로 이는 자명하다. □

이로부터 셈측도가 주어진 측도공간에서 거의 어디서나 성립하는 성질은 곧 항상 성립한다.

Theorem 2.236 집합 X를 가산이라 하면 측도공간 $(X,\mathscr{A},\#)$ 위에서 정의된 음이 아닌 가측함수 $f:X\to\overline{\mathbb{R}}_0^+$에 대해 $\int_X f\,d\#=\sum_{x\in X}f(x)$이다. 한편, 적분가능한 (혹은 $\sum_{x\in X}f(x)$가 well-define되는[35]가측함수) $f:X\to\overline{\mathbb{R}}$에 대해서도 동일한 결과가 성립하며 이때 $\sum_{x\in X}f(x)$는 well-define된다 (혹은 f는 적분가능하다).

Proof 간결한 논의를 위해 X가 무한집합인 경우에 대해서만 보이도록 하자. (집합 X가 유한집합인 경우에도 이와 비슷하게 보일 수 있다.) 집합 X의 원소를 나열하여 $\{x_i\}$라 하라 하고 함수열 $\{f_j\}$를 $f_j = \sum_{i=1}^{j} f(x_i)\mathbf{1}_{\{x_i\}}$로 두면 $\int_X f_j\, d\# = \sum_{i=1}^{j} f(x_i) \int_X \mathbf{1}_{\{x_i\}}\, d\# = \sum_{i=1}^{j} f(x_i)\#\{x_i\} = \sum_{i=1}^{j} f(x_i)$이다. 그렇다면 $f_j \uparrow f$에서 MCT로부터 $\int_X f\, d\# = \lim_{j\to\infty} \int_X f_j\, d\# = \lim_{j\to\infty} \sum_{i=1}^{j} f(x_i) = \sum_{x\in X} f(x)$에서 정리가 성립한다.

한편, 적분가능한 $f : X \to \overline{\mathbb{R}}$에 대해서는 우선 정리 2.100의 ii, 명제 2.235로부터 f가 $\pm\infty$를 그 함숫값으로 갖지 아니하여 $\infty - \infty$와 같은 부정형이 합의 과정에서 등장하지 않고, 정리 2.100의 i로부터 $|f|$가 적분가능하여 앞선 결과로부터 $\sum_{x\in X} |f(x)| = \int_X |f|\, d\# < \infty$에서 $\sum_{x\in X} f(x)$가 절대수렴하므로 $\sum_{x\in X} |f(x)|$는 well-define된다. 이제 $Y = f^{-1}([0,\infty)), Z = f^{-1}((-\infty, 0))$에 대해 $\int_X f\, d\# = \int_X f_+\, d\# - \int_X f_-\, d\# = \sum_{x\in X} f_+ - \sum_{x\in X} f_- = \sum_{x\in Y} f(x) + \sum_{x\in Z} f(x) = \sum_{x\in X} f(x)$이므로 증명이 끝나고, $\sum_{x\in X} f(x)$가 well-define되는 가측함수 $f : X \to \overline{\mathbb{R}}$에 대해서도 비슷하게 하면 된다. □

이를 조금 더 친숙한 $\mathbb{N}$에서의 결과로 옮기면 다음과 같다.

Corollary 2.237 함수 $f : \mathbb{N} \to \overline{\mathbb{R}}$에 대해 수열 $\{a_i\}$를 $a_i := f(i)$로 두자. 만약 $\sum_{i=1}^{\infty} a_i$가 양항급수이거나 절대수렴하면 $\int_{\mathbb{N}} f\, d\# = \sum_{i=1}^{\infty} a_i$이다.

Proof 이는 정리 2.236와 가측공간 $(\mathbb{N}, \mathscr{P}(\mathbb{N}))$ 위에서 정의된 함수 $f : \mathbb{N} \to \overline{\mathbb{R}}$는 모두 가측이라는 점에서 자명하다. □

따라서 셈측도가 주어진 측도공간에 있어 적분은 합과 같다. 담담하게 말했지만, 이는 정말 흥분되는 결과이다. 아마 다들 합은 이산세계의 적분이고, 적분은 연속세계의 합이라는 수학적 직관을 어느정도 가지고 있을 것인데, 위의 결과는 이러한 직관을 확인시켜주는 동시에 더 이상 합과 적분을 구별하여 다룰 필요가 없음을 뜻한다. 우리가 적분을 통해 얻는 모든 결과는 자연스럽게 합에 대한 결과를 따름정리의 형태로 함의한다.

Corollary 2.238 다음이 성립한다.

i. (MCT) 음이 아닌 수열 $\{a_{ij}\}$에 대해 임의의 $i, j \in \mathbb{N}$에 대해 $a_{ij} \le a_{ij+1}$이고, 적당한 음이 아닌 수열 $\{a_i\}$가 존재하여 각 $i \in \mathbb{N}$에 대해 $j \to \infty$이면 $a_{ij} \uparrow a_i$라 하자. 그렇다면 $j \to \infty$일때 $\sum_{i=1}^{\infty} a_{ij} \uparrow \sum_{i=1}^{\infty} a_i$이다.
ii. (Fatou의 보조정리) 음이 아닌 수열 $\{a_{ij}\}$에 대해 $\liminf_{j\to\infty} \sum_{i=1}^{\infty} a_{ij} \ge \sum_{i=1}^{\infty} \liminf_{j\to\infty} a_{ij}$이다.
iii. (DCT) 수열 $\{a_{ij}\}$에 대해 적당한 수열 $\{a_i\}$가 존재하여 각 $i \in \mathbb{N}$에 대해 $j \to \infty$이면 $a_{ij} \to a_i$라 하자. 또한, 적당한 음이 아닌 수열 $\{M_i\}$가 존재하여 $\sum_{i=1}^{\infty} M_i$가 수렴하고 모든 $i, j \in \mathbb{N}$에 대해 $|a_{ij}| \le M_i$라 하자. 그렇다면 각 $j \in \mathbb{N}$에 대해 $\sum_{i=1}^{\infty} a_{ij}$와 $\sum_{i=1}^{\infty} a_i$는 절대수렴하고 $j \to \infty$이면 $\sum_{i=1}^{\infty} a_{ij} \to \sum_{i=1}^{\infty} a_i$이다.

iv. (Fubini) 음이 아닌 수열 $\{a_{ij}\}$ 에 대해 $\sum_{(i,j)\in\mathbb{N}^2} a_{ij} = \sum_{i=1}^{\infty}\sum_{j=1}^{\infty} a_{ij} = \sum_{j=1}^{\infty}\sum_{i=1}^{\infty} a_{ij}$ 이다. 한편, $\sum_{(i,j)\in\mathbb{N}^2} a_{ij}$ 가 절대수렴하는 수열 $\{a_{ij}\}$ 에 대해서도 동일한 결과가 성립하며 이때 $\sum_{i=1}^{\infty} a_{ij}, \sum_{j=1}^{\infty} a_{ij}$ 는 절대수렴하고 수열 $\{b_j\}, \{c_i\}$ 를 각각 $b_j := \sum_{i=1}^{\infty} a_{ij}$, $c_i := \sum_{j=1}^{\infty} a_{ij}$ 라 하면 $\sum_{j=1}^{\infty} b_j, \sum_{i=1}^{\infty} c_i$ 도 절대수렴한다.

Proof 이는 셈측도에 각각 MCT, Fatou의 보조정리, DCT, Fubini의 정리를 적용한 결과로 정리 2. 236와 가측공간 $(\mathbb{N}, \mathscr{P}(\mathbb{N}))$ 위에서 정의된 함수 $f : \mathbb{N} \to \overline{\mathbb{R}}$ 는 모두 가측이라는 점에서 자명하다. □

Corollary 2.239 다음이 성립한다.

i. 열린집합 $U \subseteq \mathbb{R}^n$ 위에서 정의된 연속함수 $f_i : U \to \mathbb{R}$ 의 열 $\{f_i\}$ 가 다음 조건을 만족한다고 하자.
   a. 임의의 $x \in U$ 에 대해 $\sum_{i=1}^{\infty} f_i(x)$ 가 절대수렴한다.
   b. 적당한 음이 아닌 수열 $\{M_i\}$ 가 존재하여 $\sum_{i=1}^{\infty} M_i$ 가 수렴하고 각 $i \in \mathbb{N}_0$ 에 대해 $|f_i| \leq M_i$ 이다.

   그렇다면 함수 $x \mapsto \sum_{i=1}^{\infty} f_i(x)$ 는 연속이다.
ii. (Leibniz의 법칙) 열린집합 $U \subseteq \mathbb{R}$ 위에서 정의된 미분가능한 함수 $f_i : U \to \mathbb{R}$ 의 열 $\{f_i\}$ 가 다음 조건을 만족한다고 하자.
   a. 임의의 $x \in U$ 에 대해 $\sum_{i=1}^{\infty} f_i(x)$ 가 절대수렴한다.
   b. 적당한 음이 아닌 수열 $\{M_i\}$ 가 존재하여 $\sum_{i=1}^{\infty} M_i$ 가 수렴하고 각 $i \in \mathbb{N}_0$ 에 대해 $|f_i'| \leq M_i$ 이다.

   그렇다면 함수 $x \mapsto \sum_{i=1}^{\infty} f_i(x)$ 는 미분가능하고 임의의 $x \in U$ 에 대해 $\sum_{i=1}^{\infty} f'(x)$ 가 절대수렴하며 다음이 성립한다.

$$\frac{d}{dx}\sum_{i=1}^{\infty} f_i(x) = \sum_{i=1}^{\infty} f_i'(x)$$

Proof 이는 셈측도에 각각 따름정리 2. 108, Leibniz의 법칙을 적용한 결과로 정리 2. 236와 가측공간 $(\mathbb{N}, \mathscr{P}(\mathbb{N}))$ 위에서 정의된 함수 $f : \mathbb{N} \to \overline{\mathbb{R}}$ 는 모두 가측이라는 점에서 자명하다. □

마지막으로, Radon-Nikodým 도함수가 셈측도에 대해 어떻게 적용되는지를 살펴보자.

Proposition 2.240 가측공간 $(X, \mathscr{A})$ 위의 임의의 측도 μ 와 셈측도 $\#$에 대해 $\mu \ll \#$이다.

Proof 명제 2. 235로부터 $\#(N) = 0$ 인 집합 $N \in \mathscr{A}$ 은 공집합뿐이고, 따라서 $\mu \ll \#$이다. □

Theorem 2.241 측도공간 $(X, \mathscr{A}, \mu)$ 에 대해 X 가 가산이고 이의 모든 한원소 집합이 $\mathscr{A}$ 에 속한다면 $(d\mu/d\#)(x) = \mu\{x\}$ 이다.

PROOF 임의의 $A \in \mathscr{A}$ 에 대해 이가 가산이므로 $\mu(A) = \sum_{x \in A} \mu\{x\} = \int_A \mu\{x\}\, d\#$ 이고, Radon-Nikodým 정리와 명제 2.235, 2.240로부터 $(d\mu/d\#)(x) = \mu\{x\}$ 이다. □

2.15.3 Integration of vector/complex-valued functions

지금까지의 우리는 공역이 $\overline{\mathbb{R}}$ 인 함수의 적분만 생각해왔는데, 때로는 (주로 표기상의 편의를 위해) 공역이 $\mathbb{R}^n$ 이거나 심지어 $\mathbb{C}$ 인 함수 즉, 벡터함수나 복소함수의 적분을 생각해야 하는 경우도 있다. 함숫값으로 벡터는 그렇다 쳐도 복소수를 가진다니 복소해석학을 배우지도 않은 상태에서 이게 가능할까 싶을 수도 있지만, 여기서는 아주 간단한 내용만 다룰 것이므로 걱정할 필요는 없다.

Definition 2.242 측도공간 $(X, \mathscr{A}, \mu)$ 에서 정의된 가측함수 $f : X \to \mathbb{R}^n$ 에 대해 f 의 모든 성분이 적분가능하면 이때 f 를 **(μ−Lebesgue) 적분가능 (− integrable)** 하다고 한다. 나아가 적분가능한 $f : X \to \mathbb{R}^n$ 에 대해 μ 에 대한 f 의 **(Lebesgue) 적분 (− integral)** 을 $\int_X f(x)\, d\mu(x)$ 혹은 간단히 $\int_X f\, d\mu$ 로 쓰고 $\int_X f\, d\mu = (\int_X f_i\, d\mu)_i$ 로 정의한다.

앞서 정리 2.99에서 보였던 적분의 기본적인 성질들도 대부분 비슷하게 성립한다.

Corollary 2.243 측도공간 $(X, \mathscr{A}, \mu)$ 에서 정의된 적분가능한 함수 $f, g : X \to \mathbb{R}^n$ 에 대해 다음이 성립한다.

i. 임의의 $c \in \mathbb{R}$ 에 대해 cf 가 적분가능하고 $\int_X cf\, d\mu = c \int_X f\, d\mu$ 이다.
ii. 함수 $f + g$ 가 적분가능하고 $\int_X f + g\, d\mu = \int_X f\, d\mu + \int_X g\, d\mu$ 이다.
iii. 임의의 $x \in \mathbb{R}^n$ 에 대해 $\langle x, f \rangle$ 가 적분가능하고 $\int_X \langle x, f \rangle\, d\mu = \langle x, \int_X f\, d\mu \rangle$ 이다.
iv. $||\int_X f\, d\mu|| \le \int_X ||f||\, d\mu$.

PROOF i. ii. 이는 정의와 적분의 기본적인 성질로부터 자명하다.

iii. 각 f_i 가 적분가능하므로 $\langle x, f \rangle = \sum_{i=1}^n x_i f_i$ 가 적분가능함은 분명하고, 곧 $\int_X \langle x, f \rangle\, d\mu = \int_X \sum_{i=1}^n x_i f_i\, d\mu = \sum_{i=1}^n x_i \int_X f_i\, d\mu = \langle x, \int_X f\, d\mu \rangle$ 이다.

iv. 표기의 편의를 위해 $I = \int_X f\, d\mu$ 라 하자. 나아가 만약 $I = 0$ 이면 부등식이 자명하므로 $I \ne 0$ 이라 하자. 그렇다면 Cauchy-Schwarz 부등식으로부터 $||I||^2 = \langle I, \int_X f\, d\mu \rangle = \int_X \langle I, f \rangle\, d\mu \le \int_X ||I||||f||\, d\mu = ||I|| \int_X ||f||\, d\mu$ 가 되어 정리가 성립한다. □

Corollary 2.244 (Fundamental theorem of calculus) 다음이 성립한다.

i. 적분가능한 연속함수 $f : \mathbb{R} \to \mathbb{R}^n$ 에 대해 함수 $F : \mathbb{R} \to \mathbb{R}^n$ 를 $F : x \mapsto \int_{-\infty}^x f\, d\lambda_1$ 로 두면 이는 f 의 역도함수이다. 즉, $\nabla F = f$ 이다.

ii. 함수 $F : \mathbb{R} \to \mathbb{R}^n$ 가 $\mathscr{C}^1$ 급이고 그 도함수가 적분가능하면 $f := \nabla F$ 에 대해 $F(x) = \int_{-\infty}^{x} f \, d\lambda_1$ 이다.

Proof 이는 정의와 FTC로부터 자명하다. □

복소함수의 적분도 이와 비슷한 아이디어로 정의할 수 있다. 다만, 우선 $\mathbb{C}$ 위의 '바람직한' σ−대수를 어떻게 잡을지에 대한 질문의 답이 필요하다. 앞서 $\mathbb{R}^n$ 위의 '바람직한' σ−대수인 $\mathscr{B}_n$ 이나 $\mathscr{M}_n$ 을 구성하기 위해 얼마나 많은 지면을 할애하여 논의를 전개했는지를 생각해보면 정신이 아득해질 법하지만, 조금만 요령을 부리면 이 질문에 꽤 근사한 답을 할 수 있다. 바로 $\mathbb{C}$와 $\mathbb{R}^2$를 동일시하는 것이다. 즉, 복소수 $x + \mathbf{i}y$를 (x, y)로 생각하고, 이를 통해 우리는 $\mathbb{R}^2$에서의 표준위상, 거리함수, 심지어는 $\mathscr{B}_2$와 $\mathscr{M}_2$로 만들어지는 가측공간의 구조도 $\mathbb{C}$로 자연스럽게 옮겨올 수 있다. 곧, 우리는 가측공간 $(X, \mathscr{A})$에서 정의된 복소함수 $f : X \to \mathbb{C}$에 대해 이의 실수부와 허수부가 모두 가측이면 이때의 f를 가측함수라 한다. 비슷하게, $f : \mathbb{R} \to \mathbb{C}$의 실수부와 허수부가 모두 연속이거나 미분가능하면 이때의 f를 각각 연속 혹은 미분가능하다고 한다. (보통 표기의 편의를 위해 $f : \mathbb{R} \to \mathbb{C}$의 미분을 $f' = (\mathbf{Re} f)' + (\mathbf{Im} f)'$로 쓴다.)

Definition 2.245 측도공간 $(X, \mathscr{A}, \mu)$에서 정의된 가측함수 $f : X \to \mathbb{C}$에 대해 f의 실수부와 허수부가 적분가능하면 이때 f를 **(μ−Lebesgue) 적분가능 (− integrable)**하다고 한다. 나아가 적분가능한 $f : X \to \mathbb{C}$에 대해 μ에 대한 f의 **(Lebesgue) 적분 (− integral)**을 $\int_X f(x) \, d\mu(x)$ 혹은 간단히 $\int_X f \, d\mu$로 쓰고 $\int_X f \, d\mu = \int_X \mathbf{Re} f \, d\mu + \mathbf{i} \int_X \mathbf{Im} f \, d\mu$로 정의한다.

적분의 기본적인 성질이 비슷하게 성립하는 것은 벡터함수의 적분과 마찬가지이다.

Corollary 2.246 측도공간 $(X, \mathscr{A}, \mu)$에서 정의된 적분가능한 함수 $f, g : X \to \mathbb{R}^n$에 대해 다음이 성립한다.

i. 임의의 $c \in \mathbb{R}$에 대해 cf가 적분가능하고 $\int_X cf \, d\mu = c \int_X f \, d\mu$이다.
ii. 함수 $f + g$가 적분가능하고 $\int_X f + g \, d\mu = \int_X f \, d\mu + \int_X g \, d\mu$이다.
iii. 함수 $\overline{f}$가 적분가능하고 $\int_X \overline{f} \, d\mu = \overline{\int_X f \, d\mu}$이다.
iv. $|\int_X f \, d\mu| \le \int_X |f| \, d\mu$.

Proof 이는 정의와 적분의 기본적인 성질로부터 자명하다. □

Corollary 2.247 (Fundamental theorem of calculus) 다음이 성립한다.

i. 적분가능한 연속함수 $f : \mathbb{R} \to \mathbb{C}$에 대해 함수 $F : \mathbb{R} \to \mathbb{C}$를 $F : x \mapsto \int_{-\infty}^{x} f \, d\lambda_1$로 두면 이는 f의 역도함수이다. 즉, $F' = f$이다.
ii. 함수 $F : \mathbb{R} \to \mathbb{C}$가 $\mathscr{C}^1$ 급이고 그 도함수가 적분가능하면 $f := F'$에 대해 $F(x) = \int_{-\infty}^{x} f \, d\lambda_1$이다.

PROOF 이는 정의와 FTC로부터 자명하다. □

특히, 복소함수의 적분의 경우 Euler의 공식으로 알려진 $e^{\mathbf{i}\theta} = \cos\theta + \mathbf{i}\sin\theta$ 로부터 $e^{\mathbf{i}\theta}$ 의 θ 에 대한 미분이 $\mathbf{i}e^{\mathbf{i}\theta}$ 가 되어 그 계산이 간단해지는 경우가 많다. 예컨대 $\int_0^{\sqrt{2\pi}} \theta e^{\mathbf{i}\theta^2}\, d\lambda_1$ 를 계산하는 경우에, 이를 $\int_0^{\sqrt{2\pi}} \theta(\cos\theta^2 + \mathbf{i}\sin\theta^2)\, d\lambda_1$ 로 생각하여 계산해도 되지만, $(e^{\mathbf{i}\theta^2}/2\mathbf{i})' = \theta e^{\mathbf{i}\theta^2}$ 임에 착안하여 $\int_0^{\sqrt{2\pi}} \theta e^{\mathbf{i}\theta^2}\, d\lambda_1 = [e^{\mathbf{i}\theta^2}/2\mathbf{i}]_0^{\sqrt{2\pi}} = 0$ 로 손쉽게 계산할 수도 있다.

이로써 길고 험했던 측도론 이야기는 막을 내린다. 무려 250개 가까이 되는 생소한 정의, 정리와 씨름하며 여기까지 와 준 독자들에게 정말 감사할 따름이다. 통계학 이야기는 이제서야 본격적으로 시작되지만, 늦었다고 걱정할 필요는 없다. 측도론을 통해 보는 통계학은 분명 여태까지와는 다른 새로운 통계학일 것이다.

Notes

1 [측도론] Lebesgue 측도로서 잴 수 없는 집합을 구성하는 과정을 간단히 실어두도록 한다. 우선 $\mathbb{R}$ 위의 관계 $\sim$을 $x \sim y :\Leftrightarrow x - y \in \mathbb{Q}$로 정의하면 이가 동치관계임이 거의 자명하다. 이제 함수 $f : \mathbb{R}/\sim \to [0,1]$를 각 동치류에서 $[0,1]$에 속하는 representative를 하나 선택하는 선택함수라 하면 이는 well-define 된다. (여기서 선택공리가 필요하다.) 그렇다면 집합 $V = f(\mathbb{R}/\sim)$가 바로 우리가 찾는 Lebesgue 측도로서 잴 수 없는 집합으으로, 이를 흔히 **Vitali 집합 (− set)**이라 한다.

집합 V가 정말 비가측인지를 확인하기 위해 이가 가측이라 가정하고 임의의 $p \in \mathbb{Q}$에 대해 $V_p = V + p$라 하자. 그렇다면 Lebesgue 측도의 이동 불변성으로부터 임의의 $p \in \mathbb{Q}$에 대해 $\lambda_1(V) = \lambda_1(V_p)$이고 V의 구성으로부터 서로다른 $p, q \in \mathbb{Q}$에 대해 V_p, V_q는 서로소이다. (그렇지 않다면 어떤 $x \in \mathbb{R}$에 대해 $x - p, x - q \in V$인데, 이는 $x - p \sim x - q$에서 모순이다.) 이제 $(-1,1]$에 속하는 모든 유리수를 나열하여 $\{p_i\}$와 같이 수열의 형태로 나타내면 $(0,1] \subseteq \bigsqcup_{i=1}^{\infty} V_{p_i} \subseteq (-1,2]$이므로 (Hint: 임의의 $x \in (0,1]$에 대해 적당한 $y \in V$가 존재하여 $p := x - y \in \mathbb{Q}$이고 $p \in (-1,1]$이다.) 이상의 결과로부터 $1 \le \lambda_1(\bigsqcup_{i=1}^{\infty} V_{p_i}) = \sum_{i=1}^{\infty} \lambda_1(V_{p_i}) = \sum_{i=1}^{\infty} \lambda_1(V) \le 3$의 모순이 발생하고, 따라서 V는 비가측이다.

조금 더 충격적이지만 재미있는 사실은 양의 측도를 갖는 모든 $\mathbb{R}$의 가측집합은 Vitali 집합과 같은 비가측 집합을 부분집합으로 가진다는 사실이다. **Vitali의 정리**라 불리는 이 정리의 구성적인 증명에는 몇가지 다른 정리의 도움이 필요하다. 먼저 compact한 $K \subseteq \mathbb{R}$와 집합 $D = \{x - y \in \mathbb{R} : x, y \in K\}$에 대해 $\lambda_1(K) > 0$이면 적당한 $r > 0$이 존재하여 $B(r) \subseteq D$임을 보이자. 우선 Lebesgue 측도가 regular하므로 적당한 열린집합 $U \subseteq \mathbb{R}$가 존재하여 $K \subseteq U$이고 $\lambda_1(U) < 2\lambda_1(K)$이다. 이제 $r = d(K, U^c) > 0$라 하고 $|x| < r$이면 $K + x \subseteq U$이다. (그렇지 않다면 $x + y \notin U$인 $y \in K$가 존재하는데, $r \le |x + y - y| = |x| < r$에서 모순이 발생한다.) 여기서 만약 $(K + x) \cap K = \emptyset$이라면 $(K + x) \sqcup K \subseteq U$와 Lebesgue 측도의 이동 불변성으로부터 $\lambda_1(U) \ge \lambda_1(K + x) + \lambda_1(K) = 2\lambda_1(K)$가 되고, 곧 가정에 모순되므로 $(K + x) \cap K \ne \emptyset$이고, 따라서 적당한 $y \in K$가 존재하여 $y - x \in K$에서 $x = y - (y - x) \in D$이다. 그리고 이로부터 $B(r) \subseteq D$이다.

다음으로, 임의의 가측집합 $A \subseteq \mathbb{R}$와 집합 $D = \{x - y \in \mathbb{R} : x, y \in A\}$에 대해 $\lambda_1(A) > 0$이면 적당한 $r > 0$이 존재하여 $B(r) \subseteq D$임을 보이자. 집합열 $\{A \cap B(i)\}$를 생각하면 이는 A로 수렴하는 증가하는 집합열이므로 $\lambda_1(A \cap B(i)) \uparrow \lambda_1(A) > 0$이고, 따라서 충분히 큰 $i_0 \in \mathbb{N}$에 대해 $\lambda_1(A \cap B(i_0)) > 0$이다. 또한,

Lebesgue 측도가 regular하므로 적당한 닫힌집합 $K \subseteq A \cap B(i_0) \subseteq A$에 대해 $\lambda_1(K) > \lambda_1(A \cap B(i_0))/2$이다. 그렇다면 여기서의 K는 유계여서 compact하고, 곧 앞선 결과로부터 적당한 $r > 0$이 존재하여 $B(r) \subseteq \{x - y \in \mathbb{R} : x, y \in K\} \subseteq D$이다.
이제 임의의 가측집합 $A \subseteq \mathbb{R}$에 대해 $\lambda_1(A) > 0$이라 하고, 앞서 구성한 Vitali 집합을 $V \subseteq [0,1]$라 하자. 또한, 임의의 $p \in \mathbb{Q}$에 대해 $V_p = V + p$라 하면 서로다른 $p, q \in \mathbb{Q}$에 대해서는 V_p와 V_q가 서로소임을 이미 알고 있고, $\bigsqcup_{p \in \mathbb{Q}} V_p = \mathbb{R}$임을 쉽게 보일 수 있다. (Hint: 임의의 $x \in \mathbb{R}$에 대해 적당한 $y \in V$가 존재하여 $p := x - y \in \mathbb{Q}$이다.) 이제 모순을 유도하기 위해 모든 $p \in \mathbb{Q}$에 대해 $A \cap V_p$가 가측이라 하고 모든 유리수를 나열하여 $\{p_i\}$와 같이 수열의 형태로 나타내면 $A = \bigsqcup_{i=1}^{\infty} (A \cap V_{p_i})$에서 $0 < \lambda_1(A) = \sum_{i=1}^{\infty} \lambda_1(A \cap V_{p_i})$이다. 그런데 이는 어떤 $p \in \mathbb{Q}$에 대해 $\lambda_1(A \cap V_p) > 0$임을 함의하므로 앞선 결과로부터 적당한 $r > 0$이 존재하여 $B(r) \subseteq \{x - y \in \mathbb{R} : x, y \in A \cap V_p\}$의 모순이 발생한다. 따라서 어떤 $p \in \mathbb{Q}$에 대해 $A \cap V_p \subseteq A$는 비가측이다.

2 λ−system은 다음 세 가지 조건을 만족하는 집합족 $\mathscr{L}$로써 동등하게 정의할 수 있다.

i. $X \in \mathscr{L}$.
ii. (monotone difference) 임의의 $A, B \in \mathscr{L}$에 대해 $A \subseteq B$이면 $B \setminus A \in \mathscr{L}$이다.
iii. 집합족 $\mathscr{L}$에 속하는 임의의 증가하는 집합열 $\{A_i\}$에 대해 $\bigcup_{i=1}^{\infty} A_i \in \mathscr{L}$이다.

이와 같은 두 정의가 서로 동등함을 보이기 위해서, 먼저 정의 2.7의 조건을 만족하는 집합족 $\mathscr{L}$을 생각하자. 그렇다면 정의 2.7의 i과 ii로부터 $\emptyset \in \mathscr{L}$임을 알고, 이를 정의 2.7의 iii에 적용하면 $\mathscr{L}$이 유한번의 서로소 합집합에 대해서도 닫혀있음을 쉽게 알 수 있다. 이제 임의의 $A, B \in \mathscr{L}$에 대해 $A \subseteq B$라 하면 $B \setminus A = B \cap A^c = (B^c \sqcup A)^c \in \mathscr{L}$이 되어 위의 조건 ii가 만족된다. 비슷하게, 집합족 $\mathscr{L}$에 속하는 임의의 증가하는 집합열 $\{A_i\}$에 대해 집합열 $\{B_i\}$를 $B_i := A_i \setminus A_{i-1}$ ($B_1 := A_1$)로 정의하면 $\{B_i\}$가 $\mathscr{L}$에 속하는 서로소인 집합열이 되어 $\bigcup_{i=1}^{\infty} A_i = \bigsqcup_{i=1}^{\infty} B_i \in \mathscr{L}$에서 위의 조건 iii도 만족된다.
역으로, 위의 조건을 만족하는 집합족 $\mathscr{L}$을 생각하자. 그렇다면 임의의 $A \in \mathscr{L}$에 대해 $A^c = X \setminus A$에서 정의 2.7의 ii가 만족된다. 이제 임의의 서로소인 $A, B \in \mathscr{L}$에 대해 $B \subseteq A^c \in \mathscr{L}$이므로 $A \sqcup B = (A^c \cap B^c)^c = (A^c \setminus B)^c \in \mathscr{L}$이다. 여기에 수학적 귀납법을 적용하면 $\mathscr{L}$이 유한번의 서로소 합집합에 대해 닫혀있음을 알 수 있다. 이를 이용하여 마지막 조건인 정의 2.7의 iii을 보이기 위해 집합족 $\mathscr{L}$에 속하는 임의의 서로소인 집합열 $\{A_i\}$를 생각하고 집합열 $\{B_i\}$를 $B_i = \bigsqcup_{j=1}^{i} A_j$로 정의하자. 그렇다면 $\{B_i\}$가 $\mathscr{L}$에 속하는 증가하는 집합열이 되어 $\bigsqcup_{i=1}^{\infty} A_i = \bigcup_{i=1}^{\infty} B_i \in \mathscr{L}$에서 증명이 끝난다.
본문에서는 λ−system의 정의로 정의 2.7의 조건과 위의 세 조건을 특별한 언급 없이 모두 사용하였다.

3 **[측도론]** 임의의 $p, q \in \mathbb{Q}$에 대해 $[p, q) \cap \mathbb{Q}$를 rational semi-open interval이라 하자. 이제 $\mathscr{A}$를 $[0,1)$에 속하는 모든 rational semi-open interval의 모임이라 하면 이는 $\mathbb{R}$ 위의 semi-algebra이고 셈측도 $\#$는 $A(\mathscr{A})$ 위에서 premeasure가 된다. 또한, 그 구성으로부터 임의의 $A \in A(\mathscr{A})$에 대해 $\#(A) = 0$이거나 $\#(A) = \infty$임이 분명하다. (Hint: 대수 $A(\mathscr{A})$는 $\mathscr{A}$에 속하는 서로소인 집합들의 유한 합집합으로 만들 수 있는 모든 집합들의 모임이다.) 그렇다면 모든 $\#$−가측집합의 모임 $\mathscr{M}$에 대해 $\#$와 $2\#$ 모두 $A(\mathscr{A})$ 위의 premeasure $\#$을 $\mathscr{M}$ 위의 측도로 확장한 것이지만 이 둘은 서로 다르다. (한원소 집합이 $\sigma(\mathscr{A})$에 속하므로 $\mathscr{M}$에서 $\# = 2\#$일 수 없다.)

4 **[집합론]** 대표적인 σ−대수 $\mathscr{M}_n$과 $\mathscr{B}_n$의 기수는 각각 $2^{\mathfrak{c}}, \beth_1 (= 2^{\aleph_0})$인데, 생각하는 것과 달리 $\mathscr{M}_n$의 기수를 구하는 것이 $\mathscr{B}_n$의 기수를 구하는 것보다 더 쉽다. Cantor 집합 C를 생각하면 이는 가측이고 영집합이므로 C의 모든 부분집합도 가측인데, $|C| = \mathfrak{c}$이므로 $\mathscr{P}(C) \subseteq \mathscr{M}_n \subseteq \mathscr{P}(\mathbb{R})$에서 $|\mathscr{M}_n| = 2^{\mathfrak{c}}$임을 바로 알 수 있다. 한편, $\mathscr{B}_n$의 기수를 구하기 위해서는 준비가 조금 필요하다.
먼저 공집합이 아닌 집합 X의 부분집합의 모임 $\mathscr{C}$에 대해 이가 가산이면 $|\sigma(\mathscr{C})| < \aleph_0$이거나 $|\sigma(\mathscr{C})| \geq \beth_1$임을 보이자. 편의상 $\mathscr{C}$가 무한집합이라 생각하고 이의 원소를 나열하여 $\{A_i\}$와 같이 집합열의

형태로 쓰자. 또한, 임의의 $I \subseteq \mathbb{N}$에 대해 집합 $B_I = \bigcap_{i\in I} A_i \cap \bigcap_{i\notin I} A_i^c$를 생각하면 서로다른 $I, J \subseteq \mathbb{N}$에 대해 B_I와 B_J는 서로소이고 $\bigsqcup_{I\subseteq\mathbb{N}} B_I = X$이다. (Hint: 집합족 $\mathscr{C}$가 그려내는 Venn diagram을 생각해보면 거의 자명하다.) 이는 곧 $\mathscr{D} := \{B_I \subseteq X : I \subseteq \mathbb{N}, B_I \neq \emptyset\}$가 X의 분할임을 뜻한다. 이제 경우를 나누어 $\mathscr{D}$가 유한집합인 경우를 보자. 이 경우에는 $\mathscr{D} = \{B_{I_1}, \cdots, B_{I_k}\}$로 쓸 수 있고 $\sigma(\mathscr{C}) = \sigma(\mathscr{D}) = \{\bigsqcup_{j\in J} B_{I_j} : J \subseteq \{1, \cdots, k\}\}$이므로 임을 쉽게 보일 수 있으므로 $|\sigma(\mathscr{C})| < \aleph_0$이다. (Hint: 임의의 $A_i \in \mathscr{C}$에 대해 A_i와 교차하는 $\mathscr{D}$의 원소를 골라 이들을 서로소 합집합하면 A_i이다. 따라서 $\mathscr{C} \subseteq \sigma(\mathscr{D})$이고, 이제 나머지는 자명하다.) 반대로 $\mathscr{D}$가 무한집합인 경우에는 $\mathscr{D} \subseteq \sigma(\mathscr{C})$에서 $\{\bigsqcup_{j\in J} B_{I_j}\}_{J\subseteq\mathbb{N}} \subseteq \sigma(\mathscr{C})$이고 서로다른 $J, J' \subseteq \mathbb{N}$에 대해 $\bigsqcup_{j\in J} B_{I_j}$와 $\bigsqcup_{j\in J'} B_{I_j}$가 서로소이므로 $|\sigma(\mathscr{C})| \geq \beth_1$이다.

다음으로, $\mathscr{B}_n$의 가산 생성자 $\mathscr{C}$를 하나 적당히 구한다. (예컨대 $\mathbb{R}^n$의 모든 열린집합의 모임이 $\mathscr{B}_n$을 생성하므로 꼭짓점이 $\mathbb{Q}^n$인 열린 box의 모임을 생각하면 이는 $\mathscr{B}_n$의 가산 생성자이다.) 그렇다면 앞선 결과로부터 $|\mathscr{B}_n| \leq \beth_1$만 보이면 되는데, 이를 위해서는 초한귀납법을 써야 한다. 우선 $\mathscr{C}_0 = \mathscr{C}$로 시작하여 임의의 순서수 α에 대해 $\mathscr{C}_{\alpha+1} = \{\bigcup_{i\in I} A_i \cup \bigcup_{j\in J} B_j^c \subseteq \mathbb{R}^n : I, J \subseteq \mathbb{N}, A_i, B_j \in \mathscr{C}_\alpha\}$로 두고 극한 순서수 α에 대해서는 $\mathscr{C}_\alpha = \bigcup_{\beta<\alpha} \mathscr{C}_\beta$로 둔다. 그리고는 $\mathscr{A} = \bigcup_{\alpha<\omega_1} \mathscr{C}_\alpha$라 하여 $\mathscr{A}$가 σ−대수임을 주장한다. 이를 위해서는 이런저런 조건을 확인해봐야 하는데, 가산 합집합에 대해 닫혀있다는 조건만 확인하면 나머지는 자명하다. 이에 $\mathscr{A}$에 속하는 집합열 $\{A_i\}$를 택하면 $\{\mathscr{C}_\alpha\}$가 증가하는 초한 집합열이므로 $A_i \in \mathscr{C}_{\alpha_i}$인 증가하는 수열 $\{\alpha_i\}$를 구성할 수 있다. 그런데 $\mathscr{A}$의 구성으로부터 각 $i \in \mathbb{N}$에 대해 $\alpha_i < \omega_1$이므로 $\sup_{i\in\mathbb{N}} \alpha_i < \omega_1$에서 $\bigcup_{i=1}^\infty A_i \in \bigcup_{i=1}^\infty \mathscr{C}_{\alpha_i} \subseteq \bigcup_{\alpha<\omega_1} \mathscr{C}_\alpha = \mathscr{A}$가 되어 $\mathscr{A}$가 σ−대수임을 알고, 곧 $\mathscr{B}_n \subseteq \mathscr{A}$에서 $|\mathscr{B}_n| \leq |\mathscr{A}| \leq \sum_{\alpha<\omega_1} |\mathscr{C}_\alpha| \leq \aleph_0 \beth_1 = \beth_1$이다. (이는 임의의 $\alpha < \omega_1$에 대해 $|\mathscr{C}_\alpha| \leq \beth_1$이기 때문에 성립하고, 다시 이는 초한귀납법으로 쉽게 보일 수 있다. DIY!)

5 집합 $\mathbb{R}^n$에서 모든 꼭짓점이 유리점인 열린 box를 rational box라 하면, $\mathbb{Q}$가 가산이므로 $\mathbb{Q}^n$도 가산이고, 곧 모든 rational box도 가산개이다. 이제 임의의 열린집합 $U \subseteq \mathbb{R}^n$에 대해 U에 포함되는 rational box만을 모두 택하여 집합족 $\mathscr{U}$를 구성하자. 그렇다면 $U = \bigcup \mathscr{U}$임을 쉽게 보일 수 있고, $\mathscr{U}$가 명백히 가산이므로 $\mathbb{R}^n$의 모든 열린집합은 유계인 열린 box의 가산 합집합으로 표현됨을 안다.

6 [측도론] 각주 4에서 살펴본 바와 같이 $|\mathscr{B}_n| < |\mathscr{M}_n|$이므로 $\mathscr{B}_n$은 $\mathscr{M}_n$의 진부분집합이다. 이는 곧 가측이지만 Borel은 아닌 집합이 존재한다는 뜻인데, 이를 직접 구성해볼 수 있다. 이번에는 Cantor 함수 $c : [0,1] \to \mathbb{R}$가 중요한 역할을 한다. 우선 함수 $f : [0,1] \to [0,2]$를 $f : x \mapsto x + c(x)$로 정의하면 이는 증가하는 연속함수이므로 Borel이고, 전단사임을 쉽게 보일 수 있다. (Hint: 함수 f가 전사임은 IVT로부터 자명하다. 한편, f가 단사임을 보이기 위해서 $x < y$인 임의의 $x, y \in [0,1]$에 대해 $f(x) = f(y)$라 하면 $c(x) > c(y)$에서 모순이 발생한다.) 또한, Cantor 집합 C에 대해 $\lambda_1(f(C)) = 1$임도 쉽게 보일 수 있으므로 (Hint: 닫힌구간 $[0,1]$에 속하는 서로소인 열린구간 $I_1, I_2, \cdots$에 대해 $[0,1] \setminus C = \bigsqcup_{i=1}^\infty I_i$로 쓸 수 있고, 각 I_i에서 c는 상수함수이므로 $\lambda_1(f(I_i)) = \lambda_1(I_i)$이다.) Vitali의 정리로부터 적당한 비가측 집합 $A \subseteq f(C)$가 존재한다. 그렇다면 $f^{-1}(A) \subseteq C$는 영집합이어서 가측이지만 Borel이 아니다.

7 구체적으로 다음과 같이 할 수 있다. 먼저 $\delta = \lambda_n(A) - \sum_{i=1}^\infty \rho_n(A_i) + \varepsilon > 0$로 두고 각 $i \in \mathbb{N}$에 대해 semi-open box A_i의 중심을 고정한 채로 이의 모서리의 길이를 조금씩 늘리되 그 결과로 만들어진 새 semi-open box A_i'이 $\rho_n(A_i') < \rho_n(A_i) + \delta/2^{i+1}$를 만족하도록 한다. 그렇다면 $A_i \subseteq (A_i')^\circ \subseteq A_i'$에서 $A \subseteq \bigcup_{i=1}^\infty A_i \subseteq \bigcup_{i=1}^\infty (A_i')^\circ$이고 $\sum_{i=1}^\infty \rho_n(A_i') \leq \sum_{i=1}^\infty [\rho_n(A_i) + \delta/2^{i+1}] = \sum_{i=1}^\infty \rho_n(A_i) + \delta/2 < \lambda_n(A) + \varepsilon$이 성립한다.

8 여기서 $\lim_{x_i\to-\infty} F_\mu(x) = 0$은 임의의 $\cdots, x_{i-1}, x_{i+1}, \cdots \in \mathbb{R}$와 임의의 $\varepsilon > 0$에 대해 적당한 $M \in \mathbb{R}$이 존재하여 $x_i < M$인 임의의 $x_i \in \mathbb{R}$에 대해 $F_\mu(x) < \varepsilon$이라는 의미이다.

9 여기서 $\lim_{x_1,\cdots,x_n\to\infty} F_\mu(x) = \mu(\mathbb{R}^n)$는 임의의 $\varepsilon > 0$에 대해 적당한 $M \in \mathbb{R}$이 존재하여 $x \geq M\mathbf{1}$인 임의의 $x \in \mathbb{R}^n$에 대해 $|F_\mu(x) - \mu(\mathbb{R}^n)| < \varepsilon$이라는 의미이다.

10 [위상수학] $\mathbb{R}$에 주어진 표준위상 $\mathscr{T}$는 $\beta = \{(x,\infty)\}_{x\in\mathbb{R}} \cup \{(-\infty,y)\}_{y\in\mathbb{R}}$를 부분기저로 하는 순서위상이다. 따라서 이를 그대로 확장하여 $\overline{\mathbb{R}}$에도 $\widetilde{\beta} = \{(x,\infty]\}_{x\in\mathbb{R}} \cup \{[-\infty,y)\}_{y\in\mathbb{R}}$를 부분기저로 하는 순서위상 $\mathscr{T}_{\text{ext}}$를 줄 수 있고, 이는 곧 $\gamma = \{(x,\infty]\}_{x\in\mathbb{R}} \cup \{[-\infty,y)\}_{y\in\mathbb{R}} \cup \{(x,y)\}_{x,y\in\mathbb{R}}$이 $\mathscr{T}_{\text{ext}}$의 기저라는 것과 동치이므로 $\mathscr{T} \subseteq \mathscr{T}_{\text{ext}}$임이 자명하다. 이제 임의의 $U \in \mathscr{T}_{\text{ext}}$는 적당한 $x,y\in\mathbb{R}$와 $V \subseteq (y,x)$인 $V \in \mathscr{T}$에 대해 $[-\infty,y)\cup V \cup (x,\infty]$로 쓸 수 있고, V는 유계인 열린구간의 가산 합집합으로 쓸 수 있으므로 곧 U는 적당한 $x,y\in\mathbb{R}$에 대해 (x,y), $(x,\infty]$, $[-\infty,y)$로 표현되는 구간들의 가산 합집합으로 표현되고 그 역도 성립한다.

한편, $\overline{\mathbb{R}}$에서의 사칙연산에서는 $\infty-\infty$, ∞/∞, $0\cdot\infty$와 같은 부정형이 나올 수 있으므로 각별히 주의해야 한다. 다만, 이 중의 일부의 값을 관례를 따라 $0\cdot\infty := 0$, $0^0 := 1$로 둔다.

11 [위상수학] 일반적으로 위상공간 $(X,\mathscr{T})$에 대해 $\mathscr{T}$가 생성하는 $\sigma-$대수를 X 위의 Borel $\sigma-$대수로 정의한다. 이 책에서는 위상수학의 내용을 피하기 위해 $\mathbb{R}^n$에 한정하여 Borel $\sigma-$대수를 도입하였다.

12 [측도론] 함수 $f:[0,1]\to[0,2]$를 각주 6에서와 같이 두면 이가 전단사인 연속함수로 Cantor 집합 C에 대해 $\lambda_1(f(C)) = 1$임을 이미 알고 있다. 따라서 Vitali의 정리로부터 비가측 집합 $A \subseteq f(C)$를 적어도 하나 택할 수 있는 한편, 집합 $B = f^{-1}(A) \subseteq C$는 C가 영집합이라는 점에서 가측이다. 이제 함수 $g:[0,2]\to\mathbb{R}$를 $g = \mathbf{1}_B \circ f^{-1}$로 두면 f^{-1}가 연속이고 B가 가측이므로 f^{-1}와 $\mathbf{1}_B$ 각각은 가측이지만 $g^{-1}(1) = f(B) = A$가 가측이 아니므로 g는 가측이 아니다.

13 만약 $x\in X$에 대해 $f(x)+g(x)$가 $\infty-\infty$의 부정형이 되면, 이때의 값을 0과 같은 dummy value로 정하는 관례를 전제하였다. 가정으로부터 f와 g가 적분가능하여 이들은 거의 어디서나 유한하고, 따라서 이러한 점의 집합이 영집합을 이루므로 적분의 결과는 이때의 dummy value의 선택과는 무관하다.

14 수렴정리의 조건이 만족되지 않아 극한과 적분이 교환되지 않는 예시를 보자. 단순함수 $f_i:\mathbb{R}\to\mathbb{R}$의 열 $\{f_i\}$를 $f_i := \mathbf{1}_{(i^2-i,i^2)}/2i - \mathbf{1}_{(i^2,i^2+i)}/i$로 두면 이는 $f := 0$로 수렴함이 자명하다. (심지어 균등수렴한다.) 하지만 $\{f_i\}$가 양과 음의 값을 모두 가져 MCT를 쓸 수 없다. 실제로 임의의 $i\in\mathbb{N}$에 대해 $\int_{\mathbb{R}} f_i\,d\lambda_1 = -1/2$이므로 $\lim_{i\to\infty}\int_{\mathbb{R}} f_i\,d\lambda_1 = -1/2 \neq 0 = \int_{\mathbb{R}} f\,d\lambda_1$이다. 같은 이유에서 Fatou의 보조정리도 쓸 수 없다.

그렇다면 $\{f_i\}$ 대신 $\{|f_i|\}$를 생각해보면 어떨까? 이 경우에 $\{|f_i|\}$는 명백히 음이 아닌 함수열이지만 이가 증가하지 않아 MCT는 여전히 쓸 수 없다. 실제로 임의의 $i\in\mathbb{N}$에 대해 $\int_{\mathbb{R}}|f_i|\,d\lambda_1 = 3/2$이므로 $\lim_{i\to\infty}\int_{\mathbb{R}} f_i\,d\lambda_1 = 3/2 \neq 0 = \int_{\mathbb{R}} f\,d\lambda_1$이다.

한편, $\{f_i\}$에 DCT를 쓸 수는 없을까? DCT를 쓰려면 $\{f_i\}$를 지배하는, 즉 각 f_i에 대해 $|f_i|\le g$인 적분가능한 $g:\mathbb{R}\to\overline{\mathbb{R}}_0^+$가 필요하다. 이러한 g가 존재한다고 하면 서로다른 $i,j\in\mathbb{N}$에 대해 (i^2-i,i^2+i)와 (j^2-j,j^2+j)가 서로소이므로 (Hint: 만약 $i<j$라 하면 $j^2-j-i^2-i = (j-i)(j+i)-(j+i) > (j+i)-(j+i) = 0$이다.) 모든 $x\in\mathbb{R}$에 대해 $\sum_{i=1}^{\infty}|f_i(x)|$는 well-define되고 $\sum_{i=1}^{\infty}|f_i|\le g$이다. 그러나 이는 MCT에서 $\int_{\mathbb{R}}\sum_{i=1}^{\infty}|f_i|\,d\lambda_1 = \sum_{i=1}^{\infty}\int_{\mathbb{R}}|f_i|\,d\lambda_1 = \infty \le \int_{\mathbb{R}} g\,d\lambda_1$의 모순을 일으키므로 어떠한 적분가능한 함수도 $\{f_i\}$를 지배할 수 없고, 곧 DCT도 쓸 수 없다.

15 만약 $x\in X$에 대해 $\sum_{i=1}^{\infty} f_i(x)$가 진동하면, 이때의 값을 0과 같은 dummy value로 정하는 관례를 전제하였다. 가정으로부터 이러한 점의 집합이 영집합을 이루므로 적분의 결과는 이때의 dummy value의 선택과는 무관하다.

16 만약 $x\in X$에 대해 함수 $y\mapsto f(x,y)$가 미분가능하지 않다면, 이때의 $(\partial f/\partial y)(x,y)$의 값을 0과 같은 dummy value로 정하는 관례를 전제하였다. 가정으로부터 이러한 점의 집합이 영집합을 이루므로 따름정리의 결론은 이때의 dummy value의 선택과는 무관하다.

17 측도공간 $(X,\mathscr{A},\mu)$와 $(Y,\mathscr{B},\nu)$에 대해 $\mathscr{A}\times\mathscr{B}$는 semi-algebra이다. 이제 함수 $\rho:\mathscr{A}\times\mathscr{B}\to\overline{\mathbb{R}}_0^+$를 $\rho: A\times B\mapsto \mu(A)\nu(B)$로 두고 이가 $\mathscr{A}\times\mathscr{B}$ 위의 premeasure임을 보이면 정리 2. 30와 Carathéodory의 확장정리로부터 ρ를 $\mathscr{A}\otimes\mathscr{B}$ 위의 측도로 확장시킬 수 있으므로 존재성이 보여진다. 이를 위해

$\mathscr{A} \times \mathscr{B}$에 속하는 임의의 서로소인 집합열 $\{A_i \times B_i\}$를 생각하여 $\bigsqcup_{i=1}^{\infty}(A_i \times B_i) \in \mathscr{A} \times \mathscr{B}$라 하자. 그렇다면 적당한 $A \in \mathscr{A}, B \in \mathscr{B}$에 대해 $A \times B = \bigsqcup_{i=1}^{\infty}(A_i \times B_i)$이고, $\mathbf{1}_A(x)\mathbf{1}_B(y) = \sum_{i=1}^{\infty} \mathbf{1}_{A_i}(x)\mathbf{1}_{B_i}(y)$에서 MCT로부터 $\mathbf{1}_A(x)\nu(B) = \int_Y \mathbf{1}_A(x)\mathbf{1}_B(y)\,d\nu(y) = \sum_{i=1}^{\infty} \int_Y \mathbf{1}_{A_i}(x)\mathbf{1}_{B_i}(y)\,d\nu(y) = \sum_{i=1}^{\infty} \mathbf{1}_{A_i}(x)\nu(B_i)$이고, 다시 $\rho(A \times B) = \mu(A)\nu(B) = \int_X \mathbf{1}_A(x)\nu(B)\,d\mu(x) = \sum_{i=1}^{\infty} \int_X \mathbf{1}_{A_i}(x)\nu(B_i)\,d\mu(x) = \sum_{i=1}^{\infty} \mu(A_i)\nu(B_i) = \sum_{i=1}^{\infty} \rho(A_i \times B_i)$이다. 한편, $\rho(\emptyset) = 0$임은 자명하므로 ρ는 premeasure이다.

18 임의의 열린집합 $U \subseteq \mathbb{R}^n$를 택하여 $\mathbb{R}^n$의 rational box 중에서 U에 포함되는 것만을 뽑아 집합족 $\mathscr{U}$를 구성하면 이는 명백히 가산이고, 곧 $\{U_i\}$와 같이 집합열의 형태로 쓸 수 있다. (Rational box의 의미에 대해서는 각주 5를 참조하기 바란다.) 그렇다면, 각 $i \in \mathbb{N}$에 대해 적당한 $x_i, y_i \in \mathbb{Q}^n$가 존재하여 $U_i = \prod_{j=1}^{n}(x_i^j, y_i^j)$이고, 여기서 $V_i = \prod_{j=1}^{l}(x_i^j, y_i^j), W_i = \prod_{j=l+1}^{m}(x_i^j, y_i^j)$라 하면 $\{V_i\}, \{W_i\}$는 각각 $\mathbb{R}^l, \mathbb{R}^m$에 속하는 열린 집합열로서 $U = \bigcup_{i=1}^{\infty} U_i = \bigcup_{i=1}^{\infty}(V_i \times W_i)$이다.

19 이는 각주 5와 거의 비슷하게 보일 수 있다. 집합 $\mathbb{R}^n$에서 모든 꼭짓점이 유리점인 semi-open box를 rational semi-box라 하면, $\mathbb{Q}$가 가산이므로 $\mathbb{Q}^n$도 가산이고, 곧 모든 rational semi-box의 모임 $\mathscr{S}$도 가산이다. 여기서 임의의 $B, B' \in \mathscr{S}$에 대해 $B \setminus B'$가 적당한 서로소인 $B_1, \cdots, B_k \in \mathscr{S}$에 대해 $B \setminus B' = \bigsqcup_{i=1}^{k} B_i$로 표현된다는 점을 주목하자. 이제 임의의 열린집합 $U \subseteq \mathbb{R}^n$에 대해 U에 포함되는 rational semi-box만을 모두 택하여 집합족을 구성하면 이가 명백히 가산이므로 WLOG, 필요하다면 공집합을 추가하여 이를 집합열 $\{B_i\}$로 쓸 수 있다. 한편, 각 $i \in \mathbb{N}$에 대해 $B_i' = B_i \setminus \bigcup_{j=1}^{i-1} B_j$를 생각하면 이는 적당한 서로소인 rational semi-box $B_{i1}, \cdots, B_{ik_i}$에 대해 $B_i' = \bigsqcup_{j=1}^{k_i} B_{ij}$이므로 rational semi-box의 열 $\{B_{ij}\}_{ij}$는 $\mathscr{S}$에 속하는 서로소인 집합열이고 $U = \bigcup_{i=1}^{\infty} B_i = \bigsqcup_{i=1}^{\infty} B_i' = \bigsqcup_{i=1}^{\infty} \bigsqcup_{j=1}^{k_i} B_{ij}$ 임을 쉽게 보일 수 있다.

20 부등식 $\varphi_{\mathscr{P}_i} \leq f\mathbf{1}_A \leq \psi_{\mathscr{P}_i}$가 성립하는 않는 점은 모두 $\bigcup_{P \in \mathscr{P}_i}(P \setminus \widetilde{P})$에 속하는데, 여기서 각 $P \setminus \widetilde{P}$는 위의 보조정리에서와 같은 초평면 n개의 합집합에 속하는 집합이므로 0의 측도를 가지게 되어 이 부등식은 거의 어디서나 성립한다.

21 이 함수는 집합 $\mathbb{Q} \cap [0, 1]$이 영집합이라는 점에서 거의 어디서나 0이고, 따라서 그 Lebesgue 적분값도 0이다. 여기서 함수 $\mathbf{1}_{\mathbb{Q}}$는 **Dirichlet 함수 (– function)**라는 이름을 가지고 있다. 이는 앞서 보았듯이 임의의 비어있지 않은 구간 $I \subseteq \mathbb{R}$에서 Riemann 적분은 불가능하지만 Lebesgue 적분은 가능하며 그 적분값은 항상 0이다. 때로 Dirichlet 함수는 $\lim_{m\to\infty} \lim_{n\to\infty} \cos^{2n}(m!\pi x)$와 같이 연속함수의 극한으로 표현되기도 한다.

22 이는 각주 5와 거의 비슷하게 보일 수 있다. 각 $i \in \mathbb{N}$에 대해 $2^{-i}\mathbb{Z}^n$에 속하는 점들을 꼭짓점으로 하여 만들 수 있는 모서리의 길이가 2^{-i}인 모든 semi-open box의 모임을 $\mathscr{C}_i$라 하면 이는 가산이고 모든 원소가 서로소이다. 잠시 이에 속하는 semi-open box를 ith dyadic semi-box라 하자. 이제 1st dyadic semi-box 중에서 U에 속하는 것만을 모아 집합족 $\mathscr{D}_1$을 구성한다. 비슷하게, 2nd dyadic semi-box 중에서 $U \setminus \bigcup \mathscr{D}_1$에 속하는 것만을 모아 집합족 $\mathscr{D}_2$를 구성하고 이를 반복하여 집합족의 열 $\{\mathscr{D}_i\}$를 구성한다. 그렇다면 각 $i \in \mathbb{N}$에 대해 $\mathscr{D}_i \subseteq \mathscr{C}_i$에서 $\mathscr{D}_i$는 가산이므로 곧 $\mathscr{D} = \bigcup_{i=1}^{\infty} \mathscr{D}_i$도 가산임이 분명하고, 구성으로부터 $\mathscr{D}$의 모든 원소들은 서로소이다. 따라서 $\mathscr{D}$의 원소들을 나열하여 WLOG, 필요하다면 공집합을 추가하여 이를 집합열의 형태로 $\{B_i\}$와 같이 쓸 수 있고, 곧 이가 우리가 찾던 집합열임을 쉽게 보일 수 있다.

23 여기서 Φ가 $\mathscr{C}^1$급 함수이므로 그 편미분이 모두 연속함수로서 가측이고, 곧 이들의 곱과 합으로 표현되는 행렬식 $\det \mathbf{D}\Phi$도 가측이다. 따라서 $f \circ \Phi$가 가측이면 $(f \circ \Phi)|\det \mathbf{D}\Phi|$가 가측임은 자명하다. 한편, 가측공간 $(X, \mathscr{A}), (Z, \mathscr{B})$와 $Y \in \mathscr{A}$에 대해 함수 $f : Y \to Z$가 가측이라 함은 f가 $\mathscr{A}|_Y/\mathscr{B}$–가측임을 의미한다.

24 보조정리 2. 127로부터 임의의 닫힌 box에서 그 경계를 이루는 면 중 일부를 빼어 얻는 집합은 원래의 box와 같은 측도를 가짐을 알 수 있다. 만약 $c > 0$이라면 T^{-1}는 단순히 B를 scaling시킨 것에 불과하여

$T^{-1}(B)$가 다시 semi-open box가 되므로 위의 식이 자명하다. 그러나 $c<0$인 경우에는 $T^{-1}(B)$가 semi-open box를 그 중심을 축으로 하여 점대칭시킨 모양이 되는데, 방금 지적한 바로부터 $T^{-1}(B)$에서 원래 속한 면은 빼고, 원래 빠진 면은 더하여 이를 다시 semi-open box로 만들도 그 측도는 변함이 없다. 따라서 이 경우에도 위의 식이 성립한다.

25 함수 Ψ_x와 Φ가 모두 $\mathscr{C}^1$급 미분동형사상이므로 이들은 열린집합을 열린집합으로, 닫힌집합을 닫힌 집합으로 보내고, 곧 집합 $\Psi_x(B(x,(1-\varepsilon)r)), \Phi(B(x,r))$는 열린집합이고 집합 $\Phi(\overline{B(x,r)})$는 닫힌집합이다. 또한, 같은 이유로 Ψ_x와 Φ는 연결집합을 연결집합으로 보내며, 이러한 사실들은 이 증명의 전반에 걸쳐 별다른 언급 없이 사용되었다.

26 각주 19를 참조하면 $\mathscr{S}_n$에 속하는 적당한 서로소인 집합열 $\{B_i\}$가 존재하여 $W=\bigsqcup_{i=1}^{\infty}B_i$이다. 한편, 보조정리 2.127로부터 $\bigcup_{i=1}^{\infty}\partial B_i$가 영집합임을 알 수 있다. 이제 각 $i\in\mathbb{N}$에 대해 $B(x_i,r_i)=B_i^\circ$라 하고 WLOG, 필요하다면 B_i를 가산개의 semi-open box로 분할하여 $r_i<\delta$라 하면 $\{B(x_i,r_i)\}$가 서로소인 집합열임은 자명하고 $\lambda_n(W\setminus\bigsqcup_{i=1}^{\infty}B(x_i,r_i))\le\lambda_n(\bigcup_{i=1}^{\infty}\partial B_i)=0$이다.

27 각주 19를 참조하여 각주 26에서와 똑같은 방식으로 하되, 각주 19에서 집합열을 구성할 때 $\overline{B_i}\subseteq U$를 추가적인 조건으로 하여 $\{B_i\}$를 구성하면 된다.

28 집합 $\Phi(F)$가 유계이므로 적당한 유계인 열린집합 $V'\subseteq V$가 존재하여 $\Phi(F)\subseteq V'$이다. 이제 $W=\Phi^{-1}(V')$라 하면 Φ가 $\mathscr{C}^1$급 미분동형사상이라는 점에서 이가 열려있으며 본문의 조건을 모두 만족한다.

29 Gram-Schmidt 정규직교화를 통해 $\beta=\{\beta_1,\cdots,\beta_n\}$가 $\mathbb{R}^n$의 정규직교기저가 되는 동시에 $\gamma=\beta\setminus\{\beta_n\}$가 V의 정규직교기저가 되도록 $\beta_1,\cdots,\beta_n\in\mathbb{R}^n$을 택할 수 있다. 이제 $\mathbb{R}^n$의 표준기저 $\mathscr{E}=\{\mathbf{e}_1,\cdots,\mathbf{e}_n\}$에 대해 선형사상 $T:\mathbb{R}^n\to\mathbb{R}^n$를 $[T]_\beta^{\mathscr{E}}=I$이도록 잡으면 $T(\mathsf{V})=T(\operatorname{span}\gamma)=\operatorname{span}T(\gamma)=\operatorname{span}(\mathscr{E}\setminus\{\mathbf{e}_n\})$에서 $T(\mathsf{V})$는 $\mathbf{e}_n$과 직교하는 초평면이 되고, $|\det T|=1$이다.

30 임의의 $x_0\in\mathbb{R}$에 대해 함수 $f:\mathbb{R}\to\mathbb{R}$를 $f:x\mapsto e^{x-x_0}\mathbf{1}_{(-\infty,x_0]}$로 두면 $F:x\mapsto e^{x-x_0}\mathbf{1}_{(-\infty,x_0]}+\mathbf{1}_{(x_0,\infty)}$이므로 F는 x_0에서 연속이지만 미분불가능하다.

31 변수변환 공식으로부터 임의의 $\mathscr{C}^1$급 미분동형사상 $\Phi:U\to V$와 (여기서 $U,V\subseteq\mathbb{R}^n$는 열린집합이다.) 임의의 가측인 $A\in\mathbb{R}^n$를 택하면 $\Phi^{-1}(A)$와 $\Phi(A)$가 모두 가측이므로 $\mathscr{M}_n=\Phi_*\mathscr{M}_n$이다. 한편, 다시 변수변환 공식으로부터 $(\Phi_*\lambda_n)(A)=\lambda_n(\Phi^{-1}(A))=\int_U\mathbf{1}_{\Phi^{-1}(A)}\,d\lambda_n=\int_U\mathbf{1}_A\circ\Phi\,d\lambda_n=\int_V\mathbf{1}_A/|\det\mathbf{D}\Phi\circ\Phi^{-1}|\,d\lambda_n=\int_A d(\lambda_n)_{|\det\mathbf{D}\Phi\circ\Phi^{-1}|^{-1}}=(\lambda_n)_{|\det\mathbf{D}\Phi\circ\Phi^{-1}|^{-1}}(A)$이므로 $f\Phi_*\lambda_n=(\lambda_n)_{|\det\mathbf{D}\Phi\circ\Phi^{-1}|^{-1}}$이다. 이로부터 임의의 음이 아닌 가측함수 $f:\mathbb{R}^n\to\overline{\mathbb{R}}_0^+$에 대해 $\widetilde{f}=f|\det\mathbf{D}\Phi\circ\Phi^{-1}|$라 하면 $\int_{\Phi^{-1}(V)}\widetilde{f}\circ\Phi\,d\lambda_n=\int_U(f\circ\Phi)|\det\mathbf{D}\Phi|\,d\lambda_n$이고 $\int_V\widetilde{f}\,d\Phi_*\lambda_n=\int_V\widetilde{f}\,d(\lambda_n)_{|\det\mathbf{D}\Phi\circ\Phi^{-1}|^{-1}}=\int_V f\,d\lambda_n$이므로 위의 변수변환 정리는 앞서 본 변수변환 정리의 일반화라 할 수 있다.

32 여기서 $\{\varphi(x)\in A\}$는 조건제시법으로 주어진 집합 $\{x\in X:f(x)\in A\}$의 약식 표현이다. 측도론에서는 혼동의 우려가 없다면 어떤 조건 $P(x)$에 대해 집합 $\{x\in X:P(x)\}$를 $\{P(x)\}$로 간단히 표기하는 관례가 있다.

33 [측도론] 함수 F를 Cantor 함수 $c:[0,1]\to[0,1]$로 두면 이는 연속이고 거의 어디서나 미분가능하므로 곧 거의 어디서나 symmetric differentiable하며, $F^{(s)}=0=:f$ (ae.) 이지만 명백히 $F(x)\neq0=\int_{-\infty}^{x}f\,d\lambda_1$이다.

34 사실, 이렇게 다소 아쉬움이 남는 결론에 이른 것은 이후 확률론에서의 사용을 목적으로 일반적인 $n\in\mathbb{N}$에 대해 논의를 전개한 탓이 크다. 만약 우리가 일반성을 포기하고 $n=1$인 경우에 대해서만 생각했다면 훨씬 더 편하게 더 깔끔한 결론을 얻을 수 있었을 것이다. 예컨대 i의 경우 미분계수의 정의인 극한식을 직접 다룸으로써 보다 쉽게 F가 거의 대부분의 점에서 미분가능하다는 (symmetric differentiability가 아닌 통상의 미분가능성이다.) 더 강한 결론을 얻을 수 있고, ii의 경우 유계변동이라는 개념을 통해 F가 증가함수여야 한다는 조건을 획기적으로 완화시킬 수 있다. 이에 관심이 있는 독자들은 [3]을 참조하기 바란다. 하지만 이러한 방식들이 일반적인 n에 대하여 자명하게 일반화되지 않는 까닭에 i

은 Hardy-Littlewood의 정리를 이용하여 우회적으로 접근할 수 밖에 없었고, ii의 경우 끝까지 F에게 증가함수여야 한다는 조건을 요구해야만 했다. 게다가 절대연속의 개념은 앞서 설명한 바와 같이 정리 2. 201가 성립하는 것이 핵심인데, 이를 유지하면서 일반적인 n에 대해 절대연속의 개념을 '예쁘게' 정의하는 것은 꽤나 골치아픈 문제이다. 우리는 정리 2. 201가 성립하도록 다소 무식하게 절대연속의 개념을 정의했는데, 그 결과 절대연속이 연속성을 함의하지 못하는 '못생긴' 절대연속이 되고 말았다. 이에 대한 보다 자세한 논의는 [7] 과 [8] 을 참조하기 바란다.

35 합 $\sum_{x\in X} f(x)$가 well-define된다고 함은 합의 과정에서 $\infty-\infty$와 같은 부정형이 등장하지 않고, X가 무한집합인 경우에는 이가 절대수렴하여 합의 순서에 무관하게 그 값이 주어짐을 뜻한다.

References

[1] Jerrold E. Marsden and Michael J. Hoffman, *Elementary Classical Analysis 2nd Edition*, W. H. Freeman, 1993.
[2] Walter Rudin, *Real and Complex Analysis 3rd Edition*, McGraw-Hill, 1987.
[3] Patrick Billingsley, *Probability and Measure 3rd Edition*, Wiley, 1995.
[4] Donald L. Cohn, *Measure Theory 2nd Edition*, Springer, 2013.
[5] Piermarco Cannarsa and Teresa D'Aprile, *Introduction to Measure Theory and Functional Analysis*, Springer, 2015.
[6] 이인석,『선형대수와 군』개정판, 서울대학교출판문화원, 2015.
[7] Jan Maly, "Absolutely Continuous Functions of Several Variables", *Journal of Mathematical Analysis and Applications*, vol. 231, no. 2, 1999.
[8] Jiri Sremr, "A Note on Absolutely Continuous Functions of Two Variables in The Sense of Carathéodory", *Electronic Journal of Differential Equations*, no. 154, 2010.
[9] Alexander Grigoryan, *Measure Theory and Probability*, University of Bielefeld, 2008.
[10] Dietmar A. Salamon, *Measure and Integration*, ETH Zurich, 2016.
[11] Juha Kinnunen, *Real Analysis*, Aalto University, 2016.
[12] Unknown, *Functions of Real Variables*, Tel Aviv University, 2015.
[13] Tom Lindstrom, *Mathematical Analysis*, University of Oslo, 2013.
[14] Marco Gualtieri, *Geometry and Topology*, University of Toronto, 2009.
[15] Mathematics Stack Exchange, Available: https://math.stackexchange.com.
[16] Wikipedia, Available: https://en.wikipedia.org.
[17] 나무위키, Available: https://namu.wiki.

Chapter 3

Probability Theory

Abstract 확률론은 20세기 들어 급격하게 발전한 분야이다. 애초에 확률이라는 개념이 수학에 편입된 것이 그리 오래되지 않았다. 이는 Descart의 연역주의의 영향이 진하게 남아 있던 근대 유럽의 수학에서 불확실성을 다루기를 꺼려했기 때문이다. 오죽했으면 "거의 확실한 것은 거의 확실히 거짓이다."라고까지 했을까. 하지만 도박 문제(de Méré's problem)와 같이 불확실성을 계량하여 다루어야 할 필요성은 조금씩 늘어갔고, 이러한 현실적 요구에 확률은 Pascal, Fermat, Lagrange 등의 기라성같은 수학자들에 의해 조금씩 건드려지기 시작했다. 이때까지만 하더라도 확률이 무엇인지에 대한 수학자들의 생각은 '어떤 사건이 발생할 가능성' 정도였다. 이러한 확률의 의미가 직관적으로 분명하였기에 이에 의문을 제기하는 사람도 없었고, 그럴 필요도 느끼지 못했다. 그러나 미적분학에서 극한의 개념이 그러하였듯, 확률에 대한 연구가 계속될수록 미묘한 잡음이 발생하기 시작했고, 이는 확률의 개념에 대한 엄밀한 수학적 접근이 필요함을 암시했다. 결국 '확률은 무엇인가?' 라는 질문의 답을 찾기 위한 긴 여정이 시작되었고, Laplace가 확률에 해석학을 끼얹은 것을 시작으로 Kolomogorov가 그의 명저 *Grundbegriffe der Wahrscheinlichkeitsrechnung*(영어: *Foundations of The Theory of Probability*)에서 측도론으로 확률을 정의하면서 그 여정은 일단락되게 된다. 본 장에서는 그 여정의 끝에서 수학자들이 괘뚫어본 확률의 본질에 대해 살펴보도록 하자.

3.1 Probability Spaces

단도직입적으로 말하면, 확률은 측도의 특별한 한 종류에 불과하다. 곧 확률은 일종의 넓이나 부피와 같은 개념으로 무언가를 재는 역할을 한다. 현실적인 의미를 생각하면 '가능성' 을 잰다고도 할 수 있겠다.

Definition 3.1 측도공간 $(\Omega, \mathscr{F}, \mathbb{P})$에 대해 $\mathbb{P}(\Omega) = 1$이면 이때의 유한 측도 $\mathbb{P}$를 **확률측도(probability measure)**라 하고, 유한 측도공간 $(\Omega, \mathscr{F}, \mathbb{P})$를 **확률공간(probability space)**이라 한다. 나아가 집합 Ω를 **표본공간(sample space)**이라 하고, σ-대수 $\mathscr{F}$에 속하는 임의의 집합 E를 **사건(event)**이라 하여 $\mathbb{P}(E)$의 값을 사건 E의 **확률(probability)**이라 한다.

확률의 본질이 측도라는 위의 정의는 나름 설득력이 있다. 그렇다면 이걸로 다 된 것일까? 아쉽게도 이제부터 해야 할 일이 태산이다. 일단 위의 정의를 받아들이기로 했다면, 지금까지 우리가 배웠던 확률의 대한 모든 내용들을 측도론의 언어로 다시 써야 한다. 곧 확률론을 다루는 본 장의 내용은 기본적으로 '번역 작업'으로, 다행히 대부분의 경우 이 번역 작업은 크게 어렵지 않을 것이다. 이는 측도론이 확률의 내용들을 형식화하기에 좋은 이론이라서이기도 하지만, 앞서 우리가 측도론을 배우며 이 순간을 위해 조금씩 준비해 둔 것들이 꽤 많기 때문이다.

본격적으로 시작하기에 앞서, 맥락상 다소 뜬금없기는 하지만, 우리의 확률에 대한 인식의 근간을 이루는 equally likely outcome model을 한 번은 언급하고 지나가는 것이 좋을 것 같다. 고등학교에서 경우의 수를 세는 문제로 흔히 접하는 equally likely outcome model은 표본공간으로 항상 유한집합 Ω를 가지고, 이의 모든 부분집합은 사건으로 간주된다. 나아가 한원소 집합인 사건은 특별히 **근원사건(elementary event)**이라 불리며 각 근원사건의 확률은 정확히 $1/|\Omega|$로 주어진다. (이렇게 각 근원사건의 확률이 같으므로 'equally likely'이다. 고등학교에서는 흔히 '같은 정도로 확실하다'로 번역한다.) 이러한 setting에서 우리는 임의의 사건 E의 확률을 $|E|/|\Omega|$로 정하고, 여기서의 $|E|$를 구하기 위해 여태껏 경우의 수를 열심히 계산하여 왔다. 이상의 내용을 측도론의 언어로 담백하게 번역하면 equally likely outcome model은 '표본공간 Ω에서의 셈측도 #에 대해 $(\Omega, \mathscr{P}(\Omega), \#/|\Omega|)$로 주어진 확률공간'이 된다. 곧 측도론으로 바라보면 equally likely outcome model도 측도공간의 특별한 한 종류에 불과하다.

보통 초급 확률론 교재에서는 고등학교에서와 마찬가지로 이 equally likely outcome model에 집중하여 경우의 수를 계산하는 fancy한 trick을 소개하는 데 많은 분량을 할애하곤 하지만, 여기서는 이에 대한 논의는 Sheldon Ross의 A First Course in Probability를 참고문헌으로 실어두는 것으로 대신한다. 우리는 일반적인 이론의 전개에 보다 관심이 있기에 확률공간의 한 예시에 불과한 equally likely outcome model에 대해서는 그다지 관심이 없고, 곧 이 책에서 equally likely outcome model이 다시 등장하는 일은 (아마) 없을 것이다. 또한, 일반적으로 표본공간 위의 σ-대수 $\mathscr{F}$가 모든 한원소 집합을 포함한다는 보장이 없으므로 근원사건이라는 개념도 그다지 쓸모가 없을 것이다.

다시 원래의 이야기로 돌아와, 본격적으로 번역 작업을 시작해보자. 우선 확률의 기본적인 성질 정도는 측도의 성질들로부터 거의 자명하게 얻어진다.

Theorem 3.2 확률공간 $(\Omega, \mathscr{F}, \mathbb{P})$와 사건 E, F에 대해 다음이 성립한다.

i. $\mathbb{P}(\emptyset) = 0$.
ii. ($\sigma-$가법성) 서로소인 사건열 $\{E_i\}$에 대해 $\mathbb{P}(\bigsqcup_{i=1}^{\infty} E_i) = \sum_{i=1}^{\infty} \mathbb{P}(E_i)$이다.
iii. $0 \le \mathbb{P}(E) \le 1$.
iv. $\mathbb{P}(E^c) = 1 - \mathbb{P}(E)$.
v. (단조성) 만약 $E \subseteq F$이면 $\mathbb{P}(E) \le \mathbb{P}(F)$이다.

Proof 이는 정의와 측도의 기본적인 성질로부터 자명하다. □

Theorem 3.3 확률공간 $(\Omega, \mathscr{F}, \mathbb{P})$에 대해 다음이 성립한다.

i. (포함배제의 원리) 사건 $E_1, \cdots, E_l$에 대해

$$\mathbb{P}\left(\bigcup_{i=1}^{l} E_i\right) = \sum_{i=1}^{l} (-1)^{i-1} \sum_{1 \le j_1 < \cdots < j_i \le l} \mathbb{P}\left(\bigcap_{k=1}^{i} E_{j_k}\right)$$

이다.
ii. ($\sigma-$반가법성) 사건열 $\{E_i\}$에 대해 $\mathbb{P}(\bigcup_{i=1}^{\infty} E_i) \le \sum_{i=1}^{\infty} \mathbb{P}(E_i)$이다.

Proof i는 측도론의 포함배제의 원리를 $(\Omega, \mathscr{F}, \mathbb{P})$에 적용한 결과이고, ii는 측도의 $\sigma-$가법성이 $\sigma-$반가법성을 함의한다는 점에서 자명하다. □

Theorem 3.4 확률공간 $(\Omega, \mathscr{F}, \mathbb{P})$와 사건열 $\{E_i\}$에 대해

$$\mathbb{P}(\liminf_{i\to\infty} E_i) \le \liminf_{i\to\infty} \mathbb{P}(E_i) \le \limsup_{i\to\infty} \mathbb{P}(E_i) \le \mathbb{P}(\limsup_{i\to\infty} E_i)$$

가 성립한다. 특별히, $E_i \to E$이면 $\mathbb{P}(E_i) \to \mathbb{P}(E)$이다.

Proof 이는 정리 2.26로부터 자명하다. □

위의 정리의 단서를 흔히 **확률측도의 연속성(continuity of probability measure)**이라 한다. 한편, 확률론에서는 사건열 $\{E_i\}$에 대해 이의 상극한 $\limsup_{i\to\infty} E_i$를 간단히 E_i io. 라 쓰기도 한다. 여기서 io.는 infinitely often의 줄임말로 곧 $\omega \in E_i$ io.는 $\omega \in \Omega$가 사건 $E_1, E_2, \cdots$에 무한히 많이 속한다는 것인데, 이는 $\limsup_{i\to\infty} E_i = \bigcap_{j=1}^{\infty} \bigcup_{i=j}^{\infty} E_i$임을 생각해 보면 나름 make sense하는 표기법이다.

Definition 3.5 확률공간 $(\Omega, \mathscr{F}, \mathbb{P})$에 대해 영집합인 사건을 **영사건(null event)**이라 한다.

확률측도도 측도이기에 '($\mathbb{P}-$)거의 어디서나'라는 개념이 자주 쓰이는데, 확률론에서는 이를 **($\mathbb{P}-$)거의 확실하게(($\mathbb{P}-$)almost surely)**라 하기도 한다. 이는 영사건이 확률이 0인 사건이므로 어떤 성질이 $\mathbb{P}-$거의 어디서나 성립하면 곧 1의 확률로 성립하게 되기 때문에 생겨난 관례이다.

이와 관련하여 '확률이 0인 사건'과 '불가능한 사건'은 서로 다르다는 것에 주의할 필요가 있다. 확률인 0인 사건은 표본공간 위의 σ−대수 $\mathscr{F}$에 속하는 사건이지만 그 확률이 0일 뿐이고, 불가능한 사건은 애초에 $\mathscr{F}$에 속하지 않아 그 이름과는 달리 엄밀히는 사건이 아니다. 간단한 예시를 위해 $[0,1]$에서 임의로 점 하나를 택하는 상황을 생각해보자. 이를 확률공간으로 형식화한다면, 표본공간은 $\Omega = [0,1]$이고 $\mathscr{F} = \mathscr{B}_1|_\Omega$, 확률측도는 $\mathbb{P} = (\mu_1)_\mathscr{F}$ 정도로 둘 수 있을 것이다. 그렇다면 $\mathbb{P}\{0.5\} = 0$이므로 정확히 0.5를 뽑는 사건은 확률이 0인 사건이지만, 그렇다고 이가 일어나는 것 자체가 불가능한 것은 아니다. 반대로, 2를 뽑는 사건은 애초에 발생이 불가능한 사건으로 이 경우에 $\{2\} \notin \mathscr{F}$이므로 이는 엄밀하게는 사건이 아니어서 확률의 부여가 불가능하다. 이와 비슷하게, '확률이 1인 사건'과 '항상 발생하는 사건'도 서로 다른다.

이어서, 조건부확률을 도입하고 그 성질을 측도론으로 보이자. 다만, 이후에 조건부기댓값과 조건부분포를 엄밀히 도입하기 위해서는 조건부확률의 개념을 격변에 준할 정도로 일반화시켜야 하는데, 이에 하나의 절을 오롯이 할애해야 할 정도의 논의가 필요하므로 여기에서는 아주 간단하게만 다루도록 한다.

Definition 3.6 확률공간 $(\Omega, \mathscr{F}, \mathbb{P})$와 영사건이 아닌 사건 E에 대해 **사건 E에 대한 조건부확률(conditional probability under event E)**을 $\mathbb{P}(\cdot|E) : \mathscr{F} \to \mathbb{R}_0^+$로 쓰고 $\mathbb{P}(\cdot|E) : F \mapsto \mathbb{P}(F \cap E)/\mathbb{P}(E)$로 정의한다.

Proposition 3.7 확률공간 $(\Omega, \mathscr{F}, \mathbb{P})$와 영사건이 아닌 사건 E에 대한 조건부확률 $\mathbb{P}(\cdot|E)$는 확률측도이다. 따라서 $(\Omega, \mathscr{F}, \mathbb{P}(\cdot|E))$는 확률공간을 이룬다.

Proof 우선 $\mathbb{P}(\emptyset|E) = \mathbb{P}(\emptyset)/\mathbb{P}(E) = 0$, $\mathbb{P}(\Omega|E) = \mathbb{P}(E)/\mathbb{P}(E) = 1$임은 분명하고, 임의의 서로소인 사건열 $\{E_i\}$에 대해

$$
\begin{aligned}
\mathbb{P}\left(\bigsqcup_{i=1}^{\infty} E_i | E\right) &= \frac{\mathbb{P}(\bigsqcup_{i=1}^{\infty} E_i \cap E)}{\mathbb{P}(E)} \\
&= \frac{\mathbb{P}(\bigsqcup_{i=1}^{\infty} (E_i \cap E))}{\mathbb{P}(E)} \\
&= \sum_{i=1}^{\infty} \frac{\mathbb{P}(E_i \cap E)}{\mathbb{P}(E)} \\
&= \sum_{i=1}^{\infty} \mathbb{P}(E_i|E)
\end{aligned}
$$

이므로 $\mathbb{P}(\cdot|E)$가 확률측도임을 안다. □

Theorem 3.8 확률공간 $(\Omega, \mathscr{F}, \mathbb{P})$에 대해 다음이 성립한다.

i. 사건 E와 영사건이 아닌 사건 F에 대해 $\mathbb{P}(E \cap F) = \mathbb{P}(E|F)\mathbb{P}(F)$이다.

ii. (전확률 공식) 서로소인 가산개의 사건 $E_1, E_2, \cdots$에 대해 각 E_i가 영사건이 아니고 $\bigsqcup_{i=1}^k E_i = \Omega$이면 임의의 사건 E에 대해 $\mathbb{P}(E) = \sum_{i=1}^k \mathbb{P}(E|E_i)\mathbb{P}(E_i)$이다. (여기서 k는 유한할 수도 있고, ∞일 수도 있다.)

Proof i. 이는 조건부확률의 정의로부터 자명하다.

ii. i로부터 $\mathbb{P}(E) = \mathbb{P}(E \cap \bigsqcup_{i=1}^k E_i) = \mathbb{P}(\bigsqcup_{i=1}^k (E \cap E_i)) = \sum_{i=1}^k \mathbb{P}(E \cap E_i) = \sum_{i=1}^k \mathbb{P}(E|E_i)\mathbb{P}(E_i)$ 이다. □

Theorem 3.9 (Bayes) 확률공간 $(\Omega, \mathscr{F}, \mathbb{P})$와 서로소인 가산개의 사건 $E_1, E_2, \cdots$에 대해 각 E_i가 영사건이 아니고 $\bigsqcup_{i=1}^k E_i = \Omega$이면 임의의 사건 E에 대해

$$\mathbb{P}(E_1|E) = \frac{\mathbb{P}(E|E_1)\mathbb{P}(E_1)}{\sum_{i=1}^k \mathbb{P}(E|E_i)\mathbb{P}(E_i)}$$

이다. (여기서 k는 유한할 수도 있고, ∞일 수도 있다.)

Proof 전확률 공식으로부터 $\mathbb{P}(E_1|E) = \mathbb{P}(E \cap E_1)/\mathbb{P}(E) = \mathbb{P}(E|E_1)\mathbb{P}(E_1)/\sum_{i=1}^k \mathbb{P}(E|E_i)\mathbb{P}(E_i)$가 자명하다. □

비록 증명은 간단하지만 Bayes의 정리는 실험적으로 구하는 것이 불가능한 조건부확률을 구하게 해준다는 엄청난 실용성과 함의를 지닌다. 쉽고 즐거운 예시를 위해 유리가 예나에게 휴대전화로 0x2661(UTF-16 ♡)과 0x2665(UTF-16 ♥) 중 하나의 정보를 임의로 전송하였는데, 송수신의 과정에서 정보의 일부가 손상되어 결과적으로 예나는 0x2663(UTF-16 ♣)을 수신한 상황을 생각해보자. 이를 복원하기 위해 예나는 E를 0x2663을 수신하는 사건, F, G를 각각 0x2661과 0x2665를 전송하는 사건이라 두고 두 조건부확률 $\mathbb{P}(F|E)$와 $\mathbb{P}(G|E)$를 비교하여 전자가 더 크다면 0x2661로 복원하고 후자가 더 크다면 0x2665로 복원하며, 만약 같다면 재전송을 요청하기로 하였다. 이는 훌륭한 통계적 사고방식이지만 $\mathbb{P}(F|E)$와 $\mathbb{P}(G|E)$는 일의 선후가 뒤바뀐 확률이라 실험적으로 구할 수가 없다는 치명적인 문제가 있다. 예나가 자신의 휴대전화로 자신에게 0x2661과 0x2665를 수 회 전송하는 실험을 통해 근사하게나마 구할 수 있는 확률은 $\mathbb{P}(E|F)$와 $\mathbb{P}(E|G)$ 뿐이다. 여기서 Bayes의 정리는 $\mathbb{P}(F|E)$와 $\mathbb{P}(G|E)$를 $\mathbb{P}(E|F)$와 $\mathbb{P}(E|G)$의 조합으로써 계산할 수 있도록 하여 문제를 해결하는 결정적인 역할을 한다. 따라서 예나가 실험적으로 구한 $\mathbb{P}(E|F)$와 $\mathbb{P}(E|G)$의 근사치가 만약 1/200, 1/300이었다면 Bayes의 정리로부터 $\mathbb{P}(F|E) \approx 3/5, \mathbb{P}(G|E) \approx 2/5$가 되어 ♣를 ♡로 복원할 수 있다. 요컨대, Bayes의 정리는 선후관계나 인과관계를 역전시켜주는 마법의 공식이다.

이러한 Bayes 정리는 이후 통계학에 큰 지각변동을 일으켜 이를 기초로 하는 Bayesian이라는 독자적인 학파가 구성되기에 이르렀고, 오늘날 기계학습과 같은 분야에서 요긴하게 쓰이는 모양이다. (이와 구분하여 기존의 통계학 학파를 빈도주의라 한다.) 물론, 이런 학

파는 어디까지나 확률의 해석에 대한 차이로 구분되는 것이지 확률의 측도론적 정의나 접근방식으로 구분되는 것은 아니기에 빈도주의와 Bayesian의 구분이 본 장에서는 필요하지 않지만, 이 책에서는 특별한 언급이 없는 이상 빈도주의의 관점에서 확률을 바라본다. 여기에는 통계학 교양 정도를 들은 수준에서는 빈도주의의 관점이 Bayesian의 관점보다 조금 더 친숙하다는 것을 빼면 다른 그럴싸한 이유는 없다. 만약 자신이 Bayesian이라면 그들의 방식대로 해석하면 그만이고, 당연히 Bayesian에게도 측도론적인 엄밀한 확률론은 훌륭한 이론의 토대가 될 것이다.

이번 절에서 마지막으로 다룰 것은 바로 독립에 관한 내용인데, 앞서 확률의 기본적인 성질과 기본적인 조건부확률을 큰 어려움 없이 도입할 수 있었던 것과는 달리, 독립성을 측도론의 언어로 번역해 내는 것은 살짝 어렵다. 이는 독립이라는 개념이 측도론에서는 그 느낌조차 찾아보기 힘든, 완전히 새로운 개념이기 때문이다. 우리는 측도론에서 적분을 정의할 때와 비슷하게 독립을 그 정의를 점차 일반화시키는 방법으로 도입할 것이다. 우선 가장 기본적인 독립의 정의로 시작한다.

Definition 3.10 확률공간 $(\Omega, \mathscr{F}, \mathbb{P})$와 사건 $E_1, \cdots, E_k$를 생각하자. 만약 각 $l \leq k$와 임의의 서로다른 $i_1, \cdots, i_l \leq k$에 대해 $\mathbb{P}(\bigcap_{j=1}^{l} E_{i_j}) = \prod_{j=1}^{l} \mathbb{P}(E_{i_j})$가 성립하면 이때의 사건 $E_1, \cdots, E_k$를 **(서로) 독립((mutually) independent)**이라 한다. 한편, 만약 위의 성질이 $l = 2$에 대해서만 만족되면, 즉 임의의 서로다른 $i, j \leq k$에 대해서 $\mathbb{P}(E_i \cap E_j) = \mathbb{P}(E_i)\mathbb{P}(E_j)$가 성립하는 것에 그치면 이때의 사건 $E_1, \cdots, E_k$를 **pairwise 독립(− independent)**이라 한다.

서로 독립인 것과 pairwise 독립이 다르다는 사실은 잘 알려진 사실이다. 흔해빠진 주사위와 동전 예시를 피하기 위해 해석개론을 수강신청한 수지와 이제훈이 같이 밤새 과제를 하게 된 것을 계기로 서로 어느정도 이성으로서 호감을 가지게 된, 이른바 '썸 탄다' 불리우는 상황을 생각하자. 수업을 듣던 어느날, 수지와 이제훈의 교재가 실수로 서로 바뀌어 다음 수업시간 전에 만나 책을 다시 바꾸기로 하였는데, 둘은 이를 앞두고 책 사이에 고백편지를 살짝 끼워넣을까 고민하고 있다고 하자. 여기서 사건 E, F를 각각 수지와 이제훈이 고백편지를 끼워넣는 사건이라 하고, 수지나 이제훈이 고백편지를 넣을 확률은 $1/2$로 같으며 이는 서로 독립이라 하자. 이제 사건 G를 어느 한쪽만 고백편지를 받는 사건이라 하면 E, F, G는 pairwise 독립이지만 서로 독립은 아니다. (직접 계산해보자.)

이제 유한개의 사건 사이의 독립을 무한개의 사건 사이의 독립으로 확장하는 것은 어렵지 않다.

Definition 3.11 확률공간 $(\Omega, \mathscr{F}, \mathbb{P})$와 사건의 모임 $\mathscr{C}$를 생각하자. 만약 임의의 사건 $E_1, \cdots, E_k \in \mathscr{C}$가 서로 독립이면 이때의 집합 $\mathscr{C}$를 **독립(independent)**이라 한다.

다음으로, 유한개의 사건의 모임 사이의 독립을 정의한다.

Definition 3.12 확률공간 $(\Omega, \mathscr{F}, \mathbb{P})$와 사건의 모임 $\mathscr{C}_1, \cdots, \mathscr{C}_k$를 생각하자. 만약 각 $i \le k$에 대해 임의로 택한 사건 $E_i \in \mathscr{C}_i$가 서로 독립이면 이때의 집합족 $\mathscr{C}_1, \cdots, \mathscr{C}_k$를 **독립 (independent)**이라 한다.

마지막으로 이를 무한개의 사건의 모임 사이의 독립으로까지 확장하면 독립의 가장 일반적인 정의를 얻는다.

Definition 3.13 확률공간 $(\Omega, \mathscr{F}, \mathbb{P})$와 사건의 모임의 모임 Γ를 생각하자. 만약 임의의 사건의 모임 $\mathscr{C}_1, \cdots, \mathscr{C}_k \in \Gamma$가 서로 독립이면 이때의 집합족 Γ를 **독립(independent)**이라 한다.

이렇게까지 일반적인 형태의 독립성을 고려하는 이유는 다음 정리 때문이다.

Theorem 3.14 확률공간 $(\Omega, \mathscr{F}, \mathbb{P})$와 독립인 사건의 모임의 모임 $\{\mathscr{C}_\alpha\}$에 대해 각 $\mathscr{C}_\alpha$가 $\pi-$system이라 하면 $\{\sigma(\mathscr{C}_\alpha)\}$는 독립이다.

Proof 집합족 $\{\mathscr{C}_\alpha\}$에서 임의로 유한개의 원소를 택하여 이를 $\mathscr{C}_1, \cdots, \mathscr{C}_k$라 하고 이들이 생성하는 $\sigma-$대수 $\sigma(\mathscr{C}_1), \cdots, \sigma(\mathscr{C}_k)$가 서로 독립임을 보이면 증명은 충분하다. 이를 위해 사건 $E_2 \in \mathscr{C}_2, \cdots, E_k \in \mathscr{C}_k$를 임의로 택하고 집합족 $\mathscr{L} = \{F \in \mathscr{F} : F, E_2, \cdots, E_k$가 서로 독립$\}$을 생각하면 주어진 조건으로부터 $\mathscr{C}_1 \subseteq \mathscr{L}$임은 분명하다. 나아가 $\Omega \in \mathscr{L}$ 또한 분명하고, 임의의 $F \in \mathscr{L}$에 대해 $F, E_2, \cdots, E_k$가 서로 독립이면 $F^c, E_2, \cdots, E_k$도 서로 독립임을 쉽게 보일 수 있으므로 $F^c \in \mathscr{L}$이다. 비슷한 방법으로 $\mathscr{L}$에 속하는 임의의 서로소인 사건열 $\{F_i\}$에 대해 $\bigsqcup_{i=1}^{\infty} F_i, E_2, \cdots, E_k$가 서로 독립임을 보일 수 있으므로 $\bigsqcup_{i=1}^{\infty} F_i \in \mathscr{L}$도 성립한다. 이로부터 $\mathscr{L}$은 $\lambda-$system이 되어 $\mathscr{C}_1$이 $\pi-$system이라는 사실과 Dynkin의 $\pi-\lambda$ 정리로부터 $\sigma(\mathscr{C}_1) = \lambda(\mathscr{C}_1) \subseteq \mathscr{L}$이고, 곧 $\sigma(\mathscr{C}_1), \mathscr{C}_2, \cdots, \mathscr{C}_k$는 서로 독립이다. 이제 이를 $k-1$번 반복하면 $\sigma(\mathscr{C}_1), \cdots, \sigma(\mathscr{C}_k)$가 서로 독립임을 보일 수 있고, 증명이 끝난다. □

위의 정리는 서로 독립인 사건의 모임들이 생성하는 $\sigma-$대수도 서로 독립임을 함의하는데, 잠시 이 결과의 의미에 대해 생각해보자. 독립성은 기본적으로 '정보'에 대한 이야기이다. 가장 기본적인 형태로 두 사건 E, F가 서로 독립인 경우를 살펴보면, 이는 사건 E의 발생여부에 대한 정보가 주어지더라도 사건 F의 발생여부에 대한 정보는 일체 추론해 낼 수 없음을 의미한다. 이는 2개 이상의 사건, 나아가 무한개의 사건의 독립에 대해서도 마찬가지이다. 예컨대 앞서 든 수지와 이제훈의 예시에서 E, F, G는 서로 독립이 아니었는데, 이는 수지가 고백편지를 넣는 사건 E와 이제훈이 고백편지를 넣는 사건 F의 발생여부에 대한 정보가 사건 G의 발생여부에 대한 정보를 100% 함의하기 때문이다.

사건의 모임 사이의 독립도 이와 비슷하게 이해할 수 있다. 앞서 사건들 사이의 독립에서 각 사건은 그 사건의 발생여부에 대한 이진 정보를 의미했다. 그렇다면 사건의 모임은 그 모임에 속한 각 사건들의 발생여부에 대한 이진 정보의 조합으로 이루어진 보다 풍성한 정

보로 이해하는 것이 자연스러울 것이다. 따라서 사건의 모임 사이의 독립은 각 사건의 모임이 함의하는 정보가 서로 독립적이라는, 즉 어느 하나를 안다고 해서 다른 하나를 일체 추론해내지 못한다는 것을 의미한다. 이러한 해석의 관점에서 보면, 위의 정리의 결과는 어떤 사건의 모임들이 함의하는 정보가 서로 추론이 불가능하다면, 각 모임을 그가 생성하는 σ−대수로 확장하여 훨씬 더 많은 정보로 얻어도, 여전히 서로 추론이 불가능함을 의미한다. 뭔가 자명한 듯 자명하지 않은 이 결과를 보다 효율적으로 쓰기 위해 따름정리 하나를 소개한다.

Corollary 3.15 확률공간 $(\Omega, \mathscr{F}, \mathbb{P})$와 가산개의 무한한 행과 열을 가지는 사건의 배열

$$\begin{matrix} E_{11} & E_{12} & \cdots \\ E_{21} & E_{22} & \cdots \\ \vdots & \vdots & \ddots \end{matrix}$$

에 대해 $\{E_{ij}\}_{ij}$가 독립이라 하고, 각 $i \in \mathbb{N}$에 대해 집합족 $\mathscr{G}_i$를 $\{E_{ij}\}_j$가 생성하는 σ−대수라 하면 $\{\mathscr{G}_i\}$는 독립이다. 한편, 행이나 열의 개수가 유한한 경우에도 같은 결과가 성립하며, 나아가 각 행의 열의 개수가 달라도 가산개이기만 하면 여전히 같은 결과가 성립한다.

Proof 각 $i \in \mathbb{N}$에 대해 $\{E_{ij}\}_j$의 모든 유한 교집합의 모임 $\mathscr{P}_i$를 생각하면 $\mathscr{P}_i$는 π−system이고 $\sigma(\mathscr{P}_i) = \mathscr{G}_i$임이 거의 분명하다. 따라서 $\{\mathscr{P}_i\}$가 서로 독립이라는 사실만 보이면 앞선 정리로부터 $\{\mathscr{G}_i\}$가 독립이 되어 증명이 끝난다. 이를 위해 $\{\mathscr{P}_i\}$에서 임의로 유한개의 원소를 택하여 이를 $\mathscr{P}_1, \cdots, \mathscr{P}_k$라 하고 다시 임의로 사건 $F_1 \in \mathscr{P}_1, \cdots, F_k \in \mathscr{P}_k$를 택하면 각 $i \le k$에 대해 $\mathscr{P}_i$의 구성으로부터 적당한 사건 $E_{i1}, \cdots, E_{il_i}$가 존재하여 $F_i = \bigcap_{j=1}^{l_i} E_{ij}$이다. 그렇다면 $\{E_{ij}\}_{ij}$가 독립이라는 사실로부터 $\mathbb{P}(\bigcap_{i=1}^{k} F_i) = \mathbb{P}(\bigcap_{i=1}^{k}\bigcap_{j=1}^{l_i} E_{ij}) = \prod_{i=1}^{k}\prod_{j=1}^{l_i} \mathbb{P}(E_{ij}) = \prod_{i=1}^{k} \mathbb{P}(\bigcap_{j=1}^{l_i} E_{ij}) = \prod_{i=1}^{k} \mathbb{P}(F_i)$가 되어 $\{\mathscr{P}_i\}$가 서로 독립임을 알고, 증명은 이로써 충분하다. □

이 따름정리를 적당히 응용하면 독립성에 대한 진부한 연습문제들, 예컨대 사건 E, F, G, H가 서로 독립이라면 $E \cap F$와 $G \setminus H$가 독립임을 보이라는 식의 문제들을 아주 깔끔하게 해결할 수 있다. 예시로 든 문제의 경우 배열 $E\,F//G\,H$를 생각하면 그만이다. 한편, 앞서 사건의 모임을 그에 포함된 각 사건의 발생여부로 구성된 정보의 집합으로 해석하였는데, 이는 독립의 경우에만 한정되는 해석이 아니어서 특히 그 모임이 σ−대수 $\mathscr{G}$인 경우 모든 사건의 집합 $\mathscr{F}$를 '전체 정보'로, $\mathscr{G}$는 이의 '부분 정보'로 해석하는 것이 유용한 경우가 많다.

나중을 위해 측도론에서 잠시 등장했던 Borel-Cantelli의 정리를 조금 보강하는 것으로 이번 절을 마친다.

Theorem 3.16 (Borel-Cantelli) 확률공간 $(\Omega, \mathscr{F}, \mathbb{P})$와 사건열 $\{E_i\}$에 대해 다음이 성립한다.

i. 만약 $\sum_{i=1}^{\infty}\mathbb{P}(E_i)<\infty$이면 $\mathbb{P}(E_i \text{ io.})=0$이다.
ii. 만약 $\sum_{i=1}^{\infty}\mathbb{P}(E_i)=\infty$이고 $\{E_i\}$가 독립이면 $\mathbb{P}(E_i \text{ io.})=1$이다.

Proof i. 이는 측도론에서 배운 Borel-Cantelli의 정리로부터 자명하다.

ii. 우선 $\{E_i\}$가 독립이므로 따름정리 3. 15로부터 $\{E_i^c\}$도 독립이다. 이제 임의의 $\varepsilon>0$과 임의의 $j\in\mathbb{N}$를 택하면 $\sum_{i=j}^{\infty}\mathbb{P}(E_i)=\infty$이므로 $\lim_{k\to\infty}\exp(-\sum_{i=j}^{k}\mathbb{P}(E_i))=0$이다. 이로부터 적당한 $k_0\in\mathbb{N}$가 존재하여 $k_0\geq j$이고 $\exp(-\sum_{i=j}^{k_0}\mathbb{P}(E_i))<\varepsilon$이므로 $\mathbb{P}(\bigcap_{i=j}^{k_0}E_i^c)=\prod_{i=j}^{k_0}\mathbb{P}(E_i^c)=\prod_{i=j}^{k_0}[1-\mathbb{P}(E_i)]\leq\prod_{i=j}^{k_0}\exp(-\mathbb{P}(E_i))=\exp(-\sum_{i=j}^{k_0}\mathbb{P}(E_i))<\varepsilon$이다. 이는 곧 $\mathbb{P}(\bigcap_{i=j}^{\infty}E_i^c)\leq\mathbb{P}(\bigcap_{i=j}^{k_0}E_i^c)<\varepsilon$임을 뜻하므로 $\mathbb{P}(\bigcap_{i=j}^{\infty}E_i^c)=0$임을 알고, 이는 다시 $\mathbb{P}(E_i \text{ io.})=\mathbb{P}((\liminf_{i\to\infty}E_i^c)^c)=1-\mathbb{P}(\liminf_{i\to\infty}E_i^c)\geq 1-\mathbb{P}(\bigcap_{i=j}^{\infty}E_i^c)=1$에서 $\mathbb{P}(E_i \text{ io.})=1$임을 뜻한다. □

3.2 Random Variables and Random Vectors

앞선 절에서 사건 그 자체에 관련된 확률의 내용들을 측도론의 틀에 맞추어 열심히 옮겨 놓았으니, 이번 절에서는 확률변수라는 개념을 추가하여 보다 내용을 풍성하게 만들어보도록 하자. 확률변수를 통해 우리는 일일히 사건을 정의하지 않고도 확률의 여러 내용들을 보다 더 편리하게 사용할 수 있다.

Definition 3.17 확률공간 $(\Omega,\mathscr{F},\mathbb{P})$에 대해 가측함수 $X:\Omega\to\mathbb{R}^n$를 **(n차원) 확률벡터((n dimensional) random vector)**라 하고, 특별히 $n=1$이면 **확률변수(random variable)**라 한다.

확률변수는 본질적으로 가측함수이기에 확률변수의 기본적인 성질들이 가측함수의 성질들로부터 자명하게 성립하는 것이 전혀 이상하지 않다.

Proposition 3.18 확률공간 $(\Omega,\mathscr{F},\mathbb{P})$에서 정의된 함수 $X:\Omega\to\mathbb{R}^n$에 대해 X가 rv.일 필요충분조건은 $X_1,\cdots,X_n$이 모두 rv.인 것이다.

Proof 이는 가측함수의 성분도 가측함수라는 점에서 자명하다. □

Theorem 3.19 확률공간 $(\Omega,\mathscr{F},\mathbb{P})$에서 정의된 rv. $X:\Omega\to\mathbb{R}^m$와 Borel 함수 $g:\mathbb{R}^m\to\mathbb{R}^n$에 대해 합성 $g\circ X:\Omega\to\mathbb{R}^n$도 rv.이다.

Proof 이는 Borel 함수와 가측함수의 합성은 가측이라는 점에서 자명하다. □

비록 정의상 확률벡터는 엄연한 함수이지만 그 이름에서도 잘 드러나듯이 확률론에서는 이를 마치 변수처럼 생각하고 사용하는 경우가 많다. 이는 이론적인 이유에서라기보다 실

생활의 응용에서 이렇게 생각하는 편이 조금 더 직관적으로 편하기 때문이다. 이러한 우리의 인식은 확률벡터와 관련된 여러 표기상의 관례에 잘 나타나는데, 위의 정리에서의 합성 $g \circ X$를 마치 g에 X라는 변수를 대입한 것으로 생각하여 $g(X)$로 쓰는 관례가 대표적인 예시이다. 그러나 이런 표기법은 어디까지나 관례일 뿐, 확률벡터가 변수가 아닌 함수라는 사실은 항상 염두에 두고 있어야 한다.

이어서, 측도론에서 FTC를 일반화하는 과정에서 잠시 스쳐 지나갔던 pushfowarding이 다시 등장한다. 비록 측도론에서의 pushfowarding은 도구 역할에 그쳤지만, 확률론에서의 pushfowarding은 빼놓을 수 없는 핵심적인 개념이다.

Definition 3.20 확률공간 $(\Omega, \mathscr{F}, \mathbb{P})$에서 정의된 rv. $X : \Omega \to \mathbb{R}^n$에 대해 pushfoward 측도 $X_*\mathbb{P}$를 rv. X의 **분포(distribution)**라 하고 $\mathbf{P}_X$로 쓴다. 특별히, $n \geq 2$인 경우 $\mathbf{P}_X$를 rv. $X_1, \cdots, X_n$의 **결합분포(joint distribution)**라 하기도 한다.

교양 통계학에서 배운 PDF나 CDF와 같이 확률론에는 방금 정의한 분포와 쉽게 혼동될 법한 개념들이 많고, 실제로 각자의 정의도 서로 긴밀히 연결되어 있다. 여기에 한술 더 떠서 문헌마다 조금씩 용어를 다르게 쓰는 바람에 혼란이 가중되는 부분이 없지 않지만, 대부분 논의의 맥락으로 적당히 구별할 수 있다. 아무튼 구태여 혼란을 초래할 필요는 없기에, 이 책에서는 용어를 최대한 잘 구별하여 사용하였다.

나중에 하나씩 보겠지만, 확률벡터가 가지는 성질과 특성 중에는 그 확률벡터 자체가 아닌, 그 분포에 의존하는 성질들이 있다. 곧, 이러한 성질들은 서로다른 확률공간에서 정의된 서로다른 확률벡터라도 분포만 같다면 서로 공유하게 된다. 이를 뒤집어 생각하면 때로는 하나의 확률공간에 정의된 확률벡터를 그 분포에 의존하는 성질들을 그대로 유지하면서 다른 확률공간으로 옮길 수도 있다. 이러한 이유로 분포가 같은 확률벡터를 다루게 되는 경우가 많고, 표기의 편의를 위해 확률공간 $(\Omega, \mathscr{F}, \mathbb{P}), (\Omega', \mathscr{F}', \mathbb{P}')$에서 각각 정의된 확률벡터 $X : \Omega \to \mathbb{R}^n, Y : \Omega' \to \mathbb{R}^n$에 대해 $\mathbf{P}_X = \mathbf{P}_Y$이면 이를 간단히 $X \equiv Y$로 쓴다. 이때 $X \equiv Y$이지만 $X \neq Y$일 수 있다는 사실에 주의해야 한다.

Proposition 3.21 확률공간 $(\Omega, \mathscr{F}, \mathbb{P})$에서 정의된 rv. $X : \Omega \to \mathbb{R}^n$에 대해 $(\mathbb{R}^n, \mathscr{B}_n, \mathbf{P}_X)$는 확률공간을 이룬다.

Proof 먼저 $\mathscr{B}_n \subseteq X_*\mathscr{F}$임을 보이기 위해 임의의 $A \in \mathscr{B}_n$를 택하면 $X^{-1}(A) \in \mathscr{F}$이므로 $A \in X_*\mathscr{A}$에서 $\mathscr{B}_n \subseteq X_*\mathscr{F}$이다. 따라서 $\mathbf{P}_X$가 확률측도임을 보이면 충분한데, 이는 $\mathbf{P}_X(\mathbb{R}^n) = \mathbb{P}(X^{-1}(\mathbb{R}^n)) = \mathbb{P}(\Omega) = 1$에서 쉽게 알 수 있고, 곧 증명이 끝난다. □

위의 명제는 분포를 통해 서로다른 확률공간에 정의된 확률측도를 각각 다루는 대신 이들을 pushfowarding하여 $\mathscr{B}_n$에서 정의된 확률측도로 일관되게 다룰 수 있음을 함의한다. 그리고 생각해보면, 이가 곧 사건을 직접 정의하고 사용하는 대신 확률벡터를 도입하여 사

용하는 이유이다. 무엇이 될 지 모르는 확률공간 대신 우리가 잘 알고있는 실수공간에서의 확률측도를 다루는 것이 훨씬 편하다. 한편, 위의 정리의 역이 성립한다는 것도 꽤나 흥미로운 사실이다.

Theorem 3.22 Borel σ−대수 $\mathscr{B}_n$ 위의 확률측도 μ에 대해 적당한 확률공간 $(\Omega, \mathscr{F}, \mathbb{P})$에서 정의된 rv. $X : \Omega \to \mathbb{R}^n$가 존재하여 $\mu = \mathbf{P}_X$이다. 특별히, 이때 $(\Omega, \mathscr{F}) = (\mathbb{R}^n, \mathscr{B}_n)$이도록 잡을 수 있다.

Proof 거의 자명하다. 함수 $X : \mathbb{R}^n \to \mathbb{R}^n$를 항등함수로 두면 이는 확률공간 $(\mathbb{R}^n, \mathscr{B}_n, \mu)$에서 정의된 rv.이고, 임의의 사건 E에 대해 $\mathbf{P}_X(E) = \mu(X^{-1}(E)) = \mu(E)$에서 $\mu = \mathbf{P}_X$이다. □

Theorem 3.23 확률공간 $(\Omega, \mathscr{F}, \mathbb{P})$에서 정의된 rv. $X : \Omega \to \mathbb{R}^m$와 Borel 함수 $g : \mathbb{R}^m \to \mathbb{R}^n$에 대해 $\mathbf{P}_{g(X)} = \mathbf{P}_X \circ g^{-1}$이다.

Proof 임의의 $A \in \mathscr{B}_n$에 대해 $\mathbf{P}_{g(X)}(A) = \mathbb{P}((g \circ X)^{-1}(A)) = \mathbb{P}(X^{-1}(g^{-1}(A))) = \mathbf{P}_X(g^{-1}(A)) = (\mathbf{P}_X \circ g^{-1})(A)$이므로 $\mathbf{P}_{g(X)} = \mathbf{P}_X \circ g^{-1}$이다. □

자연스러운 다음 순서는 고등학교 시절부터 들어와 이름만은 익숙한 이산확률변수와 연속확률변수를 엄밀하게 측도론의 언어로 정의하는 것이다. 물론, 고등학교나 교양 통계학에서 각각을 정의하지 않는 것은 아니지만, 그 정의가 뭔가 어색하고 작위적이라는 느낌을 지우기 힘든데, 아래의 측도론적인 정의는 더할 나위 없이 깔끔하고 명쾌하다.

Definition 3.24 확률공간 $(\Omega, \mathscr{F}, \mathbb{P})$에서 정의된 rv. $X : \Omega \to \mathbb{R}^n$에 대해 $\mathbf{P}_X$가 이산측도이면 이때의 rv. X를 **이산확률벡터(discrete rv.)**라 한다. 또한, 만약 $\mathbf{P}_X$가 μ_n에 대해 절대연속이거나 특이연속이면 이때의 rv. X를 각각 **연속확률벡터(continuous rv.)** 혹은 **특이확률벡터(singular rv.)**라 한다. 특별히, $n = 1$인 경우 이산확률벡터, 연속확률벡터, 특이확률벡터를 각각 **이산확률변수**, **연속확률변수**, **특이확률변수**라 한다.

Proposition 3.25 확률공간 $(\Omega, \mathscr{F}, \mathbb{P})$에서 정의된 rv. $X : \Omega \to \mathbb{R}^n$에 대해 이는 이산확률벡터, 연속확률벡터, 특이확률벡터의 정의 중에서 두 개의 이상을 동시에 만족시킬 수 없다.

Proof 모순을 유도하기 위해 X가 이산확률벡터인 동시에 연속확률벡터라고 하자. 그렇다면 $\mathbf{P}_X$는 가산 지지집합 $A \in \mathscr{B}_n$를 가지는데, $\mu_n(A) = 0$에서 $\mathbf{P}_X(A) = 0$의 모순이 발생한다. 이번에는 X가 이산확률벡터인 동시에 특이확률벡터라 하자. 그렇다면 이전과 같이 $\mathbf{P}_X$는 가산 지지집합 $A \in \mathscr{B}_n$를 가지는데, 이의 가산개의 원소를 $x_1, x_2, \cdots$와 같이 나열하면 $\mathbf{P}_X(A) = \sum_{i=1}^{k} \mathbf{P}_X\{x_i\} = 0$에서 모순이 발생한다. (여기서 k는 유한할 수도 있고, ∞일 수도 있다.) 마지막으로 X가 연속확률벡터인 동시에 특이확률벡터라 하면 $\mathbf{P}_X \ll \mu_n$이고 $\mathbf{P}_X \perp \mu_n$이므로 $\mathbf{P}_X = 0$의 모순이 발생하고, 증명은 이로써 충분하다. □

일반적으로, 어떤 확률벡터가 이산인지, 연속인지, singular인지는 확률측도 $\mathbb{P}$와 확률벡터 X 모두에 의해 결정되는 것이지, 이 중 어느 하나에 의해 일방적으로 결정되는 것이 아니다. 즉, 둘 중 어느 하나만 보고서 X가 이산인지, 연속인지, singular인지는 알 수 없다. 이로 말미암아 우리가 확률벡터에 대해 당연하게 생각하던 사실들에 미묘한 혼란이 생겨나게 된다.

우선 X가 이산확률벡터이지만 그 치역은 가산이 아닐 수 있다. 물론, 대부분의 응용에서는 이산확률변수의 치역도 가산으로 주어지지만, 이는 우연의 일치 그 이상도 이하도 아니다. 극단적인 예시로 가측공간 $(\mathbb{R}, \mathscr{B}_1)$ 위의 측도 $\mathbb{P}$를

$$\mathbb{P} : A \mapsto \begin{cases} 1 & \text{when } 0 \in A \\ 0 & \text{ow.} \end{cases}$$

으로 잡아 확률공간 $(\mathbb{R}, \mathscr{B}_1, \mathbb{P})$를 구성하고 확률변수 $X : \mathbb{R} \to \mathbb{R}$를 항등함수로 두면 명백히 X는 모든 실수를 그 함숫값으로 가지고 심지어 연속이지만, 분포 $\mathbf{P}_X$가 한원소 집합 $\{0\}$을 지지집합으로 가지므로 X는 이산확률벡터이다. 다만, 정의로부터 분포가 가산 지지집합을 가지므로 X가 그 가산개의 값을 제외한 나머지 값을 가질 확률이 0이 되어 '사실상' 치역이 가산이라고 생각할 수는 있다. 하지만 영집합과 공집합이 비슷하지만 완전히 같지는 않은 것처럼 이 경우에도 '사실상' 치역이 가산인 것과 치역이 정말 가산인 것은 구분해야 할 것이다.

비슷하게, X가 연속확률벡터이지만 함수로서 연속이 아닐 수 있다. 애초에 표본공간 Ω에 위상구조가 존재한다는 보장이 없으므로 연속성을 논할 수조차 없다. 그렇다고 표본공간에 위상구조가 적당히 정의되어 있고, 이에 대해 X가 연속이라고 해서 이가 연속확률변수냐 하면 이것 또한 아니다. 앞서 든 예시를 생각해보면 표본공간이 $\mathbb{R}$이고 이 위에 표준위상을 잡더라도 X가 연속이지만 연속확률변수가 아닐 수 있다. 다만, 정의로부터 연속확률벡터와 특이확률벡터는 point mass를 가질 수 없으므로 확률이 표본공간의 어느 한 점에 집중되어 있지 않고 전체에 고르게 퍼져 있으니, 이런 의미에서 '연속'이라 생각할 수는 있다.

한편, 위의 정의에서 고등학교나 교양 통계학에서는 들어보지 못한 특이확률벡터라는 새로운 종류의 확률벡터가 등장했다. 다른 두 종류의 확률벡터는 고등학교 수준에서도 접할 수 있는 반면, 이제서야 특이확률벡터를 도입하는 것에는 그럴만한 이유가 있다. 우선 특이확률벡터는 다분히 이론적인 필요에 의한 확률벡터로 실생활의 응용에서는 거의 쓸모가 없다. 또한, 이산확률변수나 연속확률변수의 경우 기댓값이나 분산과 같은 개념의 도입과 계산이 쉬운 반면, 특이확률변수의 경우 이에 상당한 이론적 뒷받침이 필요하다. 이런 이유에서 이 책에서도 특이확률변수가 구체적인 예시로 주어지는 것은 이후에 배울 Cantor 분포 하나 뿐이다. 그렇다면 이런 단점에도 불구하고 특이확률벡터를 도입해야 할 이론적

인 필요가 대체 무엇인가? 다음 정리는 이 질문에 대한 답이자 측도론의 에필로그에서 예고한 이번 절의 클라이막스이다.

Theorem 3.26 확률공간 $(\Omega, \mathscr{F}, \mathbb{P})$ 위의 rv. $X : \Omega \to \mathbb{R}^n$에 대해 다음의 조건

i. Rv. $X_{\text{ac}} : \mathbb{R}^n \to \mathbb{R}^n$는 확률공간 $(\mathbb{R}^n, \mathscr{B}_n, \mathbb{P}_{\text{ac}})$에서 정의된 연속확률벡터이다.
ii. Rv. $X_{\text{pp}} : \mathbb{R}^n \to \mathbb{R}^n$는 확률공간 $(\mathbb{R}^n, \mathscr{B}_n, \mathbb{P}_{\text{pp}})$에서 정의된 이산확률벡터이다.
iii. Rv. $X_{\text{sc}} : \mathbb{R}^n \to \mathbb{R}^n$는 확률공간 $(\mathbb{R}^n, \mathscr{B}_n, \mathbb{P}_{\text{sc}})$에서 정의된 특이확률벡터이다.

를 만족하는 적당한 $\mathscr{B}_n$ 위의 확률측도 $\mathbb{P}_{\text{ac}}, \mathbb{P}_{\text{pp}}, \mathbb{P}_{\text{sc}}$와 rv. $X_{\text{ac}}, X_{\text{pp}}, X_{\text{sc}}$가 존재하여 $\alpha+\beta+\gamma = 1$인 적당한 $\alpha, \beta, \gamma \geq 0$에 대해 $\mathbf{P}_X = \alpha\mathbf{P}_{X_{\text{ac}}} + \beta\mathbf{P}_{X_{\text{pp}}} + \gamma\mathbf{P}_{X_{\text{sc}}}$이다.

Proof 정리 3.21로부터 $(\mathbb{R}^n, \mathscr{B}_n, \mathbf{P}_X)$가 확률공간이므로 Lebesgue의 분해정리로부터 $\mathbf{P}_X$는 절대연속성분 $(\mathbf{P}_X)_{\text{ac}}$, 순수 점 성분 $(\mathbf{P}_X)_{\text{pp}}$, 특이연속성분 $(\mathbf{P}_X)_{\text{sc}}$에 대해 $\mathbf{P}_X = (\mathbf{P}_X)_{\text{ac}} + (\mathbf{P}_X)_{\text{pp}} + (\mathbf{P}_X)_{\text{cs}}$와 같이 분해된다. 또한, $\mathbf{P}_X(\mathbb{R}^n) = 1$이므로 $\alpha = (\mathbf{P}_X)_{\text{ac}}(\mathbb{R}^n), \beta = (\mathbf{P}_X)_{\text{pp}}(\mathbb{R}^n), \gamma = (\mathbf{P}_X)_{\text{cs}}(\mathbb{R}^n)$라 하면 α, β, γ는 모두 유한하고 $\alpha+\beta+\gamma = 1$이다. 이제 α, β, γ가 모두 0이 아닌 특별한 경우를 생각해보자. 그렇다면 $\mathbb{P}_1 := (\mathbf{P}_X)_{\text{ac}}/\alpha, \mathbb{P}_2 := (\mathbf{P}_X)_{\text{pp}}/\beta, \mathbb{P}_3 := (\mathbf{P}_X)_{\text{cs}}/\gamma$가 모두 $\mathscr{B}_n$ 위의 확률측도이므로 $(\mathbf{P}_X)_{\text{ac}}, (\mathbf{P}_X)_{\text{pp}}, (\mathbf{P}_X)_{\text{cs}}$의 성질과 정리 3.22로부터 적당한 확률공간 $(\mathbb{R}^n, \mathscr{B}_n, \mathbb{P}_{\text{ac}})$에서 정의된 연속확률벡터 $X_{\text{ac}} : \mathbb{R}^n \to \mathbb{R}^n$, 적당한 확률공간 $(\mathbb{R}^n, \mathscr{B}_n, \mathbb{P}_{\text{pp}})$에서 정의된 이산확률벡터 $X_{\text{pp}} : \mathbb{R}^n \to \mathbb{R}^n$, 적당한 확률공간 $(\mathbb{R}^n, \mathscr{B}_n, \mathbb{P}_{\text{sc}})$에서 정의된 특이확률벡터 $X_{\text{cs}} : \mathbb{R}^n \to \mathbb{R}^n$가 존재하여 $\mathbb{P}_1 = \mathbf{P}_{X_{\text{ac}}}, \mathbb{P}_2 = \mathbf{P}_{X_{\text{pp}}}, \mathbb{P}_3 = \mathbf{P}_{X_{\text{cs}}}$이고, 곧 $\mathbf{P}_X = (\mathbf{P}_X)_{\text{ac}} + (\mathbf{P}_X)_{\text{pp}} + (\mathbf{P}_X)_{\text{cs}} = \alpha\mathbb{P}_1 + \beta\mathbb{P}_2 + \gamma\mathbb{P}_3 = \alpha\mathbf{P}_{X_{\text{ac}}} + \beta\mathbf{P}_{X_{\text{pp}}} + \gamma\mathbf{P}_{X_{\text{cs}}}$이다. 한편, α, β, γ 중에 일부가 0인 경우에 대해서도 이와 비슷하게 하면 된다. □

연속확률벡터, 이산확률벡터, 특이확률벡터의 세 가지 분류가 서로 배타적인 관계인 것은 맞지만 그렇다고 임의의 확률벡터가 반드시 이 셋 중 하나에 속하는 것은 아니다. 즉, 확률벡터 중에는 연속도, 이산도, singular도 아닌 골치아픈 것들이 존재한다. (이런 확률벡터를 흔히 **mixed type**이라 부르며 특이확률벡터에 비할 바는 아니지만 그 실용성은 많이 떨어지는 편이다.) 이런 상황에서 위의 정리는 임의의 확률벡터에 대해 비록 이가 mixed type이더라도 그 분포는 적당한 연속확률벡터, 이산확률벡터, 특이확률벡터의 분포의 합으로 분해할 수 있다는 놀라운 결과를 함의한다. 곧 연속확률벡터, 이산확률벡터, 특이확률벡터의 세 가지 분류는 확률벡터의 공간의 기저와 비슷한 역할을 하며, 이 세 가지 확률벡터를 정의한 순간 사실상 모든 확률벡터의 분류를 끝마친 것과 다름없다.

이러한 이유로 이론 전개에 있어서는 mixed type rv.를 고려할 필요 없이 연속확률벡터, 이산확률벡터, 특이확률벡터의 세 가지 확률벡터만 생각하면 되고, mixed type 확률벡터는 이 세 종류의 확률벡터의 성질들을 적당히 섞어 가질 뿐이다. 이러니 고등학교 시절부터 연속확률벡터, 이산확률벡터의 두 가지 종류에만 지대한 관심을 가진 것이 너무나 당연하다.

이론적으로 다루기 힘든 특이확률벡터를 제외하면 이 둘을 다룸으로써 우리는 고등학교때부터 우리도 모르는 사이에 사실상 온갖 종류의 확률벡터를 모두 다루고 있던 셈이다!

이제 클라이막스의 여운을 뒤로 하고, CDF를 살펴볼 순서이다. 흔히 교양 통계학에서는 PDF를 배운 뒤 CDF를 배우므로 PDF의 개념을 도입하지도 않고 CDF를 정의하는 것이 의아할 수 있다. 하지만, 적어도 이론적으로는 CDF가 PDF보다 더 기본적인 개념이기에 이를 먼저 도입한다.

Definition 3.27 확률공간 $(\Omega, \mathscr{F}, \mathbb{P})$에서 정의된 rv. $X : \Omega \to \mathbb{R}^n$에 대해 rv. X의 **(누적)분포함수((cumulative) distribution function)**를 $F_X : \mathbb{R}^n \to \mathbb{R}$로 쓰고 $F_X : x \mapsto \mathbf{P}_X(\prod_{i=1}^n(-\infty, x_i])$로 정의한다. 특별히, $n \geq 2$인 경우 F_X를 rv. $X_1, \cdots, X_n$의 **결합(누적)분포함수(joint (cumulative) distribution function)**라 하기도 한다.

CDF의 기본적인 성질은 정리 2. 58로부터 대부분 자명하게 유도된다.

Theorem 3.28 확률공간 $(\Omega, \mathscr{F}, \mathbb{P})$에서 정의된 rv. $X : \Omega \to \mathbb{R}^n$에 대해 다음이 성립한다.

i. $0 \leq F_X \leq 1$.
ii. 임의의 유계인 semi-open box $B \subseteq \mathbb{R}^n$에 대해 $\Delta_B F_X = \mathbf{P}_X(B) \geq 0$이다.
iii. CDF F_X는 각 변수에 대해 증가한다.
iv. CDF F_X는 오른쪽 연속이다.
v. 각 $i \leq n$에 대해 $\lim_{x_i \to -\infty} F_X(x) = 0$이다. [1]
vi. $\lim_{x_1, \cdots, x_n \to \infty} F_X(x) = 1$. [2]
vii. 임의의 $x \in \mathbb{R}^n$에 대해 $B_x = \prod_{i=1}^n(-\infty, x_i]$라 하면 $F_X(x-) = \mathbf{P}_X(B_x^\circ)$이고 $F_X(x) - F_X(x-) = \mathbf{P}_X(\partial B_x)$이다. [3]

Proof i – vi. 이는 CDF의 정의와 정리 2. 58로부터 자명하다.

vii. 임의의 $x \in \mathbb{R}^n$를 택하여 $B = \prod_{i=1}^n(-\infty, x_i]$라 하고, 집합열 $\{B_j\}$를 $B_j := \prod_{i=1}^n(-\infty, x_i - 1/j]$로 두면 이는 $\mathscr{S}_n$에 속하는 증가하는 집합열로서 $B_j \uparrow \prod_{i=1}^n(-\infty, x_i) = B^\circ$이다. 따라서 $F_X(x - \mathbf{1}/j) = \mathbf{P}_X(B_j) \uparrow \mathbf{P}_X(B^\circ)$이므로 임의의 $\varepsilon > 0$을 택하면 적당한 $j_0 \in \mathbb{N}$가 존재하여 $\mathbf{P}_X(B^\circ) - \mathbf{P}_X(B_{j_0}) < \varepsilon$이다. 이제 $\delta = 1/j_0$라 하면 $||x - y|| < \delta$이고 $x > y$인 모든 $y \in \mathbb{R}^n$에 대해 $B_{j_0} \subseteq \prod_{i=1}^n(-\infty, y_i) \subseteq B^\circ$에서 $\mathbf{P}_X(B^\circ) - \varepsilon < \mathbf{P}_X(B_{j_0}) \leq F_X(y) = \mathbf{P}_X(\prod_{i=1}^n(-\infty, y_i]) \leq \mathbf{P}_X(B^\circ)$이므로 $|F_X(y) - \mathbf{P}_X(B^\circ)| < \varepsilon$가 되어 $F_X(x-) = \mathbf{P}_X(B_x^\circ)$임을 안다. 이제 $F_X(x) - F_X(x-) = \mathbf{P}_X(B) - \mathbf{P}_X(B^\circ) = \mathbf{P}_X(\partial B)$임은 자명하다. □

위의 정리에서 $n = 1$인 경우 vii는 $F_X(x) - F_X(x-) = \mathbf{P}_X\{x\}$가 되어 이 경우에는 CDF를 분포의 point mass를 구하는 용도로 사용할 수 있다. 한편, 위의 정리의 역 비슷한 정리도 성립한다.

Theorem 3.29 함수 $F : \mathbb{R}^n \to [0, 1]$에 대해 이가

i. 함수 F는 오른쪽 연속이고 각 변수에 대해 증가한다.
ii. 유계인 semi-open box $B \subseteq \mathbb{R}^n$에 대해 $\Delta_B F \geq 0$이다.
iii. 각 $i \leq n$에 대해 $\lim_{x_i \to -\infty} F(x) = 0$이고 $\lim_{x_1, \cdots, x_n \to \infty} F_X(x) = 1$이다.

를 만족하면 적당한 확률공간 $(\Omega, \mathscr{F}, \mathbb{P})$에서 정의된 rv. $X : \Omega \to \mathbb{R}^n$가 존재하여 $F = F_X$이다. 특별히, 이때 $(\Omega, \mathscr{F}) = (\mathbb{R}^n, \mathscr{B}_n)$이도록 잡을 수 있다.

Proof 정리 2. 59로부터 적당한 $\mathscr{B}_n$ 위의 측도 μ가 존재하여 임의의 $x \in \mathbb{R}^n$에 대해 $F(x) = \mu(\prod_{i=1}^n (-\infty, x_i])$이다. 그런데 정리 2. 58의 vii와 주어진 조건으로부터 $\mu(\mathbb{R}^n) = 1$이 되어 μ는 확률측도이고, 곧 정리 3. 22로부터 적당한 확률공간 $(\Omega, \mathscr{F}, \mathbb{P})$ 위의 n차원 rv. X가 존재하여 $\mu = \mathbf{P}_X$이다. 이상으로부터 임의의 $x \in \mathbb{R}^n$에 대해 $F(x) = \mu(\prod_{i=1}^n (-\infty, x_i]) = \mathbf{P}_X(\prod_{i=1}^n (-\infty, x_i]) = F_X(x)$가 성립한다. 한편, 이때 $(\Omega, \mathscr{F}) = (\mathbb{R}^n, \mathscr{B}_n)$이도록 잡을 수 있음은 정리 3. 22로부터 자명하다. □

곧 우리는 CDF가 될 수 있는 함수를 위의 정리의 조건 i – iii으로 완벽히 characterize할 수 있다. 한편, CDF의 연속성에 대한 다음 정리들도 꽤나 흥미롭다.

Theorem 3.30 확률공간 $(\Omega, \mathscr{F}, \mathbb{P})$에서 정의된 rv. $X : \Omega \to \mathbb{R}^n$와 한 점 $x_0 \in \mathbb{R}^n$ 대해 $B_{x_0} = \prod_{i=1}^n (-\infty, x_0^i]$라 하면 TFAE.

i. CDF F_X가 x_0에서 연속이다.
ii. $F_X(x_0) = \mathbf{P}_X(B_{x_0}) = \mathbf{P}_X(B_{x_0}^\circ)$.
iii. $\mathbf{P}_X(\partial B_{x_0}) = 0$.

Proof i $\Rightarrow$ ii. 집합열 $\{A_j\}$를 $A_j = \prod_{i=1}^n (-\infty, x_0^i - 1/j]$로 두면 이는 $B_{x_0}^\circ$로 수렴하는 증가하는 집합열이므로 $F_X(x_0 - \mathbf{1}/j) = \mathbf{P}_X(A_j) \uparrow \mathbf{P}_X(B_{x_0}^\circ)$이다. 한편, 가정으로부터 F_X가 x_0에서 연속이므로 $\mathbf{P}_X(A_j) = F_X(x_0 - \mathbf{1}/j) \uparrow F_X(x_0) = \mathbf{P}_X(B_{x_0})$가 되어 $F_X(x_0) = \mathbf{P}_X(B_{x_0}) = \mathbf{P}_X(B_{x_0}^\circ)$임을 안다.

ii $\Rightarrow$ iii. 이는 $\mathbf{P}_X(\partial B_{x_0}) = \mathbf{P}_X(B_{x_0}) - \mathbf{P}_X(B_{x_0}^\circ) = 0$에서 자명하다.

iii $\Rightarrow$ i. 0으로 수렴하는 $\mathbb{R}^n$에 속하는 임의의 수열 $\{h_j\}$를 택하여 각 $j \in \mathbb{N}$에 대해 $a_j = \min_{i=1}^n h_j^i$, $b_j = \max_{i=1}^n h_j^i$라 하면 $F_X(x_0 + a_j\mathbf{1}) \leq F_X(x_0 + h_j) \leq F_X(x_0 + b_j\mathbf{1})$이고 $a_j, b_j \to 0$이다. 따라서 함수 $f : \mathbb{R} \to [0, 1]$를 $f : x \mapsto F_X(x_0 + x\mathbf{1})$로 두고 이가 0에서 연속임을 보이는 것으로 증명은 충분한데, CDF의 성질로부터 $\lim_{x \downarrow 0} f(x) = f(0)$임은 이미 알고 있으므로 $\lim_{x \uparrow 0} f(x) = f(0)$이라는 사실만 보이면 된다. 이를 위해 $x_j \uparrow 0$인 임의의 실수열 $\{x_j\}$를 택하여 집합열 $\{A_j\}$를 $A_j = \prod_{i=1}^n (-\infty, x_0^i + x_j^i]$로 두면 이는 $B_{x_0}^\circ$로 수렴하는 증가하는 집합열이다. 그렇다면 가정으로부터 $f(x_j) = F_X(x_0 + x_j\mathbf{1}) = \mathbf{P}_X(A_j) \uparrow \mathbf{P}_X(B_{x_0}^\circ) = \mathbf{P}_X(B_{x_0}) = F_X(x_0) = f(0)$이고, 증명이 끝난다. □

Lemma 3.31 단조함수 $f : \mathbb{R} \to \mathbb{R}$는 가산개의 불연속점만을 가진다.

Proof 우선 f가 증가함수라 하고 집합 $A = \{x \in \mathbb{R} : f$가 x에서 불연속$\}$를 생각하자. 그렇다면 임의의 $x \in A$에 대해 $f(x-) < f(x+)$이므로 $f(x-) < p_x < f(x+)$인 적당한 $p_x \in \mathbb{Q}$를 택할 수 있고, 이로써 함수 $g : A \to \mathbb{Q}$를 $g : x \mapsto p_x$로 두자. 한편, 임의의 $x, y \in A$에 대해 $x < y$인데 $f(y-) < f(x+)$이면 $z = (x+y)/2$에 대해 $f(z) \leq f(y-) < f(x+) \leq f(z)$의 모순이 발생하므로 $(f(x-), f(x+))$와 $(f(y-), f(y+))$는 서로소가 되어 g는 단사이고, 곧 A는 가산이다. □

Theorem 3.32 확률공간 $(\Omega, \mathscr{F}, \mathbb{P})$에서 정의된 rv. $X : \Omega \to \mathbb{R}^n$에 대해 F_X의 연속점의 집합은 조밀하다. 나아가, $n = 1$인 경우 F_X가 가산개의 불연속점만을 가진다.

Proof 임의의 $x_0 \in \mathbb{R}^n$를 택하여 함수 $f : \mathbb{R} \to [0, 1]$를 $f : h \mapsto F_X(x_0 + h\mathbf{1})$로 두자. 그렇다면 f는 증가함수가 되어 위의 보조정리로부터 가산개의 불연속점만을 가지고, 곧 0으로 수렴하면서 각 점에서 f가 연속인 실수열 $\{h_i\}$를 적당히 택할 수 있다. 이는 각 $i \in \mathbb{N}$에 대해 $\mathbf{P}_X(\prod_{i=1}^n(-\infty, x_0^i + h_i)) = F_X((x_0 + h_i\mathbf{1})-) = f(h_i-) = f(h_i) = F_X(x_0 + h_i\mathbf{1})$임을 함의하므로 정리 3. 30로부터 F_X는 $x_0 + h_i\mathbf{1}$에서 연속이고, $x_0 + h_i\mathbf{1} \to x_0$임은 자명하므로 F_X의 연속점의 집합이 조밀함을 안다. 한편, $n = 1$인 경우에는 F_X가 증가함수이므로 다시 위의 보조정리로부터 가산개의 불연속점만을 가짐이 자명하다. □

이제 PDF를 도입하는 것으로 이번 절을 마무리하자. 측도론으로 PDF를 도입하는데 핵심적인 역할을 하는 것은 다름아닌 Radon-Nikodým 도함수의 개념이다.

Definition 3.33 확률공간 $(\Omega, \mathscr{F}, \mathbb{P})$에서 정의된 rv. $X : \Omega \to \mathbb{R}^n$에 대해 적당한 $\mathscr{B}_n$ 위의 $\sigma-$유한 측도 μ가 존재하여 $\mathbf{P}_X \ll \mu$라 하자. 이때 Radon-Nikodým 도함수 $d\mathbf{P}_X/d\mu$를 X의 μ에 대한 **(확률)밀도함수((probability) density function)**라 하고 f_X로 쓴다. 특별히, $n \geq 2$인 경우 f_X를 rv. $X_1, \cdots, X_n$의 **결합(확률)밀도함수(joint (probability) density function)**라 하기도 한다.

이렇게 Radon-Nikodým 도함수로 PDF를 정의하면서 이번에도 우리가 PDF에 대해 당연하게 생각하던 사실들에 미묘한 혼란이 생겨나게 된다. 우선, Radon-Nikodým 도함수가 유일하기는 하지만 그 유일성이 '$\mu-$거의 어디서나 같은 함수를 하나로 볼 때' 성립하는 것이므로 엄밀히는 유일하지 않다. 즉, $\mu-$영집합에서 조금씩 다른 무한히 많은 함수들이 모두 하나의 확률벡터의 PDF가 될 수 있으므로 일반적으로 특정 점에서의 PDF의 값을 물어보는 것은 의미가 없다. 그러나 Radon-Nikodým 도함수를 추상적인 도구로만 사용했던 측도론에서와 달리 확률론에서는 PDF를 구체적인 함수로 다루어야 할 필요가 있으므로 혼란을 피하기 위해 PDF가 될 수 있는 무한히 많은, $\mu-$거의 어디서나 같은 함수 중 특정한 하나를 PDF의 **version**이라 한다. 물론, 표기상의 편의를 이유로 논의의 대상이 PDF인지, PDF의 version인지를 명시적으로 밝히지 않는 경우가 대부분이지만, L^p 공간에서 동치류

로서의 f와 구체적인 함수 f를 논의의 맥락으로 큰 혼란 없이 구분하였듯이 이 둘 또한 구분에 큰 어려움은 없을 것이다.

다음으로, 지금까지는 연속확률벡터와 이산확률벡터에 대해서만 PDF를 생각했지만 위의 정의에서 볼 수 있듯이 임의의 확률벡터에 대해서도 $\mathbf{P}_X \ll \mu$인 $\sigma-$유한 측도 μ만 잘 잡아주면 얼마든지 X의 PDF를 생각할 수 있다. 한편, 연속확률벡터, 이산확률벡터, 특이확률벡터의 세 가지 확률벡터만 생각해도 충분하다는 사실에 놀라워했던 것이 불과 몇 페이지 전인데 굳이 이렇게 지나칠 정도로 일반적으로 PDF를 도입하는 것이 의아하게 느껴질 수 있다. 그러나 이는 mixed-type 확률벡터의 PDF를 직접 다루기 위함이 아니라 연속확률벡터와 이산확률벡터의 PDF를 동시에 다루기 위함이다. 고등학교나 교양 통계학에서 연속확률벡터와 이산확률벡터의 기댓값이나 분산 같은 성질들을 논할 때, 각각 경우를 나누어 전자는 적분으로, 후자는 합으로 접근했던 것을 기억할 것이다. 우리는 위에서 일반적으로 정의된 PDF와 측도론의 에필로그에서 소개한 셈측도를 이용해 드디어 이런 번거로움에서 벗어날 수 있다.

한편, 정의에서 볼 수 있듯이 $\mathbf{P}_X \ll \mu$인 $\sigma-$유한 측도 μ를 무엇으로 택하는지에 따라 그 PDF가 달라지므로 μ를 반드시 명시해 주어야 하는데, 우리가 주로 다룰 연속확률벡터와 이산확률벡터의 경우 이에 대한 관례가 있다. 먼저 연속확률벡터 $X : \Omega \to \mathbb{R}^n$의 경우 그 정의상 $\mathbf{P}_X \ll \mu_n$이 항상 성립하므로 특별한 언급이 없는 이상 연속확률벡터의 PDF는 μ_n에 대한 PDF로 생각한다. 이산확률벡터 $X : \Omega \to \mathbb{R}^n$의 경우는 조금 복잡한데, 우선 정의상 $\mathbf{P}_X$는 가산 지지집합 $A \in \mathscr{B}_n$를 가진다. 물론, 이때의 가산 지지집합은 유일하지 않지만 임의의 $x \in A$에 대해 $\mathbf{P}_X\{x\} > 0$이라는 조건을 추가하면 유일하게 주어진다. 그렇다면 $(\mathbf{P}_X, \mathscr{B}_n, \#)$에 대해 $\mathbf{P}_X \ll \#_A$임이 자명하고, 이때 $\#_A$가 $\sigma-$유한임도 자명하므로 특별한 언급이 없는 이상 이산확률벡터의 PDF는 $\#_A$에 대한 PDF로 생각한다. 나아가, 이산확률벡터의 경우 PDF의 version으로 특별한 언급이 없는 이상 $\mathbb{R}^n \setminus A$에서 0인 version을 택하는 관례도 있다. 이런 관례를 따르면 이때의 PDF가 우리가 기존에 알던 PMF가 된다는 것을 잠시 후에 알게 될 것이다.

이제 PDF의 기본적인 성질들을 보자. 이는 대부분 적분의 성질들로부터 거의 자명하게 얻어진다.

Theorem 3.34 확률공간 $(\Omega, \mathscr{F}, \mathbb{P})$에서 정의된 rv. $X : \Omega \to \mathbb{R}^n$에 대해 적당한 $\mathscr{B}_n$ 위의 $\sigma-$유한 측도 μ가 존재하여 $\mathbf{P}_X \ll \mu$라 하자. 그렇다면 μ에 대한 X의 PDF f_X에 대해 다음이 성립한다.

i. $f_X \geq 0$ ($\mu-$ae.).
ii. $\int_{\mathbb{R}^n} f_X \, d\mu = 1$.
iii. 임의의 $A \in \mathscr{B}_n$에 대해 $\int_A f_X \, d\mu = \mathbf{P}_X(A) = \mathbb{P}\{X \in A\}$이다.

Proof iii. 이는 $\int_A f_X\,d\mu = \int_A (d\mathbf{P}_X/d\mu)\,d\mu = \int_A d\mathbf{P}_X = \mathbf{P}_X(A)$에서 자명하다.

i. 만약 $f_X \geq 0$ (μ−ae.)가 아니라면 집합 $A = \{x \in \mathbb{R}^n : -f_X(x) > 0\}$이 양의 측도를 가지므로 iii과 따름정리 2. 102의 iii으로부터 $\mathbf{P}_X(A) = \int_A f_X\,d\mu < 0$의 모순이 발생한다. 따라서 $f_X \geq 0$ (μ−ae.)이어야 한다.

ii. iii으로부터 $\int_{\mathbb{R}^n} f_X\,d\mu = \mathbf{P}_X(\mathbb{R}^n) = 1$이므로 자명하다. □

특별히, 이산확률벡터의 경우 위의 정리는 다음 따름정리와 같이 합의 형태로 표현된다.

Corollary 3.35 확률공간 $(\Omega, \mathscr{F}, \mathbb{P})$에서 정의된 이산확률벡터 $X : \Omega \to \mathbb{R}^n$에 대해 $A \in \mathscr{B}_n$를 $\mathbf{P}_X$의 가산 지지집합이라 하고 임의의 $x \in A$에 대해 $\mathbf{P}_X\{x\} > 0$이라 하면 다음이 성립한다.

i. 임의의 $x \in \mathbb{R}^n$에 대해 $f_X(x) = \mathbf{P}_X\{x\}$이고, 따라서 $0 \leq f_X \leq 1$이다.
ii. $\sum_{x\in\mathbb{R}^n} f_X(x) = 1$.
iii. 임의의 $B \in \mathscr{B}_n$에 대해 $\sum_{x\in B} f_X(x) = \mathbf{P}_X(B) = \mathbb{P}\{X \in B\}$이다.

Proof 표기의 편의를 위해 $\mathscr{A} = \mathscr{B}_n|_A$라 하면 $\#_A|_{\mathscr{A}} = \#|_{\mathscr{A}}$는 가측공간 $(A, \mathscr{A})$ 위의 셈측도이다.

i. 만약 $x \in A$이면 정리 2. 151의 iii으로부터 $\mathbf{P}_X\{x\} = \int_{\{x\}} f_X\,d\#_A = \int_{\{x\}} f_X|_A\,d\#|_{\mathscr{A}} = f_X(x)$이고 $x \notin A$이면 f_X의 version을 택하는 관례로부터 $\mathbf{P}_X\{x\} = 0 = f_X(x)$이다.

ii. 위의 정리와 셈측도의 성질 그리고 f_X의 version을 택하는 관례와 정리 2. 151의 iii으로부터 $1 = \mathbf{P}_X(A) = \int_A f_X\,d\#_A = \int_A f_X|_A\,d\#|_{\mathscr{A}} = \sum_{x\in A} f_X(x) = \sum_{x\in\mathbb{R}^n} f_X(x)$이다.

iii. 위의 정리와 셈측도의 성질 그리고 f_X의 version을 택하는 관례와 정리 2. 151의 iii으로부터 $\mathbf{P}_X(B) = \mathbf{P}_X(A\cap B) = \int_{A\cap B} f_X\,d\#_A = \int_{A\cap B} f_X|_A\,d\#|_{\mathscr{A}} = \sum_{x\in A\cap B} f_X(x) = \sum_{x\in B} f_X(x)$이다. □

다음 정리는 연속확률벡터의 CDF와 PDF의 관계에 대한 정리로서 연속확률벡터의 CDF와 PDF를 계산하는 데 유용하게 사용된다.

Theorem 3.36 확률공간 $(\Omega, \mathscr{F}, \mathbb{P})$에서 정의된 연속확률벡터 $X : \Omega \to \mathbb{R}^n$에 대해 다음이 성립한다.

i. $f_X(x) = F_X^{(s)}(x)$.
ii. $F_X(x) = \int_{\prod_{i=1}^n (-\infty, x_i]} f_X\,d\mu_n$.

특별히, $x_0 \in \mathbb{R}^n$의 근방에서 $\partial^n F_X/\partial x_1 \cdots \partial x_n$이 존재하고 이가 x_0에서 연속이면 $f_X(x_0) = (\partial^n F_X/\partial x_1 \cdots \partial x_n)(x_0)$이다. 이때, 편미분의 순서는 중요하지 않다.

Proof 이는 따름정리 2. 198와 정리 3. 34의 iii으로부터 자명하다 □

비슷하게, 다음 정리는 어떤 확률벡터가 연속인지를 판단할 때 유용하다.

Theorem 3.37 확률공간 $(\Omega, \mathscr{F}, \mathbb{P})$에서 정의된 rv. $X : \Omega \to \mathbb{R}^n$에 대해 TFAE.

i. Rv. X는 연속확률벡터이다.
ii. CDF F_X가 절대연속이다.
iii. 적분가능한 Borel 함수 $f : \mathbb{R}^n \to \overline{\mathbb{R}}$가 존재하여 $F_X(x) = \int_{\prod_{i=1}^n (-\infty, x_i]} f\, d\mu_n$이다.

Proof i $\Leftrightarrow$ ii. 이는 정리 2. 201로부터 자명하다.

i $\Rightarrow$ iii. 이는 $f = f_X$로 두고 정리 3. 36의 ii를 생각하면 자명하다.

iii $\Rightarrow$ ii. 함수 f가 적분가능하므로 거의 어디서나 유한하고, 따라서 WLOG, 필요다하면 영집합에서의 f의 값을 0으로 바꾸어 $|f| < \infty$라 하자. 이제 집합열 $\{A_j\}$를 $A_j = |f|^{-1}((j, \infty))$로 두면 이는 $\emptyset$으로 수렴하는 감소하는 집합열이 되어 $\int_{A_j} |f|\, d\lambda_n \to \int_{\emptyset} |f|\, d\lambda_n = 0$이다. 이로부터 임의의 $\varepsilon > 0$을 택하면 적당한 $j_0 \in \mathbb{N}$가 존재하여 $\int_{A_{j_0}} |f|\, d\lambda_n < \varepsilon/2$이고, $\mu_n(A) < \varepsilon/2j_0 =: \delta$인 임의의 $A \in \mathscr{B}_n$에 대해 $\int_A |f|\, d\mu_n \leq \int_{A \setminus A_{j_0}} |f|\, d\mu_n + \int_{A_{j_0}} |f|\, d\mu_n < \mu_n(A \setminus A_{j_0}) j_0 + \varepsilon/2 < \varepsilon$이다. 따라서 임의의 유계이고 서로소인 $B_1, \cdots, B_l \in \mathscr{S}_n$에 대해 $\sum_{k=1}^l \mu_n(B_k) < \delta$이면 $\sum_{k=1}^l \Delta_{B_k} F_X = \sum_{k=1}^l \int_{B_k} f\, d\mu_n = \int_{\bigsqcup_{k=1}^l B_k} f\, d\mu < \varepsilon$이 되어 F_X가 절대연속임을 안다. □

3.3 Expectation

이번 절에서는 확률변수의 대표적인 통계량 중 하나인 기댓값과 분산을 엄밀하게 도입하고, 이와 관련된 부등식을 증명하는 것을 목표로 한다. 통계학의 큰 연구주제 중 하나는 '정보의 단순화'라 할 것이다. 당연히, 이때 단순화의 정도와 이에 따른 정보의 손실은 서로 trade-off의 관계에 있어서 단순화를 많이 하면 할수록 정보의 전달이 용이하지만 그만큼 손실되는 정보도 많아지고, 역으로 손실되는 정보의 양을 줄이면 줄일수록 단순화의 정도가 감소하여 전달이 어려워진다. 이런 상황에서 우리는 정보의 손실을 최소한으로 유지하면서도 정보를 단순하게 전달할 수 있는 좋은 방법을 찾고자 하는데, 이번 절에서 알아볼 기댓값은 어떤 분포의 정보를 단 하나의 통계량으로 단순화하여 전달할 수 있는 좋은 방법 중 하나이다. (이와 같이 분포의 정보를 요약하는 통계량을 그 분포의 대푯값이라 하며 기댓값 외에도 중앙값, 최빈값 등이 있다. 각 대푯값은 나름의 장단점이 있기에 어떤 상황에서 무엇을 써야 하는지를 잘 판단하여 사용해야 한다.)

기댓값은 확률변수가 가질 것으로 가장 기대할 수 있는 값이라 해석할 수 있다. 기댓값은 분포의 여러 대푯값 중에서 가장 널리 이용되는 통계량으로 추정이나 관련 가설의 검정 등에 있어 많은 장점을 지니지만, outlier와 같은 극단치에 민감하여 때로는 분포에 대해 다

소 오인의 가능성이 있는 정보를 전달할 수도 있다는 단점을 가진다. 이러한 기댓값은 기본적으로 확률변수의 적분으로 정의된다.

Definition 3.38 확률공간 $(\Omega, \mathscr{F}, \mathbb{P})$에서 정의된 rv. $X : \Omega \to \mathbb{R}$에 대해 만약 X가 적분가능하다면 $\int_\Omega X\, d\mathbb{P}$를 X의 **기댓값(expectation)** 혹은 **평균(mean)**이라 하고 $\mathbf{E}(X)$, μ_X 혹은 간단히 μ로 쓴다.

기댓값이 정의되기 위해서는 확률공간과 그 위에서 정의된 rv.만 있으면 충분하다는 점에 주목하기 바란다. 고등학교나 교양 통계학에서는 기댓값이 정의되기 위해서는 확률변수가 반드시 연속이거나 이산이어서 PDF나 PMF가 있어야 했다. 그러나 위의 정의는 일반적인 mixed-type 확률변수에 대해서도 기댓값을 정의할 수 있도록 해준다. 한편, 정의에서 볼 수 있듯이 기댓값은 확률변수의 적분에 불과하므로 이에 관한 기본적인 성질들은 대부분 적분의 성질들로부터 자명하게 유도된다.

Proposition 3.39 확률공간 $(\Omega, \mathscr{F}, \mathbb{P})$에서 정의된 rv. $X : \Omega \to \mathbb{R}$에 대해 $\mathbf{E}(X)$가 존재할 필요충분조건은 $\mathbf{E}(|X|)$가 존재하는 것이다.

Proof 이는 X가 적분가능할 필요충분조건이 $|X|$가 적분가능할 것이라는 점에서 자명하다. □

Theorem 3.40 확률공간 $(\Omega, \mathscr{F}, \mathbb{P})$에서 정의된 rv. $X, Y : \Omega \to \mathbb{R}$에 대해 $\mathbf{E}(X), \mathbf{E}(Y)$가 모두 존재한다면 다음이 성립한다.

i. (선형성) 임의의 $a, b \in \mathbb{R}$에 대해 $\mathbf{E}(aX + bY)$가 존재하고 $\mathbf{E}(aX + bY) = a\mathbf{E}(X) + b\mathbf{E}(Y)$이다.
ii. 만약 $X = Y$ (as.)라면 $\mathbf{E}(X) = \mathbf{E}(Y)$이다.
iii. 만약 $X \leq Y$ (as.)라면 $\mathbf{E}(X) \leq \mathbf{E}(Y)$이다.
iv. $\mathbf{E}(1) = 1$.
v. $\inf_{\omega \in \Omega} X(\omega) \leq \mathbf{E}(X) \leq \sup_{\omega \in \Omega} X(\omega)$.
vi. $|\mathbf{E}(X)| \leq \mathbf{E}(|X|)$.

Proof i − iii, vi. 이는 적분의 성질로부터 자명하다.
iv. 이는 $\mathbf{E}(1) = \int_\Omega d\mathbb{P} = \mathbb{P}(\Omega) = 1$에서 분명하다.
v. 이는 iii과 iv로부터 자명하다. □

기댓값의 선형성으로부터 $X - \mu_X$의 기댓값은 항상 0이다. 이렇게 원래의 확률변수에서 그 기댓값을 빼는 것을 **centering**이라 하며, 기댓값이 서로다른 분포를 비교할 때에 자주 사용되는 방법이다. 한편, 다양한 수렴정리들도 동일하게 성립한다.

Theorem 3.41 (Monotone convergence theorem) 확률공간 $(\Omega, \mathscr{F}, \mathbb{P})$에서 정의된 음이 아닌 rv. $X_i : \Omega \to \mathbb{R}_0^+$의 열 $\{X_i\}$와 음이 아닌 rv. $X : \Omega \to \mathbb{R}_0^+$에 대해 $X_i \uparrow X$ (as.)이고 $\mathbf{E}(X)$가 존재한다고 하면 모든 $i \in \mathbb{N}$에 대해 $\mathbf{E}(X_i)$도 존재하고 $\mathbf{E}(X_i) \uparrow \mathbf{E}(X)$이다.

Proof 모든 $i \in \mathbb{N}$에 대해 $0 \le X_i \le X$이므로 $\mathbf{E}(X_i)$가 존재한다. 이제 정리는 측도론에서 배운 MCT에서 자명하다. □

Theorem 3.42 (Lebesgue's dominated convergence theorem) 확률공간 $(\Omega, \mathscr{F}, \mathbb{P})$에서 정의된 rv. $X_i : \Omega \to \mathbb{R}$의 열 $\{X_i\}$와 rv. $X : \Omega \to \mathbb{R}$에 대해 $X_i \to X$ (as.)라 하자. 또한 어떤 rv. $Y : \Omega \to \mathbb{R}_0^+$가 존재하여 $\mathbf{E}(Y)$가 존재하고 모든 $i \in \mathbb{N}$에 대해 $|X_i| \le Y$이면 $\mathbf{E}(X_i)$와 $\mathbf{E}(X)$가 모두 존재하고 $\mathbf{E}(X_i) \to \mathbf{E}(X)$이다.

Proof 이는 측도론에서 배운 DCT에서 자명하다. □

Corollary 3.43 (Bounded convergence theorem) 확률공간 $(\Omega, \mathscr{F}, \mathbb{P})$에서 정의된 rv. $X_i : \Omega \to \mathbb{R}$의 열 $\{X_i\}$와 rv. $X : \Omega \to \mathbb{R}$에 대해 $X_i \to X$ (as.)라 하자. 또한 $\{X_i\}$가 균등하게 유계라 하자. 즉, 어떤 $M > 0$이 존재하여 모든 $i \in \mathbb{N}$에 대해 $|X_i| \le M$이라 하자. 그렇다면 $\mathbf{E}(X_i)$와 $\mathbf{E}(X)$가 모두 존재하고 $\mathbf{E}(X_i) \to \mathbf{E}(X)$이다.

Proof 이는 측도론에서 배운 유계수렴정리로부터 자명하다. □

비록 기댓값의 정의는 충분히 일반적이지만, 이가 기댓값을 계산하는 구체적인 방법을 알려주지는 않는다. 결국 기댓값을 계산하기 위해서는 우리가 여태껏 해왔던 것처럼 PDF의 도움을 받아야 한다.

Theorem 3.44 확률공간 $(\Omega, \mathscr{F}, \mathbb{P})$에서 정의된 rv. $X : \Omega \to \mathbb{R}^n$와 Borel 함수 $g : \mathbb{R}^n \to \mathbb{R}$에 대해 $\mathbf{E}(g(X))$가 존재한다면(혹은 g가 $\mathbf{P}_X$−적분가능하다면) $\mathbf{E}(g(X)) = \int_{\mathbb{R}^n} g\, d\mathbf{P}_X$이고, 이때 g는 $\mathbf{P}_X$−적분가능하다(혹은 $\mathbf{E}(g(X))$가 존재한다).

Proof 기댓값 $\mathbf{E}(g(X))$가 존재한다면 $\mathbf{E}(g(X)) = \int_\Omega g \circ X\, d\mathbb{P} = \int_{\mathbb{R}^n} g\, d\mathbf{P}_X < \infty$이므로 이는 자명하다. 한편, $\mathbf{P}_X$−적분가능한 g에 대해서도 비슷하게 하면 된다. □

Theorem 3.45 확률공간 $(\Omega, \mathscr{F}, \mathbb{P})$에서 정의된 rv. $X : \Omega \to \mathbb{R}^n$와 Borel 함수 $g : \mathbb{R}^n \to \mathbb{R}$에 대해 적당한 $\mathscr{B}_n$ 위의 σ−유한 측도 μ가 존재하여 $\mathbf{P}_X \ll \mu$라 하자. 만약 $\mathbf{E}(g(X))$가 존재한다면(혹은 $f_X g$는 μ−적분가능하다면) $\mathbf{E}(g(X)) = \int_{\mathbb{R}^n} f_X g\, d\mu$이고, 이때 $f_X g$는 μ−적분가능하다(혹은 $\mathbf{E}(g(X))$가 존재한다).

Proof 기댓값 $\mathbf{E}(g(X))$가 존재한다면 $\mathbf{E}(g(X)) = \int_{\mathbb{R}^n} g\, d\mathbf{P}_X = \int_{\mathbb{R}^n} f_X g\, d\mu < \infty$이므로 이는 자명하다. 한편, $f_X g$가 μ−적분가능한 경우에 대해서도 비슷하게 하면 된다. □

특별히, 이산확률벡터의 경우 위의 정리는 다음 따름정리와 같이 합의 형태로 표현된다.

Corollary 3.46 확률공간 $(\Omega, \mathscr{F}, \mathbb{P})$에서 정의된 이산확률벡터 $X : \Omega \to \mathbb{R}^n$와 Borel 함수 $g : \mathbb{R}^n \to \mathbb{R}$에 대해 $\mathbf{E}(g(X))$가 존재한다면 $\mathbf{E}(g(X)) = \sum_{x \in \mathbb{R}^n} f_X(x) g(x)$이다.

Proof 가정으로부터 $\mathbf{P}_X$의 가산 지지집합 $A \in \mathscr{B}_n$가 존재하여 WLOG, 필요하다면 몇몇 원소를 제거하여 임의의 $x \in A$에 대해 $\mathbf{P}_X\{x\} > 0$이라 해도 된다. 표기의 편의를 위해 $\mathscr{A} = \mathscr{B}_n|_A$라 하면 $\#_A|_{\mathscr{A}} = \#|_{\mathscr{A}}$는 가측공간 $(A, \mathscr{A})$ 위의 셈측도이므로 위의 정리와 f_X의 version을 택하는 관례 그리고 정리 2. 151의 iii으로부터 $\mathbf{E}(g(X)) = \int_{\mathbb{R}^n} f_X g \, d\#_A = \int_A f_X g \, d\#_A = \int_A (f_X g)|_A \, d\#|_{\mathscr{A}} = \sum_{x \in A} f_X(x) g(x) = \sum_{x \in \mathbb{R}^n} f_X(x) g(x)$이다. □

만약 확률변수가 음이 아니라면 CDF를 사용하여 기댓값을 계산할 수도 있다.

Theorem 3.47 확률공간 $(\Omega, \mathscr{F}, \mathbb{P})$에서 정의된 음이 아닌 rv. $X : \Omega \to \mathbb{R}_0^+$에 대해 $\mathbf{E}(X)$가 존재하면 $1 - F_X$는 $\mathbb{R}_0^+$에서 μ_1-적분가능하고 $\mathbf{E}(X) = \int_{\mathbb{R}_0^+} 1 - F_X \, d\mu_1$이다.

Proof 함수 $\mathbf{1}_{\{0 \le x < y\}}(x, y)$가 음이 아닌 Borel이므로 Fubini의 정리로부터

$$\begin{aligned}
\mathbf{E}(X) &= \int_{\mathbb{R}} x \, d\mathbf{P}_X(x) \\
&= \int_{\mathbb{R}_0^+} x \, d\mathbf{P}_X(x) \\
&= \int_{\mathbb{R}_0^+} \left(\int_{[0,x]} d\mu_1 \right) d\mathbf{P}_X(x) \\
&= \int_{\mathbb{R}_0^+} \left(\int_{\mathbb{R}_0^+} \mathbf{1}_{[0,x)}(y) \, d\mu_1(y) \right) d\mathbf{P}_X(x) \\
&= \int_{\mathbb{R}_0^+} \left(\int_{\mathbb{R}_0^+} \mathbf{1}_{[0,x)}(y) \, d\mathbf{P}_X(x) \right) d\mu_1(y) \\
&= \int_{\mathbb{R}_0^+} \left(\int_{\mathbb{R}_0^+} \mathbf{1}_{(y,\infty)}(x) \, d\mathbf{P}_X(x) \right) d\mu_1(y) \\
&= \int_{\mathbb{R}_0^+} \mathbf{P}_X((y, \infty)) \, d\mu_1(y) \\
&= \int_{\mathbb{R}_0^+} 1 - F_X(y) \, d\mu_1(y)
\end{aligned}$$

이고 이때 $1 - F_X$는 $\mathbb{R}_0^+$에서 μ_1-적분가능하다. □

다음으로 넘어가기 전에, 기댓값에 대해 한 가지 유념해야 할 점이 있다. 기댓값의 정의에는 이가 마치 어떤 확률변수의 특성인 것처럼 쓰여 있지만 사실 기댓값은 확률변수의 특성이 아닌, 분포의 특성이어서 서로 다른 확률공간에서 정의된 서로다른 확률변수라도 그 분포가 같다면 기댓값도 같다. 실제로 확률공간 $(\Omega, \mathscr{F}, \mathbb{P})$와 $(\Omega', \mathscr{F}', \mathbb{P}')$에서

각각 정의된 확률변수 $X:\Omega\to\mathbb{R}, Y:\Omega'\to\mathbb{R}$에 대해 만약 $X\equiv Y$라면 정리 3.44로부터 $\mathbf{E}(X)=\int_{\mathbb{R}^n} x\,d\mathbf{P}_X(x)=\mathbf{E}(Y)$이다. 이러한 사실은 기댓값이 분포의 정보를 요약하여 전달하는 통계량이라는 점에서 어떻게 보면 당연한 것이다.

기댓값에 이어 살펴볼 통계량은 분산이다. 기댓값이 분포의 정보를 요약하여 전달하는 좋은 통계량인 것은 사실이지만, 이가 분포의 정보를 완벽하게 전달하지는 못한다. 이때 발생하는 정보의 손실이 크게 중요하지 않은 경우도 있지만, 때로는 기댓값만으로는 부족한 경우도 있다. 이에 기댓값만을 전달할 때 손실되는 정보를 보충할 목적으로 보조적인 통계량을 제공할 수 있는데, 보통 분포가 기댓값으로부터 퍼져있는 정도를 의미하는 분산이나 표준편차를 제공한다. (이와 같이 분포가 퍼져있는 정도를 의미하는 통계량을 그 분포의 산포도라 하며 분산과 표준편차 이외에도 평균절대편차(MAD; Mean Absolute Deviation), 사분위수범위(IQR; InterQuartile Range) 등이 있다.) 이러한 분산은 기본적으로 기댓값으로 정의된다.

Definition 3.48 확률공간 $(\Omega,\mathscr{F},\mathbb{P})$에서 정의된 rv. $X:\Omega\to\mathbb{R}$에 대해 만약 $\mathbf{E}((X-\mu_X)^2)$이 존재한다면 이를 X의 **분산(variance)**이라 하고 $\mathbf{Var}(X), \sigma_X^2$ 혹은 간단히 σ^2으로 쓴다. 나아가, $\mathbf{Var}(X)$가 존재한다면 $\sqrt{\mathbf{Var}(X)}$를 X의 **표준편차(standard deviation)**라 하고 $\mathbf{Sd}(X), \sigma_X$ 혹은 간단히 σ로 쓴다.

Proposition 3.49 확률공간 $(\Omega,\mathscr{F},\mathbb{P})$에서 정의된 rv. $X:\Omega\to\mathbb{R}$에 대해 $\mathbf{Var}(X)$가 존재할 필요충분조건은 $\mathbf{E}(X^2)$이 존재하는 것이다.

Proof 만약 $\mathbf{Var}(X)$가 존재한다면 정의로부터 $\mathbf{E}(X)$와 $\mathbf{E}((X-\mu_X)^2)$가 존재하고, 곧 $X^2=(X-\mu_X)^2+2\mu_X X-\mu_X^2$에서 $\mathbf{E}(X^2)$이 존재함을 안다. 역으로 $\mathbf{E}(X^2)$이 존재한다면 $|X|\le X^2\mathbf{1}_{\{|X|\ge 1\}}+\mathbf{1}_{\{|X|<1\}}\le X^2+1$에서 $\mathbf{E}(X)$가 존재함을 알고, 곧 $(X-\mu_X)^2=X^2-2\mu_X X+\mu_X^2$에서 $\mathbf{Var}(X)=\mathbf{E}((X-\mu_X)^2)$이 존재한다. □

분산의 기본적인 성질들은 모두 기댓값의 성질들에 기인하다. 특히 다음 정리의 i는 간단한 공식이지만 분산을 실제로 계산할 때 유용하게 사용되는 공식이다.

Theorem 3.50 확률공간 $(\Omega,\mathscr{F},\mathbb{P})$에서 정의된 rv. $X:\Omega\to\mathbb{R}$에 대해 $\mathbf{Var}(X)$가 존재한다면 다음이 성립한다.

i. $\mathbf{Var}(X)=\mathbf{E}(X^2)-\mu_X^2$.
ii. 임의의 $a,b\in\mathbb{R}$에 대해 $\mathbf{Var}(aX+b)$가 존재하고 $\mathbf{Var}(aX+b)=a^2\mathbf{Var}(X)$이다.

Proof i. 이는

$$\begin{aligned}\mathbf{Var}(X)&=\mathbf{E}((X-\mu_X)^2)\\&=\mathbf{E}(X^2-2\mu_X X+\mu_X^2)\end{aligned}$$

$$
\begin{aligned}
&= \mathbf{E}(X^2) - 2\mu_X \mathbf{E}(X) + \mu_X^2 \\
&= \mathbf{E}(X^2) - \mu_X^2
\end{aligned}
$$

에서 자명하다.

ii. 우선 $\mathbf{Var}(X)$가 존재하므로 $\mathbf{E}(X)$와 $\mathbf{E}(X^2)$이 존재하여 $\mathbf{E}((aX+b)^2) = \mathbf{E}(a^2X^2 + 2abX + b^2)$가 존재하여 $\mathbf{Var}(aX+b)$도 존재한다. 이제

$$
\begin{aligned}
\mathbf{Var}(aX+b) &= \mathbf{E}([aX+b-\mathbf{E}(aX+b)]^2) \\
&= \mathbf{E}(a^2[X-\mathbf{E}(X)]^2) \\
&= a^2\mathbf{E}((X-\mu_X)^2) \\
&= a^2\mathbf{Var}(X)
\end{aligned}
$$

에서 정리는 자명하다. □

기댓값과 분산의 성질로부터 $\mathbf{Var}(X)$가 존재하고 0이 아니라면 $(X-\mu_X)/\sigma_X$의 기댓값과 분산은 항상 각각 $0, 1$이다. 이렇게 원래의 확률변수에서 그 기댓값을 빼고, 그 표준편차로 나누는 것을 **표준화(standardization)**라 하며, 기댓값과 분산이 서로다른 분포를 비교할 때 자주 사용되는 방법이다.

한편, 기댓값과 분산은 각각 다음과 같이 $\mathbf{E}((X-x)^2)$가 최솟값을 갖도록 하는 값과 그 최솟값으로 정의할 수도 있다.

Theorem 3.51 확률공간 $(\Omega, \mathscr{F}, \mathbb{P})$에서 정의된 rv. $X : \Omega \to \mathbb{R}$에 대해 $\mathbf{Var}(X)$가 존재하면 함수 $x \mapsto \mathbf{E}((X-x)^2)$는 well-define되고 $\mathbf{E}(X)$에서 최솟값 $\mathbf{Var}(X)$를 가진다.

Proof 가정으로부터 $\mathbf{E}(X)$와 $\mathbf{E}(X^2)$이 존재하고 곧 임의의 $x \in \mathbb{R}$에 대해 $(X-x)^2 = X^2 - 2xX + x^2$에서 $\mathbf{E}((X-x)^2)$이 존재하므로 함수 $x \mapsto \mathbf{E}((X-x)^2)$가 well-define된다. 한편, 임의의 $x \in \mathbb{R}$에 대해 $(X-x)^2 = (X-\mu_X)^2 + 2(X-\mu_X)(\mu_X - x) + (\mu_X - x)^2$이므로 $\mathbf{E}((X-x)^2) = \mathbf{E}((X-\mu_X)^2) + 2(\mu_X - x)\mathbf{E}(X-\mu_X) + (\mu_X - x)^2 = \mathbf{Var}(X) + (\mu_X - x)^2$에서 정리가 성립한다. □

이제 기댓값에 관한 다양한 부등식을 알아보는 것으로 이번 절을 마무리하자. 이러한 부등식은 이후 확률론의 다양한 정리를 증명할 때에 약방의 감초와 같이 사용될 것이다.

Theorem 3.52 (Jensen's inequality) 확률공간 $(\Omega, \mathscr{F}, \mathbb{P})$에서 정의된 rv. $X, Y : \Omega \to \mathbb{R}$와 함수 $g : \mathbb{R} \to \mathbb{R}$에 대해 적당한 구간 $I \subseteq \mathbb{R}$가 존재하여 g가 I에서 볼록하고(혹은 오목하고) $X \in I$ (as.)라 하자. 만약 $\mathbf{E}(X)$와 $\mathbf{E}(g(X))$가 존재한다면 $g(\mathbf{E}(X)) \leq \mathbf{E}(g(X))$이다(혹은 $g(\mathbf{E}(X)) \geq \mathbf{E}(g(X))$이다).

Proof 먼저 I가 열린구간인 경우를 생각하여 $I=(a,b)$로 두면 가정과 따름정리 2.102의 iii으로부터 $\mathbf{E}(X)\in I$이다. 한편, g가 I에서 볼록하므로 $(\mathbf{E}(X),g(\mathbf{E}(X)))$를 지나는 supporting line $l:y=mx+n$이 존재하여 $mX+n\le g(X)$ (as.)이고, 따라서 $g(\mathbf{E}(X))=m\mathbf{E}(X)+n=\mathbf{E}(mX+n)\le\mathbf{E}(g(X))$이다.

이제 I가 닫힌구간인 경우를 생각하여 $I=[a,b]$로 두면 가정으로부터 $\mathbf{E}(X)\in I$인데, 만약 $\mathbf{E}(X)\in I^\circ$라면 이전과 같이 하여 $g(\mathbf{E}(X))\le\mathbf{E}(g(X))$임을 보일 수 있으므로 $\mathbf{E}(X)$가 I의 양 끝점인 경우만 고려하면 된다. 간결한 논의를 위해, $\mathbf{E}(X)=b$라 하자. (반대로 $\mathbf{E}(X)=a$인 경우에도 이와 비슷하게 하면 된다.) 그렇다면 정리 2.101의 ii로부터 $X=b$ (as.)이므로 $g(X)=g(b)$ (as.)에서 $g(\mathbf{E}(X))=g(b)=\mathbf{E}(g(X))$이고, 증명이 끝난다.

한편, 구간 I가 한쪽 끝점만 포함하는 경우에도 I가 닫힌구간인 경우와 비슷하게 하면 같은 결론을 얻고, g가 I에서 오목한 경우에는 $-g$를 생각하면 되므로 증명은 이로써 충분하다. □

Theorem 3.53 (Hölder's inequality) 확률공간 $(\Omega,\mathscr{F},\mathbb{P})$에서 정의된 rv. $X,Y:\Omega\to\mathbb{R}$와 $1/p+1/q=1$인 $p,q>1$에 대해 $\mathbf{E}(|X|^p)$와 $\mathbf{E}(|Y|^q)$가 존재한다면 $\mathbf{E}(|XY|)$도 존재하고 $\mathbf{E}(|XY|)\le[\mathbf{E}(|X|^p)]^{1/p}[\mathbf{E}(|Y|^q)]^{1/q}$이다.

Proof 이는 측도론에서 배운 Hölder의 부등식으로부터 자명하다. □

Theorem 3.54 (Minkowski's inequality) 확률공간 $(\Omega,\mathscr{F},\mathbb{P})$에서 정의된 rv. $X,Y:\Omega\to\mathbb{R}$와 $p\ge1$에 대해 $\mathbf{E}(|X|^p)$와 $\mathbf{E}(|Y|^p)$가 존재한다면 $\mathbf{E}(|X+Y|^p)$도 존재하고 $[\mathbf{E}(|X+Y|^p)]^{1/p}\le[\mathbf{E}(|X|^p)]^{1/p}+[\mathbf{E}(|Y|^p)]^{1/p}$이다.

Proof 이는 측도론에서 배운 Minkowski의 부등식으로부터 자명하다. □

Corollary 3.55 (Cauchy-Schwarz's inequality) 확률공간 $(\Omega,\mathscr{F},\mathbb{P})$에서 정의된 rv. $X,Y:\Omega\to\mathbb{R}$에 대해 $\mathbf{E}(X^2)$과 $\mathbf{E}(Y^2)$이 존재한다면 $\mathbf{E}(|XY|)$도 존재하고 $\mathbf{E}(|XY|)\le\sqrt{\mathbf{E}(X^2)\mathbf{E}(Y^2)}$이다.

Proof 이는 위의 Hölder의 부등식에서 $p=q=2$인 특수한 경우이다. □

Theorem 3.56 (Liapounov's inequality) 확률공간 $(\Omega,\mathscr{F},\mathbb{P})$에서 정의된 rv. $X:\Omega\to\mathbb{R}$와 $0<p<q$인 $p,q\in\mathbb{R}$에 대해 $\mathbf{E}(|X|^q)$가 존재한다면 $\mathbf{E}(|X|^p)$도 존재하고 $[\mathbf{E}(|X|^p)]^{1/p}\le[\mathbf{E}(|X|^q)]^{1/q}$이다.

Proof 우선 $|X|^p\le|X|^q\mathbf{1}_{\{|X|\ge1\}}+\mathbf{1}_{\{|X|<1\}}\le|X|^q+1$에서 $\mathbf{E}(|X|^p)$이 존재함을 안다. 이제 함수 $g:\mathbb{R}\to\mathbb{R}$를 $g:x\mapsto x^{q/p}\mathbf{1}_{\mathbb{R}^+}(x)$로 두면 이는 볼록함수이므로 Jensen의 부등식으로부터 $[\mathbf{E}(|X|^p)]^{q/p}=g(\mathbf{E}(|X|^p))\le\mathbf{E}(g(|X|^p))=\mathbf{E}(|X|^q)$이고, 곧 증명이 끝난다. □

아래의 두 부등식은 증명의 도구로 쓰일 뿐만 아니라 어떤 확률분포가 특정 값을 얼마 이상 벗어날 확률에 대한 상계가 된다. 이를 두고 아래의 두 부등식이 ‘tail probability를 bounding한다’ 고 한다.

Theorem 3.57 (Markov's inequality) 확률공간 $(\Omega, \mathscr{F}, \mathbb{P})$에서 정의된 rv. $X : \Omega \to \mathbb{R}$와 $p > 0$에 대해 $\mathbf{E}(|X|^p)$이 존재한다면 임의의 $x > 0$에 대해 $\mathbb{P}\{|X| \geq x\} \leq \mathbf{E}(|X|^p)/x^p$이다.

Proof 이는 측도론에서 배운 Markov의 부등식으로부터 자명하다. □

Corollary 3.58 (Chebyshev's inequality) 확률공간 $(\Omega, \mathscr{F}, \mathbb{P})$에서 정의된 rv. $X : \Omega \to \mathbb{R}$에 대해 $\mathbf{Var}(X)$가 존재한다면 임의의 $x > 0$에 대해 $\mathbb{P}\{|X - \mu_X| \geq x\} \leq \mathbf{Var}(X)/x^2$이다.

Proof 이는 rv. $X - \mu_X$에 $p = 2$인 경우의 Markov의 부등식을 적용한 결과이다. □

다음은 Chebyshev의 부등식의 기초적인 응용으로, 분산이 0이라는 것의 의미를 잘 나태나준다. 이 결과는 이후에 이따금씩 쓰일 것이다.

Proposition 3.59 확률공간 $(\Omega, \mathscr{F}, \mathbb{P})$에서 정의된 rv. $X : \Omega \to \mathbb{R}$에 대해 $\mathbf{Var}(X) = 0$이면 $\mathbf{E}(X)$가 존재하고 $X = \mathbf{E}(X)$ (as.)이다.

Proof 우선 $\mathbf{E}(X)$가 존재함은 자명하다. 이제 사건열 $\{E_i\}$를 $E_i = \{|X - \mu_X| > 1/i\}$로 두면 이는 $\{X \neq \mu_X\}$로 수렴하는 사건열로 $\mathbb{P}\{X = \mu_X\} = 1 - \mathbb{P}\{X \neq \mu_X\} = 1 - \lim_{i\to\infty} \mathbb{P}(E_i)$이다. 한편, 각 $i \in \mathbb{N}$에 대해 Chebyshev의 부등식으로부터 $\mathbb{P}(E_i) \leq \mathbf{Var}(X)/i^2 = 0$이므로 $\mathbb{P}(E_i) \to 0$에서 $\mathbb{P}\{X = \mu_X\} = 1$임을 안다. □

3.4 Moments

이번 절에서는 적률이라는 통계량에 대해 알아본다. 앞서 어떤 분포에 대한 정보를 간단히 전달하기 위해 기댓값이나 분산을 사용하였듯이 적률 또한 어떤 분포의 정보를 전달하기 위해 사용한다. 당장에 기댓값과 분산도 적률의 일종이다. 형식적으로 k차 적률은 확률변수 X의 k제곱의 기댓값으로 정의되는데, 기댓값이나 분산이 그 자체로 나름의 의미를 가지는 것과 달리 k차 적률 그 자체의 의미를 해석하는 것은 쉽지 않다. 그럼에도 불구하고, 하나의 절을 할애하여 적률에 대해 논하는 까닭은 많은 경우에 적률은 분포의 정보를 ‘손실없이’ 전달할 수 있기 때문이다. 즉, 모든 차수의 적률을 수열의 형태로 전달하면 그것만으로 분포를 완벽하게 복원해낼 수 있고, 정확히 언제 이가 가능한지를 밝히고, 증명하는 것이 이번 절의 주된 목표이다. 그렇다면 적률의 정의부터 시작하자.

Definition 3.60 확률공간 $(\Omega, \mathscr{F}, \mathbb{P})$에서 정의된 rv. $X : \Omega \to \mathbb{R}$와 $k \in \mathbb{N}_0$에 대해 다음을 정의한다.

i. 만약 $\mathbf{E}(X^k)$이 존재한다면 이를 X의 k**차 적률**(**kth moment**)이라 하고 $\mu_{k,X}$ 혹은 간단히 μ_k으로 쓴다.
ii. 만약 $\mathbf{E}((X-\mu_X)^k)$이 존재한다면 이를 X의 k**차 중심화된 적률**(**nth centeral moment**)이라 하고 $\tau_{k,X}$ 혹은 간단히 τ_k으로 쓴다.
iii. 만약 $\mathbf{E}([(X-\mu_X)/\sigma_X]^k)$이 존재한다면 이를 X의 k**차 표준화된 적률**(**kth standardized moment**)이라 하고 $\kappa_{k,X}$ 혹은 간단히 κ_k으로 쓴다.
iv. 만약 $\mathbf{E}(X^{\underline{k}})$이 존재한다면 이를 X의 k**차 (하향)계승적률**(**kth (falling) factorial moment**)이라 하고 $\mu_{\underline{k},X}$ 혹은 간단히 $\mu_{\underline{k}}$으로 쓴다.
v. 만약 $\mathbf{E}(X^{\overline{k}})$이 존재한다면 이를 X의 k**차 (상향)계승적률**(**nth (rising) factorial moment**)이라 하고 $\mu_{\overline{k},X}$ 혹은 간단히 $\mu_{\overline{k}}$으로 쓴다.

곧 기댓값은 1차 위에서 정의한 적률을 제외하고도 많은 종류의 적률이 존재하는데, 이는 계산상의 편의를 위한 목적이 크다. 나아가 중심화된 적률과 표준화된 적률은 그 이름에서도 알 수 있듯이 각각 원래 확률변수의 평행이동과 Affine 변환에 대해 불변이라는 유용한 특성을 가지기도 한다. 한편, 적률의 존재성을 따지는 데에는 다음 두 명제가 유용하다.

Proposition 3.61 확률공간 $(\Omega, \mathscr{F}, \mathbb{P})$에서 정의된 rv. $X : \Omega \to \mathbb{R}$에 대해 만약 X의 k차 적률(혹은 중심화된 적률, 표준화된 적률, 하향계승적률, 상향계승적률)이 존재하면 X의 $l \le k$차 적률(혹은 중심화된 적률, 표준화된 적률, 하향계승적률, 상향계승적률)이 존재한다.

Proof 적률의 경우 Liapounov의 부등식으로부터 자명하고, 이로부터 중심화된 적률과 표준화된 적률에 대해서도 명제가 자명해진다. 다음으로 하향계승적률의 경우, 만약 X의 $k > 1$차 하향계승적률이 존재한다면 $X^{\underline{k}} = X^{\underline{k-1}}(X-k+1)$에서

$$\begin{aligned}|X^{\underline{k-1}}| &= \left|\frac{X^{\underline{k}}}{X-k+1}\right| \mathbf{1}_{\{X \ne k-1\}} + (k-1)!\mathbf{1}_{\{X=k-1\}} \\ &\le |X^{\underline{k}}|\mathbf{1}_{\{|X-k+1|\ge 1\}} + |X^{\underline{k-1}}|\mathbf{1}_{\{|X-k+1|<1\}} + (k-1)! \\ &< |X^{\underline{k}}| + 2k!\end{aligned}$$

이다. 여기서 마지막 부등호는 함수 $x \mapsto |x^{\underline{k-1}}|$이 $(k-2,\infty)$에서 증가하므로 성립한다. 이로부터 X의 $k-1$차 하향계승적률이 존재함을 알고, 이를 $k-2$번 반복하면 모든 $l \le k$에 대해 X의 l차 하향계승적률이 존재함을 안다. 마지막으로 상향계승적률의 경우에는 하향계승적률과 비슷하게 하면 된다. □

Proposition 3.62 확률공간 $(\Omega, \mathscr{F}, \mathbb{P})$에서 정의된 rv. $X : \Omega \to \mathbb{R}$에 대해 TFAE.

i. Rv. X의 k차 적률이 존재한다.
ii. 만약 $\mathbf{E}(X)$가 존재한다면 X의 k차 중심화된 적률이 존재한다.
iii. 만약 $\mathbf{Var}(X)$가 존재하고 0이 아니면 X의 k차 표준화된 적률이 존재한다.
iv. Rv. X의 k차 하향계승적률이 존재한다.
v. Rv. X의 k차 상향계승적률이 존재한다.

Proof 만약 $k=0$이면 명제가 자명하므로 $k>0$이라 하자.

i ⇔ ii. 만약 X의 k차 적률이 존재하면 명제 3.61로부터 X의 모든 $l \le k$차 적률이 존재한다. 이로부터

$$(X-\mu_X)^k = \sum_{i=0}^{k}(-1)^{k-i}\binom{k}{i}X^i\mu_X^{k-i}$$

에서 X의 k차 중심화된 적률이 존재한다. 역으로, $\mathbf{E}(X)$가 존재하고 X의 k차 중심화된 적률이 존재하면 다시 명제 3.61로부터 X의 모든 $l \le k$차 중심화된 적률이 존재한다. 그렇다면

$$\begin{aligned}X^k &= (X-\mu_X+\mu_X)^k \\ &= \sum_{i=0}^{k}\binom{k}{i}(X-\mu_X)^i\mu_X^{k-i}\end{aligned}$$

에서 X의 k차 적률이 존재한다.

ii ⇔ iii. 이는 정의로부터 자명하다.

i ⇒ iv. 명제 3.61로부터 X의 모든 $l \le k$차 적률이 존재한다. 한편, 적당한 $a_0, \cdots, a_k \in \mathbb{R}$에 대해 $X^{\underline{k}} = \prod_{i=0}^{k-1}(X-i) = \sum_{i=0}^{k} a_iX^i$이므로 X의 k차 하향계승적률이 존재한다.

iv ⇒ i. 명제 3.61로부터 X의 모든 $l \le k$차 하향계승적률이 존재한다. 이제 적당한 $a_0, \cdots, a_k \in \mathbb{R}$에 대해 $X^{\underline{k}} = \prod_{i=0}^{k-1}(X-i) = \sum_{i=0}^{k} a_iX^i$이고 이때 $a_k = 1$임이 명백하므로 $X^k = X^{\underline{k}} - \sum_{i=0}^{k-1} a_iX^i$이다. 이를 $k-1$번 반복하면 다시 적당한 $b_0, \cdots, b_k \in \mathbb{R}$에 대해 $X^k = \sum_{i=0}^{k} b_iX^i$이고, 곧 X의 k차 적률도 존재한다.

i ⇔ v. 이는 i ⇔ iv의 증명과 비슷하게 하면 된다. □

앞서 적률 그 자체에 의미를 해석하는 것은 쉽지 않다고 했는데, 특별히 3차와 4차 표준화된 적률에 대해서는 어느정도 널리 받아들여지는 해석이 있어 잠시 소개한다.

Definition 3.63 확률공간 $(\Omega, \mathscr{F}, \mathbb{P})$에서 정의된 rv. $X : \Omega \to \mathbb{R}$에 대해 만약 X의 3차 표준화된 적률이 존재한다면 이를 특별히 X의 **왜도(skewness)**라 하고 $\mathbf{Skew}(X)$, τ_X 혹은 간단히 τ로 쓴다. 비슷하게, X의 4차 표준화된 적률이 존재한다면 $\kappa_{4,X}-3$을 특별히 X의 **(excess) 첨도(− kurtosis)**라 하고 $\mathbf{Kurt}(X)$, κ_X 혹은 간단히 κ로 쓴다.

왜도는 분포가 얼마나 치우쳤는가, 즉 얼마나 비대칭인가에 대한 척도이다. 정규분포와 같이 분포가 기댓값을 기준으로 정확히 대칭을 이루면 왜도는 0이 되고, 기댓값이 분포

의 중앙값보다 작아서 분포가 왼쪽으로 치우쳐 있으면 왜도는 양수가 되며 이때 그 분포는 **positively skewed**되었다고 한다. 반대로 기댓값이 분포의 중앙값보다 커서 분포가 오른쪽으로 치우쳐 있으면 왜도는 음수가 되며 이때 그 분포는 **negatively skewed**되었다고 한다.

첨도는 분포가 얼마나 뾰족한가, 즉 분포의 꼬리가 얼마나 두꺼운가에 대한 척도이다. 왜도와는 달리 4차 표준화된 적률에서 3을 뺀 것으로 정의되어 있는데, 이는 표준정규분포의 4차 표준화된 적률이 3이어서 비교의 기준이 되는 표준정규분포의 첨도를 0으로 만들기 위함이다. 표준정규분포와 같이 첨도가 0이면 이때 그 분포는 **mesokurtic**하다고 하고, 표준정규분포보다 얇은 꼬리를 가지면 첨도는 양수가 되며 이때 그 분포는 **leptokurtic**하다고 한다. 반대로 표준정규분포보다 두꺼운 꼬리를 가지면 첨도는 음수가 되며 이때 그 분포는 **platykurtic**하다고 한다. 일반적으로 꼬리의 두께는 outlier와 같은 극단적 사건의 발생확률을 의미하므로 leptokurtic한 분포와 platykurtic한 분포는 표준정규분포에 비해 극단적 사건의 확률이 각각 더 낮고 높다.

한편, 왜도와 첨도는 모두 Affine 변환에 대해 불변이다. 다만, 왜도의 경우 부호 정도의 영향은 받으며, 이는 그 정의를 생각해보면 당연하다.

Theorem 3.64 확률공간 $(\Omega, \mathscr{F}, \mathbb{P})$에서 정의된 rv. $X : \Omega \to \mathbb{R}$와 임의의 $a, b \in \mathbb{R}$에 대해 $a \neq 0$이면 다음이 성립한다.

i. 만약 $\mathbf{Skew}(X)$가 존재하면 $\mathbf{Skew}(aX+b)$도 존재하고 $\mathbf{Skew}(aX+b) = \operatorname{sgn}(a)\mathbf{Skew}(X)$이다.
ii. 만약 $\mathbf{Kurt}(X)$가 존재하면 $\mathbf{Kurt}(aX+b)$도 존재하고 $\mathbf{Kurt}(aX+b) = \mathbf{Kurt}(X)$이다.

Proof 먼저 $a \neq 0$이므로 $\mathbf{Var}(aX+b) = a^2\mathbf{Var}(X) \neq 0$이다. 또한, $\mathbf{Skew}(X)$가 존재하면 X의 3차 적률이 존재하여 곧 $\mathbf{E}(X), \mathbf{E}(X^2), \mathbf{E}(X^3)$이 모두 존재하며, 이로부터 $\mathbf{E}((aX+b)^3) = a^3\mathbf{E}(X^3) + 3a^2b\mathbf{E}(X^2) + 3ab^2\mathbf{E}(X) + b^3$가 존재하므로 $\mathbf{Skew}(aX+b)$도 존재한다. 한편,

$$\begin{aligned}\mathbf{Skew}(aX+b) &= \mathbf{E}\left(\left(\frac{aX+b-\mu_{aX+b}}{\sigma_{aX+b}}\right)^3\right)\\ &= \mathbf{E}\left(\left[\frac{a(X-\mu_X)}{|a|\sigma_X}\right]^3\right)\\ &= \mathbf{E}\left(\operatorname{sgn}(a)\left(\frac{X-\mu_X}{\sigma_X}\right)^3\right)\\ &= \operatorname{sgn}(a)\mathbf{Skew}(X)\end{aligned}$$

에서 i이 성립한다. 비슷하게 $\mathbf{Kurt}(X)$가 존재하면 $\mathbf{Kurt}(aX+b)$도 존재하고

$$\mathbf{Kurt}(aX+b) = \mathbf{E}\left(\left(\frac{aX+b-\mu_{aX+b}}{\sigma_{aX+b}}\right)^4\right) - 3$$

$$= \mathbf{E}\left(\left[\frac{a(X-\mu_X)}{|a|\sigma_X}\right]^4\right) - 3$$
$$= \mathbf{E}\left(\left(\frac{X-\mu_X}{\sigma_X}\right)^4\right) - 3$$
$$= \mathbf{Kurt}(X)$$

에서 ii가 성립한다. □

적률은 확률벡터인 경우로 보다 일반적으로 정의할 수도 있다. 물론, 이렇게 확률벡터에 대해 정의된 적률에 대해서도 명제 3.62, 3.61는 비슷하게 성립한다.

Definition 3.65 확률공간 $(\Omega, \mathscr{F}, \mathbb{P})$에서 정의된 rv. $X : \Omega \to \mathbb{R}^n$와 $k \in \mathbb{N}_0^n$에 대해 $K = \sum_{i=1}^n k_i$라 하고 다음을 정의한다.

i. 만약 $\mathbf{E}(\prod_{i=1}^n X_i^{k_i})$가 존재한다면 이를 X의 **K차 (결합)적률(Kth (joint) moment)**이라 하고 $\mu_{k,X}$ 혹은 간단히 μ_k으로 쓴다.
ii. 만약 $\mathbf{E}(\prod_{i=1}^n (X_i - \mu_{X_i})^{k_i})$가 존재한다면 이를 X의 **K차 중심화된 (결합)적률(nth (joint) centeral moment)**이라 하고 $\tau_{k,X}$ 혹은 간단히 τ_k으로 쓴다.
iii. 만약 $\mathbf{E}(\prod_{i=1}^n [(X_i - \mu_{X_i})/\sigma_{X_i}]^{k_i})$가 존재한다면 이를 X의 **K차 표준화된 (결합)적률(Kth (joint) standardized moment)**이라 하고 $\kappa_{k,X}$ 혹은 간단히 κ_k으로 쓴다.
iv. 만약 $\mathbf{E}(\prod_{i=1}^n X_i^{\underline{k_i}})$가 존재한다면 이를 X의 **K차 (하향결합)계승적률(Kth (joint falling) factorial moment)**이라 하고 $\mu_{\underline{k},X}$ 혹은 간단히 $\mu_{\underline{k}}$으로 쓴다.
v. 만약 $\mathbf{E}(\prod_{i=1}^n X_i^{\overline{k_i}})$가 존재한다면 이를 X의 **K차 (상향결합)계승적률(nth (joint rising) factorial moment)**이라 하고 $\mu_{\overline{k},X}$ 혹은 간단히 $\mu_{\overline{k}}$으로 쓴다.

Proposition 3.66 확률공간 $(\Omega, \mathscr{F}, \mathbb{P})$에서 정의된 rv. $X : \Omega \to \mathbb{R}^n$에 대해 만약 X의 K차 적률(혹은 중심화된 적률, 표준화된 적률, 하향계승적률, 상향계승적률)이 존재하면 X의 $L \le K$차 적률(혹은 중심화된 적률, 표준화된 적률, 하향계승적률, 상향계승적률)이 존재한다.

Proof 적률에 대해서는 수학적 귀납법을 사용하자. 우선 각 $i \le n$에 대해 $\mathbf{E}(X_i^K)$가 존재하므로 $\mathbf{E}(X_i)$도 존재하여 모든 1차 적률이 존재한다. 이제 귀납가정으로서 $L < K$에 대해 모든 $1, \cdots, L$차 적률이 존재한다고 하고 $L+1 = \sum_{i=1}^n l_i$인 임의의 $l \in \mathbb{N}_0^n$을 택하여 WLOG, $l_1 \ne 0$이라 하면 $\prod_{i=1}^n |X_i|^{l_i} \le |X_1|^{l_1+K-L} \prod_{i=2}^n |X_i|^{l_i} \mathbf{1}_{\{|X_1| \ge 1\}} + \prod_{i=2}^n |X_i|^{l_i} \mathbf{1}_{\{|X_1| \ge 1\}} \le |X_1|^{l_1+K-L} \prod_{i=2}^n |X_i|^{l_i} + \prod_{i=2}^n |X_i|^{l_i}$에서 $\mathbf{E}(\prod_{i=1}^n |X_i|^{l_i})$가 존재함을 알고, 곧 모든 $L+1$차 적률이 존재하여 모든 $L \le K$차 적률이 존재함을 안다. 그리고 이로부터 중심화된 적률과 표준화된 적률에 대해서는 명제가 자명해진다.

다음으로 X의 하향계승적률에 대해서도 수학적 귀납법을 사용하자. 우선 각 $i \le n$에 대해 $\mathbf{E}(X_i^{\underline{K}})$가 존재하므로 $\mathbf{E}(X_i)$도 존재하여 모든 1차 하향계승적률이 존재한다. 이제 귀납 가정으로서 $L < K$에 대해 모든 $1, \cdots, L$차 적률이 존재한다고 하고 $L+1 = \sum_{i=1}^n l_i$인 임의의 $l \in \mathbb{N}_0^n$을 택하여 WLOG, $l_1 \neq 0$이라 하면

$$\begin{aligned}
&\prod_{i=1}^n |X_i^{\underline{l_i}}| \\
&\qquad \le \frac{|X_1^{\underline{l_1+K-L}}|}{|(X_1-l_1)^{\underline{K-L}}|}\left(\prod_{i=2}^n |X_i^{\underline{l_i}}|\right)\left(\prod_{j=l_1}^{l_1+K-L-1} \mathbf{1}_{\{X_1 \neq j\}}\right) + \sum_{j=l_1}^{l_1+K-L-1} j^{\underline{l_1}}\left(\prod_{i=2}^n |X_i^{\underline{l_i}}|\right)\mathbf{1}_{\{X_1 = j\}}
\end{aligned}$$

이다. 한편, 함수 $f : \mathbb{R} \to \mathbb{R}$를 $f : x \mapsto |(x-l_1)^{\underline{K-L}}|$로 두면 이는 적당한 $M > l_1 + K - L - 1$이 존재하여 $(-\infty, -M)$와 (M, ∞)에서 f는 각각 감소하고 증가하며 $f(-M), f(M) \ge 1$이다. 이상으로부터

$$\begin{aligned}
\prod_{i=1}^n |X_i^{\underline{l_i}}| &\le |X_1^{\underline{l_1+K-L}}|\left(\prod_{i=2}^n |X_i^{\underline{l_i}}|\right)\mathbf{1}_{\{|X_1| \ge M\}} + \left(\prod_{i=1}^n |X_i^{\underline{l_i}}|\right)\left(\prod_{j=l_1}^{l_1+K-L-1} \mathbf{1}_{\{X_1 \neq j\}}\right)\mathbf{1}_{\{|X_1| < M\}} \\
&\qquad + \sum_{j=l_1}^{l_1+K-L-1} j^{\underline{l_1}}\left(\prod_{i=2}^n |X_i^{\underline{l_i}}|\right)\mathbf{1}_{\{X_1 = j\}} \\
&\le |X_1^{\underline{l_1+K-L}}|\prod_{i=2}^n |X_i^{\underline{l_i}}| + M^{\underline{l_1}}\prod_{i=2}^n |X_i^{\underline{l_i}}| + \sum_{j=l_1}^{l_1+K-L-1} M^{\underline{l_1}}\prod_{i=2}^n |X_i^{\underline{l_i}}| \\
&= |X_1^{\underline{l_1+K-L}}|\prod_{i=2}^n |X_i^{\underline{l_i}}| + (K-L+1)M^{\underline{l_1}}\prod_{i=2}^n |X_i^{\underline{l_i}}|
\end{aligned}$$

이다. 여기서 두 번째 부등호는 함수 $x \mapsto |x^{\underline{l_1}}|$이 $(l_1 - 1, \infty)$에서 증가하므로 성립한다. 이로부터 $\mathbf{E}(\prod_{i=1}^n |X_i^{\underline{l_i}}|)$가 존재함을 알고, 곧 모든 $L+1$차 계승하향적률이 존재하여 모든 $L \le K$의 하향계승적률이 존재함을 안다. □

Proposition 3.67 확률공간 $(\Omega, \mathscr{F}, \mathbb{P})$에서 정의된 rv. $X : \Omega \to \mathbb{R}^n$에 대해 TFAE.

i. Rv. X의 모든 K차 적률이 존재한다.
ii. 각 $i \le n$에 대해 $\mathbf{E}(X_i)$가 존재한다면 X의 모든 K차 중심화된 적률이 존재한다.
iii. 각 $i \le n$에 대해 $\mathbf{Var}(X_i)$가 존재하고 0이 아니면 X의 모든 K차 표준화된 적률이 존재한다.
iv. Rv. X의 모든 K차 하향계승적률이 존재한다.
v. Rv. X의 모든 K차 상향계승적률이 존재한다.

Proof 만약 $K = 0$이면 명제가 자명하므로 $K > 0$이라 하자.

i $\Leftrightarrow$ ii. 명제 3. 66로부터 X의 모든 $L \le K$차 적률이 존재한다. 이제 $K = \sum_{i=1}^{n} k_i$인 임의의 $k \in \mathbb{N}_0^n$에 대해 $\prod_{i=1}^{n}(X_i - \mu_{X_i})^{k_i}$의 전개식에서 $\prod_{i=1}^{n} X_i^{k_i}$가 최고차항임을 생각해보면 충분조건임이 자명하고, 필요조건임도 이와 비슷하게 보일 수 있다.

ii $\Leftrightarrow$ iii. 이는 정의로부터 자명하다.

i $\Rightarrow$ iv. 명제 3. 66로부터 X의 모든 $L \le K$차 적률이 존재한다. 이제 $K = \sum_{i=1}^{n} k_i$인 임의의 $k \in \mathbb{N}_0^n$에 대해 $\prod_{i=1}^{n} X_i^{\underline{k_i}} = \prod_{i=1}^{n} \prod_{j=0}^{k_i - 1}(X_i - j)$의 전개식에서 $\prod_{i=1}^{n} X_i^{k_i}$가 최고차항임을 생각해보면 명제가 자명하다.

iv $\Rightarrow$ i. 수학적 귀납법을 사용하자. 우선 $K = 1$이면 명제 3. 62로부터 자명하다. 이제 귀납가정으로서 모든 $L < K$에 대해 명제가 성립한다고 하고 모든 $L+1$차 하향계승적률이 존재한다고 하자. 또한 $L+1 = \sum_{i=1}^{n} l_i$인 임의의 $l \in \mathbb{N}_0^n$을 택하자. 그렇다면 명제 3. 66로부터 X의 모든 $M \le L+1$차 하향계승적률이 존재하고, $\prod_{i=1}^{n} X_i^{\underline{l_i}} = \prod_{i=1}^{n} \prod_{j=0}^{l_i - 1}(X_i - j)$의 전개식에서 $\prod_{i=1}^{n} X_i^{l_i}$가 최고차항이므로 $\prod_{i=1}^{n} X_i^{l_i}$는 $\prod_{i=1}^{n} X_i^{\underline{k_i}}$와 $1, \cdots, L$차항 그리고 상수항의 선형결합으로 쓸 수 있어서 $\mathbf{E}(\prod_{i=1}^{n} X_i^{l_i})$가 존재하고, 곧 명제가 성립한다.

i $\Leftrightarrow$ v. 이는 i $\Leftrightarrow$ iv의 증명과 비슷하게 하면 된다. □

보통 적률에 대한 이론을 전개할 때에는 적률을 그 자체로 다루는 것보다 아래의 MGF를 통해 다루는 것이 더 편리하다.

Definition 3.68 확률공간 $(\Omega, \mathscr{F}, \mathbb{P})$에서 정의된 rv. $X : \Omega \to \mathbb{R}^n$와 적당한 0의 근방 $U \subseteq \mathbb{R}^n$이 존재하여 임의의 $t \in U$에 대해 $\mathbf{E}(e^{t^\mathsf{T} X})$가 존재하면 rv. X의 **적률생성함수(moment generating function)**를 $M_X : U \to \mathbb{R}$로 쓰고 $M_X : t \mapsto \mathbf{E}(e^{t^\mathsf{T} X})$로 정의한다. 특별히, $n \ge 2$인 경우 M_X를 rv. $X_1, \cdots, X_n$의 **결합적률생성함수(joint moment generating function)**라 하기도 한다.

Theorem 3.69 확률공간 $(\Omega, \mathscr{F}, \mathbb{P})$에서 정의된 rv. $X : \Omega \to \mathbb{R}^n$에 대해 M_X가 적당한 0의 근방 $U \subseteq \mathbb{R}^n$ 위에서 존재하면 다음이 성립한다.

i. $M_X(0) = 1$.
ii. $M_X > 0$.
iii. MGF M_X는 log−볼록하다.
iv. 임의의 $a \in \mathbb{R}$와 임의의 $b \in \mathbb{R}^n$에 대해 M_{aX+b}가 적당한 0의 근방 $V \subseteq \mathbb{R}^n$에서 존재하고, $at \in U$인 임의의 $t \in V$에 대해 $M_{aX+b}(t) = e^{t^\mathsf{T} b} M_X(at)$이다.

Proof i. 이는 $M_X(0) = \mathbf{E}(1) = 1$에서 자명하다.

ii. 따름정리 2. 102의 iii으로부터 임의의 $t \in \mathbb{R}^n$에 대해 $M_X(t) = \mathbf{E}(e^{t^\mathsf{T} X}) > \mathbf{E}(0) = 0$이므로 이는 자명하다.

iii. 임의의 $s, t \in U$와 $\lambda \in (0, 1)$에 대해 Hölder의 부등식으로부터 $M_X(\lambda s + (1-\lambda)t) = \mathbf{E}(e^{[\lambda s + (1-\lambda)t]^\mathsf{T} X}) \le [\mathbf{E}(e^{s^\mathsf{T} X})]^\lambda [\mathbf{E}(e^{t^\mathsf{T} X})]^{1-\lambda} = M_X(s)^\lambda M_X(t)^{1-\lambda}$가 되고, ii로부터 $M_X > 0$

이므로 양변에 log를 취하면 $\log M_X(\lambda s+(1-\lambda)t) \le \lambda \log M_X(s)+(1-\lambda)\log M_X(t)$이다. 이 부등식이 $\lambda = 0, 1$인 경우에도 성립함이 자명하므로 M_X는 log−볼록하다.

iv. 집합 $V = \{t \in \mathbb{R}^n : at \in U\}$를 생각하면 이는 명백히 0의 근방이고, 임의의 $t \in V$에 대해 $\mathbf{E}(e^{t^\mathsf{T}(aX+b)}) = e^{t^\mathsf{T}b}\mathbf{E}(e^{(at)^\mathsf{T}X}) = e^{t^\mathsf{T}b}M_X(at)$에서 정리가 성립한다. □

그러나 MGF의 정의만 보아서는 이와 적률 사이의 관계를 눈치채기가 쉽지 않다. 이에 다음 정리는 MGF와 적률의 관계를 명시적으로 잘 보여준다.

Lemma 3.70 확률공간 $(\Omega, \mathscr{F}, \mathbb{P})$에서 정의된 rv. $X : \Omega \to \mathbb{R}^n$에 대해 M_X가 적당한 0의 근방 $U \subseteq \mathbb{R}^n$ 위에서 존재하면, 적당한 0의 근방 $V \subseteq U$가 존재하여 임의의 $t \in V$에 대해 $\mathbf{E}(e^{|t|^\mathsf{T}|X|})$가 존재한다. 특별히, $U = \mathbb{R}^n$이면 $V = \mathbb{R}^n$으로 잡을 수 있다.

Proof 먼저 $V := B(||t_0||) \subseteq U$인 0이 아닌 $t_0 \in U$를 택하고, 이어서 모든 성분이 음이 아닌 임의의 $t \in V$를 택하고, $\mathbb{R}^n$을 2^n개의 사분공간으로 나누어 각각을 $A_1, \cdots, A_{2^n}$이라 하자. (이때 각 사분면의 경계도 적당히 나눈다.) 이제 임의의 A_j를 고정하고 임의의 $x \in A_j^\circ$에 대해 $v = (\mathrm{sgn}(x_i))$라 하면 이는 x의 선택에 무관하게 well-define된다. 그렇다면 $s = (v_i t_i) \in V$임이 명백하므로

$$\begin{aligned}
\int_\Omega e^{t^\mathsf{T}|X|}\mathbf{1}_{A_j}(X)\,d\mathbb{P} &= \int_{A_j} e^{t^\mathsf{T}|x|}\,d\mathbf{P}_X \\
&= \int_{A_j} \exp\left(\sum_{i=1}^n v_i t_i x_i\right) d\mathbf{P}_X \\
&= \int_{A_j} e^{s^\mathsf{T}x}\,d\mathbf{P}_X \\
&\le \int_{\mathbb{R}^n} e^{s^\mathsf{T}x}\,d\mathbf{P}_X \\
&= \mathbf{E}(e^{s^\mathsf{T}X})
\end{aligned}$$

에서 $e^{t^\mathsf{T}|X|}\mathbf{1}_{A_j}(X)$가 적분가능하고, 곧 $e^{t^\mathsf{T}|X|} = \sum_{j=1}^{2^n} e^{t^\mathsf{T}|X|}\mathbf{1}_{A_j}(X)$도 적분가능하여 $\mathbf{E}(e^{t^\mathsf{T}|X|})$가 존재한다. 그렇다면 이는 임의의 $t \in V$에 대해 $\mathbf{E}(e^{|t|^\mathsf{T}|X|})$가 존재함을 함의하는 한편, $U = \mathbb{R}^n$인 경우에는 이상의 논의에서 $V = \mathbb{R}^n$으로 택할 수 있음이 자명하므로 증명이 끝난다. □

Theorem 3.71 확률공간 $(\Omega, \mathscr{F}, \mathbb{P})$에서 정의된 rv. $X : \Omega \to \mathbb{R}^n$에 대해 M_X가 적당한 0의 근방 $U \subseteq \mathbb{R}^n$ 위에서 존재하면 X의 모든 적률이 존재하고, 적당한 0의 근방 $V \subseteq U$가 존재하여 임의의 $t \in V$에 대해

$$M_X(t) = \sum_{j=0}^\infty \sum_{k_1+\cdots+k_n=j} \mathbf{E}\left(\prod_{i=1}^n X_i^{k_i}\right)\prod_{i=1}^n \frac{t_i^{k_i}}{k_i!} = \sum_{k\in\mathbb{N}_0^n} \mathbf{E}\left(\prod_{i=1}^n X_i^{k_i}\right)\prod_{i=1}^n \frac{t_i^{k_i}}{k_i!}$$

이다. 특별히, $U = \mathbb{R}^n$이면 $V = \mathbb{R}^n$으로 잡을 수 있으며, 여기서 급수는 절대수렴하여 well-define된다.

PROOF 먼저 임의의 $k \in \mathbb{N}_0^n$에 대해 $\mathbf{E}(\prod_{i=1}^n X_i^{k_i})$가 존재함을 보이자. 이를 위해 $K = \sum_{i=1}^n k_i$로 두고 각 성분이 양수인 $t \in U$를 택하면

$$\begin{aligned}
\left|\prod_{i=1}^n X_i^{k_i}\right| &= \prod_{i=1}^n t_i |X_i|^{k_i} \Big/ \prod_{i=1}^n t_i^{k_i} \\
&\le \left(\sum_{i=1}^n t_i |X_i|\right)^K \Big/ \left(\prod_{i=1}^n t_i^{k_i}\right)\binom{K}{k_1 \cdots k_n} \\
&= \left(\prod_{i=1}^n \frac{k_i!}{t_i^{k_i}}\right)\frac{(t^\top |X|)^K}{K!} \\
&\le \left(\prod_{i=1}^n \frac{k_i!}{t_i^{k_i}}\right)\sum_{j=0}^\infty \frac{(t^\top |X|)^j}{j!} \\
&= \left(\prod_{i=1}^n \frac{k_i!}{t_i^{k_i}}\right) e^{t^\top |X|}
\end{aligned}$$

인데, 보조정리로부터 t를 충분히 작게 잡아주면 $\mathbf{E}(e^{t^\top |X|})$가 존재하므로 $\mathbf{E}(\prod_{i=1}^n X_i^{k_i})$도 존재함을 알고, 곧 X의 모든 적률이 존재한다.

다음으로, 임의의 $l \in \mathbb{N}_0$와 $t \in U$에 대해

$$\begin{aligned}
\left|\sum_{j=0}^l \frac{(t^\top X)^j}{j!}\right| &\le \sum_{j=0}^l \frac{|t^\top X|^j}{j!} \\
&\le \sum_{j=0}^\infty \frac{|t^\top X|^j}{j!} \\
&= e^{|t^\top X|} \\
&\le e^{t^\top X}\mathbf{1}_{\{t^\top X \ge 0\}} + \mathbf{1}_{\{t^\top X < 0\}} \\
&\le e^{t^\top X} + 1
\end{aligned}$$

이므로 DCT에서

$$\begin{aligned}
M_X(t) &= \mathbf{E}\left(\sum_{j=0}^\infty \frac{(t^\top X)^j}{j!}\right) \\
&= \sum_{j=0}^\infty \frac{\mathbf{E}((t^\top X)^j)}{j!} \\
&= \sum_{j=0}^\infty \frac{1}{j!} \sum_{k_1+\cdots+k_n=j} \binom{j}{k_1 \cdots k_n} \mathbf{E}\left(\prod_{i=1}^n (t_i X_i)^{k_i}\right)
\end{aligned}$$

$$= \sum_{j=0}^{\infty} \sum_{k_1+\cdots+k_n=j} \mathbf{E}\left(\prod_{i=1}^{n} X_i^{k_i}\right) \prod_{i=1}^{n} \frac{t_i^{k_i}}{k_i!} \tag{$*$}$$

이다. 이제 $V := B(||t_0||) \subseteq U$인 0이 아닌 $t_0 \in U$를 택하자. 그렇다면

$$\sum_{j=0}^{\infty} \sum_{k_1+\cdots+k_n=j} \left| \mathbf{E}\left(\prod_{i=1}^{n} X_i^{k_i}\right) \prod_{i=1}^{n} \frac{t_i^{k_i}}{k_i!} \right| \leq \sum_{j=0}^{\infty} \sum_{k_1+\cdots+k_n=j} \mathbf{E}\left(\prod_{i=1}^{n} |X_i|^{k_i}\right) \prod_{i=1}^{n} \frac{|t_i|^{k_i}}{k_i!}$$

이므로 보조정리로부터 t_0를 충분히 작게 잡아주면 $\mathbf{E}(e^{|t|^\top |X|})$가 존재하고, $(*)$에서와 비슷하게 하여 $(*)$의 급수가 절대수렴함을 쉽게 보일 수 있다. 이로부터 $(*)$의 급수에서의 합의 순서를 바꿀 수 있으므로 원하는 형태의 식을 얻고, $U = \mathbb{R}^n$인 경우에는 이상의 논의에서 $V = \mathbb{R}^n$으로 택할 수 있음이 자명하므로 증명이 끝난다. □

한편, MGF는 적률을 계산하는 방법으로도 유용하게 쓰인다.

Theorem 3.72 확률공간 $(\Omega, \mathscr{F}, \mathbb{P})$에서 정의된 rv. $X : \Omega \to \mathbb{R}^n$에 대해 M_X가 적당한 0의 근방 $U \subseteq \mathbb{R}^n$ 위에서 존재한다고 하자. 그렇다면 임의의 $k \in \mathbb{N}_0^n$에 대해 $K = \sum_{i=1}^n k_i$라 하면 $\mathbf{E}(\prod_{i=1}^n X_i^{k_i}) = (\partial^K / \partial t_1^{k_1} \cdots \partial t_n^{k_n}) M_X(0)$이다. 이때, 편미분의 순서는 중요하지 않다.

Proof 정리 3. 71로부터 적당 0의 근방 $V \subseteq U$가 존재하여 임의의 $t \in V$에 대해

$$M_X(t) = \sum_{l \in \mathbb{N}_0^n} \mathbf{E}\left(\prod_{i=1}^{n} X_i^{l_i}\right) \prod_{i=1}^{n} \frac{t_i^{l_i}}{l_i!}$$

인데, 이가 멱급수이므로

$$\begin{aligned}
\frac{\partial^K}{\partial t_1^{k_1} \cdots \partial t_n^{k_n}} M_X(0) &= \sum_{l \in \mathbb{N}_0^n} \mathbf{E}\left(\prod_{i=1}^{n} X_i^{l_i}\right) \frac{\partial^K}{\partial t_1^{k_1} \cdots \partial t_n^{k_n}} \prod_{i=1}^{n} \frac{t_i^{l_i}}{l_i!} \bigg|_{t=0} \\
&= \sum_{k \leq l \in \mathbb{N}_0^n} \mathbf{E}\left(\prod_{i=1}^{n} X_i^{l_i}\right) \prod_{i=1}^{n} \frac{t_i^{l_i - k_i}}{(l_i - k_i)!} \bigg|_{t=0} \\
&= \mathbf{E}\left(\prod_{i=1}^{n} X_i^{k_i}\right)
\end{aligned}$$

이다. □

이렇게 적률과 긴밀하게 연결되어 있는 MGF이지만 이는 그 존재성이 항상 보장되지 않는다는 큰 단점을 가진다. 이론 전개에 있어 이러한 단점은 많은 귀찮음을 유발하므로 자연스럽게 항상 존재하면서도 MGF의 역할을 해 줄수 있는 대체재를 생각해내기에 이르렀는데, 그것이 바로 확률벡터의 PDF의 Fourier 변환인 CF이다. 아래의 정의에서 기댓값에 복소함수가 들어갔는데, 측도론의 에필로그에서 논한 복소함수의 적분에 대한 내용을 그대로 적용하

면 된다. 즉, 확률변수 X, Y에 대해 $\mathbf{E}(X)$와 $\mathbf{E}(Y)$가 존재한다면 $\mathbf{E}(X+\mathbf{i}Y) = \mathbf{E}(X) + \mathbf{i}\mathbf{E}(Y)$이다.

Definition 3.73 확률공간 $(\Omega, \mathscr{F}, \mathbb{P})$에서 정의된 rv. $X : \Omega \to \mathbb{R}^n$에 대해 rv. X의 **특성함수(characteristic function)**를 $\varphi_X : \mathbb{R}^n \to \mathbb{C}$로 쓰고 $\varphi_X : t \mapsto \mathbf{E}(e^{\mathbf{i}t^\top X})$로 정의한다. 특별히, $n \geq 2$인 경우 φ_X를 rv. $X_1, \cdots, X_n$의 **결합특성함수(joint characteristic function)**라 하기도 한다.

Proposition 3.74 확률공간 $(\Omega, \mathscr{F}, \mathbb{P})$에서 정의된 rv. $X : \Omega \to \mathbb{R}^n$와 임의의 $t \in \mathbb{R}^n$에 대해 $\mathbf{E}(e^{\mathbf{i}t^\top X})$가 항상 존재한다. 따라서 φ_X는 well-define된다.

Proof 함수 $e^{\mathbf{i}t^\top X}$의 실수부 $\cos t^\top X$와 허수부 $\sin t^\top X$가 모두 적분가능함을 보이면 되는데, 이는 $|\cos t^\top X|, |\sin t^\top X| \leq 1$에서 자명하다. □

CF는 항상 존재한다는 장점과 더불어 함수로서도 훌륭한 성질들을 가진다.

Theorem 3.75 확률공간 $(\Omega, \mathscr{F}, \mathbb{P})$에서 정의된 rv. $X : \Omega \to \mathbb{R}^n$에 대해 다음이 성립한다.

i. $\varphi_X(0) = 1$.
ii. $|\varphi_X| \leq 1$.
iii. φ_X는 균등연속이다.
iv. 임의의 $a \in \mathbb{R}$와 임의의 $b, t \in \mathbb{R}^n$에 대해 $\varphi_{aX+b}(t) = e^{\mathbf{i}t^\top b}\varphi_X(at)$이다.

Proof i. 이는 $\varphi_X(0) = \mathbf{E}(1) = 1$에서 자명하다.

ii. 임의의 $t \in \mathbb{R}^n$에 대해 $|\varphi_X(t)| = |\mathbf{E}(e^{\mathbf{i}t^\top X})| \leq \mathbf{E}(|e^{\mathbf{i}t^\top X}|) = \mathbf{E}(1) = 1$이므로 이는 자명하다.

iii. 함수 $f : \Omega \times \mathbb{R}^n$를 $f : (\omega, t) \mapsto |e^{\mathbf{i}t^\top X(\omega)} - 1|$로 두면 임의의 $\omega \in \Omega$에 대해 f_ω는 연속이다. 또한, 임의의 $(\omega, t) \in \Omega \times \mathbb{R}^n$에 대해 $|e^{\mathbf{i}t^\top X(\omega)} - 1| \leq |e^{\mathbf{i}t^\top X(\omega)}| + 1 = 2$이므로 임의의 $t \in \mathbb{R}^n$에 대해 f^t는 적분가능하고 $|f| \leq 2$이다. 이상으로부터 f가 따름정리 2.108의 조건을 모두 만족시킴을 알 수 있으므로 함수 $t \mapsto \int_\Omega |e^{\mathbf{i}t^\top X} - 1|\, d\mathbb{P} = \mathbf{E}(|e^{\mathbf{i}t^\top X} - 1|)$는 연속이다. 특히 이가 0에서 연속이므로 $t \to 0$이면 $\mathbf{E}(|e^{\mathbf{i}t^\top X} - 1|) \to 0$이고, 곧 임의의 $\varepsilon > 0$을 택하면 적당한 $\delta > 0$가 존재하여 $|t| < \delta$이면 $\mathbf{E}(|e^{\mathbf{i}t^\top X} - 1|) < \varepsilon$이다. 이로부터 $||s-t|| < \delta$인 임의의 $s, t \in \mathbb{R}^n$에 대해 $|\varphi_X(s) - \varphi_X(t)| = |\mathbf{E}(e^{\mathbf{i}s^\top X} - e^{\mathbf{i}t^\top X})| \leq \mathbf{E}(|e^{\mathbf{i}t^\top X}(e^{\mathbf{i}(s-t)^\top X} - 1)|) = \mathbf{E}(|e^{\mathbf{i}(s-t)^\top X} - 1|) < \varepsilon$이 되어 φ_X가 균등연속임을 안다.

iv. 이는 $\mathbf{E}(e^{\mathbf{i}t^\top(aX+b)}) = e^{\mathbf{i}t^\top b}\mathbf{E}(e^{\mathbf{i}(at)^\top X}) = e^{\mathbf{i}t^\top b}\varphi_X(at)$에서 자명하다. □

나아가, MGF와 적률의 관계를 잘 보여주었던 정리 3. 71, 3. 72가 CF에 대해서도 거의 비슷하게 (사실은 더 잘) 성립한다.

Lemma 3.76 임의의 $x, y \in \mathbb{R}^n$와 $k \in \mathbb{N}$에 대해

$$\left| e^{\mathbf{i} x^\top y} - \sum_{i=0}^{k} \mathbf{i}^i \frac{(x^\top y)^i}{i!} \right| \le \min\left\{ \frac{(|x|^\top |y|)^{k+1}}{(k+1)!}, \frac{2(|x|^\top |y|)^k}{k!} \right\}$$

이다.

Proof Taylor의 정리로부터

$$\begin{aligned} e^{\mathbf{i} x^\top y} &= \sum_{i=0}^{k-1} \mathbf{i}^i \frac{(x^\top y)^i}{i!} + \mathbf{i}^k \frac{(x^\top y)^k}{(k-1)!} \int_0^1 (1-t)^{k-1} e^{\mathbf{i} t x^\top y}\, dt \\ &= \sum_{i=0}^{k} \mathbf{i}^i \frac{(x^\top y)^i}{i!} + \mathbf{i}^k \frac{(x^\top y)^k}{(k-1)!} \int_0^1 (1-t)^{k-1} (e^{\mathbf{i} t x^\top y} - 1)\, dt \end{aligned}$$

이다. 여기서 마지막 등호는 $\int_0^1 (1-t)^{k-1}\, dt = 1/k$에서 성립한다. 이로부터

$$\begin{aligned} \left| e^{\mathbf{i} x^\top y} - \sum_{i=0}^{k} \mathbf{i}^i \frac{(x^\top y)^i}{i!} \right| &= \left| \mathbf{i}^{k+1} \frac{(x^\top y)^{k+1}}{k!} \int_0^1 (1-t)^k e^{\mathbf{i} t x^\top y}\, dt \right| \\ &\le \frac{(|x|^\top |y|)^{k+1}}{k!} \int_0^1 (1-t)^k |e^{\mathbf{i} t x^\top y}|\, dt \\ &= \frac{(|x|^\top |y|)^{k+1}}{k!} \int_0^1 (1-t)^k\, dt \\ &= \frac{(|x|^\top |y|)^{k+1}}{(k+1)!} \end{aligned}$$

이고

$$\begin{aligned} \left| e^{\mathbf{i} x^\top y} - \sum_{i=0}^{k} \mathbf{i}^i \frac{(x^\top y)^i}{i!} \right| &= \left| \mathbf{i}^k \frac{(x^\top y)^k}{(k-1)!} \int_0^1 (1-t)^{k-1} (e^{\mathbf{i} t x^\top y} - 1)\, dt \right| \\ &\le \frac{(|x|^\top |y|)^k}{(k-1)!} \int_0^1 (1-t)^{k-1} |e^{\mathbf{i} t x^\top y} - 1|\, dt \\ &\le \frac{(|x|^\top |y|)^k}{(k-1)!} \int_0^1 (1-t)^{k-1} (|e^{\mathbf{i} t x^\top y}| + 1)\, dt \\ &= \frac{(|x|^\top |y|)^k}{(k-1)!} \int_0^1 2(1-t)^{k-1}\, dt \\ &= \frac{2(|x|^\top |y|)^k}{k!} \end{aligned}$$

이므로

$$\left| e^{\mathbf{i} x^\top y} - \sum_{i=0}^{k} \mathbf{i}^i \frac{(x^\top y)^i}{i!} \right| \le \min\left\{ \frac{(|x|^\top |y|)^{k+1}}{(k+1)!}, \frac{2(|x|^\top |y|)^k}{k!} \right\}$$

이다. □

Theorem 3.77 확률공간 $(\Omega, \mathscr{F}, \mathbb{P})$에서 정의된 rv. $X : \Omega \to \mathbb{R}^n$에 대해 M_X가 $\mathbb{R}^n$ 위에서 존재하면 임의의 $t \in \mathbb{R}^n$에 대해

$$\varphi_X(t) = \sum_{j=0}^{\infty} \mathbf{i}^j \sum_{k_1+\cdots+k_n=j} \mathbf{E}\left(\prod_{i=1}^n X_i^{k_i}\right) \prod_{i=1}^n \frac{t_i^{k_i}}{k_i!} = \sum_{k\in\mathbb{N}_0^n} \mathbf{i}^{k_1+\cdots+k_n} \mathbf{E}\left(\prod_{i=1}^n X_i^{k_i}\right) \prod_{i=1}^n \frac{t_i^{k_i}}{k_i!}$$

이다. 여기서 급수는 절대수렴하여 well-define된다.

Proof 위의 보조정리로부터 임의의 $t \in \mathbb{R}^n$와 임의의 $K \in \mathbb{N}$를 택하면

$$\begin{aligned}
&\left|\varphi_X(t) - \sum_{j=0}^{K} \mathbf{i}^j \sum_{k_1+\cdots+k_n=j} \mathbf{E}\left(\prod_{i=1}^n X_i^{k_i}\right) \prod_{i=1}^n \frac{t_i^{k_i}}{k_i!}\right| \\
&\quad= \left|\varphi_X(t) - \sum_{j=0}^{K} \frac{\mathbf{i}^j}{j!} \sum_{k_1+\cdots+k_n=j} \binom{j}{k_1 \cdots k_n} \mathbf{E}\left(\prod_{i=1}^n t_i X_i\right)\right| \\
&\quad= \left|\varphi_X(t) - \sum_{j=0}^{K} \mathbf{i}^j \frac{\mathbf{E}((t^\top X)^j)}{j!}\right| \\
&\quad= \left|\mathbf{E}\left(e^{\mathbf{i}t^\top X} - \sum_{j=0}^{K} \mathbf{i}^j \frac{(t^\top X)^j}{j!}\right)\right| \\
&\quad\le \mathbf{E}\left(\left|e^{\mathbf{i}t^\top X} - \sum_{j=0}^{K} \mathbf{i}^j \frac{(t^\top X)^j}{j!}\right|\right) \\
&\quad\le \min\left\{\frac{\mathbf{E}((|t|^\top |X|)^{K+1})}{(K+1)!}, \frac{2\mathbf{E}((|t|^\top |X|)^{K})}{K!}\right\} \\
&\quad\le \min\left\{\frac{1}{(K+1)!} \sum_{k_1+\cdots+k_n=K+1} \binom{K+1}{k_1 \cdots k_n} \mathbf{E}\left(\prod_{i=1}^n |X_i|^{k_i}\right) \prod_{i=1}^n t_i^{k_i},\right. \\
&\qquad\qquad \left.\frac{2}{K!} \sum_{k_1+\cdots+k_n=K} \binom{l}{k_1 \cdots k_n} \mathbf{E}\left(\prod_{i=1}^n |X_i|^{k_i}\right) \prod_{i=1}^n t_i^{k_i}\right\}
\end{aligned}$$

인데, M_X가 $\mathbb{R}^n$ 위에서 존재하므로 정리 3. 71로부터 $l \to \infty$이면 우변의 두 항이 모두 0으로 수렴하여 $\varphi_X(t) = \sum_{j=0}^{\infty} \mathbf{i}^j \sum_{k_1+\cdots+k_n=j} \mathbf{E}(\prod_{i=1}^n X_i^{k_i}) \prod_{i=1}^n t_i^{k_i}/k_i!$임을 안다. 한편, 다시 정리 3. 71로부터 이 급수가 절대수렴함을 알고, 곧 합의 순서를 바꿀 수 있으므로 원하는 형태의 식을 얻는다. □

Theorem 3.78 확률공간 $(\Omega, \mathscr{F}, \mathbb{P})$에서 정의된 rv. $X : \Omega \to \mathbb{R}^n$와 $k \in \mathbb{N}_0^n$에 대해 X의 k차 적률이 존재한다고 하자. 그렇다면 $K = \sum_{i=1}^n k_i$와 임의의 $t \in \mathbb{R}^n$에 대해 $\mathbf{E}((\prod_{i=1}^n X_i^{k_i}) e^{\mathbf{i}t^\top X})$가 존재하고

$$\frac{\partial^K}{\partial t_1^{k_1}\cdots\partial t_n^{k_n}}\varphi_X(t)=\mathbf{i}^K\mathbf{E}\left(\left(\prod_{i=1}^n X_i^{k_i}\right)e^{\mathbf{i}t^\top X}\right)$$

이다. 이때, 편미분의 순서는 중요하지 않다.

Proof 간결한 논의를 위해 $k_n\neq 0$이라 하자. ($k_n=0$인 경우에도 이와 비슷하게 하면 된다.) 임의의 $t\in\mathbb{R}^n$를 고정하고 함수 $f:\Omega\times\mathbb{R}$를 $f:(\omega,h)\mapsto e^{\mathbf{i}(t+h\mathbf{e}_n)^\top X(\omega)}$로 두면 임의의 $(\omega,h)\in\Omega\times\mathbb{R}$에 대해 $|(\partial/\partial h)f(\omega,h)|=|\mathbf{i}X_n(\omega)e^{\mathbf{i}(t+h\mathbf{e}_n)^\top X(\omega)}|=|X_n(\omega)|$이고 $\mathbf{E}(X_n)$이 존재하므로 f가 Leibniz의 법칙의 모든 조건을 만족시킨다. 따라서 함수 $h\mapsto\int_\Omega e^{\mathbf{i}(t+h\mathbf{e}_n)^\top X}\,d\mathbb{P}=\varphi_X(t+h\mathbf{e}_n)$는 미분가능하고 $\mathbf{E}(X_ne^{\mathbf{i}t^\top X})$가 존재하며, $(\partial/\partial t_n)\varphi_X(t)=(d/dh)\varphi_X(t+h\mathbf{e}_n)|_{h=0}=\int_\Omega \mathbf{i}X_ne^{\mathbf{i}t^\top X}\,d\mathbb{P}=\mathbf{i}\mathbf{E}(X_ne^{\mathbf{i}t^\top X})$이다. 이제 이와 비슷하게 $K-1$번 반복하면 원하는 결과를 얻는다. □

Corollary 3.79 확률공간 $(\Omega,\mathscr{F},\mathbb{P})$에서 정의된 rv. $X:\Omega\to\mathbb{R}^n$와 $k\in\mathbb{N}_0^n$에 대해 X의 k차 적률이 존재한다고 하자. 그렇다면 $K=\sum_{i=1}^n k_i$에 대해 $\mathbf{E}(\prod_{i=1}^n X_i^{k_i})=(\partial^K/\partial t_1^{k_1}\cdots\partial t_n^{k_n})\varphi_X(0)/\mathbf{i}^K$이다. 이때, 편미분의 순서는 중요하지 않다.

Proof 이는 위의 정리로부터 자명하다. □

앞서 적률을 도입하며 이를 이용하면 분포의 정보를 손실 없이 전달할 수 있다고 하였는데, 지금부터는 이에 대해 조금 자세히 알아보도록 하자. 분포의 정보를 손실 없이 전달할 수 있다고 함은 곧 적률만 알면 그 분포를 다시 복원해 낼 수 있음을 의미하는데, 기껏해야 가산개인 적률로써 $\mathscr{B}_n$ 위의 측도인 분포를 유일하게 특정해 내는 것은 결코 쉬운 일이 아니며, 이와 같은 문제를 흔히 **problem of moments**라 한다. 다음의 정리는 사실상 Fourier 역변환 공식으로 이 문제를 공략하는 좋은 출발점이 된다.

Theorem 3.80 확률공간 $(\Omega,\mathscr{F},\mathbb{P})$에서 정의된 rv. $X:\Omega\to\mathbb{R}^n$와 유계인 $B=\prod_{i=1}^n(x_i,y_i)\in\mathscr{S}_n$에 대해 $\mathbf{P}_X(\partial B)=0$이면

$$\mathbf{P}_X(B)=\lim_{M\to\infty}\frac{1}{(2\pi)^n}\int_{[-M,M]^n}\left(\prod_{i=1}^n\frac{e^{-\mathbf{i}t_ix_i}-e^{-\mathbf{i}t_iy_i}}{\mathbf{i}t_i}\right)\varphi_X(t)\,d\mu_n(t)$$

이다.

Proof 0을 성분으로 가지지 않는 임의의 $t\in\mathbb{R}^n$를 고정하면 임의의 $u\in\mathbb{R}^n$에 대해

$$\begin{aligned}\left(\prod_{i=1}^n\frac{e^{-\mathbf{i}t_ix_i}-e^{-\mathbf{i}t_iy_i}}{\mathbf{i}t_i}\right)e^{\mathbf{i}t^\top u}&=\prod_{i=1}^n\frac{e^{-\mathbf{i}t_ix_i}-e^{-\mathbf{i}t_iy_i}}{\mathbf{i}t_i}e^{\mathbf{i}t_iu_i}\\&=\prod_{i=1}^n\frac{e^{\mathbf{i}t_i(u_i-x_i)}-e^{\mathbf{i}t_i(u_i-y_i)}}{\mathbf{i}t_i}\end{aligned}$$

이므로

$$\begin{aligned}\left|\left(\prod_{i=1}^{n}\frac{e^{-\mathbf{i}t_ix_i}-e^{-\mathbf{i}t_iy_i}}{\mathbf{i}t_i}\right)e^{\mathbf{i}t^\top u}\right| &= \prod_{i=1}^{n}\frac{|e^{\mathbf{i}t_i(u_i-x_i)}-e^{\mathbf{i}t_i(u_i-y_i)}|}{|\mathbf{i}t_i|}\\ &= \prod_{i=1}^{n}\frac{|e^{\mathbf{i}t_i(u_i-x_i)}||1-e^{\mathbf{i}t_i(x_i-y_i)}|}{|t_i|}\\ &\le \prod_{i=1}^{n}\frac{|t_i(x_i-y_i)|}{|t_i|}\\ &= \prod_{i=1}^{n}(y_i-x_i)\\ &= \mu_n(B)\\ &< \infty\end{aligned}$$

이다. 여기서 첫번째 부등호는 보조정리 3.76로부터 성립한다. 이제 임의의 $M>0$을 택하여 $\mathscr{A}=\mathscr{B}_n|_{[-M,M]^n}$이라 하면 $\mathbf{P}_X\otimes\mu_n|_{\mathscr{A}}$이 유한하므로 함수 $(u,t)\mapsto[\prod_{i=1}^{n}(e^{-\mathbf{i}t_ix_i}-e^{-\mathbf{i}t_iy_i})/\mathbf{i}t_i]e^{\mathbf{i}t^\top u}$가 $\mathbf{P}_X\otimes\mu_n|_{\mathscr{A}}$-적분가능하고, 곧 Fubini의 정리로부터

$$\begin{aligned}&\int_{[-M,M]^n}\left(\prod_{i=1}^{n}\frac{e^{-\mathbf{i}t_ix_i}-e^{-\mathbf{i}t_iy_i}}{\mathbf{i}t_i}\right)\varphi_X(t)\,d\mu_n(t)\\ &\quad=\int_{[-M,M]^n}\left(\prod_{i=1}^{n}\frac{e^{-\mathbf{i}t_ix_i}-e^{-\mathbf{i}t_iy_i}}{\mathbf{i}t_i}\right)\varphi_X(t)\,d\mu_n|_{\mathscr{A}}(t)\\ &\quad=\int_{\mathbb{R}^n\times[-M,M]^n}\left(\prod_{i=1}^{n}\frac{e^{-\mathbf{i}t_ix_i}-e^{-\mathbf{i}t_iy_i}}{\mathbf{i}t_i}\right)e^{\mathbf{i}t^\top u}\,d(\mathbf{P}_X\otimes\mu_n|_{\mathscr{A}})(u,t)\\ &\quad=\int_{\mathbb{R}^n}\left(\int_{[-M,M]^n}\prod_{i=1}^{n}\frac{e^{\mathbf{i}t_i(u_i-x_i)}-e^{\mathbf{i}t_i(u_i-y_i)}}{\mathbf{i}t_i}\,d\mu_n|_{\mathscr{A}}(t)\right)d\mathbf{P}_X(u)\\ &\quad=\int_{\mathbb{R}^n}\left(\int_{[-M,M]^n}\prod_{i=1}^{n}\frac{e^{\mathbf{i}t_i(u_i-x_i)}-e^{\mathbf{i}t_i(u_i-y_i)}}{\mathbf{i}t_i}\,d\mu_n(t)\right)d\mathbf{P}_X(u)\\ &\quad=\int_{\mathbb{R}^n}\prod_{i=1}^{n}\left(\int_{[-M,M]}\frac{e^{\mathbf{i}t_i(u_i-x_i)}-e^{\mathbf{i}t_i(u_i-y_i)}}{\mathbf{i}t_i}\,d\mu_1(t_i)\right)d\mathbf{P}_X(u)\end{aligned}$$

인데, 여기서 각 $i\le n$에 대해

$$\begin{aligned}&\int_{[-M,M]}\frac{e^{\mathbf{i}t_i(u_i-x_i)}-e^{\mathbf{i}t_i(u_i-y_i)}}{\mathbf{i}t_i}\,d\mu_1(t_i)\\ &\quad=\int_{[-M,M]}\frac{\cos t(u_i-x_i)+\mathbf{i}\sin t(u_i-x_i)-\cos t(u_i-y_i)-\mathbf{i}\sin t(u_i-y_i)}{\mathbf{i}t}\,d\mu_1(t)\\ &\quad=2\int_{[0,M]}\frac{\sin t(u_i-x_i)}{t}\,d\mu_1(t)-2\int_{[0,M]}\frac{\sin t(u_i-y_i)}{t}\,d\mu_1(t)\end{aligned}$$

$$= 2\mathrm{sgn}(u_i - x_i)\int_{[0,M|u_i-x_i|]} \frac{\sin t}{t}\,d\mu_1(t) - 2\mathrm{sgn}(u_i - y_i)\int_{[0,M|u_i-y_i|]} \frac{\sin t}{t}\,d\mu_1(t)$$

이다. 표기의 편의를 위해 함수 $\mathrm{Si}:\mathbb{R}_0^+ \to \mathbb{R}$를 $\mathrm{Si}: x \mapsto \int_{[0,x]} \sin t/t\,d\mu_1(t)$로 두면 이는 FTC로부터 연속이고 $x\to\infty$일 때 $\pi/2$로 수렴하므로 곧 Si는 적당한 $N>0$에 대해 유계이다. 이로부터 $M_j\to\infty$인 임의의 수열 $\{M_j\}$를 생각하여 $j\in\mathbb{N}$가 충분히 크다고 하면

$$\begin{aligned}
&\left|\int_{[-M_j,M_j]} \frac{e^{\mathbf{i}t_i(u_i-x_i)} - e^{\mathbf{i}t_i(u_i-y_i)}}{\mathbf{i}t_i}\,d\mu_1(t_i)\right| \\
&\qquad = |2\mathrm{sgn}(u_i-x_i)\mathrm{Si}(M_j|u_i-x_i|) - 2\mathrm{sgn}(u_i-y_i)\mathrm{Si}(M_j|u_i-y_i|)| \\
&\qquad \le 2|\mathrm{Si}(M_j|u_i-x_i|)| + 2|\mathrm{Si}(M_j|u_i-y_i|)| \\
&\qquad \le 4N
\end{aligned}$$

이 되어 DCT에서

$$\begin{aligned}
&\lim_{j\to\infty}\int_{[-M_j,M_j]^n}\left(\prod_{i=1}^n \frac{e^{-\mathbf{i}t_ix_i} - e^{-\mathbf{i}t_iy_i}}{\mathbf{i}t_i}\right)\varphi_X(t)\,d\mu_n(t) \\
&\quad = \int_{\mathbb{R}^n}\lim_{j\to\infty}\prod_{i=1}^n\left(\int_{[-M_j,M_j]} \frac{e^{\mathbf{i}t_i(u_i-x_i)} - e^{\mathbf{i}t_i(u_i-y_i)}}{\mathbf{i}t_i}\,d\mu_1(t_i)\right)d\mathbf{P}_X(u) \\
&\quad = \int_{\mathbb{R}^n}\lim_{j\to\infty}\prod_{i=1}^n[2\mathrm{sgn}(u_i-x_i)\mathrm{Si}(M_j|u_i-x_i|) - 2\mathrm{sgn}(u_i-y_i)\mathrm{Si}(M_j|u_i-y_i|)]\,d\mathbf{P}_X(u) \\
&\quad = \int_{\mathbb{R}^n}\prod_{i=1}^n \pi[\mathrm{sgn}(u_i-x_i) - \mathrm{sgn}(u_i-y_i)]\,d\mathbf{P}_X(u) \\
&\quad = \int_{B^\circ}(2\pi)^n\,d\mathbf{P}_X(u) + \int_{\partial B}\prod_{i=1}^n \pi[\mathrm{sgn}(u_i-x_i) - \mathrm{sgn}(u_i-y_i)]\,d\mathbf{P}_X(u) \\
&\quad = (2\pi)^n\mathbf{P}_X(B)
\end{aligned}$$

이고, 증명이 끝난다. 여기서 마지막 등호는 $\mathbf{P}_X(\partial B) = 0$이므로 성립한다. □

위의 정리는 CF를 알면 이를 통해 그 분포를 (거의) 알 수 있음을 의미한다. 여기서 이론적인 기교를 조금 더 부리면 다음 따름정리를 얻는다.

Lemma 3.81 확률공간 $(\Omega, \mathscr{F}, \mathbb{P})$에서 정의된 rv. $X:\Omega\to\mathbb{R}^n$에 대해 $\mathbb{R}^n$의 각 축에 수직인 초평면 중에 $\mathbf{P}_X$–영집합이 아닌 것은 가산개이다.

Proof 임의의 $i\le n$에 대해 CDF F_{X_i}가 $x_0\in\mathbb{R}$에서 연속이라면 $\mathbf{P}_X(\mathbb{R}^{i-1}\times\{x_0\}\times\mathbb{R}^{n-i}) = \mathbb{P}\{X_i = x_0\} = \mathbf{P}_{X_i}\{x_0\} = F_{X_i}(x_0) - F_{X_i}(x_0-) = 0$이므로 초평면 $\pi: x_i = x_0$는 $\mathbf{P}_X$–영집합이다. 그런데 3. 32로부터 각 $i\le n$에 대해 F_{X_i}는 가산개의 불연속점만을 가지고, 곧 보조정리가 성립한다. □

Corollary 3.82 확률공간 $(\Omega, \mathscr{F}, \mathbb{P})$, $(\Omega', \mathscr{F}', \mathbb{P}')$에서 각각 정의된 rv. $X : \Omega \to \mathbb{R}^n$, $Y : \Omega' \to \mathbb{R}^n$에 대해 $X \equiv Y$일 필요충분조건은 $\varphi_X = \varphi_Y$인 것이다.

Proof 만약 $X \equiv Y$이면 $\mathbf{P}_X = \mathbf{P}_Y$에서 $\varphi_X = \varphi_Y$임이 자명하므로 역만 성립함을 보이면 된다. 이를 위해 $\varphi_X = \varphi_Y$라 하면 위의 정리로부터 유계인 $B = \prod_{i=1}^n (x_i, y_i] \in \mathscr{S}_n$에 대해 $\mathbf{P}_X(\partial B) = \mathbf{P}_Y(\partial B) = 0$이면

$$\begin{aligned}\mathbf{P}_X(B) &= \lim_{M\to\infty} \frac{1}{(2\pi)^n} \int_{[-M,M]^n} \left(\prod_{i=1}^n \frac{e^{-\mathbf{i}t_i x_i} - e^{-\mathbf{i}t_i y_i}}{\mathbf{i}t_i} \right) \varphi_X(t)\, d\mu_n(t) \\ &= \lim_{M\to\infty} \frac{1}{(2\pi)^n} \int_{[-M,M]^n} \left(\prod_{i=1}^n \frac{e^{-\mathbf{i}t_i x_i} - e^{-\mathbf{i}t_i y_i}}{\mathbf{i}t_i} \right) \varphi_Y(t)\, d\mu_n(t) \\ &= \mathbf{P}_Y(B)\end{aligned}$$

이다. 따라서 만약 집합족 $\mathscr{A} = \{B \in \mathscr{S}_n : \mathbf{P}_X(\partial B) = \mathbf{P}_Y(\partial B) = 0\}$가 $\pi-$system이고 $\mathscr{B}_n$의 생성자이면 정리 2. 29로부터 증명이 끝난다. 우선 임의의 $B, B' \in \mathscr{A}$를 생각하면 $\partial(B \cap B') \subseteq \partial B \cup \partial B'$이므로 $\mathscr{A}$가 $\pi-$system임은 거의 자명하다. 한편, 임의의 $B = \prod_{i=1}^n (x_i, y_i] \in \mathscr{S}_n$를 택하면 위의 보조정리로부터 $\mathbf{P}_X-$영집합이 아니거나 $\mathbf{P}_Y-$영집합이 아닌 $\mathbb{R}^n$의 축에 수직인 초평면이 가산개이므로 적당한 수열 $\{z_j\}, \{w_j\}$가 존재하여 각 $j \in \mathbb{N}$에 대해 $\mathbf{P}_X(\partial \prod_{i=1}^n (z_j^i, w_j^i]) = \mathbf{P}_Y(\partial \prod_{i=1}^n (z_j^i, w_j^i]) = 0$이고 $z_j \downarrow x, w_j \downarrow y$이다. 이제 각 $j \in \mathbb{N}$에 대해 $B_j = \prod_{i=1}^n (z_j^i, w_j^i] \in \mathscr{A}$로 두면 $B_j \to B$이므로 $B \in \sigma(\mathscr{A})$이고, 곧 $\mathscr{S}_n \subseteq \sigma(\mathscr{A})$에서 $\mathscr{B}_n \subseteq \sigma(\mathscr{A})$인데, 그 역의 포함관계는 자명하므로 $\mathscr{A}$가 $\mathscr{B}_n$의 생성자가 되어 증명이 끝난다. □

위의 따름정리에 의하면 CF와 분포는 정확히 일대일로 대응한다. 한편, 앞서 보인 정리 3. 77에서 확률벡터의 모든 적률을 알면 (MGF가 존재한다는 가정 하에) 그로부터 CF를 구할 수 있으므로 이상을 종합하면 problem of moments에 어느 정도 만족스러운 답이 된다. 즉, 적률을 알면 CF를 알고, 그런 CF를 가지는 분포는 유일하게 특정된다. 나아가, MGF와 CF는 이론적으로 그 역할이 비슷하므로 CF 대신 MGF를 이용하는 방법도 생각해 볼 수 있다. 다음 정리와 따름정리는 이에 대한 이론적인 전개를 담고 있다.

Theorem 3.83 확률공간 $(\Omega, \mathscr{F}, \mathbb{P})$에서 정의된 rv. $X : \Omega \to \mathbb{R}^n$에 대해 이의 모든 적률이 존재하고 적당한 0의 근방 $U \subseteq \mathbb{R}^n$가 존재하여 임의의 $t \in U$에 대해 급수

$$\sum_{j=0}^{\infty} \sum_{k_1+\cdots+k_n=j} \mathbf{E}\left(\prod_{i=1}^n X_i^{k_i} \right) \prod_{i=1}^n \frac{t_i^{k_i}}{k_i!}$$

가 절대수렴한다고 하자. 이제 확률공간 $(\Omega', \mathscr{F}', \mathbb{P}')$에서 정의된 rv. $Y : \Omega' \to \mathbb{R}^n$에 대해 Y의 모든 적률이 존재하여 X의 적률과 같다면 $X \equiv Y$이다.

PROOF 먼저 $\sum_{i=1}^{n} k_i \to \infty$이면 모든 성분이 양수인 임의의 $h \in U$에 대해

$$\mathbf{E}\left(\prod_{i=1}^{n} |X_i|^{k_i}\right) \prod_{i=1}^{n} \frac{h_i^{k_i}}{k_i!} \to 0$$

임을 보이자. 간결한 논의를 위해 $n = 2$라 하고 모든 성분이 1보다 작은 양수인 임의의 $h \in U$를 택하면

$$\sum_{j=0}^{\infty} \sum_{k_1+k_2=j} \mathbf{E}(X_1^{k_1} X_2^{k_2}) \frac{h_1^{k_1} h_2^{k_2}}{k_1! k_2!}$$

가 절대수렴하므로 $j \to \infty$이면

$$\sum_{k_1+k_2=j} \left| \mathbf{E}(X_1^{k_1} X_2^{k_2}) \frac{h_1^{k_1} h_2^{k_2}}{k_1! k_2!} \right| \to 0$$

이고, 곧 $k_1 + k_2 \to \infty$이면 $\mathbf{E}(X_1^{k_1} X_2^{k_2}) h_1^{k_1} h_2^{k_2} / k_1! k_2! \to 0$이다. ($n = 1$인 경우를 포함한 일반적인 경우에도 이와 비슷하게 하면 된다.) 이로부터 적당한 $M > 0$이 존재하여 임의의 $k \in \mathbb{N}_0^2$에 대해 $|\mathbf{E}(X_1^{k_1} X_2^{k_2}) h_1^{k_1} h_2^{k_2} / k_1! k_2!| \leq M$이고, 임의의 $\varepsilon > 0$을 택하면 적당한 $K \in \mathbb{N}$가 존재하여 $k_1 + k_2 \geq K$이면 $|\mathbf{E}(X_1^{k_1} X_2^{k_2}) h_1^{k_1} h_2^{k_2} / k_1! k_2!| < h_1 h_2 \varepsilon / 16$이다. 이제 $k_1 + k_2 \geq \max\{K + 2, K + 8M/h_1 h_2 \varepsilon, 8/\varepsilon\}$인 임의의 $k \in \mathbb{N}_0^2$를 택하고 경우를 나누어 생각한다.

만약 k_1, k_2가 모두 짝수라면 $\mathbf{E}(|X_1|^{k_1} |X_2|^{k_2}) h_1^{k_1} h_2^{k_2} / k_1! k_2! < \varepsilon$임은 자명하다. 만약 k_1은 홀수이지만 k_2는 짝수라면 적당한 $l_1 \in \mathbb{N}$에 대해 $k_1 = 2l_1 - 1$로 쓸 수 있고, 임의의 $x \in \mathbb{R}$에 대해 $|x|^{k_1} \leq x^{2l_1} \mathbf{1}_{\{|x| \geq 1\}} + \mathbf{1}_{\{|x| < 1\}} \leq x^{2l_1} + 1$이므로

$$\begin{aligned}
\left| \mathbf{E}(|X_1|^{k_1} |X_2|^{k_2}) \frac{h_1^{k_1} h_2^{k_2}}{k_1! k_2!} \right| &= \left[\int_{\mathbb{R}^n} |x_1|^{k_1} x_2^{k_2} \, d\mathbf{P}_X(x) \right] \frac{h_1^{k_1} h_2^{k_2}}{k_1! k_2!} \\
&\leq \left[\int_{\mathbb{R}^n} (x_1^{2l_1} + 1) x_2^{k_2} \, d\mathbf{P}_X(x) \right] \frac{h_1^{k_1} h_2^{k_2}}{k_1! k_2!} \\
&= \frac{2l_1}{k_1 h_1} \mathbf{E}(X_1^{2l_1} X_2^{k_2}) \frac{h_1^{2l_1} h_2^{k_2}}{(2l_1)! k_2!} + \frac{h_1^{k_1}}{k_1!} \mathbf{E}(X_2^{k_2}) \frac{h_2^{k_2}}{k_2!} \\
&\leq \frac{2}{h_1} \mathbf{E}(X_1^{2l_1} X_2^{k_2}) \frac{h_1^{2l_1} h_2^{k_2}}{(2l_1)! k_2!} + \frac{1}{k_1} \mathbf{E}(X_2^{k_2}) \frac{h_2^{k_2}}{k_2!}
\end{aligned}$$

이다. 그렇다면 $2l_1 + \sum_{i=2}^{n} k_i \geq K + 1$에서 $(2/h_1) \mathbf{E}(X_1^{2l_1} X_2^{k_2}) h_1^{2l_1} h_2^{k_2} / (2l_1)! k_2! < h_2 \varepsilon / 8 < \varepsilon / 2$이고 $k_2 \geq K$인 경우에는 $(1/k_1) \mathbf{E}(X_2^{k_2}) h_2^{k_2} / k_2! < h_1 h_2 \varepsilon / 16 k_1 < \varepsilon / 2$이며 그렇지 않은 경우에는 $k_1 > 8M / h_1 h_2 \varepsilon > 2M / \varepsilon$에서 $(1/k_1) \mathbf{E}(X_2^{k_2}) h_2^{k_2} / k_2! < M / k_1 < \varepsilon / 2$이다. 이상으로부터 k_1은 홀수이지만 k_2는 짝수라면 $\mathbf{E}(|X_1|^{k_1} |X_2|^{k_2}) h_1^{k_1} h_2^{k_2} / k_1! k_2! < \varepsilon$임을 안다. 반대로 k_1은 짝수이지만 k_2는 홀수인 경우에도 위와 비슷하게 하면 같은 결론을 얻는다. 마지막으로 k_1, k_2

모두 홀수인 경우에는 위의 과정을 비슷하게 한 번 더 반복하면 되는데, 적당한 $l_1, l_2 \in \mathbb{N}$에 대해 $k_1 = 2l_1 - 1, k_2 = 2l_2 - 1$라 하면 위의 결과로부터

$$\begin{aligned}
&\mathbf{E}(|X_1|^{k_1}|X_2|^{k_2})\frac{h_1^{k_1}h_2^{k_2}}{k_1!k_2!} \\
&\le \frac{2}{h_1}\mathbf{E}(X_1^{2l_1}|X_2|^{k_2})\frac{h_1^{2l_1}h_2^{k_2}}{(2l_1)!k_2!} + \frac{1}{k_1}\mathbf{E}(|X_2|^{k_2})\frac{h_2^{k_2}}{k_2!} \\
&\le \frac{4}{h_1h_2}\mathbf{E}(X_1^{2l_1}X_2^{2l_2})\frac{h_1^{2l_1}h_2^{2l_2}}{(2l_1)!(2l_2)!} + \frac{2}{h_1k_2}\mathbf{E}(X_1^{2l_1})\frac{h_1^{2l_1}}{(2l_1)!} + \frac{2}{h_2k_1}\mathbf{E}(X_2^{2l_2})\frac{h_2^{2l_2}}{(2l_2)!} + \frac{1}{k_1k_2}
\end{aligned}$$

임을 쉽게 보일 수 있다. 그렇다면 $2l_1 + 2l_2 \ge K$에서 $(4/h_1h_2)\mathbf{E}(X_1^{2l_1}X_2^{2l_2})h_1^{2l_1}h_2^{2l_2}/(2l_1)!(2l_2)! < \varepsilon/4$이고 $2l_1 \ge K$인 경우에는 $(2/h_1k_2)\mathbf{E}(X_1^{2l_1})h_1^{2l_1}/(2l_1)! < \varepsilon/8k_2 < \varepsilon/4$이며 그렇지 않은 경우에는 $k_2 \ge 8M/h_1h_2\varepsilon > 8M/h_1\varepsilon$에서 $(2/h_1k_2)\mathbf{E}(X_1^{2l_1})h_1^{2l_1}/(2l_1)! < 2M/h_1k_2 < \varepsilon/4$이다. 비슷하게 $(2/h_2k_1)\mathbf{E}(X_2^{2l_2})h_2^{2l_2}/(2l_2)! < \varepsilon/4$임을 알고, $1/k_1k_2 \le 1/\max\{k_1, k_2\} \le \varepsilon/4$이므로 이상으로부터 k_1과 k_2가 모두 홀수라도 $\mathbf{E}(|X_1|^{k_1}|X_2|^{k_2})h_1^{k_1}h_2^{k_2}/k_1!k_2! < \varepsilon$임을 안다. 곧 어떠한 경우에도 $\mathbf{E}(|X_1|^{k_1}|X_2|^{k_2})h_1^{k_1}h_2^{k_2}/k_1!k_2! < \varepsilon$이 되어 $k_1 + k_2 \to \infty$이면 $\mathbf{E}(|X_1|^{k_1}|X_2|^{k_2})h_1^{k_1}h_2^{k_2}/k_1!k_2! \to 0$임을 안다. 나아가, $h \in U$가 0인 성분을 가지더라도 이가 성립함을 위와 비슷하게 보일 수 있다.

이제 보조정리 3.76로부터 임의의 $K \in \mathbb{N}$와 임의의 $t \in \mathbb{R}^n$, 임의의 $h \in U$에 대해

$$\begin{aligned}
&\left|\varphi_X(t+h) - \sum_{j=0}^{K}\frac{\mathbf{i}^j}{j!}\int_{\mathbb{R}^n}(h^\top x)^j e^{\mathbf{i}t^\top x}\,d\mathbf{P}_X(x)\right| \\
&\qquad = \left|\mathbf{E}(e^{\mathbf{i}(t+h)^\top X}) - \sum_{j=0}^{K}\frac{1}{j!}\mathbf{E}((\mathbf{i}h^\top X)^j e^{\mathbf{i}t^\top X})\right| \\
&\qquad = \left|\mathbf{E}\left(e^{\mathbf{i}(t+h)^\top X} - \sum_{j=0}^{K}\frac{1}{j!}(\mathbf{i}h^\top X)^j e^{\mathbf{i}t^\top X}\right)\right| \\
&\qquad \le \mathbf{E}\left(\left|e^{\mathbf{i}(t+h)^\top X} - \sum_{j=0}^{K}\frac{1}{j!}(\mathbf{i}h^\top X)^j e^{\mathbf{i}t^\top X}\right|\right) \\
&\qquad = \mathbf{E}\left(|e^{\mathbf{i}t^\top X}|\left|e^{\mathbf{i}h^\top X} - \sum_{j=0}^{K}\frac{1}{j!}(\mathbf{i}h^\top X)^j\right|\right) \\
&\qquad = \mathbf{E}\left(\left|e^{\mathbf{i}h^\top X} - \sum_{j=0}^{K}\frac{1}{j!}(\mathbf{i}h^\top X)^j\right|\right) \\
&\qquad \le \mathbf{E}\left(\frac{(|h|^\top|X|)^{K+1}}{(K+1)!}\right) \\
&\qquad = \frac{1}{(K+1)!}\sum_{k_1+\cdots+k_n=K+1}\binom{K+1}{k_1\ \cdots\ k_n}\mathbf{E}\left(\prod_{i=1}^{n}|h_i|^{k_i}|X_i|^{k_i}\right)
\end{aligned}$$

$$= \sum_{k_1+\cdots+k_n=K+1} \mathbf{E}\left(\prod_{i=1}^n |X_i|^{k_i}\right) \prod_{i=1}^n \frac{|h_i|^{k_i}}{k_i!}$$

이므로 앞선 결론으로부터 $K \to \infty$이면 위 식의 우변이 0으로 수렴하여 정리 3. 78로부터

$$\begin{aligned}
\varphi_X(t+h) &= \sum_{j=0}^\infty \frac{\mathbf{i}^j}{j!} \int_{\mathbb{R}^n} (h^\top x)^j e^{\mathbf{i}t^\top x} d\mathbf{P}_X(x) \\
&= \sum_{j=0}^\infty \frac{1}{j!} \sum_{k_1+\cdots+k_n=j} \binom{j}{k_1 \cdots k_n} \mathbf{i}^j \int_{\mathbb{R}^n} \left[\prod_{i=1}^n (h_i x_i)^{k_i}\right] e^{\mathbf{i}t^\top x} d\mathbf{P}_X(x) \\
&= \sum_{j=0}^\infty \frac{1}{j!} \sum_{k_1+\cdots+k_n=j} \binom{j}{k_1 \cdots k_n} \mathbf{i}^j \mathbf{E}\left(\left[\prod_{i=1}^n X_i^{k_i}\right] e^{\mathbf{i}t^\top X}\right) \prod_{i=1}^n h_i^{k_i} \\
&= \sum_{j=0}^\infty \frac{1}{j!} \sum_{k_1+\cdots+k_n=j} \binom{j}{k_1 \cdots k_n} \frac{\partial^j}{\partial t_1^{k_1} \cdots \partial t_n^{k_n}} \varphi_X(t) \prod_{i=1}^n h_i^{k_i} \\
&= \sum_{j=0}^\infty \frac{1}{j!} \mathbf{D}^j \varphi_X(t)(h, \cdots, h)
\end{aligned}$$

이다. 이는 Y에 대해서도 성립하므로 임의의 $t \in \mathbb{R}^n$와 임의의 $h \in U$에 대해 $\varphi_Y(t+h) = \sum_{j=0}^\infty \mathbf{D}^j \varphi_Y(t)(h, \cdots, h)/j!$인데, X와 Y의 적률이 모두 같으므로 다시 정리 3. 78로부터 임의의 $j \in \mathbb{N}$에 대해 $\mathbf{D}^j \varphi_X(0) = \mathbf{D}^j \varphi_Y(0)$이고, 곧 U 위에서 φ_X와 φ_Y가 일치한다. 그런데 위의 결론에서 $t \in \mathbb{R}^n$가 임의의 점이었음을 상기한다면 이는 곧 $\varphi_X = \varphi_Y$임을 의미하여 곧 따름정리 3. 82로부터 $X \equiv Y$이다. □

Corollary 3.84 확률공간 $(\Omega, \mathscr{F}, \mathbb{P})$, $(\Omega', \mathscr{F}', \mathbb{P}')$에서 각각 정의된 rv. $X : \Omega \to \mathbb{R}^n$, $Y : \Omega' \to \mathbb{R}^n$에 대해 적당한 0의 근방 $U \subseteq \mathbb{R}^n$에서 M_X가 존재한다고 하자. 그렇다면 $X \equiv Y$일 필요충분조건은 적당한 0의 근방 $V \subseteq \mathbb{R}^n$에서 M_Y가 존재하여 $U \cap V$에서 $M_X = M_Y$인 것이다.

Proof 만약 $X \equiv Y$이면 $\mathbf{P}_X = \mathbf{P}_Y$에서 정리가 자명하므로 역만 성립함을 보이면 된다. 이를 위해 적당한 0의 근방 $V \subseteq \mathbb{R}^n$에서 M_Y가 존재하여 $U \cap V$에서 $M_X = M_Y$라 하자. 그렇다면 정리 3. 71와 3. 72로부터 X와 Y의 모든 적률이 존재하여 서로 같으며

$$\sum_{j=0}^\infty \sum_{k_1+\cdots+k_n=j} \mathbf{E}\left(\prod_{i=1}^n X_i^{k_i}\right) \prod_{i=1}^n \frac{t_i^{k_i}}{k_i!}$$

이 적당한 0의 근방에서 절대수렴하므로 위의 정리로부터 $X \equiv Y$이다. □

Problem of moments에 대한 논의를 마무리하기 전에, 한 가지 유의해야 할 점이 있다. 우리가 적률을 통해 MGF나 CF를 계산할 수 있고, 곧 분포를 유일하게 특정할 수 있지만, 이는 어디까지나 MGF가 존재할 때에만 가능한 것이다. 따라서 어떤 확률벡터의 MGF가 존재하지 않는다면 적률을 모두 알아도 이로부터 분포를 특정지을 수가 없다. 즉, 그러한 적률을 만들어내는 서로다른 분포가 얼마든지 존재할 수 있다.

다음으로 살펴볼 CF에 대한 흥미로운 주제는 연속확률벡터와 CF의 관계이다. 앞서 정리 3. 37는 어떤 확률벡터가 연속인지를 판단할 수 있는 방법을 제시했는데, 다음 정리는 CF를 이용한 방법을 하나 새로 제시한다.

Theorem 3.85 (Riemann-Lebesgue) 확률공간 $(\Omega, \mathscr{F}, \mathbb{P})$에서 정의된 연속활률벡터 $X : \Omega \to \mathbb{R}^n$에 대해 $\lim_{|t_1|,\cdots,|t_n| \to \infty} \varphi_X(t) \to 0$이다.

Proof 임의의 $\varepsilon > 0$을 택하면 $\int_{\mathbb{R}^n} f_X \, d\mu_n = 1$이므로 정리 2. 180의 i로부터 적당한 단순함수 $g : \mathbb{R}^n \to \mathbb{R}$가 존재하여 이는 적당한 $a_1, \cdots, a_k \in \mathbb{R}$와 유계인 $B_1, \cdots, B_k \in \mathscr{S}_n$에 대해 $g = \sum_{i=1}^n a_i \mathbf{1}_{B_i}$로 쓸 수 있으며 $\int_{\mathbb{R}^n} |f_X - g| \, d\mu_n < \varepsilon$이다. 한편, 각 $i \le k$에 대해 적당한 $x_i, y_i \in \mathbb{R}^n$에 대해 $B_i = \prod_{j=1}^n (x_i^j, y_i^j]$이다. 이로부터 임의의 $t \in \mathbb{R}^n$에 대해 $|\int_{\mathbb{R}^n} [f_X(x) - g(x)] e^{\mathbf{i} t^\top x} \, d\mu_n(x)| \le \int_{\mathbb{R}^n} |[f_X(x) - g(x)] e^{\mathbf{i} t^\top x}| \, d\mu_n(x) = \int_{\mathbb{R}^n} |f_X(x) - g(x)| \, d\mu_n(x) < \varepsilon/2$이다. 또한 Fubini의 정리로부터 0을 성분으로 갖지 않는 임의의 $t \in \mathbb{R}^n$에 대해

$$
\begin{aligned}
\left| \int_{\mathbb{R}^n} g(x) e^{\mathbf{i} t^\top x} \, d\mu_n(x) \right| &= \left| \int_{\mathbb{R}^n} \sum_{i=1}^k a_i \mathbf{1}_{B_i} e^{\mathbf{i} t^\top x} \, d\mu_n(x) \right| \\
&= \left| \sum_{i=1}^k a_i \int_{B_i} \prod_{j=1}^n e^{\mathbf{i} t_j x_j} \, d\mu_n(x) \right| \\
&= \left| \sum_{i=1}^k a_i \prod_{j=1}^n \int_{x_i^j}^{y_i^j} e^{\mathbf{i} t_j x_j} \, d\mu_1(x_j) \right| \\
&= \left| \sum_{i=1}^k a_i \prod_{j=1}^n \frac{e^{\mathbf{i} t_j y_i^j} - e^{\mathbf{i} t_j x_i^j}}{\mathbf{i} t_j} \right| \\
&\le \sum_{i=1}^k |a_i| \prod_{j=1}^n \left| \frac{e^{\mathbf{i} t_j y_i^j} - e^{\mathbf{i} t_j x_i^j}}{\mathbf{i} t_j} \right| \\
&\le \sum_{i=1}^k |a_i| \prod_{j=1}^n \frac{|e^{\mathbf{i} t_j y_i^j}| + |e^{\mathbf{i} t_j x_i^j}|}{|t_j|} \\
&= 2^n \sum_{i=1}^k |a_i| \Big/ \prod_{j=1}^n |t_j|
\end{aligned}
$$

이므로 $\lim_{|t_1|,\cdots,|t_n| \to \infty} |\int_{\mathbb{R}^n} g(x) e^{\mathbf{i} t^\top x} \, d\mu_n(x)| = \infty$에서 적당한 $M > 0$이 존재하여 $|t_1|, \cdots, |t_n| \ge M$이면 $|\int_{\mathbb{R}^n} g(x) e^{\mathbf{i} t^\top x} \, d\mu_n(x)| < \varepsilon/2$이다. 이상으로부터 $|t_1|, \cdots, |t_n| \ge M$이면 $|\varphi_X(t)| = |\int_{\mathbb{R}^n} f_X(x) e^{\mathbf{i} t^\top x} \, d\mu_n(x)| \le |\int_{\mathbb{R}^n} [f_X(x) - g(x)] e^{\mathbf{i} t^\top x} \, d\mu_n(x)| + |\int_{\mathbb{R}^n} g(x) e^{\mathbf{i} t^\top x} \, d\mu_n(x)| < \varepsilon$이 되어 증명이 끝난다. □

심지어 CF를 이용해 PDF를 바로 구할 수도 있다.

Theorem 3.86 확률공간 $(\Omega, \mathscr{F}, \mathbb{P})$에서 정의된 연속확률벡터 $X : \Omega \to \mathbb{R}$에 대해 $\int_{\mathbb{R}^n} |\varphi_X| \, d\mu_n < \infty$이면

$$f_X(x) = \frac{1}{(2\pi)^n} \int_{\mathbb{R}^n} e^{-\mathbf{i} t^\top x} \varphi_X(t) \, d\mu_n(t)$$

이다.

이제 앞서 기댓값에 대한 다양한 부등식을 소개할 때 미처 소개하지 못한 부등식을 하나 소개하고 MGF와 CF에 대한 내용을 마무리하자. 이 부등식도 tail probability를 bounding 하는 부등식이다.

Theorem 3.87 (Chernoff's bound) 확률공간 $(\Omega, \mathscr{F}, \mathbb{P})$에서 정의된 rv. $X : \Omega \to \mathbb{R}$에 대해 0의 적당한 근방 $U \in \mathbb{R}$에서 M_X가 존재하면 임의의 $x \in \mathbb{R}$에 대해 $\mathbb{P}\{X \geq x\} \leq \inf_{0 \leq t \in U} e^{-xt} M_X(t)$이다. 비슷하게, 임의의 $x \in \mathbb{R}$에 대해 $\mathbb{P}\{X \leq x\} \leq \inf_{0 \geq t \in U} e^{-xt} M_X(t)$이다.

PROOF 임의의 $x, y \in \mathbb{R}$와 임의의 음이 아닌 $t \in U$에 대해 $\mathbf{1}_{[x,\infty)}(y) \leq e^{t(y-x)}$이므로

$$\begin{aligned}
\mathbb{P}\{X \geq x\} &= \mathbf{P}_X([x,\infty)) \\
&= \int_{\mathbb{R}} \mathbf{1}_{[x,\infty)} \, d\mathbf{P}_X \\
&\leq \int_{\mathbb{R}} e^{t(y-x)} \, d\mathbf{P}_X(y) \\
&= \mathbf{E}(e^{t(X-x)}) \\
&= e^{-xt} M_X(t) \\
&\leq \inf_{0 \leq t \in U} e^{-xt} M_X(t)
\end{aligned}$$

이다. 이제 임의의 $x \in \mathbb{R}$에 대해 $\mathbb{P}\{X \leq x\} \leq \inf_{0 \geq t \in U} e^{-xt} M_X(t)$임도 비슷하게 보일 수 있다. □

이번 절의 후반부에서는 MGF와 관련되거나 혹은 유사한 함수들을 간단히 살펴본다. 먼저 첫째로 볼 것은 MGF에 로그를 취하여 얻는 CGF, 그리고 이의 Taylor 전개의 계수로 주어지는 누율이다.

Definition 3.88 확률공간 $(\Omega, \mathscr{F}, \mathbb{P})$에서 정의된 rv. $X : \Omega \to \mathbb{R}^n$에 대해 $M_X : U \to \mathbb{R}$가 존재하면 rv. X의 **누율생성함수(cumulant generating function)**를 $C_X : U \to \mathbb{R}$로 쓰고 $C_X = \log M_X$로 정의한다. 특별히, $n \geq 2$인 경우 C_X를 rv. $X_1, \cdots, X_n$의 **결합누율생성함수(joint cumulant generating function)**라 하기도 한다.

Proposition 3.89 확률공간 $(\Omega, \mathscr{F}, \mathbb{P})$에서 정의된 rv. $X : \Omega \to \mathbb{R}^n$ 적당한 0의 근방 $U \subseteq \mathbb{R}^n$에서 C_X가 존재한다면 C_X는 U에서 해석적이다.

Proof 가정으로부터 U에서 M_X가 존재하고 정리 3. 71로부터 M_X는 U에서 해석적이다. 한편, 함수 $x \mapsto \log x$가 $(0, \infty)$에서 해석적이므로 $C_X = \log M_X$도 U에서 해석적이다. □

Definition 3.90 확률공간 $(\Omega, \mathscr{F}, \mathbb{P})$에서 정의된 rv. $X : \Omega \to \mathbb{R}^n$에 대해 적당한 0의 근방 $U \subseteq \mathbb{R}^n$에서 C_X가 존재한다고 하고 임의의 $k \in \mathbb{N}_0^n$에 대해 $K = \sum_{i=1}^n k_i$라 하자. 이때 $C_X(t)$의 0에서의 Taylor 전개에서 $\prod_{i=1}^n t_i^{k_i}/k_i!$의 계수를 X의 K차 **(결합)누율(kth (joint) cumulant)**이라 하고 $c_{k,X}$ 혹은 간단히 c_k로 쓴다.

Theorem 3.91 확률공간 $(\Omega, \mathscr{F}, \mathbb{P})$에서 정의된 rv. $X : \Omega \to \mathbb{R}^n$에 대해 적당한 0의 근방 $U \subseteq \mathbb{R}^n$에서 C_X가 존재한다고 하면 임의의 $k \in \mathbb{N}_0^n$에 대해 $K = \sum_{i=1}^n k_i$라 하면 $c_{X,k} = (\partial^K/\partial t_1^{k_1} \cdots \partial t_n^{k_n})C_X(0)$이다. 이때, 편미분의 순서는 중요하지 않다.

Proof 이는 누율의 정의로부터 자명하다. □

CGF의 기본적인 성질들은 MGF의 성질들로부터 거의 자명하게 얻어진다.

Theorem 3.92 확률공간 $(\Omega, \mathscr{F}, \mathbb{P})$에서 정의된 rv. $X : \Omega \to \mathbb{R}^n$에 대해 적당한 0의 근방 $U \subseteq \mathbb{R}^n$에서 C_X가 존재하면 다음이 성립한다.

i. $C_X(0) = 0$.
ii. CGF C_X는 볼록하다.
iii. 임의의 $a \in \mathbb{R}$와 임의의 $b \in \mathbb{R}^n$에 대해 C_{aX+b}가 적당한 0의 근방 $V \subseteq \mathbb{R}^n$에서 존재하고, $at \in U$인 $t \in V$에 대해 $C_{aX+b}(t) = C_X(at) + t^\mathsf{T} b$이다.

Proof i. 이는 $C_X(0) = \log M_X(0) = 0$에서 자명하다.

ii. 정리 3. 69의 iii으로부터 이는 자명하다.

iii. 정리 3. 69의 iv로부터 적당한 0의 근방 $V \subseteq \mathbb{R}^n$에서 C_{aX+b}가 존재하고 $at \in U$인 $t \in V$에 대해 $C_{aX+b}(t) = \log M_{aX+b}(t) = \log e^{t^\mathsf{T} b} M_X(at) = C_X(at) + t^\mathsf{T} b$이다. □

앞서 배운 적률이 확률변수의 거듭제곱의 기댓값으로 깔끔하게 주어진 것에 비해, 누율과 기댓값을 깔끔하게 연결짓는 것은 쉽지 않다. 다만, 몇몇 낮은 차수의 누율의 경우 특별한 의미를 가진다. 그리고 때로는 이러한 관계를 잘 사용하면 기댓값이나 분산 등을 복잡한 계산을 피해 쉽고 간단하게 구할 수 있다.

Proposition 3.93 확률공간 $(\Omega, \mathscr{F}, \mathbb{P})$에서 정의된 rv. $X : \Omega \to \mathbb{R}$에 대해 적당한 0의 근방 $U \subseteq \mathbb{R}$에서 C_X가 존재한다고 하면 $c_{X,1} = \mathbf{E}(X)$이고 $c_{X,2} = \mathbf{Var}(X)$이다.

PROOF 이는 정리 3. 72으로부터 $C'_X(0) = (\log M_X)'(0) = (M'_X/M_X)(0) = \mathbf{E}(X)$이고

$$\begin{aligned}
C''_X(0) &= (\log M_X)''(0) \\
&= \left(\frac{M'_X}{M_X}\right)'(0) \\
&= \left[\frac{M''_X M_X - (M'_X)^2}{M_X^2}\right](0) \\
&= \mathbf{E}(X^2) - [\mathbf{E}(X)]^2 \\
&= \mathbf{Var}(X)
\end{aligned}$$

이므로 자명하다. □

Proposition 3.94 확률공간 $(\Omega, \mathscr{F}, \mathbb{P})$에서 정의된 rv. $X : \Omega \to \mathbb{R}$에 대해 적당한 0의 근방 $U \subseteq \mathbb{R}$에서 C_X가 존재하며 $\mathbf{Var}(X) \neq 0$이라 하자. 이제 $Z = (X - \mu_X)/\sigma_X$라 하면 적당한 0의 근방 $V \subseteq \mathbb{R}$에서 C_Z가 존재하고 $c_{Z,1} = 0, c_{Z,2} = 1, c_{Z,3} = \mathbf{Skew}(X), c_{Z,4} = \mathbf{Kurt}(X)$이다.

PROOF 정리 3. 92의 iii으로부터 적당한 0의 근방 $V \subseteq \mathbb{R}$에서 C_Z가 존재함을 안다. 이제 정리 3. 72으로부터

$$\begin{aligned}
M_Z(0) &= 1 \\
M'_Z(0) &= \mathbf{E}\left(\frac{X-\mu_X}{\sigma_X}\right) = \frac{\mathbf{E}(X)-\mu_X}{\sigma_X} = 0 \\
M''_Z(0) &= \mathbf{E}\left(\left(\frac{X-\mu_X}{\sigma_X}\right)^2\right) = \frac{\mathbf{Var}(X)}{\sigma_X^2} = 1, \\
M'''_Z(0) &= \mathbf{E}\left(\left(\frac{X-\mu_X}{\sigma_X}\right)^3\right) = \mathbf{Skew}(X) \\
M^{(4)}_Z(0) &= \mathbf{E}\left(\left(\frac{X-\mu_X}{\sigma_X}\right)^4\right) = \mathbf{Kurt}(X) + 3
\end{aligned}$$

이다. 그렇다면 명제 3. 93로부터 $C_{Z,1} = \mathbf{E}(Z) = 0, C_{Z,2} = \mathbf{Var}(Z) = 1$이고

$$\begin{aligned}
C'''_Z(0) &= (\log M_Z)'''(0) \\
&= \left(\frac{M'_Z}{M_Z}\right)''(0) \\
&= \left[\frac{M''_Z M_Z - (M'_Z)^2}{M_Z^2}\right]'(0) \\
&= \left\{\frac{(M'''_Z M_Z + M''_Z M'_Z - 2M'_Z M''_Z)M_Z^2 - [M''_Z M_Z - (M'_Z)^2]2M_Z M'_Z}{M_Z^4}\right\}(0)
\end{aligned}$$

$$
\begin{aligned}
&= \left[\frac{M_Z'''M_Z^2 - 3M_Z''M_Z'M_Z + 2(M_Z')^3}{M_Z^3}\right](0) \\
&= \mathbf{Skew}(X)
\end{aligned}
$$

이며

$$
\begin{aligned}
C_Z^{(4)}(0) &= (\log M_Z)^{(4)}(0) \\
&= \left(\frac{M_Z'}{M_Z}\right)'''(0) \\
&= \left[\frac{M_Z''M_Z - (M_Z')^2}{M_Z^2}\right]''(0) \\
&= \left[\frac{M_Z'''M_Z^2 - 3M_Z''M_Z'M_Z + 2(M_Z')^3}{M_Z^3}\right]'(0) \\
&= \left\{\frac{\begin{array}{l}[M_Z^{(4)}M_Z^2 + M_Z'''2M_ZM_Z' - 3M_Z'''M_Z'M_Z - 3(M_Z'')^2M_Z \\ \quad - 3M_Z''(M_Z')^2 + 6(M_Z')^2M_Z'']M_Z^3 - [M_Z'''M_Z^2 - 3M_Z''M_Z'M_Z + 2(M_Z')^3]3M_Z^2M_Z'\end{array}}{M_Z^6}\right\}(0) \\
&= \left[\frac{M_Z^{(4)}M_Z^3 - 4M_Z'''M_Z'M_Z^2 + 12M_Z''(M_Z')^2M_Z - 3(M_Z'')^2M_Z^2 - 6(M_Z')^4}{M_Z^4}\right](0) \\
&= \mathbf{Kurt}(X)
\end{aligned}
$$

이다. □

Notes

1 여기서 $\lim_{x_i \to -\infty} F_X(x) = 0$은 임의의 $\cdots, x_{i-1}, x_{i+1}, \cdots \in \mathbb{R}$와 임의의 $\varepsilon > 0$에 대해 적당한 $M \in \mathbb{R}$이 존재하여 $x_i < M$인 임의의 $x_i \in \mathbb{R}$에 대해 $F_X(x) < \varepsilon$이라는 의미이다.

2 여기서 $\lim_{x_1, \cdots, x_n \to \infty} F_X(x) = 1$은 임의의 $\varepsilon > 0$에 대해 적당한 $M \in \mathbb{R}$이 존재하여 $x \geq M\mathbf{1}$인 임의의 $x \in \mathbb{R}^n$에 대해 $|F_\mu(x) - 1| < \varepsilon$이라는 의미이다.

3 집합 $A \subseteq \mathbb{R}^m$에서 정의된 함수 $f : A \to \mathbb{R}^n$와 한 점 $x_0 \in \mathrm{acc}A$에 대해 극한 $\lim_{x \uparrow x_0} f(x)$가 존재하는 경우 그 값을 $f(x_0-)$로 쓴다. 한편, 여기서 $\lim_{x \uparrow x_0} f(x) = f(x_0-)$는 임의의 $\varepsilon > 0$에 대해 적당한 $\delta > 0$가 존재하여 $||x - x_0|| < \delta$이고 $x < x_0$인 임의의 $x \in A$에 대해 $||f(x) - f(x_0-)|| < \varepsilon$이라는 의미이다. 이제 $f(x_0+)$도 비슷하게 정의된다.