信息学中的概率统计

王若松

前沿计算研究中心
北京大学

多维连续随机变量

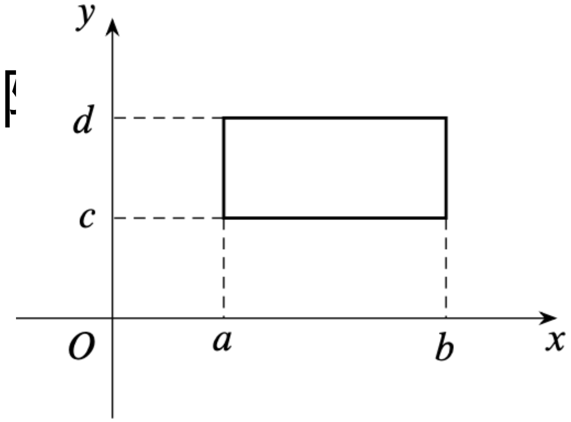
1. 多维连续随机变量的分布函数和密度函数
2. 多维连续随机变量的独立性
3. 多维连续随机变量的特征数
4. 多维连续随机变量函数的分布
5. 多维高斯分布

1. 多维连续随机变量的分布函数和密度函数

- 给定二维随机变量X,Y和实数x,y，定义$F(x,y)=P(X\le x,Y\le y)$为二维随机变量X,Y的**联合分布函数**

- 性质1(**有界性**): $0\le F(x,y)\le 1, F(-\infty,y)=0, F(x,-\infty)=0, F(+\infty,+\infty)=1$
- 性质2(**单调性**): $x_1<x_2\Rightarrow F(x_1,y)\le F(x_2,y)$，$y_1<y_2\Rightarrow F(x,y_1)\le F(x,y_2)$
- 性质3(**右连续**): $F(x+0,y)=F(x,y), F(x,y+0)=F(x,y)$
- 性质4(**非负性**): $P(a<X\le b, c<Y\le d)=F(b,d)-F(a,d)-F(b,c)+F(a,c)\ge 0$

- $F(x,y)$满足上述四条性质**等价于**$F(x,y)$是某个二维



- 性质4是否被性质1-3蕴含？
  - $F(x,y)=1$ 当$x+y\ge 0$，否则$F(x,y)=0$

1. 多维连续随机变量的分布函数和密度函数

- 给定二维随机变量X,Y，若存在$f(x,y)\geq 0$使得分布函数 $F(x,y)$可表示为 $F(x,y)=\int_{-\infty}^{x}\int_{-\infty}^{y}f(u,v)dudv$，则称二维随机变量$X,Y$为**二维连续随机变量**，称$f(x,y)$为$X,Y$的**联合密度函数**

- 联合密度函数的性质
- 性质1(**非负性**): $f(x,y)\geq 0$
- 性质2(**正则性**): $\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}f(u,v)dudv=1$

- 在$F(x,y)$ 偏导数存在的点，$\frac{\partial^2 F(x,y)}{\partial x\partial y}=f(x,y)$
- 对于区域G，$P\big((X,Y)\in G\big)=\iint_G f(x,y)\,dxdy$

1. 多维连续随机变量的分布函数和密度函数

- 例1：X,Y的联合密度函数满足
  - $f(x,y) = c \cdot e^{-2x-3y}$ 若 $x > 0, y > 0$
  - 否则$f(x,y) = 0$
- 这里c是某个常数
- 计算常数c，并计算$P(X < 1, Y > 1), P(X > Y)$

- 由正则性$\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} f(x,y)dxdy = c \cdot \frac{1}{2} \cdot \frac{1}{3} = 1 \Rightarrow c = 6$
- $P(X < 1, Y > 1) = \int_0^1 \int_1^{+\infty} 6 \cdot e^{-2x-3y} dxdy = \left(1 - e^{-2}\right)e^{-3}$
- $P(X > Y) = \int_0^{+\infty} \int_0^x 6 \cdot e^{-2x-3y} dydx = \int_0^{+\infty} 6 \cdot e^{-2x} \cdot \frac{1}{3} \cdot \left(1 - e^{-3x}\right)dx = \frac{4}{5}$

1. 多维连续随机变量的分布函数和密度函数

- 例2：给定$\mathbb{R}^2$中的一个有界区域 D。 随机变量(X,Y)表示从 D中均匀取一点的坐标。写出X,Y的联合密度函数。
  - $f(x,y) = 1/S_D$若 $(x,y) \in D$
  - 否则$f(x,y) = 0$

- 称(X,Y)服从D上的**二维均匀分布**，记为$(X,Y) \sim U(D)$

PEKING UNIVERSITY

1. 多维连续随机变量的分布函数和密度函数

- 给定二维随机变量X,Y的**联合分布函数**$F(x,y)$，如何计算X的分布函数?
  - $P(X \leq x) = P(X \leq x, Y < +\infty) = F(x, +\infty)$
- 定义$F_X(x) = F(x,+\infty) = \lim\limits_{y\to+\infty} F(x,y)$为 X的**边际分布函数**
- 类似有 $F_Y(y) = F(+\infty, y) = \lim\limits_{x\to+\infty} F(x,y)$为 Y的**边际分布函数**
- 例： X,Y的联合分布函数满足
  - $F(x,y) = 1 - e^{-x} - e^{-y} + e^{-x-y-\lambda xy}$ 当 $x > 0, y > 0$,
  - 否则$F(x,y) = 0$
  - 参数$0 \leq \lambda \leq 1$
- 计算$F_X(x)$和 $F_Y(y)$
  - $F_X(x) = 1 - e^{-x}$ 当 $x > 0$
  - $F_Y(y) = 1 - e^{-y}$ 当 $y > 0$

1. 多维连续随机变量的分布函数和密度函数

- 给定二维连续随机变量X,Y，$f(x,y)$为X,Y的联合密度函数
- 如何计算X的密度函数$f_X(x)$?
  - $P(X \leq x) = F(x,+\infty) = \int_{-\infty}^{x} \int_{-\infty}^{+\infty} f(u,v)\, dvdu = \int_{-\infty}^{x} f_X(u)du$
- $f_X(x) = \int_{-\infty}^{+\infty} f(x,y)dy$为$X$的**边际密度函数**
- 类似$f_Y(y) = \int_{-\infty}^{+\infty} f(x,y)dx$为$Y$的**边际密度函数**
- 例1： X,Y的联合密度函数满足
  - $f(x,y) = 6 \cdot e^{-2x-3y}$ 若 $x > 0, y > 0$
  - 否则$f(x,y) = 0$
- 计算X,Y的边际密度函数
  - $f_X(x) = \int_{-\infty}^{+\infty} f(x,y)dy = 2 \cdot e^{-2x}$
  - $f_Y(y) = \int_{-\infty}^{+\infty} f(x,y)dx = 3 \cdot e^{-3y}$

1. 多维连续随机变量的分布函数和密度函数

- $f_X(x) = \int_{-\infty}^{+\infty} f(x,y)dy$为$X$的边际密度函数
- $f_Y(y) = \int_{-\infty}^{+\infty} f(x,y)dx$为$Y$的边际密度函数

- 例2： X, Y的联合密度函数满足
  - $f(x,y) = 1$ 若 $0 < x < 1, |y| < x$
  - 否则$f(x,y) = 0$
- 计算Y的边际密度函数
  - 当$|y| > 1$， $f_Y(y) = 0$
  - 当$-1 < y < 0$， $-y < x < 1$，$f_Y(y) = \int_{-\infty}^{+\infty} f(x,y)dx = \int_{-y}^{1} f(x,y)dx = 1 + y$
  - 当$0 < y < 1$， $y < x < 1$，$f_Y(y) = \int_{-\infty}^{+\infty} f(x,y)dx = \int_{y}^{1} f(x,y)dx = 1 - y$

北京大学 PEKING UNIVERSITY

1. 多维连续随机变量的分布函数和密度函数

- 回顾：若$P\left(Y=y_j\right)>0$，则称$p_{i|j}=P(X=x_i\mid Y=y_j)=P(X=x_i,Y=y_j)/P(Y=y_j)$为给定 $Y=y_j$条件下 X的**条件分布列**

- 对于连续随机变量，如何定义条件分布函数和条件密度函数?
- $P(X\le x|Y=y)$可定义为 $\lim\limits_{\Delta\to 0}P(X\le x|y\le Y\le y+\Delta)$
- $P(X\le x|Y=y)=\lim\limits_{\Delta\to 0}P(X\le x|y\le Y\le y+\Delta)=\lim\limits_{\Delta\to 0}\frac{P(X\le x,y\le Y\le y+\Delta)}{P(y\le Y\le y+\Delta)}$
- $\lim\limits_{\Delta\to 0}\frac{P(X\le x,y\le Y\le y+\Delta)}{P(y\le Y\le y+\Delta)}=\lim\limits_{\Delta\to 0}\frac{\int_{-\infty}^{x}\left(\frac{1}{\Delta}\int_{y}^{y+\Delta}f(u,v)dv\right)du}{\frac{1}{\Delta}\int_{y}^{y+\Delta}f_Y(v)dv}$

1. 多维连续随机变量的分布函数和密度函数

- 当密度函数连续
  - $\lim_{\Delta\to 0}\frac{1}{\Delta}\int_y^{y+\Delta} f(u,v)dv = f(u,y)$
  - $\lim_{\Delta\to 0}\frac{1}{\Delta}\int_y^{y+\Delta} f_Y(v)dv = f_Y(y)$
- $P(X\le x|Y=y) = \lim_{\Delta\to 0}\frac{\int_{-\infty}^{x}\left(\frac{1}{\Delta}\int_y^{y+\Delta} f(u,v)dv\right)du}{\frac{1}{\Delta}\int_y^{y+\Delta} f_Y(v)dv} = \frac{\int_{-\infty}^{x} f(u,y)du}{f_Y(y)} = \int_{-\infty}^{x}\frac{f(u,y)}{f_Y(y)}du$
- $F(x|y) = P(X\le x|Y=y) = \int_{-\infty}^{x}\frac{f(u,y)}{f_Y(y)}du$ 为给定$Y=y$条件下 X的**条件分布函数**
- $f(x|y) = \frac{f(x,y)}{f_Y(y)}$为给定$Y=y$条件下 X的**条件密度函数**
  - 正则性?
  - $\int_{-\infty}^{+\infty}\frac{f(x,y)}{f_{Y(y)}}dx = \frac{\int_{-\infty}^{+\infty} f(x,y)dx}{f_Y(y)} = \frac{f_Y(y)}{f_Y(y)} = 1$

PEKING UNIVERSITY 北京大学

1. 多维连续随机变量的分布函数和密度函数

- 例： 随机变量X,Y服从单位圆$(x^2+y^2\le 1)$上的二维均匀分布。计算$f(x|y)$
- $f(x|y)=\frac{f(x,y)}{f_Y(y)}$
- $f(x,y)=\frac{1}{\pi}$若(x,y)在单位圆内
- 否则$f(x,y)=0$
- $f_Y(y)=\int_{-\infty}^{+\infty}f(x,y)dx=\int_{-\sqrt{1-y^2}}^{+\sqrt{1-y^2}}f(x,y)dx=\frac{2\sqrt{1-y^2}}{\pi}$
- $f(x|y)=\frac{1}{\pi}/\frac{2\sqrt{1-y^2}}{\pi}=\frac{1}{2\sqrt{1-y^2}}$若$|x|\le\sqrt{1-y^2}$
- 否则$f(x|y)=0$

2. 多维连续随机变量的独立性

- 回顾：给定二维离散随机变量(X,Y)，若对于任意实数x,y均有 $P(X=x,Y=y)=P(X=x)\cdot P(Y=y)$，则称$X,Y$**相互独立**

- 给定二维随机变量(X,Y)，分布函数为$F(x,y)$， 边际分布函数为$F_X(x)$和$F_Y(y)$。若对于任意实数x,y均有 $F(x,y)=F_X(x)\cdot F_Y(y)$，则称$X,Y$**相互独立**

- 对于离散随机变量， 相互独立**等价于**对于任意实数x,y均有 $P(X=x,Y=y)=P(X=x)\cdot P(Y=y)$
- 对于连续随机变量， 相互独立**等价于**对于任意实数x,y均有密度函数 $f(x,y)=f_X(x)\cdot f_Y(y)$

2. 多维连续随机变量的独立性

- 例1：X,Y的联合密度函数满足
  - $f(x,y) = 6 \cdot e^{-2x-3y}$ 若 $x > 0, y > 0$
  - 否则$f(x,y) = 0$
- 判断 X,Y是否相互独立
  - $f_X(x) = 2 \cdot e^{-2x}$
  - $f_Y(y) = 3 \cdot e^{-3y}$
  - $f(x,y) = f_X(x) \cdot f_Y(y)$

2. 多维连续随机变量的独立性

- 例2：令随机变量 X某服务器第一次发生故障的时间， Y表示另一台服务器第一次发生故障的时间。已知则$X \sim \mathrm{Exp}(\lambda_1)$， $Y \sim \mathrm{Exp}(\lambda_2)$，且$X$与$Y$相互独立。
- 计算$P(X<Y)$

- $P(X<Y)=\int_{x=0}^{+\infty}\int_{y=x}^{+\infty} f(x,y)dxdy=\int_{x=0}^{+\infty}\int_{y=x}^{+\infty} f_X(x)\cdot f_Y(y)dxdy$
- $f_X(x)\cdot f_Y(y)=\lambda_1 e^{-\lambda_1 x}\cdot\lambda_2 e^{-\lambda_2 y}$
- $P(X<Y)=\int_0^{+\infty}\lambda_1 e^{-\lambda_1 x}\cdot\int_x^{+\infty}\lambda_2 e^{-\lambda_2 y}dydx=\int_0^{+\infty}\lambda_1 e^{-\lambda_1 x}\cdot P(Y\geq x)\,dx$
- $P(Y\geq x)=e^{-\lambda_2 x}$
- $P(X<Y)=\int_{x=0}^{+\infty}\lambda_1 e^{-\lambda_1 x}\cdot e^{-\lambda_2 x}dx=\frac{\lambda_1}{\lambda_1+\lambda_2}$

2. 多维连续随机变量的独立性

- 二维连续随机变量X,Y的联合密度函数满足
- $f(x,y)=\frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}}\exp\left[-\frac{1}{2(1-\rho^2)}\cdot\left(\frac{(x-\mu_1)^2}{\sigma_1^2}+\frac{(y-\mu_2)^2}{\sigma_2^2}-\frac{2\rho(x-\mu_1)(y-\mu_2)}{\sigma_1\sigma_2}\right)\right]$
- 其中 $\mu_1,\mu_2\in\mathbb{R},\sigma_1,\sigma_2>0,|\rho|<1$
- 称 X,Y服从参数为$\mu_1,\mu_2,\sigma_1,\sigma_2,\rho$的**二维正态（高斯）分布**
- 记号： $X,Y\sim N(\mu_1,\mu_2,\sigma_1^2,\sigma_2^2,\rho)$

- 验证正则性：$\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}f(x,y)dxdy=1$
- 计算边际密度函数
- 计算条件密度函数
- 判断 X,Y是否相互独立

2. 多维连续随机变量的独立性

- $f(x,y) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}}\exp\left[-\frac{1}{2(1-\rho^2)}\cdot\left(\frac{(x-\mu_1)^2}{\sigma_1^2}+\frac{(y-\mu_2)^2}{\sigma_2^2}-\frac{2\rho(x-\mu_1)(y-\mu_2)}{\sigma_1\sigma_2}\right)\right]$
- 计算$\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} f(x,y)dxdy$?
- 换元法：定义u',v'，使得$\frac{(x-\mu_1)^2}{\sigma_1^2}+\frac{(y-\mu_2)^2}{\sigma_2^2}-\frac{2\rho(x-\mu_1)(y-\mu_2)}{\sigma_1\sigma_2}=(u')^2+(v')^2$
  - 配方得：$u'=\frac{x-\mu_1}{\sigma_1}-\rho\cdot\frac{y-\mu_2}{\sigma_2}\Rightarrow v'=\frac{y-\mu_2}{\sigma_2}\cdot\sqrt{1-\rho^2}$
- 若$u=\frac{\left(\frac{x-\mu_1}{\sigma_1}-\rho\cdot\frac{y-\mu_2}{\sigma_2}\right)}{\sqrt{1-\rho^2}}, v=\frac{y-\mu_2}{\sigma_2}$，则有$f(x,y)=\frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}}\exp\left[-\frac{1}{2}(u^2+v^2)\right]$
- $\left|\frac{\partial(u,v)}{\partial(x,y)}\right| = \begin{vmatrix} \frac{1}{\sigma_1\sqrt{1-\rho^2}} & -\frac{\rho}{\sigma_2\sqrt{1-\rho^2}} \\ 0 & \frac{1}{\sigma_2} \end{vmatrix} = \frac{1}{\sigma_1\sigma_2\sqrt{1-\rho^2}} \Rightarrow \left|\frac{\partial(x,y)}{\partial(u,v)}\right| = \sigma_1\sigma_2\sqrt{1-\rho^2}$

2. 多维连续随机变量的独立性

- $f(x,y)=\frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}}\exp\left[-\frac{1}{2}\left(u^2+v^2\right)\right]$
- $\left|\frac{\partial(x,y)}{\partial(u,v)}\right|=\sigma_1\sigma_2\sqrt{1-\rho^2}$
- $\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}f(x,y)dxdy=\iint\frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}}\exp\left[-\frac{1}{2}\left(u^2+v^2\right)\right]\cdot\sigma_1\sigma_2\sqrt{1-\rho^2}dudv$
- $=\iint\frac{1}{2\pi}\exp\left[-\frac{1}{2}\left(u^2+v^2\right)\right]dudv$
- $=\int_{-\infty}^{-\infty}\frac{1}{\sqrt{2\pi}}e^{-\frac{u^2}{2}}du\cdot\int_{-\infty}^{-\infty}\frac{1}{\sqrt{2\pi}}e^{-\frac{v^2}{2}}dv=1$
- 思考：如何从随机变量的角度理解换元 $u=\frac{\left(\frac{x-\mu_1}{\sigma_1}-\rho\cdot\frac{y-\mu_2}{\sigma_2}\right)}{\sqrt{1-\rho^2}},v=\frac{y-\mu_2}{\sigma_2}$?

2. 多维连续随机变量的独立性

- 边际密度函数$f_X(x) = \int_{-\infty}^{+\infty} f(x,y)dy$
- $f(x,y) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \exp\left[-\frac{1}{2(1-\rho^2)} \cdot \left(\frac{(x-\mu_1)^2}{\sigma_1^2} + \frac{(y-\mu_2)^2}{\sigma_2^2} - \frac{2\rho(x-\mu_1)(y-\mu_2)}{\sigma_1\sigma_2}\right)\right]$
- 换元$u = \frac{\left(\frac{y-\mu_2}{\sigma_2} - \rho \cdot \frac{x-\mu_1}{\sigma_1}\right)}{\sqrt{1-\rho^2}}, v = \frac{x-\mu_1}{\sigma_1}$
- $f(x,y) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \exp\left(-\frac{u^2}{2}\right) \cdot \exp\left(-\frac{v^2}{2}\right)$
- $f_X(x) = \int_{-\infty}^{+\infty} f(x,y)dy = \sqrt{1-\rho^2} \cdot \sigma_2 \cdot \int_{-\infty}^{+\infty} f(x,y)du$
- $f_X(x) = \frac{1}{2\pi\sigma_1} \int_{-\infty}^{+\infty} \exp\left(-\frac{u^2}{2}\right) du \cdot \exp\left(-\frac{v^2}{2}\right) = \frac{1}{\sqrt{2\pi} \cdot \sigma_1} \exp\left(-\frac{(x-\mu_1)^2}{2\sigma_1^2}\right)$
- 类似有 $f_Y(y) = \frac{1}{\sqrt{2\pi} \cdot \sigma_2} \exp\left(-\frac{(y-\mu_2)^2}{2\sigma_2^2}\right)$

PEKING UNIVERSITY

2. 多维连续随机变量的独立性

- 若$X,Y \sim N(\mu_1,\mu_2,\sigma_1^2,\sigma_2^2,\rho)$，则$X \sim N(\mu_1,\sigma_1^2)$， $Y \sim N(\mu_2,\sigma_2^2)$
  - 边际分布与参数 ρ无关 $\Rightarrow$ 具有相同边际分布的多维联合分布可以不同
- 计算条件密度函数 $f(x|y) = \frac{f(x,y)}{f_{Y(y)}}$
  - $f(x,y) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}}\exp\left[-\frac{1}{2(1-\rho^2)}\cdot\left(\frac{(x-\mu_1)^2}{\sigma_1^2}+\frac{(y-\mu_2)^2}{\sigma_2^2}-\frac{2\rho(x-\mu_1)(y-\mu_2)}{\sigma_1\sigma_2}\right)\right]$
  - $f_Y(y) = \frac{1}{\sqrt{2\pi}\cdot\sigma_2}\exp\left(-\frac{(y-\mu_2)^2}{2\sigma_2^2}\right)$
  - $f(x|y) = \frac{1}{\sqrt{2\pi}\sigma_1\sqrt{1-\rho^2}}\exp\left(-\frac{1}{2(1-\rho^2)}\cdot\left(\frac{x-\mu_1}{\sigma_1}-\rho\cdot\frac{y-\mu_2}{\sigma_2}\right)^2\right)$
  - $= \frac{1}{\sqrt{2\pi}\sigma_1\sqrt{1-\rho^2}}\exp\left(-\frac{1}{2(1-\rho^2)}\cdot\left(\frac{x-\mu_1-\rho\cdot\frac{\sigma_1}{\sigma_2}\cdot(y-\mu_2)}{\sigma_1}\right)^2\right)$
- 给定$Y = y$条件下， X的条件分布服从$N(\mu_1+\rho\cdot\frac{\sigma_1}{\sigma_2}\cdot(y-\mu_2),\sigma_1^2(1-\rho^2))$

2. 多维连续随机变量的独立性

- 给定$Y=y$条件下，X的条件分布服从$N(\mu_1+\rho\cdot\frac{\sigma_1}{\sigma_2}\cdot(y-\mu_2),\sigma_1^2(1-\rho^2))$
- 给定$X=x$条件下，Y的条件分布服从$N(\mu_2+\rho\cdot\frac{\sigma_2}{\sigma_1}\cdot(x-\mu_1),\sigma_2^2(1-\rho^2))$

- 何时有 X,Y相互独立?
  - $f(x|y)=\frac{f(x,y)}{f_Y(y)}$
  - 若X,Y相互独立，$f(x,y)=f_X(x)\cdot f_Y(y)$，也即$f(x|y)=f_X(x)$
  - $X\sim N(\mu_1,\sigma_1^2)$，$Y\sim N(\mu_2,\sigma_2^2)$
  - X,Y相互独立等价于$\rho=0$

PEKING UNIVERSITY

3. 多维连续随机变量的特征数

- 回顾： 给定离散随机变量X,Y和函数g， $Z = g(X,Y)$。
- 定理： $E(Z) = \sum_i \sum_j P\left(X = x_i, Y = y_j\right) \cdot g(x_i, y_j)$

- 给定连续随机变量X,Y和函数g， $Z = g(X,Y)$。如何计算$E(Z)$?
- 定理： $E(Z) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f(x,y) \cdot g(x,y)\, dxdy$

- 例： $X \sim U(0,1), Y \sim U(0,1)$，且$X$与$Y$相互独立。求$E(|X-Y|)$
  - $E(|X-Y|) = \int_0^1 \int_0^1 |x-y| dxdy = \frac{1}{3}$

3. 多维连续随机变量的特征数

- **数学期望的线性性：** $E(X+Y)=E(X)+E(Y)$
  - $\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} f(x,y)\cdot(x+y)dxdy$
  - $=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} f(x,y)\cdot x\, dxdy+\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} f(x,y)\cdot y\cdot dxdy$
  - $=\int_{-\infty}^{+\infty} x\cdot\int_{-\infty}^{+\infty} f(x,y)\, dydx+\int_{-\infty}^{+\infty} y\cdot\int_{-\infty}^{+\infty} f(x,y)\cdot dxdy$
  - $=\int_{-\infty}^{+\infty} x\cdot f_X(x)dx+\int_{-\infty}^{+\infty} y\cdot f_Y(y)dy=E(X)+E(Y)$

- 推广： $E(X_1+X_2+\cdots+X_n)=E(X_1)+E(X_2)+\cdots+E(X_n)$

3. 多维连续随机变量的特征数

- 定理：若连续随机变量X和Y相互独立，则有$E(XY)=E(X)\cdot E(Y)$
- $\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} f(x,y)\cdot x\cdot y\, dxdy$
- $=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} f_X(x)\cdot x\cdot f_Y(y)\cdot y\, dxdy$
- $=\int_{-\infty}^{+\infty} x\cdot f_X(x)dx\cdot\int_{-\infty}^{+\infty} y\cdot f_Y(y)dx$
- $=E(X)\cdot E(Y)$

- 推广：若连续随机变量$X_1,X_2,\ldots,X_n$相互独立，则有$E(X_1X_2\cdots X_n)=E(X_1)\cdot E(X_2)\cdots E(X_n)$
- 推论：若连续随机变量$X_1,X_2,\ldots,X_n$相互独立，则有$\mathrm{Var}(X_1\pm X_2\pm\cdots\pm X_n)=\mathrm{Var}(X_1)+\mathrm{Var}(X_2)+\cdots+\mathrm{Var}(X_n)$

3. 多维连续随机变量的特征数

- 例1: 随机向量$X = (X_1, X_2, \ldots, X_n)$满足$X_i \sim N(0,1)$
- 随机变量Y表示X的模长。计算 $E(Y^2)$

- $E\left(Y^2\right) = E\left(\sum_{i=1}^{n} X_i^2\right) = \sum_{i=1}^{n} E(X_i^2) = n$

PEKING UNIVERSITY

3. 多维连续随机变量的特征数

- 例2: 随机向量$X = (X_1, X_2, \dots, X_n)$满足$X_i \sim N(0,1)$，且$X_i$相互独立
- 给定固定向量$a = (a_1, a_2, \dots, a_n)$。令随机变量$Y$表示$X$与$a$的内积。计算$E(Y)$和$\mathrm{Var}(Y)$

- $Y = \sum_{i=1}^{n} a_i X_i$
- $E(Y) = E(\sum_{i=1}^{n} a_i X_i) = \sum_{i=1}^{n} a_i E(X_i) = 0$
- $\mathrm{Var}(Y) = \sum_{i=1}^{n} \mathrm{Var}(a_i X_i) = \sum_{i=1}^{n} a_i^2$

3. 多维连续随机变量的特征数

- 例3: $n \times n$矩阵 A每个元素均服从 $N(0,1)$，且不同元素相互独立
- 计算 $E(\det(A))$， $E(\mathrm{trace}(A))$， $E(\mathrm{trace}(A^2))$
- $\det(A) = \sum_\sigma \left(\mathrm{sgn}(\sigma) \cdot \prod_{i=1}^n A_{i,\sigma(i)}\right)$
- $E\left(\det(A)\right) = \sum_\sigma \mathrm{sgn}(\sigma) \cdot E\left(\prod_{i=1}^n A_{i,\sigma(i)}\right) = \sum_\sigma \mathrm{sgn}(\sigma) \cdot \prod_{i=1}^n E\left(A_{i,\sigma(i)}\right) = 0$

- $\mathrm{trace}(A) = \sum_{i=1}^n A_{i,i}$
- $E\left(\mathrm{trace}(A)\right) = E\left(\sum_{i=1}^n A_{i,i}\right) = 0$

- $A_{i,i}^2 = \sum_{j=1}^n A_{i,j} \cdot A_{j,i}$
- $E\left(A_{i,i}^2\right) = \sum_{j=1}^n E(A_{i,j} \cdot A_{j,i}) = 1$
- $E\left(\mathrm{trace}\left(A^2\right)\right) = E\left(\sum_{i=1}^n A_{i,i}^2\right) = n$

3. 多维连续随机变量的特征数

- 回顾：
  - $F(x|y) = \int_{-\infty}^{x} \frac{f(u,y)}{f_Y(y)} du$ 为给定$Y = y$条件下 X的**条件分布函数**
  - $f(x|y) = \frac{f(x,y)}{f_Y(y)}$为给定$Y = y$条件下 X的**条件密度函数**
  - 对于离散随机变量，$E(X|Y = y_j) = \sum_i x_i \cdot P(X = x_i \mid Y = y_j)$
- 对于二维连续随机变量X, Y，**定义条件数学期望** $E(X|Y = y) = \int_{-\infty}^{+\infty} f(x|y) \cdot x \cdot dx$
- 回顾：$E\left(E(X|Y)\right) = E(X)$

3. 多维连续随机变量的特征数

- **条件数学期望** $E(X|Y=y)=\int_{-\infty}^{+\infty} f(x|y)\cdot x\cdot dx$
- $E\big(E(X|Y)\big)=E(X)$

- 例：$X\sim \mathrm{Exp}(\lambda)$， $Y\sim U(0,X)$。计算$E(Y)$和 $\mathrm{Var}(Y)$

- $E(Y|X=x)=\frac{x}{2}$
- $E(Y)=E\big(E(Y|X)\big)=\frac{E(X)}{2}=\frac{1}{2\lambda}$
- $E\left(Y^2|X=x\right)=\frac{x^2}{3}$
- $E\left(Y^2\right)=E\left(E\left(Y^2|X\right)\right)=\frac{E\left(X^2\right)}{3}=\frac{2}{3\lambda^2}$

3. 多维连续随机变量的特征数

- 给定随机变量X和Y，定义X和Y的**协方差**
- $\mathrm{Cov}(X,Y)=E\Big(\big(X-E(X)\big)\big(Y-E(Y)\big)\Big)=E(XY)-E(X)E(Y)$
- 性质回顾：
- $\mathrm{Cov}(X,X)=\mathrm{Var}(X)=E(X^2)-\big(E(X)\big)^2$
- $\mathrm{Cov}(X,Y)=\mathrm{Cov}(Y,X)$
- $\mathrm{Cov}(aX,bY)=ab\cdot\mathrm{Cov}(X,Y)$
- $\mathrm{Cov}(X_1+X_2,Y)=\mathrm{Cov}(X_1,Y)+\mathrm{Cov}(X_2,Y)$
- 若X和Y相互独立，则$\mathrm{Cov}(X,Y)=0$
- $\mathrm{Var}(X_1+X_2+\cdots+X_n)=\sum_i\sum_j\mathrm{Cov}(X_i,X_j)=\sum_i\mathrm{Var}(X_i)+2\sum_i\sum_{j<i}\mathrm{Cov}(X_i,X_j)$
- $\mathrm{Var}(X+Y)=\mathrm{Var}(X)+\mathrm{Var}(Y)+2\mathrm{Cov}(X,Y)$

3. 多维连续随机变量的特征数

- 例： X,Y的联合密度函数满足
  - $f(x,y)=\frac{x+y}{3}$若$0<x<1,0<y<2$
  - 否则$f(x,y)=0$
- 计算$\mathrm{Cov}(X,Y),\mathrm{Var}(X),\mathrm{Var}(Y)$
- $E(X)=\int_0^1 x\cdot\int_0^2 f(x,y)\cdot dy\,dx=\int_0^1\frac{2x+2}{3}\cdot x\cdot dx=\frac{5}{9}$
- $E\left(X^2\right)=\int_0^1 x^2\cdot\int_0^2 f(x,y)\cdot dy\,dx=\int_0^1\frac{2x+2}{3}\cdot x^2\cdot dx=\frac{7}{8}$
- $E(Y)=\int_0^2 y\cdot\int_0^1 f(x,y)\cdot dx\,dy=\int_0^2\frac{2y+1}{6}\cdot y\cdot dx=\frac{11}{9}$
- $E\left(Y^2\right)=\int_0^2 y^2\cdot\int_0^1 f(x,y)\cdot dx\,dy=\int_0^2\frac{2y+1}{6}\cdot y^2\cdot dx=\frac{16}{9}$
- $E(XY)=\int_0^1 x\cdot\int_0^2 y\cdot f(x,y)\cdot dy\,dx=\int_0^1 x\cdot\frac{2x+8/3}{3}\cdot dx=\frac{2}{3}$
- $\mathrm{Var}(X)=E\left(X^2\right)-\left(E(X)\right)^2=\frac{13}{162},\mathrm{Var}(Y)=E\left(Y^2\right)-\left(E(Y)\right)^2=\frac{23}{81}$
- $\mathrm{Cov}(X,Y)=E(XY)-E(X)E(Y)=-\frac{1}{81}$

3. 多维连续随机变量的特征数

- $X,Y \sim N(\mu_1,\mu_2,\sigma_1^2,\sigma_2^2,\rho)$，计算$\text{Cov}(X,Y)$
- $E(X)=\mu_1, E(Y)=\mu_2$
- $f(x,y)=\frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}}\exp\left[-\frac{1}{2(1-\rho^2)}\cdot\left(\frac{(x-\mu_1)^2}{\sigma_1^2}+\frac{(y-\mu_2)^2}{\sigma_2^2}-\frac{2\rho(x-\mu_1)(y-\mu_2)}{\sigma_1\sigma_2}\right)\right]$
- $\text{Cov}(X,Y)=\iint f(x,y)\cdot(x-\mu_1)\cdot(y-\mu_2)dxdy$
- $u=\frac{\left(\frac{x-\mu_1}{\sigma_1}-\rho\cdot\frac{y-\mu_2}{\sigma_2}\right)}{\sqrt{1-\rho^2}}, v=\frac{y-\mu_2}{\sigma_2}$
- $y=\sigma_2 v+\mu_2, x=\left(\sqrt{1-\rho^2}\cdot u+\rho\cdot v\right)\cdot\sigma_1+\mu_1$
- $(x-\mu_1)(y-\mu_2)=\sigma_1\sigma_2 v\left(\sqrt{1-\rho^2}\cdot u+\rho\cdot v\right)$
- $f(x,y)=\frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}}\exp\left[-\frac{1}{2}\left(u^2+v^2\right)\right], \left|\frac{\partial(x,y)}{\partial(u,v)}\right|=\sigma_1\sigma_2\sqrt{1-\rho^2}$

3. 多维连续随机变量的特征数

- $(x-\mu_1)(y-\mu_2)=\sigma_1\sigma_2 v\left(\sqrt{1-\rho^2}\cdot u+\rho\cdot v\right)$
- $f(x,y)=\frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}}\exp\left[-\frac{1}{2}\left(u^2+v^2\right)\right], \left|\frac{\partial(x,y)}{\partial(u,v)}\right|=\sigma_1\sigma_2\sqrt{1-\rho^2}$
- $\mathrm{Cov}(X,Y)=\iint f(x,y)\cdot(x-\mu_1)\cdot(y-\mu_2)dxdy$
- $=\iint \frac{1}{2\pi}\exp\left[-\frac{1}{2}\left(u^2+v^2\right)\right]\cdot(x-\mu_1)\cdot(y-\mu_2)dudv$
- $=\iint \frac{1}{2\pi}\exp\left[-\frac{1}{2}\left(u^2+v^2\right)\right]\cdot\sigma_1\sigma_2 v\left(\sqrt{1-\rho^2}\cdot u+\rho\cdot v\right)\cdot dudv$
- $=\sigma_1\sigma_2\iint \frac{1}{2\pi}\exp\left[-\frac{1}{2}\left(u^2+v^2\right)\right]\cdot v\left(\sqrt{1-\rho^2}\cdot u+\rho\cdot v\right)\cdot dudv$
- $\iint \frac{1}{2\pi}\exp\left[-\frac{1}{2}\left(u^2+v^2\right)\right]\cdot vududv=0$
- $\iint \frac{1}{2\pi}\exp\left[-\frac{1}{2}\left(u^2+v^2\right)\right]\cdot v^2dudv=1$
- $\mathrm{Cov}(X,Y)=\rho\sigma_1\sigma_2$

3. 多维连续随机变量的特征数

- 给定随机变量X和Y，若$\sigma(X),\sigma(Y)>0$，定义X和Y的**相关系数** $\mathrm{Corr}(X,Y)=\frac{\mathrm{Cov}(X,Y)}{\sigma(X)\sigma(Y)}$
- 回顾：$\mathrm{Cov}(X,Y)=E\Big(\big(X-E(X)\big)\big(Y-E(Y)\big)\Big)$
- 回顾： $\tilde{X}=\frac{X-E(X)}{\sigma(X)}$为$X$的标准化随机变量
- $\mathrm{Corr}(\tilde{X},\tilde{Y})=E(\tilde{X}\tilde{Y})=E\left(\frac{X-E(X)}{\sigma(X)}\cdot\frac{Y-E(Y)}{\sigma(Y)}\right)=\frac{\mathrm{Cov}(X,Y)}{\sigma(X)\sigma(Y)}=\mathrm{Corr}(X,Y)$
  - $\mathrm{Corr}(X,Y)>0$（或$\mathrm{Cov}(X,Y)>0$）： X和Y**正相关**
  - $\mathrm{Corr}(X,Y)<0$（或$\mathrm{Cov}(X,Y)<0$）： X和Y**负相关**
  - $\mathrm{Corr}(X,Y)=0$（或$\mathrm{Cov}(X,Y)=0$）： X和Y**不相关**
- 若相互独立，一定有不相关
  - $E\Big(\big(X-E(X)\big)\big(Y-E(Y)\big)\Big)=E(XY)-E(X)E(Y)=0$
- 不相关，是否一定有相互独立?

3. 多维连续随机变量的特征数

- 给定随机变量X和Y，若$\sigma(X),\sigma(Y)>0$，定义X和Y的**相关系数** $\mathrm{Corr}(X,Y)=\dfrac{\mathrm{Cov}(X,Y)}{\sigma(X)\sigma(Y)}$

- 性质1: $|\mathrm{Corr}(X,Y)|\leq 1$

- 证明:
  - $g(t)=E\left(\left(t(X-E(X))+(Y-E(Y))\right)^2\right)=t^2\sigma(X)^2+2t\mathrm{Cov}(X,Y)+\sigma(Y)^2$
  - $g(t)\geq 0,\sigma(X),\sigma(Y)>0\Rightarrow\left(2\mathrm{Cov}(X,Y)\right)^2-4\sigma(X)^2\sigma(Y)^2\leq 0$

PEKING UNIVERSITY

3. 多维连续随机变量的特征数

- 给定随机变量X和Y，若$\sigma(X),\sigma(Y)>0$，定义X和Y的**相关系数** $\operatorname{Corr}(X,Y)=\frac{\operatorname{Cov}(X,Y)}{\sigma(X)\sigma(Y)}$
- 回顾： $\tilde{X}=\frac{X-E(X)}{\sigma(X)}$为$X$的标准化随机变量
- $\operatorname{Corr}(\tilde{X},\tilde{Y})=E(\tilde{X}\tilde{Y})=E\left(\frac{X-E(X)}{\sigma(X)}\cdot\frac{Y-E(Y)}{\sigma(Y)}\right)=\frac{\operatorname{Cov}(X,Y)}{\sigma(X)\sigma(Y)}=\operatorname{Corr}(X,Y)$
- 性质2: $\operatorname{Corr}(X,Y)=\pm1\Leftrightarrow$存在$a\neq0$与$b$，$P(Y=aX+b)=1$
- 证明（$\operatorname{Corr}(X,Y)=1$）
- $\Rightarrow$： $\operatorname{Var}(\tilde{X}-\tilde{Y})=\operatorname{Var}(\tilde{X})+\operatorname{Var}(\tilde{Y})-2\operatorname{Corr}(X,Y)=0$
- $P(\tilde{X}-\tilde{Y}=c)=1$
- $\Leftarrow$：$P(Y=aX+b)=1$， $\sigma(Y)=|a|\sigma(X)$， $\operatorname{Cov}(X,Y)=a\big(\sigma(X)\big)^2$
- $\operatorname{Corr}(X,Y)=\frac{a}{|a|}$

3. 多维连续随机变量的特征数

- 例1：X,Y的联合密度函数满足
  - $f(x,y)=\frac{x+y}{3}$ 若 $0<x<1, 0<y<2$
  - 否则$f(x,y)=0$
- 计算$\mathrm{Corr}(X,Y)$并判断相关性

- $\mathrm{Var}(X)=E\left(X^2\right)-\left(E(X)\right)^2=\frac{13}{162}, \mathrm{Var}(Y)=E\left(Y^2\right)-\left(E(Y)\right)^2=\frac{23}{81}$
- $\mathrm{Cov}(X,Y)=E(XY)-E(X)E(Y)=-\frac{1}{81}$
- $\mathrm{Corr}(X,Y)=\frac{\mathrm{Cov}(X,Y)}{\sigma(X)\sigma(Y)}=-\frac{\frac{1}{81}}{\sqrt{\frac{13}{162}}\cdot\sqrt{\frac{23}{81}}}=-\sqrt{\frac{2}{299}}$，负相关

3. 多维连续随机变量的特征数

- 例2： $X,Y\sim N(\mu_1,\mu_2,\sigma_1^2,\sigma_2^2,\rho)$证明 X,Y相互独立当且仅当 X,Y 不相关

- X,Y相互独立等价于 $\rho=0$
- X,Y不相关等价于$\rho=0$

4. 多维连续随机变量函数的分布

- 给定连续随机变量X,Y和函数$g(x,y)$，求$Z=g(X,Y)$的概率密度函数
- **卷积公式**：若X,Y相互独立，$Z=X+Y$，则 $f_Z(z)=\int_{-\infty}^{+\infty} f_X(z-y)f_Y(y)dy$
- 证明：
- $P(Z\le z)=\iint_{x+y\le z} f_X(x)f_Y(y)dxdy=\int_{-\infty}^{+\infty}\int_{-\infty}^{z-y} f_X(x)dx\cdot f_Y(y)dy$
- $\int_{-\infty}^{z-y} f_X(x)dx=P(X\le z-y)$
- $P(Z\le z)=\int_{-\infty}^{+\infty} P(X\le z-y)\cdot f_Y(y)dy$，两边对$z$求导

4. 多维连续随机变量函数的分布

- 例1: $X \sim N\left(0,\sigma_1^2\right), Y \sim N(0,\sigma_2^2)$，$X,Y$相互独立，求$Z = X + Y$的概率密度函数
- $f_Z(z) = \int_{-\infty}^{+\infty} f_X(z-y)f_Y(y)dy = \frac{1}{2\pi\sigma_1\sigma_2}\int_{-\infty}^{+\infty} \exp\left(-\frac{1}{2}\left(\frac{(z-y)^2}{\sigma_1^2}+\frac{y^2}{\sigma_2^2}\right)\right)dy$
- $\frac{(z-y)^2}{\sigma_1^2}+\frac{y^2}{\sigma_2^2} = y^2\left(\frac{1}{\sigma_1^2}+\frac{1}{\sigma_2^2}\right)-\frac{2yz}{\sigma_1^2}+\frac{z^2}{\sigma_1^2}$
- 令 $A = \frac{1}{\sigma_1^2}+\frac{1}{\sigma_2^2}$
- $\frac{(z-y)^2}{\sigma_1^2}+\frac{y^2}{\sigma_2^2} = A\left(y-\frac{z}{\sigma_1^2 A}\right)^2+\frac{z^2}{\sigma_1^2}-\frac{z^2}{\sigma_1^4 A} = A\left(y-\frac{z}{\sigma_1^2 A}\right)^2+\frac{z^2}{\sigma_1^2}\left(1-\frac{1}{\sigma_1^2 A}\right)$
- $\frac{z^2}{\sigma_1^2}\left(1-\frac{1}{\sigma_1^2 A}\right) = \frac{z^2}{\sigma_1^2+\sigma_2^2}$
- $f_Z(z) = \frac{1}{2\pi\sigma_1\sigma_2}\exp\left(-\frac{1}{2}\cdot\frac{z^2}{\sigma_1^2+\sigma_2^2}\right)\cdot\int_{-\infty}^{+\infty}\exp\left(-\frac{A}{2}\left(y-\frac{z}{\sigma_1^2 A}\right)^2\right)dy$

4. 多维连续随机变量函数的分布

- 例1: $X \sim N(0,\sigma_1), Y \sim N(0,\sigma_2)$，$X, Y$相互独立，求$Z = X + Y$的概率密度函数
- $f_Z(z) = \frac{1}{2\pi\sigma_1\sigma_2} \exp\left(-\frac{1}{2} \cdot \frac{z^2}{\sigma_1^2+\sigma_2^2}\right) \cdot \int_{-\infty}^{+\infty} \exp\left(-\frac{A}{2}\left(y - \frac{z}{\sigma_1^2 A}\right)^2\right) dy$
- $\frac{1}{\sqrt{2\pi}\cdot\sqrt{1/A}} \int_{-\infty}^{+\infty} \exp\left(-\frac{A}{2}\left(y - \frac{z}{\sigma_1^2 A}\right)^2\right) dy = 1$
- $f_Z(z) = \frac{1}{2\pi\sigma_1\sigma_2} \exp\left(-\frac{1}{2} \cdot \frac{z^2}{\sigma_1^2+\sigma_2^2}\right) \cdot \sqrt{2\pi} \cdot \sqrt{1/A}$
- $A = \frac{1}{\sigma_1^2} + \frac{1}{\sigma_2^2} \Rightarrow f_Z(z) = \frac{1}{\sqrt{2\pi}\cdot\sqrt{\sigma_1^2+\sigma_2^2}} \exp\left(-\frac{1}{2} \cdot \frac{z^2}{\sigma_1^2+\sigma_2^2}\right)$
- 也即 $Z \sim N(0, \sigma_1^2 + \sigma_2^2)$

4. 多维连续随机变量函数的分布

- 推广： $X_i \sim N(\mu_i, \sigma_i)$，且相互独立，则$\sum_{i=1}^{n} a_i X_i \sim N\left(\mu_0, \sigma_0^2\right)$
- $\mu_0 = \sum_{i=1}^{n} a_i \mu_i$ ， $\sigma_0^2 = \sum_{i=1}^{n} a_i^2 \sigma_i^2$

- 特别有，若X_i独立同分布，且$X_i \sim N(0,1)$，则$\sum_{i=1}^{n} a_i X_i \sim N\left(0, |a|^2\right)$
  - 当 $a_i = \frac{1}{n}$， $\frac{1}{n}\sum_{i=1}^{n} X_i \sim N\left(0, \frac{1}{n}\right)$

4. 多维连续随机变量函数的分布

- 例2: $X \sim \text{Exp}(1), Y \sim \text{Exp}(1)$，$X, Y$相互独立，求$Z = X + Y$的概率密度函数

- $f_Z(z) = \int_{-\infty}^{+\infty} f_X(z-y) f_Y(y) dy = \int_0^z e^{-(z-y)-y} dy = ze^{-z}$
- 也即 $Z \sim \Gamma(2,1)$

- 回顾：对于$\alpha, \lambda > 0$，定义概率密度函数
  - $f(x) = \frac{\lambda^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\lambda x}$，当$x \geq 0$
  - $f(x) = 0$，当$x < 0$

- 推广： $X \sim \Gamma(\alpha_1, \lambda), Y \sim \Gamma(\alpha_2, \lambda)$，$X, Y$相互独立， $X + Y \sim \Gamma(\alpha_1 + \alpha_2, \lambda)$

4. 多维连续随机变量函数的分布

- $X \sim U(0,1), Y \sim U(0,1)$，$X, Y$相互独立，求$Z = \max\{X, Y\}$的概率密度函数
  - $P(Z \le z) = P(X \le z)P(Y \le z) = z^2$
  - $f_Z(z) = 2z$当$z \in (0,1)$

- $X \sim U(0,1), Y \sim U(0,1)$，$X, Y$相互独立，求$Z = \min\{X, Y\}$的概率密度函数
  - $P(Z \ge z) = P(X \ge z)P(Y \ge z) = (1 - z)^2$
  - $P(Z \le z) = 1 - (1 - z)^2$
  - $f_Z(z) = 2(1 - z)$ 当 $z \in (0,1)$

4. 多维连续随机变量函数的分布

- 若 $X_1, X_2, \dots, X_n$相互独立，则$Y = \max\{X_1, X_2, \dots, X_n\}$的分布函数$F(y)$满足
  - $F_Y(y) = F_{X_1}(y) \cdot F_{X_2}(y) \cdot \cdots \cdot F_{X_n}(y)$
- 若 $X_1, X_2, \dots, X_n$相互独立，则$Y = \min\{X_1, X_2, \dots, X_n\}$的分布函数$F(y)$满足
  - $F_Y(y) = 1 - (1 - F_{X_1}(y)) \cdot (1 - F_{X_2}(y)) \cdot \cdots \cdot (1 - F_{X_n}(y))$

PEKING UNIVERSITY

4. 多维连续随机变量函数的分布

- 回顾：设X为连续随机变量，若函数$y = g(x)$严格单调，其反函数$h(y)$有连续导数，则$Y = g(X)$的概率密度函数为
  - $f_Y(y) = f_X\big(h(y)\big) \cdot |h'(y)|$ 当 $y \in (\alpha, \beta)$
  - $f_Y(y) = 0$当 $y \notin (\alpha, \beta)$

- 若连续随机变量X, Y的联合密度函数为$f(x, y)$。函数$u = u(x, y), v = v(x, y)$有连续偏导数且$x = x(u, v), y = y(u, v)$为唯一的反函数
- 则$U = u(X, Y), V = v(X, Y)$的联合密度函数为 $f(x(u, v), y(u, v)) \cdot |J|$，其中
  - $J = \left|\frac{\partial(x,y)}{\partial(u,v)}\right| = \begin{vmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} \end{vmatrix} = \left(\left|\frac{\partial(u,v)}{\partial(x,y)}\right|\right)^{-1} = \left(\begin{vmatrix} \frac{\partial u}{\partial x} & \frac{\partial u}{\partial y} \\ \frac{\partial v}{\partial x} & \frac{\partial v}{\partial y} \end{vmatrix}\right)^{-1}$

PEKING UNIVERSITY

4. 多维连续随机变量函数的分布

- 例1：若X,Y相互独立，且$X\sim N(\mu,1),Y\sim N(\mu,1)$。计算$U=X+Y$和$V=X-Y$的联合密度函数
- $u=x+y,v=x-y\Rightarrow x=\frac{u+v}{2},y=\frac{u-v}{2}$
- $J=\left|\frac{\partial(x,y)}{\partial(u,v)}\right|=\begin{vmatrix}\frac{\partial x}{\partial u} & \frac{\partial x}{\partial v}\\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v}\end{vmatrix}=\begin{vmatrix}\frac{1}{2} & \frac{1}{2}\\ \frac{1}{2} & -\frac{1}{2}\end{vmatrix}=-\frac{1}{2}\Rightarrow |J|=\frac{1}{2}$
- $U=X+Y$和$V=X-Y$的联合密度函数为
- $\frac{1}{\sqrt{2\pi}}e^{-\frac{\left(\frac{u+v}{2}-\mu\right)^2}{2}}\cdot\frac{1}{\sqrt{2\pi}}e^{-\frac{\left(\frac{u-v}{2}-\mu\right)^2}{2}}\cdot\left|-\frac{1}{2}\right|=\frac{1}{4\pi}e^{-\frac{1}{4}((u-2\mu)^2+v^2)}$
- 也即 $U\sim N(2\mu,2),V\sim N(0,2)$，且$U,V$相互独立

4. 多维连续随机变量函数的分布

- 例2：若连续随机变量X, Y相互独立，计算$U = XY$的概率密度函数
- $u = xy, v = y \Rightarrow x = u/v, y = v$
- $J = \left|\frac{\partial(x,y)}{\partial(u,v)}\right| = \begin{vmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} \end{vmatrix} = \begin{vmatrix} 1/v & -\frac{u}{v^2} \\ 0 & 1 \end{vmatrix} = \frac{1}{v} \Rightarrow |J| = \frac{1}{|v|}$
- U, V的联合密度函数为 $f\left(\frac{u}{v}, v\right) \cdot |J| = f_X\left(\frac{u}{v}\right) \cdot f_Y(v) \cdot \frac{1}{|v|}$
- U的边际密度函数为 $\int_{-\infty}^{+\infty} f_X\left(\frac{u}{v}\right) \cdot f_Y(v) \cdot \frac{1}{|v|} \cdot dv$

4. 多维连续随机变量函数的分布

- 例2：若连续随机变量X,Y相互独立，计算$U = XY$的概率密度函数
- U的概率密度函数为 $\int_{-\infty}^{+\infty} f_X\left(\frac{u}{v}\right) \cdot f_Y(v) \cdot \frac{1}{|v|} \cdot dv$

- 验证：
- $P(U \le u) = \iint_{xy\le u} f_X(x) f_Y(y)\, dxdy$
- $= \int_0^{+\infty} \int_{-\infty}^{u/y} f_X(x)dx \cdot f_Y(y)dy + \int_{-\infty}^{0} \int_{u/y}^{+\infty} f_X(x)dx \cdot f_Y(y)dy$
- 对u求导，$f_U(u) = \int_0^{+\infty} f_X(u/y) \cdot f_Y(y) \cdot \frac{1}{y} \cdot dy - \int_{-\infty}^{0} f_X\left(\frac{u}{y}\right) \cdot f_Y(y) \cdot \frac{1}{y} \cdot dy$
- $= \int_{-\infty}^{+\infty} f_X(u/y) \cdot f_Y(y) \cdot \frac{1}{|y|} \cdot dy$

线性代数入门（复习）

- 特征值/特征向量：对于$\boldsymbol{x} \neq \boldsymbol{0}$，若$\boldsymbol{A}\boldsymbol{x} = \lambda \boldsymbol{x}$，则$\lambda$是$\boldsymbol{A}$的特征值，$\boldsymbol{x}$是$\boldsymbol{A}$的特征向量
- 对角化： $\boldsymbol{A} = \boldsymbol{P}\boldsymbol{\Lambda}\boldsymbol{P}^{-1}$。$\boldsymbol{P}$的第$i$列为特征值为$\boldsymbol{\Lambda}_i$的特征向量
- $\det(\boldsymbol{A}) = \det(\boldsymbol{\Lambda}) = \prod_{i=1}^{n} \boldsymbol{\Lambda}_i$
- $\det\left(\boldsymbol{A}^2\right) = \det\left(\boldsymbol{\Lambda}^2\right) = \prod_{i=1}^{n} \boldsymbol{\Lambda}_i^2$
- $\operatorname{trace}(\boldsymbol{A}) = \operatorname{trace}(\boldsymbol{\Lambda}) = \sum_{i=1}^{n} \boldsymbol{\Lambda}_i$
- $\operatorname{rank}(\boldsymbol{A}) = \boldsymbol{\Lambda}$对角线非零元素数量

- 对称矩阵： $\boldsymbol{A}_{i,j} = \boldsymbol{A}_{j,i}$
- 正交矩阵： $\boldsymbol{U}\boldsymbol{U}^T = \boldsymbol{U}^T\boldsymbol{U} = \boldsymbol{I}$。 $|\boldsymbol{U}\boldsymbol{x}| = |\boldsymbol{x}|$
- 对称矩阵可对角化，且特征值一定为**实数**，不同特征值的特征向量**正交**
- 对称矩阵的正交分解（谱）： $\boldsymbol{A} = \boldsymbol{U}\boldsymbol{\Lambda}\boldsymbol{U}^{-1} = \boldsymbol{U}\boldsymbol{\Lambda}\boldsymbol{U}^T$

线性代数入门

- 二次型： $f(x)=\boldsymbol{x}^T\boldsymbol{A}\boldsymbol{x}$， $\boldsymbol{A}$为对称矩阵
- 半正定矩阵：对于$\boldsymbol{x}\neq\boldsymbol{0}$， $\boldsymbol{x}^T\boldsymbol{A}\boldsymbol{x}\geq 0$
- 正定矩阵，对于$\boldsymbol{x}\neq\boldsymbol{0}$， $\boldsymbol{x}^T\boldsymbol{A}\boldsymbol{x}>0$

- 半正定矩阵 $\Leftrightarrow$所有特征值均非负
- 正定矩阵 $\Leftrightarrow$所有特征值均为正数

- 对于半正定矩阵$\boldsymbol{A}$，存在对称矩阵$\boldsymbol{B}=\boldsymbol{A}^{1/2}$， $\boldsymbol{B}^T\boldsymbol{B}=\boldsymbol{B}^2=\boldsymbol{A}$
- 对于正定矩阵，$\boldsymbol{B}$可逆，且$(\boldsymbol{B}^{-1})^2=\boldsymbol{A}^{-1}$

- 对于半正定（正定）矩阵$\boldsymbol{A}$，满足$\boldsymbol{B}^T\boldsymbol{B}=\boldsymbol{A}$的$\boldsymbol{B}$不唯一

5. 多维高斯分布

- 随机向量$\boldsymbol{X}=(X_1,X_2,\dots,X_n)$。定义$E(\boldsymbol{X})=(E(X_1),E(X_2,),\dots,E(X_n))$为$\boldsymbol{X}$的数学期望向量，$\mathrm{Cov}(\boldsymbol{X})=E\left(\left(\boldsymbol{X}-E(\boldsymbol{X})\right)\left(\boldsymbol{X}-E(\boldsymbol{X})\right)^T\right)$为$\boldsymbol{X}$的协方差矩阵
- $\mathrm{Cov}(\boldsymbol{X})=\begin{pmatrix} \mathrm{Var}(X_1) & \mathrm{Cov}(X_1,X_2) & \dots & \mathrm{Cov}(X_1,X_n) \\ \mathrm{Cov}(X_2,X_1) & \mathrm{Var}(X_2) & \dots & \mathrm{Cov}(X_2,X_n) \\ \dots & \dots & \dots & \dots \\ \mathrm{Cov}(X_n,X_1) & \mathrm{Cov}(X_n,X_2) & \dots & \mathrm{Var}(X_n) \end{pmatrix}$
- 性质：$\mathrm{Cov}(\boldsymbol{X})$是半正定的对阵矩阵
  - $\boldsymbol{\alpha}^T\mathrm{Cov}(\boldsymbol{X})\boldsymbol{\alpha}=E\left(\left(\boldsymbol{\alpha}^T\left(\boldsymbol{X}-E(\boldsymbol{X})\right)\right)^2\right)=\mathrm{Var}\left(\boldsymbol{\alpha}^T\boldsymbol{X}\right)\geq 0$

PEKING UNIVERSITY 北京大学

5. 多维高斯分布

- $X_1, X_2 \sim N(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho)$，计算协方差矩阵和其逆矩阵
  - 协方差矩阵为$\boldsymbol{B} = \begin{pmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{pmatrix}$
  - $\det(\boldsymbol{B}) = (1-\rho^2)\sigma_1^2\sigma_2^2$
  - 逆矩阵为 $\boldsymbol{B}^{-1} = \frac{1}{1-\rho^2} \cdot \begin{pmatrix} \frac{1}{\sigma_1^2} & -\frac{\rho}{\sigma_1\sigma_2} \\ -\frac{\rho}{\sigma_1\sigma_2} & \frac{1}{\sigma_2^2} \end{pmatrix}$
- $f(\boldsymbol{x}) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \exp\left[-\frac{1}{2(1-\rho^2)} \cdot \left(\frac{(x_1-\mu_1)^2}{\sigma_1^2} + \frac{(x_2-\mu_2)^2}{\sigma_2^2} - \frac{2\rho(x_1-\mu_1)(x_2-\mu_2)}{\sigma_1\sigma_2}\right)\right]$
- $= \frac{1}{2\pi\cdot(\det(\boldsymbol{B}))^{\frac{1}{2}}} \exp\left(-\frac{1}{2} \cdot (\boldsymbol{x}-\boldsymbol{\mu})^T \boldsymbol{B}^{-1} (\boldsymbol{x}-\boldsymbol{\mu})\right)$

5. 多维高斯分布

- n维高斯分布的联合密度函数为
- $f(\boldsymbol{x}) = \frac{1}{(2\pi)^{\frac{n}{2}} \cdot (\det(\boldsymbol{B}))^{\frac{1}{2}}} \exp\left(-\frac{1}{2} \cdot (\boldsymbol{x}-\boldsymbol{\mu})^T \boldsymbol{B}^{-1} (\boldsymbol{x}-\boldsymbol{\mu})\right)$
- 数学期望向量：$\boldsymbol{\mu} \in \mathbb{R}^n$
- 协方差矩阵：$\boldsymbol{B} \in \mathbb{R}^{n\times n}$，$\boldsymbol{B}$ **可逆**
- 记号：$\boldsymbol{X} \sim N(\boldsymbol{\mu}, \boldsymbol{B})$
- 当 $\boldsymbol{\mu} = \boldsymbol{0}$，$\boldsymbol{B} = \boldsymbol{I}_n$
- $f(\boldsymbol{x}) = (2\pi)^{-\frac{n}{2}} \cdot \exp\left(-\frac{1}{2}|\boldsymbol{x}|^2\right) = \prod_{i=1}^n \frac{1}{\sqrt{2\pi}} \cdot \exp\left(-\frac{x_i^2}{2}\right)$
- 也即 $X_1, X_2, \ldots, X_n$相互独立且 $X_i \sim N(0,1)$

5. 多维高斯分布

- n维高斯分布的联合密度函数为
- $f(\boldsymbol{x}) = \frac{1}{(2\pi)^{\frac{n}{2}} \cdot (\det(\boldsymbol{B}))^{\frac{1}{2}}} \exp\left(-\frac{1}{2} \cdot (\boldsymbol{x}-\boldsymbol{\mu})^T \boldsymbol{B}^{-1} (\boldsymbol{x}-\boldsymbol{\mu})\right)$
- 性质：$\boldsymbol{X} \sim N(\boldsymbol{\mu}, \boldsymbol{B})$。令$\boldsymbol{Y} = \boldsymbol{AX} + \boldsymbol{b}$，则有$\boldsymbol{AX} + \boldsymbol{b} \sim N\left(A\boldsymbol{\mu} + \boldsymbol{b}, \boldsymbol{ABA}^T\right)$
- 这里要求A **行满秩。**
- 证明：若A是方阵（可逆）
- $f_Y(\boldsymbol{y}) = f_X(A^{-1}(\boldsymbol{y}-\boldsymbol{b})) \cdot \left|\frac{\partial x}{\partial y}\right| = f_X\left(\boldsymbol{A}^{-1}(\boldsymbol{y}-\boldsymbol{b})\right) \cdot \left|\frac{\partial y}{\partial x}\right|^{-1} = f_X(\boldsymbol{A}^{-1}(\boldsymbol{y}-\boldsymbol{b})) \frac{1}{|\det(\boldsymbol{A})|}$
- $= \frac{1}{(2\pi)^{\frac{n}{2}} \cdot (\det(\boldsymbol{B}))^{\frac{1}{2}}} \exp\left(-\frac{1}{2} \cdot (\boldsymbol{A}^{-1}(\boldsymbol{y}-\boldsymbol{b})-\boldsymbol{\mu})^T \boldsymbol{B}^{-1} (\boldsymbol{A}^{-1}(\boldsymbol{y}-\boldsymbol{b})-\boldsymbol{\mu})\right) \cdot \frac{1}{|\det(\boldsymbol{A})|}$
- $= \frac{1}{(2\pi)^{\frac{n}{2}} \cdot \left(\det(\boldsymbol{ABA}^T)\right)^{\frac{1}{2}}} \exp\left(-\frac{1}{2} \cdot (\boldsymbol{y}-\boldsymbol{A\mu}-\boldsymbol{b})^T \left(\boldsymbol{ABA}^T\right)^{-1} (\boldsymbol{y}-\boldsymbol{A\mu}-\boldsymbol{b})\right)$

5. 多维高斯分布

- 例：$X \sim N(\mu, \sigma^2)$，若$a \neq 0$，求$Y = aX + b$服从的分布
- 例：若X_i独立同分布，且$X_i \sim N(0,1)$，证明$\sum_{i=1}^{n} a_i X_i \sim N\left(0, |a|^2\right)$
- 例：若X, Y相互独立，且$X \sim N(\mu, 1), Y \sim N(\mu, 1)$。计算$U = X + Y$和$V = X - Y$的联合密度函数

- 例1：$\boldsymbol{X} \sim N(\boldsymbol{\mu}, \boldsymbol{B})$，给出$\boldsymbol{X}$的子向量服从的边际分布。给出$X_i$的边际分布
  - $X_i \sim N(\mu_i, B_{i,i})$

- 例2：$X, Y \sim N(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho)$，求$X + Y$服从的分布
  - $X + Y \sim N(\mu_1 + \mu_2, \sigma_1^2 + \rho_2^2 + 2\rho\sigma_1\sigma_2)$

5. 多维高斯分布

- 一般多维高斯分布可视为独立同分布标准正态分布**线性变换**后的结果
- $\boldsymbol{X} \sim N(\boldsymbol{0}, \boldsymbol{I_n})$，$\boldsymbol{Y} \sim N(\boldsymbol{\mu}, \boldsymbol{B})$。
- $\boldsymbol{Y} = \boldsymbol{B}^{1/2}\boldsymbol{X} + \boldsymbol{\mu}$
- $\boldsymbol{X} = \boldsymbol{B}^{-1/2}(\boldsymbol{Y} - \boldsymbol{\mu})$。白化（whitening）

- 对$\boldsymbol{Y} \sim N(0, I_n)$做线性变换$f(\boldsymbol{y}) = \boldsymbol{A}\boldsymbol{y}$
- 我们已经通过密度变换公式验证了$\boldsymbol{X} = \boldsymbol{A}\boldsymbol{Y} \sim N\left(\boldsymbol{0}, \boldsymbol{A}\boldsymbol{A}^T\right)$
- $f(\boldsymbol{x}) = \frac{1}{(2\pi)^{\frac{n}{2}} \cdot (\det(\boldsymbol{A}\boldsymbol{A}^T))^{\frac{1}{2}}} \exp\left(-\frac{1}{2} \cdot x^T (\boldsymbol{A}\boldsymbol{A}^T)^{-1} x\right)$

- 为什么线性变换之后的密度函数只和$\boldsymbol{B} = \boldsymbol{A}\boldsymbol{A}^T$有关?
- $\det(\boldsymbol{B})^{1/2}$的几何含义？$\boldsymbol{x}^T\boldsymbol{B}^{-1}\boldsymbol{x}$的几何含义?

5. 多维高斯分布

- 旋转不变性：$\boldsymbol{Y} \sim N(\boldsymbol{0}, \boldsymbol{I_n})$，若$\boldsymbol{V}$为正交矩阵，则$\boldsymbol{VY} \sim N(0, \boldsymbol{I_n})$
- 几何理解： $\boldsymbol{Y}$的概率密度函数正比于 $\exp\left(-\frac{1}{2}|\boldsymbol{y}|^2\right)$，正交变换$\boldsymbol{V}$不改变模长

- 奇异值分解（SVD）：对于任意矩阵$\boldsymbol{A}$，$\boldsymbol{A} = \boldsymbol{U\Sigma V}$
  - $\boldsymbol{U}$和 $\boldsymbol{V}$为正交矩阵
  - $\boldsymbol{\Sigma}$为对角矩阵，当$\boldsymbol{A}$可逆时全部对角元素非负
  - 几何理解：任何一个线性变换，都可以分解成旋转$\boldsymbol{V}$，拉伸 $\boldsymbol{\Sigma}$，再旋转$\boldsymbol{U}$

- 由旋转不变性，只需要关注拉伸$\boldsymbol{\Sigma}$和第二次旋转$\boldsymbol{U}$
- 拉伸$\boldsymbol{\Sigma}$和第二次旋转$\boldsymbol{U}$刻画了协方差矩阵$\boldsymbol{B} = \boldsymbol{AA}^T$的谱
- $\boldsymbol{B} = \boldsymbol{AA}^T = \boldsymbol{U\Sigma}^2\boldsymbol{U}^T$

5. 多维高斯分布

- $\boldsymbol{Y} \sim N(\boldsymbol{0}, \boldsymbol{I}_n)$的等密度轮廓线（equidensity contour）由欧式空间中的模长刻化
- $f_{\boldsymbol{Y}}(\boldsymbol{y}) \sim \exp\left(-\frac{1}{2}|\boldsymbol{y}|^2\right)$
- 线性变换后，$\boldsymbol{X} = \boldsymbol{A}\boldsymbol{y}$的等密度轮廓线由协方差矩阵$\boldsymbol{B} = \boldsymbol{A}\boldsymbol{A}^T$刻化
- $\boldsymbol{y}^T\boldsymbol{y} \leq 1 \Rightarrow \boldsymbol{x}^T\left(\boldsymbol{A}\boldsymbol{A}^T\right)^{-1}\boldsymbol{x} \leq 1$
- 线性变换后的单位球变为了由协方差矩阵$\boldsymbol{B}$刻化的**椭球**
- $\boldsymbol{B} = \boldsymbol{A}\boldsymbol{A}^T = \boldsymbol{U}\boldsymbol{\Sigma}^2\boldsymbol{U}^T$
  - $\boldsymbol{\Sigma}$刻化了椭球轴长
  - $\boldsymbol{U}$刻化了椭球轴的方向
  - $\left(\det(\boldsymbol{B})\right)^{1/2} = \det(\boldsymbol{\Sigma})$刻化了椭球的体积
- 正则化：概率密度函数中含有$\left(\det(\boldsymbol{B})\right)^{1/2}$