概率论与数理统计 (Fall 2024)
习题课讲义

2024 年 10 月 14 日

聂嵘琢, 祝国隽

B01GB006Y-02/Note6

---

1. 令随机变量 X 的概率函数为

$$P(X=0)=\frac{1}{3}, P(X=1)=\frac{2}{3}$$

随机变量 Y 的概率函数为

$$P(Y=-1)=P(Y=0)=P(Y=1)=\frac{1}{3}$$

且 $P\left(X^2=Y^2\right)=1$. 求 $(X,Y)^{\mathrm{T}}$ 的联合概率分布和 $Z=X+Y$ 的概率分布.

Ans: $(X,Y)^\top$ 的联合概率分布为

$X\backslash Y$	-1	0	1	行和
0	a	b	c	1/3
1	d	e	f	2/3
列和	1/3	1/3	1/3	1

由于

$$\begin{aligned} 1=P\left(X^2=Y^2\right)&=P(X=0,Y=0)+P(X=1,Y=-1)+P(X=1,Y=1) \\ &=b+d+f \end{aligned}$$

而 $a+b+c+d+e+f=1$ 且每个值非负, 故

$$a=c=e=0$$
$$d=1/3-a=1/3, b=1/3-e=1/3, f=1/3-c=1/3.$$

$(X,Y)^\top$ 的联合概率分布为

$X\backslash Y$	-1	0	1
0	0	1/3	0
1	1/3	0	1/3

$Z = X + Y$ 可取 $-1, 0, 1, 2$, 且

$$
\begin{aligned}
P(Z=-1) &= P(X=0, Y=-1) = 0 \\
P(Z=0) &= P(X=0, Y=0) + P(X=1, Y=-1) = \frac{2}{3} \\
P(Z=1) &= P(X=0, Y=1) + P(X=1, Y=0) = 0 \\
P(Z=2) &= P(X=1, Y=1) = \frac{1}{3}
\end{aligned}
$$

2. 令随机变量 $X \sim \operatorname{Poisson}(\lambda_1), Y \sim \operatorname{Poisson}(\lambda_2)$, 且 X 与 Y 相互独立. 求证:

$$X \mid X+Y=n \sim B\left(n, \lambda_1/\left(\lambda_1+\lambda_2\right)\right), n \geqslant 1$$

Ans: 由题意知: $X + Y \sim \operatorname{Poisson}(\lambda_1 + \lambda_2)$.

$$
\begin{aligned}
P(X=k \mid X+Y=n) &= \frac{P(X=k, Y=n-k)}{P(X+y=n)} \\
&= \frac{\dfrac{\lambda_1^k e-\lambda_1}{k!} \dfrac{\lambda_2^{n-k} e_e-\lambda_2}{(n-k)!}}{\dfrac{(\lambda_1+\lambda_2)^{n_e}-(\lambda_1+\lambda_2)}{n!}} \\
&= \binom{n}{k}\left(\frac{\lambda_1}{\lambda_1+\lambda_2}\right)^k\left(\frac{\lambda_1}{\lambda_1+\lambda_2}\right)^{n-k}
\end{aligned}
$$

3. 令随机变量 X, Y 相互独立, 且 $X \sim \operatorname{Exp}(\lambda), Y \sim \operatorname{Exp}(\zeta), \lambda > 0, \zeta > 0$. 记 $Z = I_{\{X \geqslant Y\}}$. 求 Z 的概率分布.

Ans: 记 $W = X - Y$, 则 W 的密度函数为

$$
\begin{aligned}
f_W(w) &= \int_{-\infty}^{\infty} f_{X,Y}(y+w, y) dy = \int_0^{\infty} f_X(y+w) f_Y(y) dy \\
&= \int_0^{\infty} \lambda e^{-\lambda(y+w)} \delta e^{-\delta y} dy = \lambda \delta e^{-\lambda w} \int_0^{\infty} e^{-(\lambda+\delta) y} dy \\
&= \frac{\lambda \delta}{\lambda+\delta} e^{-\lambda w}
\end{aligned}
$$

$Z = I_{\{X\geq Y\}} = I_{\{X-Y\geq 0\}}$ 的可能取值为 0,1 . 且

$$P(Z=1) = P(X-Y\geq 0) = P(W\geq 0) = \int_0^\infty \frac{\lambda\delta}{\lambda+\delta}e^{-\lambda w}dw = \frac{\delta}{\lambda+\delta}$$
$$P(Z=0) = \frac{\lambda}{\lambda+\delta}$$

4. 令二元随机变量 $(X,Y)^{\mathrm{T}}$ 的联合密度函数为

$$f(x,y) = \begin{cases} \mathrm{e}^{-x}, & 0<y<x, \\ 0, & \text{其他}. \end{cases}$$

求 $f_{Y|X}(y\mid x), f_{X|Y}(x\mid y)$ 和 $P(X\leqslant 1\mid Y\leqslant 1)$.

Ans: (潘天麟同学提供) 直接计算有

$$f_X(x) = \int_{-\infty}^{+\infty} f_{X,Y}(x,y)\mathrm{d}y = \int_0^x \exp(-x)\mathrm{d}y = x\exp(-x)$$
$$f_Y(y) = \int_{-\infty}^{+\infty} f_{X,Y}(x,y)\mathrm{d}x = \int_y^{+\infty} \exp(-x)\mathrm{d}x = \exp(-y), \quad y>0$$

从而由定义可知

$$f_{Y|X}(y\mid x) = \frac{f_{X,Y}(x,y)}{f_X(x)} = \frac{\exp(-x)}{x\exp(-x)} = \frac{1}{x}, \quad x>y>0$$
$$f_{X|Y}(x\mid y) = \frac{f_{X,Y}(x,y)}{f_Y(y)} = \frac{\exp(-x)}{\exp(-y)} = \exp(y-x), \quad x>y>0$$

又因为

$$P(Y\leq 1) = \int_0^1 f_Y(y)\mathrm{d}y = \int_0^1 \exp(-y)\mathrm{d}y = 1-\frac{1}{e}$$
$$P(X\leq 1, Y\leq 1) = \int_0^1\int_0^x \exp(-x)\mathrm{d}y\ \mathrm{d}x = \int_0^1 x\exp(-x)\mathrm{d}x = 1-\frac{2}{e}$$

故有

$$P(X\leq 1\mid Y\leq 1) = \frac{P(X\leq 1, Y\leq 1)}{P(Y\leq 1)} = \frac{1-\frac{2}{e}}{1-\frac{1}{e}} = \frac{e-2}{e-1}$$

5. 令二元随机变量 $(X,Y)^{\mathrm{T}}$ 的联合密度函数为

$$f(x,y)=\begin{cases}\dfrac{xy+1}{4}, & -1\leqslant x\leqslant 1,-1\leqslant y\leqslant 1,\\ 0, & \text{其他}.\end{cases}$$

求证: X,Y 不独立, 但 X^2,Y^2 相互独立.

Ans: 令二元随机变量 $(X,Y)^{\mathrm{T}}$ 的联合密度函数为

$$f(x,y)=\begin{cases}\dfrac{xy+1}{4}, & -1\leqslant x\leqslant 1,-1\leqslant y\leqslant 1,\\ 0, & \text{其他}.\end{cases}$$

求证: X,Y 不独立, 但 X^2,Y^2 相互独立. 证明: x,y 对 $f(x,y)$ 来说地位相等, 则 X,Y 的边缘密度均为

$$f_X(x)=\begin{cases}\int_{-1}^{1}\dfrac{xy+1}{4}dy=\dfrac{1}{2}, & -1\leq x\leq 1,\\ 0, & \text{其他}.\end{cases}$$

由于 $f(x,y)\neq f_X(x)f_Y(y)$, 故 X,Y 不独立.

$$\text{记}U=X^2,V=Y^2.$$

$$\begin{aligned}F_{U,V}(u,v)&=P\left(X^2\leq u,Y^2\leq v\right)\\&=\begin{cases}\int_{-\sqrt{u}}^{\sqrt{u}}\int_{-\sqrt{v}}^{\sqrt{v}}\dfrac{xy+1}{4}dydx=\sqrt{u}\sqrt{v}, & 0\leq u,v\leq 1\\ \int_{-\sqrt{u}}^{\sqrt{u}}\int_{-1}^{1}\dfrac{xy+1}{4}dydx=\sqrt{u}, & 0\leq u\leq 1,v>1\\ \int_{-\sqrt{v}}^{\sqrt{v}}\int_{-1}^{1}\dfrac{xy+1}{4}dxdy=\sqrt{v}, & u>1,0\leq v\leq 1\\ 1, & u,v>1\end{cases}\end{aligned}$$

得 $F_U(u)=\begin{cases}0, & u<0,\\ \sqrt{u}, & 0\leq u\leq 1.\end{cases}$ 可知 $F_{U,V}(u,v)=F_U(u)F_V(v)$, 故 X^2 和 Y^2 相互独立.

6. 令 X,Y 独立同分布, 服从 $N(0,1)$. 求证: X^2+Y^2 与 X/Y 相互独立.

Ans: $\begin{cases} u = x^2 + y^2 \\ v = \dfrac{x}{y} \end{cases} \rightarrow \begin{cases} x = v\sqrt{\dfrac{u}{v^2+1}} \\ y = \sqrt{\dfrac{u}{v^2+1}} \end{cases}$

考虑 $J = \begin{vmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial u}{\partial u} & \frac{\partial y}{\partial v} \end{vmatrix}$, 不妨计算 $J^{-1} = \begin{vmatrix} \frac{\partial u}{\partial x} & \frac{\partial u}{\partial y} \\ \frac{\partial v}{\partial x} & \frac{\partial v}{\partial y} \end{vmatrix} = -2\left(v^2+1\right)$

得 $|J| = \dfrac{1}{2\left(U^2+1\right)}$, 则我们有

$$\begin{aligned} g(u,v) &= f\left(v \cdot \sqrt{\frac{u}{v^2+1}}, \sqrt{\frac{u}{v^2+1}}\right)|J| \\ &= \frac{1}{2\pi}\exp\left[-\frac{1}{2}\left(v^2 \cdot \frac{u}{v^2+1} + \frac{u}{v^2+1}\right)\right\} \cdot \frac{1}{2\left(v^2+1\right)} \\ &= \frac{1}{2\pi} \cdot e^{-\frac{u}{2}} \cdot \frac{1}{2\left(v^2+1\right)} = g_1(u) \cdot g_2(v) \end{aligned}$$

因此 $X^2 + Y^2, X/Y$ 相互独立.

7. 令 $X_1, X_2, \cdots, X_n$ 独立同分布, 服从密度函数为 $f(x) = x\mathrm{e}^{-x}(x > 0)$ 的分布. 记 $Y = X_1 + X_2 + \cdots + X_n$. 求证: Y 的密度函数为 $\dfrac{y^{2n-1}\mathrm{e}^{-y}}{(2n-1)!}, y > 0$.

Ans: 当 $n = 2$ 时, $Y = X_1 + X_2$ 的密度函数为:

$$\begin{aligned} g(y) &= \int_{-\infty}^{\infty} f(y-x, x)dx = \int_0^{\infty} f(y-x)f(x)dx \\ &= \int_0^y (y-x)e^{-(y-x)} \cdot xe^{-x}dx \\ &= e^{-y} \cdot \int_0^y \left(yx - x^2\right)dx \\ &= e^{-y}\left(\frac{yx^2}{2} - \frac{x^3}{3}\right)\Bigg|_{x=0}^{y} \\ &= \frac{y^3e^{-y}}{6} = \frac{y^{2n-1}e^{-y}}{(2n-1)!}, \quad y > 0 \end{aligned}$$

因此, 当 $n = 2$ 时, 结果成立. 假设当 $n = k$ 时, $Y = X_1 + X_2 + \cdots + X_k$ 的密度

函数为:

$$g_k(y) = \frac{y^{2k-1}}{(2k-1)!}e^{-y}, \quad y > 0$$

那么, 对于 $n = k+1, Y = X_1 + X_2 + \cdots + X_{k+1}$ 的密度函数为:

$$\begin{aligned}
g_{k+1}(y) &= \int_{-\infty}^{\infty} g_k(x)f(y-x)dx \\
&= \int_0^y \frac{x^{2k-1}}{(2k-1)!}e^{-x} \cdot (y-x)e^{-(y-x)}dx \\
&= e^{-y} \cdot \int_0^y \left(\frac{x^{2k-1}}{(2k-1)!}y - \frac{x^{2k}}{(2k-1)!}\right)dx \\
&= e^{-y}\left[\left[\frac{yx^{2k}}{(2k)!} - \frac{x^{2k+1}}{(2k-1)!(2k+1)}\right]\right]_{x=0}^{y} \\
&= e^{-y} \cdot y^{2k+1}\left[\frac{1}{(2k)!} - \frac{1}{(2k-1)!(2k+1)}\right] \\
&= \frac{y^{2k+1}}{(2k+1)!}e^{-y}, \quad y > 0
\end{aligned}$$

故当 $n = k+1$ 时, 结果也成立. 由数学归纳法可知, 对 $\forall n \in N_+$, 都有

$$g(y) = \frac{y^{2n-1}}{(2n-1)!}e^{-y}, \quad y > 0.$$

这次作业 Note 补充一下耦合 (Coupling) 的知识, 参考宗语轩同学 (USTC) 的讲义.

阅读材料：耦合 (Coupling)

宗语轩

2024 年 2 月 11 日

我们先看下面一个例子:

例 0.1. 有两个玩家 A, B 投掷两枚两面硬币, 玩家 A 投掷到正面的概率是 0.75, 玩家 B 投掷到正面的概率是 0.5. 现两个玩家独立投掷硬币各 100 枚, 记 N_A, N_B 分别表示玩家 A, B 投掷到正面的个数, 证明: 对任意固定的 $k(0<k<100)$, 均有

$$\mathbb{P}(N_A \geq k) \geq \mathbb{P}(N_B \geq k)$$

问题与思考: 我们当然可以显示地写出两者的概率表达式:

$$\begin{aligned}\mathbb{P}(N_A \geq k) &= \sum_{n=k}^{100} \binom{100}{n} 0.75^n 0.25^{100-n},\\ \mathbb{P}(N_B \geq k) &= \sum_{n=k}^{100} \binom{100}{n} 0.5^n 0.5^{100-n} = \sum_{n=k}^{100} \binom{100}{n} 0.5^{100}.\end{aligned}$$

但是这样的话我们不能比较直观地得到待证命题 (读者可以进行尝试), 而命题直观上看又是对的, 那么我们需要怎么解决呢？这里我们要引入一个耦合 (Coupling) 的技巧:

注意到随机变量的定义 $X:\Omega \to \mathbb{R}$ 是 Borel 实值可测函数, 两个随机变量的分布相同并不代表随机变量相同, 也即不代表他们的样本空间相同 (反例请读者自行思考). 换句话说, 同一个的" 输出端" 对应的" 输入端" 不一定唯一. 借助这个想法, 我们通过匹配合适的样本空间来达到上述目的:

我们进行如下机制: 考察两个玩家分别掷一次硬币的情形. 先让玩家 B 投掷, 若结果为正面, 则令玩家 A 结果也为正面; 若结果为反面, 则令玩家 A 结果为正面的概率为 0.5.

按照上述机制, 我们可以计算玩家 A 投掷到正面的概率亦为 0.75. 我们记 $X_{A,i}$ 表示玩家 A 在第 i 次投掷硬币正面的个数 (结果只可能为 1 或 0), $X_{B,i}$ 表示玩家 B 在第 i 次投掷硬币正面的个数. 由上述机制可知: 对任意 $1 \leq i \leq 100$, 均有 $X_{A,i} \geq X_{B,i}$, 因此

$$\sum_{i=1}^{100} X_{A,i} \geq \sum_{i=1}^{100} X_{B,i} \Rightarrow \mathbb{P}\left(\sum_{i=1}^{100} X_{A,i} \geq k\right) \geq \mathbb{P}\left(\sum_{i=1}^{100} X_{B,i} \geq k\right).$$

而 N_A 和 $\sum_{i=1}^{100} X_{A,i}$ 两者分布相同, N_B 和 $\sum_{i=1}^{100} X_{B,i}$ 两者分布相同. 因此

$$\mathbb{P}(N_A \geq k) \geq \mathbb{P}(N_B \geq k).$$

更一般的证明见下:

---

[0]个人主页: http://home.ustc.edu.cn/~zyx240014/.

发现错误欢迎联系: zyx240014@mail.ustc.edu.cn.

证明. 设随机变量 $Z \sim U[0,1]$, 即 $[0,1]$ 上的均匀分布, 对任意 $1 \le i \le 100$ 引入随机变量 $Y_{A,i}, Y_{B,i}$, 使其相互独立且满足

$$Y_{A,i} = \begin{cases} 1, & 0 \le Z \le 0.75, \\ 0, & 0.75 < Z \le 1. \end{cases}, \quad Y_{B,i} = \begin{cases} 1, & 0 \le Z \le 0.5, \\ 0, & 0.5 < Z \le 1. \end{cases}$$

因此对任意 $1 \le i \le 100$, 均有 $Y_{A,i} \ge Y_{B,i}$, 故

$$\sum_{i=1}^{100} Y_{A,i} \ge \sum_{i=1}^{100} Y_{B,i} \Rightarrow \mathbb{P}\left(\sum_{i=1}^{100} Y_{A,i} \ge k\right) \ge \mathbb{P}\left(\sum_{i=1}^{100} Y_{B,i} \ge k\right).$$

而 N_A 和 $\sum_{i=1}^{100} Y_{A,i}$ 两者分布相同, N_B 和 $\sum_{i=1}^{100} Y_{B,i}$ 两者分布相同. 因此

$$\mathbb{P}(N_A \ge k) \ge \mathbb{P}(N_B \ge k).$$

□

上述例子帮助我们对耦合这个概念有了一个直观地了解, 现在我们给出它的严格定义:

定义 0.1. 概率空间 $\mathcal{X}$ 上两个概率测度 μ, v 的**耦合 (coupling)**, 是指 $\mathcal{X} \times \mathcal{X}$ 上的联合分布 π. 离散性下若随机向量 $(X, Y) \sim \pi$, 则对任意 x, y, 满足

$$\mathbb{P}(X = x) = \mu(x), \qquad \mathbb{P}(Y = y) = \nu(x)$$

注. 联合分布可推出边缘分布, 反之不成立.

考虑有限离散型的情形, μ, ν 的耦合可被矩阵 $q \in \mathbb{R}^{|\mu| \times |\nu|}$ 刻画, 其中 $q(x, y) = \mathbb{P}(X = x, Y = y)$ 满足

$$\sum_y q(x, y) = \mu(x), \qquad \sum_x q(x, y) = \nu(y).$$

注. 耦合的存在性: 对任意 μ, ν, 均存在耦合 $q(x, y) = \mu(x)\nu(y)$, 此时 X 和 Y 独立.

例 0.2. 耦合伯努利分布 $B(0.75)$ 和 $B(0.5)$.

解. 记耦合的矩阵

$$q = \begin{pmatrix} q(0,0) & q(0,1) \\ q(1,0) & q(1,1) \end{pmatrix} := \begin{pmatrix} a & b \\ c & d \end{pmatrix},$$

由定义知,

$$\begin{cases} a + b = 0.25, \\ c + d = 0.75, \\ a + c = 0.5, \\ b + d = 0.5, \end{cases} \Rightarrow q = \begin{pmatrix} a & 0.25 - a \\ 0.5 - a & 0.25 + a \end{pmatrix}, 0 \le a \le 0.25.$$

□

注. **例 0.1** 中用到的耦合方法即为上述取 $a = 0.25$ 的情形.

例 0.3. 耦合伯努利分布 $B(0.5)$ 和 $B(0.5)$.

解. 由定义知,

$$q = \begin{pmatrix} a & 0.5-a \\ 0.5-a & a \end{pmatrix}, 0 \leq a \leq 0.5.$$

- **Case 1:** $a = 0.5$, $q = \begin{pmatrix} 0.5 & 0 \\ 0 & 0.5 \end{pmatrix}$, identical coupling.
- **Case 2:** $a = 0$, $q = \begin{pmatrix} 0 & 0.5 \\ 0.5 & 0 \end{pmatrix}$, negative coupling.
- **Case 3:** $a = 0.25$, $q = \begin{pmatrix} 0.25 & 0.25 \\ 0.25 & 0.25 \end{pmatrix}$, independent coupling.

□

习题. 对 $0 < p \leq q < 1$, 耦合伯努利分布 $B(p)$ 和 $B(q)$, 并给出当矩阵 q 存在元素 0 时的耦合方法.

提示. 参考**例 0.1** 中的证明方法.

参考资料

[1] 王冠扬, 马尔科夫链里的耦合方法, 中科大研究生创新计划高水平和华罗庚科技英才班前沿短课程.
https://wvpn.ustc.edu.cn/http/77726476706e69737468656265737421e7fb4a8869257b447d468ca88d1
video/detail_5722_30386.htm

[2] Markov chains and mixing times, David A. Levin and Yuval Peres, American Mathematical Society, 2017.