概率论与数理统计

第一章

Jenny Zhang 2021

随机试验

随机试验的3个特征

1. 可在相同条件下重复
2. 能事先知道所有的试验结果 → 样本空间，记作S (sample)
3. 进行一次试验之前不能确定哪个结果会出现

ex. 硬币投掷试验：

我们能事先知道试验结果-非正即反；我们能重复投掷硬币-重复10次或100次；我们在投掷之前并不知道会是正面还是反面。

任何符合这三个特征的试验都叫做**随机试验**或**试验**（experiment）记作**E**！

样本空间与随机事件

一些基本定义

- 样本空间：记作S；试验中所有可能的结果组成的集合。

  举例：S_coin:{head, tail}; S_dice:{1,2,3,4,5,6}

- 样本点：也叫基本事件；试验的每个结果。

  举例：head是一个基本事件，tail是一个基本事件

- 随机事件或事件：记作A；样本空间的子集。

  举例：A_dice = S_dice1: {1，2} 是一个事件 — 投掷结果小于3

- 事件发生：一个子集中的样本点发生，则事件发生。

  举例：投掷结果是1时，我们称事件A_dice发生。

若我们把S看成一个事件，那么S是必然事件，因为S包含了该试验的所有样本点。

而Ø是不可能事件。

事件间的关系

假设两个事件A和B

- 互斥事件：

  也叫互不相容事件；

  A与B不可能同时发生；

  集合表达：$A \cap B = \emptyset$

  举例：掷骰子：A: {1,2}, B: {3,4}, C: {5,6}为互斥事件。

- 对立事件：

  也叫逆事件

  满足两个条件：互斥；A与B中必有一事件发生。

  集合表达：$A \cup B = S \quad and \quad A \cap B = \emptyset$

  举例：掷骰子：A: {1,2,3}, B: {4,5,6}为对立事件。

频率与概率

- 频率

记作$f_n(A)$

- 概率

记作$P(A)$，易证 $\lim_{n\to\infty} f_n(A) = P(A)$，因此我们常用概率来表示事件A在一次试验中发生的概率大小。

概率的性质：（略）

等可能概型/古典概型

等可能概型的2个特征

- 样本空间S中样本点有限，设为n；
- 试验中每个基本事件e_i（样本点）发生可能性相同；

$\because$ 每个基本事件发生可能性相同

$$\therefore\ np(e_i) = 1 \quad where \quad 1 \le i \le n$$

$$\therefore\ p(e_i) = \frac{1}{n}$$

设 A事件中包含k个基本事件，那么A的概率计算公式为

$$kp(e_i) = \frac{k}{n}$$

注：看等可能概型概型的例子

条件概率
条件概率公式

- 条件概率的定义：

$$p(B|A)=\frac{p(AB)}{p(A)},$$

其中A,B是两个事件；

$P(A)>0$（分母不能为零）；

$p(B|A)$为事件A发生的条件下事件B发生的条件概率。

$A\cap B$ 记作 AB

注：条件概率仍然符合概率的性质。

- 等可能概型的条件概率：

设 样本空间中基本事件总数为n，A事件中包含基本事件个数为m，AB包含基本事件个数为k，

则

$$p(B|A)=\frac{\frac{k}{n}}{\frac{m}{n}}=\frac{k}{m}=\frac{p(AB)}{p(A)}$$

条件概率

性质1 — 可列可加性

- 可列可加性：

case 1. 当B_1, B_2两事件互不相容是时：

$$p(B_1 \cup B_2 | A) = p(B_1 | A) + p(B_2 | A)$$

case 2. 更general的case，对任意B_1, B_2两事件：

$$p(B_1 \cup B_2 | A) = p(B_1 | A) + p(B_2 | A) - p(B_1 B_2 | A)$$

条件概率

性质2 – 乘法定理

- 乘法定理：

$p(A_1A_2\ldots A_n)=p(A_n|A_1A_2\ldots A_{n-1})p(A_{n-1}|A_1A_2\ldots An-2)\ldots p(A_2|A_1)p(A_1)$，其中，$A_1,A_2,\ldots,A_n$为$n$个事件，$n\geq 2$，且 $p(A_1A_2\ldots A_{n-1})>0$

注：此定理可从条件概率公式的变形推得。

条件概率

全概率公式及其证明

- 全概率公式：A事件的概率为

$p(A) = p(A \mid B_1)p(B_1) + p(A \mid B_2)p(B_2) + \ldots + p(A \mid B_n)p(B_n)$，其中$B_1, B_2, \ldots, B_n$是样本空间$S$的一个划分，且 $p(B_i) > 0 \quad where \quad 0 \leq i \leq n$。

注：划分：$B_1, B_2, \ldots, B_n$为互不相容事件，且$B_1 \cup B_2 \cup \ldots \cup B_n = S$

证

$A = A \cap S = AS = A(B_1 \cup B_2 \cup \ldots \cup B_n) = AB_1 \cup AB_2 \cup \ldots \cup AB_n$，其中$AB_i$，$AB_j$互不相容。

$\overset{***}{\rightarrow} \quad p(A) = p(AB_1) + p(AB_2) + \ldots + p(AB_n)$

$***$ 根据可列可加性

条件概率

贝叶斯公式及其证明

- 贝叶斯公式：

$$p(B_i|A)=\frac{p(A|B_i)p(B_i)}{\sum_{j=1}^{n}p(A|B_j)p(B_j)}$$

证

分母为全概率公式，分母 $=p(A)$；

分子由条件概率得：分子$p(AB_i)$；

由条件概率得：

$$p(B_i|A)=\frac{p(AB_i)}{p(A)}$$

条件概率

常用的全概率及贝叶斯公式：当 $n = 2$

将B_1记作B，B_2记作$\overline{B}$，则

- 全概率公式为

$$p(A) = p(A \mid B)p(B) + p(A \mid \overline{B})p(\overline{B})$$

- 贝叶斯公式为

$$p(B \mid A) = \frac{p(A \mid B)p(B)}{p(A \mid B)p(B) + p(A \mid \overline{B})p(\overline{B})}$$

事件的独立性

与条件概率做比较，看看有什么不同?

若A, B两事件满足：

$$p(AB) = p(A)p(B)\text{，则}A, B\text{事件独立。} \qquad \ldots \qquad (1)$$

若A, B两事件独立（互不影响），则：

$$p(B\,|\,A) = p(B)\text{，带入条件概率公式可得（1）式}$$

举例：见书同时投掷两枚硬币。

注意：若 $p(A) > 0,\quad p(B) > 0$，则显然 $p(AB) \neq 0$，那么$AB \neq \varnothing$, 因此相互独立的事件A, B不是互不相容的，不是互斥的，而是必有一定概率会同时发生的。

举例：现实生活中，甲在A地感染了病毒，乙在B地感染了病毒，A、B两地相差甚远，互不影响，那么我们称此为独立事件。用通俗的话讲，一个事件的发生不会影响另一个事件的发生，则称这两事件**相互独立**。虽然是独立事件，又必然有一定机率会同时发生。

看书中：两个定理及n个事件相互独立的概念。

第二章 随机变量及其分布

Jenny Zhang 2021

随机变量

随机变量的定义

- $X = X(e)$ 是**随机变量**，是定义在样本空间S (其中$S = \{e\}$）的实值单值函数。

离散随机变量

定义以及分布率

- 离散随机变量：随机变量取值为有限个或可列无限个。
- 分布率：用于掌握离散随机变量X的统计规律。
- 必须且仅需：
- 1、所有可能取值：$x_k(k=1,2,\ldots)$
- 2、每个可能值的概率：p_k
- 即需要知道离散型随机变量X的分布率：
- $P\{X=x_k\}=p_k,(k=1,2,\ldots)$.
- 其中，$\sum_{k=1}^{\infty}(p_k)=1$，可理解为1以一定的规律分布在各个可能值上，也是为什么上是叫做分布率的原因。我们先称X符合某分布。

离散随机变量

3种重要的离散随机变量（一）：（0-1）分布

- $(0-1)$分布的定义：

  随机变量X取值为$0,1$；

  分布率为：$$P\{X=k\}=p^k(1-p)^{1-k}, k=0,1 \quad (0<p<1)$$

- 应用：所有样本空间S只有两个元素/样本点/基本事件的随机试验，即$S=\{e_1,e_2\}$。

- 符合$(0-1)$分布的离散随机变量定义为：

$$X=X(e)=\begin{cases} 0, when \quad e=e_1 \\ 1, when \quad e=e_2 \end{cases},$$ 用于表达随机试验结果。

  我们称X符合$(0-1)$分布。

离散随机变量

3种重要的离散随机变量（二）：二项分布

- 二项分布的定义：

随机变量X取值为$0,1,\ldots,n$；

分布率为：

$$P\{X=k\}=\binom{n}{k}p^k(1-p)^{n-k}\overset{\text{q=1−p}}{=}\binom{n}{k}p^k(q)^{n-k},k=0,1,\ldots,n$$

其中，$\displaystyle\sum_{k=0}^{n}P\{X=k\}=\sum_{k=0}^{n}\binom{n}{k}p^k(q)^{n-k}=(p+q)^n=1$