第三章：随机向量及其分布

有时候仅仅用一个随机变量去描述某个样本点ω的信息是不够的，此时就需要引入随机向量。

观察下面的随机试验：

1. 研究$4-6$岁年龄段儿童的生长发育情况，考虑每个儿童(ω)的身高$(X_1(\omega))$和体重$(X_2(\omega))$
2. 研究每个家庭的支出情况，考察每个家庭(ω)的衣$(X_1(\omega))$食$(X_2(\omega))$住$(X_3(\omega))$行$(X_4(\omega))$的消费情况

随机向量的概念以及分布函数 I

当一个随机试验的结果是若干并列的随机元素的时候，我们可以用一个数组（向量）来记录随机试验的结果，这就是**随机向量**。

定义3.1.1

设$(\Omega, \mathcal{F}, P)$为概率空间，$X_i, i = 1, 2, \cdots, n$是定义在同一个样本空间Ω上的n个随机变量，则称向量$(X_1, X_2, \cdots, X_n)$为**随机向量**或者**n维随机变量**.

多维随机变量

随机向量$(X_1, X_2, \cdots, X_n)$是 某一个样本空间Ω到n维实空间的一个映射：

$$\Omega \ni \omega \mapsto (X_1(\omega), X_2(\omega), \cdots, X_n(\omega)) \in \mathbb{R}^n$$

随机向量的每个分量都是一个随机变量

联合分布

定义3.1.2

设$(\Omega, \mathcal{F}, P)$为概率空间，$(X_1, X_2, \cdots, X_n)$为其上的随机向量，它的联合分布函数定义为

$$
\begin{aligned}
& F_{X_1, X_2, \cdots, X_n}(x_1, x_2, \cdots, x_n) \\
& = P(\omega \in \Omega : X_1(\omega) \le x_1, X_2(\omega) \le x_2, \cdots, X_n(\omega) \le x_n) \\
& = P(\omega \in \Omega : \bigcap_{i=1}^{n} \{X_i(\omega) \le x_i\})
\end{aligned}
$$

联合分布的性质 I

有界性 对于$\forall(x_1, x_2, \cdots, x_n) \in \mathbb{R}^n$

$$0 \leq F_{X_1,X_2,\cdots,X_n}(x_1, x_2, \cdots, x_n) \leq 1$$

连续型 $F_{X_1,X_2,\cdots,X_n}(x_1, x_2, \cdots, x_n)$关于每个变元$x_i$单调递增并且右连续，$i = 1, 2, \cdots, n$.

边缘特性

$$\lim_{x_i \to -\infty} F_{X_1,X_2,\cdots,X_n}(x_1, x_2, \cdots, x_n) = 0,\ \forall i = 1, 2, \cdots, n$$

$$\lim_{x_1,\cdots,x_n \to +\infty} F_{X_1,X_2,\cdots,X_n}(x_1, x_2, \cdots, x_n) = 1.$$

联合分布的性质 II

事件概率 对二维随机向量，任意$(x_1, x_2) \in \mathbb{R}^n$和$h_i > 0$，$i = 1, 2$有

$$
\begin{aligned}
& P(X_1 \in (x_1, x_1 + h_1], X_2 \in (x_2, x_2 + h_2]) \\
& = F_{X_1,X_2}(x_1 + h_1, x_2 + h_2) - F_{X_1,X_2}(x_1, x_2 + h_2) \\
& - F_{X_1,X_2}(x_1 + h_1, x_2) + F_{X_1,X_2}(x_1, x_2) \geq 0
\end{aligned}
$$

证明 I

1. 概率的非负性。
2. 固定除了x_i以外的所有x_j，

$$
\begin{aligned}
G_{X_i}(x_i) =& F_{X_1,\cdots,X_i\cdots,X_n}(x_1,\cdots,x_i,\cdots,x_n) \\
&= P\left(\omega\in\Omega : \bigcap_{j\neq i}\{X_j(\omega)\le x_j\}\cap\{X_i(\omega)\le x_i\}\right)
\end{aligned}
$$

对$G_{X_i}(x_i)$仿一维随机变量分布函数的证明方法，其单调递增以及右连续性是显然的。

证明 II

3

$$0 \le \lim_{x_i \to -\infty} G_{X_i}(x_i) \le \lim_{x_i \to -\infty} F_{X_i}(x_i) = 0$$

另外

$$\lim_{x_1, x_2, \cdots, x_n \to +\infty} F_{X_1, X_2, \cdots, X_n}(x_1, x_2, \cdots, x_n) = P(\Omega) = 1$$

4 对$n = 2$的情形

$$\begin{aligned}
0 &\le P(x_1 < X_1 \le x_1 + h_1, x_2 < X_2 \le x_2 + h_2) \\
&= (P(X_1 \le x_1 + h_1, X_2 \le x_2 + h_2) - P(X_1 \le x_1, X_2 \le x_2 + h_2)) \\
&- (P(X_1 \le x_1 + h_1, X_2 \le x_2) - P(X_1 \le x_1, X_2 \le x_2)) \\
&= (F_{X_1, X_2}(x_1 + h_1, x_2 + h_2) - F_{X_1, X_2}(x_1, x_2 + h_2)) \\
&- (F_{X_1, X_2}(x_1 + h_1, x_2) - F_{X_1, X_2}(x_1, x_2))
\end{aligned}$$

柯尔莫哥洛夫存在性定理 I

柯尔莫哥洛夫存在性定理

若有定义在$\mathbb{R}^n$上的实函数F满足定理3.1.1的所有性质，那么就可以定义一个概率空间$(\Omega, \mathcal{F}, P)$和其上的随机向量$(X_1, X_2, \cdots, X_n)$，使得

$$F_{X_1, X_2, \cdots, X_n}(x_1, x_2, \cdots, x_n) = F(x_1, x_2, \cdots, x_n),$$

$\forall (x_1, x_2, \cdots, x_n) \in \mathbb{R}^n$

讨论

请说明下面的函数能不能成为某个二维随机向量的分布函数，为什么?

$$F(x,y)=\begin{cases}0, & x+y<0\\ 1, & x+y\geq 0\end{cases}$$

边缘分布函数

定义X_i，$i = 1, \cdots, n$，的**边缘分布函数**为$F_{X_i}(x_i)$，那么

$$\begin{aligned} F_{X_i}(x_i) &= P(X_i \le x_i) \\ &= P\left((X_1 \le +\infty) \cap \cdots \cap (X_i \le x_i) \cap \cdots \cap (X_n \le +\infty)\right) \\ &= \lim_{x_j \to +\infty, j \ne i} F_{X_1, \cdots, X_i, \cdots, X_n}(x_1, \cdots, x_i, \cdots, x_n) \end{aligned}$$

边缘分布反映的是随机向量每个分量**单独**的分布。

二维随机向量的边缘分布函数定义

$$F_{X_1}(x_1) = \lim_{x_2 \to +\infty} F_{X_1,X_2}(x_1, x_2) = F_{X_1,X_2}(x_1, +\infty)$$
$$F_{X_2}(x_2) = \lim_{x_1 \to +\infty} F_{X_1,X_2}(x_1, x_2) = F_{X_1,X_2}(+\infty, x_2)$$

二维离散型随机向量 I

取值为**可列个**的随机向量为**离散型随机向量**

设二维离散型随机向量(X,Y)的取值为(x_i,y_j)，$i,j=1,2,\cdots$，其联合分布列$\{P(X=x_i,Y=y_j)\}$为

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

显然有

1. $p_{ij}\geq 0$
2. $\sum_{i=1}^{+\infty}\sum_{j=1}^{+\infty}p_{ij}=1$

二维离散型随机向量 II

此时，(X,Y)的联合分布函数为

$$F_{X,Y}(x,y) = \sum_{x_i \le x, y_j \le y} p_{ij}$$

且对任意的$a < b$， $c < d$，有

$$P(a < X \le b, c < Y \le d) = \sum_{a < x_i \le b, c < y_j \le d} p_{ij}$$

即(X,Y)的统计特性完全由概率分布$\{p_{ij}, i, j = 1, 2, \cdots\}$确定。

二维离散型随机变量 I

Table : 二维离散随机向量的分布列

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

二维离散型随机变量 II

注意上面的表格中的最右列和最下行,

$$p_{i\cdot} = \sum_j p_{ij} = \sum_j P(X = x_i, Y = y_j)$$

$$= P\left(X = x_i, \bigcup_{j=1}^{+\infty}(Y = y_j)\right) = P(X = x_i), i = 1, 2, \cdots .$$

$$p_{\cdot j} = \sum_i p_{ij} = \sum_i P(X = x_i, Y = y_j)$$

$$= P\left(\bigcup_{i=1}^{+\infty}(X = x_i), Y = y_j\right) = P(Y = y_j), j = 1, 2, \cdots .$$

即$\{p_{i\cdot}, i = 1, 2, \cdots\}$和$\{p_{\cdot j}, j = 1, 2, \cdots\}$分别为$X$和$Y$的边缘分布。

例3.2.1 已知10件产品中有3件一等品，5件二等品，2件三等品。现从中任取4件，求其中一等品件数X与二等品件数Y的联合分布。

解：根据题目条件以及问题，$X=0,1,2,3$，$Y=0,1,2,3,4$，且$2\leq i+j\leq 4$，按照古典概型计算联合分布

$$p_{ij}=P(X=i,Y=j)=\frac{\binom{3}{i}\binom{5}{j}\binom{2}{4-i-j}}{\binom{10}{4}}$$

其中$i=0,1,2,3$，$j=0,1,2,3,4$，$2\leq i+j\leq 4.$

写出X,Y的联合分布表

课堂练习 II

Table：例题3.2.1

Y \ X	0	1	2	3	4	$p_{i\cdot}$
0	0	0	$10/210$	$20/210$	$5/210$	$35/210$
1	0	$15/210$	$60/210$	$30/210$	0	$105/210$
2	$3/210$	$30/210$	$30/210$	0	0	$63/210$
3	$2/210$	$5/210$	0	0	0	$7/210$
$p_{\cdot j}$	$5/210$	$50/210$	$100/210$	$50/210$	$5/210$	1

三项分布

设随机试验只有A, B, C三个结果，各结果出现的概率分别为$p, q, 1-p-q$。现将该随机试验独立的做n次，记X和Y分别为n次试验中A和B发生的次数，试求(X, Y)的联合分布和边缘分布。

解：X和Y的取值只可能是$0, 1, 2, \cdots, n$，并且$X+Y \leq n$，由于试验是独立的，按独立试验概型计算，联合分布

$$P(X=i, Y=j)=\left(\begin{array}{c} n \\ i \end{array}\right)\left(\begin{array}{c} n-i \\ j \end{array}\right) p^{i} q^{j}(1-p-q)^{n-i-j},$$

$0 \leq i+j \leq n$。

记$p_{ij} = P(X = i, Y = j)$，那么，

$$\sum_{i=0}^{n-j} p_{ij} = b(j; n, q), \quad \sum_{j=0}^{n-i} p_{ij} = b(i; n, p)$$

即

称上述(X, Y)服从三项分布，其边缘分布

$$X \sim B(n, p), \quad Y \sim B(n, q)$$

练习

设二维随机向量(X_1, X_2)的分布为$P(X_i = -1) = P(X_i = 1) = 0.25$，$P(X_i = 0) = 0.5$，且满足$P(X_1X_2 = 0) = 1$，求$(X_1, X_2)$的联合分布以及$P(X_1 = X_2)$。

二维离散型随机向量的条件分布列

已知二维随机向量(X,Y)的分布列为

$$P(X=a_i,Y=b_j)=p_{ij},i,j=1,2,\cdots.$$

在事件$\{Y=b_j\}$发生的前提下X的分布列称为**条件分布列**，记为$P(X=a_i\mid Y=b_j)$，$i=1,2,\cdots$。

条件分布列

由条件概率定义有，在事件$\{Y=b_j\}$发生条件下X的条件分布列为

$$P(X=a_i \mid Y=b_j)=\frac{P(X=a_i,Y=b_j)}{P(Y=b_j)}=\frac{p_{ij}}{p_{\cdot j}}, i=1,2,\cdots.$$

类似的，在事件$\{X=a_i\}$发生条件下Y的条件分布列为

$$P(Y=b_j \mid X=a_i)=\frac{P(X=a_i,Y=b_j)}{P(X=a_i)}=\frac{p_{ij}}{p_{i\cdot}}, j=1,2,\cdots.$$

课堂练习

在例3.2.1中，计算$\{Y=1\}$条件下X的条件分布列以及$\{X=2\}$条件下Y的条件分布列

二维连续型随机向量 I

对于二维连续性随机向量(X,Y)，其分布函数为

$$F_{X,Y}(x,y)=\int_{-\infty}^{x}\int_{-\infty}^{y}f_{X,Y}(u,v)\,\mathrm{d}u\,\mathrm{d}v$$

密度函数$f_{X,Y}(x,y)$满足下面几条性质

1. $f_{X,Y}(x,y)\geq 0,(x,y)\in\mathbb{R}^2$
2. $$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}f_{X,Y}(x,y)\,\mathrm{d}x\,\mathrm{d}y=1$$
3. 对二维平面的任何区域D有

$$P((X,Y)\in D)=\int_{D}f_{X,Y}(x,y)\,\mathrm{d}x\,\mathrm{d}y$$

4 X和Y的边缘密度分布函数分别为

$$f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x,y)\,\mathrm{d}y,$$

和

$$f_Y(y) = \int_{-\infty}^{\infty} f_{X,Y}(x,y)\,\mathrm{d}x.$$

随机变量的独立性 I

定义3.1.3

设$(\Omega, \mathcal{F}, P)$为概率空间，$(X_1, X_2, \cdots, X_n)$为其上的随机向量，如果

$$F_{X_1,X_2,\cdots,X_n}(x_1, x_2, \cdots, x_n) = F_{X_1}(x_1)F_{X_2}(x_2)\cdots F_{X_n}(x_n),$$

$\forall (x_1, x_2, \cdots, x_n) \in \mathbb{R}^n$, 则称$X_1, X_2, \cdots, X_n$相互独立。

独立性 I

设$(\Omega, \mathcal{F}, P)$为概率空间，$(X_1, X_2, \cdots, X_n)$为其上的随机向量。

1. 若$(X_1, X_2, \cdots, X_n)$为离散型随机向量，则$X_1, X_2, \cdots, X_n$相互独立的**充分必要条件**是

$$P\left(X_1 = a_1, X_2 = a_2, \cdots, X_n = a_n\right) = P\left(X_1 = a_1\right) P\left(X_2 = a_2\right) \cdots P\left(X_n = a_n\right)$$

2. 若$(X_1, X_2, \cdots, X_n)$为连续型随机向量，则$X_1, X_2, \cdots, X_n$相互独立的**充分必要**条件是

$$f_{X_1, X_2, \cdots, X_n}(x_1, x_2, \cdots, x_n) = f_{X_1}(x_1) f_{X_2}(x_2) \cdots f_{X_n}(x_n)$$

课堂练习 I

1. 设随机变量X,Y独立，且

$$P(X=-1)=P(X=1)=P(Y=-1)=P(Y=1)=1/2$$

求$P(X=Y)$

2. **例题**3.3.1　已知

$$f_{X,Y}(x,y)=\begin{cases} kxy, x^2\le y\le 1, 0\le x\le 1 \\ 0, \text{其他} \end{cases}$$

   1. 确定k的值
   2. 求(X,Y)落在区域$D=\{(x,y)\,|\,x^2\le y\le x, 0\le x\le 1\}$的概率
   3. 求X和Y的边缘分布密度

课堂练习 II

4. 请问X与Y是否相互独立?

3. 从$(0,1)$中任取两个数，求下列事件的概率：(1) 两数之和小于1.2;(2)两数之积小于1/4

4. 若二维随机向量(X,Y)的联合密度函数为

$$f(x,y)=\begin{cases}8xy, & 0\leq x\leq y\leq 1\\ 0, & \text{其他}\end{cases}$$

请问X与Y是否独立

服从特定分布的二维随机向量

本章我们将学习以下几种特殊的二维连续型随机向量

1. 二维均匀分布
2. 二维正态分布

二维均匀分布

二维均匀分布

设D为二维平面上的一个有界区域，面积为S_D，若随机向量(X, Y)的分布密度函数为

$$f_{X,Y}(x,y) = \begin{cases} \dfrac{1}{S_D}, (x,y) \in D \\ 0, \quad \text{其他} \end{cases}$$

则称(X, Y)服从D上的**均匀分布**。

例3.3.3

在某一分钟内的任何时刻，信号进入收音机是等可能的，若收到两个**独立**信号的时间间隔小于0.5秒，则信号将相互干扰。试求一分钟内两个信号相互干扰的概率。

解：设两信号进入收音机的时刻分别为X和Y，
则$X \sim U[0,60]$，$Y \sim U[0,60]$，由于信号进入收音机的时刻**相互独立**，即X和Y相互独立，知X,Y的联合分布函数为

$$f_{X,Y}(x,y) = f_X(x)f_Y(y) = \begin{cases} \dfrac{1}{3600}, (x,y) \in [0,60] \times [0,60] \\ 0, \quad \text{其他} \end{cases}$$

那么

$$P(|X-Y|<0.5) = \int_{|X-Y|<0.5} \frac{1}{3600}\,\mathrm{d}x\,\mathrm{d}y$$

二维正态分布

若二维随机向量(X,Y)的分布密度函数为

$$f_{X,Y}(x,y) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \exp\left\{-\frac{1}{2(1-\rho^2)}\left(\frac{(x-\mu_1)^2}{\sigma_1^2} - 2\rho\frac{(x-\mu_1)(y-\mu_2)}{\sigma_1\sigma_2} + \frac{(y-\mu_2)^2}{\sigma_2^2}\right)\right\}$$

则称(X,Y)服从参数为$\mu_1,\mu_2,\sigma_1^2,\sigma_2^2,\rho$的正态分布，记为$(X,Y)\sim N(\mu_1,\mu_2,\sigma_1^2,\sigma_2^2,\rho)$.

函数图像

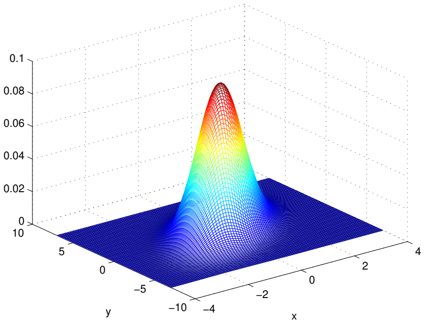

Figure : $\mu_1 = \mu_2 = 0,\ \sigma_1 = 1,\ \sigma_2 = 2,\ \rho = \frac{1}{2}$

边缘分布

二维正态分布的边缘分布为一维正态分布，即

$$f_X(x) = \frac{1}{\sqrt{2\pi}\sigma_1} e^{-\frac{(x-\mu_1)^2}{2\sigma_1^2}} \quad f_Y(y) = \frac{1}{\sqrt{2\pi}\sigma_2} e^{-\frac{(y-\mu_2)^2}{2\sigma_2^2}}$$

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

$\rho = 0$是X, Y是相互独立随机变量的充要条件。

n维正态分布（非退化情形）

设$\mu = (\mu_1, \mu_2, \cdots, \mu_n)'$，$\mathbf{\Sigma}$为$n$阶正定矩阵，记$\boldsymbol{X} = (x_1, x_2, \cdots, x_n)'$，若

$$f_{X_1,X_2,\cdots,X_n}(x_1, x_2, \cdots, x_n) = \frac{1}{(2\pi)^{\frac{n}{2}} |\mathbf{\Sigma}|^{\frac{1}{2}}} \exp\left\{ -\frac{1}{2}(\boldsymbol{X} - \boldsymbol{\mu})'\mathbf{\Sigma}^{-1}(\boldsymbol{X} - \boldsymbol{\mu}) \right\}$$

则称$(X_1, X_2, \cdots, X_n)'$服从n**维联合正态分布**，记作$\boldsymbol{X} \sim N(\boldsymbol{\mu}, \mathbf{\Sigma})$。

二维联合正态分布

对于$n=2$的情形

$$\boldsymbol{\mu} = \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix} \quad \boldsymbol{\Sigma} = \begin{pmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{pmatrix}$$

二维连续型随机向量的条件密度函数

若$f_Y(y)>0$，那么在$\{Y=y\}$发生的条件下，X的**条件密度函数**定义为

$$f_{X|Y}(x|y)=\frac{f_{X,Y}(x,y)}{f_Y(y)},\forall x\in\mathbb{R}$$

同样的，当$f_X(x)>0$，在$\{X=x\}$发生的条件下Y的**条件密度函数**定义为

$$f_{Y|X}(y|x)=\frac{f_{X,Y}(x,y)}{f_X(x)},\forall y\in\mathbb{R}$$

设(X,Y)的联合密度函数为

$$f_{X,Y}(x,y)=\begin{cases} kxy, x^2\leq y\leq 1, 0\leq x\leq 1 \\ 0, \text{其他} \end{cases}$$

求$f_{X|Y}(x|y=\frac{1}{2})$ 和 $f_{Y|X}(y|x=\frac{1}{3})$

大家可以自行验证，对于二维正态分布，X在$\{Y=y\}$下和Y在$\{X=x\}$下的条件分布分别为

$$N(\mu_1+\rho\frac{\sigma_1}{\sigma_2}(y-\mu_2),\sigma_1^2(1-\rho^2))$$

和

$$N(\mu_2+\rho\frac{\sigma_2}{\sigma_1}(x-\mu_1),\sigma_2^2(1-\rho^2))$$

二维随机向量函数的分布 I

问题

根据(X,Y)的联合分布，求$g(X,Y)=X+Y$的分布

1. 对于二维离散型随机向量的情形，设(X,Y)的分布列为

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

那么$Z=X+Y$的分布列为

$$P(Z=z_k)=\sum_{x_i+y_j=z_k} p_{ij}, k=1,2,\cdots$$

二维随机向量函数的分布 II

特别的，若$x_i = i$，$y_j = j$，那么

$$P(Z=k) = \sum_{i=0}^{k} P(X=i, Y=k-i) = \sum_{i=0}^{k} p_{i(k-i)}, k = 0, 1, 2, \cdots$$

进一步的，若X与Y独立，那么

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

二维随机向量函数的分布 III

2 对于二维连续型随机向量，设(X,Y)的联合分布密度函数为$f_{X,Y}$，那么$Z=X+Y$的分布函数为

$$\begin{aligned} F_Z(z) &= P(Z\le z) = P(X+Y\le z) \\ &= \int_{x+y\le z} f_{X,Y}(x,y)\,\mathrm{d}x\,\mathrm{d}y = \int_{-\infty}^{\infty}\mathrm{d}x\int_{-\infty}^{z-x} f_{X,Y}(x,y)\,\mathrm{d}y \\ &= \int_{-\infty}^{\infty}\mathrm{d}y\int_{-\infty}^{z-y} f_{X,Y}(x,y)\,\mathrm{d}x \end{aligned}$$

按照连续型随机变量函数的分布密度函数求法，对分布函数$F_Z(z)$关于z求导，得到Z的分布密度函数为

$$f_Z(z) = \int_{-\infty}^{\infty} f_{X,Y}(x,z-x)\,\mathrm{d}x = \int_{-\infty}^{\infty} f_{X,Y}(z-y,y)\,\mathrm{d}y$$

二维随机向量函数的分布 IV

特别的，若X, Y相互独立，那么

$$f_Z(z) = \int_{-\infty}^{\infty} f_X(x) f_Y(z-x)\,\mathrm{d}x = \int_{-\infty}^{\infty} f_X(z-y) f_Y(y)\,\mathrm{d}y$$

即两个独立的随机变量和的分布密度函数为它们各自的分布密度函数的“卷积”。

练习

设X, Y的联合密度函数为

$$f_{X,Y}(x,y) = \begin{cases} e^{-(x+y)}, & x > 0, y > 0 \\ 0, & \text{其他} \end{cases}$$

求下面的随机变量的密度函数：

$$Z = (X+Y)/2, \quad W = X - Y$$

多维正态分布的函数 I

1. (X, Y)是二维随机向量，X与Y独立，$X \sim N(\mu_1, \sigma_1^2)$，$Y \sim (\mu_2, \sigma_2^2)$，那么$Z = X + Y \sim N(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2)$
2. $(X_1, X_2, \cdots, X_n)$是n维正态分布，且$X_1, X_2, \cdots, X_n$相互独立，$X_i \sim N(\mu_i, \sigma_i^2), i = 1, \cdots, n$，那么$\sum_{i=1}^n a_i X_i \sim N(\sum_{i=1}^n a_i \mu_i, \sum_{i=1}^n a_i^2 \sigma_i^2)$
3. **命题3.4.1** 随机向量$\boldsymbol{X} = (X_1, X_2, \cdots, X_n)'$服从$n$维正态分布的充要条件是对任意的$\boldsymbol{k} = (k_1, k_2, \cdots, k_n)' \in \mathbb{R}^n$，那么$\boldsymbol{k}'\boldsymbol{X}$服从一维正态分布
4. **命题3.4.2** 设随机向量$\boldsymbol{X} = (X_1, X_2, \cdots, X_n)'$服从$n$维正态分布，期望向量为$\boldsymbol{\mu} = (\mu_1, \mu_2, \cdots, \mu_n)'$，协方差矩阵为$\Sigma$，那么对于任意的实矩阵$\boldsymbol{A}_{m \times n}$，有$\boldsymbol{AX} \sim N(\boldsymbol{A\mu}, \boldsymbol{A}\Sigma\boldsymbol{A}')$