协方差

续本达

清华大学 工程物理系

2024-10-14 清华

复习

数学期望

设连续随机变量 X 的分布函数为 $F(x)$，若广义积分 $\int_{-\infty}^{+\infty} x\mathrm{d}F(x)$ 绝对收敛，则

$$\mathrm{E}(X) = \int_{-\infty}^{+\infty} x\mathrm{d}F(x)$$

设连续随机变量的密度函数为 $f(x)$ 若广义积分 $\int_{-\infty}^{+\infty} g(x)f(x)\mathrm{d}x$ 绝对收敛，则

$$\mathrm{E}(Y) = \int_{-\infty}^{+\infty} g(x)f(x)\mathrm{d}x$$

重期望公式

设 (X, Y) 是二维随机变量，且 $\mathrm{E}(X)$ 存在，则

$$\mathrm{E}(X) = \mathrm{E}[\mathrm{E}(X|Y)]$$

方差

定义

若 $E\{[X-\mathrm{E}(X)]^2\}$ 存在，则称其为随机变量 X 的方差，记为 $\mathrm{D}(X)$ 或 $\mathrm{Var}(X)$。即

$$\mathrm{Var}(X) = \mathrm{E}\{[X-\mathrm{E}(X)]^2\}$$

称 $\sqrt{\mathrm{Var}(X)}$ 为 X 的 均方差 或 标准差 ，两者量纲相同。

释义

描述随机变量 X 的取值偏离平均值的平均偏离程度。$\mathrm{D}(\cdot)$ 含义为 deviance，$\mathrm{Var}(\cdot)$ 为 variance。

计算方差的常用公式

$$\begin{aligned}\mathrm{Var}(X) &= \mathrm{E}\{[X-\mathrm{E}(X)]^2\} = \mathrm{E}\{X^2-2X\,\mathrm{E}(X)+[\mathrm{E}(X)]^2\} \\ &= \mathrm{E}(X^2)-2[\mathrm{E}(X)]^2+[\mathrm{E}(X)]^2 = \mathrm{E}(X^2)-[\mathrm{E}(X)]^2\end{aligned}$$

定义

若 $E\{[X-\mathrm{E}(X)]^2\}$ 存在, 则称其为随机变量 X 的方差，记为 $\mathrm{D}(X)$ 或 $\mathrm{Var}(X)$。即

$$\mathrm{Var}(X) = \mathrm{E}\{[X-\mathrm{E}(X)]^2\}$$

称 $\sqrt{\mathrm{Var}(X)}$ 为 X 的 均方差 或 标准差 ，两者量纲相同。

释义

描述随机变量 X 的取值偏离平均值的平均偏离程度。$\mathrm{D}(\cdot)$ 含义为 deviance，$\mathrm{Var}(\cdot)$ 为 variance。

计算方差的常用公式

$$\begin{aligned}\mathrm{Var}(X) &= \mathrm{E}\{[X-\mathrm{E}(X)]^2\} = \mathrm{E}\{X^2-2X\,\mathrm{E}(X)+[\mathrm{E}(X)]^2\} \\ &= \mathrm{E}(X^2)-2[\mathrm{E}(X)]^2+[\mathrm{E}(X)]^2 = \mathrm{E}(X^2)-[\mathrm{E}(X)]^2\end{aligned}$$

定义

若 $E\{[X-\mathrm{E}(X)]^2\}$ 存在, 则称其为随机变量 X 的方差，记为 $\mathrm{D}(X)$ 或 $\mathrm{Var}(X)$。即

$$\mathrm{Var}(X)=\mathrm{E}\{[X-\mathrm{E}(X)]^2\}$$

称 $\sqrt{\mathrm{Var}(X)}$ 为 X 的 均方差 或 标准差 ，两者量纲相同。

释义

描述随机变量 X 的取值偏离平均值的平均偏离程度。$\mathrm{D}(\cdot)$ 含义为 deviance，$\mathrm{Var}(\cdot)$ 为 variance。

计算方差的常用公式

$$\mathrm{Var}(X)=\mathrm{E}\{[X-\mathrm{E}(X)]^2\}=\mathrm{E}\{X^2-2X\,\mathrm{E}(X)+[\mathrm{E}(X)]^2\}$$
$$=\mathrm{E}(X^2)-2[\mathrm{E}(X)]^2+[\mathrm{E}(X)]^2=\mathrm{E}(X^2)-[\mathrm{E}(X)]^2$$

定义

若 $E\{[X-\mathrm{E}(X)]^2\}$ 存在，则称其为随机变量 X 的方差，记为 $\mathrm{D}(X)$ 或 $\mathrm{Var}(X)$。即

$$\mathrm{Var}(X) = \mathrm{E}\{[X-\mathrm{E}(X)]^2\}$$

称 $\sqrt{\mathrm{Var}(X)}$ 为 X 的 均方差 或 标准差 ，两者量纲相同。

释义

描述随机变量 X 的取值偏离平均值的平均偏离程度。$\mathrm{D}(\cdot)$ 含义为 deviance，$\mathrm{Var}(\cdot)$ 为 variance。

计算方差的常用公式

$$\begin{aligned}\mathrm{Var}(X) &= \mathrm{E}\{[X-\mathrm{E}(X)]^2\} = \mathrm{E}\{X^2-2X\,\mathrm{E}(X)+[\mathrm{E}(X)]^2\} \\ &= \mathrm{E}(X^2)-2[\mathrm{E}(X)]^2+[\mathrm{E}(X)]^2 = \mathrm{E}(X^2)-[\mathrm{E}(X)]^2\end{aligned}$$

计算

若 X 为离散型随机变量，分布律为 $P(X = x_k) = p_k, k = 1, 2, \cdots$

$$\operatorname{Var}(X) = \sum_{k=1}^{+\infty} [x_k - \mathrm{E}(X)]^2 p_k$$

若 X 为连续型随机变量，概率密度为 $f(x)$

$$\operatorname{Var}(X) = \int_{-\infty}^{+\infty} [x - \mathrm{E}(X)]^2 f(x) dx$$

性质

a, b, C 为常数，X, Y 为随机变量

- $\mathrm{Var}(C) = 0$
- $\mathrm{Var}(aX) = a^2 \mathrm{Var}(X)$
- 一般地 $\mathrm{Var}(aX + b) = a^2 \mathrm{Var}(X)$
- $\mathrm{Var}(X \pm Y) = \mathrm{Var}(X) + \mathrm{Var}(Y) \pm 2\,\mathrm{E}[(X - \mathrm{E}(X))(Y - \mathrm{E}(Y))]$
- 若 X, Y 相互独立

$$\mathrm{Var}(X \pm Y) = \mathrm{Var}(X) + \mathrm{Var}(Y)$$

- $\mathrm{Var}(X) = 0 \iff P[X = \mathrm{E}(X)] = 1$
  X 以概率 1 取常数 $\mathrm{E}(X)$。
- $\forall C \{\mathrm{Var}(X) \le \mathrm{E}[(X - C)^2]\}$
  方差是 $\mathrm{E}[(X - C)^2]$ 的下界。

性质

a, b, C 为常数，X, Y 为随机变量

- $\mathrm{Var}(C) = 0$
- $\mathrm{Var}(aX) = a^2 \mathrm{Var}(X)$
- 一般地 $\mathrm{Var}(aX + b) = a^2 \mathrm{Var}(X)$
- $\mathrm{Var}(X \pm Y) = \mathrm{Var}(X) + \mathrm{Var}(Y) \pm 2\,\mathrm{E}[(X - \mathrm{E}(X))(Y - \mathrm{E}(Y))]$
- 若 X, Y 相互独立

$$\mathrm{Var}(X \pm Y) = \mathrm{Var}(X) + \mathrm{Var}(Y)$$

- $\mathrm{Var}(X) = 0 \iff P[X = \mathrm{E}(X)] = 1$
  X 以概率 1 取常数 $\mathrm{E}(X)$。
- $\forall C \{\mathrm{Var}(X) \le \mathrm{E}[(X - C)^2]\}$
  方差是 $\mathrm{E}[(X - C)^2]$ 的下界。

性质

a, b, C 为常数，X, Y 为随机变量

- $\operatorname{Var}(C) = 0$
- $\operatorname{Var}(aX) = a^2 \operatorname{Var}(X)$
- 一般地 $\operatorname{Var}(aX + b) = a^2 \operatorname{Var}(X)$
- $\operatorname{Var}(X \pm Y) = \operatorname{Var}(X) + \operatorname{Var}(Y) \pm 2\,\mathrm{E}[(X - \mathrm{E}(X))(Y - \mathrm{E}(Y))]$
- 若 X, Y 相互独立

$$\operatorname{Var}(X \pm Y) = \operatorname{Var}(X) + \operatorname{Var}(Y)$$

- $\operatorname{Var}(X) = 0 \iff P[X = \mathrm{E}(X)] = 1$
  X 以概率 1 取常数 $\mathrm{E}(X)$。
- $\forall C \{\operatorname{Var}(X) \le \mathrm{E}[(X - C)^2]\}$
  方差是 $\mathrm{E}[(X - C)^2]$ 的下界。

性质

a, b, C 为常数，X, Y 为随机变量

- $\mathrm{Var}(C) = 0$
- $\mathrm{Var}(aX) = a^2 \mathrm{Var}(X)$
- 一般地 $\mathrm{Var}(aX + b) = a^2 \mathrm{Var}(X)$
- $\mathrm{Var}(X \pm Y) = \mathrm{Var}(X) + \mathrm{Var}(Y) \pm 2\,\mathrm{E}[(X - \mathrm{E}(X))(Y - \mathrm{E}(Y))]$
- 若 X, Y 相互独立

$$\mathrm{Var}(X \pm Y) = \mathrm{Var}(X) + \mathrm{Var}(Y)$$

- $\mathrm{Var}(X) = 0 \iff P[X = \mathrm{E}(X)] = 1$
  X 以概率 1 取常数 $\mathrm{E}(X)$。
- $\forall C\{\mathrm{Var}(X) \le \mathrm{E}[(X - C)^2]\}$
  方差是 $\mathrm{E}[(X - C)^2]$ 的下界。

例：泊松分布

例

$X \sim \pi(\lambda)$，求 $\operatorname{Var} X$。

回忆

$$\mathrm{E}(X) = \sum_{k=0}^{+\infty} k \cdot \frac{\lambda^k e^{-\lambda}}{k!} = \lambda e^{-\lambda} \sum_{k=1}^{+\infty} \frac{\lambda^{k-1}}{(k-1)!} = \lambda$$

受计算 $\mathrm{E}(X)$ 的启发，$\mathrm{E}[X(X-1)]$ 容易计算，然后再利用 $\mathrm{E}[X(X-1)] = \mathrm{E}(X^2) - \mathrm{E}(X)$ 求 $\mathrm{E}(X^2)$。

$$\mathrm{E}[X(X-1)] = \sum_{k=0}^{+\infty} k(k-1) \frac{\lambda^k e^{-\lambda}}{k!} = \lambda^2 e^{-\lambda} \sum_{k=2}^{+\infty} \frac{\lambda^{k-2}}{(k-2)!} = \lambda^2$$

于是，$\mathrm{E}(X^2) = \mathrm{E}[X(X-1)] + \mathrm{E}(X) = \lambda^2 + \lambda$

$$\mathrm{Var}(X) = \mathrm{E}(X^2) - \mathrm{E}^2(X) = \lambda$$

例：泊松分布

例

$X \sim \pi(\lambda)$，求 $\operatorname{Var} X$。

回忆

$$\mathrm{E}(X) = \sum_{k=0}^{+\infty} k \cdot \frac{\lambda^k e^{-\lambda}}{k!} = \lambda e^{-\lambda} \sum_{k=1}^{+\infty} \frac{\lambda^{k-1}}{(k-1)!} = \lambda$$

受计算 $\mathrm{E}(X)$ 的启发，$\mathrm{E}[X(X-1)]$ 容易计算，然后再利用 $\mathrm{E}[X(X-1)] = \mathrm{E}(X^2) - \mathrm{E}(X)$ 求 $\mathrm{E}(X^2)$。

$$\mathrm{E}[X(X-1)] = \sum_{k=0}^{+\infty} k(k-1)\frac{\lambda^k e^{-\lambda}}{k!} = \lambda^2 e^{-\lambda} \sum_{k=2}^{+\infty} \frac{\lambda^{k-2}}{(k-2)!} = \lambda^2$$

于是，$\mathrm{E}(X^2) = \mathrm{E}[X(X-1)] + \mathrm{E}(X) = \lambda^2 + \lambda$

$$\operatorname{Var}(X) = \mathrm{E}(X^2) - \mathrm{E}^2(X) = \lambda$$

例：泊松分布

例

$X \sim \pi(\lambda)$，求 $\operatorname{Var} X$。

回忆

$$\mathrm{E}(X) = \sum_{k=0}^{+\infty} k \cdot \frac{\lambda^k e^{-\lambda}}{k!} = \lambda e^{-\lambda} \sum_{k=1}^{+\infty} \frac{\lambda^{k-1}}{(k-1)!} = \lambda$$

受计算 $\mathrm{E}(X)$ 的启发，$\mathrm{E}[X(X-1)]$ 容易计算，然后再利用 $\mathrm{E}[X(X-1)] = \mathrm{E}(X^2) - \mathrm{E}(X)$ 求 $\mathrm{E}(X^2)$。

$$\mathrm{E}[X(X-1)] = \sum_{k=0}^{+\infty} k(k-1)\frac{\lambda^k e^{-\lambda}}{k!} = \lambda^2 e^{-\lambda} \sum_{k=2}^{+\infty} \frac{\lambda^{k-2}}{(k-2)!} = \lambda^2$$

于是，$\mathrm{E}(X^2) = \mathrm{E}[X(X-1)] + \mathrm{E}(X) = \lambda^2 + \lambda$

$$\operatorname{Var}(X) = \mathrm{E}(X^2) - \mathrm{E}^2(X) = \lambda$$

例：泊松分布

例

$X \sim \pi(\lambda)$，求 $\operatorname{Var} X$。

回忆

$$\mathrm{E}(X) = \sum_{k=0}^{+\infty} k \cdot \frac{\lambda^k e^{-\lambda}}{k!} = \lambda e^{-\lambda} \sum_{k=1}^{+\infty} \frac{\lambda^{k-1}}{(k-1)!} = \lambda$$

受计算 $\mathrm{E}(X)$ 的启发，$\mathrm{E}[X(X-1)]$ 容易计算，然后再利用 $\mathrm{E}[X(X-1)] = \mathrm{E}(X^2) - \mathrm{E}(X)$ 求 $\mathrm{E}(X^2)$。

$$\mathrm{E}[X(X-1)] = \sum_{k=0}^{+\infty} k(k-1) \frac{\lambda^k e^{-\lambda}}{k!} = \lambda^2 e^{-\lambda} \sum_{k=2}^{+\infty} \frac{\lambda^{k-2}}{(k-2)!} = \lambda^2$$

于是，$\mathrm{E}(X^2) = \mathrm{E}[X(X-1)] + \mathrm{E}(X) = \lambda^2 + \lambda$

$$\mathrm{Var}(X) = \mathrm{E}(X^2) - \mathrm{E}^2(X) = \lambda$$

例：泊松分布

例

$X \sim \pi(\lambda)$，求 $\operatorname{Var} X$。

回忆

$$\mathrm{E}(X) = \sum_{k=0}^{+\infty} k \cdot \frac{\lambda^k e^{-\lambda}}{k!} = \lambda e^{-\lambda} \sum_{k=1}^{+\infty} \frac{\lambda^{k-1}}{(k-1)!} = \lambda$$

受计算 $\mathrm{E}(X)$ 的启发，$\mathrm{E}[X(X-1)]$ 容易计算，然后再利用 $\mathrm{E}[X(X-1)] = \mathrm{E}(X^2) - \mathrm{E}(X)$ 求 $\mathrm{E}(X^2)$。

$$\mathrm{E}[X(X-1)] = \sum_{k=0}^{+\infty} k(k-1) \frac{\lambda^k e^{-\lambda}}{k!} = \lambda^2 e^{-\lambda} \sum_{k=2}^{+\infty} \frac{\lambda^{k-2}}{(k-2)!} = \lambda^2$$

于是，$\mathrm{E}(X^2) = \mathrm{E}[X(X-1)] + \mathrm{E}(X) = \lambda^2 + \lambda$

$$\operatorname{Var}(X) = \mathrm{E}(X^2) - \mathrm{E}^2(X) = \lambda$$

例：泊松分布

例

$X \sim \pi(\lambda)$，求 $\operatorname{Var} X$。

回忆

$$\mathrm{E}(X) = \sum_{k=0}^{+\infty} k \cdot \frac{\lambda^k e^{-\lambda}}{k!} = \lambda e^{-\lambda} \sum_{k=1}^{+\infty} \frac{\lambda^{k-1}}{(k-1)!} = \lambda$$

受计算 $\mathrm{E}(X)$ 的启发，$\mathrm{E}[X(X-1)]$ 容易计算，然后再利用 $\mathrm{E}[X(X-1)] = \mathrm{E}(X^2) - \mathrm{E}(X)$ 求 $\mathrm{E}(X^2)$。

$$\mathrm{E}[X(X-1)] = \sum_{k=0}^{+\infty} k(k-1)\frac{\lambda^k e^{-\lambda}}{k!} = \lambda^2 e^{-\lambda} \sum_{k=2}^{+\infty} \frac{\lambda^{k-2}}{(k-2)!} = \lambda^2$$

于是，$\mathrm{E}(X^2) = \mathrm{E}[X(X-1)] + \mathrm{E}(X) = \lambda^2 + \lambda$

$$\operatorname{Var}(X) = \mathrm{E}(X^2) - \mathrm{E}^2(X) = \lambda$$

例：泊松分布

例

$X \sim \pi(\lambda)$，求 $\operatorname{Var} X$。

回忆

$$\mathrm{E}(X) = \sum_{k=0}^{+\infty} k \cdot \frac{\lambda^k e^{-\lambda}}{k!} = \lambda e^{-\lambda} \sum_{k=1}^{+\infty} \frac{\lambda^{k-1}}{(k-1)!} = \lambda$$

受计算 $\mathrm{E}(X)$ 的启发，$\mathrm{E}[X(X-1)]$ 容易计算，然后再利用 $\mathrm{E}[X(X-1)] = \mathrm{E}(X^2) - \mathrm{E}(X)$ 求 $\mathrm{E}(X^2)$。

$$\mathrm{E}[X(X-1)] = \sum_{k=0}^{+\infty} k(k-1) \frac{\lambda^k e^{-\lambda}}{k!} = \lambda^2 e^{-\lambda} \sum_{k=2}^{+\infty} \frac{\lambda^{k-2}}{(k-2)!} = \lambda^2$$

于是，$\mathrm{E}(X^2) = \mathrm{E}[X(X-1)] + \mathrm{E}(X) = \lambda^2 + \lambda$

$$\operatorname{Var}(X) = \mathrm{E}(X^2) - \mathrm{E}^2(X) = \lambda$$

离散型随机变量的方差

分布	概率分布$P(X=k)$	方差
$(0-1)$分布	$p^k(1-p)^{1-k}$ $k=0,1$	$p(1-p)$
$b(n,p)$	$C_n^k p^k(1-p)^{n-k}$ $k=0,1,2,\cdots,n$	$np(1-p)$
$\pi(\lambda)$	$\lambda^k e^{-\lambda}/k!$ $k=0,1,2,\cdots$	λ
$h(n,M)$	$C_M^k C_{N-M}^{n-k}/C_N^n$ $k=0,1,\cdots,r \quad r=\min\{n,M\}$	$\dfrac{nM(N-M)(N-n)}{N^2(N-1)}$
$Ge(p)$	$(1-p)^{k-1}p$ $k=0,1,\cdots$	$\dfrac{1-p}{p^2}$

连续型随机变量的方差

分布	概率密度$f(x)$		方差
$U(a,b)$	$1/(b-a)$,	$a<x<b$	$(b-a)^2/12$
$Exp(\lambda)$	$\lambda e^{-\lambda x}$,	$x>0$	$1/\lambda^2$
$N(\mu,\sigma^2)$	$\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(x-\mu)^2}{2\sigma^2}}$,	$-\infty<x<\infty$	σ^2
$\Gamma(\alpha,\lambda)$	$\frac{\lambda^\alpha}{\Gamma(\alpha)}x^{\alpha-1}e^{-\lambda x}$,	$x\ge 0$	α/λ^2
$\chi^2(n)$	$\frac{x^{n/2-1}}{2^{n/2}\Gamma(n/2)}e^{-x/2}$,	$x>0$	$2n$
$Be(a,b)$	$\frac{1}{B(a,b)}x^{a-1}(1-x)^{b-1}$,	$0<x<1$	$\frac{ab}{(a+b)^2(a+b+1)}$
$Cau(\mu,\lambda)$	$\frac{1}{\pi}\frac{1}{1+x^2}$,	$-\infty<x<\infty$	不存在
朗道分布	$\frac{1}{\xi}\phi(\lambda)$，没有解析表达式		不存在

例：提枪作战问题

例

一个 n 人特种小队在休整时，把枪混放在一起。此时突发敌情，队友们不假思索地提枪应战。求恰好拿到自己枪的战士个数 X 的方差。

设 $X_i = 1$ 代表第 i 名战士拿到了自己的枪，否则 $X_i = 0$。根据问题的对称性可知 X_i 同分布。

例：提枪作战问题

例

一个 n 人特种小队在休整时，把枪混放在一起。此时突发敌情，队友们不假思索地提枪应战。求恰好拿到自己枪的战士个数 X 的方差。

设 $X_i = 1$ 代表第 i 名战士拿到了自己的枪，否则 $X_i = 0$。根据问题的对称性可知 X_i 同分布。

解：提枪作战问题

配对个数 X 的概率分布不容易写出，巧妙的办法是将 X 分解成 n 个 0-1 分布的随机变量 X_i 之和，求 $\mathrm{Var}(X = \sum_{i=1}^{n} X_i)$。

$X_1, X_2, \cdots, X_n$ 并不相互独立，所以 $\mathrm{Var}(X) \neq \sum_{i=1}^{n} \mathrm{Var}(X_i)$。需要利用定义计算：

$$
\begin{aligned}
\mathrm{Var}(X) &= \mathrm{E}(X^2) - [\mathrm{E}(X)]^2 = \mathrm{E}\left[\left(\sum_{i=1}^{n} X_i\right)^2\right] - \left[\mathrm{E}\left(\sum_{i=1}^{n} X_i\right)\right]^2 \\
&= \mathrm{E}\left[\sum_{i=1}^{n} X_i^2 + 2\sum_{1\leq i<j\leq n} X_i X_j\right] - \left[\sum_{i=1}^{n} \mathrm{E}(X_i)\right]^2 \\
&= n\,\mathrm{E}(X_i^2) + n(n-1)\,\mathrm{E}(X_i X_j) - n^2 E^2(X_i) (j \neq i)
\end{aligned}
$$

解：提枪作战问题

配对个数 X 的概率分布不容易写出，巧妙的办法是将 X 分解成 n 个 0-1 分布的随机变量 X_i 之和，求 $\mathrm{Var}(X = \sum_{i=1}^n X_i)$。
$X_1, X_2, \cdots, X_n$ 并不相互独立，所以 $\mathrm{Var}(X) \neq \sum_{i=1}^n \mathrm{Var}(X_i)$。需要利用定义计算：

$$
\begin{aligned}
\mathrm{Var}(X) &= \mathrm{E}(X^2) - [\mathrm{E}(X)]^2 = \mathrm{E}\left[\left(\sum_{i=1}^n X_i\right)^2\right] - \left[\mathrm{E}\left(\sum_{i=1}^n X_i\right)\right]^2 \\
&= \mathrm{E}\left[\sum_{i=1}^n X_i^2 + 2\sum_{1\le i<j\le n} X_i X_j\right] - \left[\sum_{i=1}^n \mathrm{E}(X_i)\right]^2 \\
&= n\,\mathrm{E}(X_i^2) + n(n-1)\,\mathrm{E}(X_i X_j) - n^2 E^2(X_i) (j \neq i)
\end{aligned}
$$

解：提枪作战问题

配对个数 X 的概率分布不容易写出，巧妙的办法是将 X 分解成 n 个 0-1 分布的随机变量 X_i 之和，求 $\mathrm{Var}(X=\sum_{i=1}^n X_i)$。
$X_1, X_2, \cdots, X_n$ 并不相互独立，所以 $\mathrm{Var}(X) \neq \sum_{i=1}^n \mathrm{Var}(X_i)$。需要利用定义计算：

$$\mathrm{Var}(X) = \mathrm{E}(X^2) - [\mathrm{E}(X)]^2 = \mathrm{E}\left[\left(\sum_{i=1}^n X_i\right)^2\right] - \left[\mathrm{E}\left(\sum_{i=1}^n X_i\right)\right]^2$$

$$= \mathrm{E}\left[\sum_{i=1}^n X_i^2 + 2\sum_{1\le i<j\le n} X_i X_j\right] - \left[\sum_{i=1}^n \mathrm{E}(X_i)\right]^2$$

$$= n\,\mathrm{E}(X_i^2) + n(n-1)\,\mathrm{E}(X_i X_j) - n^2 E^2(X_i)\,(j \neq i)$$

解：提枪作战问题

配对个数 X 的概率分布不容易写出，巧妙的办法是将 X 分解成 n 个 0-1 分布的随机变量 X_i 之和，求 $\mathrm{Var}(X=\sum_{i=1}^{n} X_i)$。
$X_1, X_2, \cdots, X_n$ 并不相互独立，所以 $\mathrm{Var}(X) \neq \sum_{i=1}^{n} \mathrm{Var}(X_i)$。需要利用定义计算：

$$
\begin{aligned}
\mathrm{Var}(X) &= \mathrm{E}(X^2) - [\mathrm{E}(X)]^2 = \mathrm{E}\left[\left(\sum_{i=1}^{n} X_i\right)^2\right] - \left[\mathrm{E}\left(\sum_{i=1}^{n} X_i\right)\right]^2 \\
&= \mathrm{E}\left[\sum_{i=1}^{n} X_i^2 + 2\sum_{1\le i<j\le n} X_i X_j\right] - \left[\sum_{i=1}^{n} \mathrm{E}(X_i)\right]^2 \\
&= n\,\mathrm{E}(X_i^2) + n(n-1)\,\mathrm{E}(X_i X_j) - n^2 E^2(X_i) \quad (j \neq i)
\end{aligned}
$$

解：提枪作战问题

配对个数 X 的概率分布不容易写出，巧妙的办法是将 X 分解成 n 个 0-1 分布的随机变量 X_i 之和，求 $\mathrm{Var}(X=\sum_{i=1}^n X_i)$。
$X_1, X_2, \cdots, X_n$ 并不相互独立，所以 $\mathrm{Var}(X) \neq \sum_{i=1}^n \mathrm{Var}(X_i)$。需要利用定义计算：

$$
\begin{aligned}
\mathrm{Var}(X) &= \mathrm{E}(X^2) - [\mathrm{E}(X)]^2 = \mathrm{E}\left[\left(\sum_{i=1}^n X_i\right)^2\right] - \left[\mathrm{E}\left(\sum_{i=1}^n X_i\right)\right]^2 \\
&= \mathrm{E}\left[\sum_{i=1}^n X_i^2 + 2\sum_{1\le i<j\le n} X_i X_j\right] - \left[\sum_{i=1}^n \mathrm{E}(X_i)\right]^2 \\
&= n\,\mathrm{E}(X_i^2) + n(n-1)\,\mathrm{E}(X_i X_j) - n^2 E^2(X_i) (j \neq i)
\end{aligned}
$$

解：提枪作战问题

需要计算出 $\mathrm{E}(X_i), \mathrm{E}(X_i^2), \mathrm{E}(X_iX_j)(j \neq i)$

$$\mathrm{E}(X_i) = 0 \times P(X_i = 0) + 1 \times P(X_i = 1) = \frac{1}{n}$$

$$\mathrm{E}(X_i^2) = 0 \times P(X_i^2 = 0) + 1 \times P(X_i^2 = 1) = \frac{1}{n}$$

$$\mathrm{E}(X_iX_j) = 0 \times P(X_i^2 = 0) + 1 \times P(X_iX_j = 1) = \frac{1}{n(n-1)}$$

于是

$$\begin{aligned}\mathrm{Var}(X) &= n\,\mathrm{E}(X_i^2) + n(n-1)\,\mathrm{E}(X_iX_j) - n^2\,\mathrm{E}^2(X_i)(j \neq i)\\ &= n \cdot 1/n + n(n-1) \cdot 1/n(n-1) - n^2 \cdot (1/n)^2 = 1\end{aligned}$$

思考：解答发给助教，答对者总评 +2%

求恰好有 k 名战士拿到自己枪的概率

解：提枪作战问题

需要计算出 $\mathrm{E}(X_i), \mathrm{E}(X_i^2), \mathrm{E}(X_iX_j)(j \neq i)$

$$\mathrm{E}(X_i) = 0 \times P(X_i = 0) + 1 \times P(X_i = 1) = \frac{1}{n}$$

$$\mathrm{E}(X_i^2) = 0 \times P(X_i^2 = 0) + 1 \times P(X_i^2 = 1) = \frac{1}{n}$$

$$\mathrm{E}(X_iX_j) = 0 \times P(X_i^2 = 0) + 1 \times P(X_iX_j = 1) = \frac{1}{n(n-1)}$$

于是

$$\begin{aligned}\mathrm{Var}(X) &= n\,\mathrm{E}(X_i^2) + n(n-1)\,\mathrm{E}(X_iX_j) - n^2\,\mathrm{E}^2(X_i)(j \neq i) \\ &= n \cdot 1/n + n(n-1) \cdot 1/n(n-1) - n^2 \cdot (1/n)^2 = 1\end{aligned}$$

思考：解答发给助教，答对者总评 +2%

求恰好有 k 名战士拿到自己枪的概率

解：提枪作战问题

需要计算出 $\mathrm{E}(X_i), \mathrm{E}(X_i^2), \mathrm{E}(X_iX_j)(j \neq i)$

$$\mathrm{E}(X_i) = 0 \times P(X_i = 0) + 1 \times P(X_i = 1) = \frac{1}{n}$$

$$\mathrm{E}(X_i^2) = 0 \times P(X_i^2 = 0) + 1 \times P(X_i^2 = 1) = \frac{1}{n}$$

$$\mathrm{E}(X_iX_j) = 0 \times P(X_i^2 = 0) + 1 \times P(X_iX_j = 1) = \frac{1}{n(n-1)}$$

于是

$$\begin{aligned}\mathrm{Var}(X) &= n\,\mathrm{E}(X_i^2) + n(n-1)\,\mathrm{E}(X_iX_j) - n^2\,\mathrm{E}^2(X_i)(j \neq i) \\ &= n \cdot 1/n + n(n-1) \cdot 1/n(n-1) - n^2 \cdot (1/n)^2 = 1\end{aligned}$$

思考：解答发给助教，答对者总评 +2%

求恰好有 k 名战士拿到自己枪的概率

解：提枪作战问题

需要计算出 $\mathrm{E}(X_i), \mathrm{E}(X_i^2), \mathrm{E}(X_i X_j)(j \neq i)$

$$\mathrm{E}(X_i) = 0 \times P(X_i = 0) + 1 \times P(X_i = 1) = \frac{1}{n}$$

$$\mathrm{E}(X_i^2) = 0 \times P(X_i^2 = 0) + 1 \times P(X_i^2 = 1) = \frac{1}{n}$$

$$\mathrm{E}(X_i X_j) = 0 \times P(X_i^2 = 0) + 1 \times P(X_i X_j = 1) = \frac{1}{n(n-1)}$$

于是

$$\mathrm{Var}(X) = n\,\mathrm{E}(X_i^2) + n(n-1)\,\mathrm{E}(X_i X_j) - n^2\,\mathrm{E}^2(X_i)(j \neq i)$$

$$= n \cdot 1/n + n(n-1) \cdot 1/n(n-1) - n^2 \cdot (1/n)^2 = 1$$

思考：解答发给助教，答对者总评 +2%

求恰好有 k 名战士拿到自己枪的概率

解：提枪作战问题

需要计算出 $\mathrm{E}(X_i), \mathrm{E}(X_i^2), \mathrm{E}(X_iX_j)(j \neq i)$

$$\mathrm{E}(X_i) = 0 \times P(X_i = 0) + 1 \times P(X_i = 1) = \frac{1}{n}$$

$$\mathrm{E}(X_i^2) = 0 \times P(X_i^2 = 0) + 1 \times P(X_i^2 = 1) = \frac{1}{n}$$

$$\mathrm{E}(X_iX_j) = 0 \times P(X_i^2 = 0) + 1 \times P(X_iX_j = 1) = \frac{1}{n(n-1)}$$

于是

$$\begin{aligned}\mathrm{Var}(X) &= n\,\mathrm{E}(X_i^2) + n(n-1)\,\mathrm{E}(X_iX_j) - n^2\,\mathrm{E}^2(X_i)(j \neq i) \\ &= n \cdot 1/n + n(n-1) \cdot 1/n(n-1) - n^2 \cdot (1/n)^2 = 1\end{aligned}$$

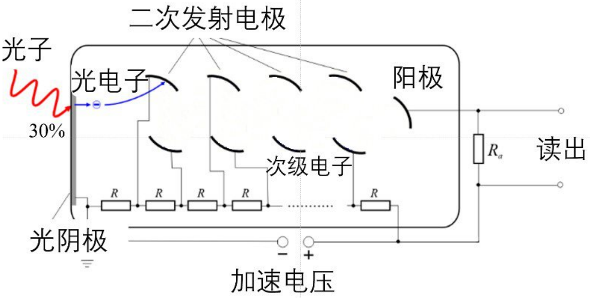
思考：解答发给助教，答对者总评 +2%

求恰好有 k 名战士拿到自己枪的概率

全方差公式

光电倍增管（Photo-Multiplier Tube, PMT）

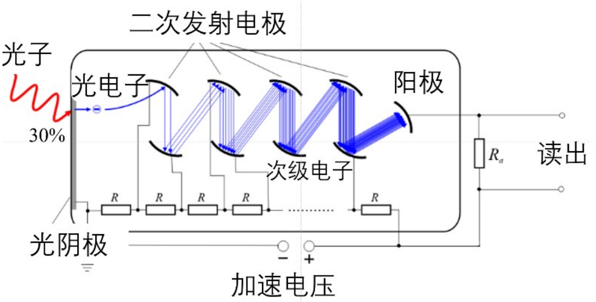
光电倍增管是检测极微弱光的器件，在辐射测量、医学影像等领域应用广泛。



固定地区的天气	固定光强下的 PMT 电荷输出
总降雨量 $S = \sum_{i=1}^{N} R_i$	总电荷量 $Y = \sum_{i=1}^{N} Q_i$
降雨次数 N	光子数 N
各场雨的降雨量 R_i	各个光子诱导的电荷量 Q_i

光电倍增管（Photo-Multiplier Tube, PMT）

光电倍增管是检测极微弱光的器件，在辐射测量、医学影像等领域应用广泛。



固定地区的天气	固定光强下的 PMT 电荷输出
总降雨量 $S = \sum_{i=1}^{N} R_i$	总电荷量 $Y = \sum_{i=1}^{N} Q_i$
降雨次数 N	光子数 N
各场雨的降雨量 R_i	各个光子诱导的电荷量 Q_i

方差如何借助条件概率计算?

总降雨量 S 的方差由降雨次数 N 的方差和每次降雨量 R 的方差共同贡献。

$$\mathrm{Var}(S) =? \mathrm{Var}(S|N)$$

猜想

$$\mathrm{Var}(S) \overset{?}{=} \cancel{\mathrm{Var}}\, \mathrm{Var}(S|N)$$

- $\mathrm{Var}(S)$ 的量纲（mm）是 S 量纲的平方，若求两次方差，则量纲不匹配。

进一步：如何组合 E 和 Var 使得量纲匹配?

方差如何借助条件概率计算?

总降雨量 S 的方差由降雨次数 N 的方差和每次降雨量 R 的方差共同贡献。

$$\mathrm{Var}(S) =? \mathrm{Var}(S|N)$$

猜想

$$\mathrm{Var}(S) \stackrel{?}{=} \cancel{\mathrm{Var}}\,\mathrm{Var}(S|N)$$

- $\mathrm{Var}(S)$ 的量纲（mm）是 S 量纲的平方，若求两次方差，则量纲不匹配。

进一步：如何组合 E 和 Var 使得量纲匹配?

1. $\mathrm{Var}(S) \stackrel{?}{=} \mathrm{E}[\mathrm{Var}(S|N)]$ ，“方差的期望”；
2. $\mathrm{Var}(S) \stackrel{?}{=} \mathrm{Var}[\mathrm{E}(S|N)]$ ，“期望的方差”；

方差如何借助条件概率计算?

总降雨量 S 的方差由降雨次数 N 的方差和每次降雨量 R 的方差共同贡献。

$$\mathrm{Var}(S) =? \mathrm{Var}(S|N)$$

猜想

$$\mathrm{Var}(S) \stackrel{?}{=} \cancel{\mathrm{Var}}\,\mathrm{Var}(S|N)$$

- $\mathrm{Var}(S)$ 的量纲（mm）是 S 量纲的平方，若求两次方差，则量纲不匹配。

进一步：如何组合 E 和 Var 使得量纲匹配?

方差如何借助条件概率计算?

总降雨量 S 的方差由降雨次数 N 的方差和每次降雨量 R 的方差共同贡献。

$$\mathrm{Var}(S) =? \mathrm{Var}(S|N)$$

猜想

$$\mathrm{Var}(S) \stackrel{?}{=} \cancel{\mathrm{Var}}\, \mathrm{Var}(S|N)$$

- $\mathrm{Var}(S)$ 的量纲（mm）是 S 量纲的平方，若求两次方差，则量纲不匹配。

进一步：如何组合 E 和 Var 使得量纲匹配?

1. $\mathrm{Var}(S) \stackrel{?}{=} \mathrm{E}[\mathrm{Var}(S|N)]$，“方差的期望”；
2. $\mathrm{Var}(S) \stackrel{?}{=} \mathrm{Var}[\mathrm{E}(S|N)]$，“期望的方差”；
3. $\mathrm{Var}(S) \stackrel{?}{=} a\,\mathrm{E}[\mathrm{Var}(S|N)] + b\,\mathrm{Var}[\mathrm{E}(S|N)]$，$a, b$ 待定。

方差如何借助条件概率计算?

总降雨量 S 的方差由降雨次数 N 的方差和每次降雨量 R 的方差共同贡献。

$$\mathrm{Var}(S) =? \mathrm{Var}(S|N)$$

猜想

$$\mathrm{Var}(S) \overset{?}{=} \cancel{\mathrm{Var}}\, \mathrm{Var}(S|N)$$

- $\mathrm{Var}(S)$ 的量纲（mm）是 S 量纲的平方，若求两次方差，则量纲不匹配。

进一步：如何组合 E 和 Var 使得量纲匹配?

1. $\mathrm{Var}(S) \overset{?}{=} \mathrm{E}[\mathrm{Var}(S|N)]$，“方差的期望”；
2. $\mathrm{Var}(S) \overset{?}{=} \mathrm{Var}[\mathrm{E}(S|N)]$，“期望的方差”；
3. $\mathrm{Var}(S) \overset{?}{=} a\,\mathrm{E}[\mathrm{Var}(S|N)] + b\,\mathrm{Var}[\mathrm{E}(S|N)]$，$a, b$ 待定。

方差如何借助条件概率计算?

总降雨量 S 的方差由降雨次数 N 的方差和每次降雨量 R 的方差共同贡献。

$$\mathrm{Var}(S) =? \mathrm{Var}(S|N)$$

猜想

$$\mathrm{Var}(S) \stackrel{?}{=} \cancel{\mathrm{Var}}\,\mathrm{Var}(S|N)$$

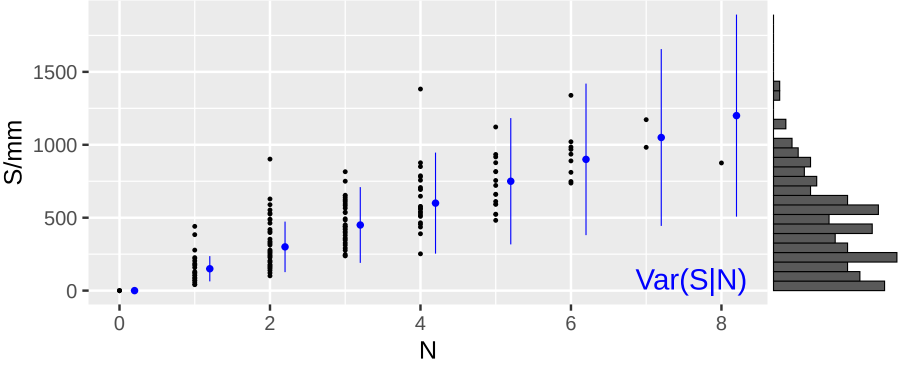
- $\mathrm{Var}(S)$ 的量纲（mm）是 S 量纲的平方，若求两次方差，则量纲不匹配。

进一步：如何组合 E 和 Var 使得量纲匹配?

1. $\mathrm{Var}(S) \stackrel{?}{=} \mathrm{E}[\mathrm{Var}(S|N)]$，“方差的期望”；
2. $\mathrm{Var}(S) \stackrel{?}{=} \mathrm{Var}[\mathrm{E}(S|N)]$，“期望的方差”；
3. $\mathrm{Var}(S) \stackrel{?}{=} a\,\mathrm{E}[\mathrm{Var}(S|N)] + b\,\mathrm{Var}[\mathrm{E}(S|N)]$，$a, b$ 待定。

方差的期望 $\mathrm{E}[\mathrm{Var}(S|N)]$

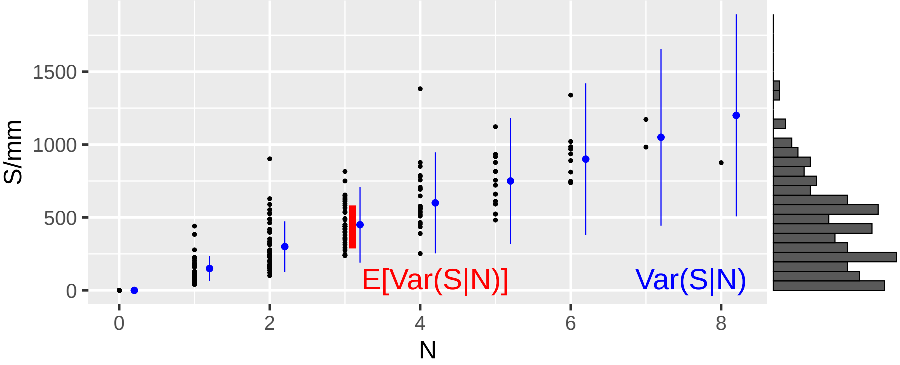
- 每一个 N 都对应一个方差 [1] $\mathrm{Var}(S|N)$
- 它的期望表征了“平均”的方差，刻画“各组之内”的差异。



[1] 图中的“方差”用开平方后的标准差表示，以保证量纲与随机变量相同。

方差的期望 $\mathrm{E}[\mathrm{Var}(S|N)]$

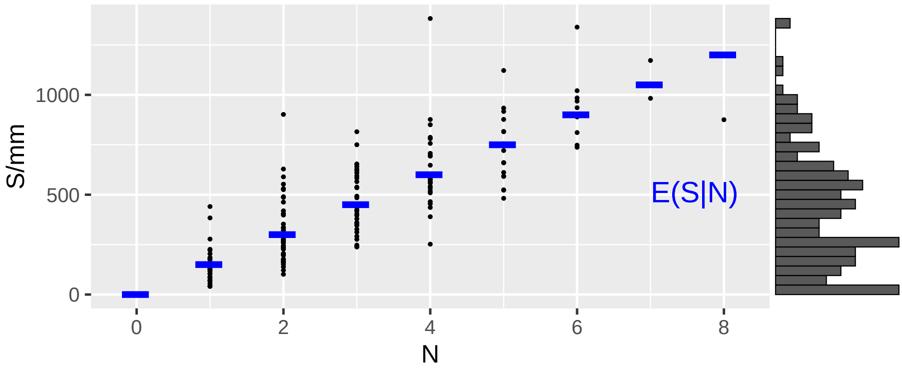
- 每一个 N 都对应一个方差 [1] $\mathrm{Var}(S|N)$
- 它的期望表征了“平均”的方差，刻画“各组之内”的差异。



[1] 图中的“方差”用开平方后的标准差表示，以保证量纲与随机变量相同。

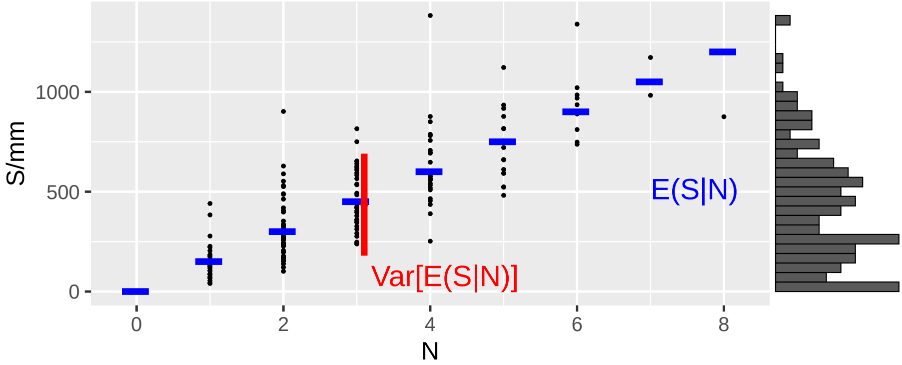
期望的方差 $\mathrm{Var}[\mathrm{E}(S|N)]$

- 每一个 N 对应的期望各不相同。
- 它们之间的差异贡献给了“期望的方差”，刻画“组与组之间”的差异。



期望的方差 $\mathrm{Var}[\mathrm{E}(S|N)]$

- 每一个 N 对应的期望各不相同。
- 它们之间的差异贡献给了“期望的方差”，刻画“组与组之间”的差异。



定量证明

- 方差的期望

$$\mathrm{E}[\mathrm{Var}(S|N)] = \mathrm{E}\left[\mathrm{E}(S^2|N) - [\mathrm{E}(S|N)]^2\right] = \mathrm{E}(S^2) - \mathrm{E}\left\{[\mathrm{E}(S|N)]^2\right\}$$

- 期望的方差

$$\mathrm{Var}[\mathrm{E}(S|N)] = \mathrm{E}\left\{[\mathrm{E}(S|N)]^2\right\} - \left\{\mathrm{E}\left[\mathrm{E}(S|N)\right]\right\}^2 = \mathrm{E}\left\{[\mathrm{E}(S|N)]^2\right\} - [\mathrm{E}(S)]^2$$

- 获得全方差公式 $\mathrm{Var}(S) = \mathrm{E}[\mathrm{Var}(S|N)] + \mathrm{Var}[\mathrm{E}(S|N)]$

定量证明

- 方差的期望

$$\begin{aligned}\mathrm{E}[\mathrm{Var}(S|N)] &= \mathrm{E}\left[\mathrm{E}(S^2|N) - [\mathrm{E}(S|N)]^2\right] \\ &= \mathrm{E}(S^2) - \mathrm{E}\left\{[\mathrm{E}(S|N)]^2\right\}\end{aligned}$$

- 期望的方差

$$\begin{aligned}\mathrm{Var}[\mathrm{E}(S|N)] &= \mathrm{E}\left\{[\mathrm{E}(S|N)]^2\right\} - \left\{\mathrm{E}\left[\mathrm{E}(S|N)\right]\right\}^2 \\ &= \mathrm{E}\left\{[\mathrm{E}(S|N)]^2\right\} - [\mathrm{E}(S)]^2\end{aligned}$$

- 获得全方差公式 $\mathrm{Var}(S) = \mathrm{E}[\mathrm{Var}(S|N)] + \mathrm{Var}[\mathrm{E}(S|N)]$

定量证明

- 方差的期望

$$
\begin{aligned}
\mathrm{E}[\mathrm{Var}(S|N)] &= \mathrm{E}\left[\mathrm{E}(S^2|N) - [\mathrm{E}(S|N)]^2\right] \\
&= \mathrm{E}(S^2) - \mathrm{E}\left\{[\mathrm{E}(S|N)]^2\right\}
\end{aligned}
$$

- 期望的方差

$$
\begin{aligned}
\mathrm{Var}[\mathrm{E}(S|N)] &= \mathrm{E}\left\{[\mathrm{E}(S|N)]^2\right\} - \left\{\mathrm{E}\left[\mathrm{E}(S|N)\right]\right\}^2 \\
&= \mathrm{E}\left\{[\mathrm{E}(S|N)]^2\right\} - [\mathrm{E}(S)]^2
\end{aligned}
$$

- 获得全方差公式 $\mathrm{Var}(S) = \mathrm{E}[\mathrm{Var}(S|N)] + \mathrm{Var}[\mathrm{E}(S|N)]$

定量证明

- 方差的期望

$$\begin{aligned}\mathrm{E}[\mathrm{Var}(S|N)] &= \mathrm{E}\left[\mathrm{E}(S^2|N) - [\mathrm{E}(S|N)]^2\right] \\ &= \mathrm{E}(S^2) - \mathrm{E}\left\{[\mathrm{E}(S|N)]^2\right\}\end{aligned}$$

- 期望的方差

$$\begin{aligned}\mathrm{Var}[\mathrm{E}(S|N)] &= \mathrm{E}\left\{[\mathrm{E}(S|N)]^2\right\} - \left\{\mathrm{E}\left[\mathrm{E}(S|N)\right]\right\}^2 \\ &= \mathrm{E}\left\{[\mathrm{E}(S|N)]^2\right\} - [\mathrm{E}(S)]^2\end{aligned}$$

- 获得全方差公式 $\mathrm{Var}(S) = \mathrm{E}[\mathrm{Var}(S|N)] + \mathrm{Var}[\mathrm{E}(S|N)]$

定量证明

- 方差的期望

$$
\begin{aligned}
\mathrm{E}[\mathrm{Var}(S|N)] &= \mathrm{E}\left[\mathrm{E}(S^2|N) - [\mathrm{E}(S|N)]^2\right] \\
&= \mathrm{E}(S^2) - \mathrm{E}\left\{[\mathrm{E}(S|N)]^2\right\}
\end{aligned}
$$

- 期望的方差

$$
\begin{aligned}
\mathrm{Var}[\mathrm{E}(S|N)] &= \mathrm{E}\left\{[\mathrm{E}(S|N)]^2\right\} - \{\mathrm{E}\left[\mathrm{E}(S|N)\right]\}^2 \\
&= \mathrm{E}\left\{[\mathrm{E}(S|N)]^2\right\} - [\mathrm{E}(S)]^2
\end{aligned}
$$

- 获得全方差公式 $\mathrm{Var}(S) = \mathrm{E}[\mathrm{Var}(S|N)] + \mathrm{Var}[\mathrm{E}(S|N)]$

定量证明

- 方差的期望

$$\begin{aligned}\mathrm{E}[\mathrm{Var}(S|N)] &= \mathrm{E}\left[\mathrm{E}(S^2|N) - [\mathrm{E}(S|N)]^2\right] \\ &= \mathrm{E}(S^2) - \mathrm{E}\left\{[\mathrm{E}(S|N)]^2\right\}\end{aligned}$$

- 期望的方差

$$\begin{aligned}\mathrm{Var}[\mathrm{E}(S|N)] &= \mathrm{E}\left\{[\mathrm{E}(S|N)]^2\right\} - \{\mathrm{E}\left[\mathrm{E}(S|N)\right]\}^2 \\ &= \mathrm{E}\left\{[\mathrm{E}(S|N)]^2\right\} - [\mathrm{E}(S)]^2\end{aligned}$$

- 获得全方差公式 $\mathrm{Var}(S) = \mathrm{E}[\mathrm{Var}(S|N)] + \mathrm{Var}[\mathrm{E}(S|N)]$

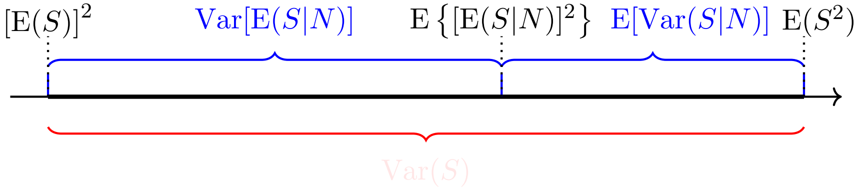
定量证明

- 方差的期望

$$
\begin{aligned}
\mathrm{E}[\mathrm{Var}(S|N)] &= \mathrm{E}\left[\mathrm{E}(S^2|N) - [\mathrm{E}(S|N)]^2\right] \\
&= \mathrm{E}(S^2) - \mathrm{E}\left\{[\mathrm{E}(S|N)]^2\right\}
\end{aligned}
$$

- 期望的方差

$$
\begin{aligned}
\mathrm{Var}[\mathrm{E}(S|N)] &= \mathrm{E}\left\{[\mathrm{E}(S|N)]^2\right\} - \{\mathrm{E}\left[\mathrm{E}(S|N)\right]\}^2 \\
&= \mathrm{E}\left\{[\mathrm{E}(S|N)]^2\right\} - [\mathrm{E}(S)]^2
\end{aligned}
$$



- 获得全方差公式 $\mathrm{Var}(S) = \mathrm{E}[\mathrm{Var}(S|N)] + \mathrm{Var}[\mathrm{E}(S|N)]$

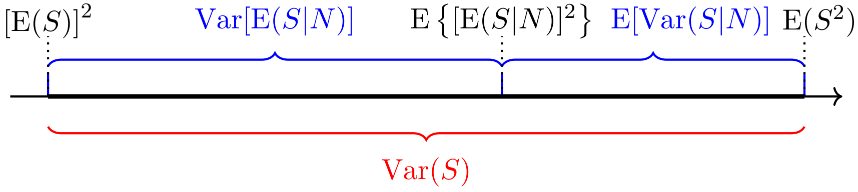
定量证明

- 方差的期望

$$
\begin{aligned}
\mathrm{E}[\mathrm{Var}(S|N)] &= \mathrm{E}\left[\mathrm{E}(S^2|N) - [\mathrm{E}(S|N)]^2\right] \\
&= \mathrm{E}(S^2) - \mathrm{E}\left\{[\mathrm{E}(S|N)]^2\right\}
\end{aligned}
$$

- 期望的方差

$$
\begin{aligned}
\mathrm{Var}[\mathrm{E}(S|N)] &= \mathrm{E}\left\{[\mathrm{E}(S|N)]^2\right\} - \{\mathrm{E}\left[\mathrm{E}(S|N)\right]\}^2 \\
&= \mathrm{E}\left\{[\mathrm{E}(S|N)]^2\right\} - [\mathrm{E}(S)]^2
\end{aligned}
$$



- 获得全方差公式 $\mathrm{Var}(S) = \mathrm{E}[\mathrm{Var}(S|N)] + \mathrm{Var}[\mathrm{E}(S|N)]$

全方差公式的意义

$$\mathrm{Var}(S) = \underbrace{\mathrm{E}[\mathrm{Var}(S|N)]}_{N\text{固定时}S\text{的方差}} + \underbrace{\mathrm{Var}[\mathrm{E}(S|N)]}_{\text{由}N\text{诱导的}S\text{方差}}$$

- 把 S （降雨量）的方差，分解为是否由 N （降雨次数）导致的两部分。
- 体现了 S 对 N 的相关关系。

高斯 Theoria Combinationis 1821, 1823

- 高斯曾使用“最小二乘法”，在 1801 年预测了谷神星（Ceres）的存在。但理论基础遭到质疑。
- 高斯在不对 N 和 S 的分布不做正态假设的前提下，证明了全方差公式；
- 发明了“最小二乘法”解线性回归问题 [2]，并奠定了严格的理论基础。

[2] Legendre 在同时期发明

全方差公式的意义

$$\mathrm{Var}(S) = \underbrace{\mathrm{E}[\mathrm{Var}(S|N)]}_{N\text{固定时}S\text{的方差}} + \underbrace{\mathrm{Var}[\mathrm{E}(S|N)]}_{\text{由}N\text{诱导的}S\text{方差}}$$

- 把 S （降雨量）的方差，分解为是否由 N （降雨次数）导致的两部分。
- 体现了 S 对 N 的相关关系。

高斯 Theoria Combinationis 1821, 1823

- 高斯曾使用“最小二乘法”，在 1801 年预测了谷神星（Ceres）的存在。但理论基础遭到质疑。
- 高斯在不对 N 和 S 的分布不做正态假设的前提下，证明了全方差公式；
- 发明了“最小二乘法”解线性回归问题 [2]，并奠定了严格的理论基础。

[2] Legendre 在同时期发明

全方差公式的意义

$$\mathrm{Var}(S) = \underbrace{\mathrm{E}[\mathrm{Var}(S|N)]}_{N\text{固定时}S\text{的方差}} + \underbrace{\mathrm{Var}[\mathrm{E}(S|N)]}_{\text{由}N\text{诱导的}S\text{方差}}$$

- 把 S （降雨量）的方差，分解为是否由 N （降雨次数）导致的两部分。
- 体现了 S 对 N 的相关关系。

高斯 Theoria Combinationis 1821, 1823

- 高斯曾使用“最小二乘法”，在 1801 年预测了谷神星（Ceres）的存在。但理论基础遭到质疑。
- 高斯在不对 N 和 S 的分布不做正态假设的前提下，证明了全方差公式；
- 发明了“最小二乘法”解线性回归问题 [2]，并奠定了严格的理论基础。

[2] Legendre 在同时期发明

全方差公式的意义

$$\mathrm{Var}(S) = \underbrace{\mathrm{E}[\mathrm{Var}(S|N)]}_{N\text{固定时}S\text{的方差}} + \underbrace{\mathrm{Var}[\mathrm{E}(S|N)]}_{\text{由}N\text{诱导的}S\text{方差}}$$

- 把 S （降雨量）的方差，分解为是否由 N （降雨次数）导致的两部分。
- 体现了 S 对 N 的相关关系。

高斯 Theoria Combinationis 1821, 1823

- 高斯曾使用“最小二乘法”，在 1801 年预测了谷神星（Ceres）的存在。但理论基础遭到质疑。
- 高斯在不对 N 和 S 的分布不做正态假设的前提下，证明了全方差公式；
- 发明了“最小二乘法”解线性回归问题 [2]，并奠定了严格的理论基础。

[2] Legendre 在同时期发明

例：光电倍增管总电荷二刷

例

$Y = \sum_{i=1}^{N} Q_i$ 求 $\mathrm{Var}(Y)$

- 总电荷 Y 的方差，关系到探测精度。

$$
\begin{aligned}
\mathrm{Var}(Y) &= \mathrm{E}[\mathrm{Var}(Y|N)] + \mathrm{Var}[\mathrm{E}(Y|N)] \\
&= \mathrm{E}[N\,\mathrm{Var}(Q)] + \mathrm{Var}[N\,\mathrm{E}(Q)] \\
&= \mathrm{E}(N)\,\mathrm{Var}(Q) + \mathrm{Var}(N)[\mathrm{E}(Q)]^2
\end{aligned}
$$

启示

- 把难以计算的 Y 方差，分解为 N 与 Q 的方差，简化问题。

例：光电倍增管总电荷二刷

例

$Y = \sum_{i=1}^{N} Q_i$ 求 $\mathrm{Var}(Y)$

- 总电荷 Y 的方差，关系到探测精度。

$$
\begin{aligned}
\mathrm{Var}(Y) &= \mathrm{E}[\mathrm{Var}(Y|N)] + \mathrm{Var}[\mathrm{E}(Y|N)] \\
&= \mathrm{E}[N\,\mathrm{Var}(Q)] + \mathrm{Var}[N\,\mathrm{E}(Q)] \\
&= \mathrm{E}(N)\,\mathrm{Var}(Q) + \mathrm{Var}(N)[\mathrm{E}(Q)]^2
\end{aligned}
$$

启示

- 把难以计算的 Y 方差，分解为 N 与 Q 的方差，简化问题。

例：光电倍增管总电荷二刷

例

$Y = \sum_{i=1}^{N} Q_i$ 求 $\mathrm{Var}(Y)$

- 总电荷 Y 的方差，关系到探测精度。

$$
\begin{aligned}
\mathrm{Var}(Y) &= \mathrm{E}[\mathrm{Var}(Y|N)] + \mathrm{Var}[\mathrm{E}(Y|N)] \\
&= \mathrm{E}[N\,\mathrm{Var}(Q)] + \mathrm{Var}[N\,\mathrm{E}(Q)] \\
&= \mathrm{E}(N)\,\mathrm{Var}(Q) + \mathrm{Var}(N)[\mathrm{E}(Q)]^2
\end{aligned}
$$

启示

- 把难以计算的 Y 方差，分解为 N 与 Q 的方差，简化问题。

例：光电倍增管总电荷二刷

例

$Y = \sum_{i=1}^{N} Q_i$ 求 $\mathrm{Var}(Y)$

- 总电荷 Y 的方差，关系到探测精度。

$$
\begin{aligned}
\mathrm{Var}(Y) &= \mathrm{E}[\mathrm{Var}(Y|N)] + \mathrm{Var}[\mathrm{E}(Y|N)] \\
&= \mathrm{E}[N\,\mathrm{Var}(Q)] + \mathrm{Var}[N\,\mathrm{E}(Q)] \\
&= \mathrm{E}(N)\,\mathrm{Var}(Q) + \mathrm{Var}(N)[\mathrm{E}(Q)]^2
\end{aligned}
$$

启示

- 把难以计算的 Y 方差，分解为 N 与 Q 的方差，简化问题。

额外噪声因子

额外噪声因子：excess noise factor 衡量了器件在泊松涨落之上的额外涨落。
对 $N \sim \pi(\lambda)$ 有 $\mathrm{E}(N) = \mathrm{Var}(N) = \lambda$，

$$\mathrm{Var}(Y) = \mathrm{E}(N)\,\mathrm{Var}(Q) + \mathrm{Var}(N)[\mathrm{E}(Q)]^2$$

$$\Longrightarrow \frac{\mathrm{Var}(Y)}{[\mathrm{E}(Q)]^2} = \mathrm{Var}(N)\left\{\underbrace{1 + \frac{\mathrm{Var}(Q)}{[\mathrm{E}(Q)]^2}}_{\text{ENF}}\right\}.$$

- 全方差公式是辐射与光电探测领域核心概念“额外噪声因子”的理论基础。

Journal of Instrumentation, 17 P06040

协方差

引子

- 二维随机变量 (X, Y)，已知联合分布可得边缘分布，但反之不可。
- 二维随机变量，除每个随机变量各自的边缘分布外，相互之间可能还有某种联系。

考虑两个随机变量 X, Y 的和或差的方差：

$$\mathrm{Var}(X \pm Y) = \mathrm{Var}(X) + \mathrm{Var}(Y) \pm 2\,\mathrm{E}[(X - \mathrm{E}(X))(Y - \mathrm{E}(Y))]$$

$\mathrm{E}([X - \mathrm{E}(X)][Y - \mathrm{E}(Y)])$ 反映了随机变量 X, Y 之间的某种关系。

引子

- 二维随机变量 (X, Y)，已知联合分布可得边缘分布，但反之不可。
- 二维随机变量，除每个随机变量各自的边缘分布外，相互之间可能还有某种联系。

考虑两个随机变量 X, Y 的和或差的方差：

$$\mathrm{Var}(X \pm Y) = \mathrm{Var}(X) + \mathrm{Var}(Y) \pm 2\,\mathrm{E}[(X-\mathrm{E}(X))(Y-\mathrm{E}(Y))]$$

$\mathrm{E}([X-\mathrm{E}(X)][Y-\mathrm{E}(Y)])$ 反映了随机变量 X, Y 之间的某种关系。

定义

称 $\mathrm{E}([X-\mathrm{E}(X)][Y-\mathrm{E}(Y)])$ 为随机变量 X, Y 的 协方差 。记为

$$\mathrm{Cov}(X,Y) = \mathrm{E}([X-\mathrm{E}(X)][Y-\mathrm{E}(Y)]) = \mathrm{E}(XY) - \mathrm{E}(X)\,\mathrm{E}(Y)$$

若 $\mathrm{Var}(X) > 0, \mathrm{Var}(Y) > 0$ ，称

$$\mathrm{E}\left(\frac{[X-\mathrm{E}(X)][Y-\mathrm{E}(Y)]}{\sqrt{\mathrm{Var}(X)}\sqrt{\mathrm{Var}(Y)}}\right) = \frac{\mathrm{Cov}(X,Y)}{\sqrt{\mathrm{Var}(X)}\sqrt{\mathrm{Var}(Y)}}$$

为随机变量 X, Y 的 相关系数 ，记为

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

若 $\rho_{XY} = 0$，称随机变量 X, Y 不相关 。

定义

称 $\mathrm{E}([X-\mathrm{E}(X)][Y-\mathrm{E}(Y)])$ 为随机变量 X,Y 的 协方差 。记为

$$\mathrm{Cov}(X,Y)=\mathrm{E}([X-\mathrm{E}(X)][Y-\mathrm{E}(Y)])=\mathrm{E}(XY)-\mathrm{E}(X)\,\mathrm{E}(Y)$$

若 $\mathrm{Var}(X)>0,\mathrm{Var}(Y)>0$ ，称

$$\mathrm{E}\left(\frac{[X-\mathrm{E}(X)][Y-\mathrm{E}(Y)]}{\sqrt{\mathrm{Var}(X)}\sqrt{\mathrm{Var}(Y)}}\right)=\frac{\mathrm{Cov}(X,Y)}{\sqrt{\mathrm{Var}(X)}\sqrt{\mathrm{Var}(Y)}}$$

为随机变量 X,Y 的 相关系数 ，记为

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

若 $\rho_{XY}=0$，称随机变量 X,Y 不相关 。

定义

称 $\mathrm{E}([X-\mathrm{E}(X)][Y-\mathrm{E}(Y)])$ 为随机变量 X,Y 的 协方差 。记为

$$\mathrm{Cov}(X,Y)=\mathrm{E}([X-\mathrm{E}(X)][Y-\mathrm{E}(Y)])=\mathrm{E}(XY)-\mathrm{E}(X)\,\mathrm{E}(Y)$$

若 $\mathrm{Var}(X)>0,\mathrm{Var}(Y)>0$ ，称

$$\mathrm{E}\left(\frac{[X-\mathrm{E}(X)][Y-\mathrm{E}(Y)]}{\sqrt{\mathrm{Var}(X)}\sqrt{\mathrm{Var}(Y)}}\right)=\frac{\mathrm{Cov}(X,Y)}{\sqrt{\mathrm{Var}(X)}\sqrt{\mathrm{Var}(Y)}}$$

为随机变量 X,Y 的 相关系数 ，记为

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

若 $\rho_{XY}=0$，称随机变量 X,Y 不相关 。

性质

计算技巧

$$\mathrm{Cov}(X,Y) = \mathrm{E}(XY) - \mathrm{E}(X)\,\mathrm{E}(Y)$$
$$= \pm\frac{1}{2}\left[\mathrm{Var}(X \pm Y) - \mathrm{Var}(X) - \mathrm{Var}(Y)\right]$$

对称性 $\mathrm{Cov}(X,Y) = \mathrm{Cov}(Y,X)$

双线性 $\mathrm{Cov}(aX,bY) = ab\,\mathrm{Cov}(X,Y)$

- $\mathrm{Cov}(X+Y,Z) = \mathrm{Cov}(X,Z) + \mathrm{Cov}(Y,Z)$

方差特例 $\mathrm{Cov}(X,X) = \mathrm{Var}(X)$

性质

计算技巧

$$\begin{aligned}\mathrm{Cov}(X,Y) &= \mathrm{E}(XY) - \mathrm{E}(X)\,\mathrm{E}(Y) \\ &= \pm\frac{1}{2}\left[\mathrm{Var}(X \pm Y) - \mathrm{Var}(X) - \mathrm{Var}(Y)\right]\end{aligned}$$

对称性 $\mathrm{Cov}(X,Y) = \mathrm{Cov}(Y,X)$

双线性 $\mathrm{Cov}(aX,bY) = ab\,\mathrm{Cov}(X,Y)$

- $\mathrm{Cov}(X+Y,Z) = \mathrm{Cov}(X,Z) + \mathrm{Cov}(Y,Z)$

方差特例 $\mathrm{Cov}(X,X) = \mathrm{Var}(X)$

性质

计算技巧

$$\begin{aligned}\operatorname{Cov}(X,Y) &= \operatorname{E}(XY) - \operatorname{E}(X)\operatorname{E}(Y)\\ &= \pm\frac{1}{2}\left[\operatorname{Var}(X\pm Y) - \operatorname{Var}(X) - \operatorname{Var}(Y)\right]\end{aligned}$$

对称性 $\operatorname{Cov}(X,Y) = \operatorname{Cov}(Y,X)$

双线性 $\operatorname{Cov}(aX,bY) = ab\operatorname{Cov}(X,Y)$

- $\operatorname{Cov}(X+Y,Z) = \operatorname{Cov}(X,Z) + \operatorname{Cov}(Y,Z)$

方差特例 $\operatorname{Cov}(X,X) = \operatorname{Var}(X)$

标准化随机变量

设随机变量 X 的期望 $\mathrm{E}(X)$、方差 $\mathrm{Var}(X)$ 都存在，且 $\mathrm{Var}(X) \neq 0$，则称

$$X^* = \frac{X - \mathrm{E}(X)}{\sqrt{\mathrm{Var}(X)}}$$

为 X 的标准化随机变量.

$$\mathrm{E}(X^*) = 0, \mathrm{Var}(X^*) = 1$$

- 相关系数的解读

$$\mathrm{Cov}(X^*, Y^*) = \rho_{XY}$$

标准化随机变量

设随机变量 X 的期望 $\mathrm{E}(X)$、方差 $\mathrm{Var}(X)$ 都存在，且 $\mathrm{Var}(X) \neq 0$，则称

$$X^* = \frac{X - \mathrm{E}(X)}{\sqrt{\mathrm{Var}(X)}}$$

为 X 的标准化随机变量.

$$\mathrm{E}(X^*) = 0, \mathrm{Var}(X^*) = 1$$

- 相关系数的解读

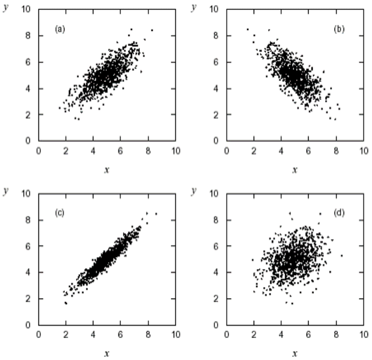
$$\mathrm{Cov}(X^*, Y^*) = \rho_{XY}$$

相关系数的性质

- $|\rho_{XY}| \le 1$
- $|\rho_{XY}| = 1 \iff$ 存在常数 a, b，使 $P(Y = aX + b) = 1$
  即 X 与 Y 以概率 1 线性相关
- 若 X, Y 相互独立，则 $\rho_{XY} = 0$，X, Y 不相关。反之不一定成立。

$\rho = 0.75$ $\rho = -0.75$

$\rho = 0.95$ $\rho = 0.25$

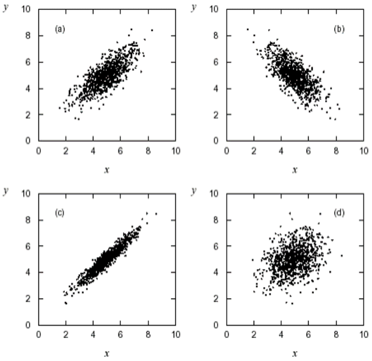


相关系数的性质

- $|\rho_{XY}| \le 1$
- $|\rho_{XY}| = 1 \iff$ 存在常数 a, b，使 $P(Y = aX + b) = 1$
  即 X 与 Y 以概率 1 线性相关
- 若 X, Y 相互独立，则 $\rho_{XY} = 0$，X, Y 不相关。反之不一定成立。

$\rho = 0.75$ $\rho = -0.75$

$\rho = 0.95$ $\rho = 0.25$

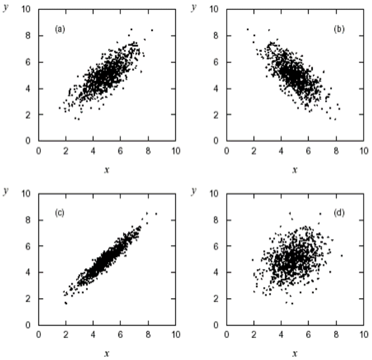


相关系数的性质

- $|\rho_{XY}| \leq 1$
- $|\rho_{XY}| = 1 \iff$ 存在常数 a, b，使 $P(Y = aX + b) = 1$
  即 X 与 Y 以概率 1 线性相关
- 若 X, Y 相互独立，则 $\rho_{XY} = 0$，X, Y 不相关。反之不一定成立。

$\rho = 0.75$ $\rho = -0.75$

$\rho = 0.95$ $\rho = 0.25$



例：柯西-施瓦兹不等式

例

证明：柯西-施瓦兹不等式 $[\mathrm{E}(XY)]^2 \le \mathrm{E}(X^2)\,\mathrm{E}(Y^2)$

设实变量 t 的二次非负函数

$$g(t) = \mathrm{E}[(X+tY)^2] = \mathrm{E}(X^2) + 2t\,\mathrm{E}(XY) + t^2\,\mathrm{E}(Y^2) \ge 0$$

从而由二次方程 $g(t)=0$ 根的判别式得到

$$\Delta = 4[\mathrm{E}(XY)]^2 - 4\,\mathrm{E}(X^2)\,\mathrm{E}(Y^2) \le 0 \implies [\mathrm{E}(XY)]^2 \le \mathrm{E}(X^2)\,\mathrm{E}(Y^2)$$

相关系数

将标准化的 X^*, Y^* 代入，得

$$|\rho_{XY}| = |\mathrm{E}(X^*Y^*)| \le \sqrt{\mathrm{E}(X^{*2})\,\mathrm{E}(Y^{*2})} = 1$$

例：柯西-施瓦兹不等式

例

证明：柯西-施瓦兹不等式 $[\mathrm{E}(XY)]^2 \le \mathrm{E}(X^2)\,\mathrm{E}(Y^2)$

设实变量 t 的二次非负函数

$$g(t) = \mathrm{E}[(X+tY)^2] = \mathrm{E}(X^2) + 2t\,\mathrm{E}(XY) + t^2\,\mathrm{E}(Y^2) \ge 0$$

从而由二次方程 $g(t)=0$ 根的判别式得到

$$\Delta = 4[\mathrm{E}(XY)]^2 - 4\,\mathrm{E}(X^2)\,\mathrm{E}(Y^2) \le 0 \implies [\mathrm{E}(XY)]^2 \le \mathrm{E}(X^2)\,\mathrm{E}(Y^2)$$

相关系数

将标准化的 X^*, Y^* 代入，得

$$|\rho_{XY}| = |\mathrm{E}(X^*Y^*)| \le \sqrt{\mathrm{E}(X^{*2})\,\mathrm{E}(Y^{*2})} = 1$$

例：柯西-施瓦兹不等式

例

证明：柯西-施瓦兹不等式 $[\mathrm{E}(XY)]^2 \leq \mathrm{E}(X^2)\,\mathrm{E}(Y^2)$

设实变量 t 的二次非负函数

$$g(t) = \mathrm{E}[(X+tY)^2] = \mathrm{E}(X^2) + 2t\,\mathrm{E}(XY) + t^2\,\mathrm{E}(Y^2) \geq 0$$

从而由二次方程 $g(t)=0$ 根的判别式得到

$$\Delta = 4[\mathrm{E}(XY)]^2 - 4\,\mathrm{E}(X^2)\,\mathrm{E}(Y^2) \leq 0 \implies [\mathrm{E}(XY)]^2 \leq \mathrm{E}(X^2)\,\mathrm{E}(Y^2)$$

相关系数

将标准化的 X^*, Y^* 代入，得

$$|\rho_{XY}| = |\mathrm{E}(X^*Y^*)| \leq \sqrt{\mathrm{E}(X^{*2})\,\mathrm{E}(Y^{*2})} = 1$$

例：柯西-施瓦兹不等式

例

证明：柯西-施瓦兹不等式 $[\mathrm{E}(XY)]^2 \le \mathrm{E}(X^2)\,\mathrm{E}(Y^2)$

设实变量 t 的二次非负函数

$$g(t) = \mathrm{E}[(X+tY)^2] = \mathrm{E}(X^2) + 2t\,\mathrm{E}(XY) + t^2\,\mathrm{E}(Y^2) \ge 0$$

从而由二次方程 $g(t)=0$ 根的判别式得到

$$\Delta = 4[\mathrm{E}(XY)]^2 - 4\,\mathrm{E}(X^2)\,\mathrm{E}(Y^2) \le 0 \implies [\mathrm{E}(XY)]^2 \le \mathrm{E}(X^2)\,\mathrm{E}(Y^2)$$

相关系数

将标准化的 X^*, Y^* 代入，得

$$|\rho_{XY}| = |\mathrm{E}(X^*Y^*)| \le \sqrt{\mathrm{E}(X^{*2})\,\mathrm{E}(Y^{*2})} = 1$$

例：柯西-施瓦兹不等式

例

证明：柯西-施瓦兹不等式 $[\mathrm{E}(XY)]^2 \le \mathrm{E}(X^2)\,\mathrm{E}(Y^2)$

设实变量 t 的二次非负函数

$$g(t) = \mathrm{E}[(X+tY)^2] = \mathrm{E}(X^2) + 2t\,\mathrm{E}(XY) + t^2\,\mathrm{E}(Y^2) \ge 0$$

从而由二次方程 $g(t)=0$ 根的判别式得到

$$\Delta = 4[\mathrm{E}(XY)]^2 - 4\,\mathrm{E}(X^2)\,\mathrm{E}(Y^2) \le 0 \implies [\mathrm{E}(XY)]^2 \le \mathrm{E}(X^2)\,\mathrm{E}(Y^2)$$

相关系数

将标准化的 X^*, Y^* 代入，得

$$|\rho_{XY}| = |\,\mathrm{E}(X^*Y^*)| \le \sqrt{\mathrm{E}(X^{*2})\,\mathrm{E}(Y^{*2})} = 1$$

协方差矩阵

定义

设 n 维随机变量 $\vec{X}=(X_1,X_2,\cdots,X_n)$ 每个分量的方差存在，任意两个分量的协方差存在，则称

$$\begin{bmatrix} \mathrm{Var}(X_1) & \mathrm{Cov}(X_1,X_2) & \cdots & \mathrm{Cov}(X_1,X_n) \\ \mathrm{Cov}(X_2,X_1) & \mathrm{Var}(X_2) & \cdots & \mathrm{Cov}(X_2,X_n) \\ \vdots & \vdots & \ddots & \vdots \\ \mathrm{Cov}(X_n,X_1) & \mathrm{Cov}(X_n,X_2) & \cdots & \mathrm{Var}(X_n) \end{bmatrix}$$

$\vec{X}$ 的 **协方差矩阵** ， 也称为 **方差-协方差** 矩阵，记为 $\mathrm{Var}(\vec{X})$。若记 $\vec{X}-\mathrm{E}(\vec{X})$ 为 n 维列向量，则

$$\mathrm{Var}(\vec{X})=\mathrm{E}\{[\vec{X}-\mathrm{E}(\vec{X})][\vec{X}-\mathrm{E}(\vec{X})]^{\mathrm{T}}\}$$

协方差矩阵 对称非负定。

定义

设 n 维随机变量 $\vec{X} = (X_1, X_2, \cdots, X_n)$ 每个分量的方差存在，任意两个分量的协方差存在，则称

$$\begin{bmatrix} \mathrm{Var}(X_1) & \mathrm{Cov}(X_1, X_2) & \cdots & \mathrm{Cov}(X_1, X_n) \\ \mathrm{Cov}(X_2, X_1) & \mathrm{Var}(X_2) & \cdots & \mathrm{Cov}(X_2, X_n) \\ \vdots & \vdots & \ddots & \vdots \\ \mathrm{Cov}(X_n, X_1) & \mathrm{Cov}(X_n, X_2) & \cdots & \mathrm{Var}(X_n) \end{bmatrix}$$

$\vec{X}$ 的 协方差矩阵 ， 也称为 方差-协方差 矩阵，记为 $\mathrm{Var}(\vec{X})$。若记 $\vec{X} - \mathrm{E}(\vec{X})$ 为 n 维列向量，则

$$\mathrm{Var}(\vec{X}) = \mathrm{E}\{[\vec{X} - \mathrm{E}(\vec{X})][\vec{X} - \mathrm{E}(\vec{X})]^\intercal\}$$

协方差矩阵 对称非负定。

相关系数矩阵

记 $\rho_{ij} = \frac{\mathrm{Cov}(X_i,X_j)}{\sqrt{\mathrm{Var}(X_i)}\sqrt{\mathrm{Var}(X_j)}}, i,j = 1,2,\cdots,n,$

$$\begin{bmatrix} \rho_{11} & \rho_{12} & \cdots & \rho_{1n} \\ \rho_{21} & \rho_{22} & \cdots & \rho_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \rho_{n1} & \rho_{n2} & \cdots & \rho_{nn} \end{bmatrix}$$

称为 相关系数矩阵 ，简称 相关矩阵 ，为对称非负定矩阵。

二维正态分布

两个正态随机变量的联合分布

$X \sim N(\mu_1, \sigma_1^2), Y \sim N(\mu_2, \sigma_2^2)$ 相互独立

$$\Longrightarrow f_X(x) = \frac{1}{\sqrt{2\pi}\sigma_1} \exp\left[-\frac{(x-\mu_1)^2}{2\sigma_1^2}\right]$$

$$f_Y(y) = \frac{1}{\sqrt{2\pi}\sigma_2} \exp\left[-\frac{(y-\mu_2)^2}{2\sigma_2^2}\right]$$

$$\Longrightarrow f(x,y) = f_X(x) f_Y(y) = \frac{1}{2\pi\sigma_1\sigma_2} \exp\left[-\frac{(x-\mu_1)^2}{2\sigma_1^2} - \frac{(y-\mu_2)^2}{2\sigma_2^2}\right]$$

定义（二维正态分布）

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

具有以上概率密度函数的分布称为二维正态分布。

两个正态随机变量的联合分布

$X \sim N(\mu_1, \sigma_1^2), Y \sim N(\mu_2, \sigma_2^2)$ 相互独立

$$\Longrightarrow f_X(x) = \frac{1}{\sqrt{2\pi}\sigma_1} \exp\left[-\frac{(x-\mu_1)^2}{2\sigma_1^2}\right]$$

$$f_Y(y) = \frac{1}{\sqrt{2\pi}\sigma_2} \exp\left[-\frac{(y-\mu_2)^2}{2\sigma_2^2}\right]$$

$$\Longrightarrow f(x,y) = f_X(x)f_Y(y) = \frac{1}{2\pi\sigma_1\sigma_2} \exp\left[-\frac{(x-\mu_1)^2}{2\sigma_1^2} - \frac{(y-\mu_2)^2}{2\sigma_2^2}\right]$$

定义 (二维正态分布)

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

具有以上概率密度函数的分布称为二维正态分布。

两个正态随机变量的联合分布

$X \sim N(\mu_1, \sigma_1^2), Y \sim N(\mu_2, \sigma_2^2)$ 相互独立

$$\Longrightarrow f_X(x) = \frac{1}{\sqrt{2\pi}\sigma_1} \exp\left[-\frac{(x-\mu_1)^2}{2\sigma_1^2}\right]$$

$$f_Y(y) = \frac{1}{\sqrt{2\pi}\sigma_2} \exp\left[-\frac{(y-\mu_2)^2}{2\sigma_2^2}\right]$$

$$\Longrightarrow f(x,y) = f_X(x) f_Y(y) = \frac{1}{2\pi\sigma_1\sigma_2} \exp\left[-\frac{(x-\mu_1)^2}{2\sigma_1^2} - \frac{(y-\mu_2)^2}{2\sigma_2^2}\right]$$

定义 (二维正态分布)

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

具有以上概率密度函数的分布称为二维正态分布。

两个正态随机变量的联合分布

$X \sim N(\mu_1, \sigma_1^2), Y \sim N(\mu_2, \sigma_2^2)$ 相互独立

$$\Longrightarrow f_X(x) = \frac{1}{\sqrt{2\pi}\sigma_1} \exp\left[-\frac{(x-\mu_1)^2}{2\sigma_1^2}\right]$$

$$f_Y(y) = \frac{1}{\sqrt{2\pi}\sigma_2} \exp\left[-\frac{(y-\mu_2)^2}{2\sigma_2^2}\right]$$

$$\Longrightarrow f(x,y) = f_X(x)f_Y(y) = \frac{1}{2\pi\sigma_1\sigma_2} \exp\left[-\frac{(x-\mu_1)^2}{2\sigma_1^2} - \frac{(y-\mu_2)^2}{2\sigma_2^2}\right]$$

定义 (二维正态分布)

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

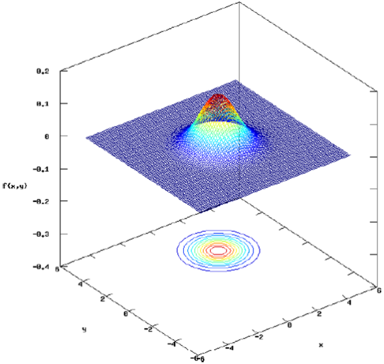
具有以上概率密度函数的分布称为二维正态分布。

$\rho = 0$ 的二维正态分布

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

X 与 Y 相互独立等价于 $\rho = 0$。如果 $\rho = 0$

$$\begin{aligned} f(x, y) &= \frac{1}{2\pi\sigma_1\sigma_2} e^{-\frac{1}{2}\left[\frac{(x-\mu_1)^2}{\sigma_1^2} + \frac{(y-\mu_2)^2}{\sigma_2^2}\right]} \\ &= \frac{1}{\sqrt{2\pi}\sigma_1} e^{-\frac{1}{2}\frac{(x-\mu_1)^2}{\sigma_1^2}} \frac{1}{\sqrt{2\pi}\sigma_2} e^{-\frac{1}{2}\frac{(y-\mu_2)^2}{\sigma_2^2}} \\ &= f_X(x) f_Y(y) \end{aligned}$$



反之，如果 $f(x, y) = f_X(x) f_Y(y)$，取 $x = \mu_1, y = \mu_2$ 则

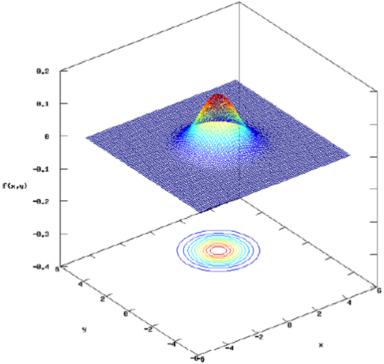
$$\frac{1}{\sqrt{1 - \rho^2}} = 1 \implies \rho = 0$$

$\rho = 0$ 的二维正态分布

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

X 与 Y 相互独立等价于 $\rho = 0$。如果 $\rho = 0$

$$
\begin{aligned}
f(x, y) =& \frac{1}{2\pi\sigma_1\sigma_2} e^{-\frac{1}{2}\left[\frac{(x-\mu_1)^2}{\sigma_1^2} + \frac{(y-\mu_2)^2}{\sigma_2^2}\right]} \\
=& \frac{1}{\sqrt{2\pi}\sigma_1} e^{-\frac{1}{2}\frac{(x-\mu_1)^2}{\sigma_1^2}} \frac{1}{\sqrt{2\pi}\sigma_2} e^{-\frac{1}{2}\frac{(y-\mu_2)^2}{\sigma_2^2}} \\
=& f_X(x) f_Y(y)
\end{aligned}
$$



反之，如果 $f(x, y) = f_X(x) f_Y(y)$，取 $x = \mu_1, y = \mu_2$ 则

$$\frac{1}{\sqrt{1-\rho^2}} = 1 \implies \rho = 0$$

$\rho = 0$ 的二维正态分布

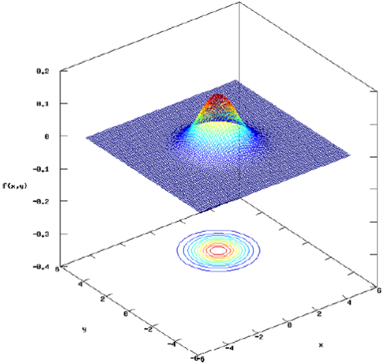
$$(X, Y) \sim N(\mu_1, \sigma_1^2; \mu_2, \sigma_2^2; \rho)$$

X 与 Y 相互独立等价于 $\rho = 0$。如果 $\rho = 0$

$$f(x, y) = \frac{1}{2\pi\sigma_1\sigma_2} e^{-\frac{1}{2}\left[\frac{(x-\mu_1)^2}{\sigma_1^2} + \frac{(y-\mu_2)^2}{\sigma_2^2}\right]}$$

$$= \frac{1}{\sqrt{2\pi}\sigma_1} e^{-\frac{1}{2}\frac{(x-\mu_1)^2}{\sigma_1^2}} \frac{1}{\sqrt{2\pi}\sigma_2} e^{-\frac{1}{2}\frac{(y-\mu_2)^2}{\sigma_2^2}}$$

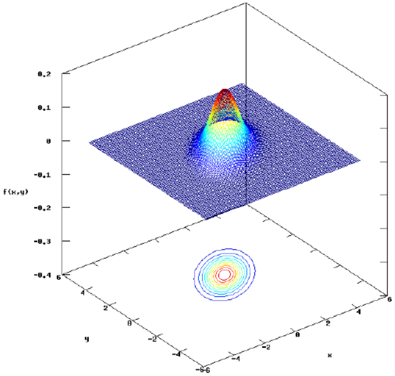
$$= f_X(x) f_Y(y)$$



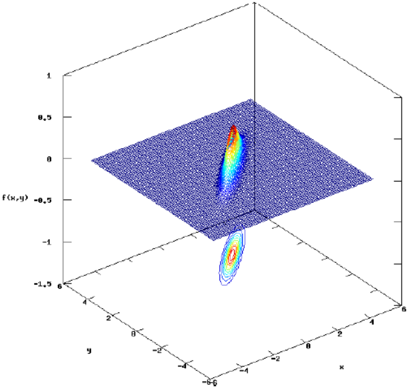
反之，如果 $f(x, y) = f_X(x) f_Y(y)$，取 $x = \mu_1, y = \mu_2$ 则

$$\frac{1}{\sqrt{1-\rho^2}} = 1 \implies \rho = 0$$

$\rho \neq 0$ 的情形



$\boldsymbol{\rho = 0.5}$

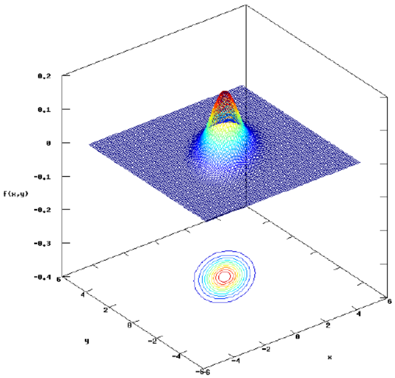


$\boldsymbol{\rho = 0.95}$

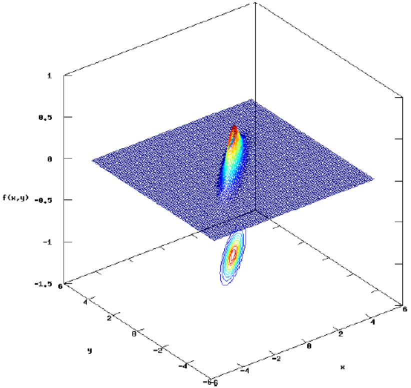
思考

$\rho \to 1$ 时，$N(\mu_1, \sigma_1^2; \mu_2, \sigma_2^2; \rho)$ 会变成什么分布？

$\rho \neq 0$ 的情形



$\boldsymbol{\rho = 0.5}$



$\boldsymbol{\rho = 0.95}$

思考

$\rho \to 1$ 时，$N(\mu_1, \sigma_1^2; \mu_2, \sigma_2^2; \rho)$ 会变成什么分布?

协方差矩阵

设 $\vec{X} \sim N(\mu_1, \sigma_1^2; \mu_2, \sigma_2^2; \rho)$，记 $\Sigma^2 = \mathrm{Var}(\vec{X}), \vec{\mu} = \mathrm{E}(\vec{X}), \vec{x} = (x_1, x_2)$

$$f(x_1, x_2) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} e^{-\frac{1}{2(1-\rho^2)}\left[\frac{(x_1-\mu_1)^2}{\sigma_1^2} - 2\rho\frac{(x_1-\mu_1)(x_2-\mu_2)}{\sigma_1\sigma_2} + \frac{(x_2-\mu_2)^2}{\sigma_2^2}\right]}$$

$$= \frac{1}{2\pi\sqrt{\begin{vmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{vmatrix}}} e^{-\frac{1}{2}(x_1-\mu_1, x_2-\mu_2)\begin{bmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{bmatrix}^{-1}\begin{pmatrix} x_1-\mu_1 \\ x_2-\mu_2 \end{pmatrix}}$$

$$= \frac{1}{2\pi\sqrt{|\Sigma^2|}} e^{-\frac{1}{2}(\vec{x}-\vec{\mu})\Sigma^{-2}(\vec{x}-\vec{\mu})^{\intercal}}$$

协方差矩阵

设 $\vec{X} \sim N(\mu_1, \sigma_1^2; \mu_2, \sigma_2^2; \rho)$，记 $\Sigma^2 = \mathrm{Var}(\vec{X}), \vec{\mu} = \mathrm{E}(\vec{X}), \vec{x} = (x_1, x_2)$

$$f(x_1, x_2) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} e^{-\frac{1}{2(1-\rho^2)}\left[\frac{(x_1-\mu_1)^2}{\sigma_1^2} - 2\rho\frac{(x_1-\mu_1)(x_2-\mu_2)}{\sigma_1\sigma_2} + \frac{(x_2-\mu_2)^2}{\sigma_2^2}\right]}$$

$$= \frac{1}{2\pi\sqrt{\begin{vmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{vmatrix}}} e^{-\frac{1}{2}(x_1-\mu_1, x_2-\mu_2)\begin{bmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{bmatrix}^{-1}\begin{pmatrix} x_1-\mu_1 \\ x_2-\mu_2 \end{pmatrix}}$$

$$= \frac{1}{2\pi\sqrt{|\Sigma^2|}} e^{-\frac{1}{2}(\vec{x}-\vec{\mu})\Sigma^{-2}(\vec{x}-\vec{\mu})^\intercal}$$

协方差矩阵

设 $\vec{X} \sim N(\mu_1, \sigma_1^2; \mu_2, \sigma_2^2; \rho)$，记 $\Sigma^2 = \mathrm{Var}(\vec{X}), \vec{\mu} = \mathrm{E}(\vec{X}), \vec{x} = (x_1, x_2)$

$$f(x_1, x_2) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} e^{-\frac{1}{2(1-\rho^2)}\left[\frac{(x_1-\mu_1)^2}{\sigma_1^2} - 2\rho\frac{(x_1-\mu_1)(x_2-\mu_2)}{\sigma_1\sigma_2} + \frac{(x_2-\mu_2)^2}{\sigma_2^2}\right]}$$

$$= \frac{1}{2\pi\sqrt{\begin{vmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{vmatrix}}} e^{-\frac{1}{2}(x_1-\mu_1, x_2-\mu_2)\begin{bmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{bmatrix}^{-1}\begin{pmatrix} x_1-\mu_1 \\ x_2-\mu_2 \end{pmatrix}}$$

$$= \frac{1}{2\pi\sqrt{|\Sigma^2|}} e^{-\frac{1}{2}(\vec{x}-\vec{\mu})\Sigma^{-2}(\vec{x}-\vec{\mu})^{\intercal}}$$

更多的高斯不变现象

1. 若 X, Y 相互独立，且 $X \sim N(\mu_1, \sigma_1^2), Y \sim N(\mu_2, \sigma_2^2)$，则 $X + Y \sim N(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2)$
   - 如果 $(X, Y) \sim N(\mu_1, \sigma_1^2; \mu_2, \sigma_2^2; \rho)$ 那么 $X + Y$ 服从什么分布?
2. 边缘分布：若 $(X, Y) \sim N(\mu_1, \sigma_1^2; \mu_2, \sigma_2^2; \rho)$，

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

3. 条件分布：若 $(X, Y) \sim N(\mu_1, \sigma_1^2; \mu_2, \sigma_2^2; \rho)$，

$$X|Y \sim N\left(\mu_1 - (Y - \mu_2)\rho\frac{\sigma_1}{\sigma_2}, (1 - \rho^2)\sigma_1^2\right)$$

$$Y|X \sim N\left(\mu_2 - (X - \mu_1)\rho\frac{\sigma_2}{\sigma_1}, (1 - \rho^2)\sigma_2^2\right)$$

这些不变性使得高斯分布成为仅有的可积分布，提供了量子场论的路径积分的唯一可计算例子。

更多的高斯不变现象

1. 若 X, Y 相互独立，且 $X \sim N(\mu_1, \sigma_1^2), Y \sim N(\mu_2, \sigma_2^2)$，则 $X + Y \sim N(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2)$
   - 如果 $(X, Y) \sim N(\mu_1, \sigma_1^2; \mu_2, \sigma_2^2; \rho)$ 那么 $X + Y$ 服从什么分布?
2. 边缘分布：若 $(X, Y) \sim N(\mu_1, \sigma_1^2; \mu_2, \sigma_2^2; \rho)$，

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

3. 条件分布：若 $(X, Y) \sim N(\mu_1, \sigma_1^2; \mu_2, \sigma_2^2; \rho)$，

$$X|Y \sim N\left(\mu_1 - (Y - \mu_2)\rho\frac{\sigma_1}{\sigma_2}, (1 - \rho^2)\sigma_1^2\right)$$

$$Y|X \sim N\left(\mu_2 - (X - \mu_1)\rho\frac{\sigma_2}{\sigma_1}, (1 - \rho^2)\sigma_2^2\right)$$

这些不变性使得高斯分布成为仅有的可积分布，提供了量子场论的路径积分的唯一可计算例子。

更多的高斯不变现象

1. 若 X, Y 相互独立，且 $X \sim N(\mu_1, \sigma_1^2), Y \sim N(\mu_2, \sigma_2^2)$，则 $X + Y \sim N(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2)$
   - 如果 $(X, Y) \sim N(\mu_1, \sigma_1^2; \mu_2, \sigma_2^2; \rho)$ 那么 $X + Y$ 服从什么分布?
2. 边缘分布：若 $(X, Y) \sim N(\mu_1, \sigma_1^2; \mu_2, \sigma_2^2; \rho)$，

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

3. 条件分布：若 $(X, Y) \sim N(\mu_1, \sigma_1^2; \mu_2, \sigma_2^2; \rho)$，

$$X|Y \sim N\left(\mu_1 - (Y - \mu_2)\rho\frac{\sigma_1}{\sigma_2}, (1 - \rho^2)\sigma_1^2\right)$$

$$Y|X \sim N\left(\mu_2 - (X - \mu_1)\rho\frac{\sigma_2}{\sigma_1}, (1 - \rho^2)\sigma_2^2\right)$$

这些不变性使得高斯分布成为仅有的可积分布，提供了量子场论的路径积分的唯一可计算例子。

更多的高斯不变现象

1. 若 X, Y 相互独立，且 $X \sim N(\mu_1, \sigma_1^2), Y \sim N(\mu_2, \sigma_2^2)$，则 $X + Y \sim N(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2)$
   - 如果 $(X, Y) \sim N(\mu_1, \sigma_1^2; \mu_2, \sigma_2^2; \rho)$ 那么 $X + Y$ 服从什么分布?
2. 边缘分布：若 $(X, Y) \sim N(\mu_1, \sigma_1^2; \mu_2, \sigma_2^2; \rho)$，

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

3. 条件分布：若 $(X, Y) \sim N(\mu_1, \sigma_1^2; \mu_2, \sigma_2^2; \rho)$，

$$X|Y \sim N\left(\mu_1 - (Y - \mu_2)\rho\frac{\sigma_1}{\sigma_2}, (1 - \rho^2)\sigma_1^2\right)$$

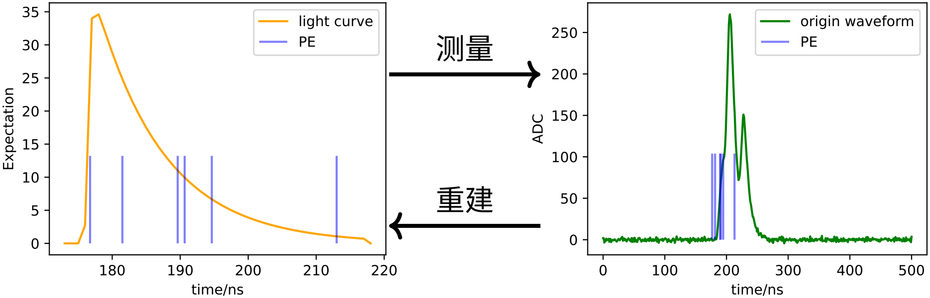
$$Y|X \sim N\left(\mu_2 - (X - \mu_1)\rho\frac{\sigma_2}{\sigma_1}, (1 - \rho^2)\sigma_2^2\right)$$

这些不变性使得高斯分布成为仅有的可积分布，提供了量子场论的路径积分的唯一可计算例子。

例：光子计数问题

设光敏仪器按一定的光强曲线 $\mu\phi(t-t_0)$ 接收光电子 (photoelectron, PE)，各光电子时间为 $\vec{z}:=(t_1,t_2,\cdots,t_N)$。波形 $\vec{w}$ 是时间序列向量。

$$\mu, t_0 \to \vec{z} \to \vec{w}$$



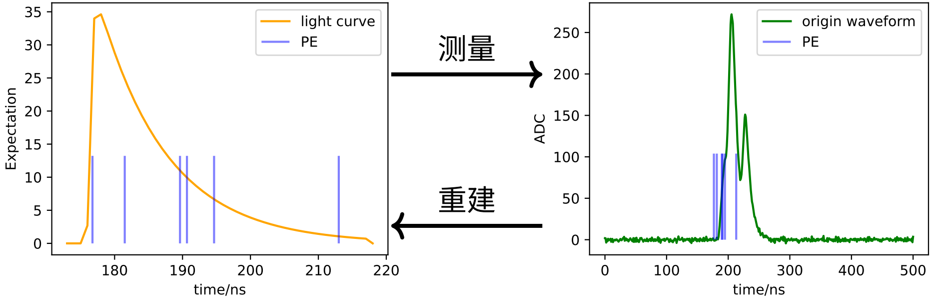
$$\mu, t_0 \xleftarrow{\text{拟合}} \vec{z} \xleftarrow{\text{分析计数}} \vec{w}$$

- 但脉冲可能有重叠，无法确切测量 $\vec{z}$，成为粒子物理实验领域的难题。

例：光子计数问题

设光敏仪器按一定的光强曲线 $\boxed{\mu}\phi(t-\boxed{t_0})$ 接收光电子 (photoelectron, PE)，各光电子时间为 $\vec{z} := (t_1, t_2, \cdots, t_N)$。波形 $\vec{w}$ 是时间序列向量。

$$\boxed{\mu, t_0} \to \vec{z} \to \vec{w}$$



$$\boxed{\mu, t_0} \xleftarrow{\text{拟合}} \vec{z} \xleftarrow{\text{分析计数}} \vec{w}$$

- 但脉冲可能有重叠，无法确切测量 $\vec{z}$，成为粒子物理实验领域的难题。

快速随机匹配追踪

- 当今最精确的中微子与暗物质实验的 PMT 波形分析方法
  - 演化自 Ghost Hunter 2019
- 多元正态分布模型

$$
\begin{aligned}
w_0(t_w)|\boldsymbol{z} &\sim \mathcal{N}\left(\boldsymbol{\mu}_\mathrm{w}(\boldsymbol{z}), \boldsymbol{\Sigma}(\boldsymbol{z})\right) \\
\boldsymbol{\mu}_\mathrm{w}(\boldsymbol{z}) &:= \sum_{i=1}^{\|\boldsymbol{z}\|_0} V_\mathrm{PE}(t_w - t_i) \\
\boldsymbol{\Sigma}(\boldsymbol{z}) &:= \sum_{i=1}^{\|\boldsymbol{z}\|_0} \Xi_\mathrm{PE}(t_w - t_i, t_v - t_i) + \sigma_\epsilon^2 I_{wv}
\end{aligned} \tag{1}
$$

Journal of Instrumentation, 17 P06040

多维伽马

多维伽马调研作业（可选）

二维正态分布是最常用的多维随机变量分布。其它的多维分布例如多维伽马分布， multivariate gamma distribution。

$$f(\vec{z}) = \frac{|\Sigma^{-1}|^{\alpha} |\vec{z}|^{\alpha - 1/2(p+1)}}{\beta^{p\alpha}\Gamma_p(\alpha)} \exp\left[-\frac{1}{\beta} \operatorname{tr} \Sigma^{-1} \vec{z}\right]$$

伽马分布适用于正实数的随机变量。

- 其中的 Σ 是否是协方差矩阵?