概率论与数理统计

第四章: 随机变量的数字特征

李立颖 (lily@sustech.edu.cn)

南方科技大学数学系

2023 秋季

数字特征

随机变量的概率特征

- 分布函数 (CDF, cumulative distribution function)
- 密度函数 (PDF, probability density function)
- 频率函数 (PMF, probability mass function)

特点

- 优点: 全面、详细、完整
- 不足: 复杂、重点不突出

问题

如何粗线条地描述随机变量的特性? 函数是一个无穷维/高维的对象, 如何用一个低维的对象去刻画?

数字特征

平均值

例子

甲、乙两射手进行打靶训练，每人各打了 100 发子弹，成绩如下：

甲：

环数	8	9	10
次数	15	40	45

乙：

环数	8	9	10
次数	35	10	55

怎样评估两人的射击水平？

分析

两人的总环环数分别为

$$\begin{aligned}\text{甲}:&8\times 15+9\times 40+10\times 45=930,\\ \text{乙}:&\ 8\times 35+9\times 10+10\times 55=920.\end{aligned}$$

每枪的平均环数为

$$\begin{aligned}\text{甲}:&\frac{8\times 15+9\times 40+10\times 45}{100}=9.3,\\ \text{乙}:&\ \frac{8\times 35+9\times 10+10\times 55}{100}=9.2\end{aligned}$$

平均值概念广泛存在

- 某班级某课程考试的平均成绩
- 电子产品的平均无故障时间
- 某地区的日平均气温和日平均降水量
- 某地区的家庭平均年收入
- 某地区的水稻的平均亩产量
- 某国家国民的平均寿命.

问题

怎样定义随机变量平均值概念?

例子

甲、乙两射手进行打靶训练，每人各打了 100 发子弹，成绩如下：

甲：

环数	8	9	10
次数	15	40	45

乙：

环数	8	9	10
次数	35	10	55

怎样评估两人的射击水平?

进一步分析

记甲每枪击中的环数为 X，因为射击次数较多，故可以认为 X 的频率函数为

X	8	9	10
p_k	0.15	0.40	0.45

则甲射手每枪的平均环数为

$$8 \times 0.15 + 9 \times 0.40 + 10 \times 0.45 = 9.3.$$

平均环数

$$\mathbb{E}(X) := \sum_{k=1}^{3} x_k p_k.$$

离散型随机变量的数学期望

定义

设随机变量 X 的频率函数为

X	x_1	x_2	$\cdots$	x_k	$\cdots$
$\mathsf{P}(X=x_k)$	p_1	p_2	$\cdots$	p_k	$\cdots$

若级数 $\sum_{k=1}^{\infty}|x_k|p_k<\infty$, 则称

$$\mathbb{E}(X):=\sum_{k=1}^{\infty}x_kp_k=\sum_{k=1}^{\infty}x_k\mathsf{P}(X=x_k)$$

为随机变量 X 的**数学期望 (期望、均值)**

“ 数学期望 (Expectation, Expected Value)” 的由来

“数学期望” 是历史上沿用下来的一个名词, 可理解为在数学上对随机变量进行计算期望到的值, 即平均值.

某商店对家用电器的销售先使用后付款. 付款额根据使用寿命 X 来确定:

寿命 (年)	$X \le 1$	$1 < X \le 2$	$2 < X \le 3$	$X > 3$
付款 (元)	1500	2000	2500	3000

假设 $X \sim \mathrm{Exp}(0.1)$, 试求该商店出售一台电器的平均收费额.

设出售一台电器的收费额为 Y, 则 Y 的频率函数为

$$\mathsf{P}(Y = 1500) = \mathsf{P}(X \le 1) = \int_0^1 \frac{1}{10} e^{-x/10} = 0.0952$$

$$\mathsf{P}(Y = 2000) = \mathsf{P}(1 < X \le 2) = \int_1^2 \frac{1}{10} e^{-x/10} = 0.0861$$

$$\mathsf{P}(Y = 2500) = \mathsf{P}(2 < X \le 3) = \int_2^3 \frac{1}{10} e^{-x/10} = 0.0779$$

$$\mathsf{P}(Y = 3000) = \mathsf{P}(3 < X \le 4) = \int_3^\infty \frac{1}{10} e^{-x/10} = 0.7408.$$

因此 $\mathbb{E}Y = 1500 \times 0.952 + 2000 \times 0.0861 + 2500 \times 0.0779 + 3000 \times 0.7408 = 2732.17$, 即商店出售一台电器平均收费额为 2732.17 元.

泊松分布的期望

若 $X \sim \pi(\lambda)$, 则 $\mathbb{E}X = \lambda$.

证明

由 $\mathsf{P}(X = k) = \frac{\lambda^k}{k!}e^{-\lambda}$, $k = 0, 1, 2, \ldots$, 我们有

$$\begin{aligned}
\mathbb{E}X &= \sum_{k=0}^{\infty} k \cdot \mathsf{P}(X = k) \\
&= \sum_{k=0}^{\infty} k \frac{\lambda^k}{k!} e^{-\lambda} \\
&= \lambda e^{-\lambda} \sum_{k=1}^{\infty} \frac{\lambda^{k-1}}{(k-1)!} \\
&= \lambda e^{-\lambda} \sum_{i=0}^{\infty} \frac{\lambda^i}{i!} = \lambda e^{-\lambda} \cdot e^{\lambda} = \lambda.
\end{aligned}$$

二项分布的期望

若 $X \sim \mathrm{Bin}(n,p)$，则 $\mathbb{E}X = np$.

证明

注意到 $k\binom{n}{k} = k\frac{n!}{k!(n-k)!} = n\frac{(n-1)!}{(k-1)!(n-k)!} = n\binom{n-1}{k-1}$.

$$
\begin{aligned}
\mathbb{E}X &= \sum_{k=0}^{n} k \cdot \binom{n}{k} p^k (1-p)^{n-k} \\
&= \sum_{k=1}^{n} n \cdot \binom{n-1}{k-1} p^k (1-p)^{n-k} \\
&= np \sum_{k'=k-1=0}^{n-1} \binom{n-1}{k'} p^{k'} (1-p)^{n-1-k'} \\
&= np[p + (1-p)]^{n-1} \\
&= np.
\end{aligned}
$$

常见分布的数学期望

分布	概率分布	期望
0-1 分布	$\mathsf{P}(X=1)=p, \mathsf{P}(X=0)=1-p$	p
$\mathrm{Bin}(n,p)$	$\mathsf{P}(X=k)=\binom{n}{k}p^k(1-p)^{n-k}, k=0,1,\dots,n$	np
$\pi(\lambda)$	$\mathsf{P}(X=k)=\frac{\lambda^k}{k!}e^{-\lambda},\ k=0,1,2,\dots$	$\lambda.$

讨论

问题

数学期望的定义中，为什么要求绝对收敛 $\sum_{k=1}^{\infty}|x_k|p_k < \infty$?

绝对收敛可以保证求和结果与顺序无关，并且使得数学期望有与可列可加性兼容的一系列极限结果.

定义

若 $\sum_{k=1}^{\infty}|x_k|p_k = \infty$，则称数学期望 $\mathbb{E}X$ 不存在.

定义

设随机变量的概率密度函数为 $f(x)$. 若

$$\int_{-\infty}^{\infty} |x| f(x)\, dx < \infty,$$

则称

$$\mathbb{E}X := \int_{-\infty}^{\infty} x f(x)\, dx$$

为随机变量 X 的**数学期望 (期望、均值)**.

注

若 $\int_{-\infty}^{\infty} |x| f(x)\, dx = \infty$, 则称 $\mathbb{E}X$ 不存在.

均匀分布的期望

若 $X \sim \mathrm{U}[a,b]$, 则 $\mathbb{E}X = \dfrac{a+b}{2}$.

证明

X 的密度函数为 $f(x) = \mathbb{1}_{(a,b)}(x) \cdot \dfrac{1}{b-a}$. 因此

$$\begin{aligned}\mathbb{E}X &= \int_a^b \frac{x}{b-a}\,dx \\ &= \frac{1}{b-a}\,\frac{b^2-a^2}{2} \\ &= \frac{a+b}{2}.\end{aligned}$$

正态分布的数学期望

若 $X \sim \mathcal{N}(\mu, \sigma^2)$, 则 $\mathbb{E}X = \mu$.

证明

$$\begin{aligned}
\mathbb{E}X &= \int_{-\infty}^{\infty} x \cdot \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}\, dx \\
&= \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^{\infty} (x-\mu+\mu) e^{-\frac{(x-\mu)^2}{2\sigma^2}}\, dx \\
&= \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^{\infty} (x-\mu) e^{-\frac{(x-\mu)^2}{2\sigma^2}}\, dx + \mu \cdot \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}\, dx \\
&= 0 + \mu = \mu.
\end{aligned}$$

指数分布的数学期望

设某元器件的寿命 X 服从指数分布 $\mathrm{Exp}(1/\theta)$, 其密度为

$$f(x) = \mathbb{1}_{(0,\infty)}(x)\frac{1}{\theta}e^{-\frac{x}{\theta}}.$$

则 $\mathbb{E}X = \theta$.

$$\begin{aligned}\mathbb{E}X &= \int_0^\infty x\frac{1}{\theta}e^{-\frac{x}{\theta}}\,dx \\ &= \theta\int_0^\infty \frac{x}{\theta}e^{-\frac{x}{\theta}}\,d\left(\frac{x}{\theta}\right) \\ &= \theta\int_0^\infty te^{-t}\,dt = \theta\left(-te^{-t} - e^{-t}\Big|_0^\infty\right) = \theta.\end{aligned}$$

工程上, 如果某件产品的平均寿命为 $\theta = 10^k$ (小时), 则称该产品为 "k 级" 产品. k 越大, 失效率 10^{-k} 越低, 则平均寿命越长, 可靠性越高. 在航空、航天、军事、医疗等领域, 通常要求元器件达 9 级以上, 即 $\theta \approx 114160$ (年).

连续型随机变量的期望

分布	概率密度	期望
$\mathrm{U}[a,b]$	$f(x)=\mathbb{1}_{(a,b)}(x)\dfrac{1}{b-a}$	$\dfrac{a+b}{2}$
$\mathrm{Exp}(\lambda)$	$f(x)=\mathbb{1}_{(0,\infty)}\lambda e^{-\lambda x}$	$\dfrac{1}{\lambda}$
$\mathcal{N}(\mu,\sigma^2)$	$f(x)=\dfrac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(x-\mu)^2}{2\sigma^2}}$	μ

例题

设随机变量 X 的密度函数为

$$f(x) = \frac{1}{2}e^{-|x|}, \quad x \in \mathbb{R}.$$

求 $\mathbb{E}X$.

解

$$\mathbb{E}X = \int_{-\infty}^{\infty} x\frac{1}{2}e^{-|x|}\,dx = 0 \quad (\text{对称性}).$$

Cauchy 分布的期望

设 $X \sim f(x) = \dfrac{1}{\pi}\dfrac{1}{1+x^2}$, 则 $\mathbb{E}X$ 不存在.

注意到

$$\begin{aligned}\int_{-\infty}^{\infty} |x| f(x)\, dx &= \lim_{M\to\infty} \int_{-M}^{M} |x| f(x)\, dx \\ &= \lim_{M\to\infty} 2\int_0^M x \cdot \frac{1}{\pi}\frac{1}{1+x^2}\, dx \\ &= \lim_{M\to\infty} \frac{1}{\pi} \ln(1+x^2)\Big|_0^M = \infty.\end{aligned}$$

故 $\mathbb{E}X$ 不存在.

Markov 不等式

设随机变量 X 满足 $\mathsf{P}(X \geq 0) = 1$, 且 $\mathbb{E}X$ 存在, 则

$$\mathsf{P}(X \geq t) \leq \frac{\mathbb{E}X}{t}.$$

证明

我们只对连续情形证明. 因为 $X \geq 0$, 所以

$$\begin{aligned}
\mathbb{E}X = \int_0^\infty xf(x)\,dx &= \int_0^t xf(x)\,dx + \int_t^\infty xf(x)\,dx \\
&\geq \int_t^\infty xf(x)\,dx \\
&\geq t\int_t^\infty f(x)\,dx = t\mathsf{P}(X \geq t).
\end{aligned}$$

移项即得结论.

背景

设已知随机变量 X 的分布，我们需要计算的不是 X 的期望，而是 X 的某个函数 $g(X)$ 的期望. 如何计算?

例子

飞机机翼受到的压力为 $W = kv^2$，其中 v 为风速，$k > 0$ 为常数. 问机翼受到的平均压力有多大?

思路

已知 $X \sim g(X)$，则 $Y = g(X) \sim f_Y(y)$. 因此

$$\mathbb{E}g(X) = \mathbb{E}Y = \int_{-\infty}^{\infty} y f_Y(y)\, dy.$$

是否可以不求 $g(X)$ 的分布而只根据 X 的分布求 $\mathbb{E}g(X)$?

定理

设 $y = g(x)$ 为一 (可测) 函数, 则

1. 若 X 为离散型, 且频率函数为 $\mathsf{P}(X = x_k) = p_k$, $k = 1, 2, \ldots$, 且 $\sum_{k=1}^{\infty} |g(x_k)| \cdot p_k < \infty$, 则

$$\mathbb{E}Y = \mathbb{E}g(X) = \sum_{k=1}^{\infty} g(x_k) p_k.$$

2. 若 X 为连续型, 且密度函数为 $f(x)$, 且 $\int_{-\infty}^{\infty} |g(x)| f(x)\, dx < \infty$, 则

$$\mathbb{E}Y = \mathbb{E}g(X) = \int_{-\infty}^{\infty} g(x) f(x)\, dx.$$

求 $g(X)$ 时只需要知道 X 的分布, 这给我们带来很大的便利性.

例题

设风速 $V \sim \mathrm{U}(0,a)$, 且飞机机翼受到的正压力 $W = kV^2$, $k > 0$ 为常数. 求 $\mathbb{E}W$.

解

V 的密度函数为 $f(v) = \mathbb{1}_{(0,a)}(v)\dfrac{1}{a}$. 因此

$$\mathbb{E}W = \int_0^a \frac{k}{a} v^2 \, dv = \frac{1}{3}ka^3.$$

因此飞机机翼受到的平均正压力为 $\dfrac{1}{3}ka^3$.

二元随机变量函数的期望

设 $z = g(x, y)$ 为二元 (可测) 函数.

1. 若 X, Y 为离散型, 且联合频率函数为
$$\mathsf{P}(X = x_i,\ Y = y_j) = p_{ij},\ i, j = 1, 2, \ldots$$
若 $\sum_{i,j=1}^{\infty} |g(x_i, y_j)| p_{ij} < \infty$, 则
$$\mathbb{E}Z = \mathbb{E}g(X, Y) = \sum_{i,j=1}^{\infty} g(x_i, y_j) p_{ij}.$$
2. 若 X, Y 为连续型, 且联合密度函数为 $f(x, y)$. 若
$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} |g(x, y)| f(x, y)\, dx dy < \infty,$$
则
$$\mathbb{E}Z = \mathbb{E}g(X, Y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(x, y) f(x, y)\, dx dy.$$

例

设 (X,Y) 的密度函数为 $f(x,y)=\mathbb{1}_{\{0\le y\le x\le 1\}}(x,y)15xy^2$. 试求 $\mathbb{E}X$, $\mathbb{E}Y$ 及 $\mathbb{E}XY$.

思路

除了可以先求 $f_X(x)$, $f_Y(y)$ 然后利用一元随机变量期望的公式, 我们还可以对 $g(x,y)=x$ 或 $g(x,y)=y$ 利用二维随机变量的公式.

$$
\begin{aligned}
\mathbb{E}X &= \iint_{\mathbb{R}\times\mathbb{R}} xf(x,y)\,dxdy = \iint_{0\le y\le x\le 1} x\cdot 15xy^2\,dxdy \\
&= \int_0^1 dx\int_0^x 15x^2y^2\,dxdy = \frac{15}{3}\int_0^1 x^5\,dx = \frac{5}{6}. \\
\mathbb{E}Y &= \iint_{\mathbb{R}^2} yf(x,y)\,dxdy = \frac{5}{8}. \\
\mathbb{E}(XY) &= \iint_{\mathbb{R}^2} xyf(x,y)\,dxdy = \iint_{0\le y\le x\le 1} xy\cdot 15xy^2\,dxdy = \frac{15}{28}.
\end{aligned}
$$

例题

设随机变量 X 的分布律为

X	-2	0	2
P	0.4	0.3	0.3

求 $\mathbb{E}(3X^2+5)$.

解

$$\begin{aligned}\mathbb{E}(3X^2+5) &= [3\cdot(-2)^2+5]\cdot 0.4+[3\cdot 0^2+5]\cdot 0.3+[3.2^2+5]\cdot 0.3\\ &= 13.4.\end{aligned}$$

例题

设 (X,Y) 在区域 A 上服从均匀分布，其中 A 为 x 轴、y 轴与直线 $x+y+1=0$ 围成的区域. 求 $\mathbb{E}X$, $\mathbb{E}(-3X+2Y)$, $\mathbb{E}XY$.

解

区域 A 可以写为

$$A=\{(x,y):-1\le x\le 0,\quad -1-x\le y\le 0\}.$$

且 A 的面积为 $1/2$. 因此密度函数为 $f(x,y)=\mathbb{1}_A(x,y)\cdot 2$,

$$\mathbb{E}X=\int_{-1}^{0}dx\int_{-1-x}^{0}x\,2dy=-\frac{1}{3},$$

$$\mathbb{E}(-3X+2Y)=\int_{-1}^{0}dx\int_{-1-x}^{0}2(-3x+2y)\,dy=\frac{1}{3},$$

$$\mathbb{E}XY=\int_{-1}^{0}dx\int_{-1-x}^{0}2xy\,dy=\frac{1}{12}.$$

数学期望的基本性质

- 设 $a \le X \le b$, 则 $a \le \mathbb{E}X \le b$.
- 若 $X = c$, (a.e.), 则 $\mathbb{E}X = c$.

线性性质

- 设 c 为常数, 则 $\mathbb{E}(cX) = c\mathbb{E}X$.
- 设 X, Y 为随机变量, $\mathbb{E}(X+Y) = \mathbb{E}X + \mathbb{E}Y$.
- 设 $a_1, \ldots, a_n$ 为常数, 则对随机变量 $X_1, \ldots, X_n$, 有

$$\mathbb{E}\left(\sum_{i=1}^{n} a_i X_i\right) = \sum_{i=1}^{n} a_i \mathbb{E}X_i.$$

独立随机变量乘积的期望

- 设 X, Y 相互独立, 则 $\mathbb{E}XY = \mathbb{E}X \cdot \mathbb{E}Y$.
- 设 $X_1, X_2, \ldots, X_n$ 相互独立, 则

$$\mathbb{E}\Big(X_1 X_2 \cdots X_n\Big) = \mathbb{E}X_1 \cdot \mathbb{E}X_2 \cdot \cdots \cdot \mathbb{E}X_n.$$

例题

已知随机变量 X 服从参数为 2 的泊松分布，令 $Z = 3X - 2$，求 $\mathbb{E}Z$.

解

由于 $X \sim \pi(2)$，因此 $\mathbb{E}X = 2$. 由线性性，我们有

$$\mathbb{E}Z = \mathbb{E}(3X - 2) = 3\mathbb{E}X - 2 = 3 \cdot 2 - 4 = 4.$$

例题

设二维随机变量 (X,Y) 的密度函数为

$$f(x,y)=\begin{cases}\frac{1}{4}x(1+3y^2), & 0<x<2,\ 0<y<1,\\ 0, & \text{其它}.\end{cases}$$

求 $\mathbb{E}X$, $\mathbb{E}Y$, $\mathbb{E}(X+Y)$, $\mathbb{E}(XY)$.

解

$$\mathbb{E}X=\int_0^2 x\cdot\frac{1}{4}x\,dx\int_0^1(1+3y^2)\,dy=\frac{4}{3},\mathbb{E}Y\quad=\int_0^2\frac{1}{4}x\,dx\int_0^1 y\cdot(1+3y^2)\,dy=\frac{5}{8}.$$

由线性性,

$$\mathbb{E}(X+Y)=\mathbb{E}X+\mathbb{E}Y=\frac{4}{3}+\frac{5}{8}=\frac{47}{24}.$$

由独立性,

$$\mathbb{E}XY=\mathbb{E}X\cdot\mathbb{E}Y=\frac{4}{3}\cdot\frac{5}{8}=\frac{5}{6}.$$

作业

- P116: 6, 15, 20, 21, 31
- 1. 设随机变量 X 的概率密度为

$$f(x) = \begin{cases} e^{-x}, & x > 0, \\ 0, & x \le 0. \end{cases}$$

  求以下随机变量的期望:

  (1) $Y = 2X$,

  (2) $Y = e^{-2X}$.

  2. 设随机变量 (X, Y) 的概率密度为

$$f(x, y) = \begin{cases} 12y^2, & 0 \le y \le x \le 1, \\ 0, & \text{其它}. \end{cases}$$

  求 $\mathbb{E}X$, $\mathbb{E}Y$, $\mathbb{E}XY$, $\mathbb{E}(X^2 + Y^2)$.

思考题

考虑把原本属于 n 个人的 n 封信打乱重新分发，每个人拿到一封信. 令 X 为拿到自己信的总人数. 对于伯努利装错信封问题，我们知道

$$\mathsf{P}(X=0)=1-\frac{1}{1!}+\frac{1}{2!}+\cdots+(-1)^n\cdot\frac{1}{n!}.$$

现在我们考察 X 的期望. 求

- $n=3$ 时，$\mathbb{E}X=?$
- $n=4$ 时，$\mathbb{E}X=?$

提示

将 X 改写为

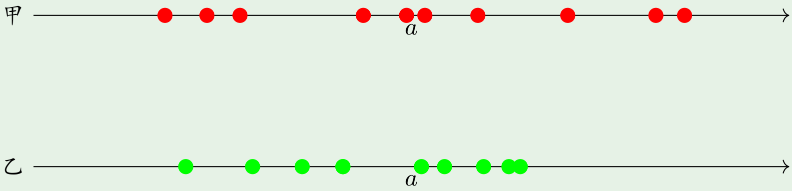
$$X=\xi_1+\xi_2+\cdots+\xi_n,\quad \xi_k=\begin{cases}1, & \text{第 } k \text{ 个人拿到属于自己的信封},\\ 0, & \text{没拿到}.\end{cases}$$

易知 $\mathsf{P}(\xi_k=1)=\dfrac{1}{n}$.

我们已经介绍了随机变量的数学期望，它体现了随机变量取值的平均水平，是随机变量的一个重要数值特征．但是在很多场合，仅仅知道平均值是不够的．

例 1：长度测量

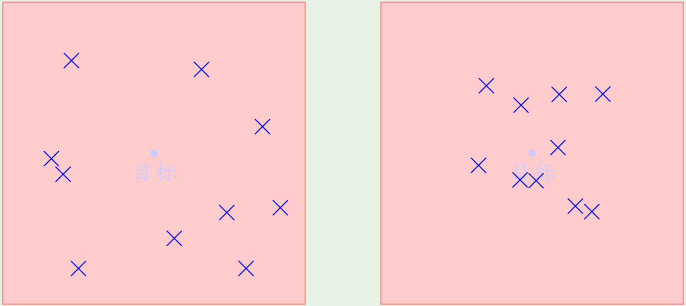
某零件的真实长度为 a，现用甲、乙两台仪器各测量 10 次，将测量结果 X 用坐标上的点表示：



我们认为乙仪器更准确一些，因为乙的结果更**集中**在均值附近．

例 2: 炮弹落点

甲、乙两门炮同时向一目标射击 10 发炮弹，其落点距目标的位置如图:



我们认为乙炮更好，因为射击结果更**集中**于中心附近.

例子

设有两个牌子的手表，其走时误差情况如下表：

日误差（秒）	-3	-2	-1	0	1	2	3
概率（甲）	0.10	0.15	0.15	0.20	0.15	0.15	0.10
概率（乙）	0.05	0.05	0.10	0.60	0.10	0.05	0.05

试问哪种牌号的手表质量较好?

分析

设两种手表的走时误差分别为 X, Y, 则

$$\mathbb{E}X = \sum_{k=1}^{7} x_k \mathsf{P}(X = x_k) = 0$$

$$\mathbb{E}Y = \sum_{k=1}^{7} y_k \mathsf{P}(Y = y_k) = 0.$$

因此两种手表的平均误差一样.

质量是否一样?

从偏离平均值的大小来考虑

对随机变量 X, 考虑偏差 $|X-\mathbb{E}X|$:

- 偏差越小, 说明质量越稳定
- 缺点: 绝对值运算不方便.

平方偏差

- 平方偏差 $\big(X-\mathbb{E}X\big)^2$ 仍是随机变量.
- 平方偏差的平均值 $\mathbb{E}\big(X-\mathbb{E}X\big)^2$ 反映了偏离平均值的大小.

定义

对随机变量 X, 若

$$\mathrm{Var}(X) := D(X) := \mathbb{E}\big(X-\mathbb{E}X\big)^2$$

存在, 则称 $D(X)$ 为随机变量 X 的**方差 (variance)**, $\sqrt{D(X)}$ 称为**标准差 (standard deviation)**.

例题

设甲、乙两射手击中环数分别为 X, Y, 频率函数为

X	8	9	10
p_k	0.15	0.40	0.45

,

Y	8	9	10
p_k	0.35	0.10	0.55

.

试从均值和方差两方面评估两个的射击技术.

解

先计算数学期望

$$\mathbb{E}X = 8 \times 0.15 + 9 \times 0.40 + 10 \times 0.45 = 9.3$$
$$\mathbb{E}Y = 8 \times 0.35 + 9 \times 0.10 + 10 \times 0.55 = 9.2.$$

再计算方差

$$\mathrm{Var}(X) = (8 - 9.3)^2 \cdot 0.15 + (9 - 9.3)^2 \cdot 0.4 + (10 - 9.3)^2 \cdot 0.45 = 0.51$$
$$\mathrm{Var}(Y) = (8 - 9.2)^2 \cdot 0.35 + (9 - 9.2)^2 \cdot 0.1 + (10 - 9.2)^2 \cdot 0.55 = 0.86.$$

可看出甲的射击水平比乙高, 且更稳定.

数学期望 随机变量的平均值

方差 随机变量与其平均值的平均偏离程度.

方差的计算

视 $\mathrm{Var}(X)=\mathbb{E}\big(X-\mathbb{E}X\big)^2$ 为 $g(X)=\big(X-\mathbb{E}X\big)^2$ 的数学期望, 则有

- 离散型: 设 X 的频率函数为 $\mathsf{P}(X=x_k)=p_k$, $k=1,2,\ldots$, 则

$$\mathrm{Var}(X)=\sum_{k=1}^{\infty}(x_k-\mathbb{E}X)^2\cdot p_k.$$

- 连续型: 设 X 的概率密度函数为 $f(x)$, 则

$$\mathrm{Var}(X)=\int_{-\infty}^{\infty}\big(x-\mathbb{E}X\big)^2 f(x)\,dx.$$

$$\begin{aligned}\mathrm{Var}(X)&=\mathbb{E}\big(X-\mathbb{E}X\big)^2 \qquad (c=\mathbb{E}X)\\&=\mathbb{E}X^2-2\mathbb{E}cX+\mathbb{E}c^2=\mathbb{E}X^2-2c\cdot c+c^2\\&=\mathbb{E}X^2-\big(\mathbb{E}X\big)^2.\end{aligned}$$

例题

对于随机变量 X, 其分布列为

X	-1	0	1
p	0.1	0.8	0.1

求 $\mathrm{Var}(X)$.

解

$$\begin{aligned} \mathbb{E}X &= 0, \\ \mathbb{E}X^2 &= (-1)^2 \cdot 0.1 + 1^2 \cdot 0.1 = 0.2 \\ \mathrm{Var}(X) &= \mathbb{E}X^2 - \left(\mathbb{E}X\right)^2 = 0.2. \end{aligned}$$

例题

设随机变量 X 的密度函数为

$$f(x) = \begin{cases} 1+x, & -1 \le x \le 0, \\ 1-x, & 0 < x \le 1. \end{cases}$$

求 $\mathrm{Var}(X)$.

解

$$\begin{aligned} \mathbb{E}X &= \int_{-1}^{0} x(1+x)\,dx + \int_{0}^{1} x(1-x)\,dx = 0, \\ \mathbb{E}X^2 &= \int_{-1}^{0} x^2(1+x) + \int_{0}^{1} x^2(1-x)\,dx = \frac{1}{6}, \\ \mathrm{Var}(X) &= \mathbb{E}X^2 - \big(\mathbb{E}X\big)^2 = \frac{1}{6}. \end{aligned}$$

例题

一批零件中有 9 个合格品 3 个次品，从这批零件中任取一个，如果每次取出的废品不再放回去. 求在取得合格品以前，已取出的废品数的期望、方差和标准差.

解

我们可以求得 X 的分布列:

X	0	1	2	3
p	$\frac{3}{4}$	$\frac{9}{44}$	$\frac{9}{44}$	$\frac{1}{220}$

因此,

$$\mathbb{E}X = \sum_{k=0}^{3} x_k p_k = 0.301,$$

$$\mathrm{Var}(X) = \sum_{k=0}^{3} x_k^2 p_k - 0.301^2 = 0.322,$$

$$\sqrt{\mathrm{Var}(X)} = \sqrt{0.322} = 0.567.$$

例题

设随机变量 X 的数学期望为 $\mathbb{E}X$, 方差为 $\mathrm{Var}(X) > 0$. 引入新的随机变量

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

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

证明

令 $\mathbb{E}X = c$, $\mathrm{Var}(X) = d$. 则

$$\begin{aligned}
\mathbb{E}X^* &= \mathbb{E}\frac{X-c}{\sqrt{d}} = \frac{1}{\sqrt{d}}\big(\mathbb{E}X - c\big) = 0 \\
\mathrm{Var}(X^*) &= \mathbb{E}\big(X^*\big)^2 = \mathbb{E}\left[\frac{X-c}{\sqrt{d}}\right]^2 \\
&= \frac{1}{d}\mathbb{E}(X-c)^2 = \frac{d}{d} = 1.
\end{aligned}$$

泊松分布的方差

若 $X \sim \pi(\lambda)$，则 $\mathrm{Var}(X) = \lambda$.

证明

已知 $\mathbb{E}X = \lambda$. 我们还有

$$\begin{aligned}\mathbb{E}X^2 &= \mathbb{E}\big[X(X-1)+X\big] \\ &= \sum_{k=0}^{\infty} k(k-1)\frac{\lambda^k}{k!}e^{-\lambda} + \mathbb{E}X \\ &= \lambda^2 e^{-\lambda}\sum_{k=2}^{\infty}\frac{\lambda^{k-2}}{(k-2)!} + \lambda \\ &= \lambda^2 + \lambda.\end{aligned}$$

因此

$$\mathrm{Var}(X) = \mathbb{E}X^2 - \big(\mathbb{E}X\big)^2 = \lambda^2 + \lambda - \lambda^2 = \lambda.$$

二项分布的方差

若 $X \sim \mathrm{Bin}(n,p)$，则 $\mathrm{Var}(X) = np(1-p)$.

引理

若 X 与 Y 独立，且 $\mathbb{E}X = \mathbb{E}Y = 0$，则 $\mathbb{E}(X+Y)^2 = \mathbb{E}X^2 + \mathbb{E}Y^2$.
特别地，若 X 与 Y 独立，则 $\mathrm{Var}(X \pm Y) = \mathrm{Var}(X) + \mathrm{Var}(Y)$.

证明

二项分布 X 可看作 n 个独立同分布的两点分布的和:

$$X = \xi_1 + \xi_2 + \cdots + \xi_n, \quad \xi_k \sim \mathrm{Ber}(p).$$

易见 $\mathbb{E}\xi = p$，$\mathbb{E}\xi^2 = p$，故 $\mathrm{Var}(\xi) = p(1-p)$. 由引理,

$$\mathrm{Var}(X) = \mathrm{Var}(\xi_1 + \xi_2 + \cdots + \xi_n) = \mathrm{Var}(\xi_1) + \mathrm{Var}(\xi_2) + \cdots + \mathrm{Var}(\xi_n) = np(1-p).$$

均匀分布的方差

若 $X \sim \mathrm{U}(a,b)$, 则 $\mathrm{Var}(X) = \dfrac{(b-a)^2}{12}$.

证明

已知 $\mathbb{E}X = \dfrac{a+b}{2}$. 我们有

$$\begin{aligned}
\mathbb{E}X^2 &= \int_a^b \frac{x^2}{b-a}\,dx \\
&= \frac{b^3-a^3}{3(b-a)} = \frac{a^2+ab+b^2}{3}. \\
\mathrm{Var}(X) &= \mathbb{E}X^2 - \left(\mathbb{E}X\right)^2 \\
&= \frac{a^2+ab+b^2}{3} - \frac{(a+b)^2}{4} \\
&= \frac{(b-a)^2}{12}.
\end{aligned}$$

指数分布的方差

若 $X \sim \text{Exp}(\lambda)$, 则 $\text{Var}(X) = \dfrac{1}{\lambda^2}$.

证明

已知 $\mathbb{E}X = \dfrac{1}{\lambda}$. 我们有

$$\begin{aligned}\text{Var}(X) &= \mathbb{E}X^2 - \frac{1}{\lambda^2} \\ &= \int_0^\infty x^2 \lambda e^{-\lambda x}\, dx - \frac{1}{\lambda^2} \\ &= \frac{2}{\lambda^2} - \frac{1}{\lambda^2} = \frac{1}{\lambda^2}.\end{aligned}$$

正态分布的方差

若 $X \sim \mathcal{N}(\mu, \sigma^2)$，则 $\mathrm{Var}(X) = \sigma^2$.

证明

$$\mathrm{Var}(X) = \int_{-\infty}^{\infty} (x-\mu)^2 \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}\, dx$$

$$\overset{t=\frac{x-\mu}{\sigma}}{=\!=\!=\!=} \int_{-\infty}^{\infty} \sigma^2 t^2 \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}}\, dt.$$

由分部积分，

$$\int_{-\infty}^{\infty} t^2 e^{-\frac{t^2}{2}}\, dt = \int_{-\infty}^{\infty} t d\left(-e^{-\frac{t^2}{2}}\right) = \int_{-\infty}^{\infty} e^{-\frac{t^2}{2}}\, dt = \sqrt{2\pi}.$$

代入得 $\mathrm{Var}(X) = \sigma^2$.

常见分布的期望与方差

分布	概率函数	期望	方差
0-1 分布	$\mathsf{P}(X=k)=\mathbb{1}_{k=1}p+\mathbb{1}_{k=0}(1-p)$	p	$p(1-p)$
$\mathrm{Ber}(n,p)$	$\mathsf{P}(X=k)=\binom{n}{k}p^k(1-p)^{n-k}$	np	$np(1-p)$
$\pi(\lambda)$	$\mathsf{P}(X=k)=\frac{\lambda^k e^{-\lambda}}{k!}$	λ	λ
$\mathrm{U}(a,b)$	$f(x)=\mathbb{1}_{(a,b)}(x)\frac{1}{b-a}$	$\frac{a+b}{2}$	$\frac{(b-a)^2}{12}$
$\mathrm{Exp}(\lambda)$	$f(x)=\mathbb{1}_{(0,\infty)}(x)\lambda e^{-\lambda}$	$\frac{1}{\lambda}$	$\frac{1}{\lambda^2}$
$\mathcal{N}(\mu,\sigma^2)$	$f(x)=\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(x-\mu)^2}{2\sigma^2}}$	μ	σ^2

方差的性质

1. 若 $X \overset{\text{a.e.}}{=} c$, 则 $\operatorname{Var}(X)=0$.
2. 设 c 为常数, 则 $\operatorname{Var}(cX)=c^2\operatorname{Var}(X)$.

$$\operatorname{Var}(cX)=\mathbb{E}\big(cX-\mathbb{E}(cX)\big)^2=\mathbb{E}(cX-c\mathbb{E}X)^2=c^2\mathbb{E}(X-\mathbb{E}X)^2=c^2\operatorname{Var}(X).$$

3. 对于随机变量 X 和 Y, 有

$$\begin{aligned}\operatorname{Var}(X+Y)&=\mathbb{E}\big((X+Y)-\mathbb{E}(X+Y)\big)^2\\&=\mathbb{E}\big((X-\mathbb{E}X)+(Y-\mathbb{E}Y)\big)^2\\&=\mathbb{E}(X-\mathbb{E}X)^2+\mathbb{E}(Y-\mathbb{E}Y)^2+2\mathbb{E}(X-\mathbb{E}X)(Y-\mathbb{E}Y)\\&=\operatorname{Var}(X)+\operatorname{Var}(Y)+2\mathbb{E}(X-\mathbb{E}X)(Y-\mathbb{E}Y).\end{aligned}$$

4. 特别地, 当 X,Y 独立时, $\mathbb{E}(X-\mathbb{E}X)(Y-\mathbb{E}Y)=\mathbb{E}(X-\mathbb{E}X)\cdot\mathbb{E}(Y-\mathbb{E}Y)=0$, 因此

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

5. 若 $\operatorname{Var}(X)=0$, 则 $X \overset{\text{a.e.}}{=} c$ (常数).

Chebyshev 不等式

3σ-原则

正态随机变量的值几乎都落在 $(\mu-3\sigma,\mu+3\sigma)$ 之间，即 $X\sim\mathcal{N}(\mu,\sigma^2)$ 时，

$$\mathsf{P}(|X-\mu|<3\sigma)=99.74\%.$$

问

对一般的随机变量 X，如何估计概率 $\mathsf{P}(|X-\mu|\le\varepsilon)$? 其中 $\mu=\mathbb{E}X$, $\varepsilon>0$ 为任意正实数.

切比雪夫 (Chebyshev) 不等式

设 $\mu=\mathbb{E}X$, $\sigma^2=\mathrm{Var}(X)$ 都存在，则 $\forall\varepsilon>0$，有

$$\mathsf{P}(|X-\mu|\ge\varepsilon)\le\frac{\sigma^2}{\varepsilon^2}.$$

切比雪夫 (Chebyshev) 不等式

设 $\mu = \mathbb{E}X$, $\sigma^2 = \mathrm{Var}(X)$ 都存在, 则 $\forall \varepsilon > 0$, 有

$$\mathsf{P}(|X-\mu| \geq \varepsilon) \leq \frac{\sigma^2}{\varepsilon^2}.$$

回顾: Markov 不等式

若随机变量 $X \geq 0$, 且 $\mathbb{E}X$ 存在, 则

$$\mathsf{P}(X \geq t) \leq \frac{\mathbb{E}X}{t}, \quad t > 0.$$

证明

由 Markov 不等式,

$$\mathsf{P}(|X-\mu| \geq \varepsilon) = \mathsf{P}(|X-\mu|^2 \geq \varepsilon^2) \leq \frac{\mathbb{E}|X-\mu|^2}{\varepsilon^2} = \frac{\mathrm{Var}(X)}{\varepsilon^2}.$$

Chebyshev 不等式推论

3σ-原则

取 $\varepsilon = 3\sigma,\ 4\sigma$, 有

$$\mathsf{P}(|X-\mu| < 3\sigma) = 1 - \mathsf{P}(|X-\mu| \geq 3\sigma) \geq 1 - \frac{1}{9} = 88.90\%,$$
$$\mathsf{P}(|X-\mu| < 4\sigma) = 1 - \mathsf{P}(|X-\mu| \geq 4\sigma) \geq 1 - \frac{1}{16} = 93.75\%.$$

因此对一般的随机变量, 3σ-原则的可信度也接近 90%.

推论 2

若 $\sigma^2 = 0$, 则 $\mathsf{P}(|X-\mu| = 0) = 1$.

例题

已知正常男性成人血液中，每一毫升白细胞数平均是 7300，标准差是 700. 利用 Chebyshev 不等式估计每毫升白细胞数在 5200 至 9400 之间的概率.

解

设每毫升白细胞数为 X，依题意，$\mathbb{E}X = 7300$，$\mathrm{Var}(X) = 700^2$，
$\{X \in (5200, 9400)\} = \{|X - 7300| \geq 3 \cdot 700\}$. 因此

$$\mathsf{P}(X \in (5200, 9400)) = 1 - \mathsf{P}(|X - 7300| \geq 3 \cdot 700) \geq 1 - \frac{700^2}{3^2 \cdot 700^2} = \frac{8}{9}.$$

- P118: 49, 50, 55
- 补充题
  1. 设 X, Y 是相互独立的随机变量, 且有 $\mathbb{E}X = 3$, $\mathbb{E}Y = 1$, $\mathrm{Var}(X) = 4$, $\mathrm{Var}(Y) = 9$. 令 $Z = 5X - 2Y + 15$. 求 $\mathbb{E}Z$ 与 $\mathrm{Var}(Z)$.
  2. 设随机变量 X_1, X_2, X_3, X_4 相互独立, 且有 $\mathbb{E}X_i = 2i$, $\mathrm{Var}(X_i) = 5 - i$, $i = 1, 2, 3, 4$. 令 $Z = 2X_1 - X_2 + 3X_3 - \frac{1}{2}X_4$. 求 $\mathbb{E}Z$ 与 $\mathrm{Var}(Z)$.

随机变量间的关系

回顾

设 $(X,Y) \sim f(x,y)$, $X \sim f_X(x)$, $Y \sim f_Y(y)$, 则

$$X, Y \text{ 相互独立} \quad \Leftrightarrow \quad f(x,y) = f_X(x) f_Y(y).$$

问: 若 X, Y 不独立, 如何刻画它们之间的关系?

若 X, Y 相互独立, 则

$$\mathbb{E}\Big[(X - \mathbb{E}X)(Y - \mathbb{E}Y)\Big] = 0.$$

反之, 若上式不为 0, 则 X, Y 必不独立, 它们之间一定存在某些联系.

定义

若 X,Y 的方差均存在, 记

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

称为 X, Y 的**协方差 (covariance)**.

协方差的基本性质

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

1. 若 X, Y 相互独立, 则 $\mathrm{Cov}(X,Y) = 0$.
2. (对称性) $\mathrm{Cov}(X,Y) = \mathrm{Cov}(Y,X)$.
3. $\mathrm{Var}(X) = \mathbb{E}(X-\mathbb{E}X)^2 = \mathrm{Cov}(X,X)$.
4. 协方差的计算

$$\begin{aligned}\mathrm{Cov}(X,Y) &= \mathbb{E}(X-\mathbb{E}X)(Y-\mathbb{E}Y) = \mathbb{E}(X-\mu_X)(Y-\mu_Y)\\ &= \mathbb{E}XY - \mathbb{E}\mu_X Y - \mathbb{E}X\mu_Y + \mu_X\mu_Y\\ &= \mathbb{E}XY - \mu_X\mu_Y = \mathbb{E}XY - \mathbb{E}X\cdot\mathbb{E}Y.\end{aligned}$$

5. $\mathrm{Cov}(X,Y) = \mathrm{Var}(X) + \mathrm{Var}(Y) + 2\,\mathrm{Cov}(X,Y)$.

协方差的双线性性

性质

1. 对任意常数 a, b，有

$$\begin{aligned}\mathrm{Cov}(aX, bY) &= \mathbb{E}\Big[(aX - \mathbb{E}(aX))(bY - \mathbb{E}(bY))\Big] \\ &= ab\mathbb{E}(X - \mathbb{E}X)(Y - \mathbb{E}Y) = ab\,\mathrm{Cov}(X, Y).\end{aligned}$$

2. $\mathrm{Cov}(X_1 + X_2, Y) = \mathrm{Cov}(X_1, Y) + \mathrm{Cov}(X_2, Y)$.

双线性性 (bilinearity)

假设 $U = a + \sum_{i=1}^{n} b_i X_i$, $V = c + \sum_{j=1}^{m} d_j Y_j$，则

$$\mathrm{Cov}(U, V) = \sum_{i=1}^{n}\sum_{j=1}^{m} b_i d_j\,\mathrm{Cov}(X_i, Y_j).$$

相关系数

协方差的意义

- X, Y 相互独立 $\Rightarrow \mathrm{Cov}(X,Y) = 0$.
- $\mathrm{Cov}(X,Y) \neq 0 \Rightarrow X, Y$ 不独立

单位化

注意到 $\forall k \in \mathbb{R}$, $\mathrm{Cov}(kX, kY) = k^2 \mathrm{Cov}(X,Y)$, 我们可以考虑 "单位化" 的随机变量

$$X^* := \frac{X - \mathbb{E}X}{\sqrt{\mathrm{Var}(X)}}, \quad Y^* := \frac{Y - \mathbb{E}Y}{\sqrt{\mathrm{Var}(Y)}}.$$

协方差的定义

称

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

为 X, Y 的**相关系数**.

相关系数与线性逼近

问题

对于随机变量 X, Y, Y 是否能表为 X 的线性函数? 若不能, **均方误差**

$$e = \mathbb{E}\Big[Y - (a + bX)\Big]^2$$

何时取得最小值?

限定 $b = 0$, $e = \mathbb{E}(Y - a)^2$ 何时最小?

由

$$\frac{\partial e}{\partial a} = \mathbb{E}2(Y - a) \cdot (-1) = 0 \quad \Rightarrow \quad a = \mathbb{E}Y.$$

此时 e 为 Y 的方差!

原问题的简化

假设 $b \neq 0$, 则 $a = \mathbb{E}(Y - bX)$. 为此, 不失一般性, 我们可以只考虑 $\mathbb{E}Y = \mathbb{E}X = 0$, 否则我们以中心化的

$$\tilde{X} = X - \mathbb{E}X, \quad \tilde{Y} = Y - \mathbb{E}Y$$

代替 X, Y.

$e=\mathbb{E}(Y-bX)^2$ 何时最小?

解

对 b 求导，得

$$\frac{\partial e}{\partial b}=\mathbb{E}2(Y-bX)\cdot(-X)=0 \quad\Rightarrow\quad b=\frac{\mathbb{E}YX}{\mathbb{E}X^2}=\frac{\mathrm{Cov}(X,Y)}{\mathrm{Var}(X)}.$$

此时

$$e_{\min}=\mathrm{Var}(Y)(1-\rho_{XY}^2).$$

定理

- $|\rho_{XY}|\le 1$ (Cauchy 不等式)
- $|\rho_{XY}|=1 \quad\Leftrightarrow Y\overset{\text{a.e.}}{=}a+bX$, a,b 为常数.

正相关与负相关

相关系数的实际意义

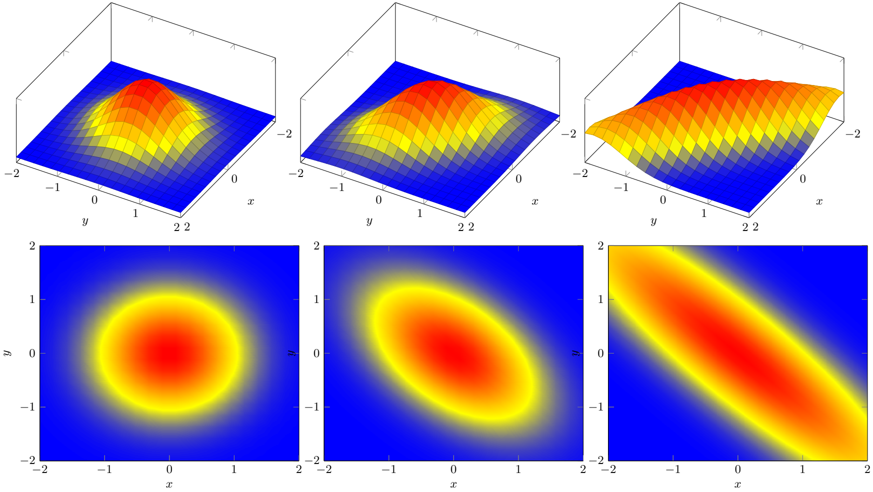
$e_{min} = \mathrm{Var}(Y) \cdot (1 - \rho_{XY}^2)$ 说明:

- $|\rho_{XY}|$ 较大时均方误差 e 较小 $\Rightarrow X, Y$ 之间的线性关系较密切.
- 当 $|\rho_{XY}| = 1$ 时, X, Y 之间以概率 1 为线性关系.
- 当 $|\rho_{XY}|$ 较小时, X, Y 之间的线性关系较弱.

定义

- 当 $0 < \rho_{XY} \le 1$ 时, 称 X 与 Y **正相关**.
- 当 $-1 \le \rho_{XY} < 0$ 时, 称 X 与 Y **负相关**.
- 当 $\rho_{XY} = 0$ 时, 称 X 与 Y **不相关**.

$\mu_1 = \mu_2 = 0,\ \sigma_1 = \sigma_2 = 1$ 的二维正态分布



(a) $\rho = 0$: 不相关

(b) $\rho = -0.5$: 负相关

(c) $\rho = -0.9$: (强) 负相关

例题

设二维随机变量 (X, Y) 的密度函数为

$$f(x,y) = \begin{cases} x+y, & 0 \le x \le 1,\ 0 \le y \le 1, \\ 0, & \text{其它}. \end{cases}$$

求 $\mathbb{E}X$, $\mathbb{E}Y$, $\mathbb{E}XY$ 及 $\mathrm{Cov}(X,Y)$.

解

$$\mathbb{E}X = \int_0^1 \int_0^1 x(x+y)\,dxdy = \frac{7}{12},$$

$$\mathbb{E}Y = \int_0^1 \int_0^1 y(x+y)\,dxdy = \frac{7}{12},$$

$$\mathbb{E}XY = \int_0^1 \int_0^1 xy(x+y)\,dxdy = \frac{1}{3},$$

$$\mathbb{E}XY = \mathbb{E}XY - \mathbb{E}X \cdot \mathbb{E}Y = -\frac{1}{144}.$$

例题

(X,Y) 在 D 上服从均匀分布，求 $\mathrm{Cov}(X,Y)$，ρ_{XY}.

解

区域 D 的面积为 $\frac{1}{2}$. 所以 (X,Y) 的联合密度为 $f(x,y) = 2\cdot \mathbb{1}_D(x,y)$. 易得

$$\mathbb{E}X = \int_0^1 dy \int_0^{1-y} 2x\,dx = \frac{1}{3} = \mathbb{E}Y, \quad (\text{对称性})$$

$$\mathbb{E}XY = \int_0^1 y\cdot dy \int_0^{1-y} 2x\,dx = \frac{1}{4},$$

$$\mathrm{Cov}(X,Y) = \mathbb{E}XY - \mathbb{E}X\cdot\mathbb{E}Y = -\frac{1}{36},$$

$$\mathbb{E}X^2 = \int_0^1 dy \int_0^{1-y} 2x^2\,dx = \frac{1}{6},$$

$$\mathrm{Var}(X) = \mathbb{E}X^2 - (\mathbb{E}X)^2 = \frac{1}{18} = \mathrm{Var}(Y), \quad (\text{对称性}),$$

$$\rho_{XY} = \frac{\mathrm{Cov}(X,Y)}{\sqrt{\mathrm{Var}(X)\cdot\mathrm{Var}(Y)}} = -\frac{1}{2}.$$

例题

设随机变量 (X,Y) 具有概率密度函数 $f(x,y)=\mathbb{1}_{(0,2)^2}(x,y)\frac{1}{8}(x+y)$. 求 $\mathbb{E}X$, $\mathbb{E}Y$ 及 ρ_{XY}.

解

$$
\begin{aligned}
\mathbb{E}X=\mathbb{E}Y &= \int_0^2 dx \int_0^2 dy \cdot \frac{1}{8}x(x+y) = \frac{7}{6},\\
\mathbb{E}X^2=\mathbb{E}Y^2 &= \int_0^2 dx \int_0^2 dy \cdot \frac{1}{8}x^2(x+y) = \frac{5}{3},\\
\mathrm{Var}(X)=\mathrm{Var}(Y) &= \frac{5}{3}-\left(\frac{7}{6}\right)^2 = \frac{11}{36},\\
\mathbb{E}XY &= \int_0^2 dx \int_0^2 dy \cdot \frac{1}{8}xy(x+y) = \frac{4}{3}\\
\mathrm{Cov}(X,Y) &= \frac{4}{3}-\left(\frac{7}{6}\right)^2 = -\frac{1}{36},\\
\rho_{XY} &= \frac{\mathrm{Cov}(X,Y)}{\sqrt{\mathrm{Var}(X)\cdot\mathrm{Var}(Y)}} = -\frac{1}{11}.
\end{aligned}
$$

例题

设随机变量 (X,Y) 服从圆域 $G: x^2+y^2 \le 1$ 上的均匀分布. 试讨论 X,Y 的独立性与相关性.

解

我们已求得

$$f_X(t) = f_Y(t) = \frac{2}{\pi}\sqrt{1-t^2}\cdot \mathbb{1}_{(-1,1)}(t),$$

而 $f(x,y) = \mathbb{1}_G(x,y)\cdot\dfrac{1}{\pi}$. 因此 X,Y 不独立. 又由积分区域和被积函数的对称性,

$$\mathbb{E}X = \mathbb{E}Y = \int_{-1}^{1} x f_X(x)\,dx = 0,$$

$$\mathbb{E}XY = \int_{x^2+y^2\le 1} \frac{xy}{\pi}\,dxdy = 0.$$

因此 $\mathrm{Cov}(X,Y)=0$, $\rho_{XY}=0$, 故 X,Y 不相关.

例题

设 (X,Y) 相互独立, 且都服从 $\mathcal{N}(\mu,\sigma^2)$. 令 $U=aX+bY$, $V=aX-bY$, a,b 为非零常数. 求 ρ_{UV}.

解

$$\operatorname{Cov}(U,V)=\mathbb{E}UV-\mathbb{E}U\cdot\mathbb{E}V=a^2\mathbb{E}X^2-b^2\mathbb{E}Y^2-\Big[a^2(\mathbb{E}X)^2-b^2(\mathbb{E}Y^2)\Big].$$

对于正态分布, 由于 $\mathbb{E}X^2=\operatorname{Var}(X)+\big(\mathbb{E}X\big)^2=\mu^2+\sigma^2=\mathbb{E}Y^2$, 代入得

$$\operatorname{Cov}(U,V)=(a^2-b^2)(\mu^2+\sigma^2)-\Big[a^2\mu^2-b^2\mu^2\Big]=(a^2-b^2)\sigma^2.$$

又 $\operatorname{Var}(U)=\operatorname{Var}(V)=a^2\operatorname{Var}(X)+b^2\operatorname{Var}(Y)=(a^2+b^2)\sigma^2$, 因此

$$\rho_{XY}=\frac{a^2-b^2}{a^2+b^2}.$$

要善用协方差的线性性!

$$\operatorname{Cov}(aX+bY,aX-bY)=a^2\operatorname{Cov}(X,X)-b^2\operatorname{Cov}(Y,Y)=(a^2-b^2)\sigma^2.$$

二维正态分布的相关系数

设 $(X,Y)\sim\mathcal{N}(\mu_1,\mu_2,\sigma_1^2,\sigma_2^2,\rho)$. 则 $\rho_{XY}=\rho$.

证明

不失一般性，我们可以假设 $\mu_1=\mu_2=0$, $\sigma_1^2=\sigma_2^2=1$. 此时 $\rho_{XY}=\mathbb{E}XY$. 我们有

$$f(x,y)=\frac{1}{2\pi\sqrt{1-\rho^2}}\cdot e^{-\frac{1}{2(1-\rho^2)}(x^2+y^2-\rho xy)}=\frac{1}{2\pi\sqrt{1-\rho^2}}\cdot e^{-\frac{1}{2(1-\rho^2)}[(y-\rho x)^2+(1-\rho^2)x^2]}.$$

令 $t=\frac{1}{\sqrt{1-\rho^2}(y-\rho x)^2},u=x$, $J=1$. 则

$$\begin{aligned}\mathbb{E}XY&=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}xyf(x,y)\,dxdy=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}(\sqrt{1-\rho^2}tu+\rho u^2)\frac{1}{2\pi}e^{-\frac{t^2+u^2}{2}}\,dtdu\\&=\frac{\rho}{2\pi}\int_{-\infty}^{\infty}e^{-t^2}dt\cdot\int_{-\infty}^{\infty}u^2e^{-u^2/2}+\frac{\sqrt{1-\rho^2}}{2\pi}\int_{-\infty}^{\infty}te^{-t^2/2}dt\cdot\int_{-\infty}^{\infty}ue^{-u^2/2}du\\&=\frac{\rho}{2\pi}\sqrt{2\pi}\sqrt{2\pi}=\rho\end{aligned}$$

独立与不相关的关系

- X,Y 相互独立 $\Rightarrow$ X,Y 互不相关:
  X,Y 没有任何关系，当然也没有线性关系.
- X,Y 互不相关 $\not\Rightarrow$ X,Y 相互独立:
  X,Y 没有线性关系，但是可能有其它关系.

特例

对于二维正态分布 $(X,Y) \sim \mathcal{N}(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho)$,

$$\begin{aligned} X,Y \text{ 相互独立} \quad &\Leftrightarrow \quad \rho = 0 \\ &\Leftrightarrow \rho_{XY} = 0 \\ &\Leftrightarrow \quad X,Y \text{ 互不相关}. \end{aligned}$$

矩 (Moment) 的定义

对于随机变量 X, Y 称

- $\mathbb{E}X^k$, $k = 1, 2, \ldots$, 为 k **阶原点矩 (k 阶矩)**;
- $\mathbb{E}(X - \mathbb{E}X)^k$, $k = 2, 3, \ldots$, 为 k **阶中心矩**;
- $\mathbb{E}X^kY^l$, $k, l = 1, 2, \ldots$, 为 $k + l$ **阶混合矩**;
- $\mathbb{E}(X - \mathbb{E}X)^k(Y - \mathbb{E}Y)^l$, $k, l = 1, 2, \ldots$, 为 $k + l$ **阶混合中心矩**.

例子

- $\mathbb{E}X$: 1 阶原点矩.
- $\mathrm{Var}(X)$: 2 阶中心矩.
- $\mathrm{Cov}(X, Y)$: 2 阶混合中心矩.

协方差矩阵

对于二维随机变量 $X=(X_1,X_2)$, 记

$$c_{11}=\mathbb{E}(X_1-\mathbb{E}X_1)^2=\mathrm{Var}(X_1), \qquad c_{12}=c_{21}$$
$$c_{21}=\mathbb{E}(X_2-\mathbb{E}X_2)(X_1-\mathbb{E}X_1)=\mathrm{Cov}(X_2,X_1), \quad c_{22}=\mathbb{E}(X_2-\mathbb{E}X_2)^2=\mathrm{Var}(X_2).$$

写作矩阵的形式

$$C:=\mathrm{Cov}(X,X):=\begin{bmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{bmatrix}$$

称为 $X=(X_1,X_2)$ 的**协方差矩阵**.

协方差矩阵的性质

- $C^T=C$, 即 C 为对称阵
- $C\geq 0$, 即 C 为正定 (非负定) 方阵

正定性证明

任取 $v=(a,b)^T\in\mathbb{R}^2$, 则

$$v^TCv=\mathrm{Cov}(aX_1+bX_2,aX_1+bX_2)=\mathrm{Var}(aX_1+bX_2)\geq 0.$$

总结: 协方差和相关系数的性质

1. 对称性: $\mathrm{Cov}(X, Y) = \mathrm{Cov}(Y, X)$
2. 数乘: $\mathrm{Cov}(aX, bY) = ab\,\mathrm{Cov}(X, Y)$
3. (双) 线性性: $\mathrm{Cov}(X + Y, Z) = \mathrm{Cov}(X, Z) + \mathrm{Cov}(Y, Z)$.
4. $\mathrm{Cov}(X, X) = \mathrm{Var}(X)$.
5. $\mathrm{Var}(X \pm Y) = \mathrm{Var}(X) + \mathrm{Var}(Y) \pm 2\,\mathrm{Cov}(X, Y)$.
6. $|\rho_{XY}| \leq 1$
7. $|\rho_{XY}| = 1 \Longleftrightarrow$ 存在常数 $a \neq 0, b$ 使得 $\P(Y = aX + b) = 1$.

作业

- P119: 54, 60
- 补充题:
  1. 设随机变量 X 的概率密度为 $f(x)=\frac{1}{2}e^{-|x|}$, $x\in\mathbb{R}$.
     (1) 求 $\mathbb{E}X$, $\mathrm{Var}(X)$
     (2) X 与 $|X|$ 是否独立? 说明理由.
     (3) X 与 $|X|$ 是否相关? 说明理由.
  2. 设随机变量 (X,Y) 的概率密度为
$$f(x,y)=\begin{cases}\frac{1}{8}(x+y), & 0\le x\le 2, 0\le y\le 2,\\ 0, & \text{其它}.\end{cases}$$
     求 $\mathbb{E}X$, $\mathbb{E}Y$, $\mathrm{Cov}(X,Y)$, ρ_{XY} 以及 $\mathrm{Var}(X+Y)$.
  3. 设随机变量 X 和 Y 独立且均服从 $\mathcal{N}(\mu,\sigma^2)$. 令 $Z=\alpha X+\beta Y$, $W=\alpha X-\beta Y$. 求 $\mathrm{Cov}(Z,W)$, ρ_{ZW}.

思考题

考虑把原本属于 n 个人的 n 封信打乱重新分发，每个人拿到一封信. 令 X 为拿到自己信的总人数. 对于伯努利装错信封问题，我们知道

$$\mathsf{P}(X=0)=1-\frac{1}{1!}+\frac{1}{2!}+\cdots+(-1)^n\cdot\frac{1}{n!}.$$

问：$\mathrm{Var}(X)=?$

条件概率复习

条件频率函数的性质

- $\mathsf{P}(X = x_i \mid Y = y_j) \geq 0$, $i = 1, 2, \ldots$
- $\displaystyle\sum_{i=1}^{\infty} \mathsf{P}(X = x_i \mid Y = y_j) = 1.$

这两条性质说明，条件频率函数也是一种频率函数.

条件密度函数的性质

- $f_{X|Y}(x \mid y) \geq 0.$
- $\displaystyle\int_{-\infty}^{\infty} f_{X|Y}(u \mid y)\, du = 1.$

这两条性质说明，条件密度函数也是一种密度函数.

总结

从条件频率/密度函数出发，我们可以定义条件期望 (conditional expectation) 和条件方差 (conditional variance)

条件期望的定义

给定 $X = x$ 的情况下, Y 的**条件期望**定义为

离散情形

$$\mathbb{E}\big(Y \mid X = x\big) = \sum_y y p_{Y|X}(y \mid x)$$

连续情形

$$\mathbb{E}\big(Y \mid X = x\big) = \int y f_{Y|X}(y \mid x)\, dy.$$

一般函数 $h(Y)$ 的条件期望

离散情形

$$\mathbb{E}\big(h(Y) \mid X = x\big) = \sum_y y p_{Y|X}(y \mid x)$$

连续情形

$$\mathbb{E}\big(h(Y) \mid X = x\big) = \int h(y) f_{Y|X}(y \mid x)\, dy.$$

例

已知 X, Y 的联合分布如下，求 $\mathbb{E}(X \mid Y)$, $\mathbb{E}(Y \mid X)$.

$X \backslash Y$	0	1	2	$p_{i\cdot}$
0	0.1	0.2	0.2	0.5
1	0.3	0.1	0.1	0.5
$p_{\cdot j}$	0.4	0.3	0.3	1.0

解

j	0	1	2
$p_{X\mid Y}(0 \mid j)$	1/4	2/3	2/3
$p_{X\mid Y}(1 \mid j)$	3/4	1/3	1/3
$\mathbb{E}(X \mid Y = j)$	3/4	1/3	1/3
$\mathsf{P}(Y = j)$	0.4	0.3	0.3

i	0	1
$p_{Y\mid X}(0 \mid i)$	1/5	3/5
$p_{Y\mid X}(1 \mid i)$	2/5	1/5
$p_{Y\mid X}(2 \mid i)$	2/5	1/5
$\mathbb{E}(Y \mid X = i)$	6/5	3/5
$\mathsf{P}(X = i)$	0.5	0.5

例

设 (X,Y) 服从 $D=\{(x,y)\mid 0<x<1,\ 0<y<x^2\}$ 上的均匀分布. 求 $\mathbb{E}(Y\mid X)$, $\mathbb{E}(X\mid Y)$.

解

易得 $|D|=\int_0^1 dx\int_0^{x^2} dy=1/3$, 所以 (X,Y) 的密度函数为 $f(x,y)=3\cdot\mathbb{1}_D(x,y)$.

由 $f_X(x)=\mathbb{1}_{(0,1)}(x)\int_0^{x^2}3dy=3x^2\mathbb{1}_{(0,1)}(x)$, 因此 $f_{Y|X}(y\mid x)=\mathbb{1}_{(0,x^2)}(y)\dfrac{1}{x^2}$. 于是

$$\mathbb{E}(Y\mid X=x)=\int_0^{x^2} y\cdot\frac{1}{x^2}\,dy=\frac{x^2}{2},\quad \text{即}\quad \mathbb{E}(Y\mid X)=\frac{X^2}{2}.$$

由 $F_Y(y)=\mathbb{1}_{(0,1)}(y)3(1-\sqrt{y})$, 得 $f_{X|Y}(x\mid y)=\mathbb{1}_{(\sqrt{y},1)}(x)\cdot\dfrac{1}{1-\sqrt{y}}$. 于是

$$\mathbb{E}(X\mid Y=y)=\int_{\sqrt{y}}^1\frac{x}{1-\sqrt{y}}\,dx=\frac{1+\sqrt{y}}{2}.$$

例: 泊松流

考虑 $[0,1]$ 区间上均值为 λ 的泊松流. 令 N 是 $[0,1]$ 上点的个数. 对于 $p<1$, 令 X 是 $[0,p]$ 上点的个数. 计算给定 $X=n$ 的情况下, X 的条件分布的期望.

分析

$\{X=x,\ N=n\}$ 是指, 在 $[0,p)$ 中有 x 个质点, $[p,1]$ 中有 $n-x$ 个质点, 而 $[0,p)$ 上的泊松流与 $[p,1]$ 上的泊松流是独立的. 又注意到在长度为 t 的区间上, 泊松粒子的个数为 $\mathsf{P}(\pi(\lambda t)=k)=e^{-\lambda t}\dfrac{(\lambda t)^k}{k!}$, $k=0,1,2,\ldots$.

解

N 和 X 的联合分布为 $\mathsf{P}(X=x,\ N=n)=\dfrac{(p\lambda)^x e^{-p\lambda}}{x!}\cdot\dfrac{[(1-p)\lambda]^{n-x}e^{-(1-p)\lambda}}{(n-x)!}$.

而 $\mathsf{P}(N=n)=\dfrac{\lambda^n e^{-\lambda}}{n!}$. 因此

$\mathsf{P}(X=x\mid N=n)=\dfrac{n!}{x!(n-x)!}p^x(1-p)^{n-x}\sim \mathrm{Bin}(n,p)$. 因而 X 的条件期望为 np.

二元正态分布

设 $(X, Y) \sim \mathcal{N}(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho)$, 则

$$f_{Y|X}(y \mid x) \sim \mathcal{N}(\mu_2 + \rho\frac{\sigma_2}{\sigma_1}(x - \mu_1), \sigma_2^2(1 - \rho^2)).$$

即给定 X, Y 的条件分布还是一维正态分布.

$$\text{条件期望: } \mu_2 + \rho\frac{\sigma_2}{\sigma_1}(x - \mu_1), \qquad \text{条件方差: } \sigma_2^2(1 - \rho^2).$$

注

- 条件期望是 x 的线性函数
- 条件方差随着 $|\rho|$ 的增加而减少.

怎么理解 $\mathbb{E}(Y \mid X)$ 与 $\mathbb{E}(Y \mid X = x)$

$\mathbb{E}(Y \mid X)$ 是一个随机变量

假设对于 X 范围内的任意 x, 有 $\mathbb{E}(Y \mid X = x)$ 存在, 记

$$\mathbb{E}(Y \mid X = x) := \varphi(x).$$

则 $\mathbb{E}(Y \mid X) := \varphi(X)$ 可看为随机变量 X 的 (可测) 函数, 因此还是一个随机变量.

$\mathbb{E}(Y \mid X)$ 作为随机变量 X 的函数, 也是一个新的随机变量, 可以再求的它的期望与方差:

$$\mathbb{E}\Big[\mathbb{E}(Y \mid X)\Big], \quad \mathrm{Var}\Big[\mathbb{E}(Y \mid X)\Big]$$

1. $\mathbb{E}(c \mid X) = c$
2. (线性性) $\mathbb{E}(aY_1 + bY_2 \mid X) = a\mathbb{E}(Y_1 \mid X) + b\mathbb{E}(Y_2 \mid X)$.

证明

只对连续情形证明. 设 (X, Y_1, Y_2) 的联合密度为 $f(x, y_1, y_2)$. 则

$$
\begin{aligned}
\mathbb{E}(aY_1 + bY_2 \mid X = x) &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} (ay_1 + by_2) f_{(Y_1, Y_2)\mid X}(y_1, y_2 \mid x)\, dy_1 dy_2 \\
&= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} (ay_1 + by_2) \frac{f(x, y_1, y_2)}{f_X(x)}\, dy_1 dy_2 \\
&= a \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} y_1 \frac{f(x, y_1, y_2)}{f_X(x)}\, dy_1 dy_2 + b \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} y_2 \frac{f(x, y_1, y_2)}{f_X(x)}\, dy_1 \\
&= a \int_{-\infty}^{\infty} y_1 \frac{f_{(X, Y_1)}(x, y_1, y_2)}{f_X(x)}\, dy_1 + b \int_{-\infty}^{\infty} y_2 \frac{f_{(X, Y_2)}(x, y_1, y_2)}{f_X(x)}\, dy_2 \\
&= a \int_{-\infty}^{\infty} y_1 f_{Y_1\mid X}(y_1 \mid x)\, dy_1 + b \int_{-\infty}^{\infty} y_2 f_{Y_2\mid X}(y_2 \mid x) dy_2
\end{aligned}
$$

若 X 与 Y 独立，则 $\mathbb{E}(Y \mid X) = \mathbb{E}Y$

证明

由独立性，$f_{Y|X}(y \mid x) = f(x,y)/f_X(x) = f_Y(y)$. 因此

$$\mathbb{E}(Y \mid X = x) = \int_{-\infty}^{\infty} y f_{Y|X}(y \mid x) dy = \int_{-\infty}^{\infty} y f_Y(y)\, dy = \mathbb{E}Y.$$

全期望公式 (Law of total expectation), 重期望公式 (Law of iterated expectation)

$$\mathbb{E}(Y) = \mathbb{E}\Big[\mathbb{E}(Y \mid X)\Big].$$

证明

只证明离散情形, 连续情形类似.

$$\begin{aligned}
\mathbb{E}\Big[\mathbb{E}(Y \mid X)\Big] &= \sum_x \mathbb{E}(Y \mid X = x) p_X(x) \\
&= \sum_x \Big[\sum_y y p_{Y\mid X}(y \mid x)\Big] p_X(x) \\
&= \sum_y y \sum_x p_{Y\mid X}(y \mid x) p_X(x) \\
&= \sum_y y p_Y(y) = \mathbb{E}Y.
\end{aligned}$$

注

Y 的期望可以通先以 X 的条件, 计算 $\mathbb{E}(Y \mid X)$, 再对 X 取期望 (加权平均) 得到. 例如

方差分解

$$\mathrm{Var}(Y) = \mathrm{Var}\left[\mathbb{E}(Y \mid X)\right] + \mathbb{E}\,\mathrm{Var}(Y \mid X).$$

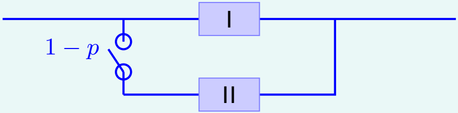
证明

利用重期望公式:

$$\begin{aligned}
\mathrm{Var}(Y) &= \mathbb{E}Y^2 - \left[\mathbb{E}Y\right]^2 \\
&= \mathbb{E}\left[\mathbb{E}(Y^2 \mid X)\right] - \mathbb{E}\left[\left[\mathbb{E}(Y \mid X)\right]^2\right] + \mathbb{E}\left[\left[\mathbb{E}(Y \mid X)\right]^2\right] - \left[\mathbb{E}\left[\mathbb{E}(Y \mid X)\right]\right]^2 \\
&= \mathbb{E}\,\mathrm{Var}(Y \mid X) + \mathrm{Var}\left[\mathbb{E}(Y \mid X)\right].
\end{aligned}$$

例题

假设在系统中，元件和备件平均寿命都是 μ. 如果元件失效，系统自动用其备件替代，但替换出错的概率为 p. 求整个系统的平均寿命.



解

令 T 是系统的寿命，用 X 标记替代是否成功，即

$$X = \begin{cases} 1, & \text{备件替代成功}, \\ 0, & \text{不成功}. \end{cases}$$

则 $\mathbb{E}(T \mid X=1) = 2\mu$, $\mathbb{E}(T \mid X=0) = \mu$. 因此

$$\mathbb{E}T = \mathbb{E}(T \mid X=1)\mathrm{P}(X=1) + \mathbb{E}(T \mid X=0)\mathrm{P}(X=0) = 2\mu(1-p) + \mu p = \mu(2-p).$$

- P120: 67, 77
- 1. 如果 X 和 Y 是两个连续的随机变量, 证明
$$\mathbb{E}\Big[\mathbb{E}(X \mid Y)\Big] = \mathbb{E}X.$$
  2. 设随机变量 (X, Y) 的概率密度为
$$f(x,y) = \begin{cases} ke^{-(x+y)}, & 0 \le y \le x, \\ 0, & \text{其它}. \end{cases}$$
     (1) 计算 $\mathrm{Cov}(X,Y)$, ρ_{XY}
     (2) 计算 $\mathbb{E}(X \mid Y = y)$ 和 $\mathbb{E}(Y \mid X = x)$.
     (3) 推导随机变量 $\mathbb{E}(X \mid X)$ 的概率密度.

生成函数/母函数 (generating function)

设 X 为取值非负整数的随机变量 $p_k = \mathsf{P}(X = k)$. 则其母函数定义为

$$g_X(t) := \sum_{k=0}^{\infty} p_k t^k = \mathbb{E} t^X, \quad t \in (-1, 1].$$

例如 $X \sim \mathrm{Geo}(p)$, 则 $g_X(t) = \frac{p}{1-p} \frac{1}{1-(1-p)t}$. 易见 $g_X(1) = 1, g_X'(1) = \mathbb{E}X$

矩母函数 (moment generating function)

设 X 为随机变量, 假设指数矩存在, 定义其矩母函数为

$$M_X(t) := \mathbb{E}e^{tX} = \sum_{k=0}^{\infty} \frac{\mathbb{E}X^k}{k!}, \quad t \in (-\delta, \delta). \qquad M_X^{(k)}(0) = \mathbb{E}X^k,\ k = 0, 1, \ldots$$

特征函数 (characteristic function)

设 X 为随机变量, 定义其特征函数为

$$\varphi_X(t) := \mathbb{E}e^{itX} \in \mathbb{C}, \quad t \in \mathbb{R}.$$

例如, $X \sim \mathcal{N}(\mu, \sigma^2)$, 则 $\varphi_X(t) = e^{i\mu t - \frac{\sigma^2}{2} t^2}$.