

## 第二章 随机过程预备知识



- 1 条件分布与条件期望
  - 条件概率与条件期望
  - 事件域与可测性
  - 条件期望的性质
  
- 2 什么是随机过程
  - 随机过程的定义
  - 样本轨道
  - 随机过程的分布
  - 常见的随机过程



数学期望是关于随机事件长期平均结果的预测;

条件期望则表明, 随着我们获得更多信息, 这个预测本身也会被动态修正和更新.

## 第2.1节 条件期望

本节将定义随机变量关于事件域的条件数学期望.

思路 —— 由浅入深:

在给出一般定义之前, 先从简单的场合给出定义.



## 条件概率与条件期望

---

给定概率空间 $(\Omega, \mathcal{F}, \mathbb{P})$ . 考虑事件 $A, B \in \mathcal{F}$ :  $\mathbb{P}(A) > 0$ ,

注. (1) 条件概率  $\mathbb{P}(B|A) := \frac{\mathbb{P}(B \cap A)}{\mathbb{P}(A)}$ ; (2) 概率  $\mathbb{P}(B)$  也可以看成期望  $\mathbb{E}[1_B]$ .

记在事件 $A$  发生的条件下, 随机变量 $X$  的条件期望为

$$\mathbb{E}[X|A] := \frac{\mathbb{E}[X1_A]}{\mathbb{P}(A)} = \sum_i x_i \mathbb{P}(X = x_i|A).$$



## 条件概率与条件期望

---

给定概率空间 $(\Omega, \mathcal{F}, \mathbb{P})$ . 考虑事件 $A, B \in \mathcal{F}$ :  $\mathbb{P}(A) > 0$ ,

注. (1) 条件概率  $\mathbb{P}(B|A) := \frac{\mathbb{P}(B \cap A)}{\mathbb{P}(A)}$ ; (2) 概率  $\mathbb{P}(B)$  也可以看成期望  $\mathbb{E}[1_B]$ .

记在事件 $A$  发生的条件下, 随机变量 $X$  的条件期望为

$$\mathbb{E}[X|A] := \frac{\mathbb{E}[X1_A]}{\mathbb{P}(A)} = \sum_i x_i \mathbb{P}(X = x_i|A).$$



## 条件期望与条件分布的关系

---

例2.1.2 假设 $\lambda > 0, \rho > 0$  是两个参数,

随机变量 $X \sim P(\lambda)$  与 $Y \sim P(\rho)$  相互独立.

给定自然数 $n$ , 求在 $X + Y = n$  的条件下 $X$  的期望.

review. 在 $X + Y = n$  的条件下,  $X$  的条件分布.



解. 首先

$$\mathbb{P}(X + Y = n) = e^{-(\lambda + \rho)} \frac{(\lambda + \rho)^n}{n!}, \quad n \geq 0,$$

推出对任意  $k \leq n$ ,

$$\begin{aligned} \mathbb{P}(X = k | X + Y = n) \\ &= \frac{n! \lambda^k \rho^{n-k}}{k!(n-k)! (\lambda + \rho)^n}. \end{aligned}$$

也就是说,

$$X | X + Y = n \sim B(n, \frac{\lambda}{\lambda + \rho}).$$

根据二项分布的性质有  $\mathbb{E}[X | X + Y = n] = \frac{\lambda n}{\lambda + \rho}$ .

# 

注释. 对于两个离散型随机变量 $X, Y$ ,

(1) 给定实数 $x$ , 条件期望 $\mathbb{E}[Y|X = x]$  是条件分布的期望:

$$\mathbb{E}[Y|X = x] = \sum_y y \mathbb{P}(Y = y|X = x),$$

其中条件分布

$$\mathbb{P}(Y = y|X = x) = \frac{\mathbb{P}(Y = y, X = x)}{\mathbb{P}(X = x)}.$$

(2) 条件期望

$\mathbb{E}[Y|X = x]$  中 $x$  替换为 $X$ , 记为 $\mathbb{E}[Y|X]$ ,

则称为 $Y$  关于 $X$  的条件期望.



例如, 若 $X$  的值域是 $\{x_1, x_2, \dots, x_n\}$ , 则

$$\mathbb{E}[Y|X] = \sum_{i=1}^n \mathbb{E}[Y|X = x_i] 1_{\{X=x_i\}},$$

两边求期望, 则有

$$\mathbb{E}(\mathbb{E}[Y|X]) = \sum_{i=1}^n \mathbb{E}[Y|X = x_i] \mathbb{P}(X = x_i) = \mathbb{E}[Y].$$

上述公式就是所谓全(重)期望公式.



例2.1.3 猫在有三个门的迷宫中,

- 如果选第一个门, 爬2 个小时就可以出去了;
  - 如果选第二个门, 爬1 个小时后回到原地;
  - 如果选第三个门, 爬3 个小时后回到原地.
- (1) 如果是个傻猫, 每次到这个地方都随机选一个门, 问平均需要多少时间爬出去?
  - (2) 如果是个聪明猫, 之前选过的门不再选, 问平均需要多少时间爬出去?



解. (1) 用  $Y$  表示笨猫出去所需时间.  $X$  是选的门. 那么

$$\begin{aligned}\mathbb{E}[Y|X = 1] &= 2, \quad \mathbb{E}[Y|X = 2] = 1 + \mathbb{E}[Y], \\ \mathbb{E}[Y|X = 3] &= 3 + \mathbb{E}[Y].\end{aligned}$$

因此

$$\mathbb{E}[Y] = \frac{1}{3}(2 + 1 + \mathbb{E}[Y] + 3 + \mathbb{E}[Y]),$$

故得  $\mathbb{E}[Y] = 6$ .



(2) 聪明猫按选门的顺序有下面的可能: 1, 21, 31, 321, 231,  
概率分别为

$$1/3, 1/6, 1/6, 1/6, 1/6.$$

且

$$\begin{aligned}\mathbb{E}[Y|1] &= 2, \quad \mathbb{E}[Y|21] = 3, \quad \mathbb{E}[Y|31] = 5, \\ \mathbb{E}[Y|321] &= \mathbb{E}[Y|231] = 6.\end{aligned}$$

因此

$$\mathbb{E}[Y] = 2/3 + 3/6 + 5/6 + 6/6 + 6/6 = 4.$$

聪明猫平均少走2 个小时.

#



## 全期望公式: 全概率公式的推广

**Review.** 为求事件 $A$ 的概率, 假设 $\{B_i, i = 1, 2, \dots, n\}$ 是 $\Omega$ 的一个划分, 即

$$B_i B_j = \emptyset, i \neq j \text{ 且 } \Omega = \bigcup_i^n B_i,$$

则  $\mathbb{P}(A) = \sum_{i=1}^n \mathbb{P}(A|B_i)\mathbb{P}(B_i)$ . 注意到概率和期望的本质相同, 全概率公式有如下推广:

假设 $\{B_i, i = 1, 2, \dots, n\}$ 是样本空间的一个划分, 则

$$\mathbb{E}[Y] = \sum_{i=1}^n \mathbb{E}[Y|B_i]\mathbb{P}(B_i).$$

(注意, 该划分与随机变量 $Y$ 有关.)



**例2.1.4** 从一个装有  $a$  个白球与  $b$  个黑球的袋子A 中随机地不放回拿出  $n$  个球, 放入另外一个袋子B 中拌匀, 再从B 中取出  $m$  个球, 求白球数的期望.

**注.** 从B 中取出  $m$  个球中的白球数记为  $Y$ , A 袋中一共  $n$  个球, 其中白球数也是随机变量, 记为  $X$ , 则由超几何分布的期望公式,

$$\mathbb{E}[Y|X] = \frac{mX}{n}.$$



解. 用 $X, Y$  分别表示从A, B 袋中取出的白球数, 则

$$\mathbb{E}[Y|X] = \frac{mX}{n}.$$

因此

$$\mathbb{E}[Y] = \mathbb{E}[E[Y|X]] = \frac{m}{n}\mathbb{E}[X].$$

再用超几何分布的期望公式, 推出

$$\mathbb{E}[Y] = \frac{m}{n}\mathbb{E}[X] = \frac{m}{n} \frac{an}{a+b} = \frac{am}{a+b}.$$

#

例2.1.5 袋子A, B 中分别有 $n$ 个白球和 $n$ 个黑球, 每次各取一球交换, 求 $k$ 次交换之后A 中的白球数之期望.

解. 用 $X_i$  表示 $i$ 次交换之后A 中的白球数,  $i \geq 1, X_0 = n$ , 则对 $\forall j \leq n$ ,

$$\mathbb{P}(X_i = j - 1 | X_{i-1} = j) = \frac{j^2}{n^2},$$

$$\mathbb{P}(X_i = j | X_{i-1} = j) = \frac{2j(n - j)}{n^2},$$

$$\mathbb{P}(X_i = j + 1 | X_{i-1} = j) = \frac{(n - j)^2}{n^2}.$$

条件期望为

$$\begin{aligned}\mathbb{E}[X_i | X_{i-1} = j] &= (j-1) \frac{j^2}{n^2} + j \frac{2j(n-j)}{n^2} + (j+1) \frac{(n-j)^2}{n^2} \\ &= \left(1 - \frac{2}{n}\right)j + 1.\end{aligned}$$

因此

$$\mathbb{E}[X_i | X_{i-1}] = \left(1 - \frac{2}{n}\right)X_{i-1} + 1.$$

然后对任意  $i \geq 1$ , 得到递推



$$\begin{aligned}\mathbb{E}[X_i] &= (1 - 2/n)\mathbb{E}[X_{i-1}] + 1 = (1 - 2/n)[(1 - 2/n)\mathbb{E}[X_{i-2}] + 1] + 1 \\ &= \dots\end{aligned}$$

$$\begin{aligned}&= (1 - 2/n)^i \mathbb{E}[X_0] + \sum_{j=0}^{i-1} (1 - 2/n)^j \\ &= (1 - 2/n)^i n + \frac{1 - (1 - 2/n)^i}{2/n} = \frac{n}{2} (1 + (1 - 2/n)^i),\end{aligned}$$

最后

$$\mathbb{E}[X_k] = \frac{n}{2} [1 + (1 - 2/n)^k].$$

#

Handwriting logo

补充例. 一名矿工陷进一个有三扇门的矿井中. 第一扇门通到一个隧道, 走两小时后他可到达安全区; 第二扇门通到又一个隧道, 走三个小时会使他回到这个矿井中; 第三扇门通到再一个隧道, 走五个小时, 也会使他回到矿井中. 假定这矿工总是等可能地在三扇门中选一扇, 让我们计算矿工到达安全区地时间 $X$  的矩母函数.



## 连续型随机变量的条件期望

**Review.** 设连续型  $(X, Y) \sim p(x, y)$ , 边缘密度函数分别为  $p_X(x), p_Y(y)$ ,  
已知  $X = x$  时,  $Y$  的条件密度函数是  $p(y|x) = \frac{p(x,y)}{p_X(x)}$ , 记为  $\mathbb{P}(Y \in dy|X = x) = p(y|x)dy$ .

条件期望  $\mathbb{E}[Y|X = x]$  是条件密度函数的期望:

$$\mathbb{E}[Y|X = x] = \int_{\mathbf{R}} yp(y|x)dy =: f(x) \quad (:\text{直观形式.})$$

则有条件期望的定义如下:

$$\mathbb{E}[Y|X] := f(X).$$



**例2.1.6** 假设一个值为 $s$ 的信号从位置 $A$ 发出, 在位置 $B$ 被接收到的信号值服从参数是 $(s, 1)$ 的正态分布.  
再假设从位置 $A$ 发出的信号 $S$ 服从参数是 $(\mu, \sigma^2)$ 的正态分布, 在位置 $B$ 被接收到的信号值 $R = r$ ,  
则对位置 $A$ 发出的信号值的最优预测为多少?



解. 首先, 计算给定 $R$ 的条件下 $S$ 的条件密度函数

$$f_{S|R}(s|r) = \frac{f_S(s)f_{R|S}(r|s)}{f_R(r)} = C'e^{-(s-\mu)^2/2\sigma^2}e^{-(r-s)^2/2},$$

其中,  $C'$  与 $s$  无关. 而

$$\begin{aligned}\frac{(s-\mu)^2}{2\sigma^2} + \frac{(r-s)^2}{2} &= s^2\left(\frac{1}{2\sigma^2} + \frac{1}{2}\right) - \left(\frac{\mu}{\sigma^2} + r\right)s + C_1 \\ &= \frac{1+2\sigma^2}{2\sigma^2} \left(s - \frac{\mu + r\sigma^2}{1+\sigma^2}\right)^2 + C_2,\end{aligned}$$

其中,  $C_1, C_2$  与 $s$  无关. 所以

$$f_{S|R}(s|r) = C \exp \left\{ - \left( s - \frac{\mu + r\sigma^2}{1 + \sigma^2} \right)^2 \middle/ \frac{2\sigma^2}{1 + 2\sigma^2} \right\},$$

其中,  $C$  与  $s$  无关. 可以断定在接收信号值  $r$  一定的情况下, 按照最小化均方误差原则, 对发出信号的最优预测是

$$\mathbb{E}[S|R = r] = \frac{1}{1 + \sigma^2}\mu + \frac{\sigma^2}{1 + \sigma^2}r.$$

上式表明条件期望等于

信号的先验均值  $\mu$  与接受信号  $r$  的加权平均,

其各自的权重之比为  $1 : \sigma^2$ .

### 例2.1.7 (Bayes 估计)

掷一个非均匀的硬币  $n$  次, 设硬币正面朝上概率是  $\theta$ ,  $X$  是正面次数, 那么

$$p(x|\theta) = C_n^x \theta^x (1 - \theta)^{n-x}, \quad 0 \leq x \leq n.$$

根据Bernoulli 大数定律, 如果观察到  $X = x$ , 那么

$\theta$  的估计是正面的频率  $x/n$ .

现假设  $\theta$  的先验分布是  $U(0, 1)$ , 求  $\theta$  的Bayes 估计.



## 事件域是分类

---

**注.** 所谓分类, 就是把一个集合分拆为互斥的子集. 比如, 奇偶, 性别, 国家, 省等都是分类. 每个子集可视为分块. 也就是说, 分类意味着信息.

针对样本空间 $\Omega$ ,

所谓事件域 $\mathcal{G}$  是指, 样本空间满足某些条件的子集组成的集合, 本质上相当于给 $\Omega$  分类.

## 随机变量诱导的事件域

**Review.** (第一章例4) 班上有6 各学生, 编号分别是1~6 号, 其中

1~4 号是男生, 5~6 号是女生,  
1, 2, 6 号戴眼镜, 3, 4, 5 号不戴眼镜.

任取一名学生,

$$X = \begin{cases} 0, & \text{女生,} \\ 1, & \text{男生,} \end{cases} \quad Y = \begin{cases} 0, & \text{不戴眼镜,} \\ 1, & \text{戴眼镜,} \end{cases}$$

写出 $\sigma(X), \sigma(Y), \sigma(X, Y)$ .

## 可测性

---

设 $X$  是 $(\Omega, \mathcal{F}, \mathbb{P})$  上的随机变量,  $\mathcal{G} \subset \mathcal{F}$ .

### 定义2.1.1

如果对任何实数 $x$ , 有  $\{X \leq x\} \in \mathcal{G}$ , 则称 $X$  关于 $\mathcal{G}$  可测.

(直观上就是,

$X$  不能区别  $\mathcal{G}$  的分类, 或者说,  $X$  是基于信息  $\mathcal{G}$  的表达.)



### 命题2.1.1

$Y$  关于  $\sigma(X)$  可测当且仅当  $Y$  是  $X$  的函数, 即

$$\exists f: \text{Borel 函数, s.t. } Y = f(X).$$

类似地, 若  $Y$  关于  $\sigma(X_1, \dots, X_k)$  可测, 则存在多元 Borel 函数  $f$ , 使得

$$Y = f(X_1, \dots, X_k).$$

这时, 经常说  $Y$  依赖于  $X_1, \dots, X_k$ .



## 条件期望的两个关键性质

### 定理2.1.2

$\xi$  是  $Y$  关于  $\mathcal{G}$  的条件期望当且仅当:

- (1) (可测性)  $\xi$  是关于  $\mathcal{G}$  可测的;
- (2) (局部均值一致性) 对任何  $A \in \mathcal{G}$ , 有

$$\mathbb{E}[(Y - \xi)1_A] = 0.$$

记为  $\xi = \mathbb{E}[Y|\mathcal{G}]$ .

## 条件期望的性质

---

(1) 线性性:

$$\mathbb{E}[c_1 Y_1 + c_2 Y_2 | \mathcal{G}] = c_1 \mathbb{E}[Y_1 | \mathcal{G}] + c_2 \mathbb{E}[Y_2 | \mathcal{G}].$$

(2) 单调性:

若  $X \leq Y$ , 则  $\mathbb{E}[X | \mathcal{G}] \leq \mathbb{E}[Y | \mathcal{G}]$ .

(3) 全期望公式:

$$\mathbb{E}[\mathbb{E}[Y | \mathcal{G}]] = \mathbb{E}[Y].$$

(4) 如果  $Y, Z$  是随机变量, 且  $Z$  是  $\mathcal{G}$  可测的随机变量, 那么

$$\mathbb{E}[ZY | \mathcal{G}] = Z \mathbb{E}[Y | \mathcal{G}].$$



所谓 $X$  与  $\mathcal{G}$  独立是指, 对任何  $A \in \mathcal{G}$ ,  $X$  与  $1_A$  独立.

(5) 若 $X$  与 $\mathcal{G}$  独立, 则 $\mathbb{E}[X|\mathcal{G}] = \mathbb{E}[X]$ , 特别的,

$$\mathbb{E}[X|\{\emptyset, \Omega\}] = \mathbb{E}[X].$$

(6)\* 设  $\mathcal{G}_1, \mathcal{G}_2$  是子事件域, 且  $\mathcal{G}_1 \subset \mathcal{G}_2$ . 那么

$$\mathbb{E}[\mathbb{E}[Y|\mathcal{G}_1]|\mathcal{G}_2] = \mathbb{E}[\mathbb{E}[Y|\mathcal{G}_2]|\mathcal{G}_1] = \mathbb{E}[Y|\mathcal{G}_1].$$

i.e.,

粗的预测与细的预测是相容的, 是全期望公式的推广.



(7) 简单的信息得到简单的预测, 完整的信息得到完整的预测:

$$\mathbb{E}[Y|\{\emptyset, \Omega\}] = \mathbb{E}[Y], \quad \mathbb{E}[Y|\mathcal{F}] = Y.$$

(8) Jensen 不等式:

若 $\phi$  是凸函数, 即 $\phi'' \geq 0$ , 并且 $\phi(\xi)$  可积, 则

$$\phi(\mathbb{E}[\xi|\mathcal{G}]) \leq \mathbb{E}[\phi(\xi)|\mathcal{G}].$$

## 例2.1.9 (二维正态分布)

$$(X, Y) \sim p(x, y) = \frac{1}{2\pi\sqrt{1-r^2}} \exp\left(-\frac{x^2 - 2rxy + y^2}{2(1-r^2)}\right),$$

给定  $X = x$  的条件密度

$$\begin{aligned} p(y|x) &= \frac{p(x, y)}{p_X(x)} = \frac{1}{\sqrt{2\pi(1-r^2)}} \exp\left(-\frac{x^2 - 2rxy + y^2}{2(1-r^2)} + \frac{x^2}{2}\right) \\ &= \frac{1}{\sqrt{2\pi(1-r^2)}} \exp\left(-\frac{r^2x^2 - 2rxy + y^2}{2(1-r^2)}\right) \\ &= \frac{1}{\sqrt{2\pi(1-r^2)}} \exp\left(-\frac{(y-rx)^2}{2(1-r^2)}\right) \leftrightarrow N(rx, 1-r^2), \end{aligned}$$

所以  $\mathbb{E}[Y|X = x] = rx$ , 严格地  $\mathbb{E}[Y|X] = rX$ .



另解. 因为 $X, Y$  是标准化的且协方差为 $r$ , 所以

$$\mathbb{E}[(Y - rX)X] = 0,$$

即 $Y - rX$  与 $X$  不相关. 由于 $X, Y$  是联合正态的, 故而

$Y - rX$  与 $X$  独立,

由条件期望的性质

$$\mathbb{E}[Y - rX|X] = \mathbb{E}[Y - rX] = 0,$$

再应用条件期望性质得

$$\mathbb{E}[Y|X] = \mathbb{E}[rX|X] = rX.$$

例2.1.10 设  $X = \{X_n, n \geq 1\}$  是一个独立同分布的随机序列,  $N$  是非负整数值的随机变量且与  $X$  相互独立, 求

随机和  $Y := \sum_{i=1}^N X_i$  的矩母函数, 期望以及方差.

解. 先对任意  $n$ ,

$$\begin{aligned}\mathbb{E}[e^{tY} | N = n] &= \mathbb{E}[e^{t \sum_{i=1}^N X_i} | N = n] = \mathbb{E}[e^{t \sum_{i=1}^n X_i} | N = n] \\ &= \mathbb{E}[e^{t \sum_{i=1}^n X_i}] = (\psi_X(t))^n,\end{aligned}$$

所以

$$\begin{aligned}\psi_Y(t) &:= \mathbb{E}[e^{tY}] = \mathbb{E}[\mathbb{E}[e^{tY} | N]] \\ &= \mathbb{E}[(\psi_X(t))^N].\end{aligned}$$



从而求导可得

$$\begin{aligned}\psi'_Y(t) &= \mathbb{E}[N(\psi_X(t))^{N-1}\psi'_X(t)], \\ \psi''_Y(t) &= \mathbb{E}[N(N-1)(\psi_X(t))^{N-2}(\psi'_X(t))^2 \\ &\quad + N(\psi_X(t))^{N-1}(\psi''_X(t))].\end{aligned}$$

取  $t = 0$  可求得

$$\begin{aligned}\mathbb{E}Y &= \mathbb{E}[NEX_1] = \mathbb{E}N \cdot \mathbb{E}X_1 \quad (: \text{Wald 等式}), \\ \mathbb{E}Y^2 &= \mathbb{E}N \cdot D(X_1) + \mathbb{E}X_1^2 \cdot (\mathbb{E}X_1)^2, \\ D(Y) &= \mathbb{E}N \cdot D(X_1) + (\mathbb{E}X_1)^2 \cdot D(N).\end{aligned}$$

#

Handwriting logo

## 第2.2节 随机过程的基本概念

随机过程是一族随机变量的集合, 用于描述随时间变化的随机现象.

其例随手可得:

- 去某银行办理业务的排队等候时间;
- 股票在一天中的价格变化;
- 某食堂一天中吃饭人数的变化;
- 某路段一天车流量的变化;
- 上海一年降雨量的变化;

... ..



设 $(\Omega, \mathcal{F}, \mathbb{P})$  是一个概率空间,  $(E, \mathcal{B}(E))$  是一个可测空间.  
设 $T$  是一个指标集.

### 定义2.2.1

若对任意 $t \in T$ ,  $X_t$  是 $(\Omega, \mathcal{F})$  到 $(E, \mathcal{B}(E))$  上的可测映射, 则称 $X = \{X_t : t \in T\}$  为

$(\Omega, \mathcal{F}, \mathbb{P})$  上以 $E$  为状态空间的随机过程,

简称 $E$ -值过程或过程.

- 当 $E$  是实数或复数空间时, 分别称过程是实值过程与复值过程. 本课程研究的都是前者.
- 如果需要, 把随机变量或随机向量都看成为随机过程.

## 分类

---

注. 指标参数集  $T$  是  $\begin{cases} \text{至多可列的, 则称为离散时间;} \\ \text{一个实数区间, 则称为连续时间.} \end{cases}$

随机过程按指标集和状态空间分类如下:

- 离散时间离散状态;
- 离散时间连续状态;
- 连续时间离散状态;
- 连续时间连续状态.



## 随机游动

例2.2.1 设  $\{\xi_n : n \geq 1\}$  是独立同分布的 Bernoulli 序列:

$$\mathbb{P}(\xi_n = 1) = p, \mathbb{P}(\xi_n = -1) = q, \quad p, q \geq 0, p + q = 1.$$

注. 比 Bernoulli 序列  $\{\xi_n : n \geq 1\}$  更有意思的是下面的过程  $\{S_n, n \geq 1\}$ :

让值 1 表示甲赢, 这时他得到 1 元钱; 值 -1 表示甲输, 这时他付出 1 元钱.

记  $S_n = S_0 + \sum_{i=1}^n \xi_i$  ( $n$  次赌博后甲所拥有的赌资),  $n \geq 1$ .

称  $\{S_n : n \geq 0\}$  为随机游动或随机游走.

如果  $S_0 = x$ , 称  $\{S_n, n \geq 0\}$  是从  $x$  出发的随机游动. #

## 注释

---

用映射来表示

$X_t(\omega) : T \times \Omega \rightarrow E$ , 是定义在  $T \times \Omega$  上的二元单值函数.

- 固定  $t$ ,  $X_t(\cdot)$  是随机变量;
- 固定  $\omega$ ,  $X_t(\cdot)$  是  $T$  上的函数.

本课程主要关注的是后者, 称为样本函数/样本轨道.



## 样本轨道

---

将随机过程作为一个整体考虑.

对任何  $\omega \in \Omega$ ,  $t \mapsto X_t(\omega)$  是  $T$  到  $E$  的映射,

称之为  $\omega$  的样本轨道.

一条样本轨道是对过程的一次具体观测结果.

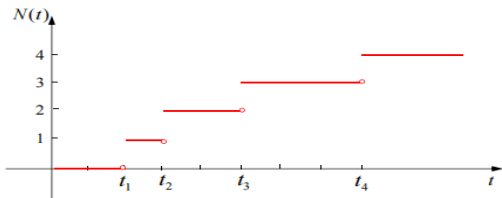
例如,

- 在新闻中看到的股票或者股指走势就是样本轨道;
- 一个花粉在液体表面的运动轨迹也是样本轨道.

补充例. 以 $N(t)$  表示 $(0, t]$  内到某保险公司理赔的人数, 则 $\{N(t), t \geq 0\}$  是连续时间离散状态的随机过程, 状态空间是

$$E = \{0, 1, 2, \dots\}.$$

假设不会有两人或两人以上同时理赔, 用 $t_i$  表示第 $i$  个人的理赔时间, 则样本函数如下图.



#  
Ka+Ting

## 一维分布, 二维分布

---

考虑过程  $X = \{X(t), t \in T\}$ .

对每个固定的  $t \in T$ , 随机变量  $X(t)$  的分布函数与  $t$  有关:

- 对任意  $t \in T, x \in \mathbf{R}$ , 称  $F_t(x) := \mathbb{P}(X(t) \leq x)$  为  $X$  的一维分布函数;
- 对任意  $s, t \in T, x, y \in \mathbf{R}$ ,  
称  $F_{s,t}(x, y) := \mathbb{P}(X(s) \leq x, X(t) \leq y)$  为  $X$  的二维分布函数.



## 均值函数与协方差函数

---

- 均值函数  $m(t) := \mathbb{E}[X(t)]$ ;
- 均方函数  $\Psi(t) := \mathbb{E}[X^2(t)]$ ;
- 方差函数  $\sigma^2(t) := D(X(t))$ ;
- 协方差函数  $K(s, t) := \text{cov}(X(s), X(t))$ .



## 随机过程的分布:

有限维分布是研究随机过程的得力工具.

对于  $n \geq 1$ ,  $t_1, t_2, \dots, t_n \in T$ ,  $x_1, x_2, \dots, x_n \in \mathbf{R}$ , 称

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

为随机过程  $X$  的一个有限维分布函数.

变化  $n$  及  $t_1, t_2, \dots, t_n$  所得  $F_{t_1, t_2, \dots, t_n}$

对应的分布的全体, 称为  $X$  的有限维分布族.

**注.** 由相容性, 通常取时间列满足  $t_1 < t_2 < \dots < t_n$ .



注. 有限维分布族完全确定过程的统计特性, 多用于理论研究.  
一般的, 不同随机过程可能有相同的有限维分布族.

### 定义2.2.2

- (1) 假设 $X, X'$  是分别在概率空间 $(\Omega, \mathcal{F}, \mathbb{P})$  和 $(\Omega', \mathcal{F}', \mathbb{P}')$  上, 有相同的状态空间 $(E, \mathcal{B}(E))$  与指标集 $T$  的随机过程.  
如果它们有相同的有限维分布族, 即对任何 $t_1, \dots, t_n \in T$ ,

$$(X_{t_1}, \dots, X_{t_n}) \text{ 与 } (X'_{t_1}, \dots, X'_{t_n}) \text{ 同分布,}$$

那么 $X, X'$  称为是等价的 / 同分布的.  
这时也称一个是另一个的版本.

## Gauss 过程

若  $X = \{X(t), t \in T\}$  的有限维分布是多维正态分布, 则称  $X$  是 Gauss 过程. 进一步地, 如果  $\mathbb{E}X_t = 0, t \in T$ , 称  $X$  是中心化的 Gauss 过程.

例2.2.4 取  $\xi \sim N(0, 1)$ , 对  $t \geq 0$ , 令  $X_t = \xi t$ , 则  $\{X_t\}$  是一个平方可积的随机过程, 且

$$\mathbb{E}X_t^2 = t^2, K(t, s) = ts.$$

另外对  $t_1, \dots, t_n \in T$ , 有限维分布的特征函数是

$$f_{t_1, \dots, t_n}(x) = \exp \left\{ -\frac{1}{2} x B x^T \right\}, x \in \mathbf{R}^n,$$

是一个正态分布. 其中,  $B = (t_1, \dots, t_n)^T (t_1, \dots, t_n)$ .



补充例. 设 $\{X(t), t \geq 0\}$  是 Gauss 过程,  $m(t) = t$ ,  $K(t, s) = ts + 1$ , 求 $X(1), X(2), X(1) + X(2)$  的分布.

解. 由于

$$\sigma^2(t) = \text{cov}(X(t), X(t)) = t^2 + 1,$$

所以 $X(t) \sim N(t, t^2 + 1)$ , 特别的,

$$X(1) \sim N(1, 2), X(2) \sim N(2, 5).$$

易知 $(X(1), X(2))$  服从正态分布, 从而 $(X(1) + X(2))$  也服从正态分布,

$$D(X(1) + X(2)) = D(X(1)) + D(X(2)) + 2\text{cov}(X(1), X(2)) = 13,$$

有 $X(1) + X(2) \sim N(3, 13)$ .

#

H+Ting

## 平稳过程

例2.2.5 如果  $X = \{X(t), t \in T\}$  的任何有限维分布是平移不变的。精确地说， $T$  是一个时间半群，即对加法封闭，且对任何  $t_1, \dots, t_n, t \in T, A_1, \dots, A_n \in E$ ，有

$$\mathbb{P}(X_{t_1+t} \in A_1, \dots, X_{t_n+t} \in A_n) = \mathbb{P}(X_{t_1} \in A_1, \dots, X_{t_n} \in A_n),$$

那么称  $X$  是(严格)平稳过程。

如果  $X$  是平方可积的，且  $m(t) = \mathbb{E}X_t$  与  $t$  无关，协方差函数  $K$  满足齐性

$$K(s+h, t+h) = K(s, t),$$

那么称  $X$  为广义(或宽)平稳过程。

#

## 独立增量过程

例2.2.7 称 $X$  有独立增量, 若对于任意的 $n \geq 1$ ,  
 $0 \leq t_0 < t_1 < t_2 < \cdots < t_n$ ,

$X(t_0), X(t_1) - X(t_0), \cdots, X(t_n) - X(t_{n-1})$  相互独立.

即, 过程在不相重叠的区间上增量独立.

进一步地,

若对于 $0 \leq s \leq t$ ,

增量 $X(t) - X(s)$  的分布只依赖于 $t - s$ ,

则为平稳独立增量过程. 即,

过程在任意两点间增量的分布仅依赖于这两点之间的距离. #

Handwriting signature: Ting

## 马氏过程

例2.2.8 随机过程  $X = \{X(t), t \in T\}$ , 若具有:

### Markov 性

对  $\forall n \geq 1, \forall t_1 < t_2 < \cdots < t_n < t_{n+1}, \forall A \in \mathcal{E}$ ,

$$\mathbb{P}(X(t_{n+1}) \in A | X(t_1), \cdots, X(t_n)) = \mathbb{P}(X(t_{n+1}) \in A | X(t_n)),$$

则称  $X$  为马氏(Markov)过程.

注. 这个定义是条件期望形式的.

#

## 更新过程

例2.2.9  $\{X_n\}$  是独立同分布随机系列, 且  $\mathbb{P}(X_n > 0) = 1$ ,  
令  $S_n = \sum_{i=1}^n X_i$ , 那么

$\{S_n\}$  是平稳独立增量过程,

再取它的右连续逆

$$Y_t := \inf\{n \geq 0 : S_n > t\}, \quad t \geq 0.$$

过程  $Y = \{Y_t : t \geq 0\}$  称为更新过程. #

思考. 更新过程一般不是马氏过程, 只有

当  $X_i \sim \text{Exp}(\lambda)$  时, 相应的  $Y$  是马氏过程.

