第三章 多维随机变量及其分布

Instructor: 郝壮

haozhuang@buaa.edu.cn
School of Economics and Management
Beihang University

December 6, 2021

第三章多维随机变量及其分布

§3.1 多维随机变量及其联合分布

3.1.1 多维随机变量

定义3.1.1 若$X(\omega), Y(\omega)$是两个定义在同一个样本空间$\Omega=\{\omega\}$上的随机变量, 则称$(X(\omega), Y(\omega))$ 是二维随机变量.

同理可定义n 维随机变量 (随机向量).

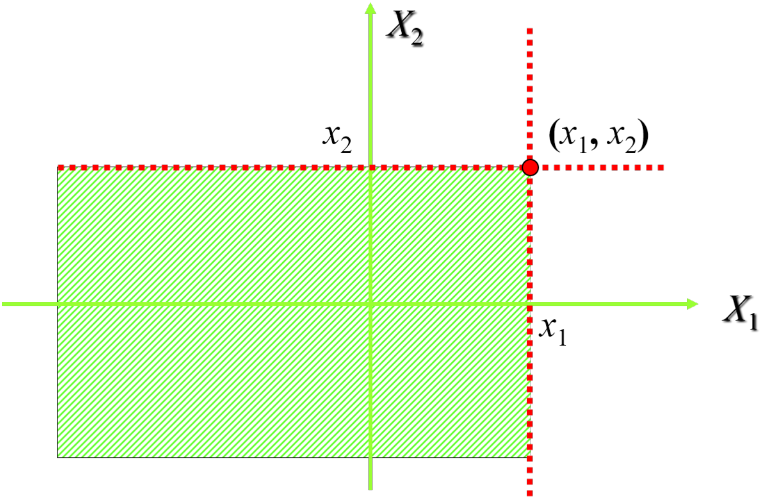
若$X_1(\omega), X_2(\omega), ..., X_n(\omega)$是定义在同一个样本空间$\Omega=\{\omega\}$上的随机变量, 则称$\mathbf{X}(\omega)=(X_1(\omega), X_2(\omega), ..., X_n(\omega))$ 是n维随机变量或随机向量.

3.1.2 联合分布函数

定义3.1.2 (n维随机变量联合分布函数) 对任意实数 $x_1, x_2, ..., x_n$, 称n个事件$\{X_1 \le x_1\}, \{X_2 \le x_2\}, ..., \{X_n \le x_n\}$同时发生的概率$F(x_1, x_2, ..., x_n) = P(X_1 \le x_1, X_2 \le x_2, ..., X_n \le x_n)$ 称为n维随机变量$(X_1, X_2, ..., X_n)$ 的**联合分布函数**.

(二维随机变量联合分布函数) 对任意实数 x 和 y, 称$F(x, y) = P(X \le x, Y \le y)$ 为(X, Y) 的联合分布函数.

- $F(x, y)$为(X, Y)落在点(x, y)的左下区域的概率.


X_2
x_2
(x_1, x_2)
X_1
x_1

定理3.1.1 (证明见教材)

(1) **(单调性)** $F(x,y)$ 分别对 x 和 y是单调非减的.

(2) **(有界性)** $0 \leq F(x,y) \leq 1$, 且

$$F(-\infty, y) = F(x, -\infty) = 0,\ F(+\infty, +\infty) = 1.$$

(3) **(右连续性)** $F(x,y)$ 关于x 和y分别右连续.

$$\begin{aligned} F(x+0, y) &= F(x,y), \\ F(x, y+0) &= F(x,y). \end{aligned}$$

(4) **(非负性)** 当$a < b, c < d$时, 有 $P(a < X \leq b, c < Y \leq d) = F(b,d) - F(b,c) - F(a,d) + F(a,c) \geq 0$.

二维离散随机变量

定义3.1.3 若二维随机变量(X, Y) 的可能取值为有限对, 或可列对, 则称(X, Y)为二维离散随机变量.

二维离散分布的联合分布列

称 $p_{ij} = P(X = x_i, Y = y_j), i, j = 1, 2, \dots$为$(X, Y)$的联合分布列, 其表格形式如下:

$X \backslash Y$	y_1	y_2	$\cdots$	y_j	$\cdots$
x_1	p_{11}	p_{12}	$\cdots$	p_{1j}	$\cdots$
x_2	p_{21}	p_{22}	$\cdots$	p_{2j}	$\cdots$
$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$
x_i	p_{i1}	p_{i2}	$\cdots$	p_{ij}	$\cdots$
$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$

联合分布列的基本性质

(1) **(非负性)** $p_{ij} \geq 0, i, j = 1, 2, \ldots$

(2) **(正则性)** $\sum_{i=1}^{\infty} \sum_{i=1}^{\infty} p_{ij} = 1$

确定联合分布列的方法

(1) 确定随机变量 (X, Y) 的所有取值数对;

(2) 计算取每个数值对的概率;

(3) 列出表格.

例3.1.2 设随机变量 X 在 $1, 2, 3, 4$四个整数中等可能地取值, 另一个随机变量 Y在1到X 中等可能地取一整数值, 试求(X, Y)的联合分布列.

解: $P(X = 1) = P(X = 2) = P(X = 3) = P(X = 4) = 1/4$
$P(Y = 1|X = 1) = 1$
$P(Y = 1|X = 2) = P(Y = 2|X = 2) = 1/2$
$P(Y = 1|X = 3) = P(Y = 2|X = 3) = P(Y = 3|X = 3) = 1/3$
$P(Y = 1|X = 4) = P(Y = 2|X = 4) = P(Y = 3|X = 4) =$
$P(Y = 4|X = 4) = 1/4$

(X, Y) 的可能取值数对及相应的概率如下:

- $P(X = 1, Y = 1) = P(Y = 1|X = 1)P(X = 1) = 1/4$
- $P(X = 2, Y = 1) = P(Y = 1|X = 2)P(X = 2) = 1/8$
- …

列表为:

$X \backslash Y$	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	$1/16$	$1/16$	$1/16$	$1/16$

课堂练习 (1')

将一枚均匀的硬币抛掷4次, X表示正面向上的次数, Y表示反面朝上次数, 求 (X, Y)的联合分布列.

解:

将一枚均匀的硬币抛掷4次, X表示正面向上的次数, Y表示反面朝上次数, 求 (X,Y)的联合分布列.

解:

概率非零的(X,Y) 可能取值对为:

X	Y	其对应的概率分别为:
0	4	$P(X=0,Y=4)=0.5^4=1/16$
1	3	$P(X=1,Y=3)=C_4^1\times 0.5\times 0.5^3=1/4$
2	2	$P(X=2,Y=2)=C_4^2\times 0.5^2\times 0.5^2=6/16$
3	1	$P(X=3,Y=1)=C_4^3\times 0.5^3\times 0.5^1=1/4$
4	0	$P(X=4,Y=0)=0.5^4=1/16$

列表为:

$X \backslash Y$	0	1	2	3	4
0	0	0	0	0	$1/16$
1	0	0	0	$1/4$	0
2	0	0	$6/16$	0	0
3	0	$1/4$	0	0	0
4	$1/16$	0	0	0	0

课堂练习 (2')

设随机变量$Y \sim N(0,1)$, 求以下随机变量的联合分布列.

$$X_1 = \begin{cases} 0, & |Y| \geq 1. \\ 1, & |Y| < 1. \end{cases} \qquad X_2 = \begin{cases} 0, & |Y| \geq 2. \\ 1, & |Y| < 2. \end{cases}$$

解:

课堂练习 (2')

设随机变量$Y \sim N(0,1)$, 求以下随机变量的联合分布列.

$$X_1 = \begin{cases} 0, & |Y| \geq 1. \\ 1, & |Y| < 1. \end{cases} \qquad X_2 = \begin{cases} 0, & |Y| \geq 2. \\ 1, & |Y| < 2. \end{cases}$$

解: (X_1, X_2) 的可能取值数对及相应的概率如下:

- $P(X_1 = 0, X_2 = 0) = P(|Y| \geq 1, |Y| \geq 2) = P(|Y| \geq 2) = 2 - 2\Phi(2) = 0.0455$
- $P(X_1 = 0, X_2 = 1) = P(|Y| \geq 1, |Y| < 2) = P(1 \leq |Y| < 2) = 2[\Phi(2) - \Phi(1)] = 0.2719$
- $P(X_1 = 1, X_2 = 0) = P(|Y| < 1, |Y| \geq 2) = 0$
- $P(X_1 = 1, X_2 = 1) = P(|Y| < 1, |Y| < 2) = P(|Y| < 1) = 0.6826$

列表为:

X_1 \ X_2	0	1
0	0.0455	0.2719
1	0	0.6826

3.1.4 联合密度函数

设二维随机变量(X, Y) 的分布函数为 $F(x, y)$, 若存在非负可积函数 $p(x, y)$, 使得

$$F(x,y) = \int_{-\infty}^{x}\int_{-\infty}^{y} p(\mu,\nu)d\nu d\mu$$

则称 (X, Y) 为二维连续型随机变量. 称$p(x, y)$ 为**联合密度函数**.

在$F(x, y)$偏导数存在的点上有

$$p(x,y) = \frac{\partial^2}{\partial x \partial y}F(x,y)$$

联合密度函数的基本性质

- (1) **(非负性)** $p(x,y) \geq 0$
- (2) **(正则性)** $\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} p(x,y)dxdy = 1$
- **注意:** 若G为平面上的一个区域, 则

$$P\{(X,Y) \in G\} = \iint\limits_{G} p(x,y)dxdy$$

3.1.5 常用多维分布

一维分布的推广

- (1) 多项分布
- (2) 多维超几何分布
- (3) 二维均匀分布
- (4) 二元正态分布

一. 多项分布 (multimonial distribution)

若每次试验有r种结果: $A_1, A_2, ..., A_r$. 记$P(A_i) = p_i$, $i = 1, 2, ..., r$. 记X_i为n次独立重复试验中A_i出现的次数. 则$(X_1, X_2, ..., X_r)$的联合分布列为:

$$P(X_1 = n_1, X_2 = n_2, ..., X_r = n_r) = \frac{n!}{n_1!n_2!...n_r!} p_1^{n_1} p_2^{n_2} ... p_r^{n_r}$$

其中$n_1 + n_2 + ... + n_r = n$. 该联合分布列称为$r$**项分布**, 又称**多项分布**, 记为$M(n, p_1, p_2, ..., p_r)$.

一. 多项分布 (multimonial distribution)

若每次试验有r种结果: $A_1, A_2, ..., A_r$. 记$P(A_i) = p_i$, $i = 1, 2, ..., r$. 记X_i为n次独立重复试验中A_i出现的次数. 则$(X_1, X_2, ..., X_r)$的联合分布列为:

$$P(X_1 = n_1, X_2 = n_2, ..., X_r = n_r) = \frac{n!}{n_1! n_2! ... n_r!} p_1^{n_1} p_2^{n_2} ... p_r^{n_r}$$

其中$n_1 + n_2 + ... + n_r = n$. 该联合分布列称为$r$**项分布**, 又称**多项分布**, 记为$M(n, p_1, p_2, ..., p_r)$.

- $\frac{n!}{n_1! n_2! ... n_r!}$是分割$n$个元素为容量为$(n_1, n_2, ..., n_r)$的$r$个集合的所有可能分割(partition)数, 称为多项式系数(multinomial coefficient): $\binom{n}{n_1, n_2, \cdots, n_r} = \binom{n}{n_1}\binom{n-n_1}{n_2}\cdots\binom{n-n_1-\cdots-n_{r-1}}{n_r} = \frac{n!}{n_1! n_2! \cdots n_r!}$

例子:

- 掷骰子10次, 求出现1次2点, 2次4点, 7次6点的概率.
- 每次抽样结果多于2个的有放回抽样

注意:

- 当$r = 2$, 多项分布褪化为二项分布.

二. 多维超几何分布(multivariate hyper-geometric distribution)

口袋中有 N 只球, 分成r类. 第i 种球有 N_i 只,
$N_1+N_2+...+N_r=N$.

从中任取 n只, 记 X_i为取出的n 只球中, 第i种球的只数.
则$(X_1,X_2,...,X_r)$的联合分布列为:

$$P(X_1=n_1,X_2=n_2,...,X_r=n_r)=\frac{\binom{N_1}{n_1}\binom{N_2}{n_2}\cdots\binom{N_r}{n_r}}{\binom{N}{n}}$$

三. 多维均匀分布(multivariate uniform distribution)

二维均匀分布: 若二维连续随机变量(X, Y) 的联合密度为:

$$p(x, y) = \begin{cases} \dfrac{1}{S_D}, & (x, y) \in D. \\ 0, & \text{其他}. \end{cases}$$

其中S_D为D的面积.

则称 (X, Y) 服从 D 上的均匀分布, 记为$(X, Y) \sim U(D)$.

多维均匀分布同理可定义.

四. 二元正态分布

若二维连续随机变量 (X, Y) 的联合密度为:

$$p(x,y) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}}$$

$$exp\left\{-\frac{1}{2(1-\rho^2)}[\frac{(x-\mu_1)^2}{\sigma_1^2} + \frac{(y-\mu_2)^2}{\sigma_2^2} - 2\rho\frac{(x-\mu_1)(y-\mu_2)}{\sigma_1\sigma_2}]\right\}$$

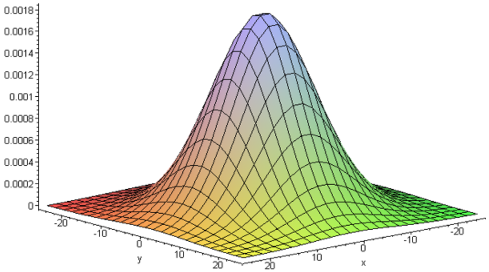
则称 (X, Y)服从二元正态分布,

记为 $(X, Y) \sim N(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho)$.

其中 $-\infty < \mu_1, \mu_2 < \infty$, $\sigma_1, \sigma_2 > 0$, $-1 \leq \rho \leq 1$.

以后将证明 $\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho$分别是$X, Y$的均值, 方差和相关系数.

二元正态密度函数



注意: 任何一截面不是正态分布密度函数! 教材pp. 133 说截面显示正态曲线不严格, 是一维正态分布密度函数曲线的比例缩放, 积分值不为1.

例3.1.3 (课堂练习 2')

若二维连续随机变量(X, Y) 的联合密度为:

$$(X, Y) \sim p(x, y) = \begin{cases} Ae^{-(2x+3y)}, & x \geq 0, y \geq 0. \\ 0, & \text{其他}. \end{cases}$$

试求常数 A.

解:

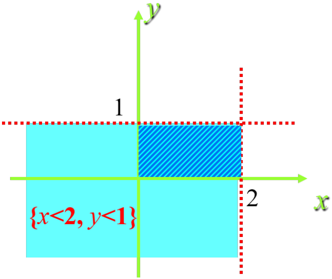
$$
\begin{aligned}
1 &= \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} p(x,y)dxdy \\
&= \int_{0}^{+\infty}\int_{0}^{+\infty} Ae^{-(2x+3y)}dxdy \\
&= A\int_{0}^{+\infty} e^{-2x}dx\int_{0}^{+\infty} e^{-3y}dy \\
&= A/6
\end{aligned}
$$

所以, $A = 6$.

若

$$(X,Y) \sim p(x,y) = \begin{cases} 6e^{-(2x+3y)}, & x \geq 0, y \geq 0. \\ 0, & \text{其他}. \end{cases}$$

试求 $P\{X < 2, Y < 1\}$.



解:

$$
\begin{aligned}
P\{X<2, Y<1\} &= \iint\limits_{\{X<2,Y<1\}} p(x,y)dxdy \\
&= \int_0^2 \int_0^1 6e^{-(2x+3y)}dydx \\
&= 6\int_0^2 e^{-2x}dx \int_0^1 e^{-3y}dy \\
&= (1-e^{-4})(1-e^{-3})
\end{aligned}
$$

例3.1.5 (课堂练习 5’)

若

$$(X,Y)\sim p(x,y)=\begin{cases}6e^{-(2x+3y)}, & x\geq 0, y\geq 0.\\ 0, & \text{其他}.\end{cases}$$

试求 $P\{(X,Y)\in D\}$. 其中D为 $2x+3y\leq 6$.

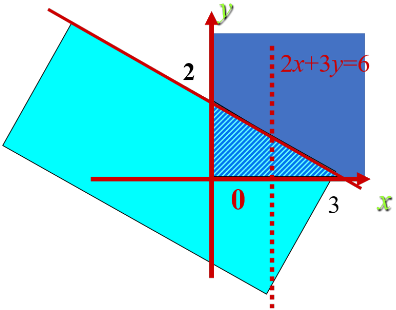
解:

若

$$(X, Y) \sim p(x, y) = \begin{cases} 6e^{-(2x+3y)}, & x \geq 0, y \geq 0. \\ 0, & \text{其他}. \end{cases}$$

试求 $P\{(X, Y) \in D\}$. 其中D为 $2x + 3y \leq 6$.

解:



解 (con’d):

$$
\begin{aligned}
P\{(X,Y)\in D\} &= \iint\limits_{\{2x+3y\le 6\}} p(x,y)dxdy \\
&= \int_0^3 \int_0^{\frac{1}{3}(6-2x)} 6e^{-(2x+3y)}dydx \\
&= 6\int_0^3 e^{-2x}(-\frac{1}{3}e^{-3y})\Big|_0^{(6-2x)/3} dx \\
&= 1-7e^{-6}
\end{aligned}
$$

课堂练习(5')

设二维随机变量 (X,Y) 的密度函数为

$$(X,Y) \sim p(x,y) = \begin{cases} e^{-y}, & 0 < x < y, \\ 0, & \text{其他}. \end{cases}$$

求概率$P(X+Y \leq 1)$.

解:

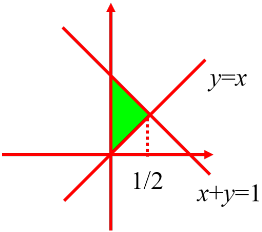
课堂练习(5')

设二维随机变量 (X, Y) 的密度函数为

$$(X, Y) \sim p(x, y) = \begin{cases} e^{-y}, & 0 < x < y, \\ 0, & \text{其他}. \end{cases}$$

求概率$P(X + Y \leq 1)$.

解:



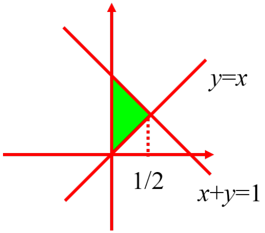
课堂练习(5')

设二维随机变量 (X,Y) 的密度函数为

$$(X,Y) \sim p(x,y) = \begin{cases} e^{-y}, & 0 < x < y, \\ 0, & \text{其他}. \end{cases}$$

求概率$P(X+Y \le 1)$.

解:



$$P(X+Y \le 1) = \int_0^{1/2} \int_x^{1-x} e^{-y} dy dx = 1 + e^{-1} - 2e^{-1/2}$$

第9次作业

习题3.1中题目1, 3, 4, 6, 7, 11, 13, 15

二维联合分布函数包含哪些主要信息?

- 每个分量的分布(边际分布);
- 两个分量的关联性;
- 给定一个分量, 另一个分量的分布, 即条件分布.

本节学习**边际分布**.

问题: 已知二维随机变量 (X, Y)的分布, 如何求出X和Y各自的分布?

3.2.1 边际分布函数 (marginal distribution function)

已知(X,Y)的联合分布函数为$F(x,y)$. 令$y\to+\infty$, 由于$\{y<\infty\}$是必然事件, 故

$$\lim_{y\to+\infty}F(x,y)=P(X\le x,y<+\infty)=P(X\le x)$$

这是由(X,Y)的联合分布求出的X的分布函数, 称为X的**边际分布**, 记为$X\sim F_X(x)=F(x,+\infty)$.

同理, $Y\sim F_Y(y)=F(+\infty,y)$.

3.2.1 边际分布函数

同理定义多维随机变量边际分布函数.

三维随机变量: 已知(X, Y, Z)的联合分布函数为$F(x, y, z)$.

则X的边际分布为$X \sim F_X(x) = F(x, +\infty, +\infty)$.

同理

$Y \sim F_Y(y) = F(+\infty, y, +\infty)$,

$Z \sim F_Z(z) = F(+\infty, +\infty, z)$,

$(X, Y) \sim F_{X,Y}(x, y) = F(x, y, +\infty)$,

...

已知(X, Y)的联合分布函数为

$$F(x, y) = \begin{cases} 1 - \mathrm{e}^{-x} - \mathrm{e}^{-y} + \mathrm{e}^{-x-y-\lambda xy}, & x > 0, y > 0 \\ 0, & \text{其他} \end{cases}$$

其中$\lambda > 0$. 这个分布被称为**二维指数分布**. 求X, Y的边际分布函数

解:

课堂练习: 二维指数分布 (1')

已知(X,Y)的联合分布函数为

$$F(x,y)=\begin{cases}1-\mathrm{e}^{-x}-\mathrm{e}^{-y}+\mathrm{e}^{-x-y-\lambda xy}, & x>0,y>0\\ 0, & \text{其他}\end{cases}$$

其中$\lambda>0$. 这个分布被称为**二维指数分布**. 求X,Y的边际分布函数

解:

$$F_X(x)=F(x,+\infty)=\begin{cases}1-\mathrm{e}^{-x}, & x>0\\ 0, & x\leqslant 0\end{cases}$$

$$F_Y(y)=F(+\infty,y)=\begin{cases}1-\mathrm{e}^{-y}, & y>0\\ 0, & y\leqslant 0\end{cases}$$

课堂练习: 二维指数分布 (1')

已知(X, Y)的联合分布函数为

$$F(x, y) = \begin{cases} 1 - \mathrm{e}^{-x} - \mathrm{e}^{-y} + \mathrm{e}^{-x-y-\lambda xy}, & x > 0, y > 0 \\ 0, & \text{其他} \end{cases}$$

其中$\lambda > 0$. 这个分布被称为**二维指数分布**. 求X, Y的边际分布函数

解:

$$F_X(x) = F(x, +\infty) = \begin{cases} 1 - \mathrm{e}^{-x}, & x > 0 \\ 0, & x \leqslant 0 \end{cases}$$

$$F_Y(y) = F(+\infty, y) = \begin{cases} 1 - \mathrm{e}^{-y}, & y > 0 \\ 0, & y \leqslant 0 \end{cases}$$

两个边际分布都是指数分布, 与λ 无关. 联合分布包含信息更多: 各分量信息+分量关系信息.

注意点

- 由联合分布可以求出边际分布.
- 但由边际分布一般无法求出联合分布: 具有相同边际分布的多维联合分布可以是不同的.
- 所以联合分布包含更多的信息.

3.2.2 边际分布列

已知 (X,Y)的联合分布列为 $p_{ij}=P(X=x_i,Y=y_i)$, 则

X 的**边际分布列为**: $p_i=P(X=x_i)=\sum\limits_{j=1}^{+\infty}p_{ij}=p_{i\cdot}$

Y 的**边际分布列为**: $p_j=P(Y=y_j)=\sum\limits_{i=1}^{+\infty}p_{ij}=p_{\cdot j}$

$X \backslash Y$	y_1	y_2	$\cdots$	y_j	$\cdots$	$p_{i\cdot}$
x_1	p_{11}	p_{12}	$\cdots$	p_{1j}	$\cdot\cdot$	$p_{1\cdot}$
x_2	p_{21}	p_{22}	$\cdots$	p_{2j}	$\cdots$	$p_{2\cdot}$
$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$
x_i	p_{i1}	p_{i2}	$\cdots$	p_{ij}	$\cdots$	$p_{i\cdot}$
$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$
$p_{\cdot j}$	$p_{\cdot 1}$	$p_{\cdot 2}$	$\cdots$	$p_{\cdot j}$	$\cdots$	

3.2.3 边际密度函数

已知 (X, Y)的联合密度函数为$p(x, y)$, 则

X**的边际密度函数为**: $p_X(x) = \int_{-\infty}^{+\infty} p(x, y)dy$

Y**的边际密度函数为**: $p_Y(y) = \int_{-\infty}^{+\infty} p(x, y)dx$

上式也可由联合分布推出

$$F_X(x) = F(x, \infty) = \int_{-\infty}^{x} \left(\int_{-\infty}^{\infty} p(x, y)\mathrm{d}y \right) \mathrm{d}x = \int_{-\infty}^{x} p_X(x)\mathrm{d}x$$

$$F_Y(y) = F(\infty, y) = \int_{-\infty}^{y} \left(\int_{-\infty}^{\infty} p(x, y)\mathrm{d}x \right) \mathrm{d}y = \int_{-\infty}^{y} p_Y(y)\mathrm{d}y$$

注意点

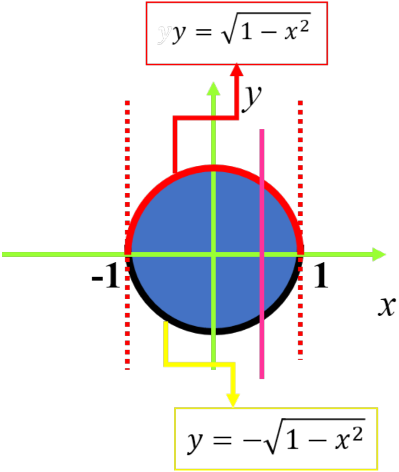
- 多项分布的一维边际分布为二项分布. (稍后证明)
- 二维正态分布的边际分布是一维正态分布:
  若$(X,Y)\sim N(\mu_1,\mu_2,\sigma_1^2,\sigma_2^2,\rho)$, 则
  $X\sim N(\mu_1,\sigma_1^2), Y\sim N(\mu_2,\sigma_2^2)$. (证明见教材例3.2.5)
- 二维均匀分布的边际分布不一定是一维均匀分布.

课堂练习(2')

设 (X, Y)服从区域$D = \{(x, y), x^2 + y^2 \leq 1\}$上的均匀分布，求$X$ 的边际密度$p(x)$.

课堂练习(2')

设 (X,Y)服从区域$D=\{(x,y),x^2+y^2\le 1\}$上的均匀分布，求X 的边际密度$p(x)$.



解:由题意得

$$(X,Y) \sim p(x,y) = \begin{cases} \frac{1}{\pi}, & x^2+y^2 \le 1. \\ 0, & \text{其他}. \end{cases}$$

当$|x| > 1$时, $p(x,y) = 0$, 所以$p(x) = 0$

当$|x| \le 1$时, $p_X(x) = \int_{-\sqrt{1-x^2}}^{\sqrt{1-x^2}} \frac{1}{\pi} dy = \frac{2}{\pi}\sqrt{1-x^2}$

不是均匀分布!

教材例3.2.4 多项分布的一维边际分布仍为二项分布

下面证明三项分布的边际分布为二项分布, 多项分布的证明同理(依次降维).

设二维随机变量X, Y服从三项分布$M(n, p_1, p_2, p_3)$, 其联合分布列为

$$P(X=i, Y=j) = \frac{n!}{i!j!(n-i-j)!} p_1^i p_2^j (1-p_1-p_2)^{n-i-j}$$

其中$i, j = 0, 1, 2, ..., n; i+j \leq n$. 对上式同时乘以和除以$(1-p_1)^{n-i}/(n-i)!$

$$P(X=i, Y=j) = \frac{(1-p_1)^{n-i}}{(1-p_1)^{n-i}} \frac{(n-i)!}{(n-i)!} \frac{n!}{i!j!(n-i-j)!} p_1^i p_2^j (1-p_1-p_2)^{n-i-j}$$

对上式同时乘以和除以$(1-p_1)^{n-i}/(n-i)!$,

$$
\begin{aligned}
& P(X=i, Y=j) \\
= & \frac{(1-p_1)^{n-i}}{(1-p_1)^{n-i}} \frac{(n-i)!}{(n-i)!} \frac{n!}{i!j!(n-i-j)!} p_1^i p_2^j (1-p_1-p_2)^{n-i-j} \\
= & \frac{n!}{i!(n-i)!} p_1^i (1-p_1)^{n-i} \cdot \frac{(n-i)!}{(1-p_1)^{n-i}} \frac{1}{j!(n-i-j)!} p_2^j (1-p_1-p_2)^{n-i-j} \\
= & \frac{n!}{i!(n-i)!} p_1^i (1-p_1)^{n-i} \cdot \frac{(n-i)!}{j!(n-i-j)!} \frac{p_2^j (1-p_1-p_2)^{n-i-j}}{(1-p_1)^{n-i}} \\
= & \frac{n!}{i!(n-i)!} p_1^i (1-p_1)^{n-i} \cdot \binom{n-i}{j} \left(\frac{p_2}{1-p_1}\right)^j \left(1-\frac{p_2}{1-p_1}\right)^{n-i-j}
\end{aligned}
$$

再对j从0到$n-i$求和

$$
\begin{aligned}
&\sum_{j=0}^{n-i} P(X=i, Y=j) = P(X=i) \\
=&\sum_{j=0}^{n-i}\left(\frac{n!}{i!(n-i)!}p_1^i(1-p_1)^{n-i}\cdot\binom{n-i}{j}(\frac{p_2}{1-p_1})^j(1-\frac{p_2}{1-p_1})^{n-i-j}\right) \\
=&\frac{n!}{i!(n-i)!}p_1^i(1-p_1)^{n-i}\cdot\sum_{j=1}^{n-i}\binom{n-i}{j}(\frac{p_2}{1-p_1})^j(1-\frac{p_2}{1-p_1})^{n-i-j} \\
=&\frac{n!}{i!(n-i)!}p_1^i(1-p_1)^{n-i}\cdot(\frac{p_2}{1-p_1}+(1-\frac{p_2}{1-p_1}))^{n-i} \\
=&\frac{n!}{i!(n-i)!}p_1^i(1-p_1)^{n-i}
\end{aligned}
$$

所以 $X \sim b(n, p_1)$

若二维随机变量(X, Y)满足以下条件之一: 对任意实数x, y有

i) $F(x, y) = F_X(x)F_Y(y)$ (联合分布函数等于边际分布函数之积)

ii) $p_{ij} = p_{i\cdot}p_{\cdot j}$ (离散随机变量情况: 联合分布列等于边际分布列之积)

iii) $p(x, y) = p_X(x)p_Y(y)$(连续随机变量情况: 联合密度函数等于边际密度函数之积)

则称 X 与Y**相互独立**. 表示二维随机变量中两个分量互不影响.

同理可定义n维随机变量的相互独立.

定义3.2.1： 若n维随机变量$(X_1, X_2, ..., X_n)$的联合分布函数为$F(x_1, x_2, ...x_n)$, $F_i(x_i)$为X_i的边际分布函数. 如果对任意n个实数$x_1, x_2, ...x_n$有

$$F(x_1, x_2, \cdots, x_n) = \prod_{i=1}^{n} F_i(x_i)$$

则称 $X_1, X_2, ..., X_n$ **相互独立**.

在离散场合, 上式等价于

$$P(X_1 = x_1, X_2 = x_2, \cdots, X_n = x_n) = \prod_{i=1}^{n} P(X_i = x_i)$$

在连续场合, 上式等价于

$$p(x_1, x_2, \cdots, x_n) = \prod_{i=1}^{n} p_i(x_i)$$

注意点

(1) 回忆1.5两个事件A, B相互独立的定义:
$P(A|B) = P(A) \Leftrightarrow P(AB) = P(A)P(B)$.

(2) 随机变量X与Y独立的本质是: 任对实数a, b, c, d, 有

$$P(a < X < b, c < Y < d) = P(a < X < b)P(c < Y < d)$$

(把$a < X < b, c < Y < d$看作事件A, B)

注意点

(3) 回忆N个事件相互独立的定义需要同时满足两两独立，三三独立,.... 为何随机变量独立只许满足

$$F\left(x_1, x_2, \cdots, x_n\right) = \prod_{i=1}^{n} F_i\left(x_i\right)$$

没有考虑所谓"两两独立, 三三独立,..."?

注意点

(3) 回忆N个事件相互独立的定义需要同时满足两两独立, 三三独立,.... 为何随机变量独立只许满足

$$F(x_1, x_2, \cdots, x_n) = \prod_{i=1}^{n} F_i(x_i)$$

没有考虑所谓"两两独立, 三三独立,..."?

这是因为

$$F(x_1, x_2, \cdots, x_n) = \prod_{i=1}^{n} F_i(x_i)$$

可以推出任意两分量, 三分量...满足以上分离条件. 如:

$$\begin{aligned} F_{X_1,X_2}(x_1, x_2) &= F(x_1, x_2, \infty, \infty, \ldots) \\ &= F(x_1) F(x_2) \prod_{i=3}^{n} F_i(\infty) = F_{X_1}(x_1) F_{X_2}(x_2) \end{aligned}$$

若X与Y 是独立的, 则任意随机变量函数$g(X)$与$h(Y)$也是独立的.

证明若X与Y 是独立的, 则任意随机变量函数$g(X)$与$h(Y)$也是独立的.

证明: 因为X与Y 是独立的, 所以$P(x, y) = P(x)P(y)$ 对任意x, y都成立.

设$U = g(X), V = h(Y)$, 只需证明
$P(U = u, V = v) = P(U = u)P(V = v)$ 对任意u, v都成立.

$$
\begin{aligned}
P(U = u, V = v) =& P(g(X) = u, h(Y) = v) \\
=& P(X = g^{-1}(u), Y = h^{-1}(v)) \\
=& P(X = g^{-1}(u))P(Y = h^{-1}(v)) \\
=& P(g(X) = u)P(h(Y) = v) \\
=& P(U = u)P(V = v)
\end{aligned}
$$

例3.2.3

(X,Y)的联合分布列为:

$X \backslash Y$	0	1
0	0.3	0.4
1	0.2	0.1

问 X与Y 是否独立?

解:

(X, Y)的联合分布列为:

$X \backslash Y$	0	1
0	0.3	0.4
1	0.2	0.1

问 X与Y 是否独立?

解: 边际分布列分别为:

X	0	1
P	0.7	0.3

Y	0	1
P	0.5	0.5

因为$P(X=0, Y=0)=0.3$,
$P(X=0)P(Y=0)=0.7\times 0.5=0.35$, 所以不独立.

利用如下边际分布列写出一个(X, Y)的联合分布列使X, Y独立.

X	0	1
P	0.7	0.3

Y	0	1
P	0.5	0.5

解:

练习

利用如下边际分布列写出一个(X,Y)的联合分布列使X,Y独立.

X	0	1
P	0.7	0.3

Y	0	1
P	0.5	0.5

解:

$X \backslash Y$	0	1
0	0.35	0.35
1	0.15	0.15

利用如下边际分布列写出一个(X,Y)的联合分布列使X,Y独立.

X	0	1
P	0.7	0.3

Y	0	1
P	0.5	0.5

解:

$X \backslash Y$	0	1
0	0.35	0.35
1	0.15	0.15

边际分布相同, 联合分布不一定相同!

例3.2.4

已知 (X, Y) 的联合密度函数为

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

问 X 与 Y 是否独立?

解:

例3.2.4

已知 (X, Y) 的联合密度函数为

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

问 X 与 Y 是否独立?

解: 边际密度函数分别为:

$$p_X(x) = \begin{cases} \int_0^{+\infty} e^{-x-y}dy = e^{-x}, & x>0. \\ 0, & x \le 0. \end{cases} \qquad p_Y(y) = \begin{cases} e^{-y}, & y>0. \\ 0, & y \le 0. \end{cases}$$

所以X 与Y 独立.

(注意: $p(x,y)$ 可分离.)

试证明: 若联合密度函数$p(x,y)$可分离, 即$p(x,y)=h(x)g(y)$, 则X与Y独立.

证明:

拓展知识(课堂练习)

试证明: 若联合密度函数$p(x,y)$可分离, 即$p(x,y)=h(x)g(y)$, 则X与Y独立.

证明: 记$k_1=\int_{-\infty}^{+\infty}h(x)dx$, $k_2=\int_{-\infty}^{+\infty}g(y)dy$. 因为$p(x,y)=h(x)g(y)$, 所以

$$\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}p(x,y)dxdy=\int_{-\infty}^{+\infty}h(x)dx\int_{-\infty}^{+\infty}g(y)dy=k_1k_2=1$$

$$\begin{aligned} p_X(x)=&h(x)\int_{-\infty}^{+\infty}g(y)dy=k_2h(x)\\ p_Y(y)=&g(y)\int_{-\infty}^{+\infty}h(x)dx=k_1g(y)\end{aligned}$$

所以$p(x,y)=h(x)g(y)=\dfrac{p_X(x)}{k_2}\dfrac{p_Y(y)}{k_1}=p_X(x)p_Y(y)$ $Q.E.D.$

注意点

(1) (X, Y)服从矩形上的均匀分布, 则X与Y独立. (习题3.2 16题)

(2) (X, Y)服从单位圆上的均匀分布, 则 X与Y不独立. (证明由前面边际分布例题可推出)

(3) 联合密度$p(x, y)$的表达式中, 若x的取值与y的取值有关系, 则X与Y不独立.

(4) 若联合密度$p(x, y)$可分离, 即$p(x, y) = h(x)g(y)$, 则X与Y独立 (已证明).

(5) 若(X, Y)服从二元正态$N(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho)$则$X$与$Y$独立的充要条件是$\rho = 0$. (由独立性定义证明, 拓展练习)

以上结论都可以由随机变量独立性定义证明.

第10次作业

习题3.2中题目1, 2, 3, 5, 9, 11, 12, 14

§3.3 多维随机变量函数的分布

问题: 已知二维随机变量(X, Y)的分布, 如何求出 $Z = g(X, Y)$的分布?

多维随机变量函数的分布在一般情况下求解需要很强的技巧, 且不存在一般化的解法, Case by case!

本节内容是本课的难点之一.

3.3.1 多维离散随机变量函数的分布

(1) 设$(X_1, X_2, ..., X_n)$是n维离散随机变量,
则$Z = g(X_1, ..., X_n)$是一维离散随机变量.

(2) 多维离散随机变量函数的分布在取值较少时是容易求的:

$i)$ 对$(X_1, X_2, ..., X_n)$的各种可能取值对, 写出Z相应的取值.

$ii)$ 对Z的相同的取值, 合并其对应的概率.

例： 设二维随机变量(X,Y)的联合分布列为

$X \backslash Y$	-1	1	2
-1	$5/20$	$2/20$	$6/20$
2	$3/20$	$3/20$	$1/20$

求如下离散随机变量函数的分布 (1) $Z_1 = X + Y$; (2) $Z_2 = X - Y$

解: 写出(X,Y)和各个函数的取值对应于同一表中

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

合并整理后

$Z_1=X+Y$	-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	$6/20$	$2/20$	$6/20$	$3/20$	$3/20$

例3.3.1 设X与Y独立, 且X, Y等可能地取值0和1. 求$Z = max(X, Y)$的分布列.

解:

例3.3.1 设X与Y独立, 且X, Y等可能地取值0和1.
求$Z = max(X, Y)$的分布列.

解:

X	0	1
P	$1/2$	$1/2$

Y	0	1
P	$1/2$	$1/2$

$Z = max(X, Y)$ 的取值为: $0, 1$

$$P(Z = 0) = P(X = 0, Y = 0) = P(X = 0)P(Y = 0) = 1/4$$

$$P(Z = 1) = P(X = 0, Y = 1) + P(X = 1, Y = 0) + P(X = 1, Y = 1) = 3/4$$

最大值分布

最大值分布 设$X_1, X_2, ..., X_n$**是相互独立的**n个随机变量，若$Y = max\{X_1, X_2, ..., X_n\}$，在以下情况求$Y$的分布.

(1) $X_i \sim F_i(x)$, $i = 1, 2, \cdots, n$

解:$Y = max\{X_1, X_2, ..., X_n\}$的分布函数为

$$\begin{aligned} F_Y(y) &= P\left(\max\left\{X_1, X_2, \cdots, X_n\right\} \leqslant y\right) \\ &= P\left(X_1 \leqslant y, X_2 \leqslant y, \cdots, X_n \leqslant y\right) \\ &= P\left(X_1 \leqslant y\right) P\left(X_2 \leqslant y\right) \cdots P\left(X_n \leqslant y\right) \\ &= \prod_{i=1}^{n} F_i(y) \end{aligned}$$

(2) 诸X_i同分布，即$X_i \sim F(x)$, $i = 1, 2, \cdots, n$

解: $F_Y(y) = [F(y)]^n$

(3) 诸X_i为连续随机变量且同分布, 即X_i的密度函数均为$p(x), i = 1, 2, \cdots, n$

解: 分布函数仍为$F_Y(y) = [F(y)]^n$, 密度函数可对y求导得$p_Y(y) = F'_Y(y) = n[F(y)]^{n-1}p(y)$

(4) $X_i \sim \mathrm{Exp}(\lambda), i = 1, 2, \cdots, n$

解: 将$\mathrm{Exp}(\lambda)$的分布函数和密度函数带入上式可得

$$F_Y(y) = \begin{cases} 0, & y < 0 \\ \left(1 - \mathrm{e}^{-\lambda y}\right)^n, & y \geqslant 0 \end{cases}$$

$$p_y(y) = \begin{cases} 0, & y < 0 \\ n\left(1 - \mathrm{e}^{-\lambda y}\right)^{n-1} \lambda \mathrm{e}^{-\lambda y}, & y \geqslant 0 \end{cases}$$

最小值分布

最小值分布 设$X_1, X_2, ..., X_n$是相互独立的n个随机变量，若$Y = min\{X_1, X_2, ..., X_n\}$, 在以下情况求$Y$的分布.

(1) $X_i \sim F_i(x), i = 1, 2, \cdots, n$

解:$Y = min\{X_1, X_2, ..., X_n\}$的分布函数为

$$\begin{aligned} F_Y(y) &= P\left(\min\left\{X_1, X_2, \cdots, X_n\right\} \leqslant y\right) \\ &= 1 - P\left(\min\{X_1, X_2, \cdots, X_n\} > y\right) \\ &= 1 - P\left(X_1 > y, X_2 > y, \cdots, X_n > y\right) \\ &= 1 - P\left(X_1 > y\right) P\left(X_2 > y\right) \cdots P\left(X_n > y\right) \\ &= 1 - \prod_{i=1}^{n}\left[1 - F_i(y)\right] \end{aligned}$$

(2) 诸X_i同分布, 即$X_i \sim F(x), i = 1, 2, \cdots, n$.

解: $F_Y(y) = 1 - [1 - F(y)]^n$

(3) 诸X_i为连续随机变量且同分布, 即X_i的密度函数均为$p(x), i = 1, 2, \cdots, n$

解: 分布函数仍为$F_Y(y) = 1 - [1 - F(y)]^n$,密度函数可对$y$求导得$p_Y(y) = F'_Y(y) = n[1 - F(y)]^{n-1}p(y)$

(4) $X_i \sim \mathrm{Exp}(\lambda), i = 1, 2, \cdots, n$

解: 将$\mathrm{Exp}(\lambda)$的分布函数和密度函数带入上式可得

$$F_Y(y) = \begin{cases} 0, & y < 0 \\ 1 - e^{-n\lambda y}, & y \geqslant 0 \end{cases}$$

$$p_Y(y) = \begin{cases} 0, & y < 0 \\ n\lambda \mathrm{e}^{-n\lambda y}, & y \geqslant 0 \end{cases}$$

课堂练习 (5')

不同的三个线路的公交车都可以从北航东门到西直门, 每个线路公交车间隔均为10分钟, 问平均等车时间为多久?

(思考：如果只有一个线路的公交车, 易知平均等车时间为5分钟. 如果有两个线路的公交车, 平均等车时间会缩短一半么? 等车时间是多久?)

解:

课堂练习 (5')

不同的三个线路的公交车都可以从北航东门到西直门, 每个线路公交车间隔均为10分钟, 问平均等车时间为多久?

(思考：如果只有一个线路的公交车, 易知平均等车时间为5分钟. 如果有两个线路的公交车, 平均等车时间会缩短一半么? 等车时间是多久?)

解: 数学抽象: 设单一线路公交车到站的时间为X_1, 这是在$(0,10)$上的均匀分布. 则3个线路公交车到站时间是在$(0,10)$间独立同均匀分布的随机变量X_1, X_2, X_3. 等车时间为$Y = min\{X_1, X_2, X_3\}$

$$p_Y(y) = n[1-F(y)]^{n-1}p(y) = 3[1-y/10]^2\frac{1}{10}, 0 \leqslant y \leqslant 10$$

$E(Y) = 2.5$分钟

总结

设$X_1, X_2, ..., X_n$独立同分布, 其分布函数和密度函数分别为$F_X(x)$和 $p_X(x)$. 若记$Y = max(X_1, X_2, ..., X_n)$, $Z = min(X_1, X_2, ..., X_n)$. 则

Y 的分布函数为: $F_Y(y) = [F_X(y)]^n$

Y 的密度函数为: $p_Y(y) = n[F_X(y)]^{n-1}p_X(y)$

Z 的分布函数为: $F_Z(z) = 1 - [1 - F_X(z)]^n$

Z 的密度函数为: $p_Z(z) = n[1 - F_X(z)]^{n-1}p_X(z)$

3.3.3 连续场合的卷积公式(Convolution formula)

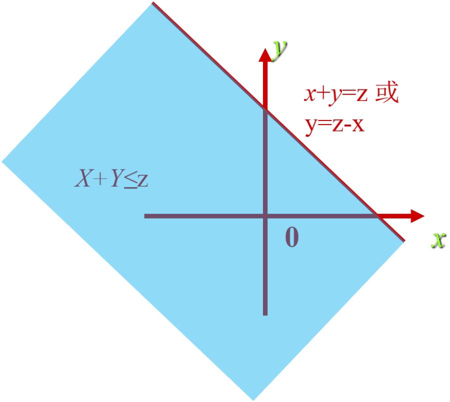
定理3.3.1 设连续随机变量X与Y独立, 其密度函数分别为$p_X(x)$和$p_Y(y)$, 则$Z = X + Y$的密度函数为

$$\begin{aligned} p_Z(z) &= \int_{-\infty}^{+\infty} p_X(x)p_Y(z-x)dx \\ &= \int_{-\infty}^{+\infty} p_X(z-y)p_Y(y)dy \end{aligned}$$

卷积公式并不要求X与Y同分布.

证明：直接求Z的p.d.f.很困难，先求Z的c.d.f.

$F_Z(z) = P(Z \leqslant z) = P(X + Y \leqslant z)$



$F_Z(z) = P(Z \leqslant z) = P(X + Y \leqslant z) =$
$\iint_{\{D_Z:X+Y\leqslant z\}} p_{X,Y}(x, y)\mathrm{d}x\mathrm{d}y$（在$X, Y$的联合pdf上，所有能使$X + Y \leqslant z$发生的积分）

已知X与Y独立，其密度函数分别为$p_X(x)$和$p_Y(y)$，则X, Y的联合密度函数为$p_{X,Y}(x, y) = p_X(x)p_Y(y)$. 所以$Z = X + Y$的分布函数为

$$F_Z(z) = P(X + Y \leqslant z) = \iint_{\{D_Z: X+Y\leqslant z\}} p_{X,Y}(x, y)\mathrm{d}x\mathrm{d}y$$

$$= \iint_{x+y\leqslant z} p_X(x)p_Y(y)\mathrm{d}x\mathrm{d}y = \int_{-\infty}^{\infty}\int_{-\infty}^{z-y} p_X(x)p_Y(y)\mathrm{d}x\mathrm{d}y$$

$$= \int_{-\infty}^{\infty}\left\{\int_{-\infty}^{z-y} p_X(x)\mathrm{d}x\right\} p_Y(y)\mathrm{d}y$$

$$= \int_{-\infty}^{\infty}\int_{-\infty}^{z} p_X(t - y)p_Y(y)\mathrm{d}t\mathrm{d}y$$

换元 $x = t - y$, 积分上限$z - y$在$t = z$时取到, $d(t - y) = dt$

$$= \int_{-\infty}^{z}\left(\int_{-\infty}^{\infty} p_X(t - y)p_Y(y)\mathrm{d}y\right)\mathrm{d}t$$

$Z=X+Y$的分布函数为

$$F_Z(z)=\int_{-\infty}^{z}\left(\int_{-\infty}^{\infty}p_X(t-y)p_Y(y)\mathrm{d}y\right)\mathrm{d}t$$

由此可得Z的密度函数为 $p_Z(z)=\int_{-\infty}^{\infty}p_X(z-y)p_Y(y)\mathrm{d}y$

上式中令$x=z-y$,可得 $p_Z(z)=\int_{-\infty}^{\infty}p_X(x)p_Y(z-x)\mathrm{d}x$

离散场合的卷积公式

设离散随机变量 X 与Y 独立, 则 $Z = X + Y$的分布列为

$$
\begin{aligned}
P(Z = k) &= \sum_i P(X = x_i)P(Y = k - x_i) \\
&= \sum_j P(X = k - y_j)P(Y = y_j)
\end{aligned}
$$

k为Z所有可能取值. 证明同连续场合.

卷积公式的应用: 证明分布的可加性

- **定义：** 若同一类分布的独立随机变量和的分布仍是此类分布, 则称此类分布具有**可加性**.

二项分布的可加性

若 $X \sim b(n, p)$, $Y \sim b(m, p)$, 且独立, 则 $Z = X + Y \sim b(n + m, p)$.

注意:

- 若 $X_i \sim b(1, p)$, 且独立, 则$Z = X_1 + X_2 + \ldots + X_n \sim b(n, p)$.

证明： $Z=X+Y$ 可取$0,1,...,n+m$共计$n+m+1$个不同值，利用离散场合的卷积公式

$$P(Z=k)=\sum_{i=0}^{k}P(X=i)P(Y=k-i)$$

在二项分布中，上式有些事件是不可能事件：

- 当$i>n$时，$\{X=i\}$是不可能事件，所以只需考虑$i\leqslant n$.
- 当$k-i>m$ 时，$\{Y=k-i\}$是不可能事件，所以只需考虑$i\geqslant k-m$.

因此记$a=max\{0,k-m\},b=min\{n,k\}$，则

$$
\begin{aligned}
P(Z=k) &= \sum_{i=a}^{b} P(X=i)P(Y=k-i) \\
&= \sum_{i=a}^{b} \binom{n}{i} p^i (1-p)^{n-i} \cdot \binom{m}{k-i} p^{k-i} (1-p)^{m-(k-i)} \\
&= p^k (1-p)^{n+m-k} \sum_{i=a}^{b} \binom{n}{i} \binom{m}{k-i}
\end{aligned}
$$

利用超几何分布可证明(思考如何推导):

$$
\sum_{i=a}^{b} \binom{n}{i} \binom{m}{k-i} = \binom{n+m}{k}
$$

所以

$$
P(Z=k) = \binom{n+m}{k} p^k (1-p)^{n+m-k}, \quad k = 0, 1, \cdots, n+m
$$

泊松分布的可加性

若 $X \sim P(\lambda_1), Y \sim P(\lambda_2)$, 且独立, 则
$Z = X + Y \sim P(\lambda_1 + \lambda_2)$.

- 可推广至有限独立泊松变量之和的分布;
- $X - Y$ 不服从泊松分布.

证明： $Z = X + Y$可取$0, 1, 2, \ldots$所有非负整数. 利用离散场合的卷积公式$P(Z = k) = \sum_{i=0}^{k} P(X = i)P(Y = k - i)$, 可得

$$
\begin{aligned}
P(Z = k) &= \sum_{i=0}^{k} \left(\frac{\lambda_1^i}{i!}\mathrm{e}^{-\lambda_1}\right)\left(\frac{\lambda_2^{k-i}}{(k-i)!}\mathrm{e}^{-\lambda_2}\right) \\
&= \frac{(\lambda_1 + \lambda_2)^k}{k!}\mathrm{e}^{-(\lambda_1+\lambda_2)} \sum_{i=0}^{k} \frac{k!}{i!(k-i)!}\left(\frac{\lambda_1}{\lambda_1 + \lambda_2}\right)^i \left(\frac{\lambda_2}{\lambda_1 + \lambda_2}\right)^{k-i} \\
&= \frac{(\lambda_1 + \lambda_2)^k}{k!}\mathrm{e}^{-(\lambda_1+\lambda_2)} \left(\frac{\lambda_1}{\lambda_1 + \lambda_2} + \frac{\lambda_2}{\lambda_1 + \lambda_2}\right)^k \\
&= \frac{(\lambda_1 + \lambda_2)^k}{k!}\mathrm{e}^{-(\lambda_1+\lambda_2)}, \quad k = 0, 1, \cdots
\end{aligned}
$$

(第二步为同时乘除$\frac{(\lambda_1+\lambda_2)^k}{k!}$构造多项式, 无$i$, 可提出加总符号)

正态分布的可加性

若 $X \sim N(\mu_1, \sigma_1^2), Y \sim N(\mu_2, \sigma_2^2)$, 且独立, 则
$Z = X + Y \sim N(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2)$.

证明: 利用卷积公式可得

$$p_Z(z) = \frac{1}{2\pi\sigma_1\sigma_2}\int_{-\infty}^{\infty} \exp\left\{-\frac{1}{2}\left[\frac{(z-y-\mu_1)^2}{\sigma_1^2} + \frac{(y-\mu_2)^2}{\sigma_2^2}\right]\right\} \mathrm{d}y$$

对上式被积函数的指数部分按y的幂次展开, 再合并同类项, 可推出

$$\frac{(z-y-\mu_1)^2}{\sigma_1^2} + \frac{(y-\mu_2)^2}{\sigma_2^2} = A\left(y-\frac{B}{A}\right)^2 + \frac{(z-\mu_1-\mu_2)^2}{\sigma_1^2+\sigma_2^2}$$

其中

$$A = \frac{1}{\sigma_1^2} + \frac{1}{\sigma_2^2}, \quad B = \frac{z-\mu_1}{\sigma_1^2} + \frac{\mu_2}{\sigma_2^2}$$

带回原式, 可得

$$p_Z(z) = \frac{1}{2\pi\sigma_1\sigma_2} \exp\left\{-\frac{1}{2}\frac{(z-\mu_1-\mu_2)^2}{\sigma_1^2+\sigma_2^2}\right\}$$
$$\cdot \int_{-\infty}^{\infty} \exp\left\{-\frac{A}{2}\left(y-\frac{B}{A}\right)^2\right\} \mathrm{d}y$$

利用正态密度函数的正则性(构造积分为1的密度函数形式), 上式中的积分应为$\sqrt{2\pi}/\sqrt{A}$, 于是

$$p_Z(z) = \frac{1}{\sqrt{2\pi\left(\sigma_1^2+\sigma_2^2\right)}} \exp\left\{-\frac{1}{2}\frac{(z-\mu_1-\mu_2)^2}{\sigma_1^2+\sigma_2^2}\right\}$$

即,$N(\mu_1+\mu_2, \sigma_1^2+\sigma_2^2)$, $Q.E.D.$

注意

- 可推广至有限独立正态分布之和;
- $X - Y$ **不服从** $N(\mu_1 - \mu_2, \sigma_1^2 - \sigma_2^2)$;
- $X - Y$ **服从** $N(\mu_1 - \mu_2, \sigma_1^2 + \sigma_2^2)$.

注意

- 可推广至有限独立正态分布之和;
- $X - Y$ **不服从** $N(\mu_1 - \mu_2, \sigma_1^2 - \sigma_2^2)$;
- $X - Y$ **服从** $N(\mu_1 - \mu_2, \sigma_1^2 + \sigma_2^2)$. **(如何证明?)**

注意

- 可推广至有限独立正态分布之和;
- $X-Y$ **不服从** $N(\mu_1-\mu_2,\sigma_1^2-\sigma_2^2)$;
- $X-Y$ **服从** $N(\mu_1-\mu_2,\sigma_1^2+\sigma_2^2)$. **(如何证明?)** (看成$X+(-Y)$)

注意

- 可推广至有限独立正态分布之和;
- $X-Y$ **不服从** $N(\mu_1-\mu_2,\sigma_1^2-\sigma_2^2)$;
- $X-Y$ **服从** $N(\mu_1-\mu_2,\sigma_1^2+\sigma_2^2)$. **(如何证明?)** (看成$X+(-Y)$)
- **注意：** 如果X与Y是独立同分布的标准正态变量, $Z=X+Y$和$Z=X-Y$的分布均为 $N(0,2)$.

推论: 独立正态变量的线性组合仍为正态变量

由于我们已知若$X \sim N(\mu, \sigma^2)$, 则$aX \sim N(a\mu, a^2\sigma^2)$, 所以

若记$X_i \sim N(\mu_i, \sigma_i^2), i = 1, 2, ..., n$. 且$X_i$ 间相互独立, 实数$a_1, a_2, ..., a_n$不全为零, 则任意正态分布的线性组合

$$\sum_{i=1}^{n} a_i X_i \sim N\left(\sum_{i=1}^{n} a_i \mu_i, \sum_{i=1}^{n} a_i^2 \sigma_i^2\right)$$

课堂练习

设$X \sim N(0,1)$, $Y \sim N(0,1)$, 且独立, 求

1) $X+X$

2) $X+Y$

3) $X-X$

4) $X-Y$

课堂练习

设$X \sim N(0,1)$, $Y \sim N(0,1)$, 且独立, 求

1) $X+X$

2) $X+Y$

3) $X-X$

4) $X-Y$

思考： $X+X \sim N(0,4)$和$X+Y \sim N(0,2)$分布不同的原因是什么?

课堂练习

设$X \sim N(0,1)$, $Y \sim N(0,1)$, 且独立, 求

1) $X+X$

2) $X+Y$

3) $X-X$

4) $X-Y$

思考： $X+X \sim N(0,4)$和$X+Y \sim N(0,2)$分布不同的原因是什么? (X和X不独立!)

伽玛分布的可加性

若$X \sim Ga(\alpha_1, \lambda), Y \sim Ga(\alpha_2, \lambda)$, 且独立,
则$Z = X + Y \sim Ga(\alpha_1 + \alpha_2, \lambda)$.

注意:

- 可推广至有限独立伽玛分布之和;
- $X - Y$不服从 $Ga(\alpha_1 - \alpha_2, \lambda)$.
- 由于指数分布是伽马分布的特例$Exp(\lambda) = Ga(1, \lambda)$, 所以$m$个独立同分布的指数变量之和为伽马分布$Exp(\lambda) * ... * Exp(\lambda) = Ga(m, \lambda)$

同样利用卷积公式证明, 见教材.

(**注意:** 教材中符号$*$ 表示分布的和.
$Exp(\lambda) * ... * Exp(\lambda)$与$X_1 + ... + X_n$且$X_1 \sim Exp(\lambda)$...相同.)

χ^2 分布的可加性

若$X \sim \chi^2(n_1), Y \sim \chi^2(n_2)$, 且独立,

则 $Z = X + Y \sim \chi^2(n_1 + n_2)$.

注意:

- $X - Y$ 不服从 χ^2 分布;
- 若 $X_i \sim N(0, 1)$, 且独立, 则$Z = \sum_{i=1}^{n} X_i^2 \sim \chi^2(n)$;
- χ^2 分布同样是伽玛分布的特例, 证明同上.

注意点

- 独立的$0-1$分布随机变量之和服从二项分布. (可从定义直接推出)

课堂练习(5')

设 X 与 Y 独立, $X \sim U(0,1)$, $Y \sim Exp(1)$. 试求 $Z = X + Y$ 的密度函数.

解:

课堂练习(5')

设 X 与 Y 独立, $X \sim U(0,1)$, $Y \sim Exp(1)$. 试求 $Z = X + Y$ 的密度函数.

解:

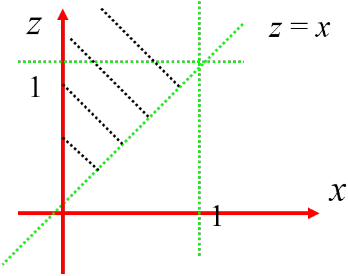
$$X \sim p_X(x) = \begin{cases} 1, & 0 \le x \le 1. \\ 0, & \text{其他}. \end{cases} \qquad Y \sim p_Y(y) = \begin{cases} e^{-y}, & y > 0. \\ 0, & y \le 0. \end{cases}$$

用卷积公式:

$$p_Z(z) = \int_{-\infty}^{+\infty} p_X(x) p_Y(z-x) dx$$

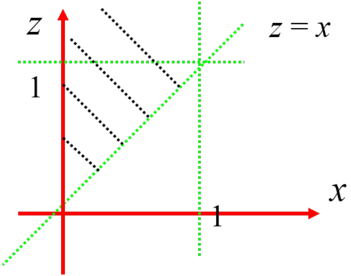
被积函数的非零区域为: $0 < x < 1$且 $z - x > 0$

(见下图)



因此有

(1) $z<0$时 $p_Z(z)=0$;
(2) $0<z<1$ 时 $p_Z(z)=\int_0^z e^{-(z-x)}dx=1-e^{-z}$
(3)$1<z$ 时$p_Z(z)=\int_0^1 e^{-(z-x)}dx=e^{-z}(e-1)$
思考： 为什么卷积上下界是0和1?



因此有

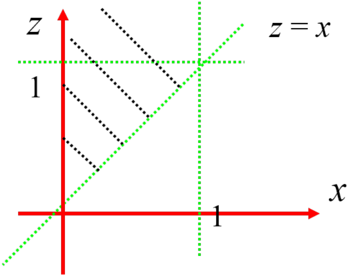
(1) $z < 0$时 $p_Z(z) = 0$;

(2) $0 < z < 1$ 时 $p_Z(z) = \int_0^z e^{-(z-x)}dx = 1 - e^{-z}$

(3)$1 < z$ 时$p_Z(z) = \int_0^1 e^{-(z-x)}dx = e^{-z}(e-1)$

思考： 为什么卷积上下界是0和1? (因为被积变量x需要满足$x \leq 1$且 $x < z$)

思考： 卷积公式的上下界如何确定?



因此有

(1) $z < 0$时 $p_Z(z) = 0$;

(2) $0 < z < 1$ 时 $p_Z(z) = \int_0^z e^{-(z-x)}dx = 1 - e^{-z}$

(3)$1 < z$ 时$p_Z(z) = \int_0^1 e^{-(z-x)}dx = e^{-z}(e-1)$

思考： 为什么卷积上下界是0和1? (因为被积变量x需要满足$x \leq 1$且 $x < z$)

思考： 卷积公式的上下界如何确定? 根据被积变量可取值.

3.3.4 变量变换法(本节难点)

已知 (X, Y) 的联合密度函数为$p(x, y)$. 如有关于(X, Y) 的函数

$$\begin{cases} U = g_1(X, Y) \\ V = g_2(X, Y) \end{cases}$$

求 (U, V) 的联合分布.

3.3.4 变量变换法(本节难点)

若

$$\begin{cases} u = g_1(x, y) \\ v = g_2(x, y) \end{cases}$$

有连续偏导, 存在反函数

$$\begin{cases} x = x(u, v) \\ y = y(u, v) \end{cases}$$

则 (U, V) 的联合密度为 $P_{UV}(u, v) = p_{XY}(x(u, v), y(u, v))|J|$

$|J|$为绝对值, 其中J为变换的雅可比行列式

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

证明略.

记忆技巧

(X, Y) 的联合密度函数为$p(x, y)$. 如有关于(X, Y) 的函数

$$\begin{cases} U = g_1(X, Y) \\ V = g_2(X, Y) \end{cases}$$

(U, V) 的联合密度为 $P_{UV}(u, v) = p_{XY}(x(u, v), y(u, v))|J|$

- 注意一维随机变量函数的分布求法在2.6已介绍, 公式形式公式形式$p_Y(y) = p_X[h(y)] \cdot |h'(y)|$, 本节介绍的二维随机变量的联合密度函数求法可以看作是对一维随机变量函数求解的拓展.
- 可对比2.6节一维随机变量函数密度函数求解公式形式$p_Y(y) = p_X[h(y)] \cdot |h'(y)|$记忆

例 3.3.10 设随机变量X与Y独立同分布, 都服从正态分布$N\left(\mu,\sigma^2\right)$. 记

$$\begin{cases} U = X + Y \\ V = X - Y \end{cases}$$

试求(U,V)的联合密度函数, 且问U与V是否独立?

独立条件?

例 3.3.10 设随机变量X与Y独立同分布, 都服从正态分布$N\left(\mu,\sigma^2\right)$. 记

$$\begin{cases} U = X + Y \\ V = X - Y \end{cases}$$

试求(U,V)的联合密度函数, 且问U与V是否独立?

独立条件? 联合分布等于边际分布之积.

解： 因为

$$\begin{cases} u = x + y \\ v = x - y \end{cases}$$

的反函数为

$$\begin{cases} x = \frac{u+v}{2} \\ y = \frac{u-v}{2} \end{cases}$$

则

$$J = \begin{vmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} \end{vmatrix} = \begin{vmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} \end{vmatrix} = -\frac{1}{2}$$

所以得 (U, V) 的联合密度函数为

$$p(u, v) = p(x(u, v), y(u, v))|J| = p_X\left(\frac{u+v}{2}\right) p_Y\left(\frac{u-v}{2}\right)\left|-\frac{1}{2}\right|$$

$$= \frac{1}{4\pi\sigma^2} \exp\left\{-\frac{(u-2\mu)^2 + v^2}{4\sigma^2}\right\}$$

例3.3.10

所以得 (U,V) 的联合密度函数为

$$p(u,v)=p(x(u,v),y(u,v))|J|=p_X\left(\frac{u+v}{2}\right)p_Y\left(\frac{u-v}{2}\right)\left|-\frac{1}{2}\right|$$
$$=\frac{1}{4\pi\sigma^2}\exp\left\{-\frac{(u-2\mu)^2+v^2}{4\sigma^2}\right\}$$

这正是二元正态分布 $N\left(2\mu,0,2\sigma^2,2\sigma^2,0\right)$ 的密度函数, 其边际分布为 $U\sim N\left(2\mu,2\sigma^2\right)$, $V\sim N\left(0,2\sigma^2\right)$, 所以由 $p(u,v)=p_U(u)p_v(v)$ 知 U 与 V 相互独立.

3.3.4 变量变换法应用: 增补变量法

若要求 $U = g_1(X,Y)$ 的密度 $p_U(u)$, 如果直接求解不便,

可增补一个变量$V = g_2(X,Y)$, 一般令$V = X$或$V = Y$.

先用变量变换法求出(U,V)的联合密度$p_{UV}(u,v)$,

然后再由联合密度$p_{UV}(u,v)$, 去求出边际密度$p_U(u)$.

用此方法可以求出随机变量积和商的分布公式.

例3.3.11 随机变量乘积的分布公式

例3.3.11(积的公式) 设随机变量X与Y相互独立，其密度函数分别为$p_X(x)$和$p_Y(y)$. 则$U=XY$的密度函数为

$$p_U(u)=\int_{-\infty}^{\infty}p_X\left(\frac{u}{v}\right)p_Y(v)\frac{1}{|v|}\mathrm{d}v$$

证明： 记 $V=Y$, 则 $\begin{cases} u=xy, \\ v=y \end{cases}$ 的反函数为 $\begin{cases} x=\frac{u}{v}, \\ y=v, \end{cases}$ 雅可比行列式为

$$J=\frac{\partial(x,y)}{\partial(u,v)}=\begin{vmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} \end{vmatrix}=\begin{vmatrix} \frac{1}{v} & -\frac{u}{v^2} \\ 0 & 1 \end{vmatrix}=\frac{1}{v}$$

所以 (U,V) 的联合密度函数为

$$p(u,v)=p_X\left(\frac{u}{v}\right)\cdot p_Y(v)|J|=p_X\left(\frac{u}{v}\right)p_Y(v)\frac{1}{|v|}$$

对 $p(u,v)$ 关于 v 积分，就可得 $U=XY$ 的密度函数为上式.

例 3.3.12 (商的公式) 设随机变量X与Y相互独立, 其密度函数分别为$p_X(x)$和$p_Y(y)$. 则$U = X/Y$的密度函数为

$$p_U(u) = \int_{-\infty}^{\infty} p_X(uv)p_Y(v)|v|\mathrm{d}v$$

证明: 记$V = Y$, 则$\begin{cases} u = x/y, \\ v = y \end{cases}$ 的反函数为 $\begin{cases} x = uv, \\ y = v, \end{cases}$ 雅可比行列式为

$$J = \begin{vmatrix} v & u \\ 0 & 1 \end{vmatrix} = v$$

所以 (U, V) 的联合密度函数为

$$p(u, v) = p_X(uv) \cdot p_Y(v)|J| = p_X(uv) \cdot p_Y(v)|v|$$

对 $p(u, v)$ 关于 v 积分, 就可得 $U = X/Y$ 的密度函数为上式.

在以上两个例子中, 如果X与Y不是相互独立的, 只需将边际密度的乘积改成联合密度即可.

第11次作业

习题3.3中题目1, 2, 4, 6, 7, 9, 10, 11, 13, 14, 16, 17, 21

§3.4 多维随机变量的特征数

本节讨论

- 多维随机变量的**数学期望**和**方差**
- X 与 Y 的**协方差**(covariance) 和**相关系数** (Pearson correlation coefficient) (反映两个随机变量的关联程度)

定理 3.4.1 设 (X,Y) 是二维随机变量, $Z=g(X,Y)$, 则

$$E(Z)=E[g(X,Y)]=\begin{cases}\sum\limits_{i}\sum\limits_{j}g(x_i,y_i)p_{ij}\\ \int\limits_{-\infty}^{+\infty}\int\limits_{-\infty}^{+\infty}g(x,y)p(x,y)dxdy\end{cases}$$

证明略.

上述定理可拓展至n维随机变量.

3.4.1 多维随机变量函数的数学期望

定理 3.4.1 设 (X, Y) 是二维随机变量, $Z = g(X, Y)$, 则

$$E(Z) = E[g(X,Y)] = \begin{cases} \sum\limits_i \sum\limits_j g(x_i, y_i) p_{ij} \\ \int\limits_{-\infty}^{+\infty} \int\limits_{-\infty}^{+\infty} g(x, y) p(x, y) dxdy \end{cases}$$

证明略.

上述定理可拓展至n维随机变量.

思考: 为何上述的定理是**定理**而非**定义**? 多维随机变量函数的数学期望如何用**定义**求解?

例3.4.1 在长为a的线段上任取两点X与 Y, 求两点间的平均长度.

解:

例3.4.1 在长为a的线段上任取两点X与 Y, 求两点间的平均长度.

解: 该问题是求$E(|X-Y|)$.

例3.4.1 在长为a的线段上任取两点X与 Y, 求两点间的平均长度.

解: 该问题是求$E(|X-Y|)$. 因为X, Y相互独立且服从均匀分布, 所以(X, Y)的联合密度函数为

$$p(x,y)=\begin{cases}\frac{1}{a^2}, & 0<x<a, \quad 0<y<a\\ 0, & \text{其他}\end{cases}$$

例3.4.1 在长为a的线段上任取两点X与 Y, 求两点间的平均长度.

解: 该问题是求$E(|X-Y|)$. 因为X, Y相互独立且服从均匀分布, 所以(X, Y)的联合密度函数为

$$p(x,y)=\begin{cases}\frac{1}{a^2}, & 0<x<a, \quad 0<y<a\\ 0, & \text{其他}\end{cases}$$

所以

$$\begin{aligned}E(|X-Y|) &= \int_0^a\int_0^a |x-y|\frac{1}{a^2}\mathrm{d}x\mathrm{d}y\\ &= \frac{1}{a^2}\left\{\int_0^a\int_0^x (x-y)\mathrm{d}y\mathrm{d}x+\int_0^a\int_x^a (y-x)\mathrm{d}y\mathrm{d}x\right\}\end{aligned}$$

(对$x>y$和$x<y$的情况分别求积分)

$$= \frac{1}{a^2}\left\{\int^a\left(x^2-ax+\frac{a^2}{2}\right)\mathrm{d}x\right\}=\frac{a}{3}$$

1. $E(X+Y)=E(X)+E(Y)$ **(性质3.4.1)**

证明1: 在连续情况下证明 $E(X+Y)=E(X)+E(Y)$, 离散情况同理

$$
\begin{aligned}
E(X+Y) &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}(x+y)p(x,y)\mathrm{d}x\mathrm{d}y \\
&= \int_{-\infty}^{\infty} x\left\{\int_{-\infty}^{\infty} p(x,y)\mathrm{d}y\right\}\mathrm{d}x + \int_{-\infty}^{\infty} y\left\{\int_{-\infty}^{\infty} p(x,y)\mathrm{d}x\right\}\mathrm{d}y \\
&= \int_{-\infty}^{\infty} xp_X(x)\mathrm{d}x + \int_{-\infty}^{\infty} yp_Y(y)\mathrm{d}y \\
&= E(X)+E(Y)
\end{aligned}
$$

2. 当X与Y独立时, $E(XY)=E(X)E(Y)$, (性质3.4.2)

证明2: 在连续情况下证明当X与Y独立时, $E(XY)=E(X)E(Y)$, 离散情况同理

$$\begin{aligned}E(XY)&=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}xyp_{XY}(x,y)\mathrm{d}x\mathrm{d}y\\&=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}xyp_X(x)p_Y(y)\mathrm{d}x\mathrm{d}y\\&=\int_{-\infty}^{\infty}xp_X(x)\mathrm{d}x\int_{-\infty}^{\infty}yp_Y(y)\mathrm{d}y\\&=E(X)E(Y)\end{aligned}$$

注意: 反之不成立! $E(XY)=E(X)E(Y)$时, X与Y不相关但不一定独立(稍后证明).

讨论$X+Y$的方差

1. $Var(X \pm Y) = Var(X) + Var(Y) \pm 2E[X - E(X)][Y - E(Y)]$
2. $E[X - E(X)][Y - E(Y)] = E(XY) - E(X)E(Y)$
3. 当X与Y独立时, $E[X - E(X)][Y - E(Y)] = 0$
4. 当X与Y独立时, $Var(X \pm Y) = Var(X) + Var(Y)$

- 3.和4.命题反之不成立.
- $E[X - E(X)][Y - E(Y)]$即为X和Y的协方差.

证明:

$$\begin{aligned}\mathrm{Var}(X \pm Y) &= E(X \pm Y - E(X \pm Y))^2 \\ &= E((X - E(X)) \pm (Y - E(Y)))^2 \\ &= \mathrm{Var}(X) + \mathrm{Var}(Y) \pm 2E(X - E(X))(Y - E(Y))\end{aligned}$$

在X与Y独立时, 最后一项因为独立性为0.

3.4.3 协方差(covariance)

定义3.4.1 称$Cov(X,Y)=E[X-E(X)][Y-E(Y)]$为$X$与$Y$的**协方差**, 又称相关(中心)矩.

- 协方差描述两个随机变量的关联程度
- $Cov(X,Y)>0$, 正相关 (随着X增大, Y 有增大的趋势)
- $Cov(X,Y)<0$, 负相关
- $Cov(X,Y)=0$, 不相关
- 不相关比独立弱.
- 相关不一定因果.

拓展内容

思考如下事件的相关性, 独立性, 和因果性:

教育水平和收入; 个人能力和收入; 出生日期和个人能力; 出生日期和教育水平.

拓展内容

思考如下事件的相关性, 独立性, 和因果性:

教育水平和收入; 个人能力和收入; 出生日期和个人能力; 出生日期和教育水平.

(拓展内容) 观测到相关是很容易的, 从相关中分离因果关系是非常难的.

协方差的性质

性质3.4.4 $Cov(X,Y)=E(XY)-E(X)E(Y)$.

证明:

$$\begin{aligned}\mathrm{Cov}(X,Y)&=E[XY-XE(Y)-YE(X)+E(X)E(Y)]\\&=E(XY)-E(X)E(Y)\end{aligned}$$

性质3.4.5 若 X与Y 独立, 则 $Cov(X,Y)=0$, 反之不然

证明: 独立性可推出$E(XY)=E(X)E(Y)$, 则 $Cov(X,Y)=0$.

反之可用反例, 设随机变量$X\sim N(0,\sigma^2)$, $Y=X^2$, 则X与Y不独立, 但

$$\mathrm{Cov}(X,Y)=\mathrm{Cov}\left(X,X^2\right)=E\left(X\cdot X^2\right)-E(X)E\left(X^2\right)=0$$

性质3.4.6 $Var(X \pm Y) = Var(X) + Var(Y) \pm 2Cov(X, Y)$

证明:

$$\begin{aligned}
&\mathrm{Var}(X \pm Y) = E\{(X \pm Y) - E(X \pm Y)\}^2 \\
&= E\{(X - E(X)) \pm (Y - E(Y))\}^2 \\
&= E\left\{(X - E(X))^2 + (Y - E(Y))^2 \pm 2(X - E(X))(Y - E(Y))\right\} \\
&= \mathrm{Var}(X) + \mathrm{Var}(Y) \pm 2\,\mathrm{Cov}(X, Y)
\end{aligned}$$

对n个随机变量有

$$\mathrm{Var}\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} \mathrm{Var}(X_i) + 2\sum_{i=1}^{n}\sum_{j=1}^{i-1} \mathrm{Cov}(X_i, X_j)$$

若X, Y不相关, 有$Var(X \pm Y) = Var(X) + Var(Y)$.

协方差的性质

性质3.4.7 $Cov(X, Y) = Cov(Y, X)$.

性质3.4.8 $Cov(X, a) = 0$.

性质3.4.9 $Cov(aX, bY) = abCov(X, Y)$.

证明:

$$\mathrm{Cov}(aX, bY) = E\{(aX - E(aX))(bY - E(bY))\} = abCov(X, Y)$$

性质3.4.10 $Cov(X + Y, Z) = Cov(X, Z) + Cov(Y, Z)$.

证明:

$$\begin{aligned}\mathrm{Cov}(X+Y, Z) &= E[(X+Y)Z] - E(X+Y)E(Z) \\ &= E(XZ) + E(YZ) - E(X)E(Z) - E(Y)E(Z) \\ &= [E(XZ) - E(X)E(Z)] + [E(YZ) - E(Y)E(Z)] \\ &= \mathrm{Cov}(X, Z) + \mathrm{Cov}(Y, Z)\end{aligned}$$

课堂练习1

X 与 Y 独立,$Var(X) = 6, Var(Y) = 3,$

则 $Var(2X - Y)$ = ( ).

解:

课堂练习1

X 与Y 独立,$Var(X)=6$,$Var(Y)=3$,

则 $Var(2X-Y)$ = ( ).

解： 27

课堂练习2

$X \sim P(2), Y \sim N(-2,4)$, X与Y独立, 求

$E(X-Y)$= ( );

$E(X-Y)^2$= ( ).

解:

课堂练习2

$X \sim P(2), Y \sim N(-2, 4), X$与$Y$独立, 求

$E(X-Y)$= ( );

$E(X-Y)^2$= ( ).

解:

因为$E(X) = 2, Var(X) = 2, E(Y) = -2, Var(Y) = 4$ 所以:
$E(X-Y) = E(X) - E(Y) = 4$

$E(X-Y)^2 = E(X^2) + E(Y^2) - 2E(XY) = E^2(X) + Var(X) + E^2(Y) + Var(Y) - 2(Cov(X, Y) + E(X)E(Y)) = 22$

拓展知识：配对模型的数学期望和方差

n个人, n件礼物, 任意取. X为拿对自己礼物的人数, 求 $E(X), Var(X)$

解:

拓展知识：配对模型的数学期望和方差

n个人, n件礼物, 任意取. X为拿对自己礼物的人数, 求 $E(X), Var(X)$

解:

解: 记 "$X_i = 1$" = "第 i 个人拿对自己的礼物", "$X_i = 0$" = "第 i 个人未拿对自己的礼物"

则$X = \sum_{i=1}^{n} X_i$, 因为 $E(X_i) = 1/n$, 所以 $E(X) = 1$.

又因为$Var(X) = \sum_{i=1}^{n} Var(X_i) + 2\sum_{i<j} Cov(X_i, X_j)$

所以先计算 $E(X_iX_j)$, X_iX_j的分布列为

X_iX_j	0	1
P	$1-1/[n(n-1)]$	$1/[n(n-1)]$

所以 $E(X_iX_j)=1/[n(n-1)]$, 由此得

$$Cov(X_i,X_j)=E(X_iX_j)-E(X_i)E(X_j)=\frac{1}{n(n-1)}-\frac{1}{n^2}=$$
$$\frac{1}{n^2(n-1)}$$

又因为

$$Var(X_i)=E(X_i^2)-[E(X_i)]^2=\frac{1}{n}-\frac{1}{n^2}=\frac{n-1}{n^2}$$

(上式推导利用随机变量函数期望性质: 因为X_i^2只能取0和1)

所以

$$
\begin{aligned}
Var(X) =& \sum_{i=1}^{n} Var(X_i) + 2\sum_{i<j} Cov(X_i, X_j) \\
=& n\frac{n-1}{n^2} + 2\sum_{i<j} \frac{1}{n^2(n-1)} \\
=& \frac{n-1}{n} + 2\binom{n}{2}\frac{1}{n^2(n-1)} \\
=& \frac{n-1}{n} + \frac{1}{n} \\
=& 1
\end{aligned}
$$

3.4.4 相关系数(correlation coefficient)

协方差有量纲, 为了消除量纲, 引入相关系数的概念.

定义3.4.2 称

$$Corr(X,Y)=\frac{Cov(X,Y)}{\sqrt{Var(X)}\sqrt{Var(Y)}}$$

为 X 与 Y 的**相关系数**.

注意点

若记

$$X^* = \frac{X - E(X)}{\sqrt{Var(X)}} = \frac{X - \mu_X}{\sigma_X}, Y^* = \frac{Y - E(Y)}{\sqrt{Var(Y)}} = \frac{Y - \mu_Y}{\sigma_Y}$$

则 X 与 Y 的**相关系数为相应标准化随机变量**X^*, Y^***的协方差**:

$$Corr(X, Y) = Cov(X^*, Y^*)$$

证明:

$$\begin{aligned}\mathrm{Cov}\left(X^*, Y^*\right) = \mathrm{Cov}\left(\frac{X - \mu_X}{\sigma_X}, \frac{Y - \mu_Y}{\sigma_Y}\right) \\ = \frac{\mathrm{Cov}(X, Y)}{\sigma_X \sigma_Y} \\ = \mathrm{Corr}(X, Y)\end{aligned}$$

相关系数的性质

1. 二维正态分布$N\left(\mu_1,\mu_2,\sigma_1^2,\sigma_2^2,\rho\right)$的相关系数就是$\rho$, 证明见教材例3.4.9.

2. **施瓦茨不等式**: $[Cov(X,Y)]^2 \leq Var(X)Var(Y)$.

证明: 当$\sigma_X^2=0$时, X几乎处处为常数, 因而与Y的协方差为0, 不等式两侧均为0, 结论成立. 当$\sigma_X^2>0$, 引入变量t考虑二次函数

$$g(t)=E[t(X-E(X))+(Y-E(Y))]^2=t^2\sigma_x^2+2t\cdot\mathrm{Cov}(X,Y)+\sigma_y^2$$

由于上述二次函数非负, 所以判别式小于或等于0, 即

$$[2\,\mathrm{Cov}(X,Y)]^2-4\sigma_X^2\sigma_Y^2\leqslant 0$$

移项后即得施瓦茨不等式.

3. **性质3.4.11** $-1\leq Corr(X,Y)\leq 1$. **证明:** 利用施瓦茨不等式和相关系数定义.

相关系数的性质

4. **性质3.4.12** $Corr(X,Y)=\pm 1 \Leftrightarrow X$ 与 Y 几乎处处有线性关系. 即存在a,b, 使得 $P(Y=aX+b)=1$.

证明见教材.

注意： $Corr(X,Y)$ 的大小反映了X与Y之间的线性关系:

- $Corr(X,Y)=1$, X 与 Y完全正相关.
- $Corr(X,Y)=-1$, X 与 Y完全负相关.
- $Corr(X,Y)=0$, X 与 Y不相关, 没有线性关系.
- $0<|Corr(X,Y)|<1$, X 与 Y有一定程度的线性关系.

例3.4.1 设(X,Y)的联合分布列为

$X \backslash Y$	-1	0	1
-1	$1/8$	$1/8$	$1/8$
0	$1/8$	0	$1/8$
1	$1/8$	$1/8$	$1/8$

求 X,Y 的相关系数.

解:

例3.4.1 设(X,Y)的联合分布列为

$X \backslash Y$	-1	0	1
-1	$1/8$	$1/8$	$1/8$
0	$1/8$	0	$1/8$
1	$1/8$	$1/8$	$1/8$

求 X,Y 的相关系数.

解:

$$E(X)=\sum_i\sum_j x_i p_{ij}=0$$

$$E(X^2)=\sum_i\sum_j x_i^2 p_{ij}=3/4$$

同理 $E(Y) = E(X) = 0$

$E(Y^2) = E(X^2) = 3/4$

另一方面$E(XY) = 1/8 - 1/8 - 1/8 + 1/8 = 0$

所以$Cov(X, Y) = E(XY) - E(X)E(Y) = 0$

即 $Corr(X, Y) = 0$

$$(X,Y)\sim p(x,y)=\begin{cases}\dfrac{1}{8}(x+y), & 0<x<2, 0<y<2\\ 0, & \text{其他}\end{cases}$$

求 X, Y 的相关系数.

解:

$$(X,Y) \sim p(x,y) = \begin{cases} \frac{1}{8}(x+y), & 0<x<2, 0<y<2 \\ 0, & \text{其他} \end{cases}$$

求 X, Y 的相关系数.

解:

$$E(X) = E(Y) = \int_0^2 \int_0^2 x \frac{1}{8}(x+y)dxdy = 7/6$$

$$E(X^2) = E(Y^2) = \int_0^2 \int_0^2 x^2 \frac{1}{8}(x+y)dxdy = 5/3$$

所以, $Var(X) = Var(Y) = 11/36$

课堂练习(5')

(continued)

$$E(XY) = \int_0^2 \int_0^2 xy\frac{1}{8}(x+y)dxdy = 4/3$$

$$Corr(X,Y) = \frac{4/3 - 7/6 \times 7/6}{11/36} = -\frac{1}{11}$$

二维正态分布的特征数总结

$(X, Y) \sim N(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho)$

(1) $X \sim N(\mu_1, \sigma_1^2), Y \sim N(\mu_2, \sigma_2^2)$

(2) 参数 ρ为 X 和 Y 的相关系数

(3) X, Y 独立 $\Leftrightarrow \rho = 0$

(4) **在二维正态分布的场合, 不相关与独立等价.** (性质3.4.13) 证明见教材.

定义3.4.3 记n维随机向量 $\mathbf{X} = (X_1, X_2, \ldots, X_n)'$, 则称 $E(\mathbf{X}) = (E(X_1), E(X_2), \ldots, E(X_n))'$ 为n维随机向量$\mathbf{X}$的期望向量, 简称为$\mathbf{X}$的数学期望. 称

$$E[(\mathbf{X} - E(\mathbf{X}))(\mathbf{X} - E(\mathbf{X}))'] =$$

$$\begin{pmatrix} Var(X_1) & Cov(X_1, X_2) & \ldots & Cov(X_1, X_n) \\ Cov(X_2, X_1) & Var(X_2) & \ldots & Cov(X_2, X_n) \\ \vdots & \vdots & \ddots & \vdots \\ Cov(X_n, X_1) & Cov(X_n, X_2) & \ldots & Var(X_n) \end{pmatrix}$$

为$\mathbf{X}$的**方差-协方差矩阵**(Variance-Covariance Matrix), 简称**协方差阵**, 记为 $Cov(\mathbf{X})$, 或Σ.

随机向量和随机向量期望为$n \times 1$, 方差协方差阵为$n \times n$.

协方差阵的性质

定理3.4.2 协方差阵$Cov(\mathbf{X}) = (Cov(X_i, X_j))_{n\times n}$对称, 非负定.

证明: 对称是显然的, 下面证非负定. 对任意n维实向量$\mathbf{c} = (c_1, c_2, \cdots, c_n)'$, 有

$$
\begin{aligned}
& c' \operatorname{Cov}(\mathbf{X}) c \\
& = (c_1, c_2, \cdots, c_n) \begin{pmatrix} \operatorname{Var}(X_1) & \cdots & \operatorname{Cov}(X_1, X_n) \\ \operatorname{Cov}(X_2, X_1) & \cdots & \operatorname{Cov}(X_2, X_n) \\ \vdots & & \vdots \\ \operatorname{Cov}(X_n, X_1) & \cdots & \operatorname{Var}(X_n) \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{pmatrix} \\
& = \sum_{i=1}^{n} \sum_{j=1}^{n} E\{[c_i (X_i - E(X_i))] [c_j (X_j - E(X_j))]\}
\end{aligned}
$$

(利用期望性质: 和的期望等于期望的和)

$$
\begin{aligned}
&= E\left\{\sum_{i=1}^{n}\sum_{j=1}^{n}\left[c_i\left(X_i - E\left(X_i\right)\right)\right]\left[c_j\left(X_j - E\left(X_j\right)\right)\right]\right\} \\
&= E\left\{\left[\sum_{i=1}^{n} c_i\left(X_i - E\left(X_i\right)\right)\right]\left[\sum_{j=1}^{n} c_j\left(X_j - E\left(X_j\right)\right)\right]\right\} \\
&= E\left[\sum_{i=1}^{n} c_i\left(X_i - E\left(X_i\right)\right)\right]^2 \geqslant 0
\end{aligned}
$$

非负定. *Q.E.D.*

n元正态分布: 记n维随机向量$\mathbf{X}=(X_1,X_2,\ldots,X_n)'$的协方差矩阵为$\mathbf{B}=Cov(\mathbf{X})$, 数学期望向量为$\mathbf{a}=(a_1,a_2,\cdots,a_n)'$. 记$\mathbf{x}=(x_1,x_2,\cdots,x_n)'$, 则由密度函数

$$p(x_1,x_2,\cdots,x_n)=p(\mathbf{x})=\frac{1}{(2\pi)^{\frac{n}{2}}|\mathbf{B}|^{\frac{1}{2}}}\exp\left\{-\frac{1}{2}(\mathbf{x}-\mathbf{a})'\mathbf{B}^{-1}(\mathbf{x}-\mathbf{a})\right\}$$

定义的分布称为n**元正态分布**, 记为$\mathbf{X}\sim N(\mathbf{a},\mathbf{B})$.

$|\mathbf{B}|$表示$\mathbf{B}$的行列式, $\mathbf{B}^{-1}$ 表示$\mathbf{B}$的逆矩阵.

当$n=2$, 若取

$$\mathbf{a}=\begin{pmatrix}\mu_1\\ \mu_2\end{pmatrix},\quad \mathbf{B}=\begin{pmatrix}\sigma_1^2 & \sigma_1\sigma_2\rho\\ \sigma_1\sigma_2\rho & \sigma_2^2\end{pmatrix}$$

上式褪化为二元正态密度函数.

课堂练习 (1')

设 $X \sim N(0,1), Y \sim N(0,1), Var(X-Y)=0$, 求$(X,Y)$的协方差阵$\Sigma$.

解:

课堂练习 (1')

设 $X \sim N(0,1), Y \sim N(0,1), Var(X-Y)=0$, 求$(X,Y)$的协方差阵$\Sigma$.

解： $Var(X-Y)=Var(X)+Var(Y)-2Cov(X,Y)$

$$\Sigma = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$$

第12次作业

习题3.4中题目2, 3, 5, 7, 10, 13, 14, 15, 16, 17, 20, 21, 25, 27, 29, 35

§3.5 条件分布与条件期望

本节学习

- 条件分布定义, 离散和连续随机变量的条件分布
- 全概率公式和贝叶斯公式
- 条件数学期望
- 重期望公式 (Law of Iterated Expectation, LIE)

3.5.1 条件分布 (conditional distribution)

对二维随机变量(X, Y),

- 称在给定Y取某个值的条件下, X的分布为X的**条件分布**(conditional distribution);
- 称在给定X取某个值的条件下, Y的分布为Y的**条件分布**.

注意区别无条件分布(unconditional distribution)

离散随机变量的条件分布

设二维随机变量(X, Y)的联合分布列为

$$p_{ij} = P\left(X = x_i, Y = y_j\right), \quad i = 1, 2, \cdots, \quad j = 1, 2, \cdots$$

定义3.5.1 对一切使$P\left(Y = y_j\right) = p_{\cdot j} = \sum_{i=1}^{\infty} p_{ij} > 0$ 的y_j, 称

$$p_{i|j} = P\left(X = x_i | Y = y_j\right) = \frac{P\left(X = x_i, Y = y_j\right)}{P\left(Y = y_j\right)} = \frac{p_{ij}}{p_{\cdot j}}, \quad i = 1, 2, \cdots$$

为给定$Y = y_j$条件下X的条件分布列.

离散随机变量的条件分布

同理, 对一切使$P(X=x_i)=p_{i\cdot}=\sum_{j=1}^{\infty}p_{ij}>0$ 的x_i, 称

$$p_{j|i}=P(Y=y_j|X=x_i)=\frac{P(X=x_i,Y=y_j)}{P(X=x_i)}=\frac{p_{ij}}{p_{i\cdot}},\quad i=1,2,\cdots$$

为给定$X=x_i$条件下Y的条件分布列.

- 注意条件概率和条件分布列的相似性与不同点

离散随机变量的条件分布函数

定义3.5.2 给定$Y=y_j$条件下X的条件分布函数为

$$F(x|y_j)=\sum_{x_i\leqslant x}P(X=x_i|Y=y_j)=\sum_{x_i\leqslant x}p_{i|j}$$

给定$X=x_i$条件下Y的条件分布为

$$F(y|x_i)=\sum_{y_j\leqslant y}P(Y=y_j|X=x_i)=\sum_{y_j\leqslant y_j}p_{j|i}$$

例： 设二维随机变量(X,Y)的联合分布列为

$X \backslash Y$	1	2	3	$p_{i\cdot}$
1	0.1	0.3	0.2	0.6
2	0.2	0.05	0.15	0.4
$p_{\cdot j}$	0.3	0.35	0.35	1

因为$P(X=1)=p_{1\cdot}=0.6$, 所以用第一行各元素除0.6就可得到在$X=1$的条件下, Y的条件分布列

$Y\|X=1$	1	2	3
P	$1/6$	$1/2$	$1/3$

同理可得$Y|X=2$, $X|Y=1$, $X|Y=2$, $X|Y=3$

问： $P(X|Y=3)$?

例: 设X,Y相互独立, 且$X\sim P(\lambda_1),Y\sim P(\lambda_2)$. 证明在$X+Y=n$的条件下, X的条件分布服从二项分布$b(n,p)$, 其中$p=\dfrac{\lambda_1}{\lambda_1+\lambda_2}$

证明: 因为泊松分布可加性(教材例3.3.2), $X+Y\sim P(\lambda_1+\lambda_2)$,

$$P(X=k|X+Y=n)=\frac{P(X=k,X+Y=n)}{P(X+Y=n)}$$

$$=\frac{P(X=k)P(Y=n-k)}{P(X+Y=n)}$$

$$=\frac{\frac{\lambda_1^k}{k!}\mathrm{e}^{-\lambda_1}\cdot\frac{\lambda_2^{n-k}}{(n-k)!}\mathrm{e}^{-\lambda_2}}{\frac{(\lambda_1+\lambda_2)^n}{n!}\mathrm{e}^{-\lambda_1-\lambda_2}}=\frac{n!}{k!(n-k)!}\frac{\lambda_1^k\lambda_2^{n-k}}{(\lambda_1+\lambda_2)^n}$$

$$=\binom{n}{k}\left(\frac{\lambda_1}{\lambda_1+\lambda_2}\right)^k\left(\frac{\lambda_2}{\lambda_1+\lambda_2}\right)^{n-k},\qquad k=0,1,\cdots,n$$

连续随机变量的条件分布

设二维随机变量(X,Y)的联合密度函数为$p(x,y)$, 边际密度函数为$p_X(x)$, $p_Y(y)$

定义3.5.2 对一切使$p_Y(y)>0$ 的y, **给定**$Y=y$**条件下**X**的条件分布函数和条件密度函数分别为**

$$F(x|y)=\int_{-\infty}^{x}\frac{p(u,y)}{p_Y(y)}\mathrm{d}u,\qquad p(x|y)=\frac{p(x,y)}{p_Y(y)}$$

同理, 对一切使$p_X(x)>0$ 的x, **给定**$X=x$**条件下**Y**的条件分布函数和条件密度函数分别为**

$$F(y|x)=\int_{-\infty}^{y}\frac{p(x,v)}{p_X(x)}\mathrm{d}v,\qquad p(y|x)=\frac{p(x,y)}{p_X(x)}$$

例3.5.4

设X,Y服从二维正态分布, $N\left(\mu_1,\mu_2,\sigma_1^2,\sigma_2^2,\rho\right)$, 由边际分布知$X$服从正态分布$N\left(\mu_1,\sigma_1^2\right)$, Y服从正态分布$N\left(\mu_2,\sigma_2^2\right)$. 求条件分布.

$$
\begin{aligned}
p(x|y) &= \frac{p(x,y)}{p_Y(y)} \\
&= \frac{\frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}}\exp\left\{-\frac{1}{2(1-\rho^2)}\left(\frac{(x-\mu_1)^2}{\sigma_1^2}-2\rho\frac{(x-\mu_1)(y-\mu_2)}{\sigma_1\sigma_2}+\frac{(y-\mu_2)^2}{\sigma_2^2}\right)\right\}}{\frac{1}{\sqrt{2\pi}\sigma_2}\exp\left\{-\frac{(y-\mu_2)^2}{2\sigma_2^2}\right\}} \\
&= \frac{1}{\sqrt{2\pi}\sigma_1\sqrt{1-\rho^2}}\exp\left\{-\frac{1}{2\sigma_1^2\left(1-\rho^2\right)}\left[x-\left(\mu_1+\rho\frac{\sigma_1}{\sigma_2}\left(y-\mu_2\right)\right)\right]^2\right\}
\end{aligned}
$$

故$X|Y\sim N(\mu_1+\rho\frac{\sigma_1}{\sigma_2}\left(y-\mu_2\right),\sigma_1^2\left(1-\rho^2\right))$ 服从正态分布: 二维正态变量的边际分布和条件分布都是一维正态分布.

同理$Y|X$也服从正态分布.

例3.5.5

设(X,Y)服从 $G=\{(x,y)|x^2+y^2\leqslant 1\}$的均匀分布, 求给定$Y=y$条件下$X$的条件密度函数$p(x|y)$.

解: 由

$$p(x,y)=\begin{cases}\frac{1}{\pi}, & x^2+y^2\leqslant 1\\ 0, & \text{其他}\end{cases}$$

故Y的边际密度函数为

$$p_Y(y)=\begin{cases}\frac{2}{\pi}\sqrt{1-y^2}, & -1\leq y\leq 1\\ 0, & \text{其他}\end{cases}$$

故当$-1<y<1$时有

$$p(x|y)=\frac{p(x,y)}{p_Y(y)}$$

$$=\begin{cases}\frac{1/\pi}{(2/\pi)\sqrt{1-y^2}}=\frac{1}{2\sqrt{1-y^2}}, & -\sqrt{1-y^2}\leqslant x\leqslant\sqrt{1-y^2}\\ 0, & \text{其他}\end{cases}$$

即$p(x|y)$当$-1<y<1$时为**一维均匀分布**.

(注意与边际密度的区别, 此处y为固定值!)

小结

(1) 条件分布列:

$$p_{i|j} = P(X = x_i | Y = y_j) = \frac{p_{ij}}{p_{\cdot j}}$$

(2) 条件密度函数:

$$p(x|y) = \frac{p(x,y)}{p_Y(y)}$$

(3) 条件分布函数:

$$F(x|y) = \begin{cases} \sum\limits_{x_i \le x} P(X = x_i | Y = y_j) \\ \int_{-\infty}^{x} p(t|y)dt = \int_{-\infty}^{x} \frac{p(t,y)}{p(y)} dt \end{cases}$$

连续场合的全概率公式和贝叶斯公式

由条件概率公式, 有

$$p(x,y)=p_X(x)p(y|x),\quad p(x,y)=p_Y(y)p(x|y)$$

再对$p(x,y)$求边际密度函数, 就得到**全概率公式**的密度函数形式:

$$p_Y(y)=\int_{-\infty}^{\infty}p_X(x)p(y|x)\mathrm{d}x$$
$$p_X(x)=\int_{-\infty}^{\infty}p_Y(y)p(x|y)\mathrm{d}y$$

和**贝叶斯公式**的密度函数形式:

$$p(x|y)=\frac{p(x,y)}{p_Y(y)}=\frac{p_X(x)p(y|x)}{\int_{-\infty}^{\infty}p_X(x)p(y|x)\mathrm{d}x}$$

或

$$p(y|x)=\frac{p_Y(y)p(x|y)}{\int_{-\infty}^{\infty}p_Y(y)p(x|y)\mathrm{d}y}$$

定义 3.5.4 条件分布的数学期望称为**条件期望**, 定义如下

$$E(X|Y=y)=\begin{cases}\sum\limits_i x_i P(X=x_i|Y=y), \text{离散}\\ \int_{-\infty}^{+\infty} xp(x|y)dx, \text{连续}\end{cases}$$

$$E(Y|X=x)=\begin{cases}\sum\limits_j y_j P(Y=y_j|X=x), \text{离散}\\ \int_{-\infty}^{+\infty} yp(y|x)dy, \text{连续}\end{cases}$$

条件期望的特例: (当随机变量褪化为常数时)常数的条件期望为常数本身, 如$E(1|X=x)=1$.

注意点

- $E(X|Y=y)$ 是 y 的函数.
- 所以记 $g(y)=E(X|Y=y)$.
- 进一步记 $g(Y)=E(X|Y)$, 表示随机变量Y的函数, 所以$E(X|Y)$也是随机变量.
- 条件期望是期望, 具有期望的一切性质. 如

$$E\left(a_1X_1+a_2X_2|Y=y\right)=a_1E\left(X_1|Y=y\right)+a_2E\left(X_2|Y=y\right)$$

- **(拓展知识) 条件期望的概念**和随机变量$E(Y|X), E(Y|X_1,...X_n)$**在计量经济学中占有重要地位.**

课堂练习(5’)

设二维连续随机变量(X, Y)的联合密度函数为

$$p(x, y) = \begin{cases} 24(1-x)y, 0 < y < x < 1 \\ 0, otherwise \end{cases}$$

试在$0 < y < 1$时, 求$E(X|Y = y)$

解:

设二维连续随机变量(X,Y)的联合密度函数为

$$p(x,y)=\begin{cases}24(1-x)y, 0<y<x<1\\ 0, otherwise\end{cases}$$

试在$0<y<1$时, 求$E(X|Y=y)$

解： 先求条件密度函数$p(x|y)$. 在$0<y<1$时,

$p_Y(y)=\int_y^1 24(1-x)ydx=12y(1-y)^2$, 所以

$$p(x\mid y)=\frac{p(x,y)}{p_Y(y)}=\begin{cases}\frac{2(1-x)}{(1-y)^2}, & 0<y<x<1,\\ 0, & \text{其他}.\end{cases}$$

由此得, 在$0<y<1$时

$$E(X\mid Y=y)=\int_{-\infty}^{+\infty}x\cdot p(x\mid y)\mathrm{d}x=\int_y^1 x\cdot\frac{2(1-x)}{(1-y)^2}\,\mathrm{d}x=\frac{2y+1}{3}$$

课后习题14

设 $E(Y), E[h(Y)]$ 存在, 试证: $E[h(Y) \mid Y] = h(Y)$.

证明:

设 $E(Y), E[h(Y)]$ 存在, 试证: $E[h(Y) \mid Y] = h(Y)$.

证明: 因为 $E[h(Y) \mid Y]$ 是随机变量Y的函数, 记 $g(Y) = E[h(Y) \mid Y]$, 它仍是随机变量.

在离散场合, 当$Y = y_i$时, $g(y)$ 以概率 $P(Y = y_i)$ 取 $g(y_i) = E[h(y_i) \mid Y = y_i]$.

由于在Y取固定值y_i时, $h(y_i)$也是常数, 故有

$$g(y_i) = E[h(y_i) \mid Y = y_i] = h(y_i) \cdot E(1 \mid Y = y_i) = h(y_i), i = 1, 2, \cdots$$

上式对 Y 的任一取值都成立, 即 $E(h(Y) \mid Y) = h(Y)$.

在连续场合也有类似解释, 所以在一般场合有 $E(h(Y) \mid Y) = h(Y)$.

(本题引申特例可定义出: $E(X|X) = X$)

课后习题15第(1)问

试证明：$E[g(X)Y \mid X] = g(X)E(Y \mid X)$

课后习题15第(1)问

试证明：$E[g(X)Y \mid X] = g(X)E(Y \mid X)$

证明： 由 $E(g(X)Y \mid X = x) = g(x)E(Y \mid X = x)$,

(对于任意固定的$X = x_i$, $g(x_i)$均为常数, 可以提出期望运算.)

可得 $E[g(X)Y \mid X] = g(X)E(Y \mid X)$

课后习题15第(1)问

试证明：$E[g(X)Y \mid X] = g(X)E(Y \mid X)$

证明： 由 $E(g(X)Y \mid X = x) = g(x)E(Y \mid X = x)$,

(对于任意固定的$X = x_i$, $g(x_i)$均为常数, 可以提出期望运算.)

可得 $E[g(X)Y \mid X] = g(X)E(Y \mid X)$

重要结论：如条件期望中有被作为条件的随机变量的函数, 可以提出期望符号.

重期望公式(Law of Iterated Expectation, LIE)

定理 3.5.1 $E(X)=E(E(X|Y))$

注意$E(X|Y)$是Y的函数, 所以外层的期望是关于Y的, 所以

- 在Y离散情况下

$$E(X)=\sum_j E\left(X|Y=y_j\right)P\left(Y=y_j\right)$$

- 在Y连续情况下

$$E(X)=\int_{-\infty}^{\infty}E(X|Y=y)p_Y(y)\mathrm{d}y$$

同理$E(Y)=E(E(Y|X))$

证明离散情况$E(Y)=E(E(Y|X))$(连续见教材)

$$
\begin{aligned}
E\left(E\left(Y \mid X\right)\right) &= \sum_{i=1}^{N_x} E(Y \mid X = x_i) Pr\left(X = x_i\right) \\
&= \sum_{i=1}^{N_x} \left[\sum_{j=1}^{N_y} y_j Pr(Y = y_j \mid X = x_i) \right] Pr\left(X = x_i\right) \\
&= \sum_{j=1}^{N_y} \sum_{i=1}^{N_x} y_j Pr(Y = y_j \mid X = x_i) Pr\left(X = x_i\right) \\
&= \sum_{j=1}^{N_y} \sum_{i=1}^{N_x} y_j Pr(Y = y_j, X = x_i) \\
&= \sum_{j=1}^{N_y} y_j \sum_{i=1}^{N_x} Pr(Y = y_j, X = x_i) = \sum_{j=1}^{N_y} y_j Pr(Y = y_j) = E\left(Y\right)
\end{aligned}
$$

课后习题15第(2)(3)问 (LIE应用)

试证明:

(1) $E[g(X)Y \mid X] = g(X)E(Y \mid X)$ (上文已证明)

(2) $E(XY) = E[XE(Y \mid X)]$;

(3) $\mathrm{Cov}[X, E(Y \mid X)] = \mathrm{Cov}(X, Y)$

证明:

试证明:

(1) $E[g(X)Y \mid X] = g(X)E(Y \mid X)$ (上文已证明)

(2) $E(XY) = E[XE(Y \mid X)]$;

(3) $\mathrm{Cov}[X, E(Y \mid X)] = \mathrm{Cov}(X, Y)$

证明:

(2) 因为 $E(XY) = E[E(XY \mid X)]$, 又由 (1) 知 $E(XY \mid X) = XE(Y \mid X)$, 所以有

$$E(XY) = E[E(XY \mid X)] = E[XE(Y \mid X)]$$

(3)
$\mathrm{Cov}(X, Y) = E(XY) - E(X)E(Y) = E[XE(Y|X)] - E(X)E(Y) = E[XE(Y|X)] - E(X)E(E(Y|X)) = \mathrm{Cov}[X, E(Y \mid X)]$

拓展知识：重期望公式(LIE)应用

试证明： 若u, X 为随机变量，且 $E(u|X) = 0$，则$E(Xu) = 0$，$Cov(X, u) = 0$，$Cov(g(X), u) = 0$

证明：

拓展知识：重期望公式(LIE)应用

试证明： 若u, X 为随机变量，且 $E(u|X) = 0$, 则$E(Xu) = 0$, $Cov(X, u) = 0$, $Cov(g(X), u) = 0$

证明：

1. $E(Xu) = E(E(Xu|X)) = E(XE(u|X)) = 0$

2.
$Cov(X, u) = E(Xu) - E(X)E(u) = E(Xu) - E(X)E(E(u|X)) = 0$

3.
$Cov(g(X), u) = E(g(X)u) - E(g(X))E(u) = E(E(g(X)u|X)) - E(g(X))E(E(u|X)) = E(g(X)E(u|X)) - E(g(X))E(E(u|X)) = 0$

(拓展知识) 计量经济学会学到：上面证明的是回归分析中解释变量与误差项间不同假设的强弱关系对比.

随机个随机变量和的数学期望

例3.5.10 随机个随机变量和的数学期望 设$X_1, X_2, \ldots$为一系列独立同分布的随机变量, 随机变量N只取正整数, N与$\{X_n\}$独立, 证明

$$E\left(\sum_{i=1}^{N} X_i\right) = E\left(X_1\right) E(N)$$

证明:

$$
\begin{aligned}
E\left(\sum_{i=1}^{N} X_i\right) &= E\left[E\left(\sum_{i=1}^{N} X_i|N\right)\right] \\
&= \sum_{n=1}^{\infty} E\left(\sum_{i=1}^{N} X_i|N=n\right) P(N=n)(\text{展开外层关于}N\text{的期望}) \\
&= \sum_{n=1}^{\infty} E\left(\sum_{i=1}^{n} X_i\right) P(N=n)(\text{条件为}N=n\text{即把}N\text{换为}n) \\
&= \sum_{n=1}^{\infty} nE\left(X_i\right) P(N=n) \\
&= E\left(X_1\right)\sum_{n=1}^{\infty} nP(N=n) = E\left(X_1\right)E(N)
\end{aligned}
$$

第13次作业

习题3.5中题目2, 3, 5, 6, 7, 8, 10, 12, 13, 16