

第 1 讲 随机过程

- 课程名：随机过程
- 任课老师：涂思铭
- 课程主页：<http://39.107.124.3/teaching/stochastic2025/>
- 教材：何书元. 随机过程, 北京大学出版社
- 评分规则：平时作业 40% + 期末考试 60%

参考书籍

- 林元烈编著. 应用随机过程, 清华大学出版社, 2002.
- 张波, 商豪, 邓军编著. 应用随机过程, 中国人民大学出版社, 2023.
- 李增沪, 张梅, 何辉编著. 概率论 (下册), 高等教育出版社, 2024.
- S. Ross. Introduction to Probability Models, thirteenth edition, Academic Press. (有中译本, 人民邮电出版社)
- Richard Durrett. Essentials of Stochastic Processes, third edition, Springer. (有中译本, 机械工业出版社)
- Robert G. Gallager. Stochastic Processes: Theory for Applications, Cambridge University Press.

- 在概率论中, 我们研究了随机变量, n 维随机变量. 在极限定理中, 涉及到了无穷多个随机变量, 但局限在它们之间是相互独立的情形.
- 将上述情形加以推广, 即研究一族无穷多个, 相互相关的随机变量, 这就是随机过程.
- 注意, 所有的随机变量都定义在相同的概率空间 $(\Omega, \mathcal{F}, P)$ 上.

课程安排

- 随机过程简介 1 周
- Poisson 过程 3 周
- 更新过程 3 周
- 离散时间 Markov 链 4 周
- 连续时间 Markov 链 3 周
- Brown 运动 2 周
- 复习 1 周

随机过程

定义 1.1

一个随机过程是指一族定义在同一个样本空间 Ω 上的随机变量 $X_T := \{X(t) : t \in T\}$, 其中 t 是参数, 它属于某个指标集 T , T 称为该随机过程的指标集或参数集.

参数 $t \in T$ 一般表示时间. 通常 T 为 $\mathbb{N} \cup \{0\}$, $\mathbb{Z}$ 或区间 $[a, b]$ (其中 a 可以取 $-\infty$, b 可以取 $-\infty$). 当 T 取可数集时, 通常称 $\{X(t) : t \in T\}$ 是一个随机序列.

注记 1.2

有时, 从另一个角度来看随机过程是很有益的, 即随机过程 $\{X(t, \omega) : t \in T, \omega \in \Omega\}$ 可以看成是定义在 $T \times \Omega$ 上的二元函数. 对于固定的样本点 $\omega \in \Omega$, $X(t, \omega)$ 就是定义在 T 上的一个函数, 称为一条样本路径或轨道. 而对于固定的时刻 $t \in T$, $X(t) = X(t, \omega)$ 是概率空间 $(\Omega, \mathcal{F}, P)$ 上的一个随机变量.

例 1.3

设一质点在时刻 $t = 0$ 时处于位置 a (整数), 以后每隔单位时间, 分别以概率 p 及概率 $q = 1 - p$ 向正或负的方向随机移动一个单位, 记 X_n 为质点在时刻 $t = n$ 的位置. 固定 n , X_n 是随机变量. 考虑不同的 n 时, $\{X_n, n \geq 0\}$ 是一随机序列.

例 1.4

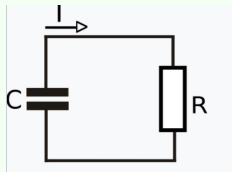
考虑某“服务站”在 $[0, t]$ 内来的“顾客”数, 记为 $N(t)$, 固定 t , $N(t)$ 就是一随机变量. 因此 $\{N(t), t \geq 0\}$ 是一个随机过程. 这里的“顾客”可以是电话的“呼叫”, 通信设备中的“信号”, 放射性物质衰变的“粒子”等.

例 1.5

考虑某输入输出系统，例如最简单的 R-C 电路. 设输入端有一个干扰信号电压，记为 $\xi(t)$. 记 $Q(t)$ 为 t 时电路的电量，则它满足

$$R \frac{dQ(t)}{dt} + \frac{1}{C} Q(t) = \xi(t).$$

由于 $\{\xi(t), t \in T\}$ 是一个随机过程，容易理解 $\{Q(t), t \in T\}$ 也是一个随机过程. 上式是一个最简单的随机微分方程.



例 1.6

我们可以用 $X(t)$ 来表示股票在 t 时刻的价格, 则 $\{X(t), t \geq 0\}$ 就可以描述股票价格随时间推移而产生的随机波动现象。

研究随机现象主要是研究其统计规律性. 对于一个或有限个随机变量来说, 掌握了分布函数就能完全了解随机变量. 对于随机过程 $\{X(t): t \in T\}$, 为了描述它的统计特性, 自然要知道对于每个 t , $X(t)$ 的分布函数 $F(t, x) = P(X(t) \leq x)$, 以及它们在任意 n 个时间点的联合分布. 对于有限个时刻 $t_1, t_2, \dots, t_n \in T$, 我们还需定义随机过程的 n 维分布:

$$F_{t_1, \dots, t_n}(x_1, \dots, x_n) = P(X(t_1) \leq x_1, X(t_2) \leq x_2, \dots, X(t_n) \leq x_n)$$

随机过程的所有有限维分布全体

$$\{F_{t_1, t_2, \dots, t_n}(x_1, x_2, \dots, x_n) : t_1, t_2, \dots, t_n \in T, n \geq 1\}$$

称为随机过程 $\{X(t), t \in T\}$ 的有限维分布族, 或称为这个随机过程的概率分布. 具有相同的有限维分布的随机过程称为是同分布的. 同分布的随机过程有相同的统计性质.

(1) **对称性:** 对 $1, 2, \dots, n$ 的任意排列 $j_1, j_2, \dots, j_n$, 有

$$\begin{aligned} & F_{t_{j_1}, t_{j_2}, \dots, t_{j_n}}(x_{j_1}, x_{j_2}, \dots, x_{j_n}) \\ &= P(X(t_{j_1}) \leq x_{j_1}, X(t_{j_2}) \leq x_{j_2}, \dots, X(t_{j_n}) \leq x_{j_n}) \\ &= P(X(t_1) \leq x_1, X(t_2) \leq x_2, \dots, X(t_n) \leq x_n) \\ &= F_{t_1, t_2, \dots, t_n}(x_1, x_2, \dots, x_n). \end{aligned}$$

(2) **相容性:** 对于 $m < n$, 有

$$F_{t_1, t_2, \dots, t_m, t_{m+1}, \dots, t_n}(x_1, x_2, \dots, x_m, \infty, \dots, \infty) \\ = F_{t_1, t_2, \dots, t_m}(x_1, x_2, \dots, x_m).$$

Kolmogorov 存在定理

定理 1.7

设分布函数族 $\{F_{t_1, t_2, \dots, t_n}(x_1, x_2, \dots, x_n), t_1, t_2, \dots, t_n \in T, n \geq 1\}$ 满足上述的对称性和相容性, 则必存在一个随机过程 $\{X(t), t \in T\}$, 使

$$\{F_{t_1, t_2, \dots, t_n}(x_1, x_2, \dots, x_n) : t_1, t_2, \dots, t_n \in T, n \geq 1\}$$

恰好是 $\{X(t), t \in T\}$ 的概率分布.

Kolmogorov 存在定理说明, 随机过程的有限维分布族是随机过程概率特征的完整描述.

定义 1.8

随机过程 $X(t)$ 和 $Y(t)$ 称为相互独立的, 如果对任意 $t_1, t_2, \dots, t_n$ 以及 $s_1, \dots, s_m$, $\mathbf{X} = (X(t_1), \dots, X(t_n))$ 和 $\mathbf{Y} = (Y(s_1), \dots, Y(s_m))$ 是独立的.

定义 1.9

- (1) 称 $X(t)$ 的期望 $\mu_X(t) = E[X(t)]$ 为过程的均值函数 (如果存在的话).
- (2) 如果对任意 $t \in T$, $E[X^2(t)]$ 存在, 则存在随机过程 $\{X(t) : t \in T\}$ 为二阶矩过程. 此时, 称函数
 $\gamma(t_1, t_2) = E[(X(t_1) - \mu_X(t_1))(X(t_2) - \mu_X(t_2))]$ ($t_1, t_2 \in T$) 为过程的协方差函数; 称 $\gamma(t, t) = \text{Var}[X(t)]$ 为过程的方差函数; 称
 $R_X(s, t) = E[X(s)X(t)]$ ($s, t \in T$) 为过程的自相关函数.

由 Schwartz 不等式知, 二阶矩过程的协方差函数和自相关函数存在, 且有 $\gamma_X(s, t) = R_X(s, t) - \mu_X(s)\mu_X(t)$.

例 1.10

设 $X(t) = X_0 + tV$ ($a \leq t \leq b$), 其中 X_0 和 V 是相互独立且均服从标准正态分布的随机变量. 于是 $X(t)$ 也服从正态分布, 并且 $(X(t_1), \dots, X(t_n))$ 服从 n 维正态分布. 所以只要知道它的一阶矩和二阶矩就完全确定了它的分布. 而它的一阶矩和二阶矩是容易求得的:

$$\begin{aligned}\mu_X(t) &= E[X(t)] = E[X_0 + tV] = E[X_0] + tE[V] = 0, \\ \gamma(t_1, t_2) &= E[X(t_1)X(t_2)] = E[(X_0 + t_1V)(X_0 + t_2V)] \\ &= E[X_0^2] + t_1t_2E[V^2] = 1 + t_1t_2.\end{aligned}$$

例 1.11

考察一个随机相位正弦波，波形为

$$X(t) = A \cos(\theta t + X), t \geq 0,$$

其中 X 服从 $[-\pi, \pi]$ 上的均匀分布，求：

- (1) X_t 的密度函数；
- (2) 随机过程 $\{X_t : t \geq 0\}$ 的均值函数和协方差函数.

(1) 设 $|x| \leq A$, 则

于是我们有

对其求导数我们可以得到 X_t 的密度函数为

◀ ◻ ▶ ◀ ◻ ▶ ◀ ≡ ▶ ◀ ≡ ▶ ≡ ▶ ↺ 🔍 ↻

解.

(2) 由 (1) 我们知 $\{X_t : t \geq 0\}$ 的均值函数为

$$\mu_X(t) = E[X(t)] = \int_{-A}^A x \frac{1}{\pi} \frac{dx}{\sqrt{A^2 - x^2}} = -\frac{1}{\pi} \sqrt{A^2 - x^2} \Big|_{-A}^A = 0.$$

我们也可以应用随机变量函数的数学期望公式来求 $\{X_t : t \geq 0\}$ 的均值函数:

$$\begin{aligned}\mu_X(t) &= E[X(t)] = E[A \cos(\theta t + X)] \\ &= A \int_{-\pi}^{\pi} \cos(\theta t + x) \frac{dx}{2\pi} \\ &= \frac{A}{2\pi} \frac{\sin(\theta t + x)}{\theta} \Big|_{-\pi}^{\pi} = 0.\end{aligned}$$

解.

类似地, 求得 $\{X_t : t \geq 0\}$ 的协方差函数

$$\begin{aligned}
 \gamma(s, t) &= R_X(s, t) = E[X_s X_t] \\
 &= A^2 E[\cos(\theta s + X) \cos(\theta t + X)] \\
 &= A^2 \int_{-\pi}^{\pi} \cos(\theta s + x) \cos(\theta t + x) \frac{dx}{2\pi} \\
 &= \frac{A^2}{2\pi} \int_{-\pi}^{\pi} \frac{\cos[\theta(t - s)] + \cos[\theta(t + s) + 2x]}{2} dx \\
 &= \frac{A^2}{2} \cos[\theta(t - s)].
 \end{aligned}$$



- 所谓平稳过程，粗略地说，是指它的概率特征不随时间推移而变化的随机过程.
- 它在电子技术、自动控制、轮船制造、管理、经济计量等领域都有广泛的应用.

定义 2.1

如果随机过程 $\{X(t), t \in T\}$ 满足对任意 $t_1, t_2, \dots, t_n$ 和任意 h ($t_i + h \in T, i = 1, 2, \dots, n$) 有 $(X(t_1 + h), X(t_2 + h), \dots, X(t_n + h))$ 与 $(X(t_1), X(t_2), \dots, X(t_n))$ 具有相同的联合分布, 则称 $\{X(t), t \in T\}$ 为 (严) 平稳的. 严平稳的随机序列也称为严平稳序列, 或简称为平稳序列.

定义 2.2

如果随机过程 $\{X(t), t \in T\}$ 的所有二阶矩都存在, 并且 $E[X(t)] = \mu$, 协方差函数 $\gamma(t, s)$ 只与时间差 $t - s$ 有关, 则称 $\{X(t), t \in T\}$ 为宽平稳过程 或 二阶平稳过程. 宽平稳的随机序列也称为宽平稳序列.

命题 2.3

设 $\{X(t), t \in T\}$ 是一个二阶矩过程. 它是宽平稳的, 当且仅当 $E[X(s)]$ 与 $E[X(s)X(s+t)]$ 都不依赖于 s .

证明.

如果 $E[X(s)]$ 与 $E[X(s)X(s+t)]$ 都不依赖于 s , 则

- (1) 均值函数 $E[X(s)]$ 不依赖于 s 知其为一个常数 μ .
- (2) 由 $E[X(s)X(s+t)]$ 不依赖于 s , 我们有 $\gamma(s, s+t) = E[X(s)X(s+t)] - \mu^2$ 仅与时间间隔 t 有关,

所以我们有 $\{X(t), t \in T\}$ 是宽平稳的.

由上述推导我们也不难看出如果 $\{X(t), t \in T\}$ 是宽平稳的, 则 $E[X(s)]$ 与 $E[X(s)X(s+t)]$ 都不依赖于 s . □

例 2.4

设 $\{\epsilon_n, n = 0, \pm 1, \dots\}$ 为一列两两互不相关的随机变量序列, 满足 $E[X_n] = 0$ ($n = 0, 1, \dots$), 且

$$E[X_m X_n] = \begin{cases} 0, & m \neq n \\ \sigma^2, & m = n. \end{cases}$$

则 $\{X_n, n = 0, 1, \dots\}$ 为宽平稳序列. 这是因为协方差函数 $\text{Cov}(X_n, X_m) = E[X_n X_m]$ 只与 $m - n$ 有关.

设 $\{\epsilon_n, n = 0, \pm 1, \dots\}$ 为一列两两互不相关的具有相同均值 μ 和相同方差 σ^2 的随机变量序列, $a_1, a_2, \dots, a_k$ 为任意 k 个实数. 考虑如下定义的序列:

$$X_n = a_1\epsilon_n + a_2\epsilon_{n-1} + \cdots + a_k\epsilon_{n-k+1}, \quad n = 0, \pm 1, \pm 2, \cdots.$$

容易看出

$$E[X_n] = \mu(a_1 + a_2 + \cdots + a_k).$$

例: (续)

令 $\xi_j = \epsilon_j - \mu$, 则由于 $\{\epsilon_j\}$ 两两不相关, 我们知协方差函数

$$\begin{aligned} & \gamma(n, n + \tau) \\ &= E[(X_n - \mu(a_1 + a_2 + \cdots + a_k))(X_{n+\tau} - \mu(a_1 + a_2 + \cdots + a_k))] \\ &= E[(a_1\epsilon_n + a_2\epsilon_{n-1} + \cdots + a_k\epsilon_{n-k+1}) \\ & \quad (a_1\epsilon_{n+\tau} + a_2\epsilon_{n+\tau-1} + \cdots + a_k\epsilon_{n+\tau-k+1})] \\ &= \begin{cases} \sigma^2(a_ka_{k-\tau} + a_{k-1}a_{k-\tau-1} + \cdots + a_{\tau+1}a_1), & 0 \leq \tau \leq k-1 \\ 0, & \tau \geq k \end{cases} \end{aligned}$$

于是协方差函数仅与时间间隔 τ 有关, 所以我们有 $\{X_n, n = 0, \pm 1, \dots\}$ 为宽平稳序列.

定义 3.1

设 $\{X(t), t \in T\}$ 是一个随机过程. 如果对任意 $t_1, t_2, \dots, t_n \in T, t_1 < t_2 < \dots < t_n$, 我们有随机变量 $X(t_2) - X(t_1), X(t_3) - X(t_2), \dots, X(t_n) - X(t_{n-1})$ 是相互独立的, 则称 $\{X(t), t \in T\}$ 是一个**独立增量过程**.

定义 3.2

设 $\{X(t), t \in T\}$ 是一个随机过程. 如果对任意 t_1, t_2, h , 其中 $t_1, t_2, t_1 + h, t_2 + h \in T$, 有 $X(t_1 + h) - X(t_1)$ 与 $X(t_2 + h) - X(t_2)$ 具有相同的分布, 则称 $\{X(t), t \in T\}$ 是一个**平稳增量过程**.

兼有独立增量和平稳增量的过程称为平稳独立增量过程.

回顾：特征函数

定义 3.3

设随机变量的分布函数为 $F(x)$, 称

$$g(t) = E[e^{itX}] = \int_{-\infty}^{\infty} e^{itx} dF(x), \quad -\infty < t < \infty$$

为 X 的**特征函数**.

特征函数 $g(t)$ 是实变量 t 的复值函数. 注意到 $|e^{itx}| = 1$, 我们知随机变量的特征函数一定存在.

- 若 X 是连续型随机变量, 概率密度为 $f(x)$, 则 X 的特征函数为

$$p_x = P(X = x_k), \quad k = 1, 2, \dots,$$

则 X 的特征函数为

$$g(t) = \sum_{k=1}^{\infty} e^{itx_k} p_k.$$

$$g(t) = \int_{-\infty}^{\infty} e^{itx} f(x) dx.$$

(1) $g(0) = 1$, $|g(t)| \leq 1$, $g(-t) = \overline{g(t)}$;

(2) $g(t)$ 在 $(-\infty, \infty)$ 上一致连续;

(3) 若随机变量 X 的 n 阶矩 $E[X^n]$ 存在, 则 X 的特征函数 $g(t)$ 可微分 n 次, 且当 $k \leq n$ 时, 有 $g^{(k)}(0) = i^k EX^k$;

(4) $g(t)$ 是非负定函数, 即对任意正整数 n 以及任意实数 $t_1, t_2, \dots, t_n$ 和复数 $z_1, z_2, \dots, z_n$, 有

$$\sum_{k,i=1}^n g(t_k - t_i) z_k z_i \geq 0;$$

- (5) 若 $X_1, X_2, \dots, X_n$ 是相互独立的随机变量, 则 $X = X_1 + X_2 + \dots + X_n$ 的特征函数为

$$g(t) = g_1(t)g_2(t) \cdots g_n(t),$$

其中 $g_i(t)$ 是随机变量 X_i 的特征函数, $i = 1, 2, \dots, n$;

- (6) 随机变量的分布函数由其特征函数唯一确定.

定义 3.4

设 $\mathbf{X} = (X_1, X_2, \dots, X_n)$ 是一个 n 维随机向量, $\mathbf{t} = (t_1, t_2, \dots, t_n) \in \mathbb{R}^n$, 则称

$$g(\mathbf{t}) = g(t_1, \dots, t_n) = E[e^{i\mathbf{t}\mathbf{X}}] = E[\exp\{i \sum_{k=1}^n t_k X_k\}]$$

为 X 的特征函数.

定理 3.5

设 $\{X(t), t \geq 0\}$ 是一个独立增量过程, 满足 $X(0) = 0$. 则 $X(t)$ 具有平稳增量的充分必要条件是: 其特征函数具有可乘性, 即

$$\phi_{X(t+s)}(a) = \phi_{X(t)}(a)\phi_{X(s)}(a).$$

因为 $X(t)$ 具有平稳增量, 所以 $X(t+s) - X(t)$ 与 $X(s) = X(s) - X(0)$ 具有相同的分布. 从而由独立增量性我们有

$$\begin{aligned}\phi_{X(t+s)}(a) &= \phi_{X(t)}(a)\phi_{X(t+s)-X(t)}(a) \\ &= \phi_{X(t)}(a)\phi_{X(s)}(a).\end{aligned}$$

必要性得证.

下证充分性，由独立增量性我们有

$$\begin{aligned}\phi_{X(t)}(a)\phi_{X(s)}(a) &= \phi_{X(t+s)}(a) \\ &= \phi_{X(t)}(a)\phi_{X(t+s)-X(t)}(a),\end{aligned}$$

所以 $\phi_{X(t)}(a) = \phi_{X(t+s)-X(t)}(a)$. 于是 $X(t+s) - X(t)$ 与 $X(s) = X(s) - X(0)$ 具有相同的分布, 即 $\{X(t)\}$ 具有平稳增量. $\square$

指数分布

定义 4.1

称连续型随机变量 X 服从参数 (或均值) $\lambda > 0$ 的指数分布, 如果它的概率密度函数为

$$f(x) = \lambda e^{-\lambda x}, \quad x > 0.$$

我们记 $X \sim \mathcal{E}(\lambda)$.

若 X 服从参数 $\lambda > 0$ 的指数分布, 我们不难求得其累积分布函数为

$$F(x) = 1 - e^{-\lambda x}, \quad x > 0.$$

我们来计算一下 X 的期望和方差. 我们有

$$\begin{aligned} E[X^n] &= \int_0^\infty x^n \lambda e^{-\lambda x} dx \\ &= -x^n \lambda e^{-\lambda x} \Big|_0^\infty + \int_0^\infty e^{-\lambda x} n x^{n-1} dx \\ &= 0 + \frac{n}{\lambda} \int_0^\infty \lambda e^{-\lambda x} x^{n-1} dx \\ &= \frac{n}{\lambda} E[X^{n-1}]. \end{aligned}$$

取 $n = 1$ 和 $n = 2$ 我们有

$$E[X] = \frac{1}{\lambda},$$
$$E[X^2] = \frac{2}{\lambda} E[X] = \frac{2}{\lambda^2}.$$

于是

$$\text{Var}(X) = \frac{2}{\lambda^2} - \left(\frac{1}{\lambda}\right)^2 = \frac{1}{\lambda^2}.$$

命题 4.2

当 $X_1, \dots, X_n$ 相互独立, $X_i \sim \mathcal{E}(\lambda_i)$ 时, 有

$$P(X_1 \leq X_2) = \frac{\lambda_1}{\lambda_1 + \lambda_2},$$
$$\min\{X_1, X_2, \dots, X_n\} \sim \mathcal{E}(\lambda_1 + \lambda_2 + \dots + \lambda_n).$$

定义 4.3

我们称一个连续型随机变量的分布具有 无记忆性, 如果对任意 $s, t \geq 0$,

$$P(X \geq s + t | X \geq s) = P(X \geq t).$$

利用条件概率的定义，我们来验证指数分布具有无记忆性. 设 $X \sim \mathcal{E}(\lambda)$. 则

$$\begin{aligned} P(X \geq s+t | X \geq s) &= \frac{P(X \geq s+t)}{P(X \geq s)} \\ &= \frac{e^{-\lambda(s+t)}}{e^{-\lambda s}} \\ &= e^{-\lambda t} = P(X \geq t). \end{aligned}$$

定理 4.4

如果取正值的连续型随机变量 X 具有无记忆性, 则 X 具有指数分布.

证明.

- 设 X 是一个取正值的连续型随机变量, 具有无记忆性.
- 设 F 是 X 的累积分布函数, 记 $G(x) = 1 - F(x)$. 我们要证明: 存在 $\lambda > 0$ 使得 $G(x) = e^{-\lambda x}$.
- 注意到由无记忆性, 我们有对任意 $s, t \geq 0$,

$$G(s+t) = G(s)G(t).$$

- 由 G 的连续性, 我们知存在 $a \geq 0$ 使得 $G(x) = a^x$.
- 由 F 是一个分布函数知 $0 < a < 1$, 从而存在 $\lambda > 0$, 使得

$$G(x) = e^{-\lambda x},$$

从而 X 服从指数分布.



定理 4.5

令 $X_1, X_2, \dots$ 是独立的参数为 λ 的指数随机变量, 那么和 $S_n = X_1 + \dots + X_n$ 服从 $\Gamma(n, \lambda)$ 分布, 也就是说, S_n 有密度函数

$$g_n(s) = \frac{\lambda^n}{\Gamma(n)} s^{n-1} e^{-\lambda s} = \frac{\lambda^n}{(n-1)!} s^{n-1} e^{-\lambda s}, \quad s \geq 0.$$

证明.

对 n 应用归纳法证明. 当 $n = 1$ 时, S_1 服从参数为 λ 的指数分布, 其密度函数即为 $g_1(s)$. 假定公式在 n 时成立, 注意到 $S_{n+1} = S_n + X_{n+1}$ 且 S_n 与 X_{n+1} 独立. 所以我们有 S_{n+1} 的概率密度函数为

$$\begin{aligned} f_{S_{n+1}}(s) &= \int_0^s f_{S_n}(t) f_{X_{n+1}}(s-t) dt \\ &= \int_0^s \lambda e^{-\lambda t} \frac{(\lambda t)^{n-1}}{(n-1)!} \cdot \lambda e^{-\lambda(s-t)} dt \\ &= e^{-\lambda s} \lambda^n \int_0^s \frac{t^{n-1}}{(n-1)!} dt = \lambda e^{-\lambda s} \frac{\lambda^n s^n}{n!}, \end{aligned}$$

即为 $g_{n+1}(s)$.



命题 4.6

设随机变量 $T_1, T_2, \dots, T_n$ 相互独立, $T_k \sim \mathcal{E}(\lambda_k)$. 我们有

$$P(T_i = \min(T_1, \dots, T_n)) = \frac{\lambda_i}{\lambda_1 + \dots + \lambda_n}.$$

令 $S = T_i$, $U := \min\{T_j | j \neq i\}$. 我们知 $U \sim \mathcal{E}((\lambda_1 + \cdots + \lambda_n) - \lambda_i)$. 因此我们有

$$\begin{aligned} &= \int_0^\infty f_S(s)P(U > s)ds \\ &= \int_0^\infty \lambda_i e^{-\lambda_i s} e^{-((\lambda_1 + \dots + \lambda_n) - \lambda_i)s} ds \\ &= \frac{\lambda_i}{\lambda_1 + \dots + \lambda_n}. \end{aligned}$$

5

命题 4.7

设随机变量 $T_1, T_2, \dots$ 相互独立, $T_k \sim \mathcal{E}(\lambda_k)$. 用 I 表示最小的随机变量 T_i 的 (随机) 下标. 则 I 和 $V = \min\{T_1, \dots, T_n\}$ 是相互独立的.

证明.

首先, 我们有

$$P(V \leq v) = 1 - e^{-\lambda v}, \quad \lambda := \lambda_1 + \cdots + \lambda_n.$$

记 $\lambda' = \lambda - \lambda_i$, $U := \min\{T_j | j \neq i\}$. 注意到 U 与 T_i 独立, 且 $U \sim \mathcal{E}(\lambda')$. 注意到

$$\begin{aligned} P(V \leq v, I = i) &= P(T_i \leq v, U \geq T_i) \\ &= \int_0^v P(U \geq t_i | T_i = t_i) \lambda_i e^{-\lambda_i t_i} dt_i \\ &= \int_0^v P(U \geq t_i) \lambda_i e^{-\lambda_i t_i} dt_i \\ &= \int_0^v e^{-\lambda' t_i} \lambda_i e^{-\lambda_i t_i} dt_i = \frac{\lambda_i}{\lambda} (1 - e^{-\lambda v}). \end{aligned}$$

于是 $P(V \leq v | I = i) = 1 - e^{-\lambda v} = P(V \leq v)$, 从而 V 和 I 相互独立. □

命题 4.8

设随机变量 $T_1, T_2, \dots$ 相互独立, $T_k \sim \mathcal{E}(\lambda_k)$. 定义 $W = \sum_{k=1}^{\infty} T_k$, $M = \min_k \{T_k\}$, $\lambda_0 = \sum_{k=1}^{\infty} \lambda_k$, 则

- (1) $P(W = \infty) = 1$ 的充分必要条件是 $\sum_{k=1}^{\infty} E[T_k] = \infty$;
- (2) $P(W = \infty)$ 只能取值 1 或 0;
- (3) $\lambda_0 < \infty$ 时, $M \sim \mathcal{E}(\lambda_0)$;
- (4) $\lambda_0 = \infty$ 时, $M = 0$ a.s..

证明.

如果 $P(W = \infty) > 0$, 利用单调收敛定理我们有

$$\begin{aligned}\sum_{k=1}^{\infty} E[T_k] &= E\left[\sum_{k=1}^{\infty} T_k\right] = E[W] \\ &= \int_0^{\infty} P(W > s) ds \geq \int_0^{\infty} P(W = \infty) ds = \infty.\end{aligned}$$

下设 $\sum_{k=1}^{\infty} E[T_k] = \infty$. 我们有

$$\begin{aligned}E[\exp(-W)] &= E\left[\exp\left(-\sum_{k=1}^{\infty} T_k\right)\right] = \prod_{k=1}^{\infty} E[\exp(-T_k)] \\ &= \prod_{k=1}^{\infty} \int_0^{\infty} \lambda_k e^{-\lambda_k s} ds = \prod_{k=1}^{\infty} \frac{\lambda_k}{1 + \lambda_k} \\ &= \prod_{k=1}^{\infty} \frac{1}{1 + 1/\lambda_k} = \left[\prod_{k=1}^{\infty} \left(1 + \frac{1}{\lambda_k}\right)\right]^{-1}.\end{aligned}$$

注意到 $E[T_k] = 1/\lambda_k$, 我们有

$$\prod_{k=1}^{\infty} (1 + \frac{1}{\lambda_k}) \geq \sum_{k=1}^{\infty} \frac{1}{\lambda_k} = \sum_{k=1}^{\infty} E[T_k] = \infty,$$

所以 $E[\exp(-W)] = 0$. 于是 $\exp(-W) = 0$ a.s., 即 $W = \infty$ a.s.. 所以我们证明了 (1).

注意到上述推导也说明了只要 $P(W = \infty) > 0$, 就有 $\sum_{k=1}^{\infty} E[T_k] = \infty$, 从而 $P(W = \infty) = 1$. 这就证明了 (2).

下面为我们来证明 (3) 和 (4). 利用概率的连续性, 得到

$$\begin{aligned} P(M \geq t) &= P\left(\bigcap_{k=1}^{\infty} \{T_k \geq t\}\right) = \lim_{n \rightarrow \infty} P\left(\bigcap_{k=1}^n \{T_k \geq t\}\right) \\ &= \lim_{n \rightarrow \infty} \prod_{k=1}^n P(T_k \geq t) = \lim_{n \rightarrow \infty} \prod_{k=1}^n e^{-\lambda_k t} = e^{-\lambda_0 t}. \end{aligned}$$

由此可知 M 服从参数为 λ_0 的指数分布. 当 $\lambda_0 = \infty$ 时, 我们知对任意 $t > 0$, $P(M \geq t) = e^{-\infty} = 0$, 所以 $P(M = 0) = 1$.

例 4.9

一艘潜水艇有三个航行设备，但只要其中两个正常工作，潜水艇仍然可以在海上。假设三个设备损坏时间分别服从均值为 1 年，1.5 年和 3 年的指数分布，那么潜水艇可以在海上平均待多长时间？

证明.

由条件我们知三个设备损坏时间服从参数 1, $2/3$ 和 $1/3$ 的指数分布. 于是第一次有设备损坏的时间服从参数 $1 + 2/3 + 1/3 = 2$ 的指数分布, 从而第一次设备损坏的时间为 $1/2$ 年. 第一次有设备损坏时, 损坏的设备是第一个设备的概率为 $1/(1 + 2/3 + 1/3) = 1/2$, 这时到下一次有设备损坏的时间服从参数为 $2/3 + 1/3 = 1$ 的指数分布, 于是平均时间为 1 年. 损坏的设备是第二个设备的概率为 $(2/3)/(1 + 2/3 + 1/3) = 1/3$, 这时到下一次有设备损坏的时间服从参数为 $1 + 1/3 = 4/3$ 的指数分布, 于是平均时间为 $3/4$ 年. 损坏的设备是第三个设备的概率为 $(1/3)/(1 + 2/3 + 1/3) = 1/6$, 这时到下一次有设备损坏的时间服从参数为 $1 + 2/3 = 5/3$ 的指数分布, 于是平均时间为 $3/5$ 年. 于是到第二个设备损坏的总平均时间为

$$1/2 + (1/2) \cdot 1 + (1/3) \cdot (3/4) + (1/6) \cdot (3/5) = 0.5 + 0.5 + 0.25 + 0.10 = 1.35 \text{ 年.}$$



下周内容

- 泊松分布 课本 2.1-2.2