概率论与数理统计

1 样本空间与概率

1. 德摩根定律:

$$(\bigcup_n S_n)^c = \bigcap_n S_n^c, \quad (\bigcap_n S_n)^c = \bigcup_n S_n^c$$

2. 概率模型的基本构成:

(1) 样本空间 Ω：一个试验的所有可能结果的集合

(2) 概率律：试验结果的集合 A（事件）确定一个非负数 $P(A)$（事件 A 的概率）

3. 条件概率：给定 B 发生之下事件 A 的条件概率

$$P(A|B) = \frac{P(A \bigcap B))}{P(B)}$$

乘法规则：

$$P(\bigcap_{i=1}^{n} A_i) = P(A_1)P(A_2|A_1)P(A_3|A_1 \bigcap A_2) \ldots P(A_n| \bigcap_{i=1}^{n-1} A_i)$$

4. 全概率定理：设 $A_1, A_2, ..., A_n$ 是一组互不相容的事件，它形成样本空间的一个分割（每个实验结果必定使得其中一个事件发生），又假定对每一个 i，$P(A_i) > 0$，则对于任何事件 B，下列公式成立：

$$\begin{aligned} P(B) &= P(A_1 \cap B) + P(A_2 \cap B) + \cdots + P(A_n \cap B) \\ &= P(A_1)P(B|A_1) + \cdots + P(A_n)P(B|A_n) \end{aligned}$$

5. 贝叶斯准则：设 $A_1, A_2, ..., A_n$ 是一组互不相容的事件，它形成样本空间的一个分割（每个实验结果必定使得其中一个事件发生），又假定对每一个 i，$P(A_i) > 0$，则对于任何事件 B，只要满足 $P(B) > 0$，下列公式成立：

$$P(A_i|B) = \frac{P(A_i)P(B|A_i)}{P(B)} = \frac{P(A_i)P(B|A_i)}{P(A_1)P(B|A_1) + \cdots + P(A_n)P(B|A_n)}$$

6. 独立性:

(1) 事件 A, B 相互独立：

$$P(A \cap B) = P(A)P(B)$$

(2) 若 A, B 相互独立，则 A 与 B^c 也相互独立

(3) 设事件 C 满足 $P(C) > 0$，A 和 B 在给定 C 的条件下条件独立：

$$P(A \cap B|C) = P(A|C)P(B|C)$$

(4) 独立性不蕴含条件独立性，反之亦然

7. 计数的乘法准则：考虑由 r 个阶段组成的一个试验假设：
(1) 在第 1 阶段有 n_1 个可能的结果
(2) 对于第 1 阶段的任何一个结果，在第 2 阶段有 n_2 个可能的结果
(3) 一般地，在前 $r-1$ 的阶段的任何一个结果，在接下来的第 r 阶段有 n_r 个结果
则在 r 个阶段的试验中一共有 $n_1 n_2 \ldots n_r$ 个试验结果
8. 计数法：
(1)n 个对象的排列数：$n!$
(2)n 个对象中取 k 个对象的排列数：$\frac{n!}{(n-k)!}$
(3)n 个对象中取 k 个对象的组合数：$\mathrm{C}_n^k = \frac{n!}{k!(n-k)!}$
(4) 将 n 个对象分成 r 个组的分割数，其中第 i 组具有 n_i 个对象：$\binom{n}{n_1,n_2,\ldots,n_r} = \frac{n!}{n_1!n_2!\ldots n_r!}$

2 离散随机变量

1. 随机变量：在一个试验的概率模型之下：
(1) 随机变量是试验结果的实值函数
(2) 随机变量的函数定义了另一个随机变量
(3) 可以在某事件或某随机变量的条件下定义一个随机变量
(4) 存在一个随机变量与某事件或某随机变量相互独立的概念
2. 离散随机变量：在一个试验的概率模型之下：
(1) 离散随机变量的取值只能是有限多个值或可数无限多个值
(2) 离散随机变量的函数也是一个离散随机变量，它的分布列可以从原随机变量的分布列得到
3. 分布列：
(1) 伯努利随机变量：

$$p_X(k) = \begin{cases} p, & k=1 \\ 1-p, & k=0 \end{cases}$$

(2) 二项随机变量：$p_X(k) = P(X=k) = \mathrm{C}_n^k p^k (1-p)^{n-k}, \quad k=0,1,\ldots,n$

$$\sum_{k=0}^{n} \mathrm{C}_n^k p^k (1-p)^{n-k} = 1$$

(3) 几何随机变量：$p_X(k) = (1-p)^{k-1} p, \quad k=1,2,\ldots$

$$\sum_{k=1}^{\infty} (1-p)^{k-1} p = 1$$

(4) 泊松随机变量：

$$p_X(k) = e^{-\lambda} \frac{\lambda^k}{k!}, \quad k=0,1,2,\ldots$$

$$\sum_{k=0}^{\infty} e^{-\lambda} \frac{\lambda^k}{k!} = 1$$

4. 期望：

$$E[X]=\sum_{n} x p_X(x)$$

随机变量的函数的期望规则：设随机变量 X 的分布列为 p_X，又设 $g(x)$ 是 X 的一个函数，则

$$E[g(x)]=\sum_{x} g(x)p_X(x), \quad E[x^n]=\sum_{x} x^n p_X(x)$$

5. 方差：

$$var(X)=E[(X-E[X])^2]=\sum_{x}(x-E[X])^2 p_X(x)=E[X^2]-(E[X])^2$$

随机变量的线性函数的均值和方差：设 X 为随机变量，令 $Y=aX+b$

$$E[Y]=aE[X]+b, \quad var(Y)=a^2 var(X)$$

6. 联合分布列：设 X 和 Y 为在某个试验中的随机变量

(1)X 和 Y 的联合分布列：

$$p_{X,Y}(x,y)=P(X=x,Y=y)$$

(2)X 和 Y 的边缘分布列：

$$p_X(x)=\sum_{y} p_{X,Y}(x,y), \quad p_Y(y)=\sum_{x} p_{X,Y}(x,y)$$

(3)X 和 Y 的函数 $g(X,Y)$ 是一个随机变量：

$$E[g(X,Y)]=\sum_{x}\sum_{y} g(x,y)p_{X,Y}(x,y)$$

若 g 是线性的，且 $g=aX+bY+c$，则 $E[aX+bY+c]=aE[X]+bE[Y]+c$

7. 独立随机变量：

设在某一试验中，A 是一个事件，满足条件 $P(A)>0$，又设 X 和 Y 是在同一个试验中的两个随机变量

(1) 称 X 相对与事件 A 独立，如果满足 $p_{X|A}(x)=p_X(x)$ 对一切成立，即对一切 x，事件 $\{X=x\}$ 与 A 相互独立

(2) 称 X 和 Y 为相互独立的随机变量，如果对一切可能的数对 (x,y)，事件 $\{X=x\}$ 和 $\{Y=y\}$ 相互独立，则 $p_{X,Y}(x,y)=p_x(x)p_Y(y)$ 对一切 x,y 成立

(3) 若 X 和 Y 相互独立，则 $E[XY]=E[X]E[Y]$ 对于任意函数 g 和 h，随机变量 $g(X)$ 和 $h(Y)$ 也是相互独立的

$$E[g(X)h(Y)]=E[g(X)]E[h(Y)]$$

(4) 若 X 和 Y 相互独立，则 $var(X+Y)=var(X)+var(Y)$

3 连续随机变量

1. 概率密度函数：对于随机变量 X，若存在一个非负函数 f_X，使得

$$P(X \in B) = \int_B f_X(x)\mathrm{d}x$$

对每一个实数轴上的集合 B 都成立，则称 X 为连续的随机变量，函数 f_X 就称为 X 的概率密度函数，当 B 是一个区间的时候

$$P(a \leqslant X \leqslant b) = \int_a^b f_X(x)\mathrm{d}x$$

区间的端点对于概率的计算不起作用，即

$$P(a \leqslant X \leqslant b) = P(a < X < b) = P(a \leqslant X < b) = P(a < X \leqslant b)$$

$$\int_{-\infty}^{\infty} f_X(x)dx = P(-\infty < X < \infty) = 1$$

设 δ 是一个充分小的正数，则 $P([x, x+\delta]) \approx f_X(x) \cdot \delta$

2. 期望和方差：记 X 为连续随机变量，相应的概率密度函数 $f_X(x)$

(1) X 的期望

$$E[X] = \int_{-\infty}^{\infty} x f_X(x)dx$$

(2) 随机变量 $g(X)$ 的期望

$$E[g(X)] = \int_{-\infty}^{\infty} g(x) f_X(x)dx$$

(3) X 的方差

$$var(X) = E[(X - E[X])^2] = \int_{-\infty}^{\infty} (x - E[X])^2 f_X(x)dx$$

(4) 关于方差，公式成立 $\quad 0 \leqslant var(X) = E[X^2] - (E[X])^2$

(5) 设 $Y = aX + b$，则 $E[Y] = aE[X] + b, \quad var(X) = a^2 var(X)$

3. 分布函数：$F_X(x) = P(X \leqslant x)$

(1) $F_X(x)$ 是 x 的单调非减函数，若 $x \leqslant y$，则 $F_X(x) \leqslant F_Y(y)$

(2) $x \to -\infty, F_X(x) \to 0; \quad x \to \infty, F_X(x) \to 1$

(3) X 是离散随机变量时，$F_X(x)$ 为阶梯函数

(4) X 是连续随机变量时，$F_X(x)$ 为 x 的连续函数

(5) 当 X 是离散随机变量并且取整数值时

$$F_X(k) = \sum_{i=0}^{k} p_X(i)$$

$$p_X(k) = P(X \leqslant k) - P(X \leqslant k-1) = F_X(k) - F_X(k-1)$$

(6) 当 X 是连续随机变量时

$$F_X(x) = \int_{-\infty}^{x} f_X(t)\mathrm{d}x \quad f_X(x) = \frac{\mathrm{d}F_X}{\mathrm{d}x}(x)$$

4. 正态随机变量：一个连续随机变量 X 称为正态的或高斯的，若它的概率密度函数为

$$f_X(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

期望和方差

$$E[X] = \mu \quad var(X) = \sigma^2$$

线性变换之下，随机变量的正态性保持不变，设 $Y = aX + b$

$$E[Y] = aE[X] + b = a\mu + b \quad var(Y) = var(aX + b) = a^2\sigma^2$$

5. 标准正态随机变量：设正态随机变量 Y 的期望为 0，方差为 1，则 Y 称为标准正态随机变量

$$\Phi(y) = P(Y \leqslant y) = P(Y < y) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{y} e^{-\frac{t^2}{2}} \mathrm{d}t$$

利用标准正态分布表计算正态随机变量 X 的分布函数
(1) 将 X 标准化，得到标准正态分布随机变量 $Y = \frac{X-\mu}{\sigma}$
(2) 从标准正态分布表查得

$$P(X \leqslant x) = P(\frac{X-\mu}{\sigma} \leqslant \frac{x-\mu}{\sigma}) = P(Y \leqslant \frac{x-\mu}{\sigma}) = \Phi(\frac{x-\mu}{\sigma})$$

6. 联合概率密度函数：令 X 和 Y 为联合连续随机变量，其联合概率密度函数为 $f_{X,Y}(x,y)$

$$P((X,Y) \in B) = \iint_{(x,y)\in B} f_{X,Y}(x,y)\mathrm{d}x\mathrm{d}y$$

(1) X 和 Y 的边缘概率密度函数为

$$f_X(x) = \int_{-\infty}^{\infty} f_{(x,y)}(x,y)\mathrm{d}y, \quad f_Y(y) = \int_{-\infty}^{\infty} f_{(x,y)}(x,y)\mathrm{d}x$$

(2) 联合分布函数：$F_{X,Y}(x,y) = P(X \leqslant x, Y \leqslant y)$ ，在联合概率密度函数的连续点上

$$f_{X,Y}(x,y) = \frac{\partial^2 F_{X,Y}}{\partial x \partial y}(x,y)$$

(3) X 和 Y 的函数 $g(X,Y)$ 定义了一个新的随机变量

$$E[g(X,Y)] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(X,Y) f_{X,Y}(x,y)\mathrm{d}x\mathrm{d}y$$

若 g 是一个线性函数 $aX + bY + c$ ，则

$$E[aX + bY + c] = aE[X] + bE[Y] + c$$