

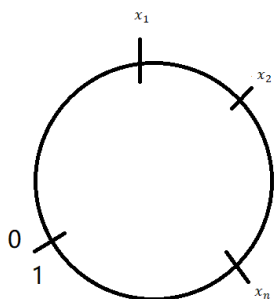
概率论笔记-19

516030910581 刘畅 only-changer@sjtu.edu.cn

例题一：

随机变量 X 满足 $x \in U[0,1]$ ，且有 x_1, x_2, \dots, x_n 为独立同分布，我们对 X 中每一项进行排序，有 $0 \leq x_1 \leq x_2 \leq \dots \leq x_n \leq 1$ ，这 n 个随机变量把 $[0,1]$ 分成了 $n+1$ 份，则每段长度的均值为？

解：



本题等价于在一个周长为 1 的圆上取 $n+1$ 个点，则有 $n+1$ 条线段的均值为 $\frac{1}{n+1}$ ，即每条线段的均值为 $\frac{1}{n+1}$

思考一：

如果变量 X 满足 $x \in U[M]$ ，且有 x_1, x_2, \dots, x_n 为独立同分布，我们仍然对 X 中每一项进行排序，有 $1 \leq x_1 \leq x_2 \leq \dots \leq x_n \leq m$ ，这 n 个随机变量把 $[1,m]$ 分成了 $n+1$ 份，求每段长度的均值？还可以用同样的方法做吗？

答案是否定的，此时的随机变量 X 有可能存在两个点大小相等，因此不再可以像第一题那样进行问题转换。

思考二：

如果变量 X 满足分布函数 $f(x)$ ，且有 x_1, x_2, \dots, x_n 为独立同分布，我们仍然对 X 中每一项进行排序，有 $0 \leq x_1 \leq x_2 \leq \dots \leq x_n \leq 1$ ，这 n 个随机变量把 $[0,1]$

分成了 $n+1$ 份，求每段长度的均值？（设累计分布函数为 $F(x)$ ）

解：

$$\begin{aligned}
 f_k(x)dx &= P(x \in dx) \\
 &= p(\text{一个 } x \in dx, \text{且恰有 } k-1 \text{ 个其他点 } < x') \\
 &= np(x_1 \in dx) * p(\text{恰有 } k-1 \text{ 个其他点 } < x') \\
 &= n * f(x)dx * C_{n-1}^{k-1} * F(x)^{k-1} * (1 - F(x))^{n-k} \\
 &\text{即有 } k-1 \text{ 个点比 } x \text{ 小, } n-k \text{ 个点比 } x \text{ 大}
 \end{aligned}$$

$$\begin{aligned}
 \text{有 } f(x) &= \begin{cases} 1, x \in [0,1] \\ 0, x \notin [0,1] \end{cases} \\
 F(x) &= \begin{cases} 0, x \in (-\infty, 0) \\ x, x \in [0,1] \\ 1, x \in (1, +\infty) \end{cases}
 \end{aligned}$$

期望公式为

$$\begin{aligned}
 E(x_k) &= \int_0^1 f_k(x)dx \\
 &= \int_0^1 n * C_{n-1}^{k-1} * x^k * (1-x)^{n-k} dx \\
 &= n * C_{n-1}^{k-1} * \int_0^1 x^k * (1-x)^{n-k} dx \\
 &= n * C_{n-1}^{k-1} * \frac{(k-1)! * (n-k)!}{n!} \\
 &= \frac{k}{n+1}
 \end{aligned}$$

所以每段长度的期望为

$$E(x_k) - E(x_{k-1}) = \frac{1}{n+1}$$

证毕，同时我们也验证了之前转换成圆的做法的正确性。

思考三：如果 X 为离散的，比如 $X_1, X_2, X_3, X_4 \in [6]$,且有 $x_1 \leq x_2 \leq x_3 \leq x_4$

试求 $E(x_2 - x_3)$

$$\begin{aligned}
 \text{解：} E(x_2 - x_3) &= \sum_{i=1}^5 P(x_2 \leq i, x_3 \geq i+1) \\
 &= \sum_{i=1}^5 C_4^2 P(x_1, x_2 \leq i, x_3, x_4 \geq i+1)
 \end{aligned}$$

$$= \frac{C_4^2}{6^4} * \sum_{i=1}^5 i^2 * (6-i)^2$$

$$= \frac{259}{216}$$

习题一：

接上题，若 $Y_0 = 1, Y_1 = X_1, \dots, Y_4 = X_4, Y_5 = 6$

求 1. $\max E(Y_i - Y_{i-1}) = ?$

2. $E(\max(Y_i - Y_{i-1})) = ?$

习题二：设点集 $V = \{0,1\}^k$ 即为长度为 k 的 01 串

设 T 为抛硬币模型下 01 串第一次出现的时间

假设两点 AB ， AB 有边当且仅当 $P(T^A < T^B) > \frac{1}{2}$

求 $\max P(T^A < T^B)$

习题三：

对于马尔科夫链 $(b, a_1, a_2, a_3, \dots, a_{k-1})$

$\exists b$, 使得 $(b, a_1, a_2, a_3, \dots, a_{k-1}) \rightarrow (a_1, a_2, a_3, \dots, a_k)$

且满足任何一点有入邻居。

该问题即著名的 De Bruijn 问题。

例题二：

关于马尔科夫链及其平稳分布，在此用矩阵重新证明。

首先给出一些定义

首次访问时间 $m_{ij} = E(\text{从 } i \text{ 出发第一次经过 } j \text{ 的时间})$

M 为 m_{ij} 构成的矩阵

首次回归时间 $r_i = E(\text{从 } i \text{ 出发第一次回到 } i \text{ 的时间})$

R 为 以 r_i 为对角元构成的矩阵

平稳分布： $W = (w_1, w_2, \dots, w_k)$

基于以上这些定义，一定有： $r_i = \frac{1}{w_i}$

证明：

先证明引理： $(I - P)M = J - R$ （其中 I 为单位矩阵，J 为全 1 矩阵）

利用全概率公式，我们可得：

$$\begin{aligned} m_{ij} &= p_{ij} + \sum_{k \neq j} p_{ik}(m_{kj} + 1) \\ &= 1 + \sum_{k \neq j} p_{ik}m_{kj} \quad \text{---①} \end{aligned}$$

（可以理解成要从 i 到 j，要经过 k，因此枚举 k 即可）

$$r_i = \sum_{k \neq i} p_{ik}(m_{ki} + 1) \quad \text{---②}$$

（可以理解成要从 i 绕回 i，要经过 k，因此枚举 k 即可）

由上述两式可得

$$M = J - R + P * M \quad \text{---③}$$

证明技巧：对于不在对角线上的点，利用①式；

在对角线上的点，利用①②相加即可。

③式经过移项易得原式

$$(I - P)M = J - R \quad \text{---④}$$

又因为 W 为平稳分布，有： $W(I - P) = 0$

④式两边同乘 W 可得

$$W(J - R) = W(I - P)M = 0$$

所以 $W(J - R) = 0$

$$\text{即 } r_i = \frac{1}{w_i}$$

证毕。

例题三：

在马尔科夫链中，证明 $(I - P + W)$ 可逆

证明：

$$\text{由平稳分布的定义： } W = \lim_{n \rightarrow \infty} P^n$$

$$\text{则 } (I - (P - W))^{-1} = I + (P - W) + (P - W)^2 + \cdots + (P - W)^n$$

$$\text{又因为 } (P - W)^n = P^n - W \rightarrow 0$$

所以 $(I - P + W)$ 可逆

注：除了这种直接根据 W 的定义证明外，还可以通过 $(I - P + W)x = 0$

推出 $x = 0$ 证明。

例题四：

在马尔科夫链中，定义基本矩阵 $Z = (I - P + W)^{-1}$

$$\text{证明 } m_{ij} = \frac{z_{jj} - z_{ij}}{w_j} = r_j(z_{jj} - z_{ij})$$

证明：

由于刚才已经证明了 $r_i = \frac{1}{w_i}$ ，因此只要证明第一个等号即可

先证明引理： $ZJ = J$

由于 $PJ = J$; $WJ = J$;

所以 $J = (I - P + W)J$

故 $ZJ = J$

我们之前证明了 $(I - P)M = J - R$

所以 $Z(I - P)M = Z(J - R) = J - ZR$ ---①

又因为 $I - P = I - P + W - W - PW + W^2 = (I - P + W)(I - W)$

即 $I - P = (I - P + W)(I - W)$

两边同时*Z 有：

$$Z(I - P) = (I - W)$$

将①代入，有 $(I - W)M = J - ZR$

即 $M = J - ZR + WM$

具体到矩阵上的点，有：

$$m_{ij} = 1 - z_{ij}r_j + (WM)_j$$

为了计算 $(WM)_j$ ，我们取 $i = j$ 的情况，由 m 的定义 $m_{jj} = 0$

$$0 = m_{jj} = 1 - z_{jj}r_j + (WM)_j$$

整理得：

$$m_{ij} = r_j(z_{jj} - z_{ij})$$

证毕。

思考题四：

对于不可约马尔科夫链

证明： $\sum_{j < i} m_{ij}w_j = \sum_{j < i} z_{jj} - 1$

例题五：

马尔科夫中心极限定理：

对于 $r < s$

有
$$P\left(r < \frac{S_j^n - w_j * n}{\sqrt{n * \sigma_j^2}} < s\right) = \frac{1}{\sqrt{2\pi}} * \int_r^s e^{-\frac{x^2}{2}} dx$$

$$\sigma^2 = 2\,w_j\,z_{jj} - {w_j}^2$$