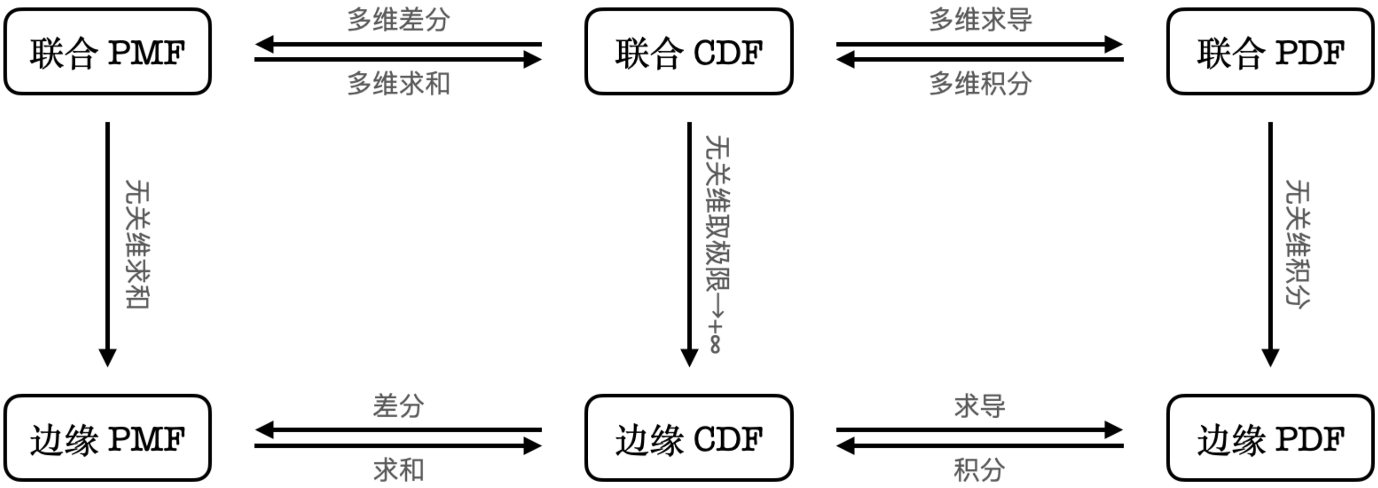
多元随机变量

xyfJASON

- 1 概览
- 2 二元正态分布相关
  - 2.1 定义
  - 2.2 密度分解
  - 2.3 边缘分布
  - 2.4 协方差与相关系数
  - 2.5 独立性

1 概览



联合 PMF
多维差分
多维求和
联合 CDF
多维求导
多维积分
联合 PDF
无关维求和
无关维取极限→+∞
无关维积分
边缘 PMF
差分
求和
边缘 CDF
求导
积分
边缘 PDF

2 二元正态分布相关

2.1 定义

若随机变量 X, Y 有如下联合概率密度函数:

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

称 X, Y 服从参数为 $\mu_1,\mu_2,\sigma_1,\sigma_2,\rho$ 的二元正态分布。

> 矩阵形式：设 $\mathbf{X}=\begin{pmatrix}x\\y\end{pmatrix},\mu=\begin{pmatrix}\mu_1\\\mu_2\end{pmatrix},\mathbf{\Sigma}=\begin{pmatrix}\sigma_1^2&\rho\sigma_1\sigma_2\\\rho\sigma_1\sigma_2&\sigma_2^2\end{pmatrix}$，则:
>
> $$p_{\mathbf{X}}(\mathbf{x})=\frac{1}{2\pi\sqrt{|\Sigma|}}\exp\left(-\frac{1}{2}(\mathbf{x}-\mu)^T\Sigma^{-1}(\mathbf{x}-\mu)\right)$$
>
> 这一形式可以推广到多元正态分布。

2.2 密度分解

对定义式进行变形可以得到:

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

注意到，前一部分是 $N(\mu_1,\sigma_1)$ 的概率密度函数，后一部分是 $N\left(\mu_2+\rho\frac{\sigma_2}{\sigma_1}(x-\mu_1),\sigma_2^2(1-\rho^2)\right)$ 的概率密度函数。又由于:

$$p_{X,Y}(x,y)=p_X(x)p_{Y|X}(y|x)$$

所以事实上后一部分是就是 $p_{Y|X}(y|x)$.

2.3 边缘分布

根据密度分解容易知道，二元正态分布的边缘分布就是正态分布，且：

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

2.4 协方差与相关系数

运用密度分解，可以计算：

$$\begin{aligned}
\operatorname{cov}(X,Y) &= \mathbb{E}[(X-\mathbb{E}X)(Y-\mathbb{E}Y)] \\
&= \iint_{\mathbb{R}^2} (x-\mu_1)(y-\mu_2) f_{X,Y}(x,y)\mathrm{d}x\mathrm{d}y \\
&= \int_{-\infty}^{+\infty} (x-\mu_1) f_X(x)\mathrm{d}x \int_{-\infty}^{+\infty} (y-\mu_2) f_{Y|X}(y|x)\mathrm{d}y \\
&= \int_{-\infty}^{+\infty} (x-\mu_1) f_X(x)\mathrm{d}x \left[\int_{-\infty}^{+\infty} y f_{Y|X}(y|x) - \mu_2\right] \\
&= \int_{-\infty}^{+\infty} (x-\mu_1) f_X(x)\mathrm{d}x \Big[\mathbb{E}[Y|X=x] - \mu_2\Big] \\
&= \int_{-\infty}^{+\infty} (x-\mu_1) f_X(x)\mathrm{d}x \Big[\rho\frac{\sigma_2}{\sigma_1}(x-\mu_1)\Big] \\
&= \rho\frac{\sigma_2}{\sigma_1} \int_{-\infty}^{+\infty} (x-\mu_1)^2 f_X(x)\mathrm{d}x \\
&= \rho\frac{\sigma_2}{\sigma_1}\operatorname{var}X \\
&= \rho\sigma_1\sigma_2
\end{aligned}$$

由此可得相关系数：

$$\rho(X,Y) = \frac{\operatorname{cov}(X,Y)}{\sqrt{\operatorname{var}X}\sqrt{\operatorname{var}Y}} = \rho$$

也即二元正态分布定义中的 ρ 就是其相关系数。

2.5 独立性

定理：设 (X,Y) 服从二元正态分布，则 X,Y 独立当且仅当 $\rho=0$.

证：由于 X,Y 独立蕴含着 X,Y 不相关，而后者等价于相关系数 $\rho=0$，所以独立 $\implies \rho=0$.

又设 $\rho=0$，则：

$$p_{X,Y}(x,y) = \frac{1}{\sqrt{2\pi}\sigma_1}\exp\left(-\frac{(x-\mu_1)^2}{2\sigma_1^2}\right)\cdot\frac{1}{\sqrt{2\pi}\sigma_2}\exp\left(-\frac{(y-\mu_2)^2}{2\sigma_2^2}\right) = p_X(x)p_Y(y)$$

所以 $\rho=0 \implies$ 独立。证毕。

由该定理可知，对于二元正态分布而言，独立和不相关是等价的。