Chapter 16: 高斯随机变量

Latest Update: 2025 年 1 月 4 日

Gaussian 积分:

$$\int \exp\left\{-a(x+b)^2\right\}dx = \sqrt{\frac{\pi}{a}}.$$

Exercise #16. 1. *设* $a \in \mathbb{R}^n, a \neq 0, \mu \in \mathbb{R}^n$. *设* H *是* $\mathbb{R}^n$ *中的一个超平面, 满足*

$$H = \{x \in \mathbf{R}^n : \langle x-\mu, a\rangle = 0\}.$$

证明: $m_n(H) = 0$, *其中* m_n *是一个* n *维欧几里得空间* $\mathbb{R}^n$ *上的 Lebesgue 测度. 并对任意* $\mathbb{R}^n$ *上的 Borel 函数推导以下的关系:*

$$\int_H f(x)dx = \int_{-\infty}^{\infty}\cdots\int_{\infty}^{\infty} f(x_1,\ldots,x_n)\,1_H(x_1,\ldots,x_n)\,dx_1\ldots dx_n = 0.$$

证明. 对第一部分, 由于 H 可以通过超平面 $x_n = 0$ 的正交和平移变换得到. 即存在正交变换使得 $Ta = e_n = (0,...,1)^\top$, 从而令 $y = Tx$ 有

$$a^\top(x-\mu) = 0 \Rightarrow y_n = a^\top\mu.$$

由于 Lebesgue 测度关于正交变换和平移变换是不变的, 从而只需证明 $m_n(H) = 0$ 对 $H = \{x_n = 0\}$ 成立. 由于 $\mathbb{Q}^n \sim \mathbb{Q}, \forall n \geq 1$, 整数. 从而记 $\mathbb{Q}^{n-1}\times\{0\} = \{\mathbf{q}_1, \mathbf{q}_2, ...\}$, $\forall\varepsilon > 0$, 取以 $\mathbf{q}_k$ 为中心, x_n 分量长为 ε, 其余分量长为 $\frac{1}{2^{k/(n-1)}}$ 的方体 A_k, 则根据有理数的稠密性,

$$H \subset \bigcup_{k=1}^{\infty} A_k.$$

于是

$$0 \leq m_n(H) \leq \sum_{k=1}^{\infty} m_n(A_k) = \sum_{k=1}^{\infty}\frac{\varepsilon}{2^k} = \varepsilon \to 0.$$

从而 $m_n(H) = 0$.

对第二部分, 第一个等式因为积分的定义. 因此, 只需证当 H 是零测集时,

$$\int_H f(x)dx = 0.$$

当 $f(x) = I_c(x)$ 是简单函数时, 由于 H 是零测集, 有

$$0 \leq \int_H f(x)dx = \int_H I_c(x)dx = m_n(H\cap C) \leq m_n(H) = 0.$$

从而根据线性性, 对于任意正 $f(x)$ 都有 $\int_H f(x)dx$. 最后, 对一般的 Borel 函数 $f(x)$, 取 f^+, f^-, 分别成立等于 0, 从而 $\int_H f(x)dx$. □

Exercise #16. 2. 记 $\mathcal{L}(Y)=\mathcal{N}(0,1)$, 令 $a>0$,

$$Z=\begin{cases} Y \text{ 若} |Y|\le a, \\ -Y \text{ 若} |Y|>a. \end{cases}$$

表明 $\mathcal{L}(Z)=N(0,1)$.

证明. 考察 Z 的特征函数,

$$\begin{aligned}
\varphi_Z(u) &= \mathbb{E}\left\{e^{iuZ}\right\} \\
&= \mathbb{E}\left\{e^{iuY} I_{\{|Y|\le a\}}\right\} + \mathbb{E}\left\{e^{-iuY} I_{\{|Y|>a\}}\right\} \\
&= \int_{|y|\le a} e^{iuy}\frac{1}{\sqrt{2\pi}}e^{-y^2/2}dy + \int_{|y|>a} e^{-iuy}\frac{1}{\sqrt{2\pi}}e^{-y^2/2}dy \\
&= \int_{|y|\le a} e^{iuy}\frac{1}{\sqrt{2\pi}}e^{-y^2/2}dy + \int_{|-z|>a} e^{iuz}\frac{1}{\sqrt{2\pi}}e^{-z^2/2}dz \quad \text{(积分方向改变和 Jacobian 抵消)} \\
&= \int_{\mathbb{R}} e^{iuy}\frac{1}{\sqrt{2\pi}}e^{-y^2/2}dy \\
&= e^{-u^2/2}.
\end{aligned}$$

由唯一性定理, $\mathcal{L}(Z)=N(0,1)$.

□

Exercise #16. 3. 设 X 服从 $\mathcal{N}(0,1)$, $Z \perp\!\!\!\perp X$, 且满足 $P(Z=1)=P(Z=-1)=\frac{1}{2}$. 令 $Y=XZ$. 证明 $Y\sim\mathcal{N}(0,1)$. 但是 (X,Y) 不是二元正态分布.

证明. 考察 Y 的特征函数, 由于 X 与 Z 独立, $\varphi_X(u)=e^{-\frac{1}{2}u^2}$. $\varphi_Z(u)=\dfrac{e^{iu}+e^{-iu}}{2}=cos(u)$. 从而用独立随机变量的特征函数

$$\begin{aligned}
\varphi_Y(u) &= \mathbb{E}\left\{e^{iuXZ}\right\} = \mathbb{E}\left\{e^{iuX}e^{iuZ}\right\} \\
&= \frac{1}{2}\mathbb{E}\{e^{iuX}\} + \frac{1}{2}\mathbb{E}\{e^{-iuX}\} \\
&= \frac{1}{2}\varphi_X(u) + \frac{1}{2}\varphi_X(-u) \\
&= e^{-\frac{1}{2}u^2}.
\end{aligned}$$

从而根据唯一性定理, $Y\sim N(0,1)$.

但是, 考察 X 与 Y 的联合特征函数,

$$\begin{aligned}\varphi_{X,Y}(u,v) &= \mathbb{E}\left\{e^{i(uX+vY)}\right\} \\ &= \mathbb{E}\left\{e^{i(uX+vXZ)}\right\} \\ &= \mathbb{E}\left\{e^{iX(u+vZ)}\right\} \\ &= \frac{1}{2}\mathbb{E}\left\{e^{iX(u+v)}\right\} + \frac{1}{2}\mathbb{E}\left\{e^{iX(u-v)}\right\} \\ &= \frac{1}{2}\varphi_X(u+v) + \frac{1}{2}\varphi_X(u-v) \\ &= e^{-\frac{1}{2}(u^2+v^2)}\cos(uv).\end{aligned}$$

从而根据唯一性定理, (X,Y) 不是二元正态分布. □

Exercise #16. 4. 设 (X,Y) 服从 *Gaussian* 分布, 均值 (μ_X,μ_Y), 协方差矩阵 Q 满足 $\det(Q)>0$. 令 ρ 是相关系数

$$\rho = \frac{\operatorname{Cov}(X,Y)}{\sqrt{\operatorname{Var}(X)\operatorname{Var}(Y)}}.$$

证明: 当 $\rho\in(-1,1)$, (X,Y) 的联合密度存在, 且为

$$f_{(X,Y)}(x,y) = \frac{1}{2\pi\sigma_X\sigma_Y\sqrt{1-\rho^2}}\exp\left\{\frac{-1}{2(1-\rho^2)}\left(\left(\frac{x-\mu_X}{\sigma_X}\right)^2 - \frac{2\rho(x-\mu_X)(y-\mu_Y)}{\sigma_X\sigma_Y} + \left(\frac{(y-\mu_Y)}{\sigma_Y}\right)^2\right)\right\}$$

再证明当 $\rho=-1,1$ 时, (X,Y) 的联合密度不存在.

证明. 根据 Cor 16.2, (X,Y) 的密度存在, 当且仅当 Q 是非退化的, 即 $\det(Q)\neq 0$.

$$\det(Q) = \begin{vmatrix}\sigma_X^2 & \rho\sigma_X\sigma_Y \\ \rho\sigma_X\sigma_Y & \sigma_Y^2\end{vmatrix} = \sigma_X^2\sigma_Y^2(1-\rho^2).$$

因此, 当 $\rho\in(-1,1)$, (X,Y) 的联合密度存在. 当 $\rho=\pm1$, $\det(Q)=0$, (X,Y) 的联合密度不存在. □

Exercise #16. 5. 设 ρ 取值在 $(-1,1)$, 给定 $\mu_j,\sigma_j^2, j=1,2$. 构造 X_1,X_2 是正态的, 均值是 μ_1,μ_2, 方差是 σ_1^2,σ_2^2, 相关系数是 ρ.

注. 令 Y_1,Y_2 独立同标准正态分布. 令 $U_1=Y_1, U_2=\rho Y_1+\sqrt{1-\rho^2}Y_2$. 接着令 $X_1=\mu_1+\sigma_1U_1, X_2=\mu_2+\sigma_2U_2$.

证明. 一个直观的想法是, 直接令 $\mu=(\mu_1,\mu_2)^\top, \Sigma=\begin{pmatrix}\sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2\end{pmatrix}$, 则 (X_1,X_2) 服从 $N(\mu,\Sigma)$.

提示中给出了一种构造方法, 令 Y_1,Y_2 独立同标准正态分布. 令 $U_1=Y_1, U_2=\rho Y_1+\sqrt{1-\rho^2}Y_2$. 则

$$\begin{pmatrix}U_1\\U_2\end{pmatrix}\sim N\left(\begin{pmatrix}0\\0\end{pmatrix},\begin{pmatrix}1 & \rho\\ \rho & 1\end{pmatrix}\right).$$

令 $X_1=\mu_1+\sigma_1 U_1, X_2=\mu_2+\sigma_2 U_2$, 则

$$\begin{pmatrix} X_1 \\ X_2 \end{pmatrix} \sim N\left(\begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}, \begin{pmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{pmatrix}\right).$$

这样就构造了满足条件的 X. □

Exercise #16. 6. *假设 X 服从 $\mathbb{R}^n$ 上的 Gaussian 分布 $N(\mu,\Sigma)$, 其中 $\det(\Sigma)>0$. 证明: 存在矩阵 B 使得 $Y=B(X-\mu)$ 服从 $N(0,I)$, 其中 I 是 n 阶单位矩阵.*

注. 这个题目表明: 任何非退化的 *Gaussian* 随机变量都可以通过线性变换得到标准正态分布.

证明. 由于 Σ 是正定的, 从而根据特征分解, 存在正交矩阵 $A^\top=A^{-1}$ 使得 $\Sigma=A\Lambda A^\top$. 其中 $\Lambda=\mathrm{diag}\{\lambda_1,...,\lambda_n\}, \lambda_i>0, \forall i$.

令 $B=A^\top\Lambda^{1/2}$, 其中 $\Lambda^{1/2}=\mathrm{diag}\{\lambda_1^{1/2},...,\lambda_n^{1/2}\}$, 于是

$$\mathbb{E}\{Y\}=\mathbb{E}\{B(X-\mu)\}=B(\mathbb{E}\{X\}-\mu)=0_n.$$

以及

$$\mathrm{Var}(Y)=B\Sigma B^\top=I_{n\times n}.$$

由定理 16.1, Y 服从 $N(0,I)$. □

Exercise #16. 7. *设 X 服从 Gaussian 分布, 令*

$$Y=\sum_{j=1}^n a_j X_j$$

其中 $X=(X_1,...,X_n)^\top$ 服从 $N(\mu,\Sigma)$. 证明: Y 服从一元 Gaussian 分布 $N(\mu,\sigma^2)$, 其中

$$\mu=\sum_{j=1}^n a_j E\left\{X_j\right\}$$

以及

$$\sigma^2=\sum_{j=1}^n a_j^2 \operatorname{Var}\left(X_j\right)+2\sum_{j<k} a_j a_k \operatorname{Cov}\left(X_j, X_k\right).$$

证明. 由于 X 服从 $N(\mu,\Sigma)$, Y 是一个线性组合, 从而根据定义, Y 也服从正态分布. 又因为 $\mathbb{E}(Y)=\mu, \mathrm{Var}(Y)=\sigma^2$, 根据特征函数, Y 服从一元 Gaussian 分布 $N(\mu,\sigma^2)$.

$$
\begin{aligned}
\mathbb{E}\{Y\} &= \mathbb{E}\left\{\sum_{j=1}^{n} a_j X_j\right\} = \sum_{j=1}^{n} a_j \mathbb{E}\left\{X_j\right\}, \\
\operatorname{Var}(Y) &= \operatorname{Cov}\left(\sum_{j=1}^{n} a_j X_j, \sum_{k=1}^{n} a_k X_k\right) \\
&= \sum_{j=1}^{n}\sum_{k=1}^{n} a_j a_k \operatorname{Cov}\left(X_j, X_k\right) \\
&= \sum_{j=1}^{n} \operatorname{Var}\left(X_j\right) + \sum_{j\neq k} a_j a_k \operatorname{Cov}\left(X_j, X_k\right) \\
&= \sum_{j=1}^{n} \operatorname{Var}\left(X_j\right) + 2\sum_{j<k} a_j a_k \operatorname{Cov}\left(X_j, X_k\right)
\end{aligned}
$$

□

注. *综合 16.6 和 16.7, 我们给出线性变换定理: 设 $X \sim N_p(\mu, \Sigma)$, B 为 $s \times p$ 常数矩阵, d 为 s 维常数向量, 则 $Y = BX + d$ 服从 $N_s(B\mu + d, B\Sigma B^\top)$.*

由于 Y 的特征函数

$$
\begin{aligned}
\varphi_Y(u) &= \mathbb{E}\{e^{iu^\top Y}\} \\
&= e^{iu^\top d}\mathbb{E}\{e^{iu^\top BX}\} \\
&= e^{iu^\top d}\varphi_X(B^\top u) \\
&= \exp\left\{i\langle u, d\rangle + i\langle u, BX\rangle - \frac{1}{2}\langle u, B\Sigma B^\top u\rangle\right\}.
\end{aligned}
$$

Exercise #16. 8. 设 (X, Y) 服从二元正态分布 $N(\mu, \Sigma)$, 其中

$$
\Sigma = \begin{pmatrix} \sigma_X^2 & \rho\sigma_X\sigma_Y \\ \rho\sigma_X\sigma_Y & \sigma_Y^2 \end{pmatrix}
$$

ρ 是相关系数, 满足 $|\rho| < 1$.(即 $\det(\Sigma) > 0$.) 则 (X, Y) 的联合密度为 f, 证明它们的条件密度 $f_{Y|X}(y)$ 是一元正态分布 $N\left(\mu_Y + \rho\frac{\sigma_Y}{\sigma_X}(x - \mu_X), \sigma_Y^2(1 - \rho^2)\right)$.

证明. 已经在 12.3 中证明过. 证明通过展开联合分布, 求出 X 的边际密度, 最后通过条件分布的定义求出条件密度. □

Exercise #16. 9. 设 X 服从 $\mathcal{N}(\mu, \Sigma)$, 其中 $\mu = (1, 1)^\top, \Sigma = \begin{pmatrix} 3 & 1 \\ 1 & 2 \end{pmatrix}$. 求出 $Y|Z$ 的分布, 其中 $Y = X_1 + X_2, Z = X_1 - X_2$.

证明. 由于

$$
\binom{Y}{Z} = \binom{X_1 + X_2}{X_1 - X_2} = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}\binom{X_1}{X_2},
$$

从而

$$\begin{pmatrix} Y \\ Z \end{pmatrix} \sim N\left(\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}\begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}\begin{pmatrix} 3 & 1 \\ 1 & 2 \end{pmatrix}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}\right).$$

计算得

$$\begin{pmatrix} Y \\ Z \end{pmatrix} \sim N\left(\begin{pmatrix} 2 \\ 0 \end{pmatrix}, \begin{pmatrix} 7 & 1 \\ 1 & 3 \end{pmatrix}\right).$$

由于条件分布为

$$Y|Z \sim N(\mu_{1\cdot 2}, \Sigma_{11\cdot 2}),$$

其中,

$$\mu_{1\cdot 2} = \mu_1 + \Sigma_{12}\Sigma_{22}^{-1}(z - \mu_2) = 2 + \frac{1}{3}(z - 0) = \frac{z}{3} + 2,$$

以及

$$\Sigma_{11\cdot 2} = \Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21} = 7 - \frac{1}{3} = \frac{20}{3}.$$

从而,

$$f_{Z=0}(y) = \frac{1}{\sqrt{2\pi}\sqrt{\frac{20}{3}}} \exp\left\{-\frac{1}{2}\frac{(y-2)^2}{\frac{20}{3}}\right\}$$

□

Exercise #16. 10. 设 X 服从 $\mathcal{N}(\mu, \Sigma)$, 满足 $\det(\Sigma) > 0$. 证明: *任意维元素的条件分布还是多元正态分布.*

注. *条件期望和条件方差的另一种表述是用 Schur 补.*

证明. 设 $X = \begin{pmatrix} X^{(1)} \\ X_2 \end{pmatrix}$, 其中, $X \in \mathbb{R}^p, X^{(1)} \in \mathbb{R}^{(r)}, X^{(2)} \in \mathbb{R}^{p-r}$. 不妨考虑 $X^{(1)}$ 是我们要求的条件变量, $X^{(2)}$ 是我们要求的条件变量. 由于 Gaussian 分布的前提是半正定的, $\det(\Sigma) > 0$, 从而 Σ 正定, 从而 Σ_{22} 正定, 从而 Σ_{22}^{-1} 存在. 从而考察非奇异线性变换. 令

$$Z = \begin{pmatrix} Z^{(1)} \\ Z^{(2)} \end{pmatrix} = \begin{pmatrix} X^{(1)} - \Sigma_{12}\Sigma_{22}^{-1}X^{(2)} \\ X^{(2)} \end{pmatrix} = \left(\begin{array}{c:c} I_r & -\Sigma_{12}\Sigma_{22}^{-1} \\ \hdashline 0 & I_{p-r} \end{array}\right)\begin{pmatrix} X^{(1)} \\ X^{(2)} \end{pmatrix} \triangleq B\begin{pmatrix} X^{(1)} \\ X^{(2)} \end{pmatrix}.$$

根据线性变换法则,

$$Z \sim N_p\left(\begin{pmatrix} \mu_1 - \Sigma_{12}\Sigma_{22}^{-1}\mu_2 \\ \mu_2 \end{pmatrix}, \left(\begin{array}{c:c} \Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21} & 0 \\ \hdashline 0 & \Sigma_{22} \end{array}\right)\right).$$

从而 $Z^{(1)}$ 与 $Z^{(2)}$ 相互独立. Z 的联合密度为

$$g(z^{(1)}, z^{(2)}) = g_1(z^{(1)})g_2(z^{(2)}) = g_1(z^{(1)})f_2(z^{(2)})$$

其中, 由于 $Z^{(2)} = X^{(2)}$, $f_2(z^{(2)})$ 是 $X^{(2)}$ 的密度函数.

由于 $Z = BX$, 使用积分变换公式:

$$\begin{aligned} f(x^{(1)}, x^{(2)}) =& g(BX) \cdot J(z \to x) \\ =& g_1(x^{(1)} - \Sigma_{12}\Sigma_{22}^{-1}x^{(2)})g_2(x^{(2)}) \cdot |\det(B^\top)|_+ \\ =& g_1(x^{(1)} - \Sigma_{12}\Sigma_{22}^{-1}x^{(2)})f_2(x^{(2)}) \end{aligned}$$

由于

$$Z^{(1)} \sim N_r(\mu^{(1)} - \Sigma_{12}\Sigma_{22}^{-1}\mu^{(2)}, \Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}),$$

从而可以写出 $f(x^{(1)}|x^{(2)})$.

$$\begin{aligned}
f_1\left(x^{(1)} \mid x^{(2)}\right) = \frac{f\left(x^{(1)}, x^{(2)}\right)}{f_2\left(x^{(2)}\right)} &= g_1\left(x^{(1)} - \Sigma_{12}\Sigma_{22}^{-1}x^{(2)}\right) \\
=&\frac{1}{(2\pi)^{r/2}\left|\Sigma_{11\cdot 2}\right|^{1/2}} \exp\left[-\frac{1}{2}\left(x^{(1)} - \Sigma_{12}\Sigma_{22}^{-1}x^{(2)}\right.\right. \\
&-\left(\mu^{(1)} - \Sigma_{12}\Sigma_{22}^{-1}\mu^{(2)}\right)\Big)' \Sigma_{11\cdot 2}^{-1}\left(x^{(1)} - \Sigma_{12}\Sigma_{22}^{-1}x^{(2)}\right. \\
&\left.\left.-\left(\mu^{(1)} - \Sigma_{12}\Sigma_{22}^{-1}\mu^{(2)}\right)\right)\right] \\
=&\frac{1}{(2\pi)^{r/2}\left|\Sigma_{11\cdot 2}\right|^{1/2}} \exp\left[-\frac{1}{2}\left(x^{(1)} - \mu_{1\cdot 2}\right)' \Sigma_{11\cdot 2}^{-1}\left(x^{(1)} - \mu_{1\cdot 2}\right)\right]
\end{aligned}$$

其中

$$\begin{aligned}
\mu_{1\cdot 2} &= \mu^{(1)} + \Sigma_{12}\Sigma_{22}^{-1}\left(x^{(2)} - \mu^{(2)}\right), \\
\Sigma_{11\cdot 2} &= \Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}.
\end{aligned}$$

□

Exercise #16. 11. 设 (X, Y) 有联合密度:

$$f_{(X,Y)}(x, y) = c\exp\left\{-\left(1 + x^2\right)\left(1 + y^2\right)\right\}, \quad -\infty < x, y < \infty,$$

其中, 选择特定的 c 使得 $f_{(X,Y)}(x, y)$ 是一个联合密度. 证明: f 不是二元正态的密度函数, 但是 X 和 Y 的边际密度都是正态密度.

注. 这表明习题 *16.10* 的逆命题不成立.

证明. 这里 c 的计算不是简单的, 这里直接取 $c^{-1} = \iint \exp\left\{-\left(1 + x^2\right)\left(1 + y^2\right)\right\} dxdy$. 考察 X 的边际密度,

$$\begin{aligned}
f_X(x) &= \int_{-\infty}^{\infty} f_{(X,Y)}(x, y)\, dy \\
&= c\int_{-\infty}^{\infty} \exp\left\{-\left(1 + x^2 + y^2 + x^2y^2\right)\right\} dy \\
&= c\exp\left(-1 - x^2\right)\int_{-\infty}^{\infty} \exp\left\{-(x^2 + 1)y^2\right\} dy \\
&= c\exp\left(-1 - x^2\right)\sqrt{\frac{\pi}{x^2 + 1}}.
\end{aligned}$$

从而条件分布是

$$\begin{aligned}
f_{Y|X}(y|x) &= \frac{f_{(X,Y)}(x, y)}{f_X(x)} \\
&= \frac{\sqrt{x^2 + 1}}{\sqrt{\pi}} \exp\left\{-\left(y^2 - x^2y^2\right)\right\} \\
&= \left\{2\pi \cdot \frac{1}{2(x^2 + 1)}\right\}^{-\frac{1}{2}} \exp\left\{-\frac{y^2}{2 \cdot \frac{1}{2(x^2+1)}}\right\}.
\end{aligned}$$

从而 $Y|X=x$ 是正态分布 $\mathcal{N}\left(0,\dfrac{1}{2(x^2+1)}\right)$. 由于对称性, $X|Y=y$ 服从 $\mathcal{N}\left(0,\dfrac{1}{2(y^2+1)}\right)$.

从而 X 和 Y 的条件密度都是正态密度. 但是 $f_{(X,Y)}(x,y)$ 不是二元正态的密度函数. 因为二元正态密度不会出现 x^2y^2 这种项. $\square$

Exercise #16. 12. 设 (X,Y) 是二元正态分布, 有相关系数 ρ, 均值为 $(0,0)^\top$. 证明*: 当* $|\rho|<1$, 则 $Z=\dfrac{X}{Y}$ 是 *Cauchy* 随机变量, 参数 $\alpha=\rho\dfrac{\sigma_X}{\sigma_Y},\beta=\dfrac{\sigma_X}{\sigma_Y}\sqrt{1-\rho^2}$. 我们总结*:* 中心化的二元正态分布的比是 *Cauchy* 分布.

注. 这个结果在例 *12.5* 中当 X,Y 是独立的情况已经证明.

以下摘抄自教材. 设 X,Y 独立同 $\mathcal{N}(0,\sigma^2)$. 令 $Z=\sqrt{X^2+Y^2},W=\begin{cases}\dfrac{X}{Y}, & Y\neq 0,\\ 0, & Y=0.\end{cases}$ 现在我们要求 Z,W 的联合密度.

考察 $g:(x,y)\mapsto(z,w),\mathbb{R}^2\to\mathbb{R}_+\times\mathbb{R}$ 的映射:

$$g(x,y)=\left(\sqrt{x^2+y^2},\frac{x}{y}\right)=(z,w).$$

显然 g 本身不是单射. 因此我们考虑 g 的单射区域. 记 $S_0=\{(x,y):x=0,y\in\mathbb{R}\}\cup\{(x,y):y=0,x\in\mathbb{R}\}$. $S_1=\{(x,y):y>0\}$ 为第一, 二象限, $S_2=\{(x,y):y<0\}$ 为第三, 四象限. 于是 $\mathbb{R}^2=\sum_{i=0}^{2}S_i$, 且 $m_2(S_0)=0$, S_1,S_2 是两片单射区域.

在 S_1 上, g 是单射, 且

$$g^{-1}(z,w)=\left(\frac{zw}{\sqrt{1+w^2}},\frac{z}{\sqrt{1+w^2}}\right).$$

此时, Jacobian:

$$J_1=\begin{vmatrix}\dfrac{w}{\sqrt{1+w^2}} & \dfrac{1}{\sqrt{1+w^2}}\\ \dfrac{z}{(1+w^2)^{\frac{3}{2}}} & -\dfrac{wz}{(1+w^2)^{\frac{3}{2}}}\end{vmatrix}=\frac{-z}{1+w^2}.$$

在 S_2 上, g 是单射, 且

$$g^{-1}(z,w)=\left(-\frac{zw}{\sqrt{1+w^2}},-\frac{z}{\sqrt{1+w^2}}\right).$$

此时, Jacobian:

$$J_2=\begin{vmatrix}-\dfrac{w}{\sqrt{1+w^2}} & -\dfrac{1}{\sqrt{1+w^2}}\\ -\dfrac{z}{(1+w^2)^{\frac{3}{2}}} & \dfrac{wz}{(1+w^2)^{\frac{3}{2}}}\end{vmatrix}=\frac{-z}{1+w^2}.$$

于是, Z,W 的联合密度为

$$f_{(Z,W)}(z,w)=\frac{z}{1+w^2}\left\{f_{(X,Y)}\left(\frac{zw}{\sqrt{1+w^2}},\frac{z}{\sqrt{1+w^2}}\right)+f_{(X,Y)}\left(\frac{-zw}{\sqrt{1+w^2}},\frac{-z}{\sqrt{1+w^2}}\right)\right\}.$$

从而,

$$f_{(Z,W)}(z,w)=\frac{2z}{1+w^2}(2\pi\sigma^2)^{-1}\exp\left\{-\frac{z^2}{2\sigma^2}\right\}\mathbb{I}_{\{Z>0\}}=\frac{1}{\pi\sigma^2}\left(\frac{1}{1+w^2}\right)\left(ze^{-\frac{z^2}{2\sigma^2}}\mathbb{I}_{\{Z>0\}}\right).$$

从而可以推断出 Z,W 独立, 且 W 服从参数为 $0,1$ 的 Cauchy 分布.

$$f_W(w)=\frac{1}{\pi(1+w^2)},\quad w\in\mathbb{R}.$$

以及 Z 服从参数为 σ^2 的 Rayleigh 分布.

$$f_Z(z)=\frac{z}{\sigma^2}e^{-\frac{z^2}{2\sigma^2}}\mathbb{I}_{\{z>0\}},\quad z\in\mathbb{R}_+.$$

证明. 使用上面的变换方法. 此时, (X,Y) 的联合密度为

$$f_{(X,Y)}(x,y)=\frac{1}{2\pi\sigma_X\sigma_Y\sqrt{1-\rho^2}}\exp\left\{\frac{-1}{2(1-\rho^2)}\left(\left(\frac{x}{\sigma_X}\right)^2-\frac{2\rho xy}{\sigma_X\sigma_Y}+\left(\frac{y}{\sigma_Y}\right)^2\right)\right\}$$

于是, (Z,W) 的联合密度为

$$\begin{aligned}
f_{(Z,W)}(z,w)&=\frac{z}{1+w^2}\left\{f_{(X,Y)}\left(\frac{zw}{\sqrt{1+w^2}},\frac{z}{\sqrt{1+w^2}}\right)+f_{(X,Y)}\left(\frac{-zw}{\sqrt{1+w^2}},\frac{-z}{\sqrt{1+w^2}}\right)\right\}\\
&=\frac{2z}{1+w^2}\left\{f_{(X,Y)}\left(\frac{zw}{\sqrt{1+w^2}},\frac{z}{\sqrt{1+w^2}}\right)\right\}\\
&=\frac{2z}{1+w^2}\left[\frac{1}{2\pi\sigma_X\sigma_Y\sqrt{1-\rho^2}}\exp\left\{\frac{-1}{2(1-\rho^2)}\left(\left(\frac{zw}{\sigma_X\sqrt{1+w^2}}\right)^2-\frac{2\rho\frac{zw}{\sqrt{1+w^2}}\frac{z}{\sqrt{1+w^2}}}{\sigma_X\sigma_Y}+\left(\frac{z}{\sigma_Y\sqrt{1+w^2}}\right)^2\right)\right\}\right]\\
&=\frac{2z}{1+w^2}\left[\frac{1}{2\pi\sigma_X\sigma_Y\sqrt{1-\rho^2}}\exp\left\{\frac{-1}{2(1-\rho^2)}\left[\frac{z^2w^2}{\sigma_X^2(1+w^2)}-\frac{2\rho z^2w}{\sigma_X\sigma_Y(1+w^2)}+\frac{z^2}{\sigma_Y^2(1+w^2)}\right]\right\}\right]\\
&=\frac{2z}{1+w^2}\left[\frac{1}{2\pi\sigma_X\sigma_Y\sqrt{1-\rho^2}}\exp\left\{\frac{-1}{2(1-\rho^2)\sigma_X^2\sigma_Y^2(1+w^2)}\left[z^2w^2\sigma_Y^2-2\rho z^2w\sigma_X\sigma_Y+z^2\sigma_X^2\right]\right\}\right]
\end{aligned}$$

关于 z 求积分. 令

$$p\triangleq\pi\sigma_X\sigma_Y\sqrt{1-\rho^2}(1+w^2),\quad q\triangleq-\frac{w^2\sigma_Y^2-2\rho w\sigma_X\sigma_Y+\sigma_X^2}{2\sigma_X^2\sigma_Y^2(1-\rho^2)(1+w^2)}$$

于是

$$f_W(w)=\int_0^\infty\frac{z}{p}\exp(qz^2)\,dz=-\frac{1}{2pq}.$$

即

$$\begin{aligned}
f_W(w)&=\frac{1}{\pi\cdot\frac{w^2\sigma_Y^2-2\rho w\sigma_X\sigma_Y+\sigma_X^2}{\sigma_X\sigma_Y\sqrt{1-\rho^2}}}\\
&=\frac{1}{\pi}\cdot\frac{\sigma_X\sqrt{1-\rho^2}}{\sigma_Y}\cdot\frac{1}{\frac{\sigma_X^2(1-\rho^2)}{\sigma_Y^2}+\left(w^2-2\rho\frac{\sigma_X}{\sigma_Y}w+\rho^2\frac{\sigma_X^2}{\sigma_Y^2}\right)}\\
&=\frac{1}{\beta\pi}\cdot\frac{1}{1+\left(\frac{x-\alpha}{\beta}\right)^2}
\end{aligned}$$

其中, $\alpha=\rho\dfrac{\sigma_X}{\sigma_Y},\beta=\dfrac{\sigma_X}{\sigma_Y}\sqrt{1-\rho^2}$. □

Exercise #16. 13. 令 (X,Y) 是二元正态变量, 均值是 0, 相关系数 ρ. 设 β 满足

$$\cos\beta=\rho\quad(0\le\beta\le\pi)$$

再证明:

$$P\{XY<0\}=\frac{\beta}{\pi}.$$

注. 在习题 *16.12* 中, 我们已经证明了若 $Z=\dfrac{X}{Y}, z=\rho\dfrac{\sigma_X}{\sigma_Y}$, 则

$$F_Z(z)=\frac{1}{2}+\frac{1}{\pi}\operatorname{Arctan}\left(\frac{z\sigma_Y-\rho\sigma_X}{\sigma_X\sqrt{1-\rho^2}}\right).$$

令 $\alpha=\arcsin(\rho)$, 其中 $\left(-\dfrac{\pi}{2}\le\alpha\le\dfrac{\pi}{2}\right)$, 并且用 $\arctan\dfrac{\rho}{\sqrt{1-\rho^2}}=\arcsin\rho$, 证明 $P(XY<0)=\dfrac{1}{2}-\dfrac{\alpha}{\pi}$.

证明. 若 Cauchy 分布的密度函数为

$$f_Z(z|\alpha,\beta)=\frac{1}{\beta\pi}\cdot\frac{1}{1+\left(\frac{z-\alpha}{\beta}\right)^2},$$

则 Cauchy 分布的分布函数为

$$F_Z(z)=\frac{1}{2}+\frac{1}{\pi}\arctan\left(\frac{z-\alpha}{\beta}\right),\quad z\in\mathbb{R}.$$

从而,

$$P(XY<0)=P\left(\frac{X}{Y}<0\right)=F_Z(0)=\frac{1}{2}-\frac{1}{\pi}\arctan\left(\frac{\alpha}{\beta}\right).$$

而由上一题,

$$\frac{\alpha}{\beta}=\frac{\rho}{\sqrt{1-\rho^2}}.$$

从而

$$P(XY<0)=\frac{1}{2}-\frac{\arcsin(\rho)}{\pi}=\frac{1}{\pi}\left\{\frac{\pi}{2}-\arcsin(\rho)\right\}=\frac{\arccos(\rho)}{\pi}.$$

但是由于 $\rho\in(-1,1)$, 于是 $\arcsin(\rho)\in(-\pi/2,\pi/2)$. 从而

$$P(XY<0)=\frac{\beta}{\pi}.$$

□

Exercise #16. 14. 设 (X,Y) 满足 *14.13* 中的条件, 证明:

$$P\{X>0,Y>0\}=P\{X<0,Y<0\}=\frac{1}{4}+\frac{\alpha}{2\pi}$$
$$P\{X>0,Y<0\}=P\{X<0,Y>0\}=\frac{1}{4}-\frac{\alpha}{2\pi}$$

证明. 由于 $1 = P(XY>0)+P(XY<0)+P(XY=0) = P(XY<0)+P(XY>0)$. 以及根据上一题,

$$P(XY<0) = P(X<0,Y>0)+P(X>0,Y<0) = \frac{1}{2}-\frac{\arcsin(\rho)}{\pi},$$

从而只需证,

$$P(X<0,Y>0) = P(X>0,Y<0), \quad P(X>0,Y>0) = P(X<0,Y<0).$$

而这, 由于 (X,Y) 的联合密度函数满足

$$f(x,y) = f(-x,-y)$$

于是

$$\int_0^\infty\int_0^\infty f(x,y)dxdy = \int_0^\infty\int_0^\infty f(-x,-y)dxdy = \int_{-\infty}^0\int_{-\infty}^0 f(x,y)dxdy.$$

于是本题得证.

□

Exercise #16. 15. 令 (X,Y) 服从二元正态分布, 密度函数是

$$f_{(X,Y)}(x,y) = \frac{1}{2\pi\sigma_X\sigma_Y\sqrt{1-\rho^2}}e^{-\frac{1}{2(1-\rho^2)}\left(\frac{x^2}{\sigma_X^2}-\frac{2\rho xy}{\sigma_X\sigma_Y}+\frac{y^2}{\sigma_Y^2}\right)}.$$

证明:

a) $\mathbb{E}\{XY\} = \rho\sigma_X\sigma_Y$.

b) $\mathbb{E}\{X^2Y^2\} = \mathbb{E}\{X^2\}\mathbb{E}\{Y^2\}+2(\mathbb{E}\{XY\})^2$.

c) $\mathbb{E}\{|XY|\} = \frac{2\sigma_X\sigma_Y}{\pi}(\cos\alpha+\alpha\sin\alpha)$, 其中 α 由 $\sin\alpha = \rho\left(-\frac{\pi}{2}\le\alpha\le\frac{\pi}{2}\right)$ 给出.

注. 问题二涉及了高阶矩的计算. 这个结果在概率论中叫 *Isserlis' theorem*, 在量子场论中叫 *Wick's theorem*.

证明. a) 由定义, $\rho\sigma_X\sigma_Y = \mathrm{cov}(X,Y) = \mathbb{E}\{XY\}-\mathbb{E}\{X\}\mathbb{E}\{Y\} = \mathbb{E}\{XY\}$.

b) 根据 (X,Y) 的特征函数,

$$\begin{aligned}\varphi_{(X,Y)}(t_1,t_2) = \mathbb{E}\{e^{i(t_1X+t_2Y)}\} &= \exp\left\{-\frac{1}{2}(t_1\quad t_2)\begin{pmatrix}\sigma_X^2 & \rho\sigma_X\sigma_Y\\ \rho\sigma_X\sigma_Y & \sigma_Y^2\end{pmatrix}\begin{pmatrix}t_1\\ t_2\end{pmatrix}\right\}\\ &= \exp\left\{-\frac{1}{2}\left(\sigma_X^2t_1^2+2\rho\sigma_X\sigma_Yt_1t_2+\sigma_Y^2t_2^2\right)\right\}.\end{aligned}$$

于是,

$$\begin{aligned}\mathbb{E}\{X^2Y^2\} &= \frac{\partial^4\varphi_{(X,Y)}(t_1,t_2)}{\partial t_1^2\partial t_2^2}\Big|_{t_1=t_2=0}\\ &= \frac{\partial^4}{\partial t_1^2\partial t_2^2}\exp\left\{-\frac{1}{2}\left(\sigma_X^2t_1^2+2\rho\sigma_X\sigma_Yt_1t_2+\sigma_Y^2t_2^2\right)\right\}\Big|_{t_1=t_2=0}\end{aligned}$$

记

$$f = \exp\left\{-\frac{1}{2}\left(\sigma_X^2t_1^2+2\rho\sigma_X\sigma_Yt_1t_2+\sigma_Y^2t_2^2\right)\right\},$$

于是
$$\frac{\partial f}{\partial t_1}=f\cdot\left(-\sigma_X^2 t_1-\rho\sigma_X\sigma_Y t_2\right),$$
再关于 t_1 求导,
$$\frac{\partial^2 f}{\partial t_1^2}=\frac{\partial f}{\partial t_1}\cdot\left(-\sigma_X^2 t_1-\rho\sigma_X\sigma_Y t_2\right)+f\cdot\left(-\sigma_X^2\right)=f\cdot\left(\sigma_X^4 t_1^2+2\rho\sigma_X^3\sigma_Y t_1 t_2+\rho^2\sigma_X^2\sigma_Y^2 t_2^2-\sigma_X^2\right).$$
由于
$$\frac{\partial f}{\partial t_2}=f\cdot\left(-\rho\sigma_X\sigma_Y t_1-\sigma_Y^2 t_2\right),$$
于是,
$$\begin{aligned}
&\frac{\partial^3 f}{\partial t_1^2 t_2}\\
=&\frac{\partial f}{\partial t_2}\cdot\left(\sigma_X^4 t_1^2+2\rho\sigma_X^3\sigma_Y t_1 t_2+\rho^2\sigma_X^2\sigma_Y^2 t_2^2-\sigma_X^2\right)+f\cdot\left(2\rho\sigma_X^3\sigma_Y t_1+2\rho^2\sigma_X^2\sigma_Y^2 t_2\right)\\
=&f\cdot\left(\sigma_X^4 t_1^2+2\rho\sigma_X^3\sigma_Y t_1 t_2+\rho^2\sigma_X^2\sigma_Y^2 t_2^2-\sigma_X^2\right)\cdot\left(-\rho\sigma_X\sigma_Y t_1-\sigma_Y^2 t_2\right)+f\cdot\left(2\rho\sigma_X^3\sigma_Y t_1+2\rho^2\sigma_X^2\sigma_Y^2 t_2\right)\\
=&f\cdot\left(-\rho\sigma_X^5\sigma_Y t_1^3-\sigma_X^4\sigma_Y^2 t_1^2 t_2-2\rho^2\sigma_X^4\sigma_Y^2 t_1^2 t_2-2\rho\sigma_X^3\sigma_Y^3 t_1 t_2^2-\rho^3\sigma_X^3\sigma_Y^3 t_1 t_2^2-\rho^2\sigma_X^2\sigma_Y^4 t_2^3\right.\\
&\left.+\rho\sigma_X^3\sigma_Y t_1+\sigma_X^2\sigma_Y^2 t_2+2\rho\sigma_X^3\sigma_Y t_1+2\rho^2\sigma_X^2\sigma_Y^2 t_2\right)
\end{aligned}$$
最后,
$$\begin{aligned}
&\frac{\partial^4 f}{\partial t_1^2\partial t_2^2}\\
=&f\cdot\left(-\rho\sigma_X\sigma_Y t_1-\sigma_Y^2 t_2\right)\cdot\left(-\rho\sigma_X^5\sigma_Y t_1^3-\sigma_X^4\sigma_Y^2 t_1^2 t_2-2\rho^2\sigma_X^4\sigma_Y^2 t_1^2 t_2-2\rho\sigma_X^3\sigma_Y^3 t_1 t_2^2-\rho^3\sigma_X^3\sigma_Y^3 t_1 t_2^2-\rho^2\sigma_X^2\sigma_Y^4 t_2^3\right.\\
&\left.+\rho\sigma_X^3\sigma_Y t_1+\sigma_X^2\sigma_Y^2 t_2+2\rho\sigma_X^3\sigma_Y t_1+2\rho^2\sigma_X^2\sigma_Y^2 t_2\right)\\
&+f\cdot\left(g(t_1,t_2)+\sigma_X^2\sigma_Y^2+2\rho^2\sigma_X^2\sigma_Y^2\right)
\end{aligned}$$
综上所述, 我们有
$$\mathbb{E}\{X^2Y^2\}=\left.\frac{\partial^4 f}{\partial t_1^2\partial t_2^2}\right|_{t_1=t_2=0}=\sigma_X^2\sigma_Y^2+2\rho^2\sigma_X^2\sigma_Y^2=\mathbb{E}\{X^2\}\mathbb{E}\{Y^2\}+2(\mathbb{E}\{XY\})^2.$$

c) 要计算 $\mathbb{E}|XY|$, 其中 (X,Y) 服从二元正态分布, 密度函数为:
$$f_{(X,Y)}(x,y)=\frac{1}{2\pi\sigma_X\sigma_Y\sqrt{1-\rho^2}}e^{-\frac{1}{2(1-\rho^2)}\left(\frac{x^2}{\sigma_X^2}-\frac{2\rho xy}{\sigma_X\sigma_Y}+\frac{y^2}{\sigma_Y^2}\right)}.$$
我们可以通过以下步骤计算 $\mathbb{E}|XY|$:

设 $U=\dfrac{X}{\sigma_X}$ 和 $V=\dfrac{Y}{\sigma_Y}$, 则 (U,V) 服从标准二元正态分布, 相关系数为 ρ. 此时, 密度函数为:
$$f_{(U,V)}(u,v)=\frac{1}{2\pi\sqrt{1-\rho^2}}e^{-\frac{1}{2(1-\rho^2)}(u^2-2\rho uv+v^2)}.$$
$\mathbb{E}|XY|$ 可以表示为:
$$\mathbb{E}|XY|=\sigma_X\sigma_Y\mathbb{E}|UV|.$$
由于 (U,V) 是标准二元正态分布, 我们可以利用其对称性和积分技巧来计算 $\mathbb{E}|UV|$.

首先, 注意到:

$$\mathbb{E}|UV| = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} |uv| f_{(U,V)}(u,v)\, du\, dv.$$

由于 $f_{(U,V)}(u,v)$ 是关于 u 和 v 的偶函数, 且 $|uv|$ 也是偶函数, 因此可以将积分限制在第一象限, 然后乘以 4:

$$\mathbb{E}|UV| = 4 \int_{0}^{\infty} \int_{0}^{\infty} uv f_{(U,V)}(u,v)\, du\, dv.$$

为了简化积分, 我们可以进行极坐标变换. 设 $u = r\cos\theta$ 和 $v = r\sin\theta$, 其中 $r \geq 0$ 且 $0 \leq \theta \leq \frac{\pi}{2}$. 雅可比行列式为 r, 因此积分变为:

$$\mathbb{E}|UV| = 4 \int_{0}^{\frac{\pi}{2}} \int_{0}^{\infty} r^2 \cos\theta \sin\theta \cdot \frac{1}{2\pi\sqrt{1-\rho^2}} e^{-\frac{r^2(1-\rho\sin 2\theta)}{2(1-\rho^2)}}\, r\, dr\, d\theta.$$

化简后得到:

$$\mathbb{E}|UV| = \frac{2}{\pi\sqrt{1-\rho^2}} \int_{0}^{\frac{\pi}{2}} \cos\theta \sin\theta \int_{0}^{\infty} r^3 e^{-\frac{r^2(1-\rho\sin 2\theta)}{2(1-\rho^2)}}\, dr\, d\theta.$$

内层积分是关于 r 的积分, 可以通过变量替换和伽马函数来计算. 设 $s = \dfrac{r^2(1-\rho\sin 2\theta)}{2(1-\rho^2)}$, 则 $ds = \dfrac{r(1-\rho\sin 2\theta)}{1-\rho^2} dr$, 积分变为:

$$\int_{0}^{\infty} r^3 e^{-\frac{r^2(1-\rho\sin 2\theta)}{2(1-\rho^2)}}\, dr = \frac{2(1-\rho^2)}{1-\rho\sin 2\theta} \int_{0}^{\infty} s e^{-s}\, ds = \frac{2(1-\rho^2)}{1-\rho\sin 2\theta}.$$

将内层积分的结果代入外层积分, 得到:

$$\mathbb{E}|UV| = \frac{2}{\pi\sqrt{1-\rho^2}} \int_{0}^{\frac{\pi}{2}} \cos\theta \sin\theta \cdot \frac{2(1-\rho^2)}{1-\rho\sin 2\theta}\, d\theta.$$

化简后得到:

$$\mathbb{E}|UV| = \frac{4(1-\rho^2)}{\pi\sqrt{1-\rho^2}} \int_{0}^{\frac{\pi}{2}} \frac{\cos\theta\sin\theta}{1-\rho\sin 2\theta}\, d\theta.$$

注意到 $\sin 2\theta = 2\sin\theta\cos\theta$, 因此可以将积分进一步简化为:

$$\mathbb{E}|UV| = \frac{4(1-\rho^2)}{\pi\sqrt{1-\rho^2}} \int_{0}^{\frac{\pi}{2}} \frac{\sin\theta\cos\theta}{1-2\rho\sin\theta\cos\theta}\, d\theta.$$

设 $t = \sin\theta$, 则 $dt = \cos\theta\, d\theta$, 积分变为:

$$\mathbb{E}|UV| = \frac{4(1-\rho^2)}{\pi\sqrt{1-\rho^2}} \int_{0}^{1} \frac{t}{1-2\rho t^2}\, dt.$$

这个积分可以通过部分分式分解或直接积分来计算. 最终结果为:

$$\mathbb{E}|UV| = \frac{2}{\pi}\left(\frac{\rho}{\sqrt{1-\rho^2}} + \arcsin\rho\right).$$

由于 $\mathbb{E}|XY| = \sigma_X\sigma_Y\mathbb{E}|UV|$, 因此:

$$\mathbb{E}|XY| = \sigma_X\sigma_Y \cdot \frac{2}{\pi}\left(\frac{\rho}{\sqrt{1-\rho^2}} + \arcsin\rho\right).$$

□

Exercise #16. 16. 设 (X,Y) 是二元正态分布, 相关系数是 ρ, $\sigma_X^2 = \sigma_Y^2$. 证明: X 和 $Y - \rho X$ 独立.

证明. 只需证: $\operatorname{cov}(X, Y-\rho X) = 0$.

$$\operatorname{cov}(X, Y-\rho X) = \operatorname{cov}(X,Y) - \rho\operatorname{cov}(X,X) = \rho\sigma_X\sigma_Y - \rho\sigma_X^2 = 0.$$

□

Exercise #16. 17. 设 X 服从 $\mathcal{N}(\mu,\Sigma)$, 满足 $\det(\Sigma) > 0$. X 是取值在 $\mathbb{R}^n$ 中的. 证明:

$$(X-\mu)^\top\Sigma^{-1}(X-\mu) \quad \sim \quad \chi^2(n).$$

证明. 由于 Σ 正定, 则存在正定的平方根矩阵 $\Sigma^{\frac{1}{2}}$, 使得 $\Sigma = \Sigma^{\frac{1}{2}}\Sigma^{\frac{1}{2}}$. 考察

$$(X-\mu)^\top\Sigma^{-1/2}\Sigma^{-1/2}(X-\mu) = \left\{\Sigma^{-1/2}(X-\mu)\right\}^\top\left\{\Sigma^{-1/2}(X-\mu)\right\},$$

而根据线性变换定理,

$$Z \triangleq \Sigma^{-1/2}(X-\mu) \sim \mathcal{N}(0, I_n).$$

从而 Z_i 独立同 $\mathcal{N}(0,1)$, 于是 Z_i^2 独立同 $\chi^2(1)$. 根据 χ^2 分布的可加性 (特征函数),

$$\sum_{i=1}^n Z_i^2 = \sum_{i=1}^n \left(\Sigma^{-1/2}(X-\mu)\right)_i^2 = (X-\mu)^\top\Sigma^{-1}(X-\mu) \sim \chi^2(n).$$

□

Exercise #16. 18. 设 $X_1, ..., X_n$ 独立同 $\mathcal{N}(0,\sigma^2)$, 并且令

$$\bar{X} = \frac{1}{n}\sum_{j=1}^n X_j \text{ 和} S^2 = \frac{1}{n-1}\sum_{j=1}^n \left(X_j - \bar{X}\right)^2.$$

在习题 *15.13* 中, 已经证明 $\bar{X}$ 和 S^2 是独立的. 证明:

$$\sum_{j=1}^n X_j^2 = \sum_{j=1}^n \left(X_j - \bar{X}\right)^2 + n\bar{X}^2$$

并推导出 $\dfrac{(n-1)S^2}{\sigma^2} \sim \chi_{n-1}^2$, $\dfrac{n\bar{X}^2}{\sigma^2} \sim \chi_1^2$.

证明.

$$\begin{aligned}\sum_{j=1}^{n} X_j^2 &= \sum_{j=1}^{n}\left\{(\bar{X}+X_j-\bar{X})^2\right\}\\ &= \sum_{j=1}^{n}\left\{(\bar{X}^2+2\bar{X}(X_j-\bar{X})+(X_j-\bar{X})^2\right\}\\ &= n\bar{X}^2+\sum_{j=1}^{n}(X_j-\bar{X})^2.\end{aligned}$$

从而根据 Cochran 定理 (或根据 Fisher 定理),

$$\frac{(n-1)S^2}{\sigma^2}=\frac{1}{\sigma^2}\sum_{j=1}^{n}(X_j-\bar{X})^2\sim\chi^2_{n-1},\quad \frac{n\bar{X}^2}{\sigma^2}\sim\chi^2_1.$$

□

Exercise #16. 19. *设 $\varepsilon_1,...,\varepsilon_n$ 独立同 $\mathcal{N}(0,\sigma^2)$, 假设 $Y_i=\alpha+\beta x_i+\varepsilon_i, 1\le i\le n$. 假设 x_i 不全相等, 令 $\bar{x}=\frac{1}{n}\sum_{i=1}^{n}x_i$. 定义回归残差是:*

$$\hat{\varepsilon}_i=Y_i-A-Bx_i.$$

a) 证明: $\mathbb{E}\{\hat{\varepsilon}_i\}=0, 1\le i\le n$.

b) 证明:

$$\operatorname{Var}\left(\hat{\varepsilon}_i\right)=\sigma^2\left(\frac{n-1}{n}-\frac{\left(x_i-\bar{x}\right)^2}{\sum_{i=1}^{n}\left(x_i-\bar{x}\right)^2}\right)$$

证明. a)

$$\begin{aligned}\mathbb{E}\{\hat{\varepsilon}_i\}&=\mathbb{E}\{Y_i-A-Bx_i\}\\ &=\mathbb{E}\{Y_i\}-A-Bx_i\\ &=\alpha+\beta x_i-\alpha-\beta x_i=0.\end{aligned}$$

b) 由于

$$\operatorname{var}\left(b_1\right)=\operatorname{var}\left(\sum_{i=1}^{n}k_iY_i\right)=\sum_{i=1}^{n}k_i^2\operatorname{var}\left(Y_i\right)=\sigma^2\sum_{i=1}^{n}k_i^2=\frac{\sigma^2}{SS_{XX}}$$

$$\operatorname{cov}\left\{b_1,Y_i\right\}=\operatorname{cov}\left\{\sum_{i=1}^{n}k_iY_i,Y_i\right\}=\sum_{j=1}^{n}\operatorname{cov}\left\{k_jY_j,Y_i\right\}=\operatorname{cov}\left\{k_iY_i,Y_i\right\}=k_i\sigma^2$$

$$\operatorname{cov}\left\{b_1,\bar{Y}\right\}=\operatorname{cov}\left\{b_1,\sum_{i=1}^{n}\frac{1}{n}Y_i\right\}=\frac{1}{n}\sum_{i=1}^{n}k_i\sigma^2=0$$

$$\begin{aligned}&\operatorname{var}\left\{b_0\right\}=\operatorname{var}\left\{\bar{Y}-b_1\bar{X}\right\}=\operatorname{var}\{\bar{Y}\}+\bar{X}^2\operatorname{var}\left\{b_1\right\}-2\bar{X}\operatorname{cov}\left\{\bar{Y},b_1\right\}\\ &=\operatorname{var}\{\bar{Y}\}+\bar{X}^2\operatorname{var}\left\{b_1\right\}=\sigma^2\left(\frac{1}{n}+\frac{\bar{X}^2}{SS_{XX}}\right)=\frac{\sum X_i^2}{nSS_{XX}}\sigma^2\end{aligned}$$

$$
\begin{aligned}
\operatorname{Var}\left(\hat{\varepsilon}_i\right) &= \operatorname{Var}\left[Y_i-\bar{Y}-b_1\left(X_i-\bar{X}\right)\right] \\
&= \operatorname{Var}\left(Y_i\right)+\operatorname{Var}(\bar{Y})+\operatorname{Var}\left(b_1\right)\left(X_i-\bar{X}\right)^2-2 \operatorname{cov}\left(Y_i, \bar{Y}\right)-2\left(X_i-\bar{X}\right)\left[\operatorname{cov}\left(Y_i, b_1\right)-\operatorname{cov}\left(\bar{Y}, b_1\right)\right] \\
&= \sigma^2+\frac{\sigma^2}{n}+\frac{\left(X_i-\bar{X}\right)^2 \sigma^2}{S_{XX}}-\frac{2 \sigma^2}{n}-\frac{2\left(X_i-\bar{X}\right)^2 \sigma^2}{S_{XX}}+0 \\
&= \frac{(n-1) \sigma^2}{n}-\frac{\left(X_i-\bar{X}\right)^2 \sigma^2}{S_{XX}}
\end{aligned}
$$

□

Exercise #16. 20. *考虑习题 16.19 定义的误差和残差. 假设 σ 是已知的, 定义*

$$
\hat{\sigma}^2=\frac{1}{n} \sum_{i=1}^n \hat{\varepsilon}_i^2
$$

证明: $\mathbb{E}\{\hat{\sigma}^2\}=\frac{n-2}{n} \sigma^2$.

注. 由于 $\mathbb{E}\{\hat{\sigma}^2\} \neq \sigma^2$, 称 $\hat{\sigma}^2$ 是 σ^2 的有偏估计. 于是 σ^2 的无偏估计是 $\frac{n}{n-2} \hat{\sigma}^2$.

证明.

$$
\begin{aligned}
E(SSE) &= E\left(\sum_{i=1}^n e_i^2\right)=\sum_{i=1}^n E\left(e_i^2\right)=\sum_{i=1}^n \operatorname{var}\left(e_i\right) \\
&= \sum_{i=1}^n\left[\frac{(n-1) \sigma^2}{n}-\frac{\left(X_i-\bar{X}\right)^2 \sigma^2}{SS_{XX}}\right]=(n-1) \sigma^2-\sigma^2=(n-2) \sigma^2 \\
E(MSE) &= \frac{E(SSE)}{n-2}=\sigma^2
\end{aligned}
$$

□

Exercise #16. 21. *在习题 16.19, 16.20 的条件下, 证明:$(A, B), S^2$ 是独立的.*

证明.

$$
\begin{aligned}
\sum\left(Y_i-\mu_i\right)^2 &= \sum\left[\left(Y_i-\hat{Y}_i\right)+\left(\hat{Y}_i-\mu_i\right)\right]^2 \\
&= \sum\left(\hat{Y}_i-\mu_i\right)^2+\sum\left(Y_i-\hat{Y}_i\right)^2 \\
&= \sum\left[\hat{\beta}_0^*+\hat{\beta}_1\left(X_i-\bar{X}\right)-\beta_0^*-\beta_1\left(X_i-\bar{X}\right)\right]^2+SS_E \\
&= n\left(\hat{\beta}_0^*-\beta_0^*\right)^2+\left(\hat{\beta}_1-\beta_1\right)^2 SS_{XX}+n \hat{\sigma}^2
\end{aligned}
$$

最后应用 Cocharn 定理即可. □

线性模型.

$$
Y_i=\alpha+\beta x_i+\varepsilon_i, 1 \leq i \leq n.
$$

其中 α, β, x_i 是常数.

一个典型的模型是考察个体在测量 $\alpha + \beta x_i$, 但是在测量时有测量误差 ε_i. 根据中心极限定理, 可以假设 $\{\varepsilon_i\}$ 服从多元正态分布. 可以称 x_i 为预测变量, Y_i 为响应变量.

假设 $\mathbb{E}\{\varepsilon_i\} = 0, 1 \le i \le n$, 对等式两边取期望, 有

$$\mathbb{E}\{Y_i\} = \alpha + \beta x_i.$$

一般来说, 我们想要知道 Y_i 与 x_i 之间的线性关系, 如果 ε 不存在当然知道这样的关系. 换句话说, 目标是在知道 x_i, Y_i 下, 找出 α, β. 首先, 我们要假设 x_i 是不全相等的, 否则我们最好给出一个常数估计 $\alpha + \beta x_1$, 但是此时无法识别出 α, β. x_i 不全相等, 当且仅当 $\sum\limits_{i=1}^{n}(x_i - \bar{x})^2 > 0$, 其中 $\bar{x} = \dfrac{1}{n}\sum\limits_{i=1}^{n} x_i$.

假设 ε_i 独立同分布于 $\mathcal{N}(0, \sigma^2)$, α 的估计量 U 和 β 的估计量 V 是随机变量 $Y_1, ..., Y_n$ 的函数, 是随机变量.

在所有这些估计量中, 考虑 "线性估计量", 即

$$U = u_0 + \sum_{i=1}^{n} u_i Y_i, \quad V = v_0 + \sum_{i=1}^{n} v_i Y_i.$$

其中 $u_0, u_1, ..., u_n, v_0, v_1, ..., v_n$ 是常数. 称 U, V 是无偏的, 如果 $\mathbb{E}\{U\} = \alpha, \mathbb{E}\{V\} = \beta$. 由于 $\mathbb{E}\{\varepsilon_i\} = 0$, 则无偏性蕴含:

$$\alpha = \mathbb{E}(U) = u_0 + \sum_{i=1}^{n} u_i(\alpha + \beta x_i) = u_0 + \alpha\left(\sum_{i=1}^{n} u_i\right) + \beta\left(\sum_{i=1}^{n} u_i x_i\right),$$
$$\beta = \mathbb{E}(V) = v_0 + \sum_{i=1}^{n} v_i(\alpha + \beta x_i) = v_0 + \alpha\left(\sum_{i=1}^{n} v_i\right) + \beta\left(\sum_{i=1}^{n} v_i x_i\right),$$

上述等式的成立应该对任意的 α, β 都成立. 从而我们有

$$\begin{aligned} u_0 = 0 \quad & \sum_{i=1}^{n} u_i = 1 \quad \text{and} \quad \sum_{i=1}^{n} u_i x_i = 0 \\ v_0 = 0 \quad & \sum_{i=1}^{n} v_i = 0 \quad \text{and} \quad \sum_{i=1}^{n} v_i x_i = 1 \end{aligned}$$

可以证明: 在均方误差最小意义下, 解为

$$v_i = \frac{x_i - \bar{X}}{\sum_{j=1}^{n}(x_j - \bar{x})^2}, \quad u_i = \frac{1}{n} - \bar{x} v_i.$$