概率统计任务 Part1,2

Yu Meng

2024 年 8 月 9 日

1 随机事件与概率

1.1 一些基本概念

样本空间 Ω：所有试验的可能集合，其中每一个可能的结果叫元素（样本点），用 ω 表示.

事件：样本空间的子集（如掷骰子为偶数）.

用集合元素关系描述：

- 事件发生为 $\omega \in A$，事件不发生为 $\omega \notin A$.
- A 发生必然导致 B 发生：$\forall \omega \in A$，都有 $\omega \in B$，即 $A \subset B$.

并事件：A 发生或 B 发生（$A \cup B$）.

交事件：A 发生且 B 发生（$A \cap B$ 或 AB）.

互不相容/互斥：A 发生时 B 一定不发生。特别的，当 $A \cup B$ 为全集时，A、B 为对立事件，记作 $B = \overline{A}$.

差事件：$A - B = A \cap \overline{AB} = A\overline{B}$

1.2 事件的运算法则

1. 交换律：$A \cap B = B \cap A$，$A \cup B = B \cup A$.
2. 结合律：$(A \cap B) \cap C = A \cap (B \cap C)$，$(A \cup B) \cup C = A \cup (B \cup C)$（必须是同种符号之间）.
3. 分配律：$(A \cap B) \cup C = (A \cup C) \cap (B \cup C)$，$(A \cup B) \cap C = (A \cap C) \cup (B \cap C)$.
4. De-Morgan(德摩根) 法则（对偶律）：$\overline{A \cap B} = \overline{A} \cup \overline{B}$，$\overline{A \cup B} = \overline{A} \cap \overline{B}$.

推广：$\overline{\bigcap_{i=1}^{n} A_i} = \bigcup_{i=1}^{n} \overline{A_i}$，$\overline{\bigcup_{i=1}^{n} A_i} = \bigcap_{i=1}^{n} \overline{A_i}$.

1.3 概型

等可能概型：样本空间中的每一个样本点等可能地发生.

性质：(1)$0 \leqslant P(\omega_i) \leqslant 1$，(2)$P(\omega_1) = P(\omega_2) = ... = P(\omega_n)$，(3)$P(\Omega) = 1 = P(\omega_1) + P(\omega_2) + ... + P(\omega_n)$.

此时，当样本点个数有限，为古典概型；当样本空间是某区域（一维直线，二维平面，三维空间等），为几何概型.

1.4 其他

求对立事件的概率：$P(\overline{A}) = 1 - P(A)$.

注意：P=0 不一定为不可能事件，P=1 不一定为必然事件.

1.5 概率的统计定义

频率和频数：称 $f_n(A) = \frac{n_A}{n}$ 为事件 A 在 n 次重复试验中出现的频率，其中 n_A 为频数.

频率稳定性：在大量重复试验中，某个事件发生的频率具有稳定性.

频率的统计定义：n 次重复试验中事件 A 发生的频率 $f_n(A)$ 随着 n 的增大将稳定到某个常数（实际应用时，往往把频率当概率使用）.

1.6 概率的公理化定义

概率的公理化定义：给定一个随机试验，Ω 是它的样本空间，对于任意一个事件 A，规定一个实数，记作 P(A)，如果 P(.) 满足下列三条公理，那么就称 P(A) 为事件 A 的概率.

1. 非负性：对于任意一个事件 A，$P(A) \geqslant 0$.
2. 规范性：$P(\Omega) = 1$.
3. 可列可加性：当可列无限个事件 $A_1, A_2, ...$ 两两互不相容时，$P(A_1 \cup A_2 \cup ...) = P(A_1) + P(A_2) + ...$

1.7 条件概率

条件概率：对于任意 2 个事件 A，B，其中 $P(B) > 0$，称 $P(A|B) = \frac{P(AB)}{P(B)}$ 为在已知事件 B 发生的条件下事件 A 的条件概率.

概率的乘法公式：当 $P(A) > 0$ 时，$P(AB) = P(A)P(B|A)$.

推广：当 $n \geqslant 2$ 且 $P(A_1...A_{n-1}) > 0$ 时，$P(A_1...A_n) = P(A_1)P(A_2|A_1)...P(A_n|A_1...A_{n-1})$.

相互独立：$P(AB) = P(A)P(B)$ 成立.

定理：如果 $P(A) > 0$，那么，事件 A 与 B 相互独立 1 充分必要条件是 P(B|A)=P(B).

相互独立的推广：对于任意 3 个事件 A，B，C，如果 $P(AB) = P(A)P(B)$，$P(BC) = P(B)P(C)$，$P(AC) = P(A)P(C)$，$P(ABC) = P(A)P(B)P(C)$ 都成立，那么，称事件 A，B，C 相互独立.

1.8 伯努利试验

Bernoulli(伯努利) 试验：在 1 个试验中只关心某个事件 A 是否发生（通常记为 $P(A) = p(0 < p < 1)$）. 如果独立地重复做 n 次，这 n 个试验合在一起称为 n 重 Bernoulli 试验.

1.9 贝叶斯公式

- 全概率公式：$P(B) = \sum_{i=1}^{n} P(A_iB) = \sum_{i=1}^{n} P(A_i)P(B|A_i)$.
- 贝叶斯公式：$P(A_i|B) = \frac{P(A_i)P(B|A_i)}{\sum_{i=1}^{n} P(A_i)P(B|A_i)}, i = 1, 2, ...$

2 离散型随机变量

2.1 随机变量与分布律

（一维）随机变量：对 Ω 中的每一个样本点 ω，有一个实数 $X(\omega)$ 与它对应.

随机变量 $X = X(\omega)$，定义域为 Ω，值域为 $\Omega_X \subset (-\infty, +\infty)$.

分布律：称表 $P(X = a_i) = p_i, i = 1, 2, ...$ 为随机变量 X 的概率（质量）函数（或分布律）.

2.2 二项分布

二项分布：在 n 重伯努利试验中，设随机变量 X 为 n 次试验中事件 A 发生的次数.

则 $\Omega_X = \{0, 1, ..., n\}$，X 的概率函数为 $P(X = k) = \begin{pmatrix} n \\ k \end{pmatrix} p^k (1-p)^{n-k}, k = 0, 1, ..., n$，记 $X \sim B(n, p)$.

2.3 二维随机变（向）量

定义：对 Ω 的每一个样本点 ω，有一对有序实数对 $(X(\omega), Y(\omega))$ 与之对应.

值域 $\Omega_{(X,Y)} = \{(a_i, b_j) : i = 1, 2, ..., j = 1, 2, ...\}$.

分布律：称 $P(X = a_i, Y = b_j) = P(\{X = a_i\} \cap \{Y = b_j\}) = P_{ij}, i, j = 1, 2, ...$ 为二维随机变量 (X, Y) 的概率（质量）函数（或分布律），或称为联合概率函数.

边缘概率函数：若 $P(X = a_i, Y = b_j) = P(\{X = a_i\} \cap \{Y = b_j\}) = P_{ij}, i, j = 1, 2, ...$，则满足 $P(X = a_i) = \sum_j P(X = a_i, Y = b_j) = \sum_j P_{ij} = P_{i.}, i = 1, 2, ...$ 的所有 $P(X)$ 称为 X 的边缘概率分布.

随机变量的独立性：如果联合概率函数恰为两个边缘概率函数的乘积，即 $P_{ij} = P_{i.} \cdot P_{.j}$ 对一切 $i, j = 1, 2, ...$ 成立，那么称随机变量 X 与 Y 相互独立.

条件概率函数：对任意一个固定的 j，j=1,2,...，称 $P(X = a_i | Y = b_j) = \frac{P_{ij}}{P_{.j}}, i = 1, 2, ...$ 为已知 $\{Y = b_j\}$ 发生的条件下 X（记作 "$X | Y = b_j$"）的条件概率.

3 连续型随机变量及其分布

3.1 连续型随机变量

连续型随机变量：值域是一个区间（或若干区间的并）.

分布函数：(关心随机变量落在某个区间内的概率，单调递增) 给定一个随机变量 X，称 $F(x) = P(X \leqslant x), -\infty < x < +\infty$ 为 x 的分布函数.

对任意随机变量 X，$F(x)$ 在 $x = x_0$ 处连续的充分必要条件为 $P(X = x_0) = 0$.

对于 n 维随机变量 $(X_1, X_2, ..., X_n)$，分布函数 $F(x_1, x_2, ..., x_n) = P(X_1 \leqslant x_1, X_2 \leqslant x_2, ..., X_n \leqslant x_n)$.

概率密度函数：给定一个连续型随机变量 X，如果存在一个定义域为 $(-\infty, +\infty)$ 的非负实值函数 $f(x)$，使得 $F(x) = \int\limits_{-\infty}^{x} f(t)dt$，则称 $f(x)$ 为 X 的概率密度函数.

性质：(1)$f(x) \geqslant 0$；(2)$\int\limits_{-\infty}^{+\infty} f(x)dx = 1$.

连续型随机变量的性质：

1. $F(x)$ 是一个连续函数，且当 $f(x)$ 在 $x = x_0$ 处连续时，$F'(x_0) = f(x_0)$.
2. 对任意一个常数 c，$P(X = c) = 0$.

3. 对任意两个常数 a, b，$P(a < x \leqslant b) = \int\limits_a^b f(x)dx$.

3.2 连续型随机变量的分布

3.2.1 连续型均匀分布

$$f(x) = \begin{cases} \frac{1}{b-a} & ,a < x < b \\ 0 & ,\text{其余} \end{cases}$$

3.2.2 指数分布（无记忆性）

$$f(x) = \begin{cases} \lambda e^{-\lambda x} & ,x > 0 \\ 0 & ,x \leqslant 0 \end{cases}$$

记作 $X \sim E(\lambda)$.

3.2.3 高斯分布（正态分布）

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, -\infty < x < +\infty$$

记作 $X \sim N(\mu, \sigma^2)$.

标准正态分布：$\mu = 0$ 且 $\sigma = 1$，记作 $\phi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}, -\infty < x < +\infty$，其分布函数记作 $\Phi(x)$.

正态概率计算公式：$P(a < X \leqslant b) = \Phi(\frac{b-\mu}{\sigma}) - \Phi(\frac{a-\mu}{\sigma})$.

3.2.4 二维连续型随机变量

联合分布：给定二维连续型随机变量 (X, Y)，如果存在一个定义域为整个平面的二元非负实值函数 $f(x, y)$，使得 (X, Y) 的分布函数 $F(x, y)$ 可以表达成 $F(x, y) = \int\limits_{-\infty}^{x} \int\limits_{-\infty}^{y} f(u, v)dudv, -\infty < x, y < +\infty$, 那么称 $f(x, y)$ 为 (X, Y) 的（概率）密度函数（或分布），或者称为 X 与 Y 的联合分布.

二维连续型均匀分布：

$$f(x, y) = \begin{cases} \frac{1}{G\text{的面积}} & ,(x, y) \in G \\ 0 & ,\text{其余} \end{cases}$$

二维正态分布的密度函数：

$$f(x, y) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \exp\Big\{-\frac{1}{2(1-\rho^2)}\Big[\frac{(x-\mu_1)^2}{\sigma_1^2} - 2\rho\frac{(x-\mu_1)(y-\mu_2)}{\sigma_1\sigma_2} + \frac{(y-\mu_2)^2}{\sigma_2^2}\Big]\Big\}, -\infty < x, y < +\infty$$

记作 $(X, Y) \sim N(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho)$，其中 $-\infty < \mu_1, \mu_2 < +\infty, \quad \sigma_1, \sigma_2 > 0, \quad |\rho| < 1$.

边缘分布函数：称 $F_X(x) = F(x, \infty), -\infty < x < +\infty$ 为 X 的边缘分布函数，称 $F_Y(y) = F(y, \infty), -\infty < y < +\infty$ 为 X 的边缘分布函数. 计算公式：

$$F_X(x) = P(X \leqslant x, -\infty < Y < +\infty) = \int\limits_{-\infty}^{x} \Big\{\int\limits_{-\infty}^{+\infty} f(x, y)dy\Big\}dx$$

边缘（概率）密度函数：$f_X(x)=\int\limits_{-\infty}^{+\infty}f(x,y)dy,-\infty<x<+\infty$ 为 X 的边缘（概率）密度函数（或边缘分布），同理有 $f_Y(y)=\int\limits_{-\infty}^{+\infty}f(x,y)dx,-\infty<y<+\infty$.

当 X 与 Y 相互独立时，有：

- $P(X\in S_1,Y\in S_2)=P(X\in S_1)P(Y\in S_2)$
- $P(X\leqslant x,Y\leqslant y)=P(X\leqslant x)P(Y\leqslant y)$

X 与 Y 相互独立的定义：$F(x,y)=F_X(x)F_Y(y)$，对一切 $-\infty<x,y<+\infty$ 都成立.

推广：若 $F(x_1,...,x_n)=\prod_{i=1}^n F_{X_i}(x_i),-\infty<x_1,...,x_n<+\infty$，则 $X_1,...,X_n$ 相互独立.

补充：也可以是 $f(x,y)=f_X(x)f_Y(y)$ 在 $f(x,y),f_X(x),f_Y(y)$ 的一切公共连续点上成立.

有关相互独立和二维正态分布的定理：设 $(X,Y)\sim N(\mu_1,\mu_2,\sigma_1^2,\sigma_2^2,\rho)$，那么，$X$ 与 Y 相互独立的充分必要条件为 $\rho=0$.

条件（概率）密度函数（或条件分布）：$f_{X|Y}(x|y)=\frac{f(x,y)}{f_Y(y)},-\infty<x<+\infty$，当 $f_Y(y)>0$ 时，称其为已知 $\{Y=y\}$ 发生时 X 的条件分布，同理也有 $f_{Y|X}(y|x)=\frac{f(x,y)}{f_X(x)},-\infty<y<+\infty$.

条件分布函数：同分布函数，对密度函数求积分即可. 有：

- $F_{X|Y}(x|y)=\int\limits_{-\infty}^{x}f_{X|Y}(u|y)du=\int\limits_{-\infty}^{x}\frac{f(u,y)}{f_Y(y)}du,-\infty<x<+\infty$
- $F_{Y|X}(y|x)=\int\limits_{-\infty}^{y}f_{Y|X}(x|v)dv=\int\limits_{-\infty}^{y}\frac{f(x,v)}{f_X(x)}dv,-\infty<y<+\infty$

最后，对于正态分布，有一个可加性：设 X 与 Y 相互独立，当 $X\sim N(\mu_1,\sigma_1^2)$，$Y\sim N(\mu_2,\sigma_2^2)$ 时，$X+Y\sim N(\mu_1+\mu_2,\sigma_1^2+\sigma_2^2)$.

4 随机变量的数字特征

4.1 数学期望

4.1.1 期望的定义

(1) 设离散型随机变量 X 的概率函数为 $P(X=a_i)=p_i,i=1,2,...$，当级数 $\sum_{i=1}^{\infty}|a_i|p_i$ 收敛时，称 $\sum_{i=1}^{\infty}a_ip_i$ 的值为 X 的期望，记作 $E(X)$.

(2) 设连续型随机变量 X 的密度函数为 $f(x)$，当积分 $\int\limits_{-\infty}^{+\infty}|x|f(x)dx$ 收敛时，称 $\int\limits_{-\infty}^{+\infty}xf(x)dx$ 的值为 X 的期望，记作 $E(X)$.

(3) 若上述的级数 $\sum_{i=1}^{\infty}|a_i|p_i$ 或积分 $\int\limits_{-\infty}^{+\infty}|x|f(x)dx$ 发散，称 X 的期望不存在.

4.1.2 期望的性质

性质如下：（注：$P(X=c)=1$ 为 X 服从参数为 c 的退化分布）

1. $E(c)=c$.
2. $E(kX+c)=kE(X)+c$.（期望具有线性性）
3. $E(kX+lY)=kE(X)+lE(Y)$.
4. 当 X 与 Y 相互独立时，$E(XY)=E(X)E(Y)$.

4.2 方差

4.2.1 定义与性质

方差与标准差的定义：设 X 是一个随机变量，称 $D(X)=E\{[X-E(X)]^2\}$ 为 X 的方差，称 $\sqrt{D(X)}$ 为 X 的标准差（或标准偏差）.

上述公式为方差的定义式，一般不用其来计算，而用 $D(X)=E(X^2)-[E(X)]^2$ 来计算方差.

方差的性质：

1. $D(c)=0$，反之，若 $D(X)=0$，则 $P(X=c)=1$，其中 $c=E(X)$.
2. $D(kX+c)=k^2D(X)$.
3. $D(X\pm Y)=D(X)+D(Y)\pm 2E\{[X-E(X)][Y-E(Y)]\}$.
4. 当 X 与 Y 相互独立时，$D(X\pm Y)=D(X)+D(Y)$.

4.2.2 协方差

协方差的定义：$cov(x,Y)=E\{[X-E(X)][Y-E(Y)]\}$ 为 X 与 Y 的协方差.

实际计算时，也往往不采用该公式，而用 $cov(X,Y)=E(XY)-E(X)E(Y)$ 来计算.

协方差的性质：

1. $cov(X,Y)=cov(Y,X)$.
2. $cov(X,c)=0$.
3. $cov(kX,lY)=klcov(X,Y)$.
4. $cov(\sum_{i=1}^m X_i,\sum_{j=1}^n Y_j)=\sum_{i=1}^m X_i\sum_{j=1}^n Y_jcov(X_i,Y_j)$.
5. $cov(X+Y,X-Y)=D(X)-D(Y)$.（了解即可）

4.3 相关系数

相关系数的定义：设 (X,Y) 是一个随机变量，当 $D(X)>0,D(Y)>0$ 时，记相关系数为 $\rho(X,Y)$，有：

$$\rho(X,Y)=E[\frac{X-E(X)}{\sqrt{D(X)}}\cdot\frac{Y-E(Y)}{\sqrt{D(Y)}}]=\frac{cov(X,Y)}{\sqrt{D(X)D(Y)}}$$

相关系数的性质：

1. $\rho(X,Y)=\rho(Y,X)$.
2. $|\rho(X,Y)|\leqslant 1$.
3. $|\rho(X,Y)=1|$ 的充分必要条件是：存在不为零的常数 k 与常数 c，使得 $P(Y=kX+c)=1$.

线性相关：当 $\rho(X,Y)=1$ 时，X 与 Y 正线性相关；当 $\rho(X,Y)=-1$ 时，X 与 Y 负线性相关；当 $\rho(X,Y)=0$ 时，X 与 Y 不相关.（注：相关和相互独立有区别）

4.4 p 分位数

p 分位数：当 $0<p<1$ 时，对任意随机变量 X，如果实数 c 满足：$\begin{cases} P(X \leqslant c) \geqslant p \\ P(X \geqslant c) \geqslant 1-p \end{cases}$，那么，称 c 为 X（或 X 所服从的分布）的 p 分位数，记作 v_p.

特别的，当 $p=\frac{1}{2}$ 时，c 称为中位数；对连续型随机变量，成立 $P(X \leqslant v_p)=p$.

4.5 众数

众数的定义：

(1) 当 X 为离散型随机变量时，假定 $P(X=a_i)=p_i, i=1,2,...$，如果存在 $a^* \in \Omega=a_1,a_2,...$，使得 $P(X=a^*) \geqslant P(X=a_i)$ 对一切 $i=1,2,...$ 均成立，那么称 a^* 为 X（或 X 所服从的分布）的众数.

(2) 当 X 为连续型随机变量时，假定 X 的密度函数为 $f(x)$，如果存在实数 x^*，使得 $f(x^*) \geqslant f(x)$ 对一切 $-\infty<x<+\infty$ 均成立，那么称 x^* 为 X（或 X 所服从的分布）的众数.