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

1 随机变量的期望
- 引言
- 离散型随机变量的数学期望
- 连续型随机变量的数学期望
- 随机变量函数的数学期望
- 数学期望的基本性质
- 作业

2 方差和标准差

3 协方差和相关系数

4 条件期望和预测

5 矩生成函数

6 近似方法

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

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

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

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

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.

例题

某商店对某种家用电器的销售采用先使用后付款的方式. 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.

解

设出售一台电器的收费额为 Y, 则 Y 的频率函数为

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

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

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

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

泊松分布的期望

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

证明

由 $\mathsf{P}(X = k) = \dfrac{\lambda^k}{k!}e^{-\lambda}$, $k = 0, 1, 2, \ldots$，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 $k\binom{n}{k} = k\dfrac{n!}{k!(n-k)!} = n\dfrac{(n-1)!}{(k-1)!(n-k)!} = n\binom{n-1}{k-1}$.

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

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
$\mathrm{Bin}(n,p)$	$\mathsf{P}(X=k)=\binom{n}{k}p^k(1-p)^{n-k}, k=0,1,\ldots,n$	np
$\pi(\lambda)$	$\mathsf{P}(X=k)=\dfrac{\lambda^k}{k!}e^{-\lambda},\ k=0,1,2,\ldots$	$\lambda.$

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

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

正态分布的数学期望

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

证明

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

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

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

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 $\dfrac{1}{3}ka^3$.

二元情形的推广

定理

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

$$\mathbb{E}Z = \mathbb{E}g(X, Y) = \int^{\infty} \int^{\infty} g(x, y) f(x, y)\, dx dy.$$

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

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

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

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

数学期望的基本性质

1. 设 $a \le X \le b$, 则 $a \le \mathbb{E}X \le b$.
2. 设 c 为常数, 则 $\mathbb{E}(cX) = c\mathbb{E}X$.
3. (线性性) 设 X, Y 为随机变量, $\mathbb{E}(X+Y) = \mathbb{E}X + \mathbb{E}Y$.
4. 设 X, Y 相互独立, 则 $\mathbb{E}XY = \mathbb{E}X \cdot \mathbb{E}Y$.

推论

1. 若 $X = c$, (a.e.), 则 $\mathbb{E}X = c$.
2. 设 $a_1, \ldots, a_n$ 为常数, 则对随机变量 $X_1, \ldots, X_n$, 有
$$\mathbb{E}\left(\sum_{i=1}^{n} a_i X_i\right) = \sum_{i=1}^{n} a_i \mathbb{E}X_i.$$
3. 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,

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

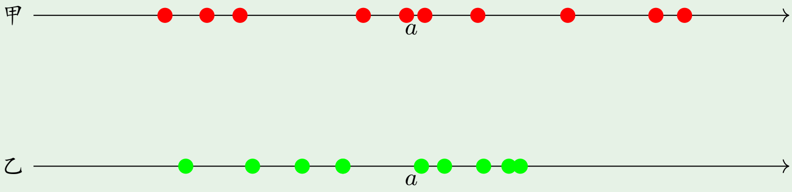
- P116: 6, 15, 20, 21, 31
- 1. 设随机变量 X 的概率密度为
$$f(x) = \begin{cases} e^{-x}, & x > 0, \\ 0, & x \le 0. \end{cases}$$
求以下随机变量的期望:
(1) $Y = 2X$,
(2) $Y = e^{-2X}$.
  2. 设随机变量 (X, Y) 的概率密度为
$$f(x,y) = \begin{cases} 12y^2, & 0 \le y \le x \le 1, \\ 0, & \text{其它}. \end{cases}$$
求 $\mathbb{E}X$, $\mathbb{E}Y$, $\mathbb{E}XY$, $\mathbb{E}(X^2 + Y^2)$.

1. 随机变量的期望
2. 方差和标准差
   - 引言
   - 定义
   - 重要分布的方差
   - 方差的基本性质
   - 作业
3. 协方差和相关系数
4. 条件期望和预测
5. 矩生成函数
6. 近似方法

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

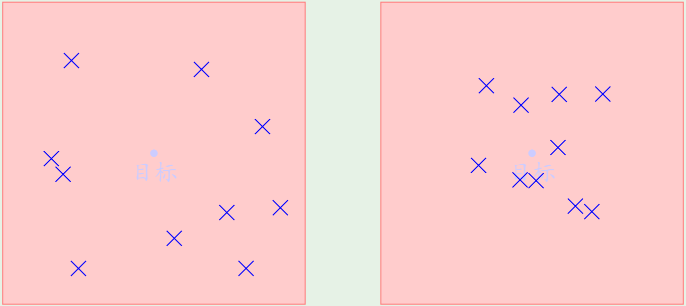
某零件的真实长度为 a，现用甲、乙两台仪器各测量 10 次，将测量结果 X 用坐标上的点表示：



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

定义

对随机变量 X, 若

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

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

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

$$\mathbb{E}X^2=\int_{-1}^{0}x^2(1+x)+\int_{0}^{1}x^2(1-x)\,dx=\frac{1}{6},$$

$$\mathrm{Var}(X)=\mathbb{E}X^2-\left(\mathbb{E}X\right)^2=\frac{1}{6}.$$

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：

$$\begin{array}{c|cccc} X & 0 & 1 & 2 & 3 \\ \hline p & \frac{3}{4} & \frac{9}{44} & \frac{9}{44} & \frac{1}{220} \end{array}.$$

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

$$\mathbb{E}X^* = \mathbb{E}\frac{X - c}{\sqrt{d}} = \frac{1}{\sqrt{d}}\big(\mathbb{E}X - c\big) = 0$$

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

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 $\mathbb{E}\xi = p$, $\mathbb{E}\xi^2 = p^2$, 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

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

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 $\operatorname{Var}(X) = \sigma^2$.

证明

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

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 $\operatorname{Var}(X) = \sigma^2$.

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

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

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

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

4. 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 $\mathsf{P}(|X-\mu|\leq\varepsilon)$？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

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

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

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

证明

由 Markov 不等式,

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

回顾: Markov 不等式

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

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

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 $\mathrm{Var}(Z)$.

1 随机变量的期望
2 方差和标准差
3 协方差和相关系数
4 条件期望和预测
5 矩生成函数
6 近似方法

1 随机变量的期望

2 方差和标准差

3 协方差和相关系数

4 条件期望和预测

5 矩生成函数

6 近似方法

1 随机变量的期望

2 方差和标准差

3 协方差和相关系数

4 条件期望和预测

5 矩生成函数

6 近似方法

1 随机变量的期望

2 方差和标准差

3 协方差和相关系数

4 条件期望和预测

5 矩生成函数

6 近似方法