Outline

二维连续型随机向量

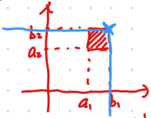
对于二维连续性随机向量 (X, Y)，其分布函数为

$P((-\infty,x)\times(-\infty,y))$

$\parallel$

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

$(-\infty,b_1)\times(-\infty,b_2) - (-\infty,a_1)\times(-\infty,b_2)$
$-(-\infty,b_1)\times(-\infty,a_2) + (-\infty,a_1)\times(-\infty,a_2)$

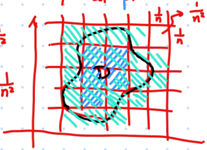


密度函数 $f_{X,Y}(x,y)$ 满足下面几条性质

- $f_{X,Y}(x,y) \geq 0, (x,y) \in \mathbb{R}^2$
- $$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{X,Y}(x,y)\,dxdy = 1$$

$\text{Area}_{\top} := |\{\text{block} \subsetneq D\}| \times \frac{1}{n^2}$

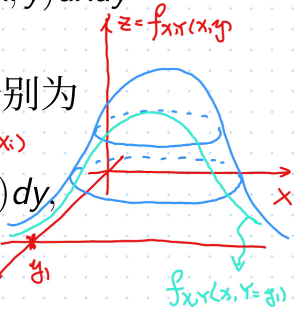
$\text{Area}_{\bot} := |\{\text{block} \cap D \neq \phi\}| \times \frac{1}{n^2}$



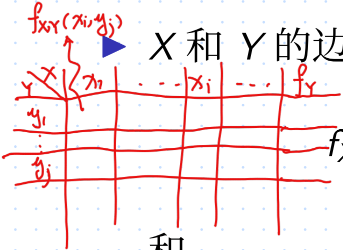
二维连续型随机向量

- 对二维平面的任何区域 D 有

$$P((X,Y)\in D)=\int_D f_{X,Y}(x,y)\,dxdy$$



- X 和 Y 的边缘密度分布函数分别为



$f_Y(y)=P(Y=y)=\sum_{i=1}^{+\infty}P(Y=y,X=x_i)$

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

和

$=P(X\in(-\infty,+\infty),Y=y_1)$

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

随机向量的独立性

事件 $\Leftrightarrow$ 子集 $\Leftrightarrow$ 映射 $A: \Omega \to \{0,1\}$

A $\quad$ A

定义 3.1.3

事件的独立

$P(A|B) = P(A) \Leftrightarrow P(A\cdot B) = P(A)\cdot P(B|A) = P(A)\cdot P(B)$

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

$P((-\infty, x_1)\times(-\infty,+\infty)) \times P((-\infty,+\infty),(-\infty,x_2))$

$\|$

$F_{X_1X_2}(x_1,x_2) = P((-\infty,x_1)\times(-\infty,x_2)) = P_{X_1}((-\infty,x_1)) \times P_{X_2}((-\infty,x_2))$

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

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

比如考察一个人的个人条件，那么他的身高和体重并不是独立的，但是他的收入水平和身高体重就是独立的。

随机向量的独立性

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。

- 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

$$
\begin{aligned}
& P(X_1 = a_1, X_2 = a_2, \cdots, X_n = a_n) \\
& = P(X_1 = a_1)\, P(X_2 = a_2) \cdots P(X_n = a_n)
\end{aligned}
$$

随机向量的独立性

- 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

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

课堂练习

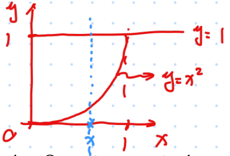
- 设随机变量 X, Y 独立，且

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

求 $P(X=Y)$ $=\frac{1}{2}$

$Y \backslash X$	-1	1	P_Y
-1	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{2}$
1	$\frac{1}{4}$	$\frac{1}{4}$	

例 3.3.1



已知

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

- 确定 k 的值

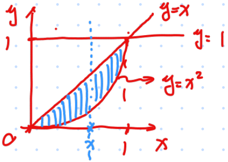
$$\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} f_{X,Y}(x,y)\,dx\,dy = 1$$

$$= \int_0^1 \boxed{\int_{x^2}^1 kxy\,dy}\,dx$$

$$= \int_0^1 g(x)\,dx$$

例 3.3.1

已知



$$f_{X,Y}(x,y) = \begin{cases} 6xy, & x^2 \leq y \leq 1, 0 \leq x \leq 1 \\ 0, & \text{其他} \end{cases}$$

- 求 (X, Y) 落在区域 $D = \{(x,y) | x^2 \leq y \leq x, 0 \leq x \leq 1\}$ 的概率

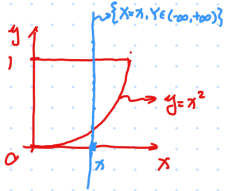
$$\int_0^1 \int_{x^2}^{x} 6xy \, dy \, dx = 6 \int_0^1 x \underbrace{\int_{x^2}^{x} y \, dy}_{\left(\frac{1}{2}y^2\right)\Big|_{x^2}^{x}} dx$$

例 3.3.1

已知

$$f_{X,Y}(x,y) = \begin{cases} 6xy, & x^2 \le y \le 1, 0 \le x \le 1 \\ 0, & \text{其他} \end{cases}$$

- 求 X 和 Y 的边缘分布密度



$$f_X(x) := \int_{-\infty}^{+\infty} f_{X,Y}(x,y)\,dy$$

$$= \begin{cases} 0 & x \notin [0,1] \\ \int_{x^2}^{1} 6xy \; dy & x \in [0,1] \end{cases}$$

例 3.3.1

已知

$$f_{X,Y}(x,y) = \begin{cases} 6xy, & x^2 \le y \le 1, 0 \le x \le 1 \\ 0, & \text{其他} \end{cases}$$

- 请问 X 与 Y 是否相互独立?

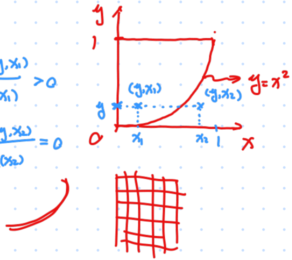
$$f_{Y|X}(y|x) = f_Y(y)$$

$$\| $$

$$\frac{f_{Y,X}(y,x)}{f_X(x)}$$

$$f_{Y|X}(y,x_1) = \frac{f_{Y,X}(y,x_1)}{f_X(x_1)} > 0$$

$$f_{Y|X}(y,x_2) = \frac{f_{Y,X}(y,x_2)}{f_X(x_2)} = 0$$



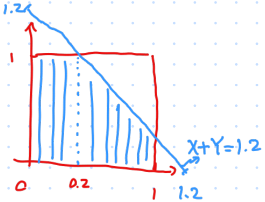
课堂练习

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

取随机变量 $X, Y \sim U(0,1)$

i.e. 随机向量 (X, Y)



1. 记 $Z = X+Y$

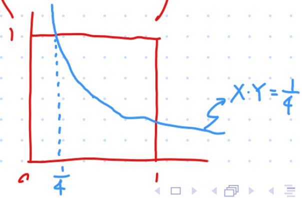
$$P(Z<1.2) = P(X+Y<1.2)$$

$$= P\begin{pmatrix} 0<X\leq 0.2 \\ 0<Y\leq 1 \end{pmatrix} + P\begin{pmatrix} 0.2<X<1 \\ 0<Y<1.2-X \end{pmatrix}$$

2. 记 $Z' = X\cdot Y$

$$P(Z'<\tfrac{1}{4}) = P(X\cdot Y<\tfrac{1}{4})$$

$$= P\begin{pmatrix} 0<X\leq \frac{1}{4} \\ 0<Y<1 \end{pmatrix} + P\begin{pmatrix} \frac{1}{4}<X<1 \\ 0<Y<\frac{1}{4X} \end{pmatrix}$$



课堂练习

若二维随机向量 (X, Y) 的联合密度函数为

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

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

- 二维均匀分布
- 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

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

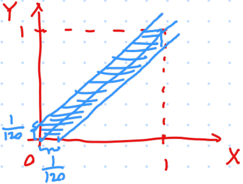
则称 (X, Y) 服从 D 上的 均匀分布 。

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。

$X, Y \sim [0,1]$

$P(|X-Y| < \frac{1}{120})$



例 3.3.3

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

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

那么

$$P(X - Y < 0.5) = \int_{X-Y<0.5} \frac{1}{3600} dxdy$$

二维正态分布

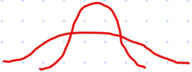
$X \sim N(\mu, \sigma)$ 位置 形状　$\frac{1}{2\pi\sigma} e^{-\frac{1}{2}\left(\frac{x-\mu^2}{\sigma^2}\right)}$

若二维随机向量 (X, Y) 的分布密度函数为

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

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



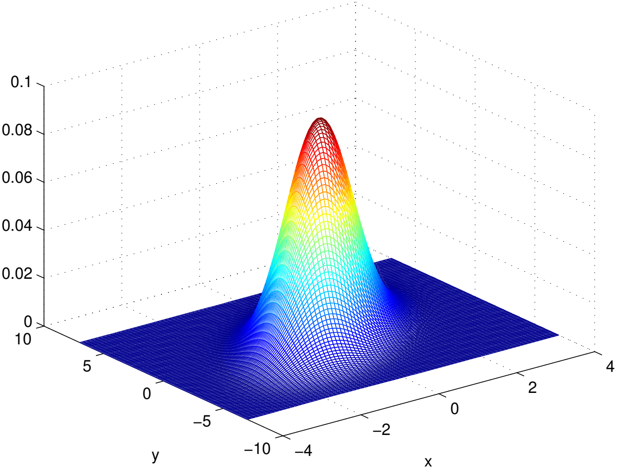


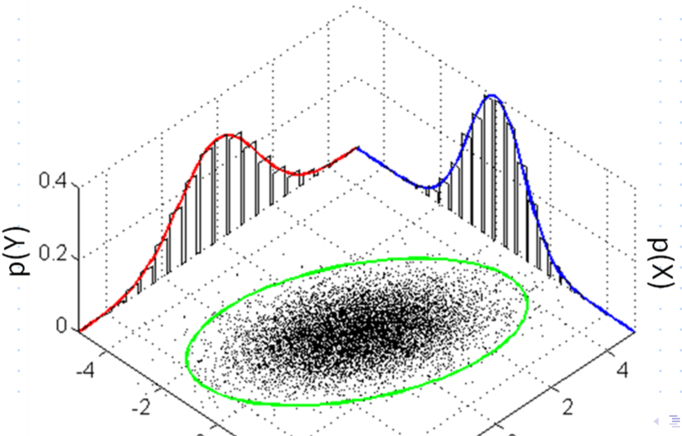
图: $\mu_1 = \mu_2 = 0,\ \sigma_1 = 1,\ \sigma_2 = 2,\ \rho = \frac{1}{2}$

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



边缘分布

$\rho = 0$ 是 X, Y 是相互独立随机变量的充要条件。

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

$$f_{X|Y}(x|y) = f(x)$$

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 $= (x_1, x_2, \cdots, x_n)'$，若

退化情形

$X \sim N(0,1)$

$(X,Y)\quad f_{X,Y}(x,y) = \begin{cases} f_X(x) & y=0 \\ 0 & y\neq 0 \end{cases}$

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

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

n 维正态分布（非退化情形）

请大家验证，对于 $n = 2$ 的情形

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

二维连续型随机变量的条件密度函数

设 (X, Y) 为二维连续型的随机向量，其联合分布密度函数为 $f_{X,Y}(x, y)$，边缘分布密度函数分别为 $f_X(x)$ 和 $f_Y(y)$，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

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

例

设 (X, Y) 的联合密度函数为

$$f_{X,Y}(x,y) = \begin{cases} 6xy, & x^2 \le y \le 1, 0 \le x \le 1 \\ 0, & 其他 \end{cases}$$

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

例

大家可以自行验证，对于二维正态分布

$$(X, Y) \sim \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \times e^{-\frac{1}{2(1-\rho^2)}((\frac{(x-\mu_1)^2}{\sigma_1^2}) - 2\rho\frac{(x-\mu_1)(y-\mu_2)}{\sigma_1\sigma_2} + \frac{(y-\mu_2)^2}{\sigma_2^2})},$$

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

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

Outline

简单情形

问题：根据 (X, Y) 的联合分布，求 $g(X, Y) = X + Y$ 的分布

离散型的简单情形

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

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

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

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

连续型的简单情形

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

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

连续型的简单情形

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

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

特别的，若 X, Y 相互独立，那么

$$f_Z(z) = \int_{-\infty}^{\infty} f_X(x)f_Y(z-x)dx = \int_{-\infty}^{\infty} f_X(z-y)f_Y(y)dy$$

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

- (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 $Z = X + Y \sim N(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2)$
- $(X_1, X_2, \cdots, X_n)$ 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 $\sum_{i=1}^n a_i X_i \sim N(\sum_{i=1}^n a_i \mu_i, \sum_{i=1}^n a_i^2 \sigma_i^2)$

命题 3.4.1

随机向量 $\boldsymbol{X} = (X_1, X_2, \cdots, X_n)^T$ 服从 n 维正态分布的充要条件是对任意的 $\boldsymbol{k} = (k_1, k_2, \cdots, k_n)^T \in \mathbb{R}^n$，那么 $\boldsymbol{k}^T\boldsymbol{X}$ 服从一维正态分布

命题 3.4.2

设随机向量 $= (X_1, X_2, \cdots, X_n)'$ 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 Σ，那么对于任意的实矩阵 $\boldsymbol{A}_{m\times n}$，有

$\boldsymbol{A} \sim N(\boldsymbol{A}\boldsymbol{\mu}, \boldsymbol{A}\Sigma\boldsymbol{A}')$

练习

相互独立的随机变量 $X \sim \exp(\lambda_1)$，$Y \sim \exp(\lambda_2)$，求 $Z = X/Y$ 的密度函数。

Outline