第十次习题课

王瑞

北京大学生物统计系

wangruiruishou@pku.edu.cn

May 13, 2022

Outline

随机向量的函数的数学期望

设随机向量 (X, Y) 的函数 $Z = g(X, Y)$ 的数学期望存在，我们一般通过以下的两种来计算数学期望 $E[Z]$:

1. 求出随机变量 Z 的分布，再使用公式 $E[Z] = \int_{\mathcal{R}} z f(z) dz$。其中 $f(z)$ 在这里是 Z 的密度函数。当 (X, Y) 是离散型随机向量时，将积分换成求和，密度函数换成对应的概率分布即可。
2. 直接使用公式

$$E[Z] = E[g(X, Y)] = \int_{\mathcal{R}} \int_{\mathcal{R}} g(x, y) f(x, y) dx dy$$

其中 $f(x, y)$ 是 (X, Y) 的联合密度函数，当 (X, Y) 是离散型随机向量时，将积分换成求和，密度函数换成对应的概率分布即可。

数学期望的进一步性质

1. 对任意两个随机变量 X, Y，如果其数学期望均存在，则 $E(X+Y) = EX + EY$
2. 进一步地，如果 X, Y 独立，那么 $E(XY)$ 存在，且 $E(XY) = EXEY$

例题

习题 1

设随机向量 (X, Y) 的概率分布为：

Table: (X, Y) 的概率分布

X \ Y	0	1
0	0.1	0.15
1	0.25	0.2
2	0.15	0.15

求 $Z = \sin[\frac{\pi}{2}(X+Y)]$ 的数学期望。

例题

习题 1 解答

$$
\begin{aligned}
E[Z] &= E[\sin[\frac{\pi}{2}(X+Y)]] \\
&= \sum_{i=0}^{2}\sum_{j=0}^{1}\sin[\frac{\pi}{2}(i+j)]P(X=i,Y=j) \\
&= 0.1\sin 0 + 0.15\sin\frac{\pi}{2} + 0.25\sin\frac{\pi}{2} + 0.2\sin\pi + 0.15\sin\pi + \\
&0.15\sin\frac{3\pi}{2} \\
&= 0.25
\end{aligned}
$$

例题

习题 2

已知 (X,Y) 为一个二维随机向量, $X_1 = X + Y$， $X_2 = X - Y$，(X_1, X_2) 的密度函数为

$$\phi(x_1, x_2) = \frac{1}{2\sqrt{3}\pi} e^{-\frac{1}{2}\left[\frac{(x_1-4)^2}{3}+(x_2-2)^2\right]}$$

请求出： X 和 Y 的密度函数以及 $E[X_1X_2]$

例题

习题 2 解答

$$\phi(x_1,x_2)=\frac{1}{2\sqrt{3}\pi}e^{-\frac{1}{2}[\frac{(x_1-4)^2}{3}+(x_2-2)^2]}$$
$$=\frac{1}{\sqrt{2\cdot 3\cdot \pi}}e^{-\frac{1}{2}\frac{(x_1-4)^2}{3}}\times\frac{1}{\sqrt{2\cdot \pi}}e^{\frac{(x_2-2)^2}{2}}$$

因此 X_1,X_2 独立，且 $X_1\sim N(4,3)$，$X_2\sim N(2,1)$。所以 $E[X_1X_2]=E[X_1]E[X_2]=8$。注意到 $X=\frac{X_1+X_2}{2},Y=\frac{X_1-X_2}{2}$，因此 $X\sim N(3,1)$，$Y\sim N(1,1)$。

Outline

多个随机变量之间的协方差

设 (X,Y) 为二维随机向量，EX 和 EY 均存在，如果 $E[(X-EX)(Y-EY)]$ 存在，则称之为随机变量 X 和 Y 之间的协方差，记作 $\text{cov}(X,Y)$

和协方差有关的性质：

1. $\text{cov}(X,Y) = EXY - EXEY$
2. $\text{cov}(X,X) = DX$
3. $\text{cov}(X,Y) = \text{cov}(Y,X)$
4. $\text{cov}(aX,bY) = ab\text{cov}(X,Y)$
5. $\text{cov}(X_1+X_2,Y) = \text{cov}(X_1,Y) + \text{cov}(X_2,Y)$
6. $\text{cov}(C,X) = 0$，其中 C 为一常数。
7. 如果 X 和 Y 独立，则 $\text{cov}(X,Y) = 0$。
8. $D(X+Y) = DX + DY + 2\text{cov}(X,Y)$，当 X 与 Y 独立时，$D(X+Y) = DX + DY$

相关系数

将协方差标准化我们便可以得到相关系数:

$$\rho_{X,Y} = \frac{\text{cov}(X, Y)}{\sqrt{DX}\sqrt{DY}}$$

可以证明:

$$|\rho_{X,Y}| \leq 1 \tag{1}$$

例题

习题 3

随机变量 (X, Y) 服从以点 $(0,1),(1,0),(1,1)$ 为顶点的三角形区域上的均匀分布，试求 $E(X+Y)$ 和 $D(X+Y)$

例题

习题 3 解答

记该三角形区域为 D，因为 D 的面积是 1/2，所以 (X, Y) 的联合密度函数是

$$f_{x,y} = 2 \cdot I_D(x, y)$$

这里 $I_D(x, y)$ 的定义是，如果 $(x, y) \in D$,$I_D(x, y) = 1$，否则 $I_D(x, y) = 0$。

而后我们可以求出 X, Y 各自的边缘密度函数：当 $0 < x < 1$ 时，有 $f_X(x) = \int_{1-x}^{1} 2dy = 2x$, 因此 $f_X(x) = 2xI_{(0,1)}(x)$，

$EX = \int_{\mathcal{R}} x \cdot 2xI_{(0,1)}(x)dx = \frac{2}{3}$，$EX^2 = \int_{\mathcal{R}} x^2 \cdot 2xI_{(0,1)}(x)dx = \frac{1}{2}$，

$DX = EX^2 - (EX)^2 = \frac{1}{18}$。同理可以求得 $f_Y(y) = 2yI_{(0,1)}(y)$，

$EY = \frac{2}{3}$，$DY = \frac{1}{18}$。

例题

习题 3 解答

$$EXY = \int_0^1 \int_{1-x}^{x} 2xydydx = \frac{5}{12}$$

由此，$\text{cov}(X, Y) = EXY - EXEY = -\frac{1}{36}$
最后可以求出

$$E(X+Y) = \frac{4}{3} \tag{2}$$

$$D(X+Y) = DX + DY + 2\text{cov}(X, Y) = \frac{1}{18} \tag{3}$$

例题

习题 4

设随机变量 X 和 Y 独立、且都服从参数为 λ 的泊松分布，令 $U = 2X + Y, V = 2X - Y$，求 U 和 V 的相关系数

例题

习题 4 解答

因为

$$D(U)=D(2X+Y)=4DX+DY=5\lambda$$
$$D(V)=D(2X-Y)=4DX+DY=5\lambda$$

所以

$$\begin{aligned}\operatorname{cov}(U,V)&=\operatorname{cov}(2X+Y,2X-Y)\\&=\operatorname{cov}(2X,2X)+\operatorname{cov}(Y,2X)+\operatorname{cov}(2X,-Y)+\operatorname{cov}(Y,-Y)\\&=4DX-DY=3\lambda\end{aligned}$$

所以

$$\rho_{U,V}=\frac{\operatorname{cov}(U,V)}{\sqrt{DUDV}}=\frac{3}{5} \tag{4}$$

例题

习题 5

已知随机变量 X 与 Y 的相关系数为 ρ，求 $X_1 = aX + b$,$X_2 = cY + d$ 的相关系数，其中 a, b, c, d 都是非零常数。

例题

习题 5 解答

按照相关系数的公式，我们应该计算 X_1,X_2 的方差和它们之间的协方差:

$$D(X_1) = a^2 DX$$
$$D(X_2) = c^2 DY$$
$$\mathrm{cov}(X_1, Y_1) = ac\mathrm{cov}(X, Y)$$

因此

$$\rho_{X_1,X_2} = \frac{ac\mathrm{cov}(X,Y)}{\sqrt{a^2DX}\sqrt{c^2DY}} \tag{5}$$

注意 $\rho = \frac{\mathrm{cov}(X,Y)}{\sqrt{DX}\sqrt{DY}}$ 因此当 a, c 同符号时，$\rho_{X_1,X_2} = \rho$，当 a, c 不同符号时，$\rho_{X_1,X_2} = -\rho$。

例题

习题 6

如果连续型随机变量 X 的密度函数 $f(x)$ 是偶函数，且 X 的三阶矩存在，请说明 X 与 X^2 不相关，但是不独立。

例题

习题 6 解答

因为 $\text{cov}(X, X^2) = EX^3 - EXEX^2 = 0$, 所以 X 与 X^2 不相关。为了说明 X 和 Y 不独立，取特定的 $a > 0$ 使得 $P(X \leq a) < 1$，考察下列概率:

$$\begin{aligned} P(X \leq a, X^2 \leq a^2) &= P(-a \leq X \leq a) \\ &> P(X \leq a)P(-a \leq X \leq a) \\ &= P(X \leq a)P(X^2 \leq a) \end{aligned}$$

例题

习题 7

试说明，如果随机变量 X, Y 都只能取两个值，X 和 Y 的独立性和不相关性时等价的。

例题

习题 7 解答

因为独立一定可以导出不相关，所以我们只需要说明如果 X 和 Y 不相关，则 X 和 Y 独立。不失一般性地假设 X 和 Y 都只能取 0 和 1，则 X 和 Y 的概率分布为

Table: (X,Y) 的概率分布

X \ Y	0	1	合计
0	p_{11}	p_{12}	$1-\alpha$
1	p_{21}	p_{22}	α
合计	$1-\beta$	β	

例题

习题 7 解答

因此 $EX=\alpha$,$EY=\beta$, $EXY=P(X=Y=1)=p_{22}$。由于 X,Y 不相关，所以 $p_{22}=\alpha\beta$，因此我们有如下的关系式:

$$p_{11}+p_{12}=1-\alpha \tag{6}$$
$$p_{21}+\alpha\beta=\alpha \tag{7}$$
$$p_{11}+p_{21}=1-\beta \tag{8}$$
$$p_{12}+\alpha\beta=\beta \tag{9}$$

因此，$p_{11}=(1-\alpha)(1-\beta)$,$p_{12}=(1-\alpha)\beta$,$p_{21}=\alpha(1-\beta)$,$p_{22}=\alpha\beta$, 即 X 和 Y 独立。