定理5.16. 若 $U = g_1(X,Y)$ 和 $V = g_2(X,Y)$ 有连续偏导, 且存在反函数

$$x = x(u,v) \quad y = y(u,v),$$

则 (U,V) 的联合密度为

$$f_{UV}(u,v) = f_{XY}(x(u,v), y(u,v))|J|$$

其中 J 为变换的雅可比行列式, 即

$$|J| = \left|\frac{\partial(x,y)}{\partial(u,v)}\right| = \left|\frac{\partial(u,v)}{\partial x \partial y}\right|^{-1} = \begin{vmatrix} \frac{\partial u}{\partial x} & \frac{\partial u}{\partial y} \\ \frac{\partial v}{\partial x} & \frac{\partial v}{\partial y} \end{vmatrix}^{-1}.$$

上述结论可推广到一般的 n 维随机变量.

5.5 多维随机变量的数学特征

5.5.1 多维随机变量的期望

定理5.17. 设二维离散型随机变量 (X,Y) 的分布列为 $p_{ij} = P(X = x_i, Y = y_j)$, 则随机变量 $Z = g(X,Y)$ 的期望为

$$E[Z] = E[g(X,Y)] = \sum_{i,j} g(x_i, y_j) p_{ij};$$

设二维连续随机变量 (X,Y) 的概率密度为 $f(x,y)$, 则随机变量 $Z = g(X,Y)$ 的期望为

$$E[Z] = E[g(X,Y)] = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} g(x,y) f(x,y) dxdy.$$

例5.13. 设随机变量 $X \sim \mathcal{N}(0,1)$ 和 $Y \sim \mathcal{N}(0,1)$ 相互独立, 求 $E[\max(X,Y)]$.

解. 根据独立性定义可得随机变量 X 和 Y 的联合概率密度为

$$f(x,y) = f_X(x) f_Y(y) = \frac{1}{2\pi} e^{-\frac{x^2+y^2}{2}}.$$

于是得到

$$\begin{aligned} E[\max(X,Y)] &= \int\int_{D_1} x f(x,y) dxdy + \int\int_{D_2} y f(x,y) dxdy \\ &= \int_{-\infty}^{+\infty} dy \int_{y}^{+\infty} x f(x,y) dx + \int_{-\infty}^{+\infty} dx \int_{x}^{+\infty} y f(x,y) dy \\ &= 2\int_{-\infty}^{+\infty} dy \int_{y}^{+\infty} x f(x,y) dx = \frac{1}{\pi}\int_{-\infty}^{+\infty} dy \int_{y}^{+\infty} x e^{-\frac{x^2+y^2}{2}} dx \\ &= \frac{1}{\pi}\int_{-\infty}^{+\infty} e^{-y^2} dy = \frac{1}{\sqrt{\pi}} \end{aligned}$$

最后一个等式成立是因为$\int_{-\infty}^{+\infty} e^{-y^2} dy = \sqrt{\pi}$. □

课题练习. 在长度为 1 米的线段上任取两点 X, Y, 求 $E[\min(X, Y)], E[|X-Y|]$.

定理5.18. 对任意随机变量 X, Y 和常数 a, b, 有

$$E[aX+bY] = aE[X] + bE[Y];$$

对独立随机变量 X 和 Y, 以及任意函数 h, g, 有

$$E[XY] = E[X]E[Y] \quad \text{和} \quad E[h(X)g(Y)] = E[h(X)]E[g(Y)];$$

对任意随机变量 X 和 Y, 有 *Cauchy-Schwartz* 不等式

$$E[XY] \leq \sqrt{E[X^2]E[Y^2]}.$$

Proof. 设随机变量 X, Y 的联合概率密度为 $f(x, y)$, 则

$$\begin{aligned} E[aX+bY] &= \int\int (ax+by)f(x,y)dxdy \\ &= a\int\int xf(x,y)dxdy + b\int\int yf(x,y)dxdy = aE(X) + bE(Y). \end{aligned}$$

若随机变量 X 与 Y 独立, 则有

$$\begin{aligned} E[XY] &= \int\int xyf(x,y)dxdy = \int\int xf_X(x)yf_Y(y)dxdy \\ &= \int_{-\infty}^{+\infty} xf_X(x)dx \int_{-\infty}^{+\infty} yf_Y(y)dy = E(X)E(Y). \end{aligned}$$

对任意随机变量 X 与Y, 以及对任意 $t \in \mathbb{R}$ 有 $E[(X+tY)^2] \geq 0$ 成立, 即任意 $t \in \mathbb{R}$,

$$t^2E[Y^2] + E[X^2] + 2tE[XY] \geq 0.$$

因此有 $\Delta = 4[E(XY)]^2 - 4E[X^2]E[Y^2] \leq 0$, 即 $E(XY) \leq \sqrt{E(X^2)E(Y^2)}$. □

5.5.2 协方差

定理5.19. 对任意随机变量 X 与 Y 有

$$Var(X+Y) = Var(X) + Var(Y) + 2E[(X-E(X))(Y-E(Y))].$$

特别地, 当 X 与 Y 独立时有

$$Var(X+Y) = Var(X) + Var(Y).$$

Proof. 令 $Z = X+Y$, 有

$$\begin{aligned} Var(Z) &= E[(Z-EZ)^2] = E[(X-EX+Y-EY)^2] \\ &= E(X-EX)^2 + E(Y-EY)^2 + 2E[(X-EX)(Y-EY)] \\ &= Var(X) + Var(Y) + 2E[(X-EX)(Y-EY)]. \end{aligned}$$

若 X 与 Y 独立, 则 $2E[(X-EX)(Y-EY)] = 0$, 所以 $Var(X+Y) = Var(X) + Var(Y)$. □

定义5.12. 定义随机变量 X 和 Y 的协方差为

$$Cov(X,Y) = E[(X - E(X))(Y - E(Y))] = E(XY) - E(X)E(Y).$$

根据协方差定义和定理 5.19 有

$$Cov(X,X) = Var(X) \quad \text{和} \quad Var(X+Y) = Var(X) + Var(Y) + 2Cov(X,Y).$$

下面研究协方差的性质.

性质5.2. 对任意随机变量 X, Y 和常数 c, 有

$$Cov(X,Y) = Cov(Y,X) \quad \text{和} \quad Cov(X,c) = 0.$$

性质5.3. 对任意常数 a 和 b, 随机变量 X 和 Y, 有

$$Cov(aX,bY) = abCov(X,Y) \quad \text{和} \quad Cov(X+a,Y+b) = Cov(X,Y).$$

Proof. 根据协方差的定义有

$$\begin{aligned} Cov(aX,bY) &= E[(aX - E(aX))(bY - E(bY))] = abE[(X - E(X))(Y - E(Y))]; \\ Cov(X+a,Y+b) &= E[(X + a - E(X+a))(Y + b - E(Y+b))] = E[(X - E(X))(Y - E(Y))]. \end{aligned}$$

□

性质5.4. 对任意随机变量 X_1, X_2, Y, 有

$$Cov(X_1 + X_2, Y) = Cov(X_1, Y) + Cov(X_2, Y).$$

Proof. 我们有

$$\begin{aligned} & Cov(X_1 + X_2, Y) = E[(X_1 + X_2 - E(X_1) - E(X_2))(Y - E(Y))] \\ = \;& E[(X_1 - E(X_1))(Y - E(Y))] + E[(X_2 - E(X_2))(Y - EY)] = Cov(X_1,Y) + Cov(X_2,Y). \end{aligned}$$

□

由此性质可进一步得到: 对随机变量$X_1, X_2, \ldots, X_n$和$Y_1, Y_2, \ldots, Y_m$, 有

$$Cov\left(\sum_{i}^{n} X_i, \sum_{j}^{m} Y_j\right) = \sum_{i}^{n}\sum_{j}^{m} Cov(X_i, Y_j),$$

以及进一步有

$$Var\left(\sum_{i=1}^{n} X_i\right) = Cov\left(\sum_{i=1}^{n} X_i, \sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} Var(X_i) + 2\sum_{i<j} Cov(X_i, X_j).$$

性质5.5. 若随机变量 X 与 Y 独立, 则有 $Cov(X,Y) = 0$; 但反之不成立.

Proof. 若 X 与 Y·独立, 则

$$Cov(X,Y)=E[(X-E(X))(Y-E(Y))]=E[X-E(X)]E[Y-E(Y)]=0.$$

反之不成立, 例如随机变量 X 的分布列为

X	-1	0	1
P_i	$1/3$	$1/3$	$1/3$

当 $X\neq 0$ 时随机变量 $Y=0$, 否则$Y=1$, 根据联合分布列可知则 X与Y 不独立, 但此时有

$$Cov(X,Y)=E[(X-E(X))(Y-E(Y))]=E[XY]-E(X)E(Y)=0.$$

□

性质5.6. *对任意随机变量 X 与 Y 有*

$$(Cov(X,Y))^2\le Var(X)Var(Y)$$

等号成立的充要条件是 $Y=aX+b$ (即 X 与 Y 之间存在线性关系).

Proof. 由 Cauchy-Schwartz 不等式有

$$\begin{aligned}Cov(X,Y) &= E[(X-E(X))(Y-E(Y))]\\ &\le \sqrt{E[(X-E(X))^2]E[(Y-E(Y))^2]}=\sqrt{Var(X)Var(Y)}.\end{aligned}$$

下面证明等号成立的充要条件. 若 $Y=aX+b$, 则

$$Cov(X,Y)=Cov(X,aX+b)=aVar(X),\quad Var(Y)=a^2Var(X),$$

所以

$$Cov^2(X,Y)=a^2Var^2(X)=Var(X)a^2Var(X)=Var(X)Var(Y).$$

另一方面, 若 $(Cov(X,Y))^2=Var(X)Var(Y)$ 则有

$$(E[(X-EX)(Y-EY)])^2=E(X-EX)^2E(Y-EY)^2,$$

设

$$\begin{aligned}f(t) &= E[t(X-EX)-(Y-EY)]^2\\ &= t^2E[X-E(X)]^2-2tE[(X-E(X))(Y-E(Y))]+E[Y-E(Y)]^2\end{aligned}$$

根据一元二次方程的性质 $\Delta=4(E[(X-EX)(Y-EY)])^2-4E(X-EX)^2E(Y-EY)^2=0$ 可得方程 $f(t)=0$恰有一重根 t_0. 由此得到

$$f(t_0)=0\equiv E[(t_0(X-EX)-(Y-EY))^2]$$

根据 $(t_0(X-EX)-(Y-EY))^2\ge 0$ 可得 $Y=t_0(X-E(X))+E(Y)=aX+b$. □

课题练习. 随机变量 X 与 Y 独立, 且 $Var(X)=6$ 和 $Var(Y)=3$, 求 $Var(2X\pm Y)$.

课题练习. 随机变量 $X\sim P(2)$ 和 $Y\sim\mathcal{N}(-2,4)$, 且 X 与 Y 独立, 则$E[(X-Y)^2]$.

根据性质 5.6可知

$$\frac{Cov(X,Y)}{\sqrt{Var(X)Var(Y)}}\leq 1.$$

等号成立的充要条件是 X 与 Y 存在线性相关. 上式一定程度上反应了随机变量 X 和 Y 的线性相关程度, 由此引入一个新概念: 相关系数.

定义5.13. 设 X 和 Y 为二维随机变量, 如果 $Var(X), Var(Y)$ 存在且不为 0, 则称

$$\rho_{XY}=\frac{Cov(X,Y)}{\sqrt{Var(X)Var(Y)}}$$

为 X 与 Y 的相关系数, 简记 ρ.

关于相关系数, 我们需要注意:

- 这里使用相关系数而不是 $Cov(X,Y)$, 主要是规范 $|\rho_{XY}|\leq 1$, 而 $Cov(X,Y)$ 受数值大小影响;
- 相关系数 $|\rho_{XY}|\leq 1$: 若 $\rho>0$, X 与 Y 正相关; 若 $\rho<0$, X 与 Y 负相关; $|\rho_{XY}|=1$ 的充要条件为 X 与 Y 有线性关系 $Y=aX+b$. 本质上 ρ_{XY} 刻画了 X,Y 的线性相关程度, 又称为“线性相关系数”;
- 相关系数 $\rho=0$ 称 X 与 Y 不相关(线性不相关). 独立 $\Rightarrow$ 不相关, 不相关 $\not\Leftarrow$ 独立;
- 随机变量 X 与 Y 不相关, 仅表示 X 与 Y 之间无线性关系, 还可能存在其他关系. 例如: $X\sim U[-\frac{1}{2},\frac{1}{2}]$, $Y=cos(X)$. 易有$E(X)=0$,

$$\begin{aligned}Cov(X,Y)&=E[(X-E(X))(Y-E(Y))]\\&=\ E[X\cdot\cos(X)-XE(\cos(X))]=E[X\cdot\cos(X)]=\int_{-\frac{1}{2}}^{\frac{1}{2}}x\cdot\cos(x)dx=0.\end{aligned}$$

定理5.20. 对方差不为零的随机变量 X 和 Y, 下述条件相互等价:

- $\rho_{XY}=0$
- $Cov(X,Y)=0$
- $E(XY)=E(X)E(Y)$
- $Var(X\pm Y)=Var(X)+Var(Y)$

定理5.21. 对二维正态分布

$$\begin{pmatrix}X\\Y\end{pmatrix}\sim\mathcal{N}\left(\begin{pmatrix}\mu_1\\\mu_2\end{pmatrix},\begin{pmatrix}\sigma_1^2&\rho\sigma_1\sigma_2\\\rho\sigma_1\sigma_2&\sigma_2^2\end{pmatrix}\right),$$

有 $X\sim\mathcal{N}(\mu_1,\sigma_1^2)$, $Y\sim\mathcal{N}(\mu_2,\sigma_2^2)$, 以及 $Cov(X,Y)=\rho\sigma_1\sigma_2$, 即参数 ρ 为 X 与 Y 的相关系数; X 与 Y 独立 $\Longleftrightarrow$ X 与 X 不相关 (此结论仅限于正太分布).

例5.14. 随机变量 (X,Y) 联合概率密度为

$$f(x,y)=\begin{cases}(x+y)/8 & 0\le x\le 2, 0\le y\le 2\\ 0 & \text{其它}\end{cases}$$

求 $Cov(X,Y), Var(X+Y)$.

解. 根据协方差的定义有 $Cov(X,Y)=E[(X-E[X])(Y-E[Y])]=E[XY]-E[X]E[Y]$, 需要计算

$$\begin{aligned}E[X] &= \int_0^2\int_0^2 x(x+y)/8dxdy=7/6,\\ E[Y] &= \int_0^2\int_0^2 y(x+y)/8dxdy=7/6,\\ E[XY] &= \int_0^2\int_0^2 xy(x+y)/8dxdy=4/3,\end{aligned}$$

由此可得 $Cov(X,Y)=4/3-(7/6)^2=-1/36$. 进一步计算

$$\begin{aligned}E[X^2] &= \int_0^2\int_0^2 x^2(x+y)/8dxdy=5/3,\\ E[Y^2] &= \int_0^2\int_0^2 y^2(x+y)/8dxdy=5/3,\end{aligned}$$

由此可得 $Var(X)=Var(Y)=5/3-(7/6)^2=11/36$, 由此可得

$$\begin{aligned}Var(X+Y) &= Var(X)+Var(Y)+2Cov(X,Y)=11/18-1/18=5/9,\\ \rho_{XY} &= \frac{Cov(X,Y)}{\sqrt{Var(X)Var(Y)}}=-1/11.\end{aligned}$$

□

例5.15. 设随机变量 $X\sim\mathcal{N}(\mu,\sigma^2)$ 和 $Y\sim\mathcal{N}(\mu,\sigma^2)$ 相互独立. 求 $Z_1=\alpha X+\beta Y$ 和 $Z_2=\alpha X-\beta Y$ 的相关系数 $(\alpha,\beta\neq 0)$.

解. 根据正态分布的定义有

$$\begin{aligned}Cov(Z_1,Z_2) &= Cov(\alpha X+\beta Y,\alpha X-\beta Y)=(\alpha^2-\beta^2)\sigma^2\\ Var(Z_1) &= Cov(\alpha X+\beta Y,\alpha X+\beta Y)=(\alpha^2+\beta^2)\sigma^2\\ Var(Z_2) &= Cov(\alpha X-\beta Y,\alpha X-\beta Y)=(\alpha^2+\beta^2)\sigma^2\end{aligned}$$

由此可知 $\rho_{XY}=(\alpha^2-\beta^2)/(\alpha^2+\beta^2)$. □

课题练习. 随机变量 $X\sim\mathcal{N}(-1,2)$ 和 $Y\sim\mathcal{N}(1,8)$, 且相关系数 $\rho_{XY}=-1/2$. 求 $Var(X+Y)$.

5.5.3 随机向量的数学期望与协方差阵

定义5.14. 设 $X=(X_1,X_2,\cdots,X_n)^\top$, 则随机向量的期望

$$E(X)=(E(X_1),E(X_2),\cdots,E(X_n))^\top,$$

称随机变量 X 的协方差矩阵为

$$Cov(X) = \Sigma = \begin{pmatrix} Cov(X_1, X_1) & \cdots & Cov(X_1, X_n) \\ Cov(X_2, X_1) & \cdots & Cov(X_2, X_n) \\ \vdots & & \vdots \\ Cov(X_n, X_1) & \cdots & Cov(X_n, X_n) \end{pmatrix}.$$

定理5.22. 随机变量 X 的协方差矩阵是对称半正定的矩阵.

Proof. 证明根据函数的

$$\begin{aligned} f(t_1, t_2, \dots, t_n) &= (t_1, t_2, \dots, t_n) \begin{pmatrix} Cov(X_1, X_1) & \cdots & Cov(X_1, X_n) \\ Cov(X_2, X_1) & \cdots & Cov(X_2, X_n) \\ \vdots & & \vdots \\ Cov(X_n, X_1) & \cdots & Cov(X_n, X_n) \end{pmatrix} (t_1, t_2, \dots, t_n)^\top \\ &= (t_1(X_1 - E[X_1]) + t_2(X_2 - E[X_2]) + \cdots + t_n(X_n - E[X_n]))^2 \geq 0. \end{aligned}$$

□

定理5.23. 设多维正态分布 $X = (X_1, X_2, \dots, X_n)^\top \sim N(\mu, \Sigma)$, 则有

$$\mu = (E[X_1], E[X_2], \dots, E[X_n])^\top \quad \text{和} \quad \Sigma = [Cov(X_i, X_j)]_{n\times n}.$$

对多维随机变量 $X = (X_1, X_2, \cdots, X_n) \sim \mathcal{N}(\mu, \Sigma)$, 有

- 每个变量X_i的边缘分布是正态分布;
- $X_1, X_2, \cdots, X_n$ 相互独立 $\Longleftrightarrow$ $X_1, X_2, \cdots, X_n$ 相互不相关;
- $X \sim \mathcal{N}(\mu, \Sigma) \Longleftrightarrow \sum_{i=1}^n a_i X_i$ 是正态分布(对任意非全为0常数$a_1, a_2, \dots, a_n$).

5.6 条件分布与条件期望

前面学过随机事件的条件概率, 即在事件 B 发生的条件下事件 A 发生的条件概率

$$P(A|B) = P(AB)/P(B).$$

相关概念可推广到随机变量: 在给定随机变量 Y 取值条件下, 求 X 的概率分布, 即条件分布.

首先考虑离散型随机变量:

定义5.15. 设 (X, Y) 为二维离散型随机变量, 其分布列为 $\{p_{ij}\}$. 若 Y 的边缘分布 $P(Y = y_j) = p_{\cdot j} > 0$, 则称

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

为在 $Y = y_j$ 条件下随机变量 X 的条件分布列.