

信息学中的概率统计

王若松

前沿计算研究中心
北京大学

多维离散随机变量

1. 多维随机变量的定义
2. 多维离散随机变量的概率分布列
3. 多维离散随机变量的独立性
4. 多维离散随机变量的特征数
5. 多维离散随机变量函数的分布

1. 多维随机变量的定义

► 多维随机变量

- 球与桶模型中，每个桶中球的数量
- 随机图模型中，每个人认识的人数
- 同一台服务器上的 n 块硬盘，每一块硬盘在一年中的故障次数

- 定义：对于样本空间 $S = \{e\}$ ，若 $X_1(e), X_2(e), \dots, X_n(e)$ 为 n 个定义在 S 上的随机变量，则称 $X(e) = (X_1(e), X_2(e), \dots, X_n(e))$ 为 **n 维随机变量或随机向量**
- 若 X_1, X_2, \dots, X_n 均为离散随机变量，则称 $X = (X_1, X_2, \dots, X_n)$ 为 **n 维离散随机变量或离散随机向量**

2. 多维离散随机变量的概率分布列

- ▶ 考虑二维随机变量 (X, Y) , X 全部可能取值为 x_1, x_2, x_3, \dots , Y 全部可能取值为 y_1, y_2, y_3, \dots , 则 $p_{ij} = P(X = x_i, Y = y_j)$ 为 (X, Y) 的**联合分布列**
 - ▶ 正则性: $\sum_i \sum_j p_{ij} = 1$
 - ▶ 非负性: $p_{ij} \geq 0$
- ▶ 例: 随机变量 X 等概率取 $\{1, 2, 3, 4\}$ 中的一个, 随机变量 Y 等概率取 $\{1, \dots, X\}$ 中的一个。求 (X, Y) 的联合分布列

X	Y			
	1	2	3	4
1	1/4	0	0	0
2	1/8	1/8	0	0
3	1/12	1/12	1/12	0
4	1/16	1/16	1/16	1/16

2. 多维离散随机变量的概率分布列

- ▶ 给定 (X, Y) 的联合分布列
- ▶ 如何求 X 和 Y 各自的分布列?
- ▶ $P(X = x_i) = \sum_j P(X = x_i, Y = y_j)$
- ▶ $P(Y = y_j) = \sum_i P(X = x_i, Y = y_j)$
- ▶ 上述分布列被称为 X 和 Y 的**边际分布列**
- ▶ 例：随机变量 X 等概率取 $\{1, 2, 3, 4\}$ 中的一个，随机变量 Y 等概率取 $\{1, \dots, X\}$ 中的一个。求 X 和 Y 的边际分布列

2. 多维离散随机变量的概率分布列

- ▶ 给定 (X, Y) 的**联合分布列**
- ▶ 给定 $Y = y_j$, 求 $X = x_i$ 的概率?
 - ▶ $P(X = x_i | Y = y_j) = P(X = x_i, Y = y_j) / P(Y = y_j)$
- ▶ 若 $P(Y = y_j) > 0$, 则称 $p_{i|j} = P(X = x_i | Y = y_j) = P(X = x_i, Y = y_j) / P(Y = y_j)$ 为给定 $Y = y_j$ 条件下 X 的**条件分布列**
- ▶ 若 $P(X = x_i) > 0$, 则称 $p_{j|i} = P(Y = y_j | X = x_i) = P(X = x_i, Y = y_j) / P(X = x_i)$ 为给定 $X = x_i$ 条件下 Y 的**条件分布列**
- ▶ 例: 随机变量 X 等概率取 $\{1, 2, 3, 4\}$ 中的一个, 随机变量 Y 等概率取 $\{1, \dots, X\}$ 中的一个。求给定 $Y = 2$ 条件下 X 的条件分布列

2. 多维离散随机变量的概率分布列

- ▶ 例：两台服务器间有一条有故障的网络连接，每次通讯有 $1 - p$ 的概率出现故障， p 的概率通讯成功。重复进行通讯直到成功通讯两次。
- ▶ X 表示首次通讯成功时的通讯次数， Y 表示第二次通讯成功时的通讯次数。
- ▶ 求联合分布列、边际分布列和条件分布列
- ▶ $P(X = m, Y = n) = p^2(1 - p)^{n-2}, \quad 1 \leq m < n, \quad n \geq 2$
- ▶ $P(X = m) = p(1 - p)^{m-1}$
- ▶ $P(Y = n) = p^2(1 - p)^{n-2} \cdot (n - 1)$
- ▶ $P(X = m | Y = n) = 1/(n - 1)$
- ▶ $P(Y = n | X = m) = p(1 - p)^{n-m-1}$

3. 多维离散随机变量的独立性

- ▶ 给定二维随机变量 (X, Y) , 若对于任意实数 x, y 均有 $P(X = x, Y = y) = P(X = x) \cdot P(Y = y)$, 则称 X, Y **相互独立**
- ▶ 例1: 随机变量 X 等概率取 $\{1, 2, 3, 4\}$ 中的一个, 随机变量 Y 等概率取 $\{1, \dots, X\}$ 中的一个。判断 X, Y 是否相互独立
- ▶ 例2: 球与桶模型中, 随机变量 X_i 表示第 i 个桶中球的数量。判断 X_i 与 X_j 是否相互独立
- ▶ 例3: 随机图模型中, 随机变量 X_i 表示第 i 个人认识的人数。判断 X_i 与 X_j 是否相互独立

3. 多维离散随机变量的独立性

- ▶ 给定 n 维随机变量 (X_1, X_2, \dots, X_n) , 若对于任意 n 个实数 x_1, x_2, \dots, x_n 均有
$$P(X_1 = x_1, X_2 = x_2, \dots, X_n = x_n) = P(X_1 = x_1) \cdot P(X_2 = x_2) \cdot \dots \cdot P(X_n = x_n),$$
则称 X_1, X_2, \dots, X_n **相互独立**
- ▶ X_1, X_2, \dots, X_n 相互独立是否蕴含 X_1, X_2, \dots, X_n 两两独立?
- ▶ X_1, X_2, \dots, X_n 两两独立是否蕴含 X_1, X_2, \dots, X_n 相互独立?

3. 多维离散随机变量的独立性

- ▶ 若 X_1, X_2, \dots, X_n 相互独立且服从相同的分布, 则称 X_1, X_2, \dots, X_n **独立同分布**
- ▶ 例1: 若 $Y \sim B(n, p)$, 则 $Y = X_1 + X_2 + \dots + X_n$, X_1, X_2, \dots, X_n 独立同分布
 - ▶ $X_i \sim B(1, p)$
- ▶ 例2: 若 $Y \sim NB(r, p)$, 则 $Y = X_1 + X_2 + \dots + X_r$, X_1, X_2, \dots, X_r 独立同分布
 - ▶ $X_i \sim G(p)$

3. 多维离散随机变量的独立性

- ▶ 随机变量 $X \sim \pi(\lambda)$ 表示一段时间内进入商店的顾客数量。每个顾客购买某商品的概率为 p ，且不同顾客是否购买均独立。
- ▶ 随机变量 Y 表示进入商店且购买商品的顾客数量，随机变量 Z 表示进入商店且没有购买商品的顾客数量。

- ▶ 判断 Y 与 Z 是否相互独立

$$\text{▶ } P(Y = k) = \sum_{m \geq k} P(Y = k | X = m) \cdot P(X = m)$$

$$\text{▶ } = \sum_{m \geq k} \binom{m}{k} p^k (1-p)^{m-k} \cdot \frac{\lambda^m}{m!} \cdot e^{-\lambda}$$

$$\text{▶ } = \frac{(\lambda p)^k e^{-\lambda}}{k!} \sum_{m \geq k} (1-p)^{m-k} \cdot \frac{\lambda^{m-k}}{(m-k)!} = \frac{(\lambda p)^k}{k!} e^{-\lambda} \cdot e^{(1-p)\lambda} = \frac{(\lambda p)^k e^{-\lambda p}}{k!}$$

$$\text{▶ } P(Z = k) = \frac{(\lambda(1-p))^k e^{-\lambda(1-p)}}{k!}$$

3. 多维离散随机变量的独立性

$$\blacktriangleright P(Y = k) = \frac{(\lambda p)^k e^{-\lambda p}}{k!}$$

$$\blacktriangleright P(Z = k) = \frac{(\lambda(1-p))^k e^{-\lambda(1-p)}}{k!}$$

$$\blacktriangleright P(Y = k, Z = l) = P(Y = k, X = k + l) = \frac{\lambda^{k+l}}{(k+l)!} e^{-\lambda} \cdot \binom{k+l}{k} p^k (1-p)^l$$

$$\blacktriangleright P(Y = k) \cdot P(Z = l) = \frac{(\lambda p)^k e^{-\lambda p}}{k!} \cdot \frac{(\lambda(1-p))^l e^{-\lambda(1-p)}}{l!} = P(Y = k, Z = l)$$

3. 多维离散随机变量的独立性

- ▶ CPU每天有 p_1 的概率出现故障，硬盘每天有 p_2 的概率出现故障。
- ▶ 每天出现故障均独立。CPU与硬盘出现故障也独立。
- ▶ 求CPU先出现故障的概率？
- ▶ 令随机变量 X 表示 CPU 第一次出现故障的天数， Y 表示硬盘第一次出现故障的天数，则 $X \sim G(p_1)$ ， $Y \sim G(p_2)$ ，且 X 与 Y 独立。
- ▶
$$P(X < Y) = \sum_{k \geq 1} P(X = k) \cdot P(Y > X \mid X = k) = \sum_{k \geq 1} P(X = k) \cdot P(Y > k)$$
- ▶
$$= \sum_{k \geq 1} p_1 (1 - p_1)^{k-1} (1 - p_2)^k$$
- ▶
$$= p_1 (1 - p_2) \sum_{k \geq 1} (1 - p_1)^{k-1} (1 - p_2)^{k-1}$$
- ▶
$$= \frac{p_1 (1 - p_2)}{1 - (1 - p_1)(1 - p_2)}$$

4. 多维随机变量的特征数

- ▶ 回顾：给定离散随机变量 X 和函数 g ，求 $Y = g(X)$ 的数学期望
- ▶ 结论： $E(g(X)) = \sum_i p_i g(x_i)$
- ▶ 给定离散随机变量 X, Y 和函数 g ，求 $Z = g(X, Y)$ 的数学期望
- ▶ 定理： $E(Z) = \sum_i \sum_j P(X = x_i, Y = y_j) \cdot g(x_i, y_j)$
- ▶ 例：随机变量 X 等概率取 $\{1, 2, 3, 4\}$ 中的一个，随机变量 Y 等概率取 $\{1, \dots, X\}$ 中的一个。求 $E(X), E(Y), E(X + Y), E(XY)$

4. 多维随机变量的特征数

- 例：随机变量 X 等概率取 $\{1, 2, 3, 4\}$ 中的一个，随机变量 Y 等概率取 $\{1, \dots, X\}$ 中的一个。求 $E(X), E(Y), E(X + Y), E(XY)$

X	Y			
	1	2	3	4
1	1/4	0	0	0
2	1/8	1/8	0	0
3	1/12	1/12	1/12	0
4	1/16	1/16	1/16	1/16

- $E(X) = \frac{1+2+3+4}{4} = \frac{5}{2}$
- $E(Y) = 1 \cdot \frac{1}{4} + 1 \cdot \frac{1}{8} + 2 \cdot \frac{1}{8} + 1 \cdot \frac{1}{12} + \dots = \frac{7}{4}$
- $E(X + Y) = 2 \cdot \frac{1}{4} + 3 \cdot \frac{1}{8} + 4 \cdot \frac{1}{8} + 4 \cdot \frac{1}{12} + \dots = \frac{17}{4}$
- $E(XY) = 1 \cdot \frac{1}{4} + 2 \cdot \frac{1}{8} + 4 \cdot \frac{1}{8} + 3 \cdot \frac{1}{12} + \dots = 5$

4. 多维随机变量的特征数

- ▶ **数学期望的线性性:** $E(X + Y) = E(X) + E(Y)$
- ▶ **证明:**
 - ▶ $E(X + Y) = \sum_i \sum_j P(X = x_i, Y = y_j) \cdot (x_i + y_j)$
 - ▶ $\sum_i \sum_j P(X = x_i, Y = y_j) \cdot x_i = \sum_i P(X = x_i) \cdot x_i = E(X)$
 - ▶ $\sum_i \sum_j P(X = x_i, Y = y_j) \cdot y_j = \sum_j P(Y = y_j) \cdot y_j = E(Y)$
- ▶ **推广:** $E(X_1 + X_2 + \cdots + X_n) = E(X_1) + E(X_2) + \cdots + E(X_n)$

4. 多维随机变量的特征数

- ▶ **数学期望的线性性:** $E(X_1 + X_2 + \cdots + X_n) = E(X_1) + E(X_2) + \cdots + E(X_n)$
- ▶ 例1: $Y \sim B(n, p)$, 求 $E(Y)$ 和 $E(Y^2)$
- ▶ $Y = X_1 + X_2 + \cdots + X_n$, $X_i \sim B(1, p)$
- ▶ $E(X_i) = p \Rightarrow E(Y) = np$
- ▶ $E(Y^2) = E((X_1 + X_2 + \cdots + X_n)^2) = \sum_i \sum_j E(X_i X_j)$
- ▶ $E(X_i^2) = p$
- ▶ 当 $i \neq j$, $E(X_i X_j) = p^2 \cdot 1 = p^2$
- ▶ $E(Y^2) = n \cdot p + (n - 1)np^2$

4. 多维随机变量的特征数

- ▶ **数学期望的线性性:** $E(X_1 + X_2 + \cdots + X_n) = E(X_1) + E(X_2) + \cdots + E(X_n)$
- ▶ 例2: $Y \sim NB(r, p)$, 求 $E(Y)$
 - ▶ $Y = X_1 + X_2 + \cdots + X_r, \quad X_i \sim G(p)$
 - ▶ $E(X_i) = \frac{1}{p} \Rightarrow E(Y) = r/p$

4. 多维随机变量的特征数

- ▶ 例3：有 n 个人，其中有 m 对人认识。将每个人等概率地加入第一组或第二组，用随机变量 X 表示属于不同组且认识的人的对数。求 $E(X)$
- ▶ $X = \sum_{i=1}^m 1_{A_i}$ 。事件 A_i 表示第 i 对认识的人被分在不同组
- ▶ $P(A_i) = \frac{1}{2} \Rightarrow E(1_{A_i}) = \frac{1}{2} \Rightarrow E(X) = \frac{m}{2}$

4. 多维随机变量的特征数

- ▶ 例4：有 n 个有编号的球，将他们随机打乱，使得每种排列均等概率出现。令随机变量 X 表示位置没有改变的球的数量。求 $E(X)$
- ▶ $X = \sum_{i=1}^n 1_{A_i}$ 。事件 A_i 表示第 i 个球的位置没有改变。
- ▶ $P(A_i) = \frac{1}{n} \Rightarrow E(1_{A_i}) = \frac{1}{n} \Rightarrow E(X) = 1$

4. 多维随机变量的特征数

- ▶ 例5：有 n 个有编号的球，有放回地取 m 次。令随机变量 X 表示取到的球中不同编号的数量。求 $E(X)$
- ▶ $X = \sum_{i=1}^n 1_{A_i}$ 。事件 A_i 表示编号为 i 的球被取出过。
- ▶ $P(A_i) = 1 - \left(1 - \frac{1}{n}\right)^m \Rightarrow E(1_{A_i}) = 1 - \left(1 - \frac{1}{n}\right)^m \Rightarrow E(X) = n \cdot \left(1 - \left(1 - \frac{1}{n}\right)^m\right)$

4. 多维随机变量的特征数

- ▶ 例6：有 n 个有编号的球，每次等概率随机取一个，直到所有的球至少被取到过一次。令随机变量 X 表示取球的总次数。求 $E(X)$
- ▶ 如何将 X 写成 $X_1 + X_2 + \cdots + X_n$?
- ▶ X_i 表示在取到 $i - 1$ 个不同的球的基础上，取到 i 个不同的球所用的**额外次数**
- ▶ $X_2 \sim G\left(\frac{n-1}{n}\right)$
- ▶ $X_3 \sim G\left(\frac{n-2}{n}\right)$
- ▶ $X_i \sim G\left(\frac{n-i+1}{n}\right), E(X_i) = \frac{n}{n-i+1}$
- ▶ $E(X) = \sum_i E(X_i) = n \cdot \left(\frac{1}{n} + \frac{1}{n-1} + \frac{1}{n-2} + \cdots + 1\right) = \Theta(n \log n)$

4. 多维随机变量的特征数

- ▶ $E(X + Y) = E(X) + E(Y)$
- ▶ 是否有 $E(XY) = E(X) \cdot E(Y)$?
- ▶ 定理：若离散随机变量 X 和 Y 相互独立，则有 $E(XY) = E(X) \cdot E(Y)$
- ▶ 证明：
 - ▶ $E(XY) = \sum_i \sum_j P(X = x_i, Y = y_j) \cdot x_i y_j$
 - ▶ 根据独立性, $E(XY) = \sum_i \sum_j P(X = x_i) \cdot x_i \cdot P(Y = y_j) \cdot y_j$
 - ▶ $\sum_i \sum_j P(X = x_i) \cdot x_i \cdot P(Y = y_j) \cdot y_j = (\sum_i P(X = x_i) \cdot x_i) (\sum_j P(Y = y_j) \cdot y_j)$
 - ▶ $E(XY) = (\sum_i P(X = x_i) \cdot x_i) (\sum_j P(Y = y_j) \cdot y_j) = E(X) \cdot E(Y)$
- ▶ 若 $E(XY) = E(X) \cdot E(Y)$ ，是否一定有 X 和 Y 相互独立？

4. 多维随机变量的特征数

- ▶ 定理：若离散随机变量 X 和 Y 相互独立，则有 $E(XY) = E(X) \cdot E(Y)$
- ▶ 推广：若离散随机变量 X_1, X_2, \dots, X_n 相互独立，则有 $E(X_1 X_2 \cdots X_n) = E(X_1) \cdot E(X_2) \cdots E(X_n)$
- ▶ 例： $Y \sim NB(r, p)$ ，求 $E(Y^2)$ 和 $\text{Var}(Y)$
- ▶ $Y = X_1 + X_2 + \cdots + X_r, \quad X_i \sim G(p)$
- ▶ $E(Y^2) = E((X_1 + X_2 + \cdots + X_r)^2) = \sum_i \sum_j E(X_i X_j)$
- ▶ $E(X_i^2) = \frac{2-p}{p^2}$
- ▶ 当 $i \neq j$ ， $E(X_i X_j) = E(X_i)E(X_j) = \frac{1}{p^2}$
- ▶ $E(Y^2) = \frac{r(2-p)}{p^2} + \frac{r(r-1)}{p^2} = \frac{r^2+r-rp}{p^2} \Rightarrow \text{Var}(Y) = \frac{r(1-p)}{p^2}$

4. 多维随机变量的特征数

- ▶ 定理：若离散随机变量 X 和 Y 相互独立，则有 $E(XY) = E(X) \cdot E(Y)$
- ▶ 推广：若离散随机变量 X_1, X_2, \dots, X_n 相互独立，则有 $E(X_1 X_2 \cdots X_n) = E(X_1) \cdot E(X_2) \cdots E(X_n)$
- ▶ 推论：若离散随机变量 X 和 Y 相互独立，则有 $\text{Var}(X \pm Y) = \text{Var}(X) + \text{Var}(Y)$
- ▶ 证明： $\text{Var}(X - Y) = E((X - Y)^2) - (E(X - Y))^2$
- ▶ $E((X - Y)^2) = E(X^2) + E(Y^2) - 2E(XY) = E(X^2) + E(Y^2) - 2E(X) \cdot E(Y)$
- ▶ $(E(X - Y))^2 = (E(X) - E(Y))^2 = (E(X))^2 + (E(Y))^2 - 2E(X) \cdot E(Y)$
- ▶ $\text{Var}(X - Y) = E(X^2) - (E(X))^2 + E(Y^2) - (E(Y))^2 = \text{Var}(X) + \text{Var}(Y)$

4. 多维随机变量的特征数

- ▶ 推论：若离散随机变量 X 和 Y 相互独立，则有 $\text{Var}(X \pm Y) = \text{Var}(X) + \text{Var}(Y)$
- ▶ 推广：若离散随机变量 X_1, X_2, \dots, X_n 相互独立，则有 $\text{Var}(X_1 \pm X_2 \pm \dots \pm X_n) = \text{Var}(X_1) + \text{Var}(X_2) + \dots + \text{Var}(X_n)$
- ▶ 例1： $Y \sim B(n, p)$ ，求 $\text{Var}(Y)$
 - ▶ $Y = X_1 + X_2 + \dots + X_n$ ， $X_i \sim B(1, p)$
 - ▶ $\text{Var}(Y) = n \cdot \text{Var}(X_i) = np(1 - p)$
- ▶ 例2： $Y \sim NB(r, p)$ ，求 $\text{Var}(Y)$
 - ▶ $Y = X_1 + X_2 + \dots + X_r$ ， $X_i \sim G(p)$
 - ▶ $\text{Var}(Y) = r \cdot \text{Var}(X_i) = \frac{r(1-p)}{p^2}$

4. 多维随机变量的特征数

- ▶ 推广：若离散随机变量 X_1, X_2, \dots, X_n 相互独立，则有 $\text{Var}(X_1 \pm X_2 \pm \dots \pm X_n) = \text{Var}(X_1) + \text{Var}(X_2) + \dots + \text{Var}(X_n)$
- ▶ $\text{Var}(X_1 + X_2 + \dots + X_n) = E((X_1 + X_2 + \dots + X_n)^2) - (E(X_1 + X_2 + \dots + X_n))^2$
- ▶ $E((X_1 + X_2 + \dots + X_n)^2) = \sum_i \sum_j E(X_i X_j)$
- ▶ $(E(X_1 + X_2 + \dots + X_n))^2 = (E(X_1) + E(X_2) + \dots + E(X_n))^2 = \sum_i \sum_j E(X_i)E(X_j)$
- ▶ $\text{Var}(X_1 + X_2 + \dots + X_n) = \sum_i \sum_j (E(X_i X_j) - E(X_i)E(X_j))$
- ▶ $E((X_i - E(X_i))(X_j - E(X_j)))$
- ▶ $= E(X_i X_j - E(X_i)X_j - E(X_j)X_i + E(X_i)E(X_j)) = E(X_i X_j) - E(X_i)E(X_j)$
- ▶ $\text{Var}(X_1 + X_2 + \dots + X_n) = \sum_i \sum_j E((X_i - E(X_i))(X_j - E(X_j)))$

4. 多维随机变量的特征数

- ▶ 给定离散随机变量 X 和 Y ，定义 X 和 Y 的**协方差**
- ▶
$$\text{Cov}(X, Y) = E\left((X - E(X))(Y - E(Y))\right) = E(XY) - E(X)E(Y)$$
- ▶ 性质：
 - ▶
$$\text{Cov}(X, X) = \text{Var}(X) = E(X^2) - (E(X))^2$$
 - ▶
$$\text{Cov}(X, Y) = \text{Cov}(Y, X)$$
 - ▶
$$\text{Cov}(aX, bY) = ab \cdot \text{Cov}(X, Y)$$
 - ▶
$$\text{Cov}(X_1 + X_2, Y) = \text{Cov}(X_1, Y) + \text{Cov}(X_2, Y)$$
 - ▶ 若 X 和 Y 相互独立，则 $\text{Cov}(X, Y) = 0$
 - ▶
$$\text{Var}(X_1 + X_2 + \cdots + X_n) = \sum_i \sum_j \text{Cov}(X_i, X_j) = \sum_i \text{Var}(X_i) + 2 \sum_i \sum_{j < i} \text{Cov}(X_i, X_j)$$
 - ▶
$$\text{Var}(X + Y) = \text{Var}(X) + \text{Var}(Y) + 2\text{Cov}(X, Y)$$

4. 多维随机变量的特征数

- ▶ 给定离散随机变量 X 和 Y ，定义 X 和 Y 的**协方差**
- ▶
$$\text{Cov}(X, Y) = E\left((X - E(X))(Y - E(Y))\right) = E(XY) - E(X)E(Y)$$
- ▶ 例1: 随机变量 X 等概率取 $\{1, 2, 3, 4\}$ 中的一个，随机变量 Y 等概率取 $\{1, \dots, X\}$ 中的一个。求 $\text{Cov}(X, Y)$
- ▶
$$\text{Cov}(X, Y) = E(XY) - E(X)E(Y) = 5 - \frac{5}{2} \cdot \frac{7}{4} = \frac{5}{8}$$

4. 多维随机变量的特征数

- ▶ 例2: 随机变量 $X \sim \pi(\lambda)$ 表示一段时间内进入商店顾客数量。每个顾客购买某商品的概率为 p ，且不同顾客是否购买均独立。随机变量 Y 表示进入商店且购买商品的数量。计算 $\text{Cov}(X, Y)$

- ▶ $E(XY) = \sum_{m=1}^{+\infty} \sum_{k=1}^m mk \cdot P(X = m, Y = k)$
- ▶ $= \sum_{m=1}^{+\infty} \sum_{k=1}^m mk \cdot \binom{m}{k} p^k (1-p)^{m-k} \cdot \frac{\lambda^m}{m!} \cdot e^{-\lambda}$
- ▶ $= \sum_{m=1}^{+\infty} \frac{\lambda^m}{m!} \cdot e^{-\lambda} \cdot m \cdot \sum_{k=1}^m k \cdot \binom{m}{k} p^k (1-p)^{m-k}$
- ▶ $\sum_{k=1}^m k \cdot \binom{m}{k} p^k (1-p)^{m-k} = mp$
- ▶ $E(XY) = \sum_{m=1}^{+\infty} \frac{\lambda^m}{m!} \cdot e^{-\lambda} \cdot m \cdot mp = p \cdot E(X^2) = p(\lambda^2 + \lambda)$
- ▶ $\text{Cov}(X, Y) = E(XY) - E(X)E(Y) = p(\lambda^2 + \lambda) - \lambda \cdot \lambda p = p\lambda$

4. 多维随机变量的特征数

- ▶ 例2: 随机变量 $X \sim \pi(\lambda)$ 表示一段时间内进入商店顾客数量。每个顾客购买某商品的概率为 p ，且不同顾客是否购买均独立。随机变量 Y 表示进入商店且购买商品的数量。计算 $\text{Cov}(X, Y)$
- ▶ $\text{Cov}(X, Y) = \text{Cov}(Y + Z, Y) = \text{Var}(Y) + \text{Cov}(Y, Z) = \lambda p$

4. 多维随机变量的特征数

- ▶ 例3：随机图模型中，随机变量 X_i 表示第 i 个人认识的人数。求 $\text{Cov}(X_i, X_j)$
- ▶ 事件 A_{ab} 表示 a 与 b 认识。则 $X_i = \sum_{k \neq i} 1_{A_{ik}}$
- ▶ $P(A_{ab}) = \frac{1}{2} \Rightarrow E(1_{A_{ik}}) = \frac{1}{2} \Rightarrow E(X_i) = \frac{n-1}{2}$
- ▶ 如何求 $E(X_i X_j)$?
- ▶ 当 $i = j$, $\text{Cov}(X_i, X_j) = \text{Var}(X_i) = (n-1)/4$

4. 多维随机变量的特征数

▶ 例3：随机图模型中，随机变量 X_i 表示第 i 个人认识的人数。求 $\text{Cov}(X_i, X_j)$

$$\text{▶ } X_i X_j = \left(\sum_{k \neq i} 1_{A_{ik}} \right) \left(\sum_{l \neq j} 1_{A_{jl}} \right) = \sum_{k \neq i} \sum_{l \neq j} 1_{A_{ik}} \cdot 1_{A_{jl}} = \sum_{k \neq i} \sum_{l \neq j} 1_{A_{ik} \cap A_{jl}}$$

$$\text{▶ } E(1_{A_{ik} \cap A_{jl}}) = P(A_{ik} \cap A_{jl})$$

$$\text{▶ 当 } k = j \text{ 且 } l = i, P(A_{ik} \cap A_{jl}) = P(A_{ij}) = 1/2$$

$$\text{▶ 否则 } P(A_{ik} \cap A_{jl}) = P(A_{ik}) \cdot P(A_{jl}) = 1/4$$

$$\text{▶ } E(X_i X_j) = 1 \cdot \frac{1}{2} + ((n-1)^2 - 1) \cdot \frac{1}{4} = \frac{(n-1)^2}{4} + \frac{1}{4}$$

$$\text{▶ } \text{Cov}(X_i, X_j) = E(X_i X_j) - E(X_i)E(X_j) = \frac{(n-1)^2}{4} + \frac{1}{4} - \frac{(n-1)^2}{4} = \frac{1}{4}$$

4. 多维随机变量的特征数

- ▶ 例3：随机图模型中，随机变量 X_i 表示第 i 个人认识的人数。求 $\text{Cov}(X_i, X_j)$
- ▶ 事件 A_{ab} 表示 a 与 b 认识。则 $X_i = \sum_{k \neq i} 1_{A_{ik}}$
- ▶ 令 $Y = \sum_i X_i$, 则 $Y/2 \sim B\left(\frac{(n-1)n}{2}, \frac{1}{2}\right)$
- ▶ $\text{Var}(Y) = \text{Var}\left(\frac{Y}{2}\right) \cdot 4 = \frac{(n-1)n}{2}$
- ▶ $\text{Var}(Y) = \sum_i \text{Var}(X_i) + \sum_i \sum_{j \neq i} \text{Cov}(X_i, X_j) = \frac{n(n-1)}{4} + \sum_i \sum_{j \neq i} \text{Cov}(X_i, X_j)$
- ▶ 对于 $i \neq j$, 所有 $\text{Cov}(X_i, X_j)$ 均相等
- ▶ $\text{Cov}(X_i, X_j) = \frac{\text{Var}(Y) - \sum_i \text{Var}(X_i)}{n(n-1)} = \frac{(n-1) \cdot n/2 - (n-1) \cdot n/4}{n(n-1)} = 1/4$

4. 多维随机变量的特征数

- ▶ 回顾：若 $P(Y = y_j) > 0$ ，则称 $p_{i|j} = P(X = x_i | Y = y_j) = P(X = x_i, Y = y_j) / P(Y = y_j)$ 为给定 $Y = y_j$ 条件下 X 的**条件分布列**
- ▶ **条件数学期望**： $E(X|Y = y_j) = \sum_i x_i \cdot P(X = x_i | Y = y_j)$
- ▶ 例：随机变量 A 等概率取 $\{1, 2, 3, 4\}$ 中的一个，随机变量 B 等概率取 $\{1, \dots, A\}$ 中的一个。求给定 $A = a$ 条件下 B 和 $2B + 1$ 的条件数学期望
 - ▶ $E(B|A = a) = (a + 1)/2$
 - ▶ $E(2B + 1|A = a) = a + 2$
- ▶ $E(X|Y = y)$ 是关于 y 的函数。改变 y 的取值则得到不同的条件期望
- ▶ 条件期望是条件分布的数学期望，具有数学期望的一切性质
 - ▶ $E(aX_1 + bX_2|Y = y) = a \cdot E(X_1|Y = y) + b \cdot E(X_2|Y = y)$

4. 多维随机变量的特征数

- ▶ 推广：事件 E 满足 $P(E) > 0$ ，则称 $P(X = x_i | E) = P(X = x_i \cap E)/P(E)$ 为给定条件 E 下 X 的**条件分布列**
- ▶ **条件数学期望**： $E(X|E) = \sum_i x_i \cdot P(X = x_i | E)$
- ▶ 例：随机变量 A 等概率取 $\{1,2,3,4\}$ 中的一个，随机变量 B 等概率取 $\{1, \dots, A\}$ 中的一个。求给定 $A \leq 2$ 条件下 B 的条件数学期望
- ▶ $P(B = 1|A \leq 2) = \frac{P(B=1 \cap A \leq 2)}{P(A \leq 2)} = \frac{3}{4}$
- ▶ $P(B = 2|A \leq 2) = \frac{P(B=2 \cap A \leq 2)}{P(A \leq 2)} = \frac{1}{4}$
- ▶ $E(B|A \leq 2) = \frac{5}{4}$

4. 多维随机变量的特征数

- ▶ $E(X|Y = y)$ 是关于 y 的函数。改变 y 的取值则得到不同的条件期望
- ▶ 若将 $E(X|Y = y)$ 记为 $g(y)$, 则 $E(X|Y) = g(Y)$ 是一个随机变量
- ▶ 例：随机变量 A 等概率取 $\{1,2,3,4\}$ 中的一个，随机变量 B 等概率取 $\{1, \dots, A\}$ 中的一个。求随机变量 $E(B|A)$ 的概率分布列和数学期望
 - ▶ $g(a) = E(B|A = a) = (a + 1)/2$
 - ▶ A 等概率取 $\{1,2,3,4\}$ 中的一个
 - ▶ $E(B|A)$ 的数学期望为 $\frac{7}{4}$, 与 $E(B)$ 相同
- ▶ 重期望公式： $E(E(X|Y)) = E(X)$

4. 多维随机变量的特征数

- ▶ 重期望公式: $E(E(X|Y)) = E(X)$
- ▶ 证明: $E(E(X|Y)) = \sum_j P(Y = y_j) \cdot E(X|Y = y_j)$
- ▶ $E(X|Y = y_j) = \sum_i x_i \cdot P(X = x_i|Y = y_j) = \sum_i x_i \cdot P(X = x_i \cap Y = y_j)/P(Y = y_j)$
- ▶ $E(E(X|Y)) = \sum_j P(Y = y_j) \cdot E(X|Y = y_j)$
- ▶ $= \sum_j \sum_i x_i \cdot P(X = x_i \cap Y = y_j) = \sum_i x_i \cdot P(X = x_i) = E(X)$
- ▶ 例: 随机变量 $X \sim \pi(\lambda)$ 表示一段时间内进入商店的顾客数量。每个顾客购买某商品的概率为 p , 且不同顾客是否购买均独立。随机变量 Y 表示进入商店且购买商品的顾客数量。求 $E(Y)$
 - ▶ $E(Y|X = x) = p \cdot x$
 - ▶ $E(Y) = E(E(Y|X)) = p \cdot E(X) = \lambda p$

4. 多维随机变量的特征数

- ▶ 例：随机变量 $X \sim G(p)$ 。证明 $E(X) = p + (1 - p)(E(X) + 1)$
- ▶ 令 $Y = 1_{X=1}$
- ▶ 给定 $Y = 0$ 和 $Y = 1$ 条件下 X 的分布列？
- ▶ $P(X = 1|Y = 1) = P(X = 1|X = 1) = 1$
- ▶ 无记忆性： $P(X > x + 1|X > 1) = P(X > x)$
- ▶ $E(X|Y = 0) = E(X|X \neq 1) = E(X|X > 1)$
- ▶ $E(X|X > 1) = \sum_{k \geq 0} P(X > k|X > 1) = \sum_{k \geq 0} P(X > k - 1) = \sum_{k \geq 0} P(X + 1 > k) = E(X + 1) = E(X) + 1$
- ▶ $E(X) = E(E(X|Y)) = P(Y = 0) \cdot E(X|Y = 0) + P(Y = 1) \cdot 1$
- ▶ $= p + (1 - p)(E(X) + 1)$

4. 多维随机变量的特征数

- ▶ 例： $X_1, X_2 \dots$ 为一列独立同分布的随机变量，随机变量 N 取正整数值且与 $\{X_n\}$ 独立。证明 $E(\sum_{i \leq N} X_i) = E(X_1)E(N)$
- ▶ $E(\sum_{i \leq N} X_i | N = n) = E(\sum_{i \leq n} X_i | N = n) = \sum_{i \leq n} E(X_i | N = n) = n \cdot E(X_1)$
- ▶ $E(\sum_{i \leq N} X_i) = E\left(E(\sum_{i \leq N} X_i | N)\right) = E(N) \cdot E(X_1)$
- ▶ 例：随机变量 $X \sim \pi(\lambda)$ 表示一段时间内进入商店的顾客数量。每个顾客购买某商品的概率为 p ，且不同顾客是否购买均独立。随机变量 Y 表示进入商店且购买商品的顾客数量。求 $E(Y)$
- ▶ $E(Y) = E(X) \cdot p = \lambda p$

5. 多维随机变量函数的分布

- ▶ 回顾：给定离散随机变量 X 和函数 g ，求 $Y = g(X)$ 的分布列
 - ▶ 若 $P(X = x_i) = p_i$ ，则对于 $Y = g(X)$ ，有 $P(Y = g(x_i)) = p_i$
 - ▶ 若某些 $g(x_i)$ 相等，将概率相加
- ▶ 给定离散随机变量 X, Y 和函数 $g(x, y)$ ，求 $Z = g(X, Y)$ 的分布列
- ▶ 例：给定随机变量 X, Y 的联合分布列

X	Y		
	-1	1	2
-1	5/20	2/20	6/20
2	3/20	3/20	1/20

- ▶ 求 $X + Y, X - Y, \max\{X, Y\}$ 的分布列

5. 多维随机变量函数的分布

- 例1：给定随机变量 X, Y 的联合分布列。求 $X + Y, X - Y, \max\{X, Y\}$ 的分布列

X	Y		
	-1	1	2
-1	5/20	2/20	6/20
2	3/20	3/20	1/20

- 对于所有 x_i, y_j , 列出 $P(X = x, Y = y_j)$ 和 $g(x_i, y_j)$

P	5/20	2/20	6/20	3/20	3/20	1/20
(X, Y)	$(-1, -1)$	$(-1, 1)$	$(-1, 2)$	$(2, -1)$	$(2, 1)$	$(2, 2)$
$Z_1 = X + Y$	-2	0	1	1	3	4
$Z_2 = X - Y$	0	-2	-3	3	1	0
$Z_3 = \max\{X, Y\}$	-1	1	2	2	2	2

- 进行合并、整理

5. 多维随机变量函数的分布

- ▶ 例1：给定随机变量 X, Y 的联合分布列
- ▶ 对于所有 x_i, y_j , 列出 $P(X = x, Y = y_j)$ 和 $g(x_i, y_j)$

P	5/20	2/20	6/20	3/20	3/20	1/20
(X, Y)	$(-1, -1)$	$(-1, 1)$	$(-1, 2)$	$(2, -1)$	$(2, 1)$	$(2, 2)$
$Z_1 = X + Y$	-2	0	1	1	3	4
$Z_2 = X - Y$	0	-2	-3	3	1	0
$Z_3 = \max\{X, Y\}$	-1	1	2	2	2	2

$Z_1 = X + Y$	-2	0	1	3	4
P	5/20	2/20	9/20	3/20	1/20

$Z_2 = X - Y$	-3	-2	0	1	3
P	6/20	2/20	6/20	3/20	3/20

$Z_3 = \max\{X, Y\}$	-1	1	2
P	5/20	2/20	13/20

5. 多维随机变量函数的分布

- ▶ 例2: $X \sim B(n, p), Y \sim B(m, p)$, 且 X 与 Y 独立。求 $X + Y$ 的概率分布列
- ▶ $P(X + Y = k) = \sum_{l \leq k} P(X = l) \cdot P(Y = k - l)$
- ▶ $P(X = l) \cdot P(Y = k - l) = \binom{n}{l} \cdot p^l \cdot (1 - p)^{n-l} \cdot \binom{m}{k-l} \cdot p^{k-l} \cdot (1 - p)^{m-(k-l)}$
- ▶ $\sum_{l \leq k} \binom{n}{l} \binom{m}{k-l} = \binom{n+m}{k}$
- ▶ $P(X + Y = k) = \sum_{l \leq k} P(X = l) \cdot P(Y = k - l) = p^k \cdot (1 - p)^{n+m-k} \cdot \binom{n+m}{k}$
- ▶ 推广: 若 $X_1 \sim B(n_1, p), X_2 \sim B(n_2, p), \dots, X_k \sim B(n_k, p)$, 且 X_1, X_2, \dots, X_k 相互独立, 则 $\sum_i X_i \sim B(\sum_i n_i, p)$

5. 多维随机变量函数的分布

- ▶ 例2: $X \sim \pi(\lambda_1), Y \sim \pi(\lambda_2)$, 且 X 与 Y 独立。求 $X + Y$ 的概率分布列
- ▶ $P(X + Y = k) = \sum_{l \leq k} P(X = l) \cdot P(Y = k - l)$
- ▶ $P(X = l) \cdot P(Y = k - l) = \frac{\lambda_1^l}{l!} \cdot e^{-\lambda_1} \cdot \frac{\lambda_2^{k-l}}{(k-l)!} \cdot e^{-\lambda_2}$
- ▶ $= \frac{(\lambda_1 + \lambda_2)^k}{k!} e^{-(\lambda_1 + \lambda_2)} \cdot \binom{k}{l} \cdot \left(\frac{\lambda_1}{\lambda_1 + \lambda_2}\right)^l \cdot \left(\frac{\lambda_2}{\lambda_1 + \lambda_2}\right)^{k-l}$
- ▶ $P(X + Y = k) = \frac{(\lambda_1 + \lambda_2)^k}{k!} e^{-(\lambda_1 + \lambda_2)} \cdot \sum_{l \leq k} \binom{k}{l} \cdot \left(\frac{\lambda_1}{\lambda_1 + \lambda_2}\right)^l \cdot \left(\frac{\lambda_2}{\lambda_1 + \lambda_2}\right)^{k-l}$
- ▶ $= \frac{(\lambda_1 + \lambda_2)^k}{k!} \cdot e^{-(\lambda_1 + \lambda_2)} \cdot \left(\frac{\lambda_1}{\lambda_1 + \lambda_2} + \frac{\lambda_2}{\lambda_1 + \lambda_2}\right)^k = \frac{(\lambda_1 + \lambda_2)^k}{k!} \cdot e^{-(\lambda_1 + \lambda_2)}$
- ▶ 推广: 若 $X_1 \sim \pi(\lambda_1), X_2 \sim \pi(\lambda_2), \dots, X_k \sim \pi(\lambda_k)$, 且 X_1, X_2, \dots, X_k 则相互独立, 则 $\sum_i X_i \sim \pi(\sum_i \lambda_i)$