

Probability: Theory and Examples

概率论：理论与例子

更新日期：2025 年 03 月 11 日

aytony

目录

1. 测度论	3
1.1. 测度空间	3
1.2. 分布	7
1.3. 随机变量	9
1.4. 积分	13
1.5. 积分的性质	17
1.6. 期望	22
1.7. 乘积测度和 Fubini 定理	31
2. 大数定律	36
2.1. 独立性	36
2.2. 弱大数定律	44
2.3. Borel-Cantelli 引理	52
2.4. 强大数定律	63
2.5. 随机级数的收敛性	66
2.6. 更新过程	74
2.7. 大偏差	76
3. 中心极限定理	77
3.1. De Moivre-Laplace 定理	77
3.2. 弱收敛	79
3.3. 特征函数	86
3.3.1. 定义, 逆转公式	86
3.3.2. 弱收敛	89
3.3.3. 矩和导数	90
3.3.4. Polya 准则	91
3.3.5. 矩问题	92
4. 鞅	93
4.1. 条件期望	93
4.2. 鞅、几乎必然收敛	97

1. 测度论

1.1. 测度空间

定义 1 (σ -域、 σ -代数): 设 \mathcal{F} 是集合 Ω 上的集合系, 称 \mathcal{F} 是 Ω 上的 σ -域或 σ -代数, 如果

1. 对 $A \in \mathcal{F}$ 有 $A^c \in \mathcal{F}$;
2. 如果 $\{A_i \in \mathcal{F}\}$ 是 Ω 的可数个子集, 那么 $\bigcup_i A_i \in \mathcal{F}$ 。

定义 2 (测度空间): 设 Ω 是集合, \mathcal{F} 是 Ω 上的 σ -域, 称二元组 (Ω, \mathcal{F}) 是一个测度空间。

定义 3 (测度、概率测度): 设 \mathcal{F} 是集合 Ω 上的集合系。

称 μ 是集合系 \mathcal{F} 上的测度, 如果

1. μ 是从 \mathcal{F} 到 \mathbb{R} 的函数;
2. $\mu(A) \geq \mu(\emptyset) = 0$ 对 $\forall A \in \mathcal{F}$ 成立;
3. 对 \mathcal{F} 中的可数无交集列 $\{A_i\}$ 有

$$\mu\left(\bigcup_i A_i\right) = \sum_i \mu(A_i).$$

称 μ 为集合系 \mathcal{F} 上的概率测度, 如果 μ 是集合系 \mathcal{F} 的测度且 $\mu(\Omega) = 1$ 。

定义 4 (概率空间): 设 Ω 是集合, \mathcal{F} 是 Ω 上的 σ -域, P 是 \mathcal{F} 上的概率测度, 称三元组 (Ω, \mathcal{F}, P) 是一个概率空间。

命题 1 (测度的性质): 设 μ 是测度空间 (Ω, \mathcal{F}) 上的测度。

1. μ 有单调性, 即对 $A \subset B$ 有 $\mu(A) \leq \mu(B)$;
2. μ 有次可数可加性, 即对 $A \subset \bigcup_{i=1}^{\infty} A_i$ 有

$$\mu(A) \leq \sum_{i=1}^{\infty} \mu(A_i),$$

3. μ 有下连续性, 即对 $A_i \uparrow A$ 有 $\mu(A_i) \uparrow \mu(A)$;
4. μ 有上连续性, 即对 $A_i \downarrow A$, 如果 $\mu(A_1) < \infty$ 则有 $\mu(A_i) \downarrow \mu(A)$ 。

命题 2 (σ -域对交的封闭性): 设 $\{\mathcal{F}_i\}_{i \in I}$ 是 Ω 上的 σ -域, 那么 $\bigcap_{i \in I} \mathcal{F}_i$ 也是 Ω 上的 σ -域。

定义 5 (生成的 σ -域): 设 Ω 是集合, \mathcal{A} 是 Ω 上的集合系, 由命题 2 知道

$$\bigcap \mathcal{F}, \quad \mathcal{F} \text{ 是包含 } \mathcal{A} \text{ 的 } \sigma\text{-域}$$

也是 σ -域, 称为 \mathcal{A} 生成的 σ -域, 记作 $\sigma(\mathcal{A})$ 。

定义 6 (Borel 集): 称 \mathbb{R}^d 上所有开集生成的 σ -域为 **Borel 集**, 记作 \mathcal{R}^d 。

定义 7 (半环、半代数): 设 Ω 是集合, \mathcal{S} 是 Ω 上的集合系, 称 \mathcal{S} 是 Ω 上的半代数或半环, 如果

1. 对 $\forall A, B \in \mathcal{S}$ 有 $A \cap B \in \mathcal{S}$;
2. 对 $\forall A \in \mathcal{S}$ 那么 A^c 可以表示成 \mathcal{S} 中有限个元素的无交并。

定义 8 (代数、域): 设 Ω 是集合, \mathcal{F} 是 Ω 上的集合系, 称 \mathcal{F} 是 Ω 上的代数或域, 如果

1. 对 $\forall A, B \in \mathcal{F}$ 有 $A \cup B \in \mathcal{F}$;
2. 对 $\forall A \in \mathcal{F}$ 有 $A^c \in \mathcal{F}$ 。

命题 3 (代数对交的封闭性): 设 \mathcal{F} 是 Ω 上的代数, 那么 $\forall A, B \in \mathcal{F}$ 有 $A \cap B \in \mathcal{F}$ 。

命题 4 (σ -代数是代数): 设 \mathcal{F} 是 Ω 上的 σ -代数, 那么 \mathcal{F} 是 Ω 上的代数。

命题 5 (半环可以生成环): 设 \mathcal{S} 是 Ω 上的半环, 那么

$$\overline{\mathcal{S}} = \{\mathcal{S} \text{ 上有限元素的无交并}\}$$

是 Ω 上的环。

定义 9 (生成的环): 设 \mathcal{S} 是 Ω 上的半环, 那么 $\overline{\mathcal{S}}$ 称为由 \mathcal{S} 生成的环。

定义 10 (σ -有限): 设 \mathcal{A} 是 Ω 上的集合系, μ 是 \mathcal{A} 上的测度, 称 μ 是 σ -有限的, 如果存在可数集合列 $\{A_i \in \mathcal{A}\}$ 使得 $\mu(A_i) < \infty$ 且 $\bigcup_i A_i = \Omega$ 。

命题 6 (测度的扩张): 设 \mathcal{S} 是半代数, μ 是 \mathcal{S} 上的函数, 满足 $\mu(\emptyset) = 0$. 如果

1. 对 \mathcal{S} 中的有限个无交元素 $S_i \in \mathcal{S}$ 满足 $\bigsqcup_i S_i = S \in \mathcal{S}$ 均有

$$\sum_i \mu(S_i) = \mu(S),$$

2. 对 \mathcal{S} 中的可数个无交元素 $S_i \in \mathcal{S}$ 满足 $\bigsqcup_i S_i = S \in \mathcal{S}$ 均有

$$\sum_i \mu(S_i) \leq \mu(S),$$

那么 μ 能够唯一地扩张到 \mathcal{S} 生成的代数 $\overline{\mathcal{S}}$ 上. 如果 $\overline{\mathcal{S}}$ 上的测度 $\bar{\mu}$ 还是 σ -有限的, 那么它还能够唯一地扩张到 $\sigma(\mathcal{S})$ 上.

练习 1.1.1: Let $\Omega = \mathbb{R}$, $\mathcal{F} =$ all subsets so that A or A^c is countable, $P(A) = 0$ in the first case and $= 1$ in the second. show that (Ω, \mathcal{F}, P) is a probability space.

证明: 只要证明 \mathcal{F} 是 Ω 上的 σ -域且 P 是 \mathcal{F} 上的测度.

首先有 \emptyset 可数, 从而 $\emptyset, \mathbb{R} \in \mathcal{F}$. 考虑 \mathcal{F} 中的集合列 $\{A_i \in \mathcal{F}\}$. 若存在 k 使得 A_k^c 可数, 那么 $|\bigcup A_i| \geq |A_k|$ 从而 $\bigcup A_i$ 的补集可数; 若不存在上述 A_k^c , 那么所有的 A_i 均为可数集合, 从而它们的可数并可数. 所以 \mathcal{F} 是 Ω 上的 σ -域.

设 $\{A_i\}$ 是 \mathcal{F} 上的可数个不交集合列. 容易看出 $\{A_i\}$ 中存在至多一个不可数集 A_k , 那么若 A_k 存在则 $P(\bigcup A_i) = \sum P(A_i) = 1$; 否则 $P(\bigcup A_i) = \sum P(A_i) = 0$. \square

练习 1.1.2: Rescri the definition of \mathcal{S}_d from example 1.1.5. show that $\sigma(\mathcal{S}_d) = \mathcal{R}^d$, the borel subsets of \mathbb{R}^d .

证明: 对于 \mathcal{R}^d 的一组拓扑基

$$\{(x_1, y_1) \times (x_2, y_2) \cdots \times (x_d, y_d) : x_i, y_i \in \mathbb{R}\}$$

中的一个开集

$$(x_1, y_2) \times \cdots \times (x_d, y_d),$$

有

$$(x_1, y_2) \times \cdots \times (x_d, y_d) = \bigcup_j \left(\left(x_1, y_1 - \frac{1}{j} \right] \times \cdots \times \left(x_d, y_d - \frac{1}{j} \right] \right) \in \mathcal{S}_d,$$

从而 $\sigma(\mathcal{S}_d) \subset \sigma(\mathcal{R}^d)$. 对 \mathcal{S}_d 中的拓扑基

$$(x_1, y_2] \times \cdots \times (x_d, y_d],$$

有

$$(x_1, y_2] \times \cdots \times (x_d, y_d] = \bigcap_j \left(\left(x_1, y_1 + \frac{1}{j} \right) \times \cdots \times \left(x_d, y_d + \frac{1}{j} \right) \right) \in \mathcal{R}_d,$$

从而 $\sigma(\mathcal{R}^d) \subset \sigma(\mathcal{S}_d)$ 。

□

练习 1.1.3: A σ -field \mathcal{F} is said to be countably generated if there is a countable collection $\mathcal{C} \subset \mathcal{F}$ so that $\sigma(\mathcal{C}) = \mathcal{F}$. Show that \mathcal{R}^d is countably generated.

证明: 只要证明 \mathcal{R}^d 可以由 $\{(a, b] : a, b \in \mathbb{Q}\}$ 生成, 那么只要证明 $\{(a, b] : a, b \in \mathbb{R}\}$ 可以被它生成, 而对于 $\forall (a, b]$ 为实数区间, 一定存在单调下降的有理数列 $b_n \downarrow b$, 从而

$$(a, b] = \bigcap_n (a, b_n].$$

那么 $(a, b]$ 可以被生成。

□

练习 1.1.4:

1. Show that if $\mathcal{F}_1 \subset \mathcal{F}_2 \subset \cdots$ are σ -algebras, then $\bigcap_i \mathcal{F}_i$ is an algebra.
2. Give an example to show that $\bigcup_i \mathcal{F}_i$ need not be σ -algebra.

证明:

1. 设 \mathcal{F}_i 是 Ω 上的 σ -代数, 那么一定有 $\emptyset, \Omega \in \mathcal{F}_i$ 对每个 i 成立。对 $\forall A \in \bigcap_i \mathcal{F}_i$, 一定有 $A^c \in \bigcap_i \mathcal{F}_i$ 。对 $\forall \{A_n\} \in \mathcal{F}_i$, 一定有 $\bigcup_n A_n \in \mathcal{F}_i$ 。
2. 设 $\Omega = \mathbb{N}^+$, \mathcal{F}_i 为由 $2^{\{1, 2, \dots, i\}}$ 生成的 σ -代数。那么这时有 $\bigcup_i \mathcal{F}_i = 2^{\mathbb{N}}$ 。令 A_i 为 $[1, i]$ 中所有偶数的集合, 那么 $A_i \in \mathcal{F}_i$ 对所有 $i \in \mathbb{N}^+$ 成立, 从而 $A_i \in \bigcup_n \mathcal{F}_n$ 。这时 $\bigcup_i A_i$ 为 \mathbb{N}^+ 中所有偶数的集合, 但是其并不包含于 $\bigcup_i \mathcal{F}_i$, 因为它是无限集, 而后者的所有元素都是有限集。

□

练习 1.1.5: A set $A \subset \{1, 2, \dots\}$ is said to have asymptotic density θ if

$$\lim_{n \rightarrow \infty} \frac{|A \cap \{1, 2, \dots, n\}|}{n} = \theta.$$

Let \mathcal{A} be the collection of sets for which the asymptotic density exists. Is \mathcal{A} a σ -algebra? an algebra?

证明: \mathcal{A} 不是代数。令集合 A 为 \mathbb{N}^+ 中所有奇数的集合, 那么显然有 A 具有渐进密度 $1/2$ 。再构造具有渐进密度为 $1/2$ 的集合 B 如下:

$$B = \{1, 4, 5, 7, 10, 12, 14, 16, 17, 19, 21, 23, \dots\},$$

B 在所有形如 $(2^{2k-1}, 2^{2k}]$ 的区间上取所有奇数, 在所有形如 $(2^{2k}, 2^{2k+1}]$ 的区间上取所有偶数。这时可以证明 $A \cup B$ 不具有渐进密度, 因为在前 2^{2k-1} 个数上的密度小于 $3/4$, 而在前 2^{2k} 个数上的密度大于 $3/4$ 。□

1.2. 分布

练习 1.2.1: Suppose X and Y are random variables on (Ω, \mathcal{F}, P) and let $A \in \mathcal{F}$. Show that if we let $Z(\omega) = X(\omega)$ for $\omega \in A$ and $Z(\omega) = Y(\omega)$ for $\omega \in A^c$, then Z is a random variable.

证明: 对任意 Borel 集 $B \in \mathcal{R}$,

$$\begin{aligned} Z^{-1}(B) &= (Z^{-1}(B) \cap A) \cup (Z^{-1}(B) \cap A^c) \\ &= (X^{-1}(B) \cap A) \cup (Y^{-1}(B) \cap A^c) \\ &\in \mathcal{F}. \end{aligned}$$

从而 Z 也是随机变量。□

练习 1.2.2: Let χ have the standard normal distribution. Use Theorem 1.2.6 to get upper and lower bounds on $P(\chi \geq 4)$.

证明:

$$\begin{aligned} P(\chi \geq 4) &= \frac{1}{\sqrt{2\pi}} \int_4^\infty \exp\left(-\frac{x^2}{2}\right) dx \leq \frac{1}{4\sqrt{2\pi}} \exp(-8), \\ P(\chi \geq 4) &= \frac{1}{\sqrt{2\pi}} \int_4^\infty \exp\left(-\frac{x^2}{2}\right) dx \geq \left(\frac{1}{4} - \frac{1}{64}\right) \frac{1}{\sqrt{2\pi}} \exp(-8). \end{aligned}$$

□

练习 1.2.3: Show that a distribution function has at most countably many discontinuities.

证明: 由于分布函数是单调有界函数, 从数学分析相关知识知道其不连续点一定都是跳跃间断点, 且不连续点至多可数。□

练习 1.2.4: Show that if $F(x) = P(X \leq x)$ is continuous then $Y = F(X)$ has a uniform distribution on $(0, 1)$, that is, if $y \in [0, 1]$, $P(Y \leq y) = y$.

证明: 对 $\forall y \in [0, 1]$,

$$P(Y \leq y) = P(F(X) \leq y) = P(X \leq F^{-1}(y)) = F(F^{-1}(y)) = y.$$

□

练习 1.2.5: Suppose X has continuous density f , $P(\alpha \leq X \leq \beta) = 1$ and g is a function that is strictly increasing and differentiable on (α, β) . Then $g(X)$ has density

$$\frac{f(g^{-1}(y))}{g'(g^{-1}(y))}$$

for $y \in (g(\alpha), g(\beta))$ and 0 otherwise. When $g(x) = ax + b$ with $a > 0$,

$$g^{-1}(y) = \frac{y - b}{a}$$

so the answer is

$$\frac{1}{a} f\left(\frac{y - b}{a}\right).$$

证明: 对 $\forall y \in [\alpha, \beta]$, 有

$$P(g(X) \leq x) = P(X \leq g^{-1}(x)) = \int_{\alpha}^{g^{-1}(x)} f(t) dt.$$

令 $s = g^{-1}(t)$ 换元得到

$$P(g(X) \leq x) = \int_{\alpha}^{g^{-1}(x)} f(t) dt = \int_{g(\alpha)}^x \frac{f(g^{-1}(s))}{g'(g^{-1}(s))} ds.$$

□

练习 1.2.6: Suppose X has a normal distribution. Use the previous exercise to compute the density of $\exp(X)$. (The answer is the lognormal distribution.)

证明: 令 $g(X) = \exp(X)$, $g'(X) = \exp(X)$, $g^{-1}(y) = \log(y)$ 得到 $\exp(X)$ 的概率密度函数

$$\frac{f(g^{-1}(y))}{g'(g^{-1}(y))} = \frac{1}{\sqrt{2\pi}} \frac{\exp(-\frac{1}{2}(\log y)^2)}{\exp(\log(y))} = \frac{1}{\sqrt{2\pi}y} \exp\left(-\frac{1}{2} \log^2 y\right).$$

□

练习 1.2.7:

1. Suppose X has density function f . Compute the distribution function of X^2 and then differentiate to find its density function.
2. Work out the answer when X has a standard normal distribution to find the density of the chi-square distribution.

证明:

1. 记 X^2 的分布函数为 $G(x)$, 那么

$$G(x) = P(X^2 \leq x) = P(-\sqrt{x} \leq X \leq \sqrt{x}) = \int_{-\sqrt{x}}^{\sqrt{x}} f(t) dt.$$

即为其分布函数。记其概率密度函数为 g , 则

$$g(x) = \frac{dG}{dx} = \frac{f(\sqrt{x}) + f(-\sqrt{x})}{2\sqrt{x}}.$$

2. 代入正态分布概率密度函数得到

$$f_{\chi^2}(x) = \frac{1}{\sqrt{2\pi x}} \exp\left(-\frac{x}{2}\right).$$

□

1.3. 随机变量

练习 1.3.1: Show that if \mathcal{A} generates \mathcal{S} , then $X^{-1}(\mathcal{A}) \equiv \{\{X \in A\} : A \in \mathcal{A}\}$ generates $\sigma(X) = \{\{X \in B\} : B \in \mathcal{S}\}$.

证明: 因为 $X^{-1}(\mathcal{A}) \subset \sigma(X)$, 从而 $\sigma(X^{-1}(\mathcal{A})) \subset \sigma(X)$ 。对于任意包含 $X^{-1}(\mathcal{A})$ 的 σ -域 \mathcal{B} , $X(\mathcal{B})$ 是包含 \mathcal{A} 的 σ -域, 从而 $\mathcal{A} \subset \mathcal{S} \subset X(\mathcal{B})$, 那么 $X^{-1}(\mathcal{A}) \subset \sigma(X) \subset \mathcal{B}$, 那么 $\sigma(X^{-1}(\mathcal{A})) \subset \sigma(X) \subset \mathcal{B}$, 那么 $\sigma(X)$ 是包含 $X^{-1}(\mathcal{A})$ 的最小 σ -域, 从而 $\sigma(X) = \sigma(X^{-1}(\mathcal{A}))$. □

练习 1.3.2: Prove Theorem 1.3.6 when $n = 2$ by checking

$$\{X_1 + X_2 < x\} \in \mathcal{F}.$$

证明: 因为 $\{(-\infty, x) : x \in \mathbb{R}\}$ 可以生成 \mathcal{B} , 所以只要证明 $\{X_1 + X_2 < x\} \in \mathcal{F}$ 。容易看出 $\{x' = (x_1, x_2) \in \mathbb{R}^2 : x_1 + x_2 < x\} \in \mathcal{B}^2$, 从而 $f(x_1, x_2) = x_1 + x_2$ 是可测函数, 那么 $\{X_1 + X_2 < x\} \in \mathcal{F}$ 。 □

练习 1.3.3: Show that if f is continuous and $X_n \rightarrow X$ almost surely then $f(X_n) \rightarrow f(X)$ almost surely.

证明: 设 $X_n \rightarrow X$ 在 Ω' 上满足, 其中 $P(\Omega') = 1$ 。那么对 $\forall \omega \in \Omega'$ 有 $X_n(\omega) \rightarrow X(\omega)$, 又知道 f 连续, 从而 $f(X_n(\omega)) \rightarrow f(X(\omega))$ 。那么 $f(X_n) \rightarrow f(X)$ 在 Ω' 上成立。 □

练习 1.3.4:

1. Show that a continuous function from $\mathbb{R}^d \rightarrow \mathbb{R}$ is a measurable map from $(\mathbb{R}^d, \mathcal{R}^d)$ to $(\mathbb{R}, \mathcal{R})$.
2. Show that \mathcal{R}^d is the smallest σ -field that makes all the continuous functions measurable.

证明:

1. 设 $f: \mathbb{R}^d \rightarrow \mathbb{R}$ 是连续函数, 那么因为 $\{(-\infty, x)\}$ 生成 \mathcal{R} , 只要对 $\forall x \in \mathbb{R}$, 证明 $f^{-1}((-\infty, x)) \in \mathcal{R}^d$ 即可。因为 f 连续, 那么 $f^{-1}((-\infty, x))$ 为开集, 从而是 Borel 集。
2. 只要证明

$$\mathcal{R}^d = \sigma(\{f^{-1}((-\infty, x)) : x \in \mathbb{R}, f \in C^\infty(\mathbb{R}^d)\}).$$

由上一问已经知道

$$\sigma(\{f^{-1}((-\infty, x)) : x \in \mathbb{R}, f \in C^\infty(\mathbb{R}^d)\}) \subset \mathcal{R}^d.$$

考虑任意包含 $\{f^{-1}((-\infty, x)) : x \in \mathbb{R}, f \in C^\infty(\mathbb{R}^d)\}$ 的 σ -域 \mathcal{A} 。取定 \mathcal{R}^d 的一组生成元 $\{(-\infty, x_1) \times \cdots \times (-\infty, x_d) : x_1, \dots, x_d \in \mathbb{R}\}$, 取其中的一个元素为 $(-\infty, x_1) \times \cdots \times (-\infty, x_d)$ 。那么可以构造连续函数

$$f(x) = \text{dis}(x, (-\infty, x_1) \times \cdots \times (-\infty, x_d)),$$

其中 dis 表示距离函数。那么有 $f^{-1}((-\infty, 0)) = (-\infty, x_1) \times \cdots \times (-\infty, x_d)$, 从而 $\mathcal{R}^d \subset \mathcal{A}$ 。那么命题得证。

□

练习 1.3.5: A function f is said to be lower semicontinuous or l.s.c. if

$$\liminf_{y \rightarrow x} f(y) \geq f(x)$$

and upper semicontinuous (u.s.c.) if $-f$ is l.s.c. Show that f is l.s.c. if and only if $\{x : f(x) \leq a\}$ is closed for each $a \in \mathbb{R}$ and conclude that semicontinuous functions are measurable.

证明: 先证必要性, 假设 f 是下半连续的, 只要对 $\forall a \in \mathbb{R}$, 证 $\{x : f(x) > a\}$ 是开集。从中任取 x , 那么有

$$\liminf_{y \rightarrow x} f(y) \geq f(x) > a,$$

从而存在 x 的邻域 $U(x, \delta(x))$ 使得对 $\forall y \in (x, \delta(x))$ 有

$$f(y) \geq \frac{f(x) + a}{2} > a,$$

从而 x 是 $\{x : f(x) > a\}$ 的内点, 那么 $\{x : f(x) > a\}$ 是开集。

再证充分性。假设对 $\forall a \in \mathbb{R}$, $\{x : f(x) \leq a\}$ 是闭集, 从而 $\{x : f(x) > a\}$ 是开集。 $\forall x_0 \in \mathbb{R}$, 考虑令 $y \rightarrow x_0$ 。对 $\forall a < f(x_0)$, 知道 $f(x_0)$ 附近存在邻域 U 使得 $\forall y \in U$ 有 $f(y) > a$ 。令 $a \uparrow f(x_0)$ 即得到 $\liminf f(y) \geq f(x_0)$ 。

从而上半连续函数都是可测函数, 类似可以证明下半连续函数都是可测函数。

□

练习 1.3.6: Let $f : \mathbb{R}^d \rightarrow \mathbb{R}$ be an arbitrary function and let

$$f^\delta(x) = \sup\{f(y) : |y - x| < \delta\}$$

and

$$f_\delta(x) = \inf\{f(y) : |y - x| < \delta\}$$

where

$$|z| = \sqrt{z_1^2 + \cdots + z_d^2}.$$

Show that f^δ is l.s.c. and f_δ is u.s.c. Let $f^0 = \lim_{\delta \downarrow 0} f^\delta$, $f_0 = \lim_{\delta \downarrow 0} f_\delta$, and conclude that the set of points at which f is discontinuous $= \{f^0 \neq f_0\}$ is measurable.

follows from the fact that $f^0 - f_0$ is.

证明: 直接按定义验证得到 f^δ 是上半连续函数而 f_δ 是下半连续函数。接下来只要证明 f 的不连续点集就是 $\{f^0 \neq f_0\}$ 。

假设 $x \in \mathbb{R}^d$ 是 f 的不连续点, 那么存在 $\varepsilon_0 > 0$, 存在收敛到 x 的点列 $\{x'_n\}$ 和 $\{x''_n\}$ 使得 $|\lim f(x'_n) - \lim f(x''_n)| < \varepsilon_0$ 。不妨设

$$\lim f(x'_n) > \lim f(x''_n) + \varepsilon_0.$$

对 $\forall \delta > 0$, 总能找到 x'_{n_0}, x''_{n_0} 使得

$$|x'_{n_0} - x|, |x''_{n_0} - x| < \delta$$

且

$$f(x'_{n_0}) \geq f(x''_{n_0}) + \frac{\varepsilon_0}{2},$$

从而一定有

$$f^\delta(x) \geq f(x'_{n_0}) \geq f(x''_{n_0}) + \frac{\varepsilon_0}{2} \geq f_\delta(x).$$

令 $\delta \downarrow 0$ 得到 $f^0(x) \geq f_0(x) + \varepsilon_0/2$ 。

再假设 x_0 是 f 的连续点, 那么对 $\forall \varepsilon > 0$, 存在 $\delta > 0$ 使得 $\forall x \in U(x_0, \delta)$ 有 $|f(x) - f(x_0)| < \varepsilon$, 从而

$$|f^\delta - f_\delta| \leq |f^\delta - f(x_0)| + |f_\delta - f(x_0)| < 2\varepsilon.$$

令 $\varepsilon \rightarrow 0$ 即可知道 $f^0 = f_0$ 。

□

练习 1.3.7: A function $\varphi : \Omega \rightarrow \mathbb{R}$ is said to be simple if

$$\varphi(\omega) = \sum_{m=1}^n c_m 1_{A_m}(\omega)$$

where the c_m are real numbers and $A_m \in \mathcal{F}$. Show that the class of \mathcal{F} measurable functions is the smallest class containing the simple functions and closed under pointwise limits.

证明: 设 P 是由从 Ω 到 \mathbb{R} 上的部分函数构成的集合, 满足所有的简单函数都属于 P 且对点态极限运算封闭。那么只要证明所有的 \mathcal{F} 可测函数都在 P 中。对 $\forall f : \Omega \rightarrow \mathbb{R}$ 为 \mathcal{F} 可测函数, 由 $f = f^+ - f^-$ 可以不妨设 $f \geq 0$ 。那么直接构造简单函数列

$$\varphi_n(\omega) = \begin{cases} n, & f(\omega) \geq n \\ \frac{k}{2^n}, & \frac{k}{2^n} \leq f(\omega) < \frac{k+1}{2^n}, k = 0, 1, \dots, n2^n - 1 \end{cases}$$

容易看出 φ_n 点态收敛于 f , 从而 $f \in P$ 。

□

练习 1.3.8: Use the previous exercise to conclude that Y is measurable with respect to $\sigma(X)$ if and only if $Y = f(X)$ where $f : \mathbb{R} \rightarrow \mathbb{R}$ is measurable.

证明: 充分性是显然的, 下证明必要性。设 Y 在 $\sigma(X)$ 上可测。那么由上一问题知道 Y 可以写成 $\sigma(X)$ 上的简单函数的点态极限, 记为 $Y = \lim \varphi_n$ 。同样可以将 X 写为 $X = \lim \psi_n$ 。记

$$\begin{aligned} \varphi_n &= \sum_{m=1}^{a_n} c_{n,m} 1_{A_{n,m}}, \\ \psi_n &= \sum_{m=1}^{a_n} d_{n,m} 1_{A_{n,m}}, \end{aligned}$$

可以不妨设 $\{d_{n,m}\}_m$ 两两不同。那么令

$$f_n : \mathbb{R} \rightarrow \mathbb{R}, f_n(d_{n,m}) = c_{n,m}, \text{ 否则 } f_n = 0.$$

那么就有 $f_n(\psi_n) = \varphi_n$ 。令 $n \rightarrow \infty$, 则有 $f(X) = Y$, 且 f 为可测函数。

□

练习 1.3.9: To get a constructive proof of the last result, note that

$$\{\omega : m2^{-n} \leq Y < (m+1)2^{-n}\} = \{X \in B_{m,n}\}$$

for some $B_{m,n} \in \mathcal{R}$ and set $f_n(x) = m2^{-n}$ for $x \in B_{m,n}$ and show that as $n \rightarrow \infty$ $f_n(x) \rightarrow f(x)$ and $Y = f(X)$.

证明：同上题。

□

1.4. 积分

练习 1.4.1: Show that if $f \geq 0$ and $\int f d\mu = 0$ then $f = 0$ a.e.

证明：设 $A = \{f > 0\} = \bigcup_n \{f > 1/n\}$ 。若 $\mu(A) > 0$ ，那么由测度的连续性知道，存在足够大的 $n_0 \in \mathbb{N}^*$ 使得 $\mu(\{f > 1/n\}) > 0$ 。那么

$$\begin{aligned} \int f d\mu &= \int_A f d\mu + \int_{A^c} f d\mu \\ &= \int_A f d\mu \\ &= \int_{\{f > \frac{1}{n}\}} f d\mu \\ &\geq \int_{\{f > \frac{1}{n}\}} \frac{1}{n} d\mu \\ &= \frac{1}{n} \mu\left(\left\{f > \frac{1}{n}\right\}\right) \\ &> 0, \end{aligned}$$

从而矛盾。

□

练习 1.4.2: Let $f \geq 0$ and

$$E_{n,m} = \left\{x : \frac{m}{2^n} \leq f(x) < \frac{m+1}{2^n}\right\}.$$

As $n \uparrow \infty$,

$$\sum_{m=1}^{\infty} \frac{m}{2^n} \mu(E_{n,m}) \uparrow \int f d\mu.$$

证明：不妨设 $\lambda = \sum_m \mu(E_{n,m})$ 有限，否则命题化为 $\infty \uparrow \infty$ 显然成立。那么有

$$\begin{aligned}
& \int f \, d\mu - \sum_{m=1}^{\infty} \frac{m}{2^n} \mu(E_{n,m}) \\
&= \sum_{m=1}^{\infty} \int_{E_{n,m}} \left(f - \frac{m}{2^n}\right) d\mu \\
&\leq \sum_{m=1}^{\infty} \int_{E_{n,m}} \frac{1}{2^n} d\mu \\
&\leq \sum_{m=1}^{\infty} \frac{\mu(E_{n,m})}{2^n} \\
&= \frac{\lambda}{2^n} \downarrow 0 \quad (n \rightarrow \infty).
\end{aligned}$$

从而命题得证。 □

练习 1.4.3: Let g be an integrable function on \mathbb{R} and $\varepsilon > 0$.

1. Use the definition of the integral to conclude there is a simple function $\varphi = \sum_k b_k 1_{A_k}$ with

$$\int |g - \varphi| \, d\mu < \varepsilon.$$

2. Use Exercise A.2.1 to approximate the A_k by finite unions of intervals to get a step function

$$q = \sum_{j=1}^k c_j 1_{(a_{j-1}, a_j)}$$

with $a_0 < a_1 < \dots < a_k$, so that

$$\int |\varphi - q| < \varepsilon.$$

3. Round the corners of q to get a continuous function r so that

$$\int |q - r| \, d\mu < \varepsilon.$$

To make a continuous function replace each $c_j 1_{(a_{j-1}, a_j)}$ by a function that is 0 on $(a_{j-1}, a_j)^c$, c_j on $[a_{j-1} + \delta_j, a_j - \delta_j]$, and linear otherwise. If the δ_j are small enough and we let

$$r(x) = \sum_{j=1}^k r_j(x)$$

then

$$\int |q(x) - r(x)| d\mu = \sum_{j=1}^k \delta_j c_j < \varepsilon.$$

证明:

1. 由于 $g = g^+ - g^-$ 不妨先证明 g^+ , 然后类似证明 g^- . 由积分的定义知道存在可测函数 h^+ , 使得 $0 \leq h^+ \leq g$, h^+ 有界, 且支集测度有限, 且

$$\int |g - h^+| d\mu < \frac{\varepsilon}{4}.$$

对 h^+ , 再由积分定义知道存在非负简单函数 $\varphi^+ \leq h^+$, 使得

$$\int |h^+ - \varphi^+| d\mu < \frac{\varepsilon}{4}.$$

那么

$$\int |g^+ - \varphi^+| d\mu < \int |g^+ - h^+| d\mu + \int |h^+ - \varphi^+| d\mu < \frac{\varepsilon}{2}.$$

类似得到 φ^- 满足

$$\int |g^- - \varphi^-| d\mu < \int |g^- - h^-| d\mu + \int |h^- - \varphi^-| d\mu < \frac{\varepsilon}{2}.$$

令 $\varphi = \varphi^+ - \varphi^-$ 直接得到结果。

2. 记

$$\varphi = \sum_{i=1}^n b_i 1_{A_i}.$$

对每个 A_i , 由测度近似定理知道存在有限个开区间 $B_{i,1}, \dots, B_{i,b_i}$ 使得对 $B_i = \bigcup_j B_{i,j}$ 有 $A_i \subset B_i$ 且

$$\mu(B_i \setminus A_i) < \frac{\varepsilon}{b_i n}.$$

那么就有

$$\varphi = \sum_{i=1}^n \sum_{j=1}^{b_i} b_i 1_{B_{i,j}},$$

重新排列就能够得到欲证命题。

□

练习 1.4.4: Prove the Riemann-Lebesgue lemma. If g is integrable then

$$\lim_{n \rightarrow \infty} \int g(x) \cos nx \, dx = 0.$$

Hint: If g is a step function, this is easy. Now use the previous exercise.

证明：先证明对于任意开区间 (l, r) 成立

$$\lim_{n \rightarrow \infty} \int_{(l, r)} \cos nx \, dx = 0.$$

考虑到 $\cos nx$ 是以 $2/n$ 为周期的周期函数，且在一个周期内的积分是 0，从而知道

$$\left| \int_{(l, r)} \cos nx \, dx \right| = \left| \int_{(l, l+\delta)} \cos nx \, dx \right| \leq \int_{(l, l+\delta)} |\cos nx| \, dx \leq \delta,$$

其中 $\delta < 2/n$. 令 $n \rightarrow \infty$ 就得到积分趋近于 0. 那么对任意的 *step function* g , 记

$$g = \sum_{i=1}^k c_i 1_{(a_i, b_i)},$$

那么就有

$$\lim_{n \rightarrow \infty} \int g(x) \cos nx \, dx = \sum_{i=1}^k c_i \lim_{n \rightarrow \infty} \int_{(a_i, b_i)} \cos nx \, dx = 0.$$

从而原命题在 g 为 *step function* 时成立。对于 g 为任意可积函数， $n \in \mathbb{N}^*$ 的情况，存在 *step function* 构成的函数列 $\{h_k\}$ 使得

$$\begin{aligned} & \int |g(x) \cos nx - h_k(x) \cos nx| \, dx \\ & \leq \int |g(x) - h_k(x)| \, dx \\ & \rightarrow 0 \quad (k \rightarrow \infty). \end{aligned}$$

从而有

$$\begin{aligned} & \left| \lim_{n \rightarrow \infty} \int g(x) \cos nx \, dx \right| \\ & \leq \lim_{n \rightarrow \infty} \int |g(x) \cos nx| \, dx \\ & \leq \lim_{n \rightarrow \infty} \lim_{k \rightarrow \infty} \int |g(x) \cos nx - h_k(x) \cos nx| \, dx + \lim_{n \rightarrow \infty} \lim_{k \rightarrow \infty} \int |h_k(x) \cos nx| \, dx \\ & \leq 0 + 0 \\ & = 0. \end{aligned}$$

□

1.5. 积分的性质

练习 1.5.1: Let

$$\|f\|_{\infty} = \inf\{M : \mu(\{x : |f(x)| > M\}) = 0\}.$$

Prove that

$$\int |fg| d\mu \leq \|f\|_1 \|g\|_{\infty}.$$

证明:

$$\begin{aligned} \int |fg| d\mu &\leq \int |f| |g| d\mu \\ &\leq \int |f| \|g\|_{\infty} d\mu \\ &= \|g\|_{\infty} \mu(\Omega) \int |f| d\mu \\ &= \|f\|_1 \|g\|_{\infty}. \end{aligned}$$

□

练习 1.5.2: Show that if μ is a probability measure then

$$\|f\|_{\infty} = \lim_{p \rightarrow \infty} \|f\|_p.$$

证明: 对任意的 $\varepsilon > 0$, 有

$$\begin{aligned} \|f\|_p &= \left[\int |f|^p d\mu \right]^{\frac{1}{p}} \\ &= \left[\int_{\{|f|^p > \|f\|_{\infty} - \varepsilon\}} |f|^p d\mu \right]^{\frac{1}{p}} \\ &\geq [\mu(\{|f|^p > \|f\|_{\infty} - \varepsilon\}) (\|f\|_{\infty} - \varepsilon)^p]^{\frac{1}{p}} \\ &= [\mu(\{|f|^p > \|f\|_{\infty} - \varepsilon\})]^{\frac{1}{p}} (\|f\|_{\infty} - \varepsilon). \end{aligned}$$

令 $p \rightarrow \infty$, 再令 $\varepsilon \rightarrow 0$ 得到 $\|f\|_p \geq \|f\|_{\infty}$.

另一方面有

$$\begin{aligned}
\|f\|_p &= \left[\int_{\{|f|^p > \|f\|_\infty\}} |f|^p d\mu + \int_{\{|f|^p \leq \|f\|_\infty\}} |f|^p d\mu \right]^{\frac{1}{p}} \\
&= \left[\int_{\{|f|^p \leq \|f\|_\infty\}} |f|^p d\mu \right]^{\frac{1}{p}} \\
&\leq \left[\int_{\{|f|^p \leq \|f\|_\infty\}} \|f\|_\infty d\mu \right]^{\frac{1}{p}} \\
&= \|f\|_\infty.
\end{aligned}$$

□

练习 1.5.3:

1. Suppose $p \in (1, \infty)$. The inequality

$$|f + g|^p \leq 2^p(|f|^p + |g|^p)$$

shows that if $\|f\|_p$ and $\|g\|_p$ are $< \infty$ then

$$\|f + g\|_p < \infty.$$

Apply Hölder's inequality to

$$|f||f + g|^{p-1}$$

and

$$|g||f + g|^{p-1}$$

to show

$$\|f + g\|_p \leq \|f\|_p + \|g\|_p.$$

2. Show that the last result remains true when $p = 1$ or $p = \infty$.

证明：直接用 Hölder 不等式知道

$$\begin{aligned}
\int |f| |f + g|^{p-1} d\mu &\leq \|f\|_p \|(f + g)^{p-1}\|_q, \\
\int |g| |f + g|^{p-1} d\mu &\leq \|g\|_p \|(f + g)^{p-1}\|_q.
\end{aligned}$$

那么有

$$\begin{aligned}
\int |f + g|^p d\mu &\leq \int (|f||f + g|^{p-1} + |g||f + g|^{p-1}) d\mu \\
&\leq (\|f\|_p + \|g\|_p) \|(f + g)^{p-1}\|_q.
\end{aligned}$$

从而有

$$\|f\|_p + \|g\|_p \geq \frac{\int |f+g|^p d\mu}{(\int |f+g|^{(p-1)q})^{\frac{1}{q}}} = \frac{\int |f+g|^p d\mu}{(\int |f+g|^p)^{1-\frac{1}{p}}} = \|f+g\|_p.$$

再由之前练习, Hölder 不等式在 $p = 1$ 或 $p = \infty$ 时也成立, 所以可以将 p 扩展到 $[1, \infty]$ 。□

练习 1.5.4: If f is integrable and E_m are disjoint sets with union E then

$$\sum_{m=0}^{\infty} \int_{E_m} f d\mu = \int_E f d\mu.$$

So if $f \geq 0$, then $\nu(E) = \int_E f d\mu$ defines a measure.

证明: 由于可以将 f 分解为 $f^+ - f^-$, 故可以不妨设 $f \geq 0$ 。令

$$g = f \cdot 1_E, \quad g_n = f \cdot \sum_{i=0}^n 1_{E_i},$$

那么只要证明

$$\lim_{n \rightarrow \infty} \int g_n d\mu = \int g d\mu.$$

注意到 $g_n \uparrow g$ 且 $g = f \cdot 1_E$ 可积, 由单调收敛定理直接推导出以上结论。□

练习 1.5.5: If $g_n \uparrow g$ and $\int g_1^- d\mu < \infty$ then

$$\int g_n d\mu \uparrow \int g d\mu.$$

证明: 令 $g'_n = g_n + g_1^-$, 那么有 $g'_n \geq g'_1 = g_1^+ \geq 0$ 且 $g'_n \uparrow (g - g_1^-)$, 由单调收敛定理知道

$$\lim \int g_n d\mu - \int g_1^- d\mu = \lim \int g'_n d\mu = \int (g - g_1^-) d\mu = \int g d\mu - \int g_1^- d\mu,$$

即

$$\int g_n d\mu \uparrow \int g d\mu.$$

□

练习 1.5.6: If $g_m \geq 0$ then

$$\int \sum_{m=0}^{\infty} g_m d\mu = \sum_{m=0}^{\infty} \int g_m d\mu.$$

证明: 记

$$f_n = \sum_{m=1}^n g_m, \quad f = \lim_{n \rightarrow \infty} f_n.$$

那么有 $f_n \geq 0$ 且 $f_n \uparrow f$, 从而由单调收敛定理知道

$$\sum_{m=0}^{\infty} \int g_m d\mu = \lim \int f_n d\mu \uparrow \int f d\mu = \int \sum_{m=0}^{\infty} g_m d\mu.$$

□

练习 1.5.7: Let $f \geq 0$.

1. Show that

$$\int f \wedge n d\mu \uparrow \int f d\mu$$

as $n \rightarrow \infty$.

2. Use (1) to conclude that if g is integrable and $\varepsilon > 0$ then we can pick $\delta > 0$ so that $\mu(A) < \delta$ implies

$$\int_A |g| d\mu < \varepsilon.$$

证明:

1. 由于 $f \wedge n \geq 0$ 且 $f \wedge n \uparrow f$ 直接应用单调收敛定理知道

$$\int f \wedge n d\mu \uparrow \int f d\mu.$$

2. 由上一问结论知道存在足够大的 $N \in \mathbb{N}^*$ 使得

$$\int_{|g| \geq N} |g| d\mu = \int |g| d\mu - \int |g| \wedge N d\mu < \frac{\varepsilon}{2}.$$

那么令 $\delta = \varepsilon/(2N)$ 直接对 $\forall A \subset \Omega$ 且 $\mu(A) < \delta$ 有

$$\begin{aligned}
\int_A |g| d\mu &= \int_{A \cap \{|g| \geq N\}} |g| d\mu + \int_{A \cap \{|g| < N\}} |g| d\mu \\
&\leq \int_{\{|g| \geq N\}} |g| d\mu + \int_A N d\mu \\
&\leq \frac{\varepsilon}{2} + \frac{\varepsilon}{2N} \cdot N \\
&= \varepsilon.
\end{aligned}$$

□

练习 1.5.8: Show that if f is integrable on $[a, b]$,

$$g(x) = \int_{[a, x]} f(y) dy$$

is continuous on (a, b) .

证明: $\forall x_0 \in (a, b)$. 考虑 $\forall x \in (x_0, b)$, 有

$$|g(x) - g(x_0)| = \int_{(x_0, x]} f(y) dy \rightarrow 0 \quad (x \downarrow x_0)$$

其中趋近的结论由上一题目得出。类似可以证明 $x \uparrow x_0$ 的情况。

□

练习 1.5.9: Show that if f has

$$\|f\|_p = \left(\int |f|^p d\mu \right)^{\frac{1}{p}} < \infty,$$

then there are simple functions φ_n so that

$$\|\varphi_n - f\|_p \rightarrow 0.$$

证明: 由于

$$\|\varphi_n - f\|_p \leq \|\varphi_n^+ - f^+\|_p + \|\varphi_n^- - f^-\|_p,$$

故可以不妨设 $f \geq 0$ 。由简单逼近定理, 存在简单函数列 φ_n 使得 $\varphi_n \leq f$ 并且有 $\varphi_n \uparrow f$ 几乎处处。那么

$$\int |\varphi_n - f|^p d\mu \leq \int |2f|^p d\mu \leq 2^p \int |f|^p d\mu < \infty,$$

由 Lebesgue 逼近定理知道

$$\lim \int |\varphi_n - f|^p d\mu = \int \lim |\varphi_n - f|^p d\mu = 0,$$

两边同时开 p 次方就得到结论。

□

练习 1.5.10: Show that if

$$\sum_n \int |f_n| d\mu < \infty$$

then

$$\sum_n \int f_n d\mu = \int \sum_n f_n d\mu.$$

证明: 因为可以将 f_n 分解为 $f_n^+ - f_n^-$ 并分别证明, 所以可以直接不妨设 $f_n \geq 0$. 令

$$g_n = \sum_{m=1}^n f_m, \quad g = \sum_m f_m.$$

那么相当于已知

$$\sum_n \int |f_n| d\mu = \lim \sum_{m=1}^n \int f_m d\mu = \lim \int g_n d\mu < \infty,$$

只要求证

$$\lim \int g_n d\mu = \sum_n \int f_n d\mu = \int \sum_n f_n d\mu = \int g d\mu.$$

由于 $f_n \geq 0$, 从而 $g_n \uparrow g$, 从而直接应用单调收敛定理知道

$$\lim \int g_n d\mu = \int g d\mu < \infty.$$

□

1.6. 期望

练习 1.6.1: Suppose φ is strictly convex, i.e., $>$ holds for $\lambda \in (0, 1)$. Show that, under the assumptions of Theorem 1.6.2, $\varphi(\mathbb{E}X) = \mathbb{E}\varphi(X)$ implies $X = \mathbb{E}X$ a.s. .

证明: 使用反证法。设 φ 严格凸, $\varphi(\mathbb{E}X) = \mathbb{E}\varphi(X)$ 且 $P(A = \{X \neq \mathbb{E}X\}) > 0$ 。那么对 $\forall \omega \in A$ 由严格凸函数性质有

$$\varphi(X(\omega)) > \varphi(\mathbb{E}X) + \varphi'(\mathbb{E}X)(X(\omega) - \mathbb{E}X),$$

也即

$$\varphi(X)1_A > (\varphi(\mathbb{E}X) + \varphi'(\mathbb{E}X)(X - \mathbb{E}X))1_A,$$

左右同时取期望得到

$$\begin{aligned}
\text{LHS} &= E(\varphi(X)1_A) \\
&= E(\varphi(X)) - E(\varphi(X)1_{A^c}) \\
&= E(\varphi(X)) - E(\varphi(EX)1_{A^c}) \\
&= E(\varphi(X)) - \varphi(EX)P(A^c) \\
&> \text{RHS} \\
&= E((\varphi(EX) + \varphi'(EX)(X - EX))1_A) \\
&= E(\varphi(EX) + \varphi'(EX)(X - EX)) - E((\varphi(EX) + \varphi'(EX)(X - EX))1_{A^c}) \\
&= E(\varphi(EX) + \varphi'(EX)(X - EX)) - E(\varphi(EX)1_{A^c}) \\
&= E(\varphi(EX) + \varphi'(EX)(X - EX)) - \varphi(EX)P(A^c),
\end{aligned}$$

从而得到

$$\begin{aligned}
E(\varphi(X)) &> E((\varphi(EX) + \varphi'(EX)(X - EX))) \\
&= \varphi(EX) + \varphi'(EX)(EX - EX) \\
&= \varphi(EX) \\
&= E(\varphi(X)),
\end{aligned}$$

矛盾。

□

练习 1.6.2: Suppose $\varphi : \mathbb{R}^n \rightarrow \mathbb{R}$ is convex. Imitate the proof of Theorem 1.5.1 to show

$$E\varphi(X_1, \dots, X_n) \geq \varphi(EX_1, \dots, EX_n)$$

provided $E|\varphi(X_1, \dots, X_n)| < \infty$ and $E|X_i| < \infty$ for all i .

证明：由凸函数性质知道

$$\varphi(X_1, \dots, X_n) \geq \varphi(EX_1, \dots, EX_n) + \varphi'(EX_1, \dots, EX_n) \cdot (X_1 - EX_1, \dots, X_n - EX_n),$$

其中 $x \cdot y$ 为向量点积。直接取期望得到

$$\begin{aligned}
E\varphi(X_1, \dots, X_n) &\geq \varphi(EX_1, \dots, EX_n) + \\
&\quad \varphi'(EX_1, \dots, EX_n) \cdot (EX_1 - EX_1, \dots, EX_n - EX_n) \\
&\geq \varphi(EX_1, \dots, EX_n).
\end{aligned}$$

□

练习 1.6.3: Chebyshev's inequality is and is not sharp.

1. Show that Theorem 1.6.4 is sharp by showing that if $0 < b \leq a$ are fixed there is an X with $EX^2 = b^2$ for which

$$P(|X| \geq a) = \frac{b^2}{a^2}.$$

2. Show that Theorem 1.6.4 is not sharp by showing that if X has $0 < EX^2 < \infty$ then

$$\lim_{a \rightarrow \infty} \frac{a^2 P(|X| \geq a)}{EX^2} = 0$$

证明:

1. 令 X 为离散分布, 其中 $P(X=0) = 1 - b^2/a^2$, $P(X=a) = b^2/a^2$ 。这时有

$$P(|X| \geq a) = P(X=a) = \frac{b^2}{a^2},$$

且

$$EX^2 = \frac{b^2}{a^2} \cdot a^2 = b^2.$$

2. 只要证明 $a^2 P(|X| \geq a) \rightarrow 0$ 。有

$$\begin{aligned} a^2 P(|X| \geq a) &= a^2 P(X^2 \geq a^2) \\ &= \int_{\{X^2 \geq a^2\}} a^2 dP \\ &\leq \int_{\{X^2 \geq a^2\}} X^2 dP \\ &\rightarrow 0 \quad (a \rightarrow \infty) \quad (\text{因为 } EX^2 < \infty). \end{aligned}$$

□

练习 1.6.4:

1. Let $a > b > 0$, $0 < p < 1$ and let X have $P(X=a) = p$ and $P(X=-b) = 1-p$. Apply Theorem 1.6.4 to $\varphi(x) = (x+b)^2$ and conclude that if Y is any random variable with $EY = EX$ and $\text{var}(Y) = \text{var}(X)$, then $P(Y \geq a) \leq p$ and equality holds when $Y = X$.
2. Suppose $EY = 0$, $\text{var}(Y) = \sigma^2$, and $a > 0$. Show that

$$P(Y \geq a) \leq \frac{\sigma^2}{a^2 + \sigma^2},$$

and there is a Y for which equality holds.

证明:

1. 直接将 $A = \{Y \geq a\}$, $\varphi(Y) = (Y+b)^2$ 代入 Chebyshev 不等式得到

$$E((Y+b)^2; Y \geq a) \leq E(Y+b)^2.$$

其中又有

$$\begin{aligned}
E(Y+b)^2 &= EY^2 + 2bEY + b^2 \\
&= \text{var } Y + EY + 2bEY + b^2 \\
&= \text{var } X + EX + 2bEX + b^2 \\
&= EX^2 + 2bEX + b^2 \\
&= pa^2 + (1-p)b^2 + 2b(pa + (1-p)(-b)) + b^2 \\
&= p(a+b)^2,
\end{aligned}$$

而

$$E((Y+b)^2; Y \geq a) \geq (a+b)^2 P(Y \geq a).$$

从而直接代入 *Chebyshev* 不等式就得到结论。容易验证 $Y = X$ 的时候取等。

2. 代入 $p = \sigma^2/(a^2 + \sigma^2)$, $b = \sigma^2/a$ 直接得出结论。

□

练习 1.6.5: Show that:

1. if $\varepsilon > 0$, $\inf\{P(|X| > \varepsilon) : EX = 0, \text{var}(X) = 1\} = 0$.
2. if $y \geq 1$, $\sigma^2 \in (0, \infty)$,

$$\inf\{P(|X| > y) : EX = 1, \text{var}(X) = \sigma^2\} = 0.$$

证明:

1. 令 X_n 满足 $P(X_n = n) = P(X_n = -n) = 1/(2n^2)$, $P(X_n = 0) = 1 - 1/(2n^2)$.
这时容易验证 $\text{var } X_n = 1$, $EX_n = 0$ 。且在 $n \rightarrow \infty$ 时有 $P(|X_n| > \varepsilon) \rightarrow 0$ 对 $\forall \varepsilon > 0$ 成立。
2. 令 X_n 满足

$$P(X_n = 1+n) = P(X_n = 1-n) = \frac{\sigma^2}{2n^2}, \quad P(X_n = 1) = 1 - \frac{\sigma^2}{n^2}.$$

这时 $EX_n = 1$, $\text{var } X_n = \sigma^2$, 且 $P(|X_n| > y) \rightarrow 0$ 在 $n \rightarrow \infty$ 时对 $\forall y \geq 1$ 成立。

□

练习 1.6.6: A useful lower bound. Let $Y \geq 0$ with $EY^2 < \infty$. Apply the Cauchy – Schwarz inequality to $Y1_{\{Y>0\}}$ and conclude

$$P(Y > 0) \geq \frac{(EY)^2}{EY^2}.$$

证明: 由于 $Y \geq 0$ 从而 $Y1_{\{Y>0\}} = Y$ 。直接使用 *Cauchy-Schwarz* 不等式得到

$$EY = EY1_{\{Y>0\}} \leq \sqrt{\int Y^2 dP \int 1_{\{Y>0\}}^2 dP} \leq \sqrt{EY^2 P(Y > 0)},$$

即

$$P(Y > 0) \geq \frac{(EY)^2}{EY^2}.$$

□

练习 1.6.7: Let $\Omega = (0, 1)$ equipped with the Borel sets and Lebesgue measure. Let $\alpha \in (1, 2)$ and

$$X_n = n^\alpha 1_{(\frac{1}{n+1}, \frac{1}{n})} \rightarrow 0 \quad \text{a.s.}$$

Show that Theorem 1.6.8 can be applied with $h(x) = x$ and $g(x) = |x|^{2/\alpha}$, but the X_n are not dominated by an integrable function.

证明: 首先验证三个条件。

1. $g(x) \geq 0$ 显然, 而且当 $|x| \rightarrow \infty$ 时容易看出 $g(x) \rightarrow \infty$;

2.
$$\frac{|h(x)|}{g(x)} = |x| \cdot |x|^{-\frac{2}{\alpha}} = |x|^{1-\frac{2}{\alpha}} \rightarrow 0 \quad (x \rightarrow \infty);$$

3.
$$\begin{aligned} Eg(X_n) &= n^2 E\left(1_{(\frac{1}{n+1}, \frac{1}{n})}\right)^{\frac{2}{\alpha}} \\ &\leq n^2 E\left(\frac{1}{n^2}\right)^{\frac{2}{\alpha}} \\ &= n^{2-\frac{4}{\alpha}} \rightarrow 0 \quad (n \rightarrow \infty), \end{aligned}$$

从而一定有 $Eg(X_n) \leq K < \infty$.

那么由定理就知道

$$EX_n \rightarrow EX = 0.$$

从另一方面看, X_n 不能被任何可积函数控制, 这是因为如果要 $|X_n| \leq Y$ 对 $\forall n = 1, 2, \dots$ 成立, 那么有 Y 在 $\{Y \in (\frac{1}{n+1}, \frac{1}{n})\}$ 上取值大于等于 $X_n = n^\alpha$. 这时

$$\begin{aligned} EY &\geq E \sum_n n^\alpha 1_{(\frac{1}{n+1}, \frac{1}{n})} \\ &= \sum_n n^\alpha E 1_{(\frac{1}{n+1}, \frac{1}{n})} \quad (\text{单调收敛定理}) \\ &\geq \sum_n \frac{n^\alpha}{2n^2} \\ &= \frac{1}{2} \sum_n n^{\alpha-2} \end{aligned}$$

不收敛。

□

练习 1.6.8: Suppose that the probability measure μ has $\mu(A) = \int_A f(x)dx$ for all $A \in \mathcal{R}$. Use the proof technique of Theorem 1.6.9 to show that for any g with $g \geq 0$ or

$$\int |g(x)|\mu(dx) < \infty$$

we have

$$\int g(x)\mu(dx) = \int g(x)f(x)dx.$$

证明:

1. 当 g 是指示函数时, 设 $g(x) = 1_A$, 其中 $A \in \mathcal{R}$ 。这时

$$\int g(x)\mu(dx) = \int 1_A\mu(dx) = \mu(A) = \int 1_A f(x) dx = \int g(x)f(x) dx;$$

2. 当 g 是简单函数时, 设 $g(x) = \sum_{i=1}^n a_i 1_{A_i}$, 其中 $A_i \in \mathcal{R}$ 。这时

$$\begin{aligned} \int g(x)\mu(dx) &= \int \sum_{i=1}^n a_i 1_{A_i} \mu(dx) \\ &= \sum_{i=1}^n a_i \int 1_{A_i} \mu(dx) \\ &= \sum_{i=1}^n a_i \int 1_{A_i} f(x) dx \\ &= \int \left(\sum_{i=1}^n a_i 1_{A_i} \right) f(x) dx \\ &= \int g(x)f(x) dx; \end{aligned}$$

3. 当 g 是非负函数时, 令

$$g_n(x) = \frac{[2^n g(x)]}{2^n} \wedge n,$$

那么 $g_n(x)$ 是简单函数且 $g_n \uparrow g, g_n f \uparrow gf$, 从而由单调收敛定理知道

$$\begin{aligned} \int g(x)\mu(dx) &= \lim \int g_n(x)\mu(dx) \\ &= \lim \int g_n(x)f(x) dx \\ &= \int g(x)f(x) dx; \end{aligned}$$

4. 当 g 为任意可积函数时, 有

$$\begin{aligned}
\int g(x)\mu(\mathrm{d}x) &= \int (g^+(x) - g^-(x))\mu(\mathrm{d}x) \\
&= \int g^+(x)\mu(\mathrm{d}x) - \int g^-(x)\mu(\mathrm{d}x) \\
&= \int g^+(x)f(x) \mathrm{d}x - \int g^-(x)f(x) \mathrm{d}x \\
&= \int (g^+(x) - g^-(x))f(x) \mathrm{d}x \\
&= \int g(x)f(x) \mathrm{d}x.
\end{aligned}$$

□

练习 1.6.9: Inclusion - exclusion formula. Let A_1, A_2, \dots, A_n be events and $A = \bigcup_{i=1}^n A_i$. Prove that $1_A = 1 - \prod_{i=1}^n (1 - 1_{A_i})$. Expand out the right hand side, then take expected value to conclude

$$\begin{aligned}
P\left(\bigcup_{i=1}^n A_i\right) &= \sum_{i=1}^n P(A_i) - \sum_{i < j} P(A_i \cap A_j) \\
&\quad + \sum_{i < j < k} P(A_i \cap A_j \cap A_k) - \dots + (-1)^{n-1} P\left(\bigcap_{i=1}^n A_i\right).
\end{aligned}$$

证明:

$$\begin{aligned}
1 - \prod_{i=1}^n (1 - 1_{A_i}) &= 1 - \prod_{i=1}^n 1_{A_i^c} \\
&= 1 - \bigcap_{i=1}^n 1_{A_i^c} \\
&= 1 - 1_{\bigcap_{i=1}^n A_i^c} \\
&= 1_{\left(\bigcap_{i=1}^n A_i^c\right)^c} \\
&= 1_{\bigcup_{i=1}^n A_i} \\
&= 1_A.
\end{aligned}$$

又知道

$$\begin{aligned}
P\left(\bigcup_{i=1}^n A_i\right) &= E\left(1_{\bigcup_{i=1}^n A_i}\right) \\
&= E\left(1 - \prod_{i=1}^n (1 - 1_{A_i})\right) \\
&= 1 - \prod_{i=1}^n (1 - E1_{A_i}) \\
&= \sum_{i=1}^n EA_i - \sum_{i<j} EA_i EA_j + \cdots + (-1)^{n-1} E\left(\bigcap_{i=1}^n A_i\right).
\end{aligned}$$

□

练习 1.6.10: Bonferroni inequalities. Let A_1, A_2, \dots, A_n be events and $A = \bigcup_{i=1}^n A_i$. Show that $1_A \leq \sum_{i=1}^n 1_{A_i}$, etc. and then take expected values to conclude

$$\begin{aligned}
P\left(\bigcup_{i=1}^n A_i\right) &\leq \sum_{i=1}^n P(A_i) \\
P\left(\bigcup_{i=1}^n A_i\right) &\geq \sum_{i=1}^n P(A_i) - \sum_{i<j} P(A_i \cap A_j) \\
P\left(\bigcup_{i=1}^n A_i\right) &\leq \sum_{i=1}^n P(A_i) - \sum_{i<j} P(A_i \cap A_j) + \sum_{i<j<k} P(A_i \cap A_j \cap A_k).
\end{aligned}$$

In general, if we stop the inclusion – exclusion formula after an even (odd) number of sums, we get an lower (upper) bound.

证明: 先证明 $1_A \leq \sum 1_{A_i}$ 。这很容易证明, 因为对 $\forall \omega \in \{1_A = 1\}$ 一定有 $\omega \in A$, 从而至少存在一个 A_{i_0} 使 $\omega \in A_{i_0}$, 从而

$$\sum_{i=1}^n 1_{A_i}(\omega) \geq 1_{A_{i_0}}(\omega) = 1 = 1_A(\omega).$$

对不等式两侧直接取期望就得到

$$P\left(\bigcup_{i=1}^n A_i\right) \leq \sum_{i=1}^n P(A_i).$$

剩余不等式可以类似证明。

□

练习 1.6.11: If $E|X|^k < \infty$ then for $0 < j < k$, $E|X|^j < \infty$, and furthermore

$$E|X|^j \leq (E|X|^k)^{\frac{j}{k}}.$$

证明: 令 $\varphi(x) = x^{\frac{j}{k}}$ 为凹函数, 那么利用 *Jensen* 不等式直接得到

$$E|X|^j \leq (E|X|^k)^{\frac{j}{k}} < \infty.$$

□

练习 1.6.12: Apply Jensen's inequality with $\varphi(x) = e^x$ and $P(X = \log y_m) = p(m)$ to conclude that if

$$\sum_{m=1}^n p(m) = 1$$

and $p(m), y_m > 0$ then

$$\sum_{m=1}^n p(m)y_m \geq \prod_{m=1}^n y_m^{p(m)}.$$

When $p(m) = 1/n$, this says the arithmetic mean exceeds the geometric mean.

证明: 有

$$\varphi(EX) = \prod_{m=1}^n y_m^{p(m)},$$

$$E\varphi(X) = \sum_{m=1}^n y_m p(m),$$

从而直接用 *Jensen* 不等式得出结论。

□

练习 1.6.13: If $EX_1^- < \infty$ and $X_n \uparrow X$ then $EX_n \uparrow EX$.

证明: 直接令 $X'_n = X_n + X_1^- \geq X_n + X_n^- = X_n^+$, 还有 $X'_n \uparrow X^+$, 那么就由单调收敛定理知道 $EX'_n \uparrow EX^+$. 对于 X_n^- , 注意到 $X_n^- \leq X_1^-$, 用 *Lebesgue* 收敛定理知道 $X_n^- \downarrow X^-$. 两个式子相减就得到结论。

□

练习 1.6.14: Let $X \geq 0$ but do NOT assume $E(1/X) < \infty$. Show

$$\lim_{y \rightarrow \infty} yE\left(\frac{1}{X}; X > y\right) = 0, \quad \lim_{y \downarrow 0} yE\left(\frac{1}{X}; X > y\right) = 0.$$

证明:

$$\begin{aligned}
yE\left(\frac{1}{X}; X > y\right) &\leq yE\left(\frac{1}{y}; X > y\right) \\
&\leq y \cdot \frac{1}{y} \cdot P(X > y) \\
&\leq P(X > y) \rightarrow 0 \quad (y \rightarrow \infty).
\end{aligned}$$

另一方面, 对 $\forall \varepsilon > y$, 有

$$\begin{aligned}
yE\left(\frac{1}{X}; X > y\right) &= E\left(\frac{y}{X}; y < X < \varepsilon\right) + E\left(\frac{y}{X}; X \geq \varepsilon\right) \\
&\leq P(y < X < \varepsilon) + E\left(\frac{y}{X}; X \geq \varepsilon\right).
\end{aligned}$$

令 $y \rightarrow 0$, 由有界收敛定理知道 $E(y/X; X \geq \varepsilon) \rightarrow 0$; 再令 $\varepsilon \rightarrow 0$, 得到 $P(y < X < \varepsilon) \rightarrow 0$, 从而得到不等式的右侧均收敛于 0. \square

练习 1.6.15: If $X_n \geq 0$ then

$$E\left(\sum_{n=0}^{\infty} X_n\right) = \sum_{n=0}^{\infty} EX_n.$$

证明: 令 $Y_n = \sum_{i=0}^n X_i$, 那么 $Y_n \uparrow Y = \sum_{i=0}^{\infty} X_i$, 只要证明

$$E(Y) = \lim E(Y_n),$$

这可以显然由单调收敛定理得出. \square

练习 1.6.16: If X is integrable and A_n are disjoint sets with union A then

$$\sum_{n=0}^{\infty} E(X; A_n) = E(X; A) \quad \text{i.e. ,}$$

the sum converges absolutely and has the value on the right.

证明: 对 $X^+1_{A_n}$ 和 $X^-1_{A_n}$ 分别使用上一题的结论, 然后相减 (因为 X 可积所以可以相减) 直接得出. \square

1.7. 乘积测度和 Fubini 定理

练习 1.7.1: If

$$\int_X \int_Y |f(x, y)| \mu_2(dy) \mu_1(dx) < \infty$$

then

$$\int_X \int_Y f(x, y) \mu_2(dy) \mu_1(dx) = \int_{X \times Y} f d(\mu_1 \times \mu_2) = \int_Y \int_X f(x, y) \mu_1(dx) \mu_2(dy).$$

Corollary Let $X = \{1, 2, \dots\}$, \mathcal{A} = all subsets of X , and μ_1 = counting measure. If

$$\sum_n \int |f_n| d\mu < \infty$$

then

$$\sum_n \int f_n d\mu = \int \sum_n f_n d\mu.$$

证明：因为 $|f| \geq 0$ ，所以直接对 $|f|$ 用 Fubini 定理知道

$$\int_X \int_Y |f(x, y)| \mu_2(dy) \mu_1(dx) = \int_{X \times Y} |f(x, y)| d(\mu_1 \times \mu_2) < \infty,$$

从而 f 满足 Fubini 定理条件，再对 f 用 Fubini 定理直接得到结论。 \square

练习 1.7.2: Let $g \geq 0$ be a measurable function on (X, \mathcal{A}, μ) . Use Theorem 1.7.2 to conclude that

$$\int_X g d\mu = (\mu \times \lambda)(\{(x, y) : 0 \leq y < g(x)\}) = \int_0^\infty \mu(\{x : g(x) > y\}) dy.$$

证明：注意到 $f(x, y) = 1_{\{g(x) > y\}} \geq 0$ 符合 Fubini 定理条件，从而有

$$\int_X \int_0^\infty 1_{\{g(x) > y\}} d\lambda d\mu = \int_{X \times [0, \infty)} 1_{\{g(x) > y\}} d(\mu \times \lambda) = \int_0^\infty \int_X 1_{\{g(x) > y\}} d\mu d\lambda.$$

容易验证这就是要证的等式。 \square

练习 1.7.3: Let F, G be Stieltjes measure functions and let μ, ν be the corresponding measures on $(\mathbb{R}, \mathcal{R})$. Show that

$$1. \quad \int_{(a, b]} \{F(y) - F(a)\} dG(y) = (\mu \times \nu)(\{(x, y) : a < x \leq y \leq b\});$$

$$2. \quad \begin{aligned} & \int_{(a, b]} F(y) dG(y) + \int_{(a, b]} G(y) dF(y) \\ &= F(b)G(b) - F(a)G(a) + \sum_{x \in (a, b]} \mu(\{x\})\nu(\{x\}); \end{aligned}$$

3. If $F = G$ is continuous then

$$\int_{(a,b]} 2F(y) dF(y) = F^2(b) - F^2(a).$$

To see the second term in (2) is needed, let $F(x) = G(x) = 1_{[0,\infty)}(x)$ and $a < 0 < b$.

证明:

1. 注意到 $f(x, y) = 1_{\{x \leq y\}} \geq 0$ 符合 *Fubini* 定理条件, 从而有

$$\int_{(a,b]} \int_{(a,b]} 1_{\{x \leq y\}} dF dG = \int_{(a,b]^2} 1_{\{x \leq y\}} d(\mu \times \nu),$$

也即欲证的等式。

2. 由于 F, G 是 *Stieltjes* 测度函数, 所以 F, G 均只在至多可数个点处不连续, 所以

$$\sum_{x \in (a,b]} \mu(\{x\})\nu(\{x\})$$

为至多可数个点相加, 从而有意义。设 F, G 分别对应定义在 $(\Omega, \mathcal{F}, \gamma)$ 上的随机变量 X, Y 。设 F, G 在 $(a, b]$ 上共同的不连续点集为 A , 那么在 $(a, b] \setminus A$ 上有

$$\int_{(a,b] \setminus A} F dG + \int_{(a,b] \setminus A} G dF = \int_{(a,b] \setminus A} d(FG).$$

再考虑在 A 上的积分情况。有

$$\begin{aligned} \int_A F dG &= \int_{G^{-1}(A)} F \circ G d\gamma \\ &= \sum_{\omega \in G^{-1}(A)} \gamma(\{\omega\}) \cdot F \circ G(\omega) \\ &= \sum_{x \in A} \nu(\{x\}) \cdot \mu(\{x\}). \end{aligned}$$

同理对 $\int_A G dF$ 得到相同的结果。那么相加就得到所求结论。

3. 因为 $F = G$ 连续所以它们 *Riemann* 可积, 那么直接应用 *Riemann* 积分结论直接得出。

□

练习 1.7.4: Let μ be a finite measure on \mathbb{R} and $F(x) = \mu((-\infty, x])$. Show that

$$\int (F(x+c) - F(x)) dx = c\mu(\mathbb{R}).$$

证明: 注意到 $1_{\{x < y \leq x+c\}} \geq 0$ 满足 *Fubini* 定理的条件, 从而直接应用 *Fubini* 定理得到

$$\int_{\mathbb{R}} \int_{\mathbb{R}} 1_{\{x < y \leq x+c\}} \mu(dy) \lambda(dx) = \int_{\mathbb{R}} \int_{\mathbb{R}} 1_{\{x < y \leq x+c\}} \lambda(dx) \mu(dy),$$

其中 λ 为 $(\mathbb{R}, \mathcal{R})$ 上的 Lebesgue 测度。上式中

$$\begin{aligned} \text{LHS} &= \int_{\mathbb{R}} \int_{(x, x+c]} dF \lambda(dx) \\ &= \int_{\mathbb{R}} (F(x+c) - F(x)) dx, \\ \text{RHS} &= \int_{\mathbb{R}} \int_{[y-c, y)} dx \mu(dy) \\ &= c \int_{\mathbb{R}} \mu(dy) \\ &= c \int_{\mathbb{R}} dF \\ &= c\mu(\mathbb{R}). \end{aligned}$$

□

练习 1.7.5: Show that $e^{-xy} \sin x$ is integrable in the strip $0 < x < a, 0 < y$. Perform the double - integral in the two orders to get:

$$\int_0^a \frac{\sin x}{x} dx = \arctan(a) - (\cos a) \int_0^\infty \frac{e^{-ay}}{1+y^2} dy - (\sin a) \int_0^\infty \frac{ye^{-ay}}{1+y^2} dy$$

and replace $1+y^2$ by 1 to conclude

$$\left| \int_0^a \frac{\sin x}{x} dx - \arctan(a) \right| \leq \frac{2}{a}$$

for $a \geq 1$.

证明：有

$$\begin{aligned} \int_0^a \int_0^\infty e^{-xy} \sin x dy dx &= - \int_0^a \frac{\sin x}{x} \int_0^\infty e^{-xy} d(-xy) dx \\ &= \int_0^a \frac{\sin x}{x} dx. \end{aligned}$$

由于

$$\int_0^a \sin x dx \text{ 有界, } \frac{1}{x} \downarrow 0,$$

由 Dirichlet 判别法知道

$$\int_0^a \frac{\sin(x)}{x} dx$$

Riemann 可积, 从而 *Lebesgue* 可积。那么可以对 $e^{-xy} \sin x$ 使用 *Fubini* 定理得到

$$\int_0^a \int_0^\infty e^{-xy} \sin x dy dx = \int_0^\infty \int_0^a e^{-xy} \sin x dx dy.$$

上式中

$$\begin{aligned} \text{LHS} &= \int_0^a \frac{\sin x}{x} dx, \\ \text{RHS} &= \int_0^\infty \left. \frac{-e^{-xy}(\cos x + y \sin x)}{1 + y^2} \right|_0^a dy \\ &= \int_0^\infty \left(\frac{1}{1 + y^2} - \frac{e^{-ay} \cos a}{1 + y^2} - \frac{e^{-ay} y \sin a}{1 + y^2} \right) dy \end{aligned}$$

即为欲证之结论。

□

2. 大数定律

2.1. 独立性

定义 11 (独立):

1. 称事件 A, B 独立, 如果 $P(A \cap B) = P(A)P(B)$ 。
2. 称随机变量 X, Y 独立, 如果对 $\forall C, D \in \mathcal{R}$ 有

$$P(X \in C, Y \in D) = P(X \in C)P(Y \in D).$$

3. 称 σ -域 \mathcal{F}, \mathcal{G} 独立, 如果对 $\forall A \in \mathcal{F}$ 和 $B \in \mathcal{G}$ 事件 A, B 独立。

练习 2.1.1: Suppose (X_1, \dots, X_n) has density $f(x_1, x_2, \dots, x_n)$, that is

$$P((X_1, X_2, \dots, X_n) \in A) = \int_A f(x) dx \quad \text{for } A \in \mathcal{R}^n.$$

If $f(x)$ can be written as $g_1(x_1) \cdots g_n(x_n)$ where the $g_m \geq 0$ are measurable, then X_1, X_2, \dots, X_n are independent. Note that the g_m are not assumed to be probability densities.

证明: 对 $\Omega = \Omega_1 \times \dots \times \Omega_n$ 中的一个长方体 $A = A_1 \times \dots \times A_n$, 有

$$\begin{aligned} P(X_1 \in A_1, \dots, X_n \in A_n) &= \int_A f(x_1, \dots, x_n) dx_1 \cdots dx_n \\ &= \int_A \prod_{i=1}^n g_i(x_i) dx_1 \cdots dx_n \\ &= \prod_{i=1}^n \int_{A_i} g_i(x_i) dx_i \quad (\text{Fubini 定理}). \end{aligned}$$

对 $\forall k \in \{1, 2, \dots, n\}$ 有

$$\begin{aligned} P(X_k \in A_k) &= P(X_1 \in \Omega_1, \dots, X_k \in A_k, \dots, X_n \in \Omega_n) \\ &= \left(\int_{A_k} g_k(x_k) dx_k \right) \cdot \prod_{i \neq k} \int_{\Omega_i} g_i(x_i) dx_i. \end{aligned}$$

又知道

$$1 = \int_{\Omega} f dx_1 \cdots dx_n = \prod_{i=1}^n \int_{\Omega_i} g_i(x_i) dx_i,$$

那么

$$\begin{aligned}
\prod_{i=1}^n P(X_i \in A_i) &= \left(\prod_{i=1}^n \int_{A_i} g_i(x_i) dx_i \right) \cdot \left(\prod_{i=1}^n \int_{\Omega_i} g_i(x_i) dx_i \right)^{n-1} \\
&= \left(\prod_{i=1}^n \int_{A_i} g_i(x_i) dx_i \right) \cdot 1^{n-1} \\
&= \prod_{i=1}^n \int_{A_i} g_i(x_i) dx_i \\
&= P(X_1 \in A_1, \dots, X_n \in A_n).
\end{aligned}$$

□

练习 2.1.2: Suppose X_1, \dots, X_n are random variables that take values in countable sets S_1, \dots, S_n . Then in order for X_1, \dots, X_n to be independent, it is sufficient that whenever $x_i \in S_i$,

$$P(X_1 = x_1, \dots, X_n = x_n) = \prod_{i=1}^n P(X_i = x_i).$$

证明: 令 \mathcal{A}_i 为 S_i 中所有单元素集和 \emptyset 构成的集合。那么 \mathcal{A}_i 是 π -系, 只要证明 $\mathcal{A}_1, \mathcal{A}_2, \dots, \mathcal{A}_n$ 独立, 那么 $\sigma(\mathcal{A}_1), \dots, \sigma(\mathcal{A}_n)$ 自然互相独立, 从而得证。而从题目条件自然得出 \mathcal{A}_i 是独立的。 □

练习 2.1.3: Let $\rho(x, y)$ be a metric.

1. Suppose h is differentiable with $h(0) = 0$, $h'(x) > 0$ for $x > 0$ and $h'(x)$ decreasing on $[0, \infty)$. Then $h(\rho(x, y))$ is a metric.
2. $h(x) = x/(x+1)$ satisfies the hypotheses in (1).

证明:

1. 条件说明 h 为单调增的凹函数。 $\forall x, y, z$, 只要证明

$$h(\rho(x, y)) + h(\rho(y, z)) \geq h(\rho(x, z)).$$

这是因为

$$\begin{aligned}
h(\rho(x, y)) + h(\rho(y, z)) &\geq h(\rho(x, y) + \rho(y, z)) \\
&\geq h(\rho(x, z)).
\end{aligned}$$

2. 直接验证即可。

□

练习 2.1.4: Let $\Omega = (0, 1)$, \mathcal{F} = Borel sets, P = Lebesgue measure. Then

$$X_n(\omega) = \sin(2\pi n\omega), \quad n = 1, 2, \dots$$

are uncorrelated but not independent.

证明: 对 $\forall n, m \in \mathbb{N}^+$ 且 $n \neq m$ 有

$$\begin{aligned} EX_n &= \int_0^1 \sin(2\pi nx) dx = 0, \\ EX_m &= 0, \\ EX_n X_m &= \int_0^1 \sin(2\pi nx) \sin(2\pi mx) dx \\ &= \int_0^1 \frac{\cos(2\pi(n-m)x) + \cos(2\pi(n+m)x)}{2} dx = 0. \end{aligned}$$

所以有 $EX_n X_m = EX_n EX_m$, 从而它们无关。但是又考虑 X_2, X_3 在 $A = (0, 1/3)$ 上的期望有

$$\begin{aligned} E(X_2; A) &= \frac{3}{8\pi}, \\ E(X_3; A) &= 0, \\ E(X_2 X_3; A) &= \frac{3\sqrt{3}}{20\pi}, \end{aligned}$$

从而它们无关。 □

练习 2.1.5:

1. Show that if X and Y are independent with distributions μ and ν then

$$P(X + Y = 0) = \sum_y \mu(\{-y\})\nu(\{y\}).$$

2. Conclude that if X has continuous distribution $P(X = Y) = 0$.

证明:

1.

$$\begin{aligned} P(X + Y = 0) &= \iint 1_{\{X+Y=0\}} \mu(dx) \nu(dy) \\ &= \int \mu(\{-y\}) \nu(dy) \\ &= \sum_y \mu(\{-y\}) \nu(\{y\}). \end{aligned}$$

2. 与上一问同理可以证明

$$P(X = Y) = \sum_y \mu(\{y\}) \nu(\{y\}),$$

等式右侧只有在 μ, ν 都不为 0, 即 X, Y 的分布都不连续的地方才不等于 0, 而这永不成立, 因为 X 的分布是连续的。

□

练习 2.1.6: Prove directly from the definition that if X and Y are independent and f and g are measurable functions then $f(X)$ and $g(Y)$ are independent.

证明: 对 $\forall A, B \in \mathcal{R}$, 有

$$\begin{aligned} P(f(X) \in A)P(g(Y) \in B) &= P(X \in f^{-1}(A))P(Y \in g^{-1}(B)) \\ &= P(X \in f^{-1}(A), Y \in g^{-1}(B)) \\ &= P(f(X) \in A, g(Y) \in B). \end{aligned}$$

□

练习 2.1.7: Let $K \geq 3$ be a prime and let X and Y be independent random variables that are uniformly distributed on $\{0, 1, \dots, K-1\}$. For $0 \leq n < K$, let

$$Z_n = X + nY \bmod K.$$

Show that Z_0, Z_1, \dots, Z_{K-1} are pairwise independent, i.e., each pair is independent. They are not independent because if we know the values of two of the variables then we know the values of all the variables.

证明: 由数论相关知识知道对于 $\forall 0 \leq n < K$ 有 $Z_n = X + nY = p$ 当且仅当 $Y = q(X)$ 存在 $q(X)$, 即此时的 q 是唯一的, 所以

$$\begin{aligned} P(Z_n = p) &= \sum_x P(X = x, Y = q(x)) \\ &= \sum_x P(X = x)P(Y = q(x)) \\ &= \sum_x \frac{1}{K^2} = \frac{1}{K} \end{aligned}$$

对 $\forall p \in \{0, 1, \dots, K-1\}$ 成立。那么对 $\forall a, b \in \{0, 1, \dots, K-1\}$ 且 $a \neq b$, 考虑 Z_a 和 Z_b 之间的独立性。有

$$P(Z_a = p) = P(Z_b = q) = \frac{1}{K}, \quad P(Z_a = p, Z_b = q) = \frac{1}{K^2}$$

从而 Z_a, Z_b 独立。

□

练习 2.1.8: Find four random variables taking values in $\{-1, 1\}$ so that any three are independent but all four are not. Hint: Consider products of independent random variables.

证明: 记 X, Y, Z, W 为满足题意的随机变量, 且 $A = \{X = 1\}, B = \{Y = 1\}, C = \{Z = 1\}, D = \{W = 1\}, P(A) = a, P(B) = b, P(C) = c, P(D) = d$ 。以下推导所有符合题意的随机变量。

设 $P(A \cap B \cap C \cap D) = x$, 那么有

1. $P(A^c \cap B \cap C \cap D) = bcd - x;$
2. $P(A \cap B^c \cap C \cap D) = acd - x,$
 $P(A^c \cap B^c \cap C \cap D) = (1 - a)cd - bcd + x;$
3. $P(A \cap B \cap C^c \cap D) = abd - x,$
 $P(A \cap B^c \cap C^c \cap D) = a(1 - b)d - acd + x,$
 $P(A^c \cap B \cap C^c \cap D) = (1 - a)bd - bcd + x,$
 $P(A^c \cap B^c \cap C^c \cap D) = (1 - a)(1 - b)d - (1 - a)cd + bcd - x;$
4. 类似推导所有带有 D^c 的情况。这样只要令 $x \neq abcd$ 的值即可保证

$$P(A \cap B \cap C \cap D) = x \neq abcd = P(A)P(B)P(C)P(D),$$

但是任意三个集合是相互独立的。

□

练习 2.1.9: Let $\Omega = \{1, 2, 3, 4\}$, \mathcal{F} = all subsets of Ω , and $P(\{i\}) = 1/4$. Give an example of two collections of sets \mathcal{A}_1 and \mathcal{A}_2 that are independent but whose generated σ -fields are not.

证明: 令 $\mathcal{A}_1 = \{\{1, 2\}, \{1, 3\}\}, \mathcal{A}_2 = \{\{2, 3\}\}$, 那么

$$P(\{1, 2\} \cap \{2, 3\}) = \frac{1}{2} = P(\{1, 2\})P(\{2, 3\}),$$

$$P(\{1, 3\} \cap \{2, 3\}) = \frac{1}{2} = P(\{1, 3\})P(\{2, 3\}),$$

从而 \mathcal{A}_1 和 \mathcal{A}_2 独立。但是有 $\{2, 3\} \in \sigma(\mathcal{A}_1)$, 这时

$$P(\{2, 3\} \cap \{2, 3\}) = P(\{2, 3\}) = \frac{1}{2} \neq P(\{2, 3\})^2,$$

从而 $\sigma(\mathcal{A}_1)$ 和 $\sigma(\mathcal{A}_2)$ 不独立。

□

练习 2.1.10: Show that if X and Y are independent, integer-valued random variables, then

$$P(X + Y = n) = \sum_m P(X = m)P(Y = n - m).$$

证明: 因为 $1_{\{X+Y=n\}} \geq 0$, 所以可以直接应用 Fubini 定理, 从而有

$$\begin{aligned}
P(X + Y = n) &= \iint 1_{\{x+y=n\}} d\mu d\nu \\
&= \int \left(\int 1_{\{x+y=n\}} d\mu \right) d\nu \\
&= \int \mu(n - y) d\nu \\
&= \sum_m \mu(n - m) \nu(m) \\
&= \sum_m P(X = m) P(Y = n - m).
\end{aligned}$$

□

练习 2.1.11: In Example 1.6.13, we introduced the Poisson distribution with parameter λ , which is given by

$$P(Z = k) = \frac{e^{-\lambda} \lambda^k}{k!}$$

for $k = 0, 1, 2, \dots$. Use the previous exercise to show that if $X = \text{Poisson}(\lambda)$ and $Y = \text{Poisson}(\mu)$ are independent then $X + Y = \text{Poisson}(\lambda + \mu)$.

证明: $\forall n \in \mathbb{N}$, 则有

$$\begin{aligned}
P(X + Y = n) &= \sum_m P(X = m) P(Y = n - m) \\
&= \sum_{m=0}^n \frac{e^{-\lambda} \lambda^m}{m!} \frac{e^{-\mu} \mu^{n-m}}{(n-m)!} \\
&= e^{-(\lambda+\mu)} \sum_{m=0}^n \frac{\lambda^m \mu^{n-m}}{m!(n-m)!} \\
&= e^{-(\lambda+\mu)} \frac{(\lambda + \mu)^n}{n!}.
\end{aligned}$$

最后一步利用了二项式定理。

□

练习 2.1.12: X is said to have a Binomial(n, p) distribution if

$$P(X = m) = \binom{n}{m} p^m (1 - p)^{n-m}.$$

1. Show that if $X = \text{Binomial}(n, p)$ and $Y = \text{Binomial}(m, p)$ are independent then $X + Y = \text{Binomial}(n + m, p)$.
2. Look at Example 1.6.12 and use induction to conclude that the sum of n independent Bernoulli(p) random variables is Binomial(n, p).

证明:

1.

$$\begin{aligned}
 P(X + Y = x) &= \sum_{y=0}^x P(X = y)P(Y = x - y) \\
 &= \sum_{y=0}^x \binom{n}{y} p^y (1-p)^{n-y} \binom{m}{x-y} p^{x-y} (1-p)^{m-x+y} \\
 &= p^x (1-p)^{n+m-x} \sum_{y=0}^x \binom{n}{y} \binom{m}{x-y} \\
 &= \binom{n+m}{x} p^x (1-p)^{n+m-x}.
 \end{aligned}$$

2. $n = 1$ 的情况是显然的。对 $n > 1$ 的情况, 记 $X = X_1 + \cdots + X_{n-1}$ 为 $(n-1)$ 个 $Bernoulli(p)$ 分布的随机变量的和, 归纳假设 X 服从 $\text{Binomial}(n-1, p)$. 设 X_n 也服从 $Bernoulli(p)$ 分布, 只要证明 $X + X_n$ 服从 $\text{Binomial}(n, p)$ 分布。有

$$\begin{aligned}
 P(X_n + X = x) &= P(X = x)P(X_n = 0) + P(X = x-1)P(X_n = 1) \\
 &= \binom{n-1}{x} p^x (1-p)^{n-x-1} (1-p) + \binom{n-1}{x-1} p^{x-1} (1-p)^{n-x} p \\
 &= \left(\binom{n-1}{x} + \binom{n-1}{x-1} \right) p^x (1-p)^{n-x} \\
 &= \binom{n}{x} p^x (1-p)^{n-x}.
 \end{aligned}$$

□

练习 2.1.13: It should not be surprising that the distribution of $X + Y$ can be $F * G$ without the random variables being independent. Suppose $X, Y \in \{0, 1, 2\}$ and take each value with probability $1/3$.

1. Find the distribution of $X + Y$ assuming X and Y are independent.
2. Find all the joint distributions (X, Y) so that the distribution of $X + Y$ is the same as the answer to (1).

证明:

1.

$$\begin{aligned}
 P(X + Y = 0) &= P(X = 0)P(Y = 0) = \frac{1}{9}, \\
 P(X + Y = 1) &= P(X = 0)P(Y = 1) + P(X = 1)P(Y = 0) = \frac{2}{9}, \\
 P(X + Y = 2) &= \frac{3}{9}, \\
 P(X + Y = 3) &= \frac{2}{9}, \\
 P(X + Y = 4) &= \frac{1}{9}.
 \end{aligned}$$

分布函数略。

2. 直接设 $P(X = 1, Y = 0) = x \leq \frac{2}{9}$, 然后根据题意和上一问列方程求解得到

$X \setminus Y$	0	1	2
0	$\frac{1}{9}$	$\frac{2}{9} - x$	x
1	x	$\frac{1}{9}$	$\frac{2}{9} - x$
2	$\frac{2}{9} - x$	x	$\frac{1}{9}$

即为所求的所有联合分布。

□

练习 2.1.14: Let $X, Y \geq 0$ be independent with distribution functions F and G . Find the distribution function of XY .

证明: 有

$$\begin{aligned}
 P(XY \leq p) &= \iint 1_{\{xy \leq p\}} dF dG \\
 &= \int_0^\infty \int_0^{p/y} dF dG \\
 &= \int_0^\infty F\left(\frac{p}{y}\right) dG.
 \end{aligned}$$

□

练习 2.1.15: If we want an infinite sequence of coin tossings, we do not have to use Kolmogorov's theorem. Let Ω be the unit interval $(0, 1)$ equipped with the Borel sets \mathcal{F} and Lebesgue measure P . Let $Y_n(\omega) = 1$ if $[2^n \omega]$ is odd and 0 if $[2^n \omega]$ is even. Show that Y_1, Y_2, \dots are independent with $P(Y_k = 0) = P(Y_k = 1) = 1/2$.

证明: $P(Y_k = 0) = 1/2$ 容易证明。下证明 Y_1, Y_2, \dots 独立。对 $\forall A \in \{1, 2, \dots\}$, 其中 A 为有限集, 只要证明 $\{Y_a : a \in A\}$ 独立。记 $A = \{a_1, \dots, a_m\}$, 其中 $a_1 < \dots < a_m$, 那么只要证明

$$P(Y_{a_1} = 1, Y_{a_2} = 1, \dots, Y_{a_m} = 1) = P(Y_{a_1} = 1)P(Y_{a_2} = 1, \dots, Y_{a_m} = 1),$$

然后归纳即可。而

$$\begin{aligned}
P(Y_{a_1} = 1, Y_{a_2} = 1, \dots, Y_{a_m} = 1) &= \frac{1}{2} P(Y_{a_1} = 1, Y_{a_2} = 1, \dots, Y_{a_{m-1}} = 1) \\
&= \frac{1}{2^2} P(Y_{a_1} = 1, Y_{a_2} = 1, \dots, Y_{a_{m-2}} = 1) \\
&= \dots \\
&= \frac{1}{2^m}, \\
P(Y_{a_2} = 1, Y_{a_3} = 1, \dots, Y_{a_m} = 1) &= \frac{1}{2^{m-1}},
\end{aligned}$$

那么自然有

$$\begin{aligned}
P(Y_{a_1} = 1, Y_{a_2} = 1, \dots, Y_{a_m} = 1) &= \frac{1}{2} P(Y_{a_2} = 1, \dots, Y_{a_m} = 1) \\
&= P(Y_{a_1} = 1) P(Y_{a_2} = 1, \dots, Y_{a_m} = 1).
\end{aligned}$$

□

2.2. 弱大数定律

练习 2.2.1: Let X_1, X_2, \dots be uncorrelated with $EX_i = \mu_i$ and $\text{var}(X_i)/i \rightarrow 0$ as $i \rightarrow \infty$. Let $S_n = X_1 + \dots + X_n$ and $\nu_n = ES_n/n$ then as $n \rightarrow \infty$,

$$\frac{S_n}{n} - \nu_n \rightarrow 0$$

in L^2 and in probability.

证明:

$$\begin{aligned}
E\left(\frac{S_n}{n} - E\frac{S_n}{n}\right)^2 &= \text{var} \frac{S_n}{n} \\
&= \frac{\text{var} S_n}{n^2} \\
&= \frac{\sum_i \text{var} X_i}{n^2} \\
&\leq \frac{\sum_i \frac{\text{var} X_i}{i}}{n}.
\end{aligned}$$

令 $n \rightarrow \infty$ 并用 *Stolz* 定理知道

$$E\left(\frac{S_n}{n} - E\frac{S_n}{n}\right)^2 \rightarrow 0,$$

从而题目中随机变量 L^2 收敛到 0, 进而依测度收敛。

□

练习 2.2.2: The L^2 weak law generalizes immediately to certain dependent sequences. Suppose $EX_n = 0$ and $EX_n X_m \leq r(n-m)$ for $m \leq n$ (no absolute value on the left-hand side!) with $r(k) \rightarrow 0$ as $k \rightarrow \infty$. Show that

$$\frac{X_1 + \cdots + X_n}{n} \rightarrow 0$$

in probability.

证明:

$$\begin{aligned} E\left(\frac{(X_1 + \cdots + X_n)^2}{n}\right) &= \frac{\sum_{i \leq j} E(X_i X_j)}{n^2} \\ &\leq \frac{\sum_{i \leq j} r(j-i)}{n^2} \\ &\leq \frac{nr(0) + (n-1)r(1) + \cdots + r(n-1)}{n^2} \\ &\rightarrow \frac{r(0) + r(1) + \cdots + r(n)}{2n+1} \\ &\rightarrow \frac{r(n)}{2} \rightarrow 0 \quad (n \rightarrow \infty). \end{aligned}$$

其中两次使用了 *Stolz* 定理。那么它 L^2 收敛，从而依测度收敛。 □

练习 2.2.3: Monte Carlo integration.

1. Let f be a measurable function on $[0, 1]$ with

$$\int_0^1 |f(x)| dx < \infty.$$

Let U_1, U_2, \dots be independent and uniformly distributed on $[0, 1]$, and let

$$I_n = \frac{f(U_1) + \cdots + f(U_n)}{n}.$$

Show that

$$I_n \rightarrow I \equiv \int_0^1 f dx$$

in probability.

2. Suppose

$$\int_0^1 |f(x)|^2 dx < \infty.$$

Use Chebyshev's inequality to estimate

$$P\left(|I_n - I| > \frac{a}{n^{\frac{1}{2}}}\right).$$

证明:

1. 因为 U_1, U_2, \dots 是独立同分布序列, 所以 $f(U_1), f(U_2), \dots$ 独立同分布, 又因为 $E(f(U_n)) < \infty$, 所以可以直接应用弱大数定律得到结论。

2.

$$\begin{aligned} P\left(|I_n - I| > \frac{a}{\sqrt{n}}\right) &= P\left((I_n - I)^2 > \frac{a^2}{n}\right) \\ &\leq \frac{n \operatorname{var}(I_n)}{a^2} \\ &= \frac{n \cdot \frac{n \sigma_{f(U_i)}}{n^2}}{a^2} \\ &= \frac{\sigma_{f(U_i)}^2}{a^2}. \end{aligned}$$

□

练习 2.2.4: Let X_1, X_2, \dots be i.i.d. with

$$P(X_i = (-1)^k k) = \frac{C}{k^2 \log k}$$

for $k \geq 2$ where C is chosen to make the sum of the probabilities = 1. Show that $E|X_i| = \infty$, but there is a finite constant μ so that $S_n/n \rightarrow \mu$ in probability.

证明: 容易看出

$$E|X_i| = \sum_{k=0}^{\infty} \frac{Ck}{k^2 \log k} = \sum_{k=0}^{\infty} \frac{c}{k \log k} = \infty.$$

再证明下一结论, 考虑使用弱大数定律。对足够大的 n , 有

$$\begin{aligned} nP(|X_i| > n) &= n \sum_{k=n}^{\infty} \frac{1}{k^2 \log k} \\ &\leq \frac{n}{\log n} \sum_{k=n}^{\infty} \frac{1}{k^2} \\ &\leq \frac{n}{\log n} \int_{n-1}^{\infty} \frac{1}{x^2} dx \\ &= \frac{n}{(n-1) \log n} \rightarrow 0 \quad (n \rightarrow \infty), \end{aligned}$$

从而满足弱大数定律条件，从而有

$$\left| \frac{S_n}{n} - \mu_n \right| \xrightarrow{P} 0,$$

其中

$$\begin{aligned} \mu_n &= \mathbb{E} X_i 1_{\{|X_i| \leq n\}} \\ &= \sum_{k=0}^n \frac{(-1)^k k}{k^2 \log k} \\ &= \sum_{k=0}^n \frac{(-1)^k}{k \log k} \end{aligned}$$

为交错级数，从而收敛。设 $\mu_n \rightarrow \mu$ ，则有

$$\left| \frac{S_n}{n} - \mu \right| \leq \left| \frac{S_n}{n} - \mu_n \right| + |\mu_n - \mu| \xrightarrow{P} 0.$$

□

练习 2.2.5: Let X_1, X_2, \dots be i.i.d. with

$$P(X_i > x) = \frac{e}{x \log x}$$

for $x \geq e$. Show that $\mathbb{E}|X_i| = \infty$, but there is a sequence of constants $\mu_n \rightarrow \infty$ so that

$$\frac{S_n}{n} - \mu_n \rightarrow 0$$

in probability.

证明：

$$\begin{aligned} \mathbb{E}|X_i| &= \mathbb{E} X_i \\ &= \int_0^\infty P(X_i > y) dy \\ &= \int_0^\infty \frac{e}{y \log y} dy \\ &= \infty. \end{aligned}$$

因为

$$xP(X_i > x) = \frac{e}{x} \rightarrow 0 \quad (x \rightarrow \infty),$$

所以 X_1, X_2, \dots 满足弱大数定律条件，那么取

$$\begin{aligned}\mu_n &= EX_1 1_{\{|X_1| \leq n\}} \\ &\rightarrow EX_1 = \infty \quad (\text{单调收敛定理})\end{aligned}$$

即可，由弱大数定律直接得到

$$\frac{S_n}{n} - \mu_n \rightarrow 0.$$

□

练习 2.2.6:

1. Show that if $X \geq 0$ is integer - valued

$$EX = \sum_{n \geq 1} P(X \geq n).$$

2. Find a similar expression for EX^2 .

证明:

1.

$$\begin{aligned}EX &= \int_0^{\infty} P(X > y) dy \\ &= \sum_{n=0}^{\infty} \int_n^{n+1} P(X \geq n+1) dy \\ &= \sum_{n=0}^{\infty} P(X \geq n+1) \\ &= \sum_{n=1}^{\infty} P(X \geq n).\end{aligned}$$

2.

$$\begin{aligned}EX^2 &= \int_0^{\infty} yP(X > y) dy \\ &= \sum_{n=0}^{\infty} P(X \geq n+1) \int_n^{n+1} y dy \\ &= \sum_{n=0}^{\infty} \left(n + \frac{1}{2}\right) P(X \geq n+1) \\ &= \sum_{n=1}^{\infty} \left(n - \frac{1}{2}\right) P(X \geq n).\end{aligned}$$

□

练习 2.2.7: Generalize Lemma 2.2.13 to conclude that if

$$H(x) = \int_{(-\infty, x]} h(y) dy$$

with $h(y) \geq 0$, then

$$EH(X) = \int_{-\infty}^{\infty} h(y)P(X \geq y) dy.$$

An important special case is $H(x) = \exp(\theta x)$ with $\theta > 0$.

证明:

$$\begin{aligned} EH(X) &= \int_{\Omega} H(X(\omega)) dP \\ &= \int_{\Omega} \int_{-\infty}^{\infty} h(t)1_{\{t \leq X(\omega)\}} dt dP \\ &= \int_{-\infty}^{\infty} \int_{\Omega} h(t)1_{\{t \leq X(\omega)\}} dP dt \quad (\text{Fubini 定理}) \\ &= \int_{-\infty}^{\infty} h(t)P(X \geq t) dt. \end{aligned}$$

□

练习 2.2.8: An unfair “fair game”. Let

$$p_k = \frac{1}{2^k k(k+1)},$$

$k = 1, 2, \dots$ and $p_0 = 1 - \sum_{k \geq 1} p_k$.

$$\sum_{k=1}^{\infty} 2^k p_k = \left(1 - \frac{1}{2}\right) + \left(\frac{1}{2} - \frac{1}{3}\right) + \dots = 1,$$

so if we let X_1, X_2, \dots be i.i.d. with $P(X_n = -1) = p_0$ and

$$P(X_n = 2^k - 1) = p_k \quad \text{for } k \geq 1$$

then $EX_n = 0$. Let $S_n = X_1 + \dots + X_n$. Use Theorem 2.2.11 with $b_n = 2^{m(n)}$ where $m(n) = \min\{m : 2^{-m} m^{-3/2} \leq n^{-1}\}$ to conclude that

$$\frac{S_n}{n/\log_2 n} \rightarrow -1 \quad \text{in probability.}$$

证明: 要使用三角随机变量列的弱大数定律, 需要先验证 $\sum_{k=1}^n P(X_k > b_n) \rightarrow 0$ 。这是因为

$$\begin{aligned}
\sum_{i=1}^n P(X_k > b_n) &= nP(X_1 > 2^m) \\
&= n \sum_{k=m+1}^{\infty} \frac{1}{2^k k(k+1)} \\
&\leq 2^m m^{3/2} \sum_{k=m+1}^{\infty} \frac{1}{2^k k^2} \\
&\leq \frac{1}{\sqrt{m}} \sum_{k=m+1}^{\infty} \frac{1}{2^{k-m}} \\
&= \frac{2}{\sqrt{m}} \\
&\rightarrow 0 \quad (n \rightarrow \infty).
\end{aligned}$$

上式最后一步中 $n \rightarrow \infty$ 时 $m \rightarrow \infty$ 。另外还需要验证 $b_n^{-2} \sum_{k=1}^n E\bar{X}_k^2 \rightarrow 0$ ，这是因为

$$\begin{aligned}
\frac{\sum_{k=1}^n E\bar{X}_k^2}{b_n^2} &= \frac{n(P(X_1 = -1) + \sum_{k=1}^m (2^{2k} - 1)p_k)}{2^{2m}} \\
&\leq \frac{n(1 + \sum_{k=1}^m 2^k/k^2)}{2^{2m}} \\
&= Cn \frac{2^{m+1}}{m^2 2^{2m}} \\
&\leq 2^m C m^{3/2} \frac{2^{m+1}}{m^2 2^{2m}} \\
&= \frac{2}{\sqrt{m}} \rightarrow 0 \quad (n \rightarrow \infty).
\end{aligned}$$

其中 C 为一常数。 □

练习 2.2.9: Weak law for positive variables. Suppose X_1, X_2, \dots are i.i.d., $P(0 \leq X_i < \infty) = 1$ and $P(X_i > x) > 0$ for all x . Let

$$\mu(s) = \int_0^s x \, dF(x)$$

and

$$\nu(s) = \frac{\mu(s)}{s(1 - F(s))}.$$

It is known that there exist constants a_n so that $S_n/a_n \rightarrow 1$ in probability, if and only if $\nu(s) \rightarrow \infty$ as $s \rightarrow \infty$. Pick $b_n \geq 1$ so that $n\mu(b_n) = b_n$ (this works for large n), and use Theorem 2.2.11 to prove that the condition is sufficient.

证明：先解释如何取到的 $n\mu(b_n) = b_n$ 。三角形式的弱大数定律的结论为

$$\frac{S_n - \mathbb{E}\bar{S}_n}{a_n} \rightarrow 0,$$

与题目中所要证明的形式对照得到

$$a_n = \mathbb{E}\bar{S}_n = n\mathbb{E}\bar{X}_n = n\mu(a_n).$$

所以只要取 $n\mu(b_n) = b_n = a_n$ 并使用弱大数定律，即可导出欲求结论。再证明 $n \rightarrow \infty$ 时 $b_n \rightarrow \infty$ 。这是容易的，因为

$$\nu(b_n) = \frac{\mu(b_n)}{b_n(1 - F(b_n))} = \frac{1}{n(1 - F(b_n))} \rightarrow \infty,$$

那么 $F(b_n) \rightarrow 1$ ，结合 $P(X_i > x) > 0$ 知道 $b_n \rightarrow \infty$ 。

再证明弱大数定律的下一条件。

$$\sum_{k=1}^n P(X_k > b_n) = nP(X_1 > b_n) = n(1 - F(b_n)) = \frac{1}{\nu(b_n)} \rightarrow 0.$$

再证明弱大数定律的另一个条件。

$$\frac{\sum_{k=1}^n \mathbb{E}\bar{X}_k^2}{b_n^2} = \frac{n\mathbb{E}\bar{X}_1^2}{b_n^2}.$$

下考虑对 b_n^2/n 进行放缩（重要），有

$$\int_0^{b_n} \mu(x) dx \leq b_n \mu(b_n) = \frac{b_n^2}{n},$$

那么有

$$\begin{aligned} \frac{\sum_{k=1}^n \mathbb{E}\bar{X}_k^2}{b_n^2} &= \frac{n\mathbb{E}\bar{X}_1^2}{b_n^2} \\ &\leq \frac{\int_0^{b_n} 2x(1 - F(x)) dx}{\int_0^{b_n} \mu(x) dx} \\ &= \frac{2b_n(1 - F(b_n))}{\mu(b_n)} \\ &= 2n(1 - F(b_n)) \\ &\rightarrow 0 \quad (n \rightarrow \infty). \end{aligned}$$

上式先后使用了洛必达法则、等式 $n\mu(b_n) = b_n$ 和 $\nu(b_n) \rightarrow \infty$ 。

综上，弱大数定律成立，从而明所欲证。 □

2.3. Borel-Cantelli 引理

练习 2.3.1: Prove that

$$P(\limsup A_n) \geq \limsup P(A_n)$$

and

$$P(\liminf A_n) \leq \liminf P(A_n).$$

证明: 有 $\liminf 1_{A_n} = 1_{\liminf A_n}$, 然后直接用 *Fatou* 引理即可。另一个结论同理。

□

练习 2.3.2: Prove the first result in Theorem 2.3.4 directly from the definition.

证明: 对 $\forall \varepsilon > 0$, 有

$$\begin{aligned} & P(|f(X_n) - f(X)| > \varepsilon) \\ &= P(|f(X_n) - f(X)| > \varepsilon, |X_n| \leq M) + P(|f(X_n) - f(X)| > \varepsilon, |X_n| > M) \end{aligned}$$

对 $\forall M > 0$ 成立。考虑分别使两部分趋于 0。对于第一部分有 f 在 $\{X(\omega) : |X_n(\omega)| \leq M\}$ 上有界, 从而一致连续。那么存在 $\delta > 0$ 使得 $\forall \omega, |X_n(\omega) - X(\omega)| < \delta$ 有 $|f(X_n(\omega)) - f(X(\omega))| < \varepsilon$ 。那么

$$\begin{aligned} P(|f(X_n) - f(X)| > \varepsilon, |X_n| \leq M) &= P(|X_n - X| > \delta, |X_n| \leq M) \\ &= P(|X_n - X| > \delta) \\ &\rightarrow 0 \quad (n \rightarrow \infty). \end{aligned}$$

再考虑第二部分, $\forall \varepsilon' > 0$, 因为 $P(\Omega) = 1$, 可以取 M 足够大使 $P(|X_n| > M) < \varepsilon'$, 从而有

$$P(|f(X_n) - f(X)| > \varepsilon, |X_n| > M) \leq P(|X_n| > M) < \varepsilon'.$$

令 $\varepsilon' \rightarrow 0$ 再令 $n \rightarrow \infty$ 即得到结果。

□

练习 2.3.3: Let ℓ_n be the length of the head run at time n . See Example 8.3.1 for the precise definition. Show that

$$\limsup_{n \rightarrow \infty} \frac{\ell_n}{\log_2 n} = 1, \quad \liminf_{n \rightarrow \infty} \ell_n = 0 \text{ a.s. .}$$

练习 2.3.4: Suppose $X_m \geq 0$ and $X_n \rightarrow X$ in probability. Show that

$$\liminf_{n \rightarrow \infty} EX_n \geq EX.$$

证明: 反设 $\liminf EX_n < EX$, 那么一定存在 X_n 的子列 X_{n_k} 使得 EX_{n_k} 收敛, 且 $\lim EX_{n_k} < EX$. 那么 X_{n_k} 一定存在一个子列 $X_{n_{k_l}} \xrightarrow{\text{a.s.}} X$, 从而由 *Fatou* 引理知道 $\lim EX_{n_{k_l}} \geq EX$, 产生矛盾。 \square

练习 2.3.5: Suppose $X_n \rightarrow X$ in probability, and:

1. $|X_n| \leq Y$ with $EY < \infty$, or
2. There is a continuous function g with $g(x) > 0$ for large x with

$$\frac{|x|}{g(x)} \rightarrow 0 \quad (x \rightarrow \infty)$$

so that $Eg(x) \leq C < \infty$ for all n . Show that $EX_n \rightarrow EX$.

证明:

1. 对 X_n 的任意子列 X_{n_k} 它存在子列 $X_{n_{k_l}} \xrightarrow{\text{a.s.}} X$, 从而由 *Lebesgue* 控制收敛定理知道 $EX_{n_{k_l}} \rightarrow EX$, 那么由定理 2.3.3 直接得到 $EX_n \rightarrow EX$.
2. 考虑题目中的极限条件, 对足够大的 M 有 $g(X) > |X|$ 在 $\{|X| > M\}$ 上恒成立。这时有

$$EX_n = E(X_n; |X_n| \leq M) + E(X_n; |X_n| > M).$$

还是分别讨论两个部分。对第一部分, 直接由 $|X_n| \leq M$ 用有界收敛定理得到

$$E(X_n; |X_n| \leq M) \rightarrow E(X; |X_n| \leq M).$$

对第二部分, 有 $|X_n| \leq g(X_n)$ 且

$$E(g(X_n); |X_n| > M) \leq E(g(X_n)) \leq C < \infty,$$

然后用 (1) 中的控制收敛定理直接得到

$$E(X_n; |X_n| > M) \rightarrow E(X; |X_n| > M).$$

\square

练习 2.3.6: Show:

1.

$$d(X, Y) = E \frac{|X - Y|}{1 + |X - Y|}$$

defines a metric on the set of random variables;

2. $d(X_n, X) \rightarrow 0$ as $n \rightarrow \infty$ if and only if $X_n \xrightarrow{P} X$.

证明:

1. 分别证明度量的三个性质。

1. 正定性。 $X = Y$ a.s. 时显然 $d(X, Y) = 0$ 。当 $d(X, Y) = 0$ 时有

$$\frac{|X-Y|}{1+|X-Y|} = 0 \text{ a.s. },$$

从而有 $X = Y$ a.s. .

2. 对称性显然。

3. 三角不等式。对 \forall r.v. X, Y, Z , 只要证明

$$E\left(\frac{|X-Y|}{1+|X-Y|} + \frac{|Y-Z|}{1+|Y-Z|}\right) \geq E\left(\frac{|X-Z|}{1+|X-Z|}\right),$$

那么只要证明

$$\frac{|X-Y|}{1+|X-Y|} + \frac{|Y-Z|}{1+|Y-Z|} \geq \frac{|X-Z|}{1+|X-Z|}.$$

上式代数等价于

$$|X-Z| \leq |X-Y| + |Y-Z| + 2|X-Y||Y-Z| + |X-Y||Y-Z||X-Z|,$$

而由 $|X-Z| \leq |X-Y| + |Y-Z|$, 上式显然成立。

2. 先证必要性。设 $d(X_n, X) \rightarrow 0$ 。令

$$\varphi(x) = \frac{x}{1+x}$$

为在 $[0, \infty)$ 上的严格单调增函数。 $\forall \varepsilon > 0$, 那么

$$P(|X_n - X| > \varepsilon) = P(\varphi(|X_n - X|) > \varphi(\varepsilon)).$$

只要证 $\varphi(|X_n - X|) \xrightarrow{P} 0$ 即可。这是显然的, 因为 $d(X_n, X) \rightarrow 0$ 直接说明 $E\varphi(|X_n - X|) \rightarrow 0$, 从 L_1 收敛直接推导出依概率收敛。

再证明充分性。由于 $\varphi(|X_n - X|) \leq 1$, 直接用依概率收敛的控制(有界)收敛定理得到 $E\varphi(|X_n - X|) \rightarrow E\varphi(|X - X|) = 0$ 。

□

练习 2.3.7: Show that random variables are a complete space under the metric defined in the previous exercise, i.e., if $d(X_m, X_n) \rightarrow 0$ whenever $m, n \rightarrow \infty$ then there is a r.v. X_∞ so that $X_n \rightarrow X_\infty$ in probability.

证明: 也即证明所有的随机变量构成的集合对依测度收敛封闭。设随机变量列 X_n 满足 $|X_n - X_m| \xrightarrow{P} 0$ ($n, m \rightarrow \infty$) (由上一题知道这和题目条件等价), 那么可以构造子列 X_{n_k} , 使得对 $\forall k \in \mathbb{N}^*$ 有

$$P\left(|X_{n_{k+1}} - X_{n_k}| > \frac{1}{k^2}\right) < \frac{1}{k^2}.$$

由于 $\sum 1/k^2 < \infty$, 所以由第一 Borel-Cantelli 定理知道

$$P\left(\limsup\left\{|X_{n_{k+1}} - X_{n_k}| > \frac{1}{k^2}\right\}\right) = 0.$$

令上式中集合为 A , 那么

$$\begin{aligned} A^c &= \liminf\left\{|X_{n_{k+1}} - X_{n_k}| \leq \frac{1}{k^2}\right\} \\ &\subset \{X_{n_k}(\omega) \text{ 为 } \mathbb{R} \text{ 上基本列}\} \\ &= \{X_{n_k}(\omega) \text{ 极限存在}\}. \end{aligned}$$

又知道 $P(A^c) = 1$, 那么 X_{n_k} 几乎处处极限存在, 在 A^c 上点态定义 r.v. X 为 X_{n_k} 的极限, 则有 $X_{n_k} \xrightarrow{\text{a.s.}} X$. 容易证明把上面的变量列 X_n 换成 X_n 的任意子列 X_{n_k} 结论仍然成立, 所以由定理 2.3.2 知道 $X_n \xrightarrow{P} X$. \square

练习 2.3.8: Let A_n be a sequence of independent events with $P(A_n) < 1$ for all n . Show that $P(\bigcup A_n) = 1$ implies $\sum_n P(A_n) = \infty$ and hence $P(A_n \text{ i.o.}) = 1$.

证明:

$$0 = P\left(\bigcap A_n^c\right) = \prod P(A_n^c) = \prod (1 - P(A_n)),$$

从而有

$$\sum P(A_n) = \infty.$$

\square

练习 2.3.9:

1. If $P(A_n) \rightarrow 0$ and

$$\sum_{n=1}^{\infty} P(A_n^c \cap A_{n+1}) < \infty$$

then $P(A_n \text{ i.o.}) = 0$.

2. Find an example of a sequence A_n to which the result in (1) can be applied but the Borel - Cantelli lemma cannot.

证明:

1.

$$P\left(\bigcup_k A_k\right) \leq P(A_n) + \sum_{k=n}^{\infty} P(A_n^c \cap A_{n+1}) \rightarrow 0.$$

2. 令概率空间为 $([0, 1], \mathcal{R}_{[0,1]}, \lambda)$, $A_n = [0, 1/n]$. 那么

$$\sum P(A_n) = \infty$$

从而不能应用 Borel-Cantelli 引理, 但是

$$\sum_{n=1}^{\infty} P(A_n^c \cap A_{n+1}) = 0.$$

□

练习 2.3.10: Kochen - Stone lemma. Suppose $\sum P(A_k) = \infty$. Use Exercises 1.6.6 and 2.3.1 to show that if

$$\limsup_{n \rightarrow \infty} \frac{\left(\sum_{k=1}^n P(A_k) \right)^2}{\sum_{1 \leq j, k \leq n} P(A_j \cap A_k)} = \alpha > 0$$

then $P(A_n \text{ i.o.}) \geq \alpha$. The case $\alpha = 1$ contains Theorem 2.3.7.

证明: 定义随机变量

$$X_n = \sum_{k=1}^n 1_{A_k}.$$

那么有

$$\begin{aligned} EX_n &= \sum_{k=1}^n E1_{A_k} = \sum_{k=1}^n P(A_k), \\ EX_n^2 &= \sum_{1 \leq j, k \leq n} P(A_j \cap A_k) < \infty. \end{aligned}$$

由练习 1.6.6 知道

$$P(X_n > 0) \geq \frac{\left(\sum_{k=1}^n P(A_k) \right)^2}{\sum_{1 \leq j, k \leq n} P(A_j \cap A_k)},$$

那么结合练习 2.3.1 有

$$\begin{aligned} P(\limsup A_n) &= P\left(\limsup \bigcup_{k=1}^n A_k\right) \\ &= P(\limsup \{X_n > 0\}) \\ &\geq \limsup P(X_n > 0) = \alpha, \end{aligned}$$

□

练习 2.3.11: Let X_1, X_2, \dots be independent with $P(X_n = 1) = p_n$ and $P(X_n = 0) = 1 - p_n$. Show that:

1. $X_n \rightarrow 0$ in probability if and only if $p_n \rightarrow 0$, and
2. $X_n \rightarrow 0$ a.s. if and only if $\sum p_n < \infty$.

证明:

1. 先证必要性, 设 $X_n \xrightarrow{P} 0$, 由依测度收敛的有界收敛定理知道 $p_n = EX_n \rightarrow 0$ 。再证充分性。设 $p_n \rightarrow 0$, 那么 $EX_n = p_n \rightarrow 0$, 从而 $X_n \xrightarrow{L_1} 0$, 那么 $X_n \xrightarrow{P} 0$ 。
2. 先证必要性。对足够大的 n , 有 $|X_n| < 1/2$ a.e., 那么有 $X_n = 0$ a.e., 即 $p_n = 1$ 。那么只有有限个 p_n 不为 0, 从而 $\sum p_n < \infty$ 。再证充分性。由

$$\sum p_n = \sum P(\{X_n = 1\}) < \infty,$$

且 X_n 相互独立, 直接应用 *Borel-Cantelli* 引理得到 $P(\limsup\{X_n = 1\}) = 0$, 即 $P(\liminf\{X_n = 0\}) = 1$ 。对 $\forall \omega \in \liminf\{X_n = 0\}$ 知道至多有限个 X_n 中随机变量满足 $X_n(\omega) = 1$, 从而一定有 $X_n(\omega) \rightarrow 0$ 。那么 $X_n \xrightarrow{\text{a.s.}} 0$ 。

□

练习 2.3.12: Let X_1, X_2, \dots be a sequence of r.v.'s on (Ω, \mathcal{F}, P) where Ω is a countable set and \mathcal{F} consists of all subsets of Ω . Show that $X_n \rightarrow X$ in probability implies $X_n \rightarrow X$ a.s.

证明: 记 $\Omega = \{\omega_n\}$, 反设 $P(X_n \text{ 不收敛于 } X) > 1$, 记 $\{X_n \text{ 不收敛于 } X\} = \{\omega_{n_k}\}$, 那么其中一定存在一个 $\omega' = \omega_{n_k}$ 使得 $P(\omega') > 0$, 且这时有 $X_n(\omega')$ 不收敛于 $X(\omega')$ 。这个不收敛性等价于

$$\limsup |X_n(\omega') - X(\omega')| = a > 0.$$

直接取 X_n 的子列 X_{n_p} 使得

$$\lim |X_{n_p}(\omega') - X(\omega')| = a.$$

那么 p 足够大时有

$$\omega' \in \left\{ |X_{n_p} - X| > \frac{a}{2} \right\},$$

从而

$$P\left(|X_{n_p} - X| > \frac{a}{2}\right) \geq P(\omega') > 0,$$

又由于 $X_n \xrightarrow{P} X$, 一定有 $X_{n_p} \xrightarrow{P} X$, 那么对足够大的 p 有

$$P\left(|X_{n_p} - X| > \frac{a}{2}\right) < P(\omega'),$$

从而产生矛盾。 □

练习 2.3.13: If X_n is any sequence of random variables, there are constants $c_n \rightarrow \infty$ so that $X_n/c_n \rightarrow 0$ a.s.

证明: 对 $\forall n \in \mathbb{N}^*$, 可以取 $c_n > 0$ 使

$$P\left(\frac{|X_n|}{c_n} > \frac{1}{n}\right) < \frac{1}{2^n}.$$

这时有

$$\sum_{n=1}^{\infty} P\left(\frac{|X_n|}{c_n} > \frac{1}{n}\right) < \infty,$$

从而由 *Borel-Cantelli* 引理知道

$$P\left(\limsup\left\{\frac{|X_n|}{c_n} > \frac{1}{n}\right\}\right) = 0,$$

取补集即有

$$P\left(\frac{X_n}{c_n} \rightarrow 0\right) \geq P\left(\liminf\left\{\frac{|X_n|}{c_n} \leq \frac{1}{n}\right\}\right) = 1.$$

□

练习 2.3.14: Let X_1, X_2, \dots be independent. Show that $\sup_n X_n < \infty$ a.s. if and only if

$$\sum_n P(X_n > A) < \infty$$

for some A .

证明: 先证必要性。设 $\sup_n X_n < \infty$, 反设 $\sum P(X_n > A) = \infty$, 那么对 $\forall A > 0$, 由 *Borel-Cantelli* 定理知道 $P(\limsup\{X_n > A\}) = 1$, 从而 $\sup_n X_n > A$ a.s. 对 $\forall A > 0$ 成立, 矛盾。

再证充分性。直接使用 *Borel-Cantelli* 定理知道 $P(\limsup\{X_n > A\}) = 0$, 即 $P(\liminf\{X_n \leq A\}) = 1$, 对这个集合中的任意元素 ω , 只有有限个随机变量 X_{n_1}, \dots, X_{n_k} 在 ω 处的取值大于 A , 那么有

$$\sup_n X_n(\omega) = \max\{X_{n_1}(\omega), \dots, X_{n_k}(\omega)\} < \infty.$$

□

练习 2.3.15: Let X_1, X_2, \dots be i.i.d. with $P(X_i > x) = e^{-x}$, let

$$M_n = \max_{1 \leq m \leq n} X_m.$$

Show that

$$1. \quad \limsup_{n \rightarrow \infty} \frac{X_n}{\log n} = 1 \text{ a.s.}$$

and

$$2. \quad \frac{M_n}{\log n} \rightarrow 1 \text{ a.s. .}$$

证明:

1. 以下证明需要用到

$$\begin{aligned} \{\limsup X_n \leq K\} &= \liminf \{X_n \leq K\}, \\ \{\limsup X_n \geq K\} &= \limsup \{X_n \geq K\}, \\ \{\liminf X_n \leq K\} &= \liminf \{X_n \leq K\}, \\ \{\liminf X_n \geq K\} &= \limsup \{X_n \geq K\}. \end{aligned}$$

对 $\forall \varepsilon > 0$, 只要证明

$$P\left(\left|\frac{\limsup X_n}{\log n} - 1\right| \geq \varepsilon \text{ i.o.}\right) = 0.$$

即

$$P(\limsup X_n \geq (1 + \varepsilon) \log n \text{ i.o.}) = 0 \quad (1)$$

且

$$P(\limsup X_n \leq (1 - \varepsilon) \log n \text{ i.o.}) = 0. \quad (2)$$

对 (1) 有

$$\begin{aligned} P(\limsup X_n \geq (1 + \varepsilon) \log n \text{ i.o.}) &= P(\limsup \{\limsup X_n \geq (1 + \varepsilon) \log n\}) \\ &= P(\limsup \limsup \{X_n \geq (1 + \varepsilon) \log n\}) \\ &= P(\limsup \{X_n \geq (1 + \varepsilon) \log n\}), \end{aligned}$$

由 *Borel-Cantelli* 引理, 只要证明

$$\sum P(X_n \geq (1 + \varepsilon) \log n) < \infty.$$

这是容易的, 因为

$$\sum P(X_n \geq (1 + \varepsilon) \log n) = \sum \frac{1}{n^{1+\varepsilon}} < \infty.$$

对 (2) 有

$$\begin{aligned}
P(\limsup X_n \leq (1-\varepsilon) \log n \text{ i.o.}) &= P(\limsup \{\limsup X_n \leq (1-\varepsilon) \log n\}) \\
&= P(\limsup \liminf \{X_n \leq (1-\varepsilon) \log n\}) \\
&= P(\liminf \{X_n \leq (1-\varepsilon) \log n\}) \\
&= 1 - P(\limsup \{X_n > (1-\varepsilon) \log n\}),
\end{aligned}$$

由 *Borel-Cantelli* 引理, 只要证明

$$\sum P(\{X_n > (1-\varepsilon) \log n\}) = \infty.$$

这也是容易的, 因为

$$\sum P(\{X_n > (1-\varepsilon) \log n\}) = \sum \frac{1}{n^{1-\varepsilon}} = \infty.$$

2. $\forall \varepsilon > 0$. 由 (I) 知道存在 $N \in \mathbb{N}^*$, 对 $\forall n > N$ 有 $X_n \leq (1+\varepsilon) \log n$ a.e.。令 $X_m = \max\{X_1, \dots, X_N\}$, 那么令 $N' = \lceil \exp(X_m) \rceil$, 那么对 $\forall n > N'$ 有

$$\begin{aligned}
M_n &= \max\{X_m, X_{N+1}, \dots, X_n\} \\
&\leq \max\{\log N', (1+\varepsilon) \log n\} \\
&\leq (1+\varepsilon) \log n.
\end{aligned}$$

由 ε 的任意性知道 $\lim M_n / \log n \leq 1$ a.s.。再证明

$$P(M_n \leq (1-\varepsilon) \log n \text{ i.o.}) = 0.$$

这是因为

$$\begin{aligned}
\sum P(M_n \leq (1-\varepsilon) \log n) &= \sum P^n(X_1 \leq (1-\varepsilon) \log n) \\
&= \sum \left(1 - \frac{1}{n^{1-\varepsilon}}\right)^n \\
&\leq \sum \exp(n^{-\varepsilon}) \\
&< \infty,
\end{aligned}$$

从而由 *Borel-Cantelli* 引理得到结论。

□

练习 2.3.16: Let X_1, X_2, \dots be i.i.d. with distribution F , let $\lambda_n \uparrow \infty$, and let

$$A_n = \left\{ \max_{1 \leq m \leq n} X_m > \lambda_n \right\}.$$

Show that $P(A_n \text{ i.o.}) = 0$ or 1 according as $\sum_{n \geq 1} (1 - F(\lambda_n)) < \infty$ or $= \infty$.

证明: 直接应用 *Borel-Cantelli* 引理知道

$$P(X_n > \lambda_n \text{ i.o.}) < \infty \text{ 或 } = \infty$$

取决于

$$\sum P(X_n > \lambda_n) = \sum (1 - F(\lambda_n)) < \infty \text{ 或 } = \infty.$$

那么只要证明 $\{X_n > \lambda_n \text{ i.o.}\} = A_n$ 。 $\forall \omega \in \{X_n > \lambda_n \text{ i.o.}\}$, 那么有

$$\max_{1 \leq m \leq n} X_m \geq X_n > \lambda_n$$

对 $\forall n \in \mathbb{N}^*$ 成立, 从而 $\omega \in A_n$ 。对 $\forall \omega \in A_n$, 反设只有有限个 n 满足 $X_n(\omega) > \lambda_n$, 设可以取到的最大的 n 为 p , 那么有 $X_p(\omega) > \lambda_p$ 。令 q 足够大使得

$$\lambda_q > \max\{X_1(\omega), \dots, X_p(\omega)\},$$

且 $M_q(\omega) > \lambda_q(\omega)$ (由 $\omega \in A_n$ 知可以这样取), 那么存在 $r \in \{p+1, \dots, q\}$ 使得

$$X_r = \max\{X_{p+1}, \dots, X_q\} > \lambda_q > \lambda_r,$$

从而 $X_r(\omega) > \lambda_r$ 且 $r > p$, 矛盾。 □

练习 2.3.17: Let Y_1, Y_2, \dots be i.i.d. Find necessary and sufficient conditions for:

1. $Y_n/n \rightarrow 0$ almost surely;

2.
$$\frac{\max_{m \leq n} Y_m}{n} \rightarrow 0$$

almost surely;

3.
$$\frac{\max_{m \leq n} Y_m}{n} \rightarrow 0$$

in probability;

4.
$$\frac{Y_n}{n} \rightarrow 0$$

in probability.

证明: 四个结论分别为: $E|Y_i| < \infty$, $EY_i^+ < \infty$, $nP(Y_i > n) \rightarrow 0$, $P(Y_i < \infty) = 1$ 。以下分别证明。

1. 无妨设 $Y_n \geq 0$ 。先证必要性。由 $Y_n/n \xrightarrow{\text{a.s.}} 0$ 知道

$$P(Y_n > n \text{ i.o.}) = 0.$$

那么又知道 Y_1, Y_2, \dots 互相独立, 使用 *Borel-Cantelli* 引理可以反过来得到

$$\sum_{n=1}^{\infty} P(Y_n > n) < \infty.$$

那么

$$\begin{aligned}
EY_n &= \int_0^\infty P(Y_n > x) dx \\
&\leq P(Y_n > 0) + \sum_{n=1}^\infty P(Y_n > n) \\
&< \infty.
\end{aligned}$$

再证充分性。\$\forall \varepsilon > 0\$, 有

$$\begin{aligned}
\sum_{n=1}^\infty P(Y_n > \varepsilon n) &\leq \int_0^\infty P(Y_n > \varepsilon x) dx \\
&= E \frac{Y_n}{\varepsilon} < \infty.
\end{aligned}$$

那么 \$P(Y_n > \varepsilon n \text{ i.o.}) = 0\$, 从而 \$Y_n/n \xrightarrow{\text{a.s.}} 0\$。

2. 先证必要性。由一致收敛性知道对 \$\forall \varepsilon > 0\$, 有

$$P(Y_m^+ > \varepsilon n \text{ i.o.}) = P\left(\max_{m=1}^n Y_m^+ > \varepsilon n \text{ i.o.}\right) \leq P\left(\left|\max_{m=1}^n Y_m\right| > \varepsilon n \text{ i.o.}\right) = 0.$$

再由上一问同理得到

$$EY_n^+ < \infty.$$

再证充分性。由 \$EY_n^+ < \infty\$ 知道

$$\sum_{n=1}^\infty P(Y_n^+ > \varepsilon n) < \infty,$$

从而 \$P(Y_n^+ > \varepsilon n \text{ i.o.}) = 0\$, 那么

$$\limsup \frac{Y_n}{n} \leq 0 \text{ a.s. ,}$$

从而得到结论。

□

练习 2.3.18: Let \$0 \leq X_1 \leq X_2 \dots\$ be random variables with \$EX_n \sim an^\alpha\$ with \$a, \alpha > 0\$, and \$\text{var}(X_n) \leq Bn^\beta\$ with \$\beta < 2\alpha\$. Show that

$$\frac{X_n}{n^\alpha} \rightarrow a \text{ a.s. .}$$

练习 2.3.19: Let \$X_n\$ be independent Poisson r.v.'s with \$EX_n = \lambda_n\$, and let \$S_n = X_1 + \dots + X_n\$. Show that if \$\sum \lambda_n = \infty\$ then

$$\frac{S_n}{ES_n} \rightarrow 1 \text{ a.s. .}$$

练习 2.3.20: Show that if X_n is the outcome of the n th play of the St. Petersburg game (Example 2.2.16) then

$$\limsup_{n \rightarrow \infty} \frac{X_n}{n \log_2 n} = \infty \text{ a.s.}$$

and hence the same result holds for S_n . This shows that the convergence

$$\frac{S_n}{n \log_2 n} \rightarrow 1$$

in probability proved in Section 2.2 does not occur a.s. .

2.4. 强大数定律

练习 2.4.1: Lazy janitor. Suppose the i th light - bulb burns for an amount of time X_i and then remains burned out for time Y_i before being replaced. Suppose the X_i, Y_i are positive and independent with the X 's having distribution F and the Y 's having distribution G , both of which have finite mean. Let R_t be the amount of time in $[0, t]$ that we have a working light - bulb. Show that

$$\frac{R_t}{t} \rightarrow \frac{EX_i}{EX_i + EY_i}$$

almost surely.

证明: 设 $EX_i = \mu < \infty, EY_i = \nu < \infty$, 再设 $T_n = \sum_{k=1}^n (X_k + Y_k)$. 那么对 $\forall t \in \mathbb{R}^+$, 一定存在 $T_n \leq t < T_{n+1}$. 考虑分类讨论:

1. 当 $0 \leq t - T_n < X_{n+1}$ 时, 这时 t 时刻灯是亮的, 有

$$\frac{R_t}{t} = \frac{R_{T_n} + t - T_n}{T_n + t - T_n} \in \left[\frac{R_{T_n}}{T_n}, \frac{R_{T_n} + X_{n+1}}{T_n + X_{n+1}} \right];$$

2. 当 $X_{n+1} \leq t - T_n < X_{n+1} + Y_{n+1}$ 时, 这时 t 时刻灯是灭的, 有

$$\frac{R_t}{t} = \frac{R_{T_n} + X_{n+1}}{T_n + t - T_n} \in \left[\frac{R_{T_{n+1}}}{T_{n+1}}, \frac{R_{T_n} + X_{n+1}}{T_n + X_{n+1}} \right].$$

设 $H_n = \sum_{k=1}^n X_k$ 为 $[0, T_n]$ 时间内亮灯的时间, 那么由强大数定律知道

$$\frac{T_n}{n} \xrightarrow{\text{a.s.}} \mu + \nu, \quad \frac{H_n}{n} \xrightarrow{\text{a.s.}} \mu,$$

那么

$$\frac{R_{T_n}}{T_n} = \frac{H_n}{T_n} \xrightarrow{\text{a.s.}} \frac{\mu}{\mu + \nu}, \quad \frac{R_{T_n} + X_{n+1}}{T_n + X_{n+1}} = \frac{H_n + X_{n+1}}{T_n + X_{n+1}} \xrightarrow{\text{a.s.}} \frac{\mu}{\mu + \nu},$$

利用夹逼定理得到结论。

□

练习 2.4.2: Let $X_0 = (1, 0)$ and define $X_n \in \mathbb{R}^2$ inductively by declaring that X_{n+1} is chosen at random from the ball of radius $|X_n|$ centered at the origin, i.e., $X_{n+1}/|X_n|$ is uniformly distributed on the ball of radius 1 and independent of X_1, \dots, X_n . Prove that $n^{-1} \log|X_n| \rightarrow c$ a.s. and compute c .

证明: 设 $Y_n = X_{n+1}/(|X_n|)$, 那么 Y_n 服从在单位圆上的均匀分布, 且 Y_1, Y_2, \dots 为独立同分布随机变量列, 且有

$$E \log|Y_n| = \frac{1}{\pi} \int_0^1 2\pi r \log r \, dr = -\frac{1}{2} > -\infty,$$

从而由强大数定律知道

$$\frac{\log|X_n|}{n} = \frac{\sum_{k=1}^n \log|Y_k|}{n} \xrightarrow{\text{a.e.}} E \log|Y_n| = -\frac{1}{2}.$$

□

练习 2.4.3: Investment problem. We assume that at the beginning of each year you can buy bonds for \$1 that are worth \$ a at the end of the year or stocks that are worth a random amount $V \geq 0$. If you always invest a fixed proportion p of your wealth in bonds, then your wealth at the end of year $n+1$ is $W_{n+1} = (ap + (1-p)V_n)W_n$. Suppose V_1, V_2, \dots are i.i.d. with $EV_n^2 < \infty$ and $E(V_n^{-2}) < \infty$.

1. Show that $n^{-1} \log W_n \rightarrow c(p)$ a.s. .
2. Show that $c(p)$ is concave. (Use Theorem A.5.1 in the Appendix to justify differentiating under the expected value.)
3. By investigating $c'(0)$ and $c'(1)$, give conditions on V that guarantee that the optimal choice of p is in $(0, 1)$.
4. Suppose $P(V=1) = P(V=4) = 1/2$. Find the optimal p as a function of a .

证明:

1. 令 $U_n = ap + (1-p)V_n$ 。因为 $EV_n^2 < \infty$ 所以 $EV_n < \infty$, 同理 $EV_n^{-1} < \infty$ 。

$$E|\log U_n| = E(\log U_n; U_n \geq 1) - E(\log U_n; 0 \leq U_n < 1),$$

其中

$$\begin{aligned}
\mathbb{E}(\log U_n; U_n \geq 1) &= \mathbb{E} \log(ap + (1-p)V_n; U_n \geq 1) \\
&\leq \mathbb{E}(ap + (1-p)V_n; U_n \geq 1) \\
&\leq \mathbb{E}(ap + (1-p)V_n) \\
&= ap + (1-p)\mathbb{E}V_n \\
&< \infty, \\
-\mathbb{E}(\log U_n; 0 \leq U_n < 1) &= \mathbb{E} \left(\log \left(\frac{1}{ap + (1-p)V_n} \right); 0 \leq U_n < 1 \right) \\
&\leq \mathbb{E} \left(\log \left(\frac{1}{ap} \right); 0 \leq U_n < 1 \right) \\
&\leq -\log ap \\
&< \infty,
\end{aligned}$$

从而可以应用强大数定律,

$$\frac{\log W_n}{n} = \frac{\sum_{k=1}^n \log U_k}{n} \xrightarrow{\text{a.e.}} \mathbb{E} \log U_n = \mathbb{E} \log(ap + (1-p)V_n) := c(p).$$

2. 对 $c(p)$ 应用定理 A.5.1。分别验证四个条件。

1. $\mathbb{E}|\log U_n| < \infty$ 上一问已证;

2.
$$\frac{\partial}{\partial p} \log(ap + (1-p)V_n) = \frac{a - V_n}{ap + (1-p)V_n}$$

对 p 连续;

3. 令

$$v(p) = \int_{\Omega} \frac{a - V_n}{ap + (1-p)V_n} dP,$$

对 $\forall \delta p$ 足够小, 有

$$\begin{aligned}
&v(p + \delta p) - v(p) \\
&= \int_{\Omega} (a - V_n) \left(\frac{1}{a(p + \delta p) + (1-p - \delta p)V_n} - \frac{1}{ap + (1-p)V_n} \right) dP \\
&= \int_{\Omega} (a - V_n) \left(\frac{\delta p(V_n - a)}{(a(p + \delta p) + (1-p - \delta p)V_n)(ap + (1-p)V_n)} \right) dP \\
&= \delta p \int_{\Omega} \left(\frac{-(V_n - a)^2}{(a(p + \delta p) + (1-p - \delta p)V_n)(ap + (1-p)V_n)} \right) dP \\
&\leq \frac{\delta p}{\min\{a^2, V_n^2\}} \int_{\Omega} -(V_n - a)^2 dP \\
&\rightarrow 0 \quad (\delta p \rightarrow 0),
\end{aligned}$$

从而 $v(p)$ 连续, 其中最后一步用到了 $\mathbb{E}V_n^2 < \infty$ 。

$$\begin{aligned}
4. \quad & \int_{\Omega} \int_{-\delta}^{\delta} \left| \frac{a - V_n}{a(p + \theta) + (1 - p - \theta)V_n} \right| d\theta dP \\
& \leq \int_{\Omega} \int_{-\delta}^{\delta} \left| \frac{a - V_n}{\min\{a, V_n\}} \right| d\theta dP \\
& \leq E \left| \frac{2\delta(a - V_n)}{\min\{a, V_n\}} \right| \\
& < \infty.
\end{aligned}$$

其中最后一步用到了 $EV_n, EV_n^{-1} < \infty$ 。

那么 $c'(p) = v(p)$ 。可以类似对 $v(p)$ 应用定理得到

$$c''(p) = v'(p) = -E \left(\frac{(V_n - a)^2}{(ap + (1 - p)V_n)^2} \right) \leq 0,$$

从而 $c(p)$ 是凹函数。

3. 容易看出需要 $c'(0) \geq 0, c'(1) \leq 0$ 。这就是

$$\begin{aligned}
c'(0) &= E \left(\frac{a - V_n}{V_n} \right) \geq 0 \quad \text{即} \quad E \left(\frac{1}{V_n} \right) \geq \frac{1}{a} \\
&\quad \text{和} \\
c'(1) &= E \left(\frac{a - V_n}{a} \right) \leq 0 \quad \text{即} \quad E(V_n) \geq a.
\end{aligned}$$

4. 只要令

$$\begin{aligned}
c'(p) &= E \left(\frac{a - V_n}{ap + (1 - p)V_n} \right) \\
&= \frac{1}{2} \frac{a - 1}{ap + (1 - p)} + \frac{1}{2} \frac{a - 4}{ap + 4 - 4p} \\
&= 0
\end{aligned}$$

即可。解方程得到

$$p = \frac{8 - 5a}{2(a - 1)(a - 4)}.$$

□

2.5. 随机级数的收敛性

练习 2.5.1: Suppose X_1, X_2, \dots are i.i.d. with $EX_i = 0$, $\text{var}(X_i) = C < \infty$. Use Theorem 2.5.5 with $n = m^\alpha$ where $\alpha(2p - 1) > 1$ to conclude that if $S_n = X_1 + \dots + X_n$ and $p > 1/2$ then $S_n/n^p \rightarrow 0$ almost surely.

证明: 只要证明对 $1/2 < q < p$ 的 q 满足 $\limsup |S_n|/n^q \leq 1$ a.s. 即可, 由 *Borel-Cantelli* 引理, 只要证明

$$\sum_{m=1}^{\infty} P\left(\max_{k=1}^n |S_k| \geq n^p\right) = \sum_{m=1}^{\infty} P\left(\max_{1 \leq k \leq m^\alpha} |S_k| \geq m^{\alpha p}\right) < \infty.$$

而

$$\begin{aligned} P\left(\max_{1 \leq k \leq m^\alpha} |S_k| \geq m^{\alpha p}\right) &\leq \frac{\text{var } S_{m^\alpha}}{m^{2\alpha p}} \\ &\leq \frac{C}{m^{\alpha(2p-1)}} \end{aligned}$$

从而级数收敛。 □

练习 2.5.2: The converse of Theorem 2.5.12 is much easier. Let $p > 0$. If $S_n/n^{1/p} \rightarrow 0$ a.s. then $E|X_1|^p < \infty$.

证明: 反设 $E|X_1|^p = \infty$, 那么

$$\infty = \int_0^\infty P(|X_n|^p > x) dx \leq \sum_{n=1}^\infty P(|X_n|^p > n) = \infty.$$

再由 BC 引理知道 $P(|X_n| > n^{1/p} \text{ i.o.}) = 1$. 这与 $S_n/n^{1/p} \xrightarrow{\text{a.s.}} 0$ 矛盾。 □

练习 2.5.3: Let X_1, X_2, \dots be i.i.d. standard normals. Show that for any t

$$\sum_{n=1}^\infty X_n \cdot \frac{\sin(nt)}{n} \text{ converges a.s.}$$

We will see this series again at the end of Section 8.1.

证明: 由于

$$\sum_n \text{var}\left(X_n \frac{\sin(nt)}{n}\right) = \sum_n \frac{\sigma^2 \sin^2(nt)}{n^2} < \infty$$

直接由定理 2.5.6 得证。 □

练习 2.5.4: Let X_1, X_2, \dots be independent with $EX_n = 0$, $\text{var}(X_n) = \sigma_n^2$.

1. Show that if

$$\sum_n \frac{\sigma_n^2}{n^2} < \infty$$

then

$$\sum_n \frac{X_n}{n}$$

converges a.s. and hence

$$n^{-1} \sum_{m=1}^n X_m \rightarrow 0$$

a.s. ;

2. Suppose

$$\sum_n \frac{\sigma_n^2}{n^2} = \infty$$

and without loss of generality that $\sigma_n^2 \leq n^2$ for all n . Show that there are independent random variables X_n with $EX_n = 0$ and

$$\text{var}(X_n) \leq \sigma_n^2$$

so that X_n/n and hence

$$\frac{\sum_{m \leq n} X_m}{n}$$

does not converge to 0 a.s. .

证明:

1. 直接由定理 2.5.6 得证。
2. 直接令 X_n 服从 $\mathcal{N}(0, n^2)$ 的正态分布。

□

练习 2.5.5: Let $X_n \geq 0$ be independent for $n \geq 1$. The following are equivalent:

1.
$$\sum_{n=1}^{\infty} X_n < \infty \text{ a.s. ;}$$
2.
$$\sum_{n=1}^{\infty} \left[P(X_n > 1) + E(X_n 1_{\{X_n \leq 1\}}) \right] < \infty;$$
3.
$$\sum_{n=1}^{\infty} E\left(\frac{X_n}{1 + X_n}\right) < \infty.$$

证明:

- (1) \implies (2):

$$P(X_n > 1) + E(X_n 1_{\{X_n \leq 1\}}) \leq EX_n 1_{\{X_n > 1\}} + EX_n 1_{\{X_n \leq 1\}} = EX_n.$$

由 Kolmogorov 三级数定理直接知道 $\sum EX_n < \infty$ 。

- (2) \implies (3):

$$\begin{aligned} E \frac{X_n}{1+X_n} &= E \left(\frac{X_n}{1+X_n}; X_n > 1 \right) + E \left(\frac{X_n}{1+X_n}; X_n \leq 1 \right) \\ &\leq P(X_n > 1) + EX_n 1_{\{X_n \leq 1\}}. \end{aligned}$$

• (3) \Rightarrow (1): 反设 $A = \{\sum X_n < \infty\}$ 满足 $P(A) > 0$ 。只要证明

$$\begin{aligned} \sum_{n=1}^{\infty} E \left(\frac{X_n}{1+X_n} \right) &= \sum_{n=1}^{\infty} E \left(\frac{X_n}{1+X_n}; A \right) + E \left(\frac{X_n}{1+X_n}; A^c \right) \\ &\geq \sum_{n=1}^{\infty} E \left(\frac{X_n}{1+X_n}; A \right) \\ &\geq E \left(\sum_{n=1}^{\infty} \frac{X_n}{1+X_n}; A \right) \\ &= \infty \quad (\star) \end{aligned}$$

即可。那么只要证明在 A 上 $\sum X_n/(1+X_n) = \infty$ 。 $\forall \omega \in A$, 若 $X_n(\omega) \not\rightarrow 0$ 那么显然成立, 若 $X_n(\omega) \rightarrow 0$ 那么有 n 足够大时

$$X_n(1+X_n)^{-1} \sim X_n(1-X_n) \sim X_n,$$

从而 $\sum X_n/(1+X_n)$ 的收敛性与 $\sum X_n$ 相同。

□

练习 2.5.6: Let $\psi(x) = x^2$ when $|x| \leq 1$ and $= |x|$ when $|x| \geq 1$. Show that if X_1, X_2, \dots are independent with $EX_n = 0$ and

$$\sum_{n=1}^{\infty} E\psi(X_n) < \infty$$

then

$$\sum_{n=1}^{\infty} X_n$$

converges a.s. .

证明: 令 $Y_n = X_n 1_{\{|X_n| \leq 1\}}$, 然后用 *Kolmogorov* 三级数定理。下分别证明三个条件:

$$\begin{aligned} 1. \quad \sum P(|X_n| > 1) &\leq \sum EX_n 1_{\{|X_n| > 1\}} \\ &\leq \sum (EX_n^2 1_{\{|X_n| \leq 1\}} + EX_n 1_{\{|X_n| > 1\}}) \\ &= \sum E\psi(X_n) < \infty; \end{aligned}$$

$$2. \quad \sum \text{var } Y_n \leq \sum EY_n^2 = \sum EX_n^2 1_{\{|X_n| \leq 1\}} < \infty;$$

$$\begin{aligned}
3. \quad \sum EY_n &= \sum (EX_n - EX_n 1_{\{|X_n| > 1\}}) \\
&= - \sum EX_n 1_{\{|X_n| > 1\}} \\
&> -\infty.
\end{aligned}$$

□

练习 2.5.7: Let X_n be independent. Suppose

$$\sum_{n=1}^{\infty} E|X_n|^{p(n)} < \infty$$

where $0 < p(n) \leq 2$ for all n and $EX_n = 0$ when $p(n) > 1$. Show that

$$\sum_{n=1}^{\infty} X_n$$

converges a.s. .

证明: 令 $Y_n = X_n 1_{\{|X_n| \leq 1\}}$, 然后使用 *Kolmogorov* 三级数定理证明, 下分别说明三个条件.

$$1. \quad \sum P(|X_n| > 1) \leq \sum E|X_n|^{p(n)} < \infty;$$

2. $0 < p(n) \leq 1$ 时,

$$E|Y_n| \leq E|Y_n|^{p(n)} \leq E|X_n|^{p(n)}.$$

$1 < p(n) \leq 2$ 时,

$$E|Y_n| = E(|X_n|; |X_n| > 1) \leq E(|X_n|^{p(n)}).$$

那么

$$\begin{aligned}
&\sum E|Y_n| \leq E|X_n|^{p(n)} < \infty; \\
3. \quad \sum E \text{ var}(Y_n) &\leq \sum EY_n^2 \leq \sum EY_n^{p(n)} \leq \sum E|X_n|^{p(n)} < \infty.
\end{aligned}$$

□

练习 2.5.8: Let X_1, X_2, \dots be i.i.d. and not $\equiv 0$. Then the radius of convergence of the power series

$$\sum_{n \geq 1} X_n(\omega) z^n$$

(i.e.

$$r(\omega) = \sup \left\{ c : \sum |X_n(\omega)| c^n < \infty \right\}$$

) is 1 a.s. or 0 a.s., according as $E \log^+ |X_1| < \infty$ or $= \infty$ where $\log^+ x = \max(\log x, 0)$.

证明: 当 $E \log^+ |X_1| = \infty$ 时, 这等价于

$$\sum_n P(\log^+ |X_1| > Kn) = \sum_n P(|X_n| > e^{Kn}) = \infty,$$

对 $\forall K \in \mathbb{N}$ 成立, 那么 $|X_n| > e^{Kn}$ i.o., 从而 $\limsup \sqrt[n]{|X_n|} \geq e^K$. 那么级数的收敛半径为 0.

当 $E \log^+ |X_1| < \infty$ 时, 这等价于

$$\sum_n P(\log^+ |X_1| < \varepsilon n) = \sum_n P(|X_n| < e^{\varepsilon n}) < \infty$$

对 $\forall \varepsilon > 0$ 成立, 那么 $|X_n| < e^{\varepsilon n}$ i.o., 那么 $\liminf \sqrt[n]{|X_n|} \leq e^\varepsilon$. 那么级数的收敛半径 $r \geq e^{-\varepsilon}$ a.s., 从而 $r \geq 1$ a.s.. 又注意到代入 $|z| = 1$ 时级数发散, 那么级数的收敛半径 $r = 1$ a.s. \square

练习 2.5.9: Let X_1, X_2, \dots be independent and let $S_{m,n} = X_{m+1} + \dots + X_n$. Then

$$(\star) \quad P\left(\max_{m < j \leq n} |S_{m,j}| > 2a\right) \min_{m < k \leq n} P(|S_{k,n}| \leq a) \leq P(|S_{m,n}| > a).$$

证明: 采用与 Kolmogorov 最大值不等式类似的方法证明。令事件

$$A_k = \{|S_{m,m+1}|, \dots, |S_{m,k+1}| \leq 2a, |S_{m,k}| > 2a\}, \quad m < k \leq n,$$

$$B_k = \{|S_{k,n}| \leq a\}, \quad m < k \leq n,$$

那么事件 A_k, B_k 独立, A_k 之间无交, 从而有

$$\begin{aligned} P\left(\max_{m < j \leq n} |S_{m,j}| > 2a\right) \min_{m < k \leq n} P(|S_{k,n}| \leq a) &= \sum_{m < k \leq n} P(A_k) \cdot \min_{m < k \leq n} P(B_k) \\ &\leq \sum_{m < k \leq n} P(A_k) P(B_k) \\ &\leq \sum_{m < k \leq n} P(A_k \cap B_k) \\ &= P\left(\bigcup_{m < k \leq n} A_k \cap B_k\right) \\ &\leq \text{RHS}. \end{aligned}$$

\square

练习 2.5.10: Use (\star) to prove a theorem of P. Lévy: Let X_1, X_2, \dots be independent and let $S_n = X_1 + \dots + X_n$. If $\lim_{n \rightarrow \infty} S_n$ exists in probability then it also exists a.s. .

证明: 对 $\forall \varepsilon > 0$,

$$P\left(\max_{m < j \leq n} |S_{m,j}| > 2\varepsilon\right) \min_{m < k \leq n} P(|S_{k,n}| \leq \varepsilon) \leq P(|S_{m,n}| > \varepsilon) \rightarrow 0 \quad (m, n \rightarrow \infty).$$

又知道 $\min P(|S_{k,n}| \leq \varepsilon) \rightarrow 1$, 仿照上一题设 $A_k(m, n)$, 从而

$$P\left(\max_{m < j \leq n} |S_{m,j}| > 2\varepsilon\right) = \sum_{m < k \leq n} P(A_k(m, n)) \rightarrow 0.$$

这里 $A_k(m, n)$ 强调 A_k 由 m, n 决定。由 A_k 关于 m 的单调性知道

$$\sum_{k=1}^{\infty} P(A_k(1, \infty)) \leq \sum_{k=1}^m P(A_k(1, \infty)) + \sum_{k=m}^{\infty} P(A_k(m, \infty)),$$

那么令 $m \rightarrow \infty$ 知道不等式右侧有界, 从而有限。由 *Borel-Cantelli* 引理知道 $P(A_k(1, \infty) \text{ i.o.}) = 0$, 从而 S_n 几乎处处为 *Cauchy* 列, 从而收敛。 \square

练习 2.5.11: Let X_1, X_2, \dots be i.i.d. and $S_n = X_1 + \dots + X_n$. Use (\star) to conclude that if $S_n/n \rightarrow 0$ in probability then

$$\frac{\max_{1 \leq m \leq n} S_m}{n} \rightarrow 0$$

in probability.

证明: 对 $\forall \varepsilon > 0$, 直接应用 (\star) 并令 $m = 0$ 得到

$$P\left(\max_{k \leq n} |S_k| > 2\varepsilon n\right) \min_{k \leq n} P(|S_{k,n}| \leq \varepsilon n) \leq P(|S_n| > \varepsilon n).$$

令 $n \rightarrow \infty$, 那么 RHS $\rightarrow 0$, 对每个 k 有 $P(|S_{k,n}| \leq \varepsilon n) \rightarrow 1$ 从而

$$\min_{k \leq n} P(|S_{k,n}| \leq \varepsilon n) \rightarrow 1,$$

那么

$$P\left(\max_{k \leq n} |S_k| > 2\varepsilon n\right) \rightarrow 0,$$

从而直接得出结论。 \square

练习 2.5.12: Let X_1, X_2, \dots be i.i.d. and $S_n = X_1 + \dots + X_n$. Suppose $a_n \uparrow \infty$ and

$$\frac{a(2^n)}{a(2^{n-1})}$$

is bounded.

1. Use (\star) to show that if $S_n/a(n) \rightarrow 0$ in probability and

$$\frac{S_{2^n}}{a(2^n)} \rightarrow 0$$

a.s. then $S_n/a(n) \rightarrow 0$ a.s. ;

2. Suppose in addition that $EX_1 = 0$ and $EX_1^2 < \infty$. Use the previous exercise and Chebyshev's inequality to conclude that

$$\frac{S_n}{n^{1/2}(\log_2 n)^{1/2+\varepsilon}} \rightarrow 0 \text{ a.s. .}$$

证明:

1. 使用 (\star) 得到

$$\begin{aligned} P\left(\max_{2^{n-1} < k \leq 2^n} |S_{2^{n-1}, k}| > 2\varepsilon a(2^n)\right) &\leq \min_{2^{n-1} < k \leq 2^n} P(|S_{k, 2^n}| \leq \varepsilon a(2^n)) \\ &\leq P(|S_{2^{n-1}, 2^n}| > \varepsilon a_{2^n}). \end{aligned}$$

对每个 $2^{n-1} < k \leq 2^n$, 由于 $S_n/a(n) \rightarrow 0$ 所以

$$P(|S_{k, 2^n}| > \varepsilon a(2^n)) \leq P(|S_k| > \varepsilon a(2^n)) + P(|S_{2^n}| > \varepsilon a(2^n)) \rightarrow 0,$$

从而

$$P(|S_{k, 2^n}| \leq \varepsilon a(2^n)) \rightarrow 1.$$

那么在 $n \rightarrow \infty$ 的情况下有

$$P\left(\max_{2^{n-1} < k \leq 2^n} |S_{2^{n-1}, k}| > 2\varepsilon a(2^n)\right) \leq P(|S_{2^{n-1}, 2^n}| > \varepsilon a_{2^n}).$$

由 $S_{2^n}/a(2^n) \xrightarrow{\text{a.s.}} 0$ 知道 $P(|S_{2^{n-1}, 2^n}| > \varepsilon a(2^n) \text{ i.o.}) = 0$, 再应用 *Borel-Cantelli* 引理知道

$$\sum_n P\left(\max_{2^{n-1} < k \leq 2^n} |S_{2^{n-1}, k}| > 2\varepsilon a(2^n)\right) \leq \sum_n P(|S_{2^{n-1}, 2^n}| > \varepsilon a_{2^n}) < \infty.$$

再应用 *Borel-Cantelli* 引理有

$$P\left(\max_{2^{n-1} < k \leq 2^n} \frac{|S_{2^{n-1}, k}|}{a(2^n)} > 2\varepsilon \text{ i.o.}\right) = 0.$$

又考虑到 $a(2^n) \leq M \cdot a(2^{n-1}) \leq M \cdot a(k)$, 那么有

$$P\left(\max_{2^{n-1} < k \leq 2^n} \frac{|S_{2^{n-1},k}|}{a(k)} > 2\varepsilon M \text{ i.o.}\right) \leq P\left(\max_{2^{n-1} < k \leq 2^n} \frac{|S_{2^{n-1},k}|}{a(2^n)} > 2\varepsilon \text{ i.o.}\right) = 0.$$

结合题目中 $a.s.$ 收敛的条件直接得到结论。

2. 对 $2^{n-1} < k < 2^n$, 有 $a_k \leq a_n \leq [\log_2 k]M$, 那么 $\forall \delta > 0$, 有

$$\begin{aligned} P\left(\frac{S_n}{\sqrt{n}(\log_2 n)^{1/2+\varepsilon}} > \delta\right) &\leq \frac{nEX_1^2}{\delta^2 n(\log_2 n)^{1+2\varepsilon}} \\ &\leq \frac{M \cdot EX_1^2}{\delta^2 a_n (\log_2 n)^{2\varepsilon}} \rightarrow 0 \quad (n \rightarrow \infty), \end{aligned}$$

从而

$$\frac{S_n}{n^{1/2}(\log_2 n)^{1/2+\varepsilon}} \xrightarrow{P} 0.$$

要进一步证明 $a.s.$ 收敛, 先证明子列 $a.s.$ 收敛。

$$\begin{aligned} P\left(\frac{S_{2^n}}{2^{n/2}n^{1/2+\varepsilon}} > \delta\right) &\leq \frac{2^n EX_1^2}{\delta^2 2^n n^{1+2\varepsilon}} \\ &= \frac{EX_1^2}{\delta^2 n^{1+2\varepsilon}} \end{aligned}$$

无穷级数收敛, 从而由 *Borel-Cantelli* 引理知道 2^n 这个子列是 $a.s.$ 收敛的。然后套用上一问结论即可得证。

□

2.6. 更新过程

练习 2.6.1: Show that $t/E(\xi_i \wedge t) \leq U(t) \leq 2t/E(\xi_i \wedge t)$.

练习 2.6.2: Deduce Theorem 2.6.3 from Theorem 2.6.1 by showing

$$\limsup_{t \rightarrow \infty} E\left(\frac{N_t}{t}\right)^2 < \infty.$$

Hint: Use a comparison like the one in the proof of Theorem 2.6.3.

练习 2.6.3: Customers arrive at times of a Poisson process with rate 1. If the server is occupied, they leave. (Think of a public telephone or prostitute.) If not, they enter service and require a service time with a distribution F that has mean μ . Show that the times at

which customers enter service are a renewal process with mean $\mu + 1$, and use Theorem 2.6.1 to conclude that the asymptotic fraction of customers served is $1/(\mu + 1)$.

In the remaining problems we assume that F is non-arithmetic, and in problems where the mean appears we assume it is finite.

练习 2.6.4: Let $A_t = t - T_{N(t)-1}$ be the “age” at time t , i.e., the amount of time since the last renewal. If we fix $x > 0$ then $H(t) = P(A_t > x)$ satisfies the renewal equation

$$H(t) = (1 - F(t)) \cdot 1_{(x, \infty)}(t) + \int_0^t H(t-s) dF(s)$$

so

$$P(A_t > x) \rightarrow \frac{1}{\mu} \int_{(x, \infty)} (1 - F(t)) dt,$$

which is the limit distribution for the residual lifetime $B_t = T_{N(t)} - t$.

练习 2.6.5: Use the renewal equation in the last problem and Theorem 2.6.9 to conclude that if T is a rate λ Poisson process A_t has the same distribution as $\xi_i \wedge t$.

练习 2.6.6: Let $A_t = t - T_{N(t)-1}$ and $B_t = T_{N(t)} - t$. Show that

$$P(A_t > x, B_t > y) \rightarrow \frac{1}{\mu} \int_{x+y}^{\infty} (1 - F(t)) dt.$$

练习 2.6.7: Alternating renewal process. Let $\xi_1, \xi_2, \dots > 0$ be i.i.d. with distribution F_1 and let $\eta_1, \eta_2, \dots > 0$ be i.i.d. with distribution F_2 . Let $T_0 = 0$ and for $k \geq 1$ let $S_k = T_{k-1} + \xi_k$ and $T_k = S_k + \eta_k$. In words, we have a machine that works for an amount of time ξ_k , breaks down, and then requires η_k units of time to be repaired. Let $F = F_1 * F_2$ and $H(t)$ be the probability the machine is working at time t . Show that if F is non-arithmetic then as $t \rightarrow \infty$

$$H(t) \rightarrow \frac{\mu_1}{\mu_1 + \mu_2}$$

where μ_i is the mean of F_i .

练习 2.6.8: Write a renewal equation for $H(t) = P$ (number of renewals in $[0, t]$ is odd) and use the renewal theorem to show that $H(t) \rightarrow 1/2$. Note: This is a special case of the previous exercise.

练习 2.6.9: Renewal densities. Show that if $F(t)$ has a directly Riemann integrable density function $f(t)$, then the $V = U - 1_{[0, \infty)}$ has a density v that satisfies

$$v(t) = f(t) + \int_0^t v(t-s) dF(s)$$

Use the renewal theorem to conclude that if f is directly Riemann integrable then $v(t) \rightarrow 1/\mu$ as $t \rightarrow \infty$.

2.7. 大偏差

3. 中心极限定理

3.1. De Moivre-Laplace 定理

练习 3.1.1: Generalize the proof of Lemma 3.1.1 to conclude that if

$$\max_{1 \leq j \leq n} |c_{j,n}| \rightarrow 0, \quad \sum_{j=1}^n c_{j,n} \rightarrow \lambda,$$

and

$$\sup_n \sum_{j=1}^n |c_{j,n}| < \infty$$

then

$$\prod_{j=1}^n (1 + c_{j,n}) \rightarrow e^\lambda.$$

证明: 由题意容易知道 $\sum_j |c_{j,n}| \rightarrow 0$, 那么

$$\sum_j \left(c_{j,n} - \frac{1}{2} c_{j,n}^2 \right) \leq \sum_j \log(1 + c_{j,n}) \leq \sum_j c_{j,n},$$

夹逼得到 $\sum_j \log(1 + c_{j,n}) \rightarrow \lambda$, 从而

$$\prod_{j=1}^n (1 + c_{j,n}) = \exp \left(\sum_j \log(1 + c_{j,n}) \right) \rightarrow e^\lambda.$$

□

The next three exercises illustrate the use of Stirling's formula. In them, X_1, X_2, \dots are i.i.d. and $S_n = X_1 + \dots + X_n$.

练习 3.1.2: If the X_i have a Poisson distribution with mean 1, then S_n has a Poisson distribution with mean n , i.e., $P(S_n = k) = e^{-n} n^k / k!$. Use Stirling's formula to show that if $(k - n) / \sqrt{n} \rightarrow x$ then

$$\sqrt{2\pi n} P(S_n = k) \rightarrow \exp \left(-\frac{x^2}{2} \right).$$

As in the case of coin flips it follows that

$$P \left(a \leq \frac{S_n - n}{\sqrt{n}} \leq b \right) \rightarrow \int_a^b (2\pi)^{-1/2} e^{-x^2/2} dx$$

but proving the last conclusion is not part of the exercise.

证明：逆用 *Stirling* 公式有

$$\begin{aligned}\sqrt{2\pi n}P(S_n = n + m) &= \sqrt{2\pi n}n^n e^{-n} \cdot \frac{n^m}{(n + m)!} \\ &\rightarrow \frac{n^m n!}{(n + m)!} \\ &= \prod_i^m \left(1 + \frac{i}{n + i}\right).\end{aligned}$$

再应用上一问结论知道它趋近于 $\exp(x^2/2)$ 。 \square

In the next two examples you should begin by considering $P(S_n = k)$ when $k/n \rightarrow a$ and then relate $P(S_n = j + 1)$ to $P(S_n = j)$ to show $P(S_n \geq k) \leq CP(S_n = k)$.

练习 3.1.3: Suppose $P(X_i = 1) = P(X_i = -1) = 1/2$. Show that if $a \in (0, 1)$

$$\frac{1}{2n} \log P(S_{2n} \geq 2na) \rightarrow -\gamma(a),$$

where

$$\gamma(a) = \frac{1}{2} \{ (1 + a) \log(1 + a) + (1 - a) \log(1 - a) \}.$$

证明：

$$\begin{aligned}\frac{1}{2n} \log P(S_{2n} = 2k) &= \frac{1}{2n} \left(\frac{1}{4n} \log(n\pi) - \frac{n + k}{2n} \log(1 + a) - \frac{n - k}{2n} \log(1 - a) \right) \\ &\rightarrow -\frac{1 + a}{2} \log(1 + a) - \frac{1 - a}{2} \log(1 - a).\end{aligned}$$

再考虑 $2k + 2$ 的情况：

$$P(S_{2n} = 2k + 2) = \frac{n - k}{n + k + 1} P(S_{2n} = 2k) \leq \frac{1 - a}{1 + a} P(S_{2n} = 2k),$$

然后对所有偶数 $k' \geq 2k$ 求级数和即可。 \square

练习 3.1.4: Suppose $P(X_i = k) = e^{-1}/k!$ for $k = 0, 1, \dots$. Show that if $a > 1$

$$\frac{1}{n} \log P(S_n \geq na) \rightarrow a - 1 - a \log a.$$

证明：

$$\begin{aligned}\frac{1}{n} \log P(S_n = k) &\rightarrow \frac{1}{n} \left(k \log n - n - \frac{1}{2} \log(2\pi k) - k \log k + k \right) \\ &= a - a \log a - 1.\end{aligned}$$

对 $k > n$ 的情况, 考虑 $k+1$:

$$P(S_n = k+1) = \frac{n}{k+1} P(S_n = k) \leq \frac{n}{k} P(S_n = k),$$

然后对 $k' \geq k$ 级数求和即可。 \square

3.2. 弱收敛

练习 3.2.1: Give an example of random variables X_n with densities f_n so that $X_n \Rightarrow a$ uniform distribution on $(0, 1)$ but $f_n(x)$ does not converge to 1 for any $x \in [0, 1]$.

证明: 对 $\forall n > 0$, 一定能找到 k, i 使得 $n = k(k-1)/2 + i$, 其中 $i \leq k+1$ 。构造 X_n 的分布函数 F_n :

$$F_n = \begin{cases} \frac{i}{k+2}, & \frac{i}{k+2} \leq x \leq \frac{i+1}{k+2}. \\ \text{线性}, & \text{否则} \end{cases}$$

这样一定有 $F_n \rightarrow F = x$, 但是 $f_n(x)$ 不收敛, 因为 $f_n(x) = 0$ i.o. \square

练习 3.2.2: Convergence of maxima. Let X_1, X_2, \dots be independent with distribution F , and let $M_n = \max_{m \leq n} X_m$. Then $P(M_n \leq x) = F(x)^n$. Prove the following limit laws for M_n :

1. If $F(x) = 1 - x^{-\alpha}$ for $x \geq 1$ where $\alpha > 0$ then for $y > 0$

$$P\left(\frac{M_n}{n^{1/\alpha}} \leq y\right) \rightarrow \exp(-y^{-\alpha}).$$

2. If $F(x) = 1 - |x|^\beta$ for $-1 \leq x \leq 0$ where $\beta > 0$ then for $y < 0$

$$P(n^{1/\beta} M_n \leq y) \rightarrow \exp(-|y|^\beta).$$

3. If $F(x) = 1 - e^{-x}$ for $x \geq 0$ then for all $y \in (-\infty, \infty)$

$$P(M_n - \log n \leq y) \rightarrow \exp(-e^{-y}).$$

The limits that appear above are called the extreme-value distributions. The last one is called the double exponential or Gumbel distribution. Necessary and sufficient conditions for $(M_n - b_n)/a_n$ to converge to these limits were obtained by Gnedenko (1943). For a recent treatment, see Resnick (1987).

证明:

1. $P(M_n \leq yn^{1/\alpha}) = (1 - (yn^{1/\alpha})^{-\alpha})^n = \left(1 - \frac{y^{-\alpha}}{n}\right)^n \rightarrow \exp(-y^{-\alpha}).$
2. $P(M_n \leq yn^{-1/\beta}) = (1 - |yn^{-1/\beta}|^\beta)^n \rightarrow \exp(-|y|^\beta).$
3. $P(M_n \leq y + \log n) = (1 - \exp(-y - \log n))^n \rightarrow \exp(-e^{-y}).$

□

练习 3.2.3: Let X_1, X_2, \dots be i.i.d. and have the standard normal distribution.

1. From Theorem 1.2.6, we know

$$P(X_i > x) \sim \frac{1}{\sqrt{2\pi}x} e^{-x^2/2} \quad \text{as } x \rightarrow \infty.$$

Use this to conclude that for any real number θ

$$\frac{P(X_i > x + (\theta/x))}{P(X_i > x)} \rightarrow e^{-\theta}.$$

2. Show that if we define b_n by $P(X_i > b_n) = 1/n$

$$P(b_n(M_n - b_n) \leq x) \rightarrow \exp(-e^{-x}).$$

3. Show that $b_n \sim (2 \log n)^{1/2}$ and conclude $M_n/(2 \log n)^{1/2} \rightarrow 1$ in probability.

证明:

$$\begin{aligned} 1. \quad \frac{P(X_i > x + (\theta/x))}{P(X_i > x)} &\rightarrow \sqrt{\frac{x}{x + \theta/x}} \exp\left(\frac{x^2 - (x + \theta/x)^2}{2}\right) \\ &\rightarrow e^{-\theta}. \end{aligned}$$

$$\begin{aligned} 2. \quad P\left(M_n \leq \frac{x}{b_n} + b_n\right) &= \left(1 - P\left(X_n > \frac{x}{b_n} + b_n\right)\right)^n \\ &\rightarrow (1 - P(X_n > b_n)e^{-x})^n \\ &\rightarrow \left(1 - \frac{e^{-x}}{n}\right)^n \\ &\rightarrow \exp(-e^{-x}). \end{aligned}$$

$$3. \quad P(X_i > \sqrt{2 \log n}) = \frac{1}{\sqrt{2\pi}\sqrt{2 \log n}} \cdot \frac{1}{n} < \frac{1}{n},$$

从而 $b_n < \sqrt{2 \log n}$ 。而

$$P(X_i > \sqrt{2 \log n - 2 \log \log n}) = \frac{1}{\sqrt{2\pi}\sqrt{2 \log n - 2 \log \log n}} \cdot \frac{\log n}{n} > \frac{1}{n},$$

从而 $b_n > \sqrt{2 \log n - 2 \log \log n}$ 。对于另外一问,

$$\begin{aligned}
P(M_n > \varepsilon \sqrt{2 \log n}) &= 1 - \left(1 - P(X_1 > \varepsilon \sqrt{2 \log n})\right)^n \\
&= 1 - \left(1 - \frac{1}{\sqrt{2\pi\varepsilon\sqrt{2\log n}}} \cdot \frac{1}{n^\varepsilon}\right)^n \\
&\rightarrow 1 - e^0 = 0.
\end{aligned}$$

□

练习 3.2.4: Fatou's lemma. Let $g \geq 0$ be continuous. If $X_n \Rightarrow X_\infty$ then

$$\liminf_{n \rightarrow \infty} \mathbb{E}g(X_n) \geq \mathbb{E}g(X_\infty).$$

证明: 令 Y_n, Y 分别为和 X_n, X_∞ 同分布的随机变量, 且满足 $Y_n \xrightarrow{\text{a.s.}} Y$, 然后利用 Fatou 引理有

$$\liminf \mathbb{E}g(X_n) = \liminf \mathbb{E}g(Y_n) \geq \mathbb{E}g(Y) = \mathbb{E}g(X_\infty).$$

□

练习 3.2.5: Integration to the limit. Suppose g, h are continuous with $g(x) > 0$, and $|h(x)|/g(x) \rightarrow 0$ as $|x| \rightarrow \infty$. If $F_n \Rightarrow F$ and

$$\int g(x) dF_n(x) \leq C < \infty$$

then

$$\int h(x) dF_n(x) \rightarrow \int h(x) dF(x).$$

证明: 设 X_n, X 分别为服从 F_n, F 的随机变量, 满足 $X_n \xrightarrow{\text{a.s.}} X$. 那么由控制收敛定理得到 $\mathbb{E}h(X_n) \rightarrow \mathbb{E}h(X)$, 也即所求。□

练习 3.2.6: The Lévy Metric. Show that

$$\rho(F, G) = \inf\{\varepsilon : F(x - \varepsilon) - \varepsilon \leq G(x) \leq F(x + \varepsilon) + \varepsilon \text{ for all } x\}$$

defines a metric on the space of distributions and $\rho(F_n, F) \rightarrow 0$ if and only if $F_n \Rightarrow F$.

证明: 正定性和对称性显然。设 $\rho(F, G) = \rho_1, \rho(G, H) = \rho_2$, 那么对 $\forall x$ 有

$$\begin{aligned}
F(x - \rho_1 - \rho_2) - \rho_1 - \rho_2 &\leq G(x - \rho_2) - \rho_2 \leq \\
&H(x) \\
&\leq G(x + \rho_2) + \rho_2 \leq F(x + \rho_1 + \rho_2) + \rho_1 + \rho_2
\end{aligned}$$

成立, 从而 $\rho(H, F) \leq \rho_1 + \rho_2$ 。

先证必要性, 假设 $\rho(F_n, F) \rightarrow 0$. 取连续点 $x \in \mathbb{R}$, 那么对 $\forall \varepsilon > 0$, 存在足够大的 n 使得 $\rho(F_n, F) < \varepsilon$, 那么有

$$F_n(x - \varepsilon) - \varepsilon \leq F(x) \leq F_n(x + \varepsilon) + \varepsilon.$$

$n \rightarrow \infty$ 时 $\varepsilon \rightarrow 0$, 得到 $\limsup F_n(x) \leq F(x) \leq \liminf F_n(x)$, 从而 $\lim F_n(x) = F(x)$, 那么 $F_n \Rightarrow F$.

再证充分性, 假设 $F_n \Rightarrow F$,

TODO

□

练习 3.2.7: The Ky Fan metric on random variables is defined by

$$\alpha(X, Y) = \inf\{\varepsilon \geq 0 : P(|X - Y| > \varepsilon) \leq \varepsilon\}.$$

Show that if $\alpha(X, Y) = \alpha$ then the corresponding distributions have Lévy distance $\rho(F, G) \leq \alpha$.

证明: 只要证明 α 满足 $F(x - \alpha) - \alpha \leq G(x) \leq F(x + \alpha) + \alpha$ 对 $\forall x$ 成立即可。对第一个不等式有

$$\begin{aligned} F(x - \alpha) - G(x) &= (P(X \leq x - \alpha, Y \leq x) + P(X \leq x - \alpha, Y > x)) \\ &\quad - (P(X \leq x - \alpha, Y \leq x) + P(X > x - \alpha, Y \leq x)) \\ &\leq P(X \leq x - \alpha, Y > x) \\ &\leq P(|X - Y| > \alpha) \leq \alpha. \end{aligned}$$

可以类似证明另一半。

□

练习 3.2.8: Let $\alpha(X, Y)$ be the metric in the previous exercise and let

$$\beta(X, Y) = E\left(\frac{|X - Y|}{1 + |X - Y|}\right)$$

be the metric of Exercise 2.3.6. If $\alpha(X, Y) = a$ then

$$\frac{a^2}{1 + a} \leq \beta(X, Y) \leq a + (1 - a)\frac{a}{1 + a}.$$

证明:

$$\begin{aligned}
\beta(X, Y) &= 1 - E\left(\frac{1}{1 + |X + Y|}\right) \\
&\leq 1 - E\left(\frac{1}{1 + |X + Y|}; |X - Y| \leq a\right) \\
&\leq 1 - \frac{1}{1 + a}(1 - a) \\
&= a + (1 - a)\frac{a}{1 + a}.
\end{aligned}$$

对于另一半不等式, 有

$$\begin{aligned}
\beta(X, Y) &\geq E\left(\frac{|X - Y|}{1 + |X - Y|}; |X - Y| > a\right) \\
&\geq \frac{a}{1 + a} \cdot a = \frac{a^2}{1 + a}.
\end{aligned}$$

□

练习 3.2.9: If $F_n \Rightarrow F$ and F is continuous then $\sup_x |F_n(x) - F(x)| \rightarrow 0$.

证明: 对 $\forall \varepsilon > 0$, 取正数 M 足够大使得 $|F(M) - 1| < \varepsilon$ 且 $|F(-M) + 1| < \varepsilon$ 。可以取足够大的 n 使得

$$\begin{aligned}
\sup_{x > M} |F_n(x) - F(x)| &\leq \sup_{x > M} |F_n(x) - 1| + \sup_{x > M} |F(x) - 1| \\
&\leq |F_n(M) - 1| + |F(M) - 1| \\
&\leq 2\varepsilon
\end{aligned}$$

且

$$\begin{aligned}
\sup_{x < -M} |F_n(x) - F(x)| &\leq \sup_{x < -M} |F_n(x) + 1| + \sup_{x < -M} |F(x) + 1| \\
&\leq |F_n(-M) + 1| + |F(-M) + 1| \\
&\leq 2\varepsilon.
\end{aligned}$$

再考虑 $[-M, M]$ 上的行为, 在这个紧集上容易用有限覆盖定理证明一致收敛性。

□

练习 3.2.10: If F is any distribution function there is a sequence of distribution functions of the form

$$\sum_{m=1}^n a_{n,m} 1_{(x_{n,m} \leq x)}$$

with $F_n \Rightarrow F$.

证明: 直接构造这样的 F_n 。令 $x_1 = -\sqrt{n}/2, x_2, \dots, x_n = \sqrt{n}/2$ 为一个平均划分, 然后令

$$F_n(x) = \begin{cases} 0, & x < x_1 \\ F(x_1), & x_1 \leq x < x_2 \\ \vdots \\ F(x_n), & x \geq x_n \end{cases}$$

那么只要证明对 F 的连续点 x 有 $F_n(x) \rightarrow F(x)$ 。对 $\forall \varepsilon > 0$, 考虑 x 的一个 F 连续邻域 $U(x, \delta)$ 满足 $|F(x - \delta) - F(x)| < \varepsilon$, 令 n 足够大使得两个相邻划分点的距离小于 δ , 那么 $F(x) - F_n(x) \leq F(x) - F(x - \delta) < \varepsilon$ 。□

练习 3.2.11: Let $X_n, 1 \leq n \leq \infty$, be integer-valued. Show that $X_n \Rightarrow X_\infty$ if and only if $P(X_n = m) \rightarrow P(X_\infty = m)$ for all m .

证明: 记 X_n, X_∞ 的分布函数为 F_n, F 。先证必要性, 设 $X_n \Rightarrow X_\infty$ 。对 $\forall m \in \mathbb{Z}$, $m \pm 1/2$ 是 F 的连续点, 从而

$$\begin{aligned} P(X_n = m) &= F_n(m + 1/2) - F_n(m - 1/2) \\ &\rightarrow F(m + 1/2) - F(m - 1/2) \\ &= P(X_\infty = m). \end{aligned}$$

再证充分性, 设 $P(X_n = m) \rightarrow P(X_\infty = m)$, 那么对 $\forall x \in \mathbb{R} \setminus \mathbb{Z}$, 取 $m \in \mathbb{Z}$ 使得 $x \in [m, m + 1)$ 。

$$\begin{aligned} F_n(x) &= \sum_{-\infty}^m P(X_n = k) \\ &\rightarrow \sum_{-\infty}^m P(X = k) \\ &= F(x), \end{aligned}$$

其中运用了有界收敛定理, 或者级数理论中的对应定理。□

练习 3.2.12: Show that if $X_n \rightarrow X$ in probability then $X_n \Rightarrow X$ and that, conversely, if $X_n \Rightarrow c$, where c is a constant then $X_n \rightarrow c$ in probability.

证明: 先证第一条结论, 设 $X_n \xrightarrow{P} X$ 。由前面习题知道 $E(|X_n - X|/(1 + |X_n - X|)) \rightarrow 0$, 那么 $Ky Fan$ 度量 $\alpha(X_n, X) \rightarrow 0$, 那么 $Levy$ 度量 $\rho(X_n, X) \rightarrow 0$, 那么 $X_n \Rightarrow X$ 。

再证第二条结论。由 $X_n \Rightarrow c$ 知道对 $\forall \varepsilon > 0$ 有 $F_n(c + \varepsilon) \rightarrow 1, F_n(c - \varepsilon) \rightarrow 0$, 那么

$$P(|X_n - c| > \varepsilon) = F_n(c - \varepsilon) + 1 - F_n(c + \varepsilon) \rightarrow 0.$$

□

练习 3.2.13: Converging together lemma. If $X_n \Rightarrow X$ and $Y_n \Rightarrow c$, where c is a constant then $X_n + Y_n \Rightarrow X + c$. A useful consequence of this result is that if $X_n \Rightarrow X$ and $Z_n - X_n \Rightarrow 0$ then $Z_n \Rightarrow X$.

证明: 取 X'_n, X', Y'_n 分别与 X_n, X, Y_n 同分布且 $X'_n \xrightarrow{\text{a.s.}} X', Y'_n \xrightarrow{\text{a.s.}} c$, 那么 $X'_n + Y'_n \xrightarrow{\text{a.s.}} X' + c$, 从而 $X'_n + Y'_n \Rightarrow X' + c$, 又考虑到同分布从而得到结论。 □

练习 3.2.14: Suppose $X_n \Rightarrow X$, $Y_n \geq 0$, and $Y_n \Rightarrow c$, where $c > 0$ is a constant then $X_n Y_n \Rightarrow cX$. This result is true without the assumptions $Y_n \geq 0$ and $c > 0$. We have imposed these only to make the proof less tedious.

证明: 与上一题同理。 □

练习 3.2.15: Show that if $X_n = (X_n^1, \dots, X_n^n)$ is uniformly distributed over the surface of the sphere of radius \sqrt{n} in \mathbb{R}^n then $X_n^1 \Rightarrow$ a standard normal. Hint: Let Y_1, Y_2, \dots be i.i.d. standard normals and let

$$X_n^i = Y_i \sqrt{\frac{n}{\sum_{m=1}^n Y_m^2}}.$$

证明: 按照题意构造 Y_i 和 X_n^i 。先证明这个构造是超球面上的均匀分布。有 $Y_n^i \sim \mathcal{N}(0, 1)$, 那么 $Y_n = (Y_n^i)_i \sim \mathcal{N}(\mathbf{0}, \mathbf{I}_n)$ 满足 $AY_n \sim \mathcal{N}(\mathbf{0}, \mathbf{I}_n)$, 从而 Y_n 分布是球对称的, 那么 Y_n 服从均匀分布。那么 $X_n = \sqrt{n}Y_n/\|Y_n\|$ 也服从均匀分布。

再证明这个均匀分布的分量弱收敛于标准正态分布。由正态分布性质知道 $EY_i^2 = 1$, 那么由强大数定律知道

$$\frac{\sum_{m=1}^n Y_m^2}{n} \xrightarrow{\text{a.s.}} EY_1^2 = 1,$$

那么 $X_n^i \xrightarrow{\text{a.s.}} Y_i \sim \mathcal{N}(0, 1)$, 从而弱收敛。 □

练习 3.2.16: Suppose $Y_n \geq 0$, $EY_n^\alpha \rightarrow 1$ and $EY_n^\beta \rightarrow 1$ for some $0 < \alpha < \beta$. Show that $Y_n \rightarrow 1$ in probability.

证明: **TODO** □

练习 3.2.17: For each $K < \infty$ and $y < 1$ there is a $c_{y,K} > 0$ so that $EX^2 = 1$ and $EX^4 \leq K$ implies $P(|X| > y) \geq c_{y,K}$.

证明: *TODO*

□

3.3. 特征函数

3.3.1. 定义, 逆转公式

练习 3.3.1: Show that if φ is a ch.f. then $\operatorname{Re} \varphi$ and $|\varphi|^2$ are also.

证明:

$$\operatorname{Re} \varphi = \frac{1}{2}(\varphi + \bar{\varphi}) = \frac{1}{2}(\varphi(t) + \varphi(-t))$$

对应分布函数为 $1/2 \cdot (F(x) + F(-x))$ 的特征函数。

$$|\varphi|^2 = \varphi(t)\varphi(-t)$$

对应随机变量为 $X \cdot (-Y)$ 的特征函数, 其中 X, Y 独立同分布且特征函数均为 φ 。

□

练习 3.3.2:

1. Imitate the proof of Theorem 3.3.11 to show that

$$\mu(\{a\}) = \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T e^{-ita} \varphi(t) dt;$$

2. If $P(X \in h\mathbb{Z}) = 1$ where $h > 0$ then its ch.f. has $\varphi(2\pi/h + t) = \varphi(t)$ so

$$P(X = x) = \frac{h}{2\pi} \int_{-\pi/h}^{\pi/h} e^{-itx} \varphi(t) dt \quad \text{for } x \in h\mathbb{Z};$$

3. If $X = Y + b$ then $E \exp(itX) = e^{itb} E \exp(itY)$. So if $P(X \in b + h\mathbb{Z}) = 1$, the inversion formula in (2) is valid for $x \in b + h\mathbb{Z}$.

证明:

1. 令

$$\begin{aligned} I_T &= \frac{1}{2T} \int_{-T}^T e^{-ita} \varphi(t) dt \\ &= \frac{1}{2T} \int \int_{-T}^T e^{it(x-a)} dt \mu(dx) \\ &= \frac{1}{2T} \int \int_{-T}^T (\cos t(x-a) + i \sin t(x-a)) dt \mu(dx) \\ &= \int \frac{1}{2T} \int_{-T}^T \cos t(x-a) dt \mu(dx). \end{aligned}$$

注意到积分号里面的部分绝对值小于等于 1 从而有界, 那么用有界收敛定理得到

$$\begin{aligned} I_T &= \int \frac{1}{2T} \int_{-T}^T \cos t(x-a) dt \mu(dx) \\ &\rightarrow \int \left[f(x) = \begin{cases} 1, & x = a \\ 0, & x \neq a \end{cases} \right] \mu(dx) \\ &= \mu(\{a\}). \end{aligned}$$

2. 因为 $P(X \in h\mathbb{Z}) = 1$ 那么

$$\varphi\left(\frac{2\pi}{h} + t\right) - \varphi(t) = \sum_{k \in \mathbb{Z}} P(X = kh)(e^{2\pi i k} - 1)e^{ithk} = 0.$$

再由第一问知道

$$\begin{aligned} P(X = x) &= \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T e^{-itx} \varphi(t) dt \\ &= \lim_{\substack{T \rightarrow \infty \\ T \in (2\pi/h)\mathbb{Z}}} \frac{1}{2T} \int_{-T}^T e^{-itx} \varphi(t) dt \quad (\text{单调性}) \\ &= \lim_{n \rightarrow \infty} \frac{h}{4n\pi} \int_{-2\pi n/h}^{2\pi n/h} e^{-itx} \varphi(t) dt \\ &= \frac{h}{2\pi} \int_{-\pi/h}^{\pi/h} e^{-itx} \varphi(t) dt. \end{aligned}$$

□

练习 3.3.3: Suppose X and Y are independent and have ch.f. φ and distribution μ . Apply Exercise 3.3.2 to $X - Y$ and use Exercise 2.1.5 to get

$$\lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T |\varphi(t)|^2 dt = P(X - Y = 0) = \sum_x \mu(\{x\})^2.$$

Remark. The last result implies that if $\varphi(t) \rightarrow 0$ as $t \rightarrow \infty$, μ has no point masses. Exercise 3.3.11 gives an example to show that the converse is false. The Riemann-Lebesgue Lemma (Exercise 1.4.4) shows that if μ has a density, $\varphi(t) \rightarrow 0$ as $t \rightarrow \infty$.

证明: 有 $X - Y$ 的特征函数为

$$\mathbb{E}e^{it(X-Y)} = \mathbb{E}e^{itX}\mathbb{E}e^{-itY} = \varphi\bar{\varphi} = |\varphi|^2,$$

然后应用上一题结论即可得到左半等式。直接应用练习 2.1.5 得到右半等式。 □

练习 3.3.4: Give an example of a measure μ with a density but for which

$$\int |\varphi(t)| dt = \infty.$$

Hint: Two of the examples above have this property.

证明: 指数分布的特征函数 $\varphi(t) = 1/(1 - it)$, 均匀分布的特征函数

$$\varphi(t) = \frac{e^{itb} - e^{ita}}{it(b - a)}.$$

□

练习 3.3.5: Show that if X_1, \dots, X_n are independent and uniformly distributed on $(-1, 1)$, then for $n \geq 2$, $X_1 + \dots + X_n$ has density

$$f(x) = \frac{1}{\pi} \int_0^\infty \left(\frac{\sin t}{t} \right)^n \cos tx \, dt.$$

Although it is not obvious from the formula, f is a polynomial in each interval $(k, k + 1)$, $k \in \mathbb{Z}$ and vanishes on $[-n, n]^c$.

证明: X_1 分布的特征函数为

$$\varphi(t) = \frac{e^{it} - e^{-it}}{2it} = \frac{\sin t}{t},$$

从而 $X_1 + \dots + X_n$ 的特征函数为 $\varphi^n(t)$, 看出

$$\int |\varphi(t)| dt < \infty,$$

再用逆转公式

$$f(x) = \frac{1}{2\pi} \int \left(\frac{\sin t}{t} \right)^n e^{-itx} dx = \frac{1}{\pi} \int_0^\infty \left(\frac{\sin t}{t} \right)^n \cos(tx) dx,$$

最后一步把积分中的奇函数部分去掉。

□

练习 3.3.6: Use the result in Example 3.3.16 to conclude that if X_1, X_2, \dots are independent and have the Cauchy distribution, then $(X_1 + \dots + X_n)/n$ has the same distribution as X_1 .

证明: Cauchy 分布的特征函数为 $\varphi = \exp(-|t|)$, 而 $(X_1 + \dots + X_n)/n$ 分布的特征函数为

$$\exp\left(-\left|\frac{t}{n}\right|\right)^n = \exp(-|t|) = \varphi.$$

根据反转公式，它们的分布相同。 \square

3.3.2. 弱收敛

练习 3.3.7: Suppose that $X_n \Rightarrow X$ and X_n has a normal distribution with mean 0 and variance σ_n^2 . Prove that $\sigma_n^2 \rightarrow \sigma^2 \in [0, \infty)$.

证明: 由于 $X_n \Rightarrow X$ 那么 X_n 的特征函数 $\exp(-\sigma_n^2 t^2/2) \rightarrow \exp(-\sigma^2 t^2/2)$ 对 $\forall x \in \mathbb{R}$ 成立，那么 $\sigma_n^2 \rightarrow \sigma^2$ 。 \square

练习 3.3.8: Show that if X_n and Y_n are independent for $1 \leq n \leq \infty$, $X_n \Rightarrow X_\infty$, and $Y_n \Rightarrow Y_\infty$, then $X_n + Y_n \Rightarrow X_\infty + Y_\infty$.

证明: 分别记 X_n, Y_n 分布函数为 φ_n, ψ_n 。那么由 $X_n \Rightarrow X_\infty$ 知道 $\varphi_n \rightarrow \varphi$ ，由 $Y_n \Rightarrow Y_\infty$ 知道 $\psi_n \rightarrow \psi$ ，那么 $\varphi_n \psi_n \rightarrow \varphi \psi$ ，从而 $X_n + Y_n \Rightarrow X_\infty + Y_\infty$ 。 \square

练习 3.3.9: Let X_1, X_2, \dots be independent and let $S_n = X_1 + \dots + X_n$. Let φ_j be the ch.f. of X_j and suppose that $S_n \rightarrow S_\infty$ a.s. Then S_∞ has ch.f.

$$\prod_{j=1}^{\infty} \varphi_j(t).$$

证明: 由于 $S_n \xrightarrow{\text{a.s.}} S$ 从而 $S_n \Rightarrow S$ ，那么 $\prod \varphi_j$ 收敛，从而 S_∞ 有特征函数

$$\prod_{j=1}^{\infty} \varphi_j(t).$$

\square

练习 3.3.10: Using the identity $\sin t = 2 \sin(t/2) \cos(t/2)$ repeatedly leads to

$$\frac{\sin t}{t} = \prod_{m=1}^{\infty} \cos\left(\frac{t}{2^m}\right).$$

Prove the last identity by interpreting each side as a characteristic function.

证明: 令 X_m 满足 $P(X_m = 1/2^m) = P(X_m = -1/2^m) = 1/2$ ，那么 X_m 的特征函数就是 $\cos(t/2^m)$ 。考虑 $\sum X_m$ 的分布，令 $Y_m = (2^m X_m + 1)/2$ 满足 $P(Y_m = 1) = P(Y_m = 0) = 1/2$ 且相互独立，从而

$$\sum X_m = \sum \frac{2Y_m - 1}{2^m} = -1 + \sum \frac{Y_m}{2^{m-1}}.$$

其中最后一项可以看作是对二进制每一位随机取 0 或 1, 从而 $\sum X_m$ 服从 $[-1, 1]$ 上的均匀分布, 其特征函数为 $\sin t/t$. \square

练习 3.3.11: Let X_1, X_2, \dots be independent taking values 0 and 1 with probability 1/2 each.

$$X = 2 \sum_{j \geq 1} \frac{X_j}{3^j}$$

has the Cantor distribution. Compute the ch.f. φ of X and notice that φ has the same value at $t = 3^k \pi$ for $k = 0, 1, 2, \dots$.

证明: 直接计算得到

$$\varphi = \prod_{j=1}^{\infty} \frac{\exp(2it/3^j) + 1}{2}.$$

容易看出它是收敛的, 且在 $3^k \pi$ 处取值相同. \square

3.3.3. 矩和导数

练习 3.3.12: Use Theorem 3.3.18 and the series expansion for $e^{-t^2/2}$ to show that the standard normal distribution has

$$EX^{2n} = (2n)!/2^n n! = (2n-1)(2n-3)\cdots 3 \cdot 1 \equiv (2n-1)!!.$$

证明:

\square

练习 3.3.13:

1. Suppose that the family of measures $\{\mu_i, i \in I\}$ is tight, i.e., $\sup_i \mu_i([-M, M]^c) \rightarrow 0$ as $M \rightarrow \infty$. Use (4) in Theorem 3.3.1 and (3.3.3) with $n = 0$ to show that their ch.f.'s φ_i are equicontinuous, i.e., if $\varepsilon > 0$ we can pick $\delta > 0$ so that $|h| < \delta$ then

$$|\varphi_i(t+h) - \varphi_i(t)| < \varepsilon.$$

2. Suppose $\mu_n \Rightarrow \mu_\infty$. Use Theorem 3.3.17 and equicontinuity to conclude that the ch.f.'s $\varphi_n \rightarrow \varphi_\infty$ uniformly on compact sets. [Argue directly. You don't need to go to AA.]
3. Give an example to show that the convergence need not be uniform on the whole real line.

练习 3.3.14: Let X_1, X_2, \dots be i.i.d. with characteristic function φ .

1. If $\varphi'(0) = ia$ and $S_n = X_1 + \dots + X_n$ then $S_n/n \rightarrow a$ in probability.

2. If $S_n/n \rightarrow a$ in probability then $(\varphi(t)/n)^n \rightarrow e^{iat}$ as $n \rightarrow \infty$ through the integers.
3. Use (2) and the uniform continuity established in (4) of Theorem 3.3.1 to show that $(\varphi(h) - 1)/h \rightarrow -ia$ as $h \rightarrow 0$ through the positive reals. Thus the weak law holds if and only if $\varphi'(0)$ exists. This result is due to E.J.G. Pitman (1956), with a little help from John Walsh who pointed out that we should prove (3).

练习 3.3.15:

$$2 \int_0^\infty \frac{1 - \operatorname{Re} \varphi(t)}{\pi t^2} dt = \int |y| dF(y).$$

Hint: Change variables $x = |y|/t$ in the density function of Example 3.3.15, which integrates to 1.

练习 3.3.16: Show that if

$$\lim_{t \downarrow 0} \frac{\varphi(t) - 1}{t^2} = c > -\infty$$

then $EX = 0$ and $E|X|^2 = -2c < \infty$. In particular, if $\varphi(t) = 1 + o(t^2)$ then $\varphi(t) \equiv 1$.

练习 3.3.17: If Y_n are r.v.'s with ch.f.'s φ_n then $Y_n \Rightarrow 0$ if and only if there is a $\delta > 0$ so that $\varphi_n(t) \rightarrow 1$ for $|t| \leq \delta$.

练习 3.3.18: Let X_1, X_2, \dots be independent. If

$$S_n = \sum_{m \leq n} X_m$$

converges in distribution then it converges in probability (and hence a.s. by Exercise 2.5.10). Hint: The last exercise implies that if $m, n \rightarrow \infty$ then $S_m - S_n \rightarrow 0$ in probability. Now use Exercise 2.5.11.

3.3.4. Polya 准则

练习 3.3.19: Show that $\exp(-|t|^\alpha)$ is a characteristic function for $0 < \alpha \leq 1$.

练习 3.3.20: If X_1, X_2, \dots are independent and have characteristic function $\exp(-|t|^\alpha)$ then $(X_1 + \dots + X_n)/n^{1/\alpha}$ has the same distribution as X_1 .

We will return to this topic in Section 3.8.

练习 3.3.21: Let φ_1 and φ_2 be ch.f.'s. Show that $A = \{t : \varphi_1(t) = \varphi_2(t)\}$ is closed, contains 0, and is symmetric about 0. Show that if A is a set with these properties and $\varphi_1(t) = e^{-|t|}$ there is a φ_2 so that $\{t : \varphi_1(t) = \varphi_2(t)\} = A$.

练习 3.3.22: Find independent r.v.'s X, Y , and Z so that Y and Z do not have the same distribution but $X + Y$ and $X + Z$ do.

练习 3.3.23: Show that if X and Y are independent and $X + Y$ and X have the same distribution then $Y = 0$ a.s. .

3.3.5. 矩问题

练习 3.3.24: Let $G(x) = P(|X| < x)$, $\lambda = \sup\{x : G(x) < 1\}$, and $\nu_k = E|X|^k$. Show that $\nu_k^{1/k} \rightarrow \lambda$, so the assumption of Theorem 3.3.26 holds if $\lambda < \infty$.

练习 3.3.25: Suppose $|X|$ has density $Cx^\alpha \exp(-x^\lambda)$ on $(0, \infty)$. Changing variables $y = x^\lambda$, $dx = (1/\lambda)x^{1/\lambda-1} dx$

$$E|X|^n = \int_0^\infty C\lambda^{-1}y^{(n+\alpha)/\lambda} \exp(-y)y^{1/\lambda-1} dy = C\lambda^{-1}\Gamma\left(\frac{n+\alpha+1}{\lambda}\right).$$

Use the identity $\Gamma(x+1) = x\Gamma(x)$ for $x \geq 0$ to conclude that the assumption of Theorem 3.3.26 is satisfied for $\lambda \geq 1$ but not for $\lambda < 1$. This shows the normal ($\lambda = 2$) and gamma ($\lambda = 1$) distributions are determined by their moments.

4. 鞅

4.1. 条件期望

练习 4.1.1: Let $G \in \mathcal{G}$ and show that

$$P(G|A) = \frac{\int_G P(A|\mathcal{G}) dP}{\int_{\Omega} P(A|\mathcal{G}) dP}.$$

When \mathcal{G} is the σ -field generated by a partition, this reduces to the usual Bayes' formula

$$P(G_i|A) = \frac{P(A|G_i)P(G_i)}{\sum_j P(A|G_j)P(G_j)}.$$

证明:

$$\begin{aligned}\int_G P(A|\mathcal{G}) dP &= \int_G E(1_A|\mathcal{G}) dP \\ &= \int_G 1_A dP \\ &= P(A \cap G),\end{aligned}$$

$$\int_{\Omega} P(A|\mathcal{G}) dP = E(E(1_A|\mathcal{G})) = E(1_A) = P(A).$$

那么

$$P(G|A) = \frac{P(A \cap G)}{P(A)}.$$

□

练习 4.1.2: If $a > 0$ then

$$P(|X| \geq a|\mathcal{F}) \leq a^{-2}E(X^2|\mathcal{F}).$$

证明: 对 $\forall A \in \mathcal{G}$ 有

$$\begin{aligned}\frac{1}{a^2} \int_A E(X^2|\mathcal{F}) dP &= \frac{1}{a^2} \int_A X^2 dP \\ &\geq \int_A 1_{\{|X| \geq a\}} dP \\ &= \int_A P(|X| \geq a|\mathcal{F}) dP.\end{aligned}$$

对 $\forall \varepsilon > 0$, 令 $A = \{E(X^2|\mathcal{F})/a^2 - P(|X| \geq a|\mathcal{F}) \leq -\varepsilon\} \in \mathcal{G}$, 那么一定有 $P(A) = 0$ 对 $\forall \varepsilon > 0$ 成立, 那么欲证不等式是 *a.s.* 成立的。 \square

练习 4.1.3:

$$E(XY|\mathcal{G})^2 \leq E(X^2|\mathcal{G})E(Y^2|\mathcal{G}).$$

证明: 对 $\forall \lambda \in \mathbb{R}$ 有

$$E((X + \lambda Y)^2|\mathcal{G}) = E(X^2|\mathcal{G}) + 2\lambda E(XY|\mathcal{G}) + \lambda^2 E(Y^2|\mathcal{G}) \geq 0,$$

将其看作对 λ 的二次函数, 那么有判别式

$$0 \geq \Delta = 4E(XY|\mathcal{G})^2 - 4E(X^2|\mathcal{G})E(Y^2|\mathcal{G})$$

直接得到结论。 \square

练习 4.1.4: If $p, q \in (1, \infty)$ with $1/p + 1/q = 1$, then

$$E(|XY| | \mathcal{G}) \leq E(|X|^p | \mathcal{G})^{1/p} E(|Y|^q | \mathcal{G})^{1/q}.$$

证明: 令 $\tilde{X} = |X|/E(|X|^p|\mathcal{G})^{1/p}$, $\tilde{Y} = |Y|/E(|Y|^q|\mathcal{G})^{1/q}$. 由 Young 不等式知道

$$ab \leq a^p/p + b^q/q$$

对 $\forall a, b \geq 0$ 成立, 那么

$$E(\tilde{X}\tilde{Y}|\mathcal{G}) \leq \frac{E(\tilde{X}^p|\mathcal{G})}{p} + \frac{E(\tilde{Y}^q|\mathcal{G})}{q},$$

而

$$E(\tilde{X}^p|\mathcal{G}) = \frac{E(|X|^p|\mathcal{G})}{E(|X|^p|\mathcal{G})} = 1,$$

$$E(\tilde{Y}^q|\mathcal{G}) = \frac{E(|Y|^q|\mathcal{G})}{E(|Y|^q|\mathcal{G})} = 1$$

那么 $E(\tilde{X}\tilde{Y}|\mathcal{G}) \leq 1/p + 1/q = 1$, 从而证明题目。 \square

练习 4.1.5: Provide an example where

$$E(E(X|\mathcal{F}_1)|\mathcal{F}_2) \neq E(E(X|\mathcal{F}_2)|\mathcal{F}_1).$$

证明: 对 $\Omega = \{1, 2, 3\}$, $\mathcal{F}_1 = \sigma(\{1, 2\}, \{3\})$, $\mathcal{F}_2 = \sigma(\{1\}, \{2, 3\})$ 和

$$P(\{1\}) = \frac{1}{7}, P(\{2\}) = \frac{2}{7}, P(\{3\}) = \frac{4}{7},$$

考虑

$$P(E(E(X|\mathcal{F}_1)|\mathcal{F}_2))(1, 2, 3) = \left(\frac{3}{14}, \frac{11}{28}, \frac{11}{28}\right),$$

$$P(E(E(X|\mathcal{F}_2)|\mathcal{F}_1))(1, 2, 3) = \left(\frac{2}{7}, \frac{2}{7}, \frac{3}{7}\right).$$

□

练习 4.1.6: If $\mathcal{G} \subset \mathcal{F}$ and $EX^2 < \infty$, then

$$E(\{X - E(X|\mathcal{F})\}^2) + E(\{E(X|\mathcal{F}) - E(X|\mathcal{G})\}^2) = E(\{X - E(X|\mathcal{G})\}^2).$$

Dropping the second term gives an inequality reflecting that larger subspaces yield closer projections (or statistiscrly, more information reduces mean square error).

证明: 令 $Y = X - E(X|\mathcal{G})$, 那么

$$\begin{aligned} X - E(X|\mathcal{F}) &= Y + E(X|\mathcal{G}) - E(Y|\mathcal{F}) - E(E(X|\mathcal{G})|\mathcal{F}) \\ &= Y - E(Y|\mathcal{F}), \end{aligned}$$

$$\begin{aligned} E(X|\mathcal{F}) - E(X|\mathcal{G}) &= E(Y|\mathcal{F}) - E(X|\mathcal{G}) - E(Y|\mathcal{G}) + E(E(X|\mathcal{G})|\mathcal{G}) \\ &= E(Y|\mathcal{F}) - E(Y|\mathcal{G}) \\ &= E(Y|\mathcal{F}) - E(X - E(X|\mathcal{G})|\mathcal{G}) \\ &= E(Y|\mathcal{F}). \end{aligned}$$

那么只要证明

$$E((Y - E(Y|\mathcal{F}))^2) + E(E(Y|\mathcal{F})^2) = E(Y^2)$$

即可。恒等变换只要证明

$$E(YE(Y|\mathcal{F})) = E(E(Y|\mathcal{F})^2)$$

即可, 这是因为

$$E(YE(Y|\mathcal{F})) = E(E(YE(Y|\mathcal{F})|\mathcal{F})) = E(E(Y|\mathcal{F})^2).$$

□

练习 4.1.7: Let $\mathcal{G} = \{\emptyset, \Omega\}$. Define $\text{var}(X|\mathcal{F}) = E(X^2|\mathcal{F}) - E(X|\mathcal{F})^2$. Show that

$$\text{var}(X) = E(\text{var}(X|\mathcal{F})) + \text{var}(E(X|\mathcal{F})).$$

证明:

$$\begin{aligned}
\text{RHS} &= E\left(E(X^2|\mathcal{F}) - E(X|\mathcal{F})^2\right) + E\left(E(X|\mathcal{F})^2\right) - (E(E(X|\mathcal{F})))^2 \\
&= EX^2 - (EX)^2 \\
&= \text{LHS}.
\end{aligned}$$

□

练习 4.1.8: Let Y_1, Y_2, \dots be i.i.d. with mean μ and variance σ^2 , N an independent positive integer-valued r.v. with $EN^2 < \infty$, and $X = Y_1 + \dots + Y_N$. Show that

$$\text{var}(X) = \sigma^2 EN + \mu^2 \text{var}(N).$$

Interpretation: Consider special cases where N or Y_i is constant.

证明：利用条件期望推导全概率公式，然后进行计算。

$$\begin{aligned}
EX^2 &= E(E(X^2|N)) \\
&= \int_{\{N \in \mathbb{N}\}} E(X^2|N) dP \\
&= \int_{\{N \in \mathbb{N}\}} E(X^2) dP \\
&= \sum_{n \in \mathbb{N}} n(\sigma^2 + \mu^2) P(N = n) \\
&= (\sigma^2 + \mu^2) EN.
\end{aligned}$$

另一方面，

$$\begin{aligned}
EX &= E(E(X|N)) \\
&= \int_{\{N \in \mathbb{N}\}} E(X|N) dP \\
&= \int_{\{N \in \mathbb{N}\}} E(X) dP \\
&= \sum_{n \in \mathbb{N}} n\mu P(N = n) \\
&= \mu EN.
\end{aligned}$$

那么

$$\begin{aligned}
\text{var}(X) &= EX^2 - (EX)^2 \\
&= (\sigma^2 + \mu^2) EN - \mu^2 (EN)^2 \\
&= \sigma^2 EN + \mu^2 \text{var}(N).
\end{aligned}$$

□

练习 4.1.9: Show that if X and Y are random variables with $E(Y|\mathcal{G}) = X$ and $EY^2 = EX^2 < \infty$, then $X = Y$ almost surely.

证明：由

$$E(E(Y|\mathcal{G})^2) + E(Y - E(Y|\mathcal{G}))^2 = EY^2 = EX^2 = E(E(Y|\mathcal{G})^2)$$

直接得到

$$E(Y - X)^2 = E(Y - E(Y|\mathcal{G}))^2 = 0$$

然后结论是显然的。

□

练习 4.1.10: Under the assumption $E|Y| < \infty$, prove that if $E(Y|\mathcal{G})$ has the same distribution as Y , then $E(Y|\mathcal{G}) = Y$ a.s. .

Hint: Show $\text{sgn}(X) = \text{sgn}(E(X|\mathcal{G}))$ a.s. for $X = Y - c$.

证明：TODO

□

4.2. 鞅、几乎必然收敛

练习 4.2.1: Suppose X_n is a martingale w.r.t. \mathcal{G}_n and let $\mathcal{F}_n = \sigma(X_1, \dots, X_n)$. Then $\mathcal{G}_n \supset \mathcal{F}_n$, and X_n is a martingale w.r.t. \mathcal{F}_n .

练习 4.2.2: Give an example of a submartingale X_n such that X_n^2 is a supermartingale.
Hint: X_n does not have to be random.

练习 4.2.3: Generalize part 1 of Theorem 4.2.7 by showing that if X_n and Y_n are submartingales w.r.t. \mathcal{F}_n , then $X_n \vee Y_n$ is also a submartingale.

练习 4.2.4: Let $X_n, n \geq 0$, be a submartingale with $\sup X_n < \infty$. Let $\xi_n = X_n - X_{n-1}$ and suppose $E(\sup \xi_n^+) < \infty$. Show that X_n converges almost surely.

练习 4.2.5: Give an example of a martingale X_n with $X_n \rightarrow -\infty$ a.s. *Hint:* Let $X_n = \xi_1 + \dots + \xi_n$, where ξ_i are independent (but not identically distributed) with $E\xi_i = 0$.

练习 4.2.6: Let Y_1, Y_2, \dots be nonnegative i.i.d. random variables with $EY_m = 1$ and $P(Y_m = 1) < 1$. By Example 4.2.3, $X_n = \prod_{m \leq n} Y_m$ defines a martingale.

1. Use Theorem 4.2.12 and an argument by contradiction to show $X_n \rightarrow 0$ a.s.
2. Use the strong law of large numbers to conclude $(1/n) \log X_n \rightarrow c < 0$.

练习 4.2.7: Suppose $y_n > -1$ for all n and $\sum |y_n| < \infty$. Show that $\prod_{m=1}^{\infty} (1 + y_m)$ converges to a finite limit.

练习 4.2.8: Let X_n and Y_n be positive integrable processes adapted to \mathcal{F}_n . Suppose

$$E(X_{n+1} | \mathcal{F}_n) \leq (1 + Y_n) X_n$$

with $\sum Y_n < \infty$ a.s. Prove that X_n converges a.s. to a finite limit by constructing a related supermartingale and applying Theorem 4.2.12.

练习 4.2.9: Suppose X_n^1 and X_n^2 are supermartingales w.r.t. \mathcal{F}_n , and N is a stopping time such that $X_N^1 \geq X_N^2$. Define

$$Y_n = X_n^1 \mathbf{1}_{\{N > n\}} + X_n^2 \mathbf{1}_{\{N \leq n\}},$$

where $\mathbf{1}$ is the indicator function. Show that Y_n is a supermartingale.

练习 4.2.10: Dubins' inequality. For every positive supermartingale X_n , $n \geq 0$, the number of upcrossings U of $[a, b]$ satisfies

$$P(U \geq k) \leq \left(\frac{a}{b}\right)^k E \min(X_0/a, 1).$$

To prove this, we let $N_0 = -1$ and for $j \geq 1$ let

$$\begin{aligned} N_{2j-1} &= \inf\{m > N_{2j-2} : X_m \leq a\} \\ N_{2j} &= \inf\{m > N_{2j-1} : X_m \geq b\}. \end{aligned}$$

Let $Y_n = 1$ for $0 \leq n < N_1$ and for $j \geq 1$

$$Y_n = \begin{cases} (b/a)^{j-1} (X_n/a) & \text{for } N_{2j-1} \leq n < N_{2j} \\ (b/a)^j & \text{for } N_{2j} \leq n < N_{2j+1} \end{cases}$$

1. Use the switching principle in the previous exercise and induction to show that $Z_n^j = Y_{n \wedge N_j}$ is a supermartingale.

2. Use $EY_{n \wedge N_{2k}} \leq EY_0$ and let $n \rightarrow \infty$ to get Dubins' inequality.