随机过程+马尔科夫链

2025年4月6日　17:11

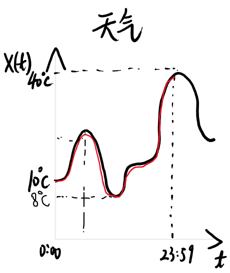
# 1、初步了解随机过程

## （一）随机过程

设 T 是一无限实数集. 我们把依赖于参数 $t\in T$ 的一族（无限多个）随机变量称为**随机过程**，记为 $\{X(t),t\in T\}$，这里对每一个 $t\in T$，$X(t)$ 是一随机变量. T 叫做**参数集**. 我们常把 t 看作为时间，称 $X(t)$ 为时刻 t 时过程的**状态**，而 $X(t_1)=x$（实数）说成是 $t=t_1$ 时过程处于状态 x. 对于一切 $t\in T$，$X(t)$ 所有可能取的一切值的全体称为随机过程的**状态空间**.

对随机过程 $\{X(t),t\in T\}$ 进行一次试验（即在 T 上进行一次全程观测），其结果是 t 的函数，记为 $x(t),t\in T$，称它为随机过程的一个**样本函数或样本曲线**. 所有不同的试验结果构成一族（可以只包含有限个结果，见本节例 1）样本函数.

随机过程可以看作是多维随机变量的延伸. 随机过程与其样本函数的关系就像数理统计中总体与样本的关系一样.




$t\in T$

参数集T →一天中的每个时刻

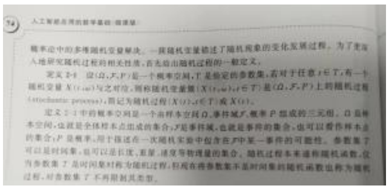
随机变量X(t)→该时刻对应温度

样本函数 →某一天温度变化曲线

状态空间 →8°C~40°C

**注意：**

- **无限随机变量：** 每个时刻的温度都是一个独立的随机变量。
- **不确定性：** 虽然今天温度从10℃涨到40℃（样本函数），但明天可能完全不同的走势（另一样本函数）。
- **总体与样本：** 所有可能的走势构成随机过程，而你观测到的只是冰山一角。



随机试验→多个可能的确定结果　可能性(概率)

↳抛硬币 正 / 反 →基本事件→只有1个元素

{0,1} 映射 数轴 ⇐ {正,反}⇒结果的集合

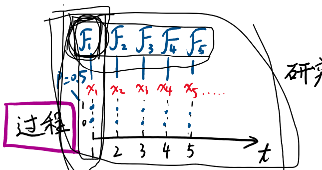
事件→子集　∅　{正}　{反}　{正,反}

不可能　随机　必然

n种可能,就是$2^n$个子集

$\{\varnothing,\{0\},\{1\},\{0,1\}\}\to\mathcal{F}$

$\Omega$




研究自变量$x_n$之间的关系

关系→作运算

$x_2-x_1$

$x_2+x_1$

$x_2\cdot x_1$

5天投5次硬币

$x_1\begin{cases}0\\1\end{cases}$ $x_2\begin{cases}0\\1\end{cases}$ $x_3\begin{cases}0\\1\end{cases}$ $x_4\begin{cases}0\\1\end{cases}$ $x_5\begin{cases}0\\1\end{cases}$

## （二）随机函数的分布函数族

给定随机过程 $\{X(t),t\in T\}$. 对于每一个固定的 $t\in T$，随机变量 $X(t)$ 的分布函数一般与 t 有关，记为

$$F_X(x,t)=P\{X(t)\leqslant x\},x\in\mathbb{R}.$$

称它为随机过程 $\{X(t),t\in T\}$ 的**一维分布函数**，而 $\{F_X(x,t),t\in T\}$ 称为**一维分布函数族**.

一维分布函数族刻画了随机过程在各个个别时刻的统计特性. 为了描述随





机过程在不同时刻状态之间的统计联系，一般可对任意 $n\ (n=2,3,\cdots)$ 个不同的时刻 $t_1,t_2,\cdots,t_n\in T$，引入 n 维随机变量 $(X(t_1),X(t_2),\cdots,X(t_n))$，它的分布函数记为

$$F_X(x_1,x_2,\cdots,x_n;t_1,t_2,\cdots,t_n)=P\{X(t_1)\leqslant x_1,X(t_2)\leqslant x_2,\cdots,X(t_n)\leqslant x_n\},$$
$$x_i\in\mathbb{R},\quad i=1,2,\cdots,n.$$

对于固定的 n，我们称 $\{F_X(x_1,x_2,\cdots,x_n;t_1,t_2,\cdots,t_n),t_i\in T\}$ 为随机过程 $\{X(t),t\in T\}$ 的 **n 维分布函数族**.

当 n 充分大时，n 维分布函数族能够近似地描述随机过程的统计特性. 显然，n 取得愈大，则 n 维分布函数族描述随机过程的特性也愈趋完善. 一般，可以指出（科尔莫戈罗夫定理）：**有限维分布函数族**，即 $\{F_X(x_1,x_2,\cdots,x_n;t_1,t_2,\cdots,t_n),n=1,2,\cdots,t_i\in T\}$，完全地确定了随机过程的统计特性.

在上一节，我们曾将随机过程按其状态或时间的连续或离散进行了分类. 然而，随机过程的本质的分类方法乃是按其分布特性进行分类. 具体地说，就是依照过程在不同时刻的状态之间的特殊统计依赖方式，抽象出一些不同类型的模型，如独立增量过程，马尔可夫过程，平稳过程等. 我们将在以后的章节中对它们作不同程度的介绍.

### 一维分布函数族

**定义：**
对某一天t，一维分布函数FX(x,t)=P(X(t)≤x) 表示第 t 天气温不超过 x° C 的概率。
**例子：**
夏季某天 t=180（6月29日），气温可能集中在 30° C 左右，分布函数显示
P(X(180)≤35)=0.9。
冬季某天 .t=30（1月30日），气温可能分布在 -5° C 到 10° C，P(X(30)≤0)=0.6。
**意义：** 一维分布刻画了单日气温的统计规律，例如夏季高温概率高，冬季低温概率高。

### 二维分布函数族

$(\Omega,\mathcal{F},P)$

**定义：**
对两天t1 和t2，二维分布函数FX(x1,x2;t1,t2)=P(X(t1)≤x1,X(t2)≤x2)
表示两天气温的联合概率。
**例子：**
若 t1=180（夏季）和t2=181，二维分布可能显示连续两天高温的正相关性，例如
P(X(180)≤35,X(181)≤35)=0.8。
若t1=180（夏季）和 t2=30（冬季），二维分布可能显示气温独立，如
P(X(180)≤35,X(30)≤0)=0.54（≈0.9×0.6）。
**意义：** 二维分布揭示了不同日期气温之间的关联性（如冷锋过境导致气温骤降）。

## （三）随机过程的部分数字特征

1. 均值函数（Mean Function）
   ~~定义：~~
   μX(t)=E[X(t)]，表示在时刻 t 所有可能状态的平均值。
   **例子：**
   假设某股票在每天上午10点的平均价格为μX(10)=100 元，反映市场对该时刻价格的“预期”。

2. 方差函数（Variance Function）
   - **定义：**
     σX2(t)=E[(X(t)−μX(t))2]，表示时刻 t 价格的波动程度。
   - **例子：**
     若股票在下午2点的方差为 σX2(14)=25，说明此时价格波动较大（标准差5元）。

3. 自相关函数（Autocorrelation Function，ACF）
   - **定义：**
     RX(t1,t2)=E[X(t1)X(t2)]，描述不同时刻价格的相关性。
   - **例子：**
     若 RX(1)=0.8，说明相邻时刻价格高度正相关（如趋势延续）。

E→一种算法

D→一种记法

$D(X)=E(X-EX)^2$

记为

$\frac{dy}{dx}$ 一阶

order

k 阶 — $\begin{vmatrix}a&c\\d&b\end{vmatrix}$ 二阶：ab-cd　强调顺序

次 — $x^2$ 二次

矩：E

原点矩：$E(X-0)^k=E(X)^k\to$ k阶原点矩 → E(X)，$E(X)^2$

中心矩：$E(X-E(X))^k\xrightarrow{k=2}E(X-E(X))^2=D(X)\Rightarrow E(X^2-2XE(X)+(E(X))^2)$

$=E(X^2)-E(2XE(X))+E((E(X))^2)$

$=E(X^2)-2E(X)E(X)+(E(X))^2$

$=E(X^2)-(E(X))^2$

$\xrightarrow{k=1}E(X-E(X))=0$

$E(X+a)$

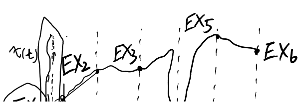
$=E(X)+Ea$

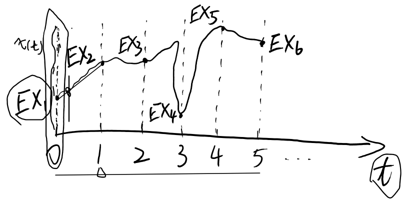
$=E(X)+a$

$\Rightarrow$

$=E(X)-E(E(X))$

$=E(X)-E(X)$






## （四）粗略谈泊松过程

### 1、泊松分布（离散）

$n, p$

二项分布：$P(x=k)=C_n^k p^k (1-p)^{n-k}$，$C_n^k \to \frac{n!}{k!(n-k)!}$

$E(x)=np$

泊松分布：二项分布的极限情况

$n\to\infty \quad p\to 0 \quad np=\lambda$

$p=\frac{\lambda}{n}$ 代入 得到

$P(x=k)=\boxed{\frac{n!}{k!(n-k)!}}\left(\frac{\lambda}{n}\right)^k\left(1-\frac{\lambda}{n}\right)^{n-k}$

当 $n\to\infty$，$\frac{n!}{(n-k)!}=\frac{n(n-1)(n-2)\cdots(n-k+1)\boxed{(n-k)!}}{\boxed{(n-k)!}}$

$=n(n-1)(n-2)\cdots(n-k+1)\approx n^k$

$\Rightarrow P(x=k)=\boxed{\frac{n^k}{k!}\left(\frac{\lambda}{n}\right)^k}\left(1-\frac{\lambda}{n}\right)^{n-k}$

$=\frac{\lambda^k}{k!}\boxed{\left(1-\frac{\lambda}{n}\right)^n}\boxed{\left(1-\frac{\lambda}{n}\right)^{-k}}$

$n\to\infty$，$\frac{\lambda}{n}\to 0$，$\left(1-\frac{\lambda}{n}\right)^{-k}\to 1$

$\lim\limits_{x\to\infty}\left(1+\frac{1}{x}\right)^x=e$

$\lim\limits_{n\to\infty}\left(1+\frac{1}{-\frac{n}{\lambda}}\right)^{n\left(-\frac{\lambda}{n}\right)}$

$=\lim\limits_{n\to\infty}\left(1+\frac{1}{-\frac{n}{\lambda}}\right)^{\left(-\frac{n}{\lambda}\right)(-\lambda)}$

$=e^{-\lambda}$

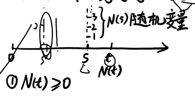
$$\boxed{=e^{-\lambda}\frac{\lambda^k}{k!}}$$

特点：$E(X)=D(X)=\lambda$

### 2、计数过程

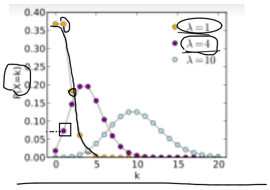
$t\to$ 连续时间

$N(t)$：从0开始到时刻 $t$ 的总人数




① $N(t)\geqslant 0$

② $S\leq t$，$N(s)\leq N(t)$

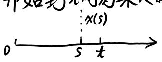



### 3、泊松过程

$x(t)$：$t\geqslant 0$ 时间

$x(t)$：从0时刻开始到 $t$ 时刻某处的总人数

① $x(0)=0$




② $x(t)$ 服从泊松分布

| $x(t)$ | 0 | 1 | 2 ⋯ n ⋯ |
|---|---|---|---|
| P | | | $P(x(t)=n)$ |

$$P(x(t)=n)=\frac{e^{-\lambda t}(\lambda t)^n}{n!}$$

$$E(x(t))=D(x(t))=\lambda t$$

## 2、初步了解马尔科夫链(条件概率)

每次心烦意乱的时候，就来听听乌龟大师这段话吧！_哔哩哔哩_bilibili

简介

俄国数学家 Andrey Andreyevich Markov 研究并提出一个用数学方法就能解释自然变化的一般规律模型，被命名为马尔科夫链（Markov Chain）。马尔科夫链为状态空间中经过从一个状态到另一个状态的转换的随机过程，该过程要求具备 **“无记忆性 ”**，即下一状态的概率分布只能由当前状态决定，在时间序列中它前面的事件均与之无关。这种特定类型的 **“无记忆性 ”** 称作马尔可夫性质。


如果求 2 次运动后的状态概率分别是多少？初始状态和终止状态未知时怎么办呢？这是就要引入**转移概率矩阵** ，可以非常直观的描述所有的概率。

$$P=\begin{array}{c}\\ A\\ B\end{array}\begin{pmatrix} A & B\\ 0.3 & 0.7\\ 0.9 & 0.1\end{pmatrix}$$

两次运动后：

$$P=\begin{pmatrix} & A & B\\ A & 0.3 & 0.7\\ B & 0.9 & 0.1\end{pmatrix}\begin{pmatrix} & A & B\\ A & 0.3 & 0.7\\ B & 0.9 & 0.1\end{pmatrix}$$

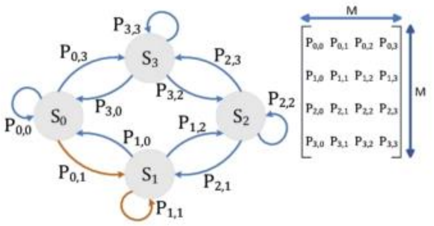
$$=\begin{pmatrix} & A & B\\ A & 0.3*0.3+0.7*0.9 & 0.3*0.7+0.7*0.1\\ B & 0.9*0.3+0.1*0.9 & 0.9*0.7+0.1*0.1\end{pmatrix}$$

$$=\begin{pmatrix} & A & B\\ A & 0.72 & 0.28\\ B & 0.36 & 0.64\end{pmatrix}$$

有了**状态矩阵**，我们可以轻松得出以下结论：
初始状态 A，2 次运动后状态为 A 的概率是 0.72；
初始状态 A，2 次运动后状态为 B 的概率是 0.28；
初始状态 B，2 次运动后状态为 A 的概率是 0.36；
初始状态 B，2 次运动后状态为 B 的概率是 0.64；

- 有了概率矩阵，即便求运动 n 次后的各种概率，也能非常方便求出。

来看一个多个状态更复杂的情况：




**状态转移矩阵的稳定性**

**状态转移矩阵**有一个非常重要的特性，经过一定有限次数序列的转换，最终一定可以得到一个**稳定的**概率分布 ，且与初始状态概率分布**无关**。例如：
假设我们当前股市的概率分布为：，[0.3,0.4,0.3] ，即 30% 概率的牛市，40% 概率的熊盘与 30% 的横盘。然后这个状态作为序列概率分布的初始状态t0 ，将其带入这个状态转移矩阵计算t1,t2,t3,... 的状态。

```
import numpy as np
matrix = np.matrix([[0.9, 0.075, 0.025],
                    [0.15, 0.8, 0.05],
                    [0.25, 0.25, 0.5]],dtype=float)

vector1 = np.matrix([[0.3, 0.4, 0.3]], dtype=float)
for i in range(100):
    vector1 = vector1 * matrix
    print('Courrent round: {}'.format(i+1))
    print(vector1)
```

输出结果：

```
Current round: 1
[[ 0.405   0.4175  0.1775]]
Current round: 2
[[ 0.4715   0.40875  0.11975]]
Current round: 3
[[ 0.5156  0.3923  0.0921]]
Current round: 4
[[ 0.54591   0.375535
0.078555]]
…         …           …
Current round: 58
[[ 0.62499999  0.31250001
0.0625    ]]
Current round: 59
[[ 0.62499999  0.3125
0.0625    ]]
Current round: 60
[[ 0.625   0.3125  0.0625]]
…         …            …
Current round: 99
[[ 0.625   0.3125  0.0625]]
Current round: 100
[[ 0.625   0.3125  0.0625]]
```

可以发现，从第 60 轮开始，我们的状态概率分布就不变了，一直保持[0.625,0.3125,0.0625] ，即 62.5% 的牛市，31.25% 的熊市与 6.25% 的横盘。这个性质不仅对状态转移矩阵有效，对于绝大多数的其他的马尔可夫链模型的状态转移矩阵也有效。同时不光是离散状态，连续状态时也成立。