概率论和数理统计

2025-2026 秋冬

黄炜
(hwmath2003@zju.edu.cn)

浙江大学数学科学学院

Chapter 4 随机变量的数字特征

关键词 (Keywords):

- (数学) 期望 ((Mathematical) Expectation)
- 方差、变异系数 (Variance, Coefficient of Variation)
- 协方差、相关系数 (Covariance, Correlation Coefficient)
- 其他数字特征, 如: 矩 (Moment), 分位数 (Quantile), 等等

在一些实际问题中，我们需要了解随机变量的分布函数外，更关心的是随机变量的某些特征.

Example 1

- 在评定某地区粮食产量的水平时，最关心的是平均产量.
- 在检查一批棉花的质量时，既需要注意纤维的平均长度，又需要注意纤维长度与平均长度的偏离程度.
- 考察杭州市区居民的家庭收入情况，我们既要知家庭的年平均收入，又要研究贫富之间的差异程度.

§4.1 数学期望 (Mathematical Expectation)

Example 2 (谁的技术比较好?)

有甲、乙两个射手，他们的某次射击成绩如下表所示:

甲射手	击中环数	8	9	10		乙射手	击中环数	8	9	10
	次数	10	80	10			次数	20	65	15

试问哪个射手技术较好?

Solution

甲的平均成绩: $\dfrac{8\times10+9\times80+10\times10}{100}=9$; 乙的平均成绩: $\dfrac{8\times20+9\times65+10\times15}{100}=8.95.$

所以甲的成绩好于乙的成绩.

设射手的平均击中的环数为 $\overline{x}$, 即

$$\overline{x}=\frac{1}{n}\sum_k x_k n_k=\sum_k x_k\frac{n_k}{n}=\sum_k x_k f_k,\quad \text{其中} f_k \text{为击中 } x_k \text{ 环的频率.}$$

设射手的射击环数为 X, f_k 的稳定值为 $p_k=\mathrm{P}(X=x_k)$.
因此 $\overline{x}$ 的稳定值为 $\sum\limits_k x_k p_k$ (当 n 充分大时).

即, $\sum\limits_k x_k f_k$ 稳定于 $\sum\limits_k x_k p_k$.
后者是一个确定值, 反映了射手的平均水平, 也就是给出了 X 的平均值.

§4.1 数学期望 (Mathematical Expectation)

Definition 1 (离散型随机变量的数学期望)

设离散型随机变量 X 的分布律为 $P(X=x_k)=p_k, k=1,2,\ldots$. 若级数 $\sum_{k=1}^{+\infty} x_k p_k$ 绝对收敛 (即 $\sum_{k=1}^{+\infty} |x_k| p_k < +\infty$), 则称 X 的 (数学) 期望*存在*, 其值为 $\sum_{k=1}^{+\infty} x_k p_k$, 即随机变量 X 的*数学期望 (mathematical expectation)*(通常记为 $E(X)$) 为

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

Definition 2 (连续型随机变量的数学期望)

设连续型随机变量 X 的概率密度函数为 $f(x)$. 若积分 $\int_{-\infty}^{+\infty} xf(x)\mathrm{d}x$ 绝对收敛 (即 $\int_{-\infty}^{+\infty} |x|f(x)\mathrm{d}x < +\infty$), 则称 X 的 (数学) 期望存在, 其值为 $\int_{-\infty}^{+\infty} xf(x)\mathrm{d}x$, 即随机变量 X 的数学期望 (mathematical expectation)(通常记为 $E(X)$) 为

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

数学期望也常简称为期望, 又称为均值 (mean).

Example 3

设随机变量 X 的分布律为

$$\mathsf{P}\left(X=(-1)^{k+1}\frac{3^k}{k}\right)=\frac{2}{3^k},\quad k=1,2,\ldots.$$

证明: X 不存在数学期望.

Proof

注意到

$$\sum_{k=1}^{+\infty}|x_k|p_k=\sum_{k=1}^{+\infty}\frac{3^k}{k}\cdot\frac{2}{3^k}=\sum_{k=1}^{+\infty}\frac{2}{k}=+\infty.$$

即该无穷级数是发散的. 因此由期望定义知, X 不存在数学期望.

Example 4

设随机变量 X 的概率密度函数为

$$f(x) = \frac{1}{\pi(1+x^2)}, \quad -\infty < x < +\infty.$$

证明: X 不存在数学期望.

Proof

由于

$$\int_{-\infty}^{+\infty} |x| f(x) \mathrm{d}x = \int_{-\infty}^{+\infty} \frac{|x|}{\pi(1+x^2)} \mathrm{d}x = \int_{0}^{+\infty} \frac{2x}{\pi(1+x^2)} \mathrm{d}x = \frac{1}{\pi} \ln(1+x^2)\Big|_0^{+\infty} = +\infty.$$

即该积分是发散的. 因此由定义知, X 不存在数学期望.

Example 5

一种常见的赌博游戏, 其规则为: 投掷一颗均匀的骰子, 赌客猜精确的骰子点数, 凡猜中者以 1 比 5 得到奖金, 否则其押金归庄家所有. 问此规则对庄家还是赌客更有利?

Solution

由题意知, 猜中点数的概率为 1/6. 不妨设一赌徒押了 10 元, 那么根据规则, 他收回 50 元的可能性为 1/6, 血本无归的可能性为 5/6. 因此, 经过一次赌博, 他能 "期望" 得到的金额 (即平均所得) 为

$$50 \times \frac{1}{6} + 0 \times \frac{5}{6} = \frac{50}{6} \approx 8.33(\text{元}) < 10(\text{元}).$$

所以可得此规则对庄家有利.

“期望”名称的起源——分赌本问题

Example 6

17 世纪中叶，有甲、乙两赌徒，赌技相同，各出赌注 50 法郎．两人约定无平局，谁先赢 3 局就获得全部赌注．现因一些原因，当甲赢 2 局、乙赢 1 局时，赌博中止．问：如何分赌本才算公平?

Solution

均分，对甲不公平；全部归甲，对乙不公平!

帕斯卡

balabalabalabala

$\rightleftarrows$

费马

按已赌局数和再赌下去的“期望”分!

Solution (Cont.)

因为最多再赌两局, 必然可以分出胜负, 共三种情况:

1. 第四局甲赢; (1/2)
2. 第四局乙赢, 第五局甲赢; (1/4)
3. 第四局乙赢, 第五局乙赢. (1/4)

所以甲获得 100 法郎的可能性为 3/4, 乙获得 100 法郎的可能性为 1/4.

Solution (Cont.)

因此，甲的所得 X 可视为一个可能取值为 0 或 100 的随机变量，其分布律为

X	0	100
P	1/4	3/4

那么甲的“期望”所得是：

$$100 \times 3/4 + 0 \times 1/4 = 75(\text{占总赌本的 } 3/4).$$

所以最终决定：甲分总赌本的 3/4，乙分总赌本的 1/4.

这种分法既考虑了已赌局数，又包含了再赌下去的“期望”，因此更为合理一些.

常见分布的数学期望 (附表一, 第二版 Page 362, 第一版 Page 293)

(1) 两点分布 $0-1(p)$

Example 7

设随机变量 X 服从参数为 p 的 $0-1$ 分布, 其分布律为:

$$\mathrm{P}(X=0)=1-p, \quad \mathrm{P}(X=1)=p,$$

则 $\mathrm{E}(X)=0\times(1-p)+1\times p=p$.

$0-1$ 分布的期望即为它的参数 p.

(2) 二项分布 $B(n,p)$

Example 8

设随机变量 X 服从二项分布 $B(n,p)$，其分布律为:

$$\mathrm{P}(X=k)=\mathrm{C}_n^k p^k(1-p)^{n-k},\quad k=0,1,\ldots,n,$$

则

$$\begin{aligned}\mathrm{E}(X)&=\sum_{k=0}^{n}k\cdot\mathrm{C}_n^k p^k(1-p)^{n-k}=\sum_{k=1}^{n}n\cdot\mathrm{C}_{n-1}^{k-1}p^k(1-p)^{n-k}\\&=np\sum_{k=1}^{n}\mathrm{C}_{n-1}^{k-1}p^{k-1}(1-p)^{(n-1)-(k-1)}=np\sum_{j=0}^{n-1}\mathrm{C}_{n-1}^{j}p^{j}(1-p)^{(n-1)-j}=np.\end{aligned}$$

二项分布的期望即为它的两个参数之乘积 np.

(3) 泊松分布 $P(\lambda)$

Example 9

设随机变量 X 服从泊松分布 $P(\lambda)$, 其分布律为:

$$\mathsf{P}(X=k)=\frac{\lambda^k \mathrm{e}^{-\lambda}}{k!}, \quad k=0,1,\ldots,$$

则

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

泊松分布的期望即为它的参数 λ.

(4) 几何分布 $Geom(p)$

Example 10

设随机变量 X 服从几何分布 $Geom(p)$, 其分布律为:

$$\mathsf{P}(X=k)=p(1-p)^{k-1},\quad k=1,2,\ldots,$$

则

$$\mathsf{E}(X)=\sum_{k=1}^{+\infty}kp(1-p)^{k-1}=p\cdot\sum_{k=1}^{+\infty}\Big((1-p)^k\Big)'=p\cdot\left(\frac{x}{1-x}\right)'\Big|_{x=1-p}=p\cdot\frac{1}{p^2}=\frac{1}{p}.$$

几何分布的期望即为它的参数的倒数 $1/p$.

(5) 均匀分布 $U(a, b)$

Example 11

设随机变量 X 服从均匀分布 $U(a, b)$, 其概率密度为:

$$f(x) = \begin{cases} \dfrac{1}{b-a}, & a < x < b, \\ 0, & \text{其他}, \end{cases}$$

则

$$\mathsf{E}(X) = \int_{-\infty}^{+\infty} x f(x) \mathrm{d}x = \int_a^b \frac{x}{b-a} \mathrm{d}x = \frac{a+b}{2}.$$

均匀分布的期望即为它的支撑区间 (a, b) 的中点 $(a+b)/2$.

(6) 指数分布 $E(\lambda)$

Example 12

设随机变量 X 服从指数分布 $E(\lambda)$, 其概率密度为:

$$f(x)=\begin{cases}\lambda \mathrm{e}^{-\lambda x}, & x>0,\\ 0, & x\leq 0,\end{cases}$$

则

$$\begin{aligned}\mathsf{E}(X)&=\int_{-\infty}^{+\infty} xf(x)\mathrm{d}x=\int_{0}^{+\infty} x\cdot\lambda \mathrm{e}^{-\lambda x}\mathrm{d}x\\ &=-x\mathrm{e}^{-\lambda x}\Big|_0^{+\infty}+\int_0^{+\infty}\mathrm{e}^{-\lambda x}\mathrm{d}x=-\frac{1}{\lambda}\mathrm{e}^{-\lambda x}\Big|_0^{+\infty}=\frac{1}{\lambda}.\end{aligned}$$

指数分布的期望即为它的参数的倒数 $1/\lambda$.

若已知 X 服从期望为 2 的指数分布，则其概率密函数为

$$f(x)=\begin{cases}\frac{1}{2}\mathrm{e}^{-x/2}, & x>0,\\ 0, & x\leq 0,\end{cases}$$

(7) 正态分布 $N(\mu,\sigma^2)$

Example 13

设随机变量 Z 服从标准正态分布 $N(0,1)$，则其期望为0.
设随机变量 X 服从正态分布 $N(\mu,\sigma^2)$，其概率密度函数为:

$$f(x)=\frac{1}{\sqrt{2\pi}\sigma}\mathrm{e}^{-\frac{(x-\mu)^2}{2\sigma^2}},\quad -\infty<x<+\infty,$$

则

$$\mathsf{E}(X)=\int_{-\infty}^{+\infty}xf(x)\mathrm{d}x=\int_{-\infty}^{+\infty}\frac{x}{\sqrt{2\pi}\sigma}\mathrm{e}^{-\frac{(x-\mu)^2}{2\sigma^2}}\mathrm{d}x=\mu+\int_{-\infty}^{+\infty}\frac{x-\mu}{\sqrt{2\pi}\sigma}\mathrm{e}^{-\frac{(x-\mu)^2}{2\sigma^2}}\mathrm{d}x$$

$$\overset{令 t=\frac{x-\mu}{\sigma}}{=\!=\!=\!=}\mu+\frac{\sigma}{\sqrt{2\pi}}\int_{-\infty}^{+\infty}t\mathrm{e}^{-\frac{t^2}{2}}\mathrm{d}t=\mu+0=\mu.$$

正态分布的期望即为它的位置参数 μ.

Example 14

某厂生产的电子产品的寿命 (单位: 年) 服从均值为 3 的指数分布. 若每件产品的生产成本为 350 元, 出售价格为 500 元, 且该厂向顾客承诺, 如果售出一年之内发生故障, 则可免费调换一件, 如果在三年之内发生故障, 则予以免费维修, 维修成本为 50 元. 在这样的价格体系下, 问该厂每售出一件产品, 其平均净收入是多少?

Solution

记某件产品的寿命为 X (年), 则 $X \sim E(1/3)$, 售出一件产品的净收入为 Y (元), 则

$$Y = \begin{cases} 500 - 350 \times 2 = -200, & 若\ 0 < X \leq 1; \\ 500 - 350 - 50 = 100, & 若\ 1 < X \leq 3; \\ 500 - 350 = 150, & 若\ X > 3. \end{cases}$$

Solution (Cont.)

由于 $X \sim E(1/3)$, 那么 Y 的分布律为

$$\mathrm{P}(Y=-200)=\mathrm{P}(0<X\le 1)=\int_0^1 \frac{1}{3}\mathrm{e}^{-\frac{x}{3}}\mathrm{d}x = 1-\mathrm{e}^{-\frac{1}{3}},$$

$$\mathrm{P}(Y=100)=\mathrm{P}(1<X\le 3)=\int_1^3 \frac{1}{3}\mathrm{e}^{-\frac{x}{3}}\mathrm{d}x = \mathrm{e}^{-\frac{1}{3}}-\mathrm{e}^{-1},$$

$$\mathrm{P}(Y=150)=\mathrm{P}(X>3)=\int_3^{+\infty} \frac{1}{3}\mathrm{e}^{-\frac{x}{3}}\mathrm{d}x = \mathrm{e}^{-1}.$$

所以售出一件产品的平均净收入为

$$\begin{aligned}\mathrm{E}(Y) &= -200\times(1-\mathrm{e}^{-\frac{1}{3}})+100\times(\mathrm{e}^{-\frac{1}{3}}-\mathrm{e}^{-1})+150\times\mathrm{e}^{-1}\\ &= -200+300\mathrm{e}^{-\frac{1}{3}}+50\mathrm{e}^{-1}\approx 33.35(\text{元}).\end{aligned}$$

Example 15

设一台机器一天内发生故障的概率为 0.2, 机器发生故障时全天停工. 若一周 5 个工作日里无故障可获利 10 万元, 发生 1 次故障获利 5 万元, 发生 2 次故障获利 0 元, 发生 3 次或以上故障亏损 2 万元. 求一周内的期望利润.

Solution

设 X 表示一周 5 天内机器发生故障天数, 则 $X \sim B(5, 0.2)$. 设 Y 表示一周内所获利润, 则 Y 的分布律为

$$\begin{aligned}
\mathsf{P}(Y=10) &= \mathsf{P}(X=0) = 0.8^5 = 0.32768,\\
\mathsf{P}(Y=5) &= \mathsf{P}(X=1) = \mathsf{C}_5^1 0.2 \times 0.8^4 = 0.4096,\\
\mathsf{P}(Y=0) &= \mathsf{P}(X=2) = \mathsf{C}_5^2 0.2^2 \times 0.8^3 = 0.2048,\\
\mathsf{P}(Y=-2) &= \mathsf{P}(X=3) + \mathsf{P}(X=4) + \mathsf{P}(X=5) = 0.05792.
\end{aligned}$$

$$\Rightarrow \mathsf{E}(Y) = 10 \times 0.32768 + 5 \times 0.4096 + 0 \times 0.2048 - 2 \times 0.05792 = 5.20896(\text{万元}).$$

Example 16

设 X 与 Y 独立同分布，均服从 $E(\lambda), \lambda > 0$. 令 $N = \min(X, Y), M = \max(X, Y)$，求 $\mathrm{E}(N), \mathrm{E}(M)$.

Solution

由题意可知 X 与 Y 的分布函数均为 $F(x) = \begin{cases} 1 - \mathrm{e}^{-\lambda x}, & x > 0, \\ 0, & x \leq 0. \end{cases}$ 故 N 的分布函数为 $F_N(x) = 1 - [1 - F(x)]^2$，因此 N 的概率密度函数为

$$f_N(x) = \begin{cases} 2\lambda \mathrm{e}^{-2\lambda x}, & x > 0, \\ 0, & x \leq 0. \end{cases}$$

即 $N \sim E(2\lambda)$，所以 $\mathrm{E}(N) = \dfrac{1}{2\lambda}$.

Solution (Cont.)

M 的分布函数为 $F_M(x) = [F(x)]^2$, 因此 M 的概率密度函数为

$$f_M(x) = \begin{cases} 2\lambda e^{-\lambda x} - 2\lambda e^{-2\lambda x}, & x > 0, \\ 0, & x \le 0. \end{cases}$$

所以

$$\mathsf{E}(M) = \int_{-\infty}^{+\infty} x f_M(x) \mathrm{d}x = 2\int_0^{+\infty} \lambda x e^{-\lambda x} \mathrm{d}x - \int_0^{+\infty} 2\lambda x e^{-2\lambda x} \mathrm{d}x = \frac{2}{\lambda} - \frac{1}{2\lambda} = \frac{3}{2\lambda}.$$

* 一般定义 (仅作参考)

设随机变量 X 的分布函数为 $F(x)$, 若 $\int_{-\infty}^{+\infty}|x|dF(x)<\infty$ (斯梯尔吉斯 (Stieltjes) 积分), 则随机变量 X 的期望存在, 其值为 $\int_{-\infty}^{+\infty}xdF(x)$. 一般可用以下公式进行计算

$$\mathrm{E}(X)=\int_0^{+\infty}(1-F(x))\mathrm{d}x-\int_{-\infty}^{0}F(x)\mathrm{d}x.$$

特别地, 当 X 为非负随机变量 (即 $\mathrm{P}(X\geq 0)=1$) 时, 有 $\mathrm{E}(X)=\int_0^{+\infty}(1-F(x))\mathrm{d}x.$

当 X 为取非负整数值的随机变量时, 有 $\mathrm{E}(X)=\sum_{k=1}^{+\infty}\mathrm{P}(X\geq k).$

随机变量函数的数学期望

Example 17

设随机变量 X 的概率分布律如下，令 $Y=X^2$，求 Y 的数学期望 $\mathrm{E}(Y)$.

X	-1	0	1	2
P	0.1	0.4	0.2	0.3

Solution

由题意可知 Y 的概率分布律为

Y	0	1	4
P	0.4	0.3	0.3

所以 $\mathrm{E}(Y)=0\times 0.4+1\times 0.3+4\times 0.3=1.5.$

在上例中，事实上，Y 的期望还可以通过如下方式计算：

$$\mathsf{E}(Y) = (-1)^2 \times 0.1 + 0^2 \times 0.4 + 1^2 \times 0.2 + 2^2 \times 0.3 = 1.5.$$

也就是说，对于随机变量 X 的函数有时可以根据 X 的分布以及函数表达式来直接得到其期望.

说明：
下文中，若无特殊说明，当使用 $\mathsf{E}(\cdot)$ 时，则说明对应的随机变量其期望是存在的.

Theorem 1

设 Y 是随机变量 X 的函数: $Y = g(X)$, 其中 $g(\cdot)$ 是连续函数.

- X 是离散型随机变量, 它的分布律为 $P(X = x_k) = p_k, k = 1, 2, \ldots$. 若 $\sum_{k=1}^{+\infty} g(x_k)p_k$ 绝对收敛, 则有
$$E(Y) = E[g(X)] = \sum_{k=1}^{+\infty} g(x_k)p_k.$$
- X 是连续型随机变量, 它的概率密度为 $f(x)$. 若 $\int_{-\infty}^{+\infty} g(x)f(x)\mathrm{d}x$ 绝对收敛, 则有
$$E(Y) = E[g(X)] = \int_{-\infty}^{+\infty} g(x)f(x)\mathrm{d}x.$$

Example 18

设 $X\sim U(-1,2)$，令 $Y=\max(X,0)$，求 $\mathsf{E}(Y)$.

Solution

设 $g(x)=\max(x,0)$，则 $Y=g(X)$ 的期望为

$$
\begin{aligned}
\mathsf{E}(Y) &= \mathsf{E}[g(X)]=\int_{-\infty}^{+\infty} g(x)f_X(x)\mathrm{d}x \\
&= \int_{-\infty}^{+\infty}\max(x,0)f_X(x)\mathrm{d}x=\int_{-1}^{2}\max(x,0)\cdot\frac{1}{3}\mathrm{d}x+0 \\
&= \int_{-1}^{0}0\cdot\frac{1}{3}\mathrm{d}x+\int_{0}^{2}x\cdot\frac{1}{3}\mathrm{d}x=\frac{2}{3}.
\end{aligned}
$$

定理1的重要意义在于求 $\mathsf{E}(Y)$ 时，不必算出 Y 的分布律或概率密度，而只要利用 X 的分布律或概率密度即可进行计算.

这一结果也可以推广到两个或两个以上随机变量的函数的情况.

Theorem 2

设 Z 是随机变量 X, Y 的函数: $Z = h(X, Y)$, 其中 $h(\cdot,\cdot)$ 是连续函数.

- 若二维离散型随机变量 (X, Y) 的分布律为 $P(X = x_i, Y = y_j) = p_{ij}$, $i, j = 1, 2, \ldots$, 则有

$$E(Z) = E[h(X, Y)] = \sum_{i=1}^{+\infty}\sum_{j=1}^{+\infty} h(x_i, y_j)p_{ij},$$

这里设上式右边的级数绝对收敛.

Theorem 2 (Cont.)

- 若二维连续型随机变量 (X, Y) 的概率密度为 $f(x, y)$, 则有

$$\mathrm{E}(Z) = \mathrm{E}[h(X, Y)] = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} h(x, y)f(x, y)\mathrm{d}x\mathrm{d}y,$$

这里设上式右边的积分绝对收敛.

特别地,

$$\mathrm{E}(X) = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} xf(x, y)\mathrm{d}x\mathrm{d}y, \quad \mathrm{E}(Y) = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} yf(x, y)\mathrm{d}x\mathrm{d}y.$$

Example 19

设二维随机变量 (X, Y) 的联合分布律为

$X \backslash Y$	0	1	2
0	0.1	0.25	0.15
1	0.15	0.2	0.15

求随机变量 $Z = \sin \dfrac{\pi(X+Y)}{2}$ 的数学期望.

Solution

由定理2可知

$$
\begin{aligned}
\mathrm{E}(Z) &= \mathrm{E}\left[\sin\frac{\pi(X+Y)}{2}\right] \\
&= \sin\frac{\pi(0+0)}{2}\times 0.1+\sin\frac{\pi(0+1)}{2}\times 0.25+\sin\frac{\pi(0+2)}{2}\times 0.15+ \\
&\qquad \sin\frac{\pi(1+0)}{2}\times 0.15+\sin\frac{\pi(1+1)}{2}\times 0.2+\sin\frac{\pi(1+2)}{2}\times 0.15 \\
&= 0.25.
\end{aligned}
$$

Example 20

设随机变量 (X, Y) 的概率密度为 $f(x,y)=\begin{cases}\dfrac{3}{2x^3y^2}, & \dfrac{1}{x}<y<x, x>1,\\ 0, & \text{其他}.\end{cases}$ 求数学期望 $\mathsf{E}(Y)$ 和 $\mathsf{E}\left(\dfrac{1}{XY}\right)$.

Solution

$$\mathsf{E}(Y)=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} yf(x,y)\mathrm{d}x\mathrm{d}y=\int_{1}^{+\infty}\left[\int_{\frac{1}{x}}^{x}\frac{3}{2x^3y}\mathrm{d}y\right]\mathrm{d}x$$

$$=\int_{1}^{+\infty}\frac{3\ln x}{x^3}\mathrm{d}x=-\frac{3}{2}\frac{\ln x}{x^2}\Big|_{1}^{+\infty}+\frac{3}{2}\int_{1}^{+\infty}\frac{1}{x^3}\mathrm{d}x=\frac{3}{4}.$$

Solution (Cont.)

$$\mathsf{E}\Big(\frac{1}{XY}\Big) = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\frac{1}{xy}f(x,y)\mathrm{d}x\mathrm{d}y = \int_{1}^{+\infty}\left[\int_{\frac{1}{x}}^{x}\frac{3}{2x^4y^3}\mathrm{d}y\right]\mathrm{d}x$$
$$= \int_{1}^{+\infty}\frac{3}{2x^4}\Big(-\frac{1}{2y^2}\Big)\Big|_{\frac{1}{x}}^{x}\mathrm{d}x = \frac{3}{4}\int_{1}^{+\infty}\Big(-\frac{1}{x^6}+\frac{1}{x^2}\Big)\mathrm{d}x$$
$$= = \frac{3}{4}\Big(-\frac{1}{5}+1\Big) = \frac{3}{5}.$$

思考: 如果先求 $f_Y(y)$ 再求 $\mathsf{E}(Y)$, 或先求 $\dfrac{1}{XY}$ 的密度函数再求其期望, 复杂度是否更高?

Example 21

某商店经销某种商品，每周进货量 X 与需求量 Y 是相互独立的随机变量，且都服从区间 $[10, 20]$ 上的均匀分布. 商店每售出一单位商品可获利 1000 元; 若需求量超过进货量，商店可从他处调剂供应，这时每单位商品可获利 500 元. 试计算出此商店经销该种商品每周所获得利润的数学期望.

Solution

设 Z 表示此商店经销该种商品每周所获得利润，则

$$Z = g(X, Y) = \begin{cases} 1000\,Y, & Y \le X, \\ 1000X + 500(Y - X) = 500(X + Y), & Y > X. \end{cases}$$

又 X 和 Y 相互独立，因此 (X, Y) 的联合概率密度函数为

$$f(x, y) = f_X(x) f_Y(y) = \begin{cases} 1/100, & 10 \le x, y \le 20, \\ 0, & \text{其他}. \end{cases}$$

Solution (Cont.)

于是

$$\begin{aligned}
\mathrm{E}(Z) &= \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} g(x,y)f(x,y)\mathrm{d}x\mathrm{d}y \\
&= \int_{10}^{20}\left(\int_{10}^{x} 10y\mathrm{d}y\right)\mathrm{d}x + \int_{10}^{20}\left(\int_{x}^{20} 5(x+y)\mathrm{d}y\right)\mathrm{d}x \\
&= \int_{10}^{20}\left[5x^2 - 500 + 5x(20-x) + \frac{5}{2}(20^2 - x^2)\right]\mathrm{d}x \\
&= \int_{10}^{20}\left(-\frac{5}{2}x^2 + 100x + 500\right)\mathrm{d}x = \frac{42500}{3}(\text{元}) \approx 14166.67(\text{元}).
\end{aligned}$$

Example 22 (极值问题)

设按季节出售的某种应时产品的销售量 X (单位: 吨) 是一个服从 $[5, 10]$ 上的均匀分布的随机变量. 若销售出一吨产品可盈利 2 万元; 但若在销售季节未能售完, 造成积压, 则每吨产品将会净亏损 0.5 万元. 若该厂家需要提前生产该种商品, 为使厂家能获得最大的期望利润, 问在该季生产多少吨产品最为合适?

Solution

设应在该季生产 a 吨产品 $(5 \leq a \leq 10)$, 所获利润为 Y 万元. 则 Y 依赖于销售量 X 及产量 a, 即

$$Y = g(X, a) = \begin{cases} 2a, & X \geq a, \\ 2X - 0.5(a - X) = 2.5X - 0.5a, & X < a. \end{cases}$$

Solution (Cont.)

于是

$$
\begin{aligned}
\mathsf{E}(Y) &= \int_{-\infty}^{+\infty} g(x,a) f_X(x)\mathrm{d}x = \int_5^{10} g(x,a)\frac{1}{5}\mathrm{d}x \\
&= \int_5^a \frac{2.5x-0.5a}{5}\mathrm{d}x + \int_a^{10}\frac{2a}{5}\mathrm{d}x = -\frac{a^2}{4}+\frac{9a}{2}-\frac{25}{4}.
\end{aligned}
$$

令 $\frac{\mathrm{d}}{\mathrm{d}a}\mathsf{E}(Y)=0$ 得 $a=9$, 又由于此时 $\frac{\mathrm{d}^2}{\mathrm{d}a^2}\mathsf{E}(Y)=-\frac{1}{2}<0$, 所以 $a=9$ 时, $\mathsf{E}(Y)$ 达到最大. 因此, 在该季生产 9 吨产品最为合适.

数学期望的性质

(1) 设 C 是常数，则有 $\mathsf{E}(C) = C$.

Proof

因为 C 为常数，随机变量 X 满足 $\mathsf{P}(X = C) = 1$，则 $\mathsf{E}(X) = \mathsf{E}(C) = 1 \times C = C$.

(2) 设 X 是一个随机变量，C 是常数，则有 $\mathsf{E}(CX) = C \cdot \mathsf{E}(X)$.

Proof (以连续型随机变量为例)

$$\mathsf{E}(CX) = \int_{-\infty}^{+\infty} Cxf(x)\mathrm{d}x = C\int_{-\infty}^{+\infty} xf(x)\mathrm{d}x = C \cdot \mathsf{E}(X).$$

(3) 设 X, Y 是两个随机变量, 则有 $\mathsf{E}(X+Y)=\mathsf{E}(X)+\mathsf{E}(Y)$.

Proof (以连续型随机变量为例)

$$\begin{aligned}\mathsf{E}(X+Y) &= \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}(x+y)f(x,y)\mathrm{d}x\mathrm{d}y \\ &= \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}xf(x,y)\mathrm{d}x\mathrm{d}y+\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}yf(x,y)\mathrm{d}x\mathrm{d}y=\mathsf{E}(X)+\mathsf{E}(Y)\end{aligned}$$

综合上面三条可以推广得到任意有限个随机变量线性组合的情形.

Proposition 1

若 n *个随机变量* $X_1, X_2, \ldots, X_n (n \geq 1)$ *的数学期望都存在, 则对任意* $n+1$ *个实数* $c_0, c_1, c_2, \ldots, c_n$, *随机变量* $c_0 + \sum_{i=1}^{n} c_i X_i$ *的数学期望也存在, 且*

$$\mathsf{E}\left(c_0 + \sum_{i=1}^{n} c_i X_i\right) = c_0 + \sum_{i=1}^{n} c_i \mathsf{E}(X_i).$$

(4) 设 X, Y 是两个相互独立的随机变量, 则有 $\mathsf{E}(XY) = \mathsf{E}(X)\mathsf{E}(Y)$.

Proof (以连续型随机变量为例)

$$
\begin{aligned}
\mathsf{E}(XY) &= \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} xyf(x, y)\mathrm{d}x\mathrm{d}y \\
&= \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} xyf_X(x)f_Y(y)\mathrm{d}x\mathrm{d}y \\
&= \int_{-\infty}^{+\infty} xf_X(x)\mathrm{d}x\int_{-\infty}^{+\infty} yf_Y(y)\mathrm{d}y = \mathsf{E}(X)\mathsf{E}(Y)
\end{aligned}
$$

注: 此性质的条件可以减弱.

性质 (4) 可以推广到任意有限个相互独立的随机变量之积的情形.

Proposition 2

若 n 个随机变量 $X_1, X_2, \ldots, X_n (n \geq 1)$ 相互独立, 且数学期望都存在, 则 $\prod_{i=1}^{n} X_i$ 的数学期望也存在, 且

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

下面从另一方式来得到：正态分布 $N(\mu,\sigma^2)$ 的均值为 μ

Example 23

设随机变量 X 服从正态分布 $N(\mu,\sigma^2), -\infty<\mu<+\infty, \sigma>0$.

由于 $Z=\dfrac{X-\mu}{\sigma}\sim N(0,1)$, 故有 $X=\sigma Z+\mu$, 其中 Z 服从标准正态分布.

已知 $\mathsf{E}(Z)=0$, 于是

$$\mathsf{E}(X)=\mathsf{E}(\sigma Z+\mu)=\sigma\mathsf{E}(Z)+\mu=\mu.$$

下面从另一方式来得到: 二项分布 $B(n,p)$ 的均值为 np

Example 24

设随机变量 X 服从二项分布 $B(n,p), 0<p<1, n\geq 1$. 则随机变量 X 可看成是 n 重 Bernoulli 试验中 A 发生的次数, 且 $\mathsf{P}(A)=p$. 引入随机变量

$$Y_i=\begin{cases} 1, & A \text{ 在第 } i \text{ 次试验发生}, \\ 0, & A \text{ 在第 } i \text{ 次试验不发生}, \end{cases} \qquad i=1,2,\ldots,n.$$

于是 $Y_1, Y_2, \ldots, Y_n$ 相互独立, 服从同一分布 $0-1(p)$, 且 $\mathsf{E}(Y_i)=p, \quad i=1,\ldots,n.$

注意到 $X=Y_1+\cdots+Y_n=\sum\limits_{i=1}^{n}Y_i$, 因此 $\mathsf{E}(X)=\mathsf{E}\left(\sum\limits_{i=1}^{n}Y_i\right)=\sum\limits_{i=1}^{n}\mathsf{E}(Y_i)=np.$

注: 参数为 n,p 的二项分布的随机变量可分解为 n 个相互独立且都服从以 p 为参数的 $0-1$ 分布的随机变量之和.

Example 25

计算机程序随机产生 0~9 中的数字，记 X_i 为第 i 次产生的数字，$i = 1, 2, \ldots, n$. 将这 n 个数依次排列，得到一数，记为 Y，求 $\mathsf{E}(Y)$.

Solution

由题意知，诸 X_i 独立同分布，其分布律为 $\mathsf{P}(X_i = k) = \dfrac{1}{10}$，$k = 0, 1, \ldots, 9$，故

$\mathsf{E}(X_i) = \displaystyle\sum_{k=0}^{9} k \cdot \frac{1}{10} = 4.5$. 又 $Y = \displaystyle\sum_{i=1}^{n} 10^{i-1} X_i$，从而

$$\mathsf{E}(Y) = \sum_{i=1}^{n} 10^{i-1} \mathsf{E}(X_i) = 4.5 \sum_{i=1}^{n} 10^{i-1} = \frac{10^n - 1}{2}.$$

Example 26

一民航送客车载有 20 位旅客自机场出发，旅客有 10 个车站可以下车，如到达一个车站没有旅客下车就不停车，以 X 表示停车的次数，求 $\mathrm{E}(X)$. (设每位旅客在各个车站下车是等可能的，并设各旅客是否下车相互独立.)(思考：X 是否服从二项分布？)

Solution

引入随机变量 $Y_i = \begin{cases} 1, & \text{第 } i \text{ 站有人下车}, \\ 0, & \text{第 } i \text{ 站没有人下车}, \end{cases}$ $i = 1, 2, \ldots, 10$, 则 $X = \sum_{i=1}^{10} Y_i$.

且

$$\mathrm{E}(Y_i) = \mathrm{P}(Y_i = 1) = 1 - \mathrm{P}(\text{第 } i \text{ 站没有人下车}) = 1 - 0.9^{20}.$$

因此 $\mathrm{E}(X) = \mathrm{E}(Y_1) + \mathrm{E}(Y_2) + \cdots + \mathrm{E}(Y_{10}) = 10(1 - 0.9^{20}) \approx 8.784(\text{次})$.

(注意：这里诸 Y_i 不是相互独立的，而且 X 不服从二项分布)

在上例中，我们将 X 分解成数个随机变量之和，然后利用随机变量和的数学期望等于随机变量数学期望之和来求数学期望，这种处理方法具有一定的普遍意义. (注意，它并不要求随机变量之间需要满足独立性假设.)

Example 27 (配对问题, Chapter 1 Example 21)

一个小班有 n 个同学, 编号为 $1,2,\ldots,n$ 号, 中秋节前每人准备一件礼物, 相应编号为 $1,2,\ldots,n$. 将所有礼物集中放在一起, 然后每个同学随机取一件, 若同学拿到自己准备的礼物, 则称配对成功. 记 X 为总的配对数, 求 $\mathsf{E}(X)$.

Solution

引入随机变量 $Y_i=\begin{cases}1, & \text{第 } i \text{ 人拿到自己的礼物},\\ 0, & \text{第 } i \text{ 人未拿到自己的礼物},\end{cases}$ $i=1,2,\ldots,n$, 则 $X=\sum\limits_{i=1}^{n}Y_i$.

注意到

$$\mathsf{E}(Y_i)=\mathsf{P}(Y_i=1)=1/n,$$

因此

$$\mathsf{E}(X)=\sum_{i=1}^{n}\mathsf{E}(Y_i)=1.$$

Example 28

设随机变量 X_1, X_2, X_3, X_4 相互独立, 且 $X_i \sim U(0, 2i)$, 求行列式 $Y = \begin{vmatrix} X_1 & X_2 \\ X_3 & X_4 \end{vmatrix}$ 的数学期望 $\mathsf{E}(Y)$.

Solution

由题设, $\mathsf{E}(X_i) = i, i = 1, 2, 3, 4$, 行列式 $Y = X_1X_4 - X_2X_3$, 因此

$$\mathsf{E}(Y) = \mathsf{E}(X_1X_4) - \mathsf{E}(X_2X_3) = \mathsf{E}(X_1)\mathsf{E}(X_4) - \mathsf{E}(X_2)\mathsf{E}(X_3) = -2.$$

Page168, 思考题四, 第 1 题

* 条件数学期望 (Conditional Expectation)(仅作参考)

Definition 3 (条件数学期望)

若 (X, Y) 为二维离散型随机变量，且在给定 $X=x$ 下，Y 的条件分布律为 $P(Y=y_j|X=x)=p_j(x), j=1,2,\ldots$；或者 (X, Y) 为二维连续型随机变量，且在给定 $X=x$ 下，Y 的条件分布概率密度函数为 $f_{Y|X}(y|x)$，则在给定了随机变量 X 取值为 x 的条件下，Y 的条件期望为

$$E(Y|X=x)=\begin{cases}\displaystyle\sum_{j=1}^{+\infty} y_j p_j(x), & \text{若 } (X, Y) \text{ 为离散型随机变量,}\\ \displaystyle\int_{-\infty}^{+\infty} y\cdot f_{Y|X}(y|x)\mathrm{d}y, & \text{若 } (X, Y) \text{ 为连续型随机变量.}\end{cases}$$

Theorem 3 (全期望公式, Total Expectation Formula)

设 (X, Y) 为二维随机变量, 若 $E(Y)$ 存在, 则有

$$E(Y) = E[E(Y|X)].$$

Proof

若 (X, Y) 为二维连续型随机变量, 则

$$\begin{aligned}
\mathrm{E}[\mathrm{E}(Y|X)] &= \int_{-\infty}^{+\infty} \mathrm{E}(Y|X=x) f_X(x)\mathrm{d}x = \int_{-\infty}^{+\infty} \left(\int_{-\infty}^{+\infty} y f_{Y|x}(y|x)\mathrm{d}y \right) f_X(x)\mathrm{d}x \\
&= \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} y \frac{f(x, y)}{f_X(x)} f_X(x)\mathrm{d}x\mathrm{d}y = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} y f(x, y)\mathrm{d}x\mathrm{d}y = \mathrm{E}(Y).
\end{aligned}$$

Proof (Cont.)

若 (X, Y) 为二维离散型随机变量，则

$$\begin{aligned}
\mathsf{E}[\mathsf{E}(Y|X)] &= \sum_{i=1}^{+\infty} \mathsf{E}(Y|X = x_i)\mathsf{P}(X = x_i) \\
&= \sum_{i=1}^{+\infty}\sum_{j=1}^{+\infty} y_j \mathsf{P}(Y = y_j|X = x_i)\mathsf{P}(X = x_i) \\
&= \sum_{i=1}^{+\infty}\sum_{j=1}^{+\infty} y_j \mathsf{P}(Y = y_j, X = x_i) = \mathsf{E}(Y),
\end{aligned}$$

其中 $x_i, i = 1, 2, \ldots$ 为 X 的所有可能取值.

Example 29

游戏迷宫的入口处有四个门 (无明显标记), 其中一个是自由之门, 选择此门 5 分钟后即可走出迷宫, 剩下的三门将分别在走了 10 分钟, 20 分钟, 30 分钟后重回该四门选择处. 设游戏者每次都是随机地从四个门中选择, 求他走出迷宫所需的平均时间.

Solution

设一个游客走出迷宫的时间为 T 分钟, 并设他对门的选择情况为 X, 用 $1, 2, 3, 4$ 分别表示自由之门及走了 10 分钟, 20 分钟, 30 分钟后重回该处的门号, 那么

$$T = \begin{cases} 5, & X = 1, \\ 10 + T_1, & X = 2, \\ 20 + T_1, & X = 3, \\ 30 + T_1, & X = 4, \end{cases} \quad \text{且 } \mathrm{P}(X = i) = \frac{1}{4}, i = 1, 2, 3, 4.$$

其中 T_1 为选择非自由之门返回该处后到走出迷宫的时间.

Solution (Cont.)

由题意知, T_1 的分布与 T 的分布是一致的. 因此

$$\mathsf{E}(T_1|X=i)=\mathsf{E}(T),\quad i=2,3,4.$$

利用全期望公式, 得

$$\begin{aligned}\mathsf{E}(T)=\mathsf{E}[\mathsf{E}(T|X)]&=\sum_{i=1}^{4}\mathsf{E}(T|X=i)\mathsf{P}(X=i)\\&=\frac{1}{4}\left[5+\mathsf{E}(10+T_1|X=2)+\mathsf{E}(20+T_1|X=3)+\mathsf{E}(30+T_1|X=4)\right]\\&=\frac{1}{4}\left[5+10+20+30+3\mathsf{E}(T)\right],\end{aligned}$$

由此解得 $\mathsf{E}(T)=65$ (分钟).

§4.2 方差，变异系数 (Variance, Coefficient of Variation)

方差 (Variance)

Example 30

设有一批灯泡寿命为：一半约 950 小时，另一半约 1050 小时.

$\rightarrow$ 平均寿命约为 1000 小时；

另一批灯泡寿命为：一半约 1300 小时，另一半约 700 小时.

$\rightarrow$ 平均寿命约为 1000 小时.

问：哪一批灯泡的质量更好？

单从平均寿命这一指标无法判断，通常会进一步考察灯泡质量的稳定性，也就是灯泡寿命 X 与期望 (均值)1000 小时的偏离程度.

随机变量 X 的期望 (均值) 为 $\mathsf{E}(X)$

随机变量 X 对于期望 (均值)$\mathsf{E}(X)$ 的离差: $X - \mathsf{E}(X)$ (仍然随机)

随机变量 X 对于期望的平均离差: $\mathsf{E}(X - \mathsf{E}(X)) = 0$
(因为出现了正负离差的抵消, 故用此值衡量偏离程度无意义)

常常用$\mathsf{E}[X - \mathsf{E}(X)]^2$来衡量随机变量的波动性.
(这一数字特征称为方差)

Definition 4 (方差)

设 X 是一个随机变量，若 $E\{[X-E(X)]^2\}$ 存在，则称其为 X 的方差 (variance)，记为 $Var(X)$ 或 $D(X)$，即

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

将 $\sqrt{Var(X)}$ 记为 $\sigma(X)$ 或 $SD(X)$，称为 X 的标准差 (standard deviation) 或均方差，它是与随机变量 X 具有相同量纲的量.

方差 $\mathrm{Var}(X)$ 和标准差 $SD(X)$ 均可刻画 X 取值的分散程度，是衡量 X 取值分散程度的常用标准.

若 X 的取值比较集中，则 $\mathrm{Var}(X)(SD(X))$ 较小；若 X 的取值比较分散，则 $\mathrm{Var}(X)(SD(X))$ 较大.

方差的计算可以看成随机变量的函数 $g(X)$ 的数学期望，其中函数

$$g(x) = [x - \mathsf{E}(X)]^2.$$

因此由定理1，可得

- 对于**离散型**随机变量 X，其分布律为 $\mathsf{P}(X = x_i) = p_i, i = 1, 2, \ldots,$ 当其方差存在时，有

$$\mathsf{Var}(X) = \sum_{i=1}^{+\infty} [x_i - \mathsf{E}(X)]^2 p_i.$$

- 对于**连续型**随机变量 X，其概率密度函数为 $f(x)$，当其方差存在时，有

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

在实际操作中，我们常常采用下面的公式来计算方差.

Theorem 4

若随机变量 X 的方差存在，则

$$Var(X) = E(X^2) - [E(X)]^2. \qquad \text{——方差的常用计算公式}$$

Proof

由方差的定义和期望的性质，可得

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

Example 31

设随机变量 X 具有数学期望 $\mathsf{E}(X)=\mu$, 方差 $\mathsf{Var}(X)=\sigma^2\neq 0$, 记 $X^*=\dfrac{X-\mu}{\sigma}$ (称为 X 的标准化变量). 证明: $\mathsf{E}(X^*)=0$, $\mathsf{Var}(X^*)=1$.

Proof

由期望的性质, 可得

$$\mathsf{E}(X^*)=\mathsf{E}\left(\frac{X-\mu}{\sigma}\right)=\frac{1}{\sigma}\mathsf{E}(X-\mu)=\frac{1}{\sigma}(\mathsf{E}(X)-\mu)=0.$$

由方差的计算公式和期望的性质, 可得

$$\mathsf{Var}(X^*)=\mathsf{E}(X^{*2})-[\mathsf{E}(X^*)]^2=\mathsf{E}\left[\left(\frac{X-\mu}{\sigma}\right)^2\right]-0=\frac{1}{\sigma^2}\mathsf{E}\left[(X-\mu)^2\right]=\frac{\sigma^2}{\sigma^2}=1.$$

Example 32

设随机变量 X 服从 $0-1(p)$ 分布 $(0<p<1)$, 其分布律为 $\mathsf{P}(X=0)=1-p, \mathsf{P}(X=1)=p$, 求方差 $\mathsf{Var}(X)$.

Solution

之前已知期望 $\mathsf{E}(X)=p$.

(法一)

$$\mathsf{Var}(X)=\mathsf{E}\{[X-\mathsf{E}(X)]^2\}=\mathsf{P}(X=0)\cdot(0-p)^2+\mathsf{P}(X=1)\cdot(1-p)^2=p(1-p).$$

(法二) 由于 $\mathsf{E}(X^2)=0^2\cdot(1-p)+1^2\cdot p=p$, 故利用方差的计算公式, 可得

$$\mathsf{Var}(X)=\mathsf{E}(X^2)-[\mathsf{E}(X)]^2=p-p^2=p(1-p).$$

Example 33

设随机变量 X 服从泊松分布 $P(\lambda)(\lambda>0)$, 求方差 $\mathrm{Var}(X)$.

Solution

泊松分布 $P(\lambda)$ 的分布律为: $\mathrm{P}(X=k)=\frac{\lambda^k e^{-\lambda}}{k!}$, $\quad k=0,1,\ldots$. 之前也已得到期望 $\mathrm{E}(X)=\lambda$, 利用方差的计算公式, 先计算 $\mathrm{E}(X^2)$.

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

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

泊松分布的均值和方差相等, 都等于参数 λ.

从泊松分布方差的计算过程中不难推广得到如下结论.

Proposition 3

设随机变量 X 服从泊松分布 $P(\lambda)(\lambda > 0)$, 则

$$\mathsf{E}[X(X-1)\cdots(X-m)] = \lambda^{m+1}, \quad m \in \mathbb{N}.$$

Proof

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

$$= \lambda^{m+1}\mathrm{e}^{-\lambda} \sum_{k=m+1}^{+\infty} \frac{\lambda^{k-(m+1)}}{(k-(m+1))!} = \lambda^{m+1}\mathrm{e}^{-\lambda}\mathrm{e}^{\lambda} = \lambda^{m+1}.$$

Example 34

设随机变量 X 具有几何分布 $Geom(p)(0<p<1)$, 求方差 $\mathsf{Var}(X)$.

Solution

之前已得期望 $\mathsf{E}(X)=1/p$, 而

$$\begin{aligned}
\mathsf{E}(X^2) &= \mathsf{E}[X(X-1)+X] = \mathsf{E}[X(X-1)] + \mathsf{E}(X) \\
&= \sum_{k=2}^{+\infty} k(k-1)p(1-p)^{k-1} + \frac{1}{p} = p(1-p)\cdot\left(\frac{x^2}{1-x}\right)''\Big|_{x=1-p} + \frac{1}{p} \\
&= p(1-p)\frac{2}{p^3} + \frac{1}{p} = \frac{2}{p^2} - \frac{1}{p}
\end{aligned}$$

$$\Rightarrow \quad \mathsf{Var}(X) = \mathsf{E}(X^2) - [\mathsf{E}(X)]^2 = \frac{1-p}{p^2}.$$

Example 35

设随机变量 X 具有均匀分布 $U(a,b)(a<b)$, 求方差 $\mathrm{Var}(X)$.

Solution

之前已得期望 $\mathrm{E}(X)=\dfrac{a+b}{2}$, 而

$$\mathrm{E}(X^2)=\int_{-\infty}^{+\infty}x^2f(x)\mathrm{d}x=\int_a^b\frac{x^2}{b-a}\mathrm{d}x=\frac{b^3-a^3}{3(b-a)}=\frac{a^2+b^2+ab}{3}.$$

$$\begin{aligned}\Rightarrow\ \mathrm{Var}(X)&=\mathrm{E}(X^2)-[\mathrm{E}(X)]^2\\&=\frac{a^2+b^2+ab}{3}-\frac{a^2+b^2+2ab}{4}=\frac{(b-a)^2}{12}.\end{aligned}$$

Example 36

设随机变量 X 服从指数分布 $E(\lambda)(\lambda > 0)$, 求方差 $\mathrm{Var}(X)$.

Solution

之前已得期望 $\mathrm{E}(X) = 1/\lambda$, 而

$$\mathrm{E}(X^2) = \int_{-\infty}^{+\infty} x^2 f(x)\mathrm{d}x = \int_0^{+\infty} x^2\lambda \mathrm{e}^{-\lambda x}\mathrm{d}x$$
$$= -x^2\mathrm{e}^{-\lambda x}\Big|_0^{+\infty} + \int_0^{+\infty} 2x\mathrm{e}^{-\lambda x}\mathrm{d}x = \frac{2}{\lambda^2}.$$
$$\Rightarrow \mathrm{Var}(X) = \mathrm{E}(X^2) - [\mathrm{E}(X)]^2 = \frac{1}{\lambda^2}.$$

指数分布的方差是其期望的平方.

方差的性质

(1) 设 C 是常数, 则有 $\mathrm{Var}(C)=0$.

Proof

由定义, $\mathrm{Var}(C)=\mathrm{E}\{[C-\mathrm{E}(C)]^2\}=\mathrm{E}\big[(C-C)^2\big]=0$.

(2) 设 X 是一个随机变量, C 是常数, 则有 $\mathrm{Var}(CX)=C^2\mathrm{Var}(X)$.

Proof

由方差的计算公式和期望的性质, 可得

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

(3) 设 X, Y 是两个随机变量，则有

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

特别地，若 X, Y 相互独立，则有 $\mathrm{Var}(X+Y)=\mathrm{Var}(X)+\mathrm{Var}(Y)$.

Proof

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

特别地，当 X, Y 相互独立时，$X-\mathrm{E}(X)$ 与 $Y-\mathrm{E}(Y)$ 相互独立，此时第三项

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

综合上面三条可以推广得到任意有限个**独立**随机变量线性组合的情形.

Proposition 4

若 n 个随机变量 $X_1, X_2, \ldots, X_n (n \geq 2)$ 相互独立, 且方差都存在. 对任意 $n+1$ 个有限实数 $c_0, c_1, c_2, \ldots, c_n$, 则 $c_0 + \sum_{i=1}^{n} c_i X_i$ 的方差也存在, 且

$$Var\left(c_0 + \sum_{i=1}^{n} c_i X_i\right) = \sum_{i=1}^{n} c_i^2 Var(X_i).$$ 注: 没有 c_0

(4) $\mathrm{Var}(X) = 0 \;\Leftrightarrow\; \mathrm{P}(X = C) = 1$ 且 $C = \mathrm{E}(X)$.

必要性显然, 但充分性的证明需要用到 Chebyshev 不等式 (Chapter 5), 此处从略.

Example 37

设随机变量 X 具有二项分布 $B(n,p)$, $n\geq 1, 0<p<1$. 求方差 $\mathrm{Var}(X)$.

Solution

(与求期望采用相同的方法) 随机变量 X 可视为 n 重贝努里试验中事件 A 发生的次数, 其中 $\mathrm{P}(A)=p$ 引入随机变量

$$Y_i=\begin{cases}1, & A\ 在第\ i\ 次试验发生,\\ 0, & A\ 在第\ i\ 次试验不发生,\end{cases}\qquad i=1,2,\ldots,n.$$

于是 $Y_1,Y_2,\ldots,Y_n$ 相互独立, 服从同一分布 $0-1(p)$, 且 $\mathrm{Var}(Y_i)=p(1-p),\forall i$.
又 $X=Y_1+\cdots+Y_n=\sum_{i=1}^n Y_i$, 结合 $Y_1,Y_2,\ldots,Y_n$ 的独立性, 可得

$$\mathrm{Var}(X)=\mathrm{Var}\left(\sum_{i=1}^n Y_i\right)=\sum_{i=1}^n \mathrm{Var}(Y_i)=np(1-p).$$

Example 38

设随机变量 X 服从正态分布 $N(\mu,\sigma^2)$, 求方差 $\mathrm{Var}(X)$.

Solution

先求标准正态随机变量 $Z=\frac{X-\mu}{\sigma}$ 的方差. 已知 $\mathrm{E}(Z)=0$, 故

$$\mathrm{Var}(Z)=\mathrm{E}(Z^2)-0^2=\int_{-\infty}^{+\infty} z^2\varphi(z)\mathrm{d}z=\int_{-\infty}^{+\infty}\frac{z^2}{\sqrt{2\pi}}\mathrm{e}^{-\frac{z^2}{2}}\mathrm{d}z$$

$$=-\frac{1}{\sqrt{2\pi}}z\mathrm{e}^{-\frac{z^2}{2}}\Big|_{-\infty}^{+\infty}+\int_{-\infty}^{+\infty}\frac{1}{\sqrt{2\pi}}\mathrm{e}^{-\frac{z^2}{2}}\mathrm{d}z=1.$$

注意到 $X=\mu+\sigma Z$, 利用方差的性质可得

$$\mathrm{Var}(X)=\mathrm{Var}(\mu+\sigma Z)=\sigma^2\mathrm{Var}(Z)=\sigma^2.$$

正态分布 $N(\mu,\sigma^2)$ 中的 μ,σ^2 分别是该分布的数学期望和方差.

独立的 n 个正态变量的线性组合仍服从正态分布.

若 $X_i \sim N(\mu_i, \sigma_i^2), i = 1, 2, \ldots, n$, 且它们相互独立, 则它们的线性组合

$$c_0 + \sum_{i=1}^{n} c_i X_i \sim N\left(c_0 + \sum_{i=1}^{n} c_i \mu_i, \sum_{i=1}^{n} c_i^2 \sigma_i^2\right),$$

其中 $c_1, c_2, \ldots, c_n$ 是不全为零的常数.

Example 39

设 $X \sim N(1, 3)$, $Y \sim N(2, 4)$, 且 X, Y 相互独立, 则

$$Z_1 = 2X - 3Y \sim N(-4, 48), \quad Z_2 = 2X - 3Y + 4 \sim N(0, 48).$$

Example 40

设活塞的直径 (单位: cm) $X \sim N(22.40, 0.03^2)$, 汽缸的直径 (单位: cm) $Y \sim N(22.50, 0.04^2)$, X, Y 相互独立. 现任取一只活塞, 任取一只汽缸, 求活塞能装入汽缸的概率.

Solution

由题意需求 $\mathrm{P}(X < Y) = \mathrm{P}(X - Y < 0)$, 由于 $X - Y \sim N(-0.10, 0.05^2)$, 所以

$$\mathrm{P}(X < Y) = \mathrm{P}(X - Y < 0) = \Phi\left(\frac{0 - (-0.10)}{0.05}\right) = \Phi(2) = 0.9772.$$

Table: 几种常见分布的均值与方差

分布	分布律或密度函数	期望	方差
0-1 分布 $0-1(p)$	$\mathrm{P}(X=k)=p^k(1-p)^{1-k}, k=0,1$	p	$p(1-p)$
二项分布 $B(n,p)$	$\mathrm{P}(X=k)=\mathrm{C}_n^k p^k(1-p)^{n-k}, k=0,1,\ldots,n$	np	$np(1-p)$
泊松分布 $P(\lambda)$	$\mathrm{P}(X=k)=\lambda^k\mathrm{e}^{-\lambda}/k!, k=0,1,\ldots$	λ	λ
均匀分布 $U(a,b)$	$f(x)=\begin{cases}1/(b-a), & a<x<b,\\ 0, & \text{其他}\end{cases}$	$\dfrac{a+b}{2}$	$\dfrac{(b-a)^2}{12}$
指数分布 $E(\lambda)$	$f(x)=\begin{cases}\lambda\mathrm{e}^{-\lambda x}, & x>0,\\ 0, & x\le 0\end{cases}$	$\dfrac{1}{\lambda}$	$\dfrac{1}{\lambda^2}$
正态分布 $N(\mu,\sigma^2)$	$f(x)=\dfrac{1}{\sqrt{2\pi}\sigma}\mathrm{e}^{-\frac{(x-\mu)^2}{2\sigma^2}}, -\infty<x<+\infty$	μ	σ^2

* 变异系数 (Coefficient of Variation)

Definition 5 (变异系数)

设随机变量 X 具有数学期望 $E(X)=\mu$, 方差 $Var(X)=\sigma^2\neq 0$, 则称

$$C_v(X)=\frac{\sqrt{Var(X)}}{E(X)}=\frac{\sigma}{\mu}$$

为 X 的变异系数 (coefficient of variation).

变异系数是一个无量纲的指标，它反映了随机变量 X 在以它的中心位置为标准时，取值的离散程度.

§4.3 协方差，相关系数 (Covariance, Correlation Coefficient)

对于二维随机变量 (X, Y)，除了讨论 X 与 Y 各自的数学期望和方差外，还需要讨论描述 X 和 Y 之间相互关系的数字特征.

协方差 (Corvariance)

Definition 6 (协方差)

对于数学期望都存在的随机变量 X 和 Y，当 $(X - E(X))(Y - E(Y))$ 的数学期望存在时，称 $E[(X - E(X))(Y - E(Y))]$ 为随机变量 X 与 Y 的协方差 (covariance)，记为 $Cov(X, Y)$，即

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

协方差的计算可以看成二维随机变量函数 $h(X, Y)$ 的数学期望，其中

$$h(x, y) = [x - \mathsf{E}(X)][y - \mathsf{E}(Y)].$$

因此由定理2，可得

- 对于二维**离散型**随机变量 (X, Y)，其联合分布律为 $\mathsf{P}(X = x_i, Y = y_j) = p_{ij}$, $i, j = 1, 2, \ldots,$ 则

$$\mathsf{Cov}(X, Y) = \sum_{i=1}^{+\infty}\sum_{j=1}^{+\infty}[x_i - \mathsf{E}(X)][y_i - \mathsf{E}(Y)]p_{ij}.$$

- 对于二维**连续型**随机变量 (X, Y)，其联合概率密度函数为 $f(x, y)$，则

$$\mathsf{Cov}(X, Y) = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}[x - \mathsf{E}(X)][y - \mathsf{E}(Y)]f(x, y)\mathrm{d}x\mathrm{d}y.$$

在实际操作中，我们常常采用下面的公式来计算协方差.

Theorem 5

若随机变量 X 与 Y 的期望存在，则

$$Cov(X, Y) = E(XY) - E(X)E(Y) \quad \text{——协方差的计算公式.}$$

Proof

由协方差的定义，

$$\begin{aligned}
\mathrm{Cov}(X, Y) &= \mathrm{E}\{[X - \mathrm{E}(X)][Y - \mathrm{E}(Y)]\} \\
&= \mathrm{E}\{XY - X\mathrm{E}(Y) - Y\mathrm{E}(X) + \mathrm{E}(X)\mathrm{E}(Y)\} \\
&= \mathrm{E}(XY) - \mathrm{E}(X)\mathrm{E}(Y) - \mathrm{E}(Y)\mathrm{E}(X) + \mathrm{E}(X)\mathrm{E}(Y) \\
&= \mathrm{E}(XY) - \mathrm{E}(X)\mathrm{E}(Y).
\end{aligned}$$

引入协方差后，方差的性质 (3) 及其推广可重新书写.

Theorem 6

当随机变量 X 与 Y 的方差存在时，有

$$Var(X+Y)=Var(X)+Var(Y)+2Cov(X,Y).$$

这一性质也可推广到任意有限个随机变量之和的情形:

若 n 个随机变量 $X_1,X_2,\ldots,X_n(n\geq 2)$ 的方差存在，则 $\sum\limits_{i=1}^{n}X_i$ 的方差也存在，且

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

协方差的性质

(1) $\text{Cov}(X, Y) = \text{Cov}(Y, X)$;
(2) $\text{Cov}(X, X) = \text{Var}(X)$;
(3) $\text{Cov}(aX, bY) = ab \cdot \text{Cov}(X, Y)$, 其中 a, b 为两个实数;
(4) $\text{Cov}(X_1 + X_2, Y) = \text{Cov}(X_1, Y) + \text{Cov}(X_2, Y)$.

思考题

对随机变量 X, Y, 及实数 a, b, c, d, 协方差 $\mathrm{Cov}(aX+bY, cX+dY)=?$

Solution

$$\begin{aligned}&\mathrm{Cov}(aX+bY, cX+dY)\\=&\mathrm{Cov}(aX, cX)+\mathrm{Cov}(aX, dY)+\mathrm{Cov}(bY, cX)+\mathrm{Cov}(bY, dY)\\=&ac\mathrm{Var}(X)+bd\mathrm{Var}(Y)+(ad+bc)\mathrm{Cov}(X, Y).\end{aligned}$$

协方差性质的证明

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

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

(3)

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

(4)

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

相关系数 (Correlation Coefficient)

Definition 7 (相关系数)

对于随机变量 X 和 Y, 当 $E(X^2)$ 与 $E(Y^2)$ 均存在且 $Var(X), Var(Y)$ 均为非零实数时, 称

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

为 X 与 Y 的相关系数 (correlation coefficient), 有时也简记为 ρ. 且

$$\rho_{XY} = Cov(X^*, Y^*), \quad \text{其中} \quad X^*, Y^* \text{分别是} X, Y \text{标准化变量}.$$

ρ_{XY} 是一个无量纲的数字特征.

相关系数是刻画两个变量之间相依关系的一种数字特征.

相关系数的性质

(1) $|\rho_{XY}| \leq 1$
(2) $|\rho_{XY}| = 1 \Leftrightarrow$ 存在常数 $a, b(b \neq 0)$, 使得 $\mathrm{P}(Y = a + bX) = 1$. 特别地, $\rho_{XY} = 1$ 时, $b > 0$; $\rho_{XY} = -1$ 时, $b < 0$.

也可将 (1)(2) 一并写成:

Proposition 5

对方差存在的随机变量 X, Y, 当 $\mathrm{Var}(X)\mathrm{Var}(Y) \neq 0$ 时, 有

$$[\mathrm{Cov}(X, Y)]^2 \leq \mathrm{Var}(X)\mathrm{Var}(Y),$$

其中等号成立当且仅当 X 与 Y 之间在概率意义下具有严格的线性关系, 即存在常数 $a, b(b \neq 0)$, 使得

$$\mathrm{P}(Y = a + bX) = 1.$$

Proof

考虑以 X 的线性函数 $a+bX$ 来近似表示 Y, 以均方误差 $e(a,b)=\mathrm{E}\{[Y-(a+bX)]^2\}$ 来衡量以 $a+bX$ 近似表示 Y 的好坏程度. 如果 $e(a,b)$ 越小, 则 $a+bX$ 与 Y 的近似程度越好.

下面来求最佳近似式 $e(a_0,b_0)=\min\limits_{a,b} e(a,b)$, 计算得

$$e(a,b)=\mathrm{E}(Y^2)+b^2\mathrm{E}(X^2)+a^2-2a\mathrm{E}(Y)-2b\mathrm{E}(XY)+2ab\mathrm{E}(X).$$

由 $\begin{cases} \dfrac{\partial e(a,b)}{\partial a}=2a+2b\mathrm{E}(X)-2\mathrm{E}(Y)=0, \\ \dfrac{\partial e(a,b)}{\partial b}=2a\mathrm{E}(X)+2b\mathrm{E}(X^2)-2\mathrm{E}(XY)=0, \end{cases}$ 可得 $\begin{cases} a_0=\mathrm{E}(Y)-b_0\mathrm{E}(X), \\ b_0=\dfrac{\mathrm{Cov}(X,Y)}{\mathrm{Var}(X)}. \end{cases}$

又 Hessen 矩阵正定, 因此 (a_0,b_0) 为 $e(a,b)$ 的极 (最) 小值点.

Proof (Cont.)

注意到 $a_0 = \mathsf{E}(Y) - b_0\mathsf{E}(X),\ b_0 = \frac{\mathsf{Cov}(X,Y)}{\mathsf{Var}(X)}$，则有

$$\begin{aligned} e(a_0,b_0) &= \mathsf{E}\{[Y-(a_0+b_0X)]^2\} = \mathsf{Var}[Y-(a_0+b_0X)] + \{\mathsf{E}[Y-(a_0+b_0X)]\}^2 \\ &= \mathsf{Var}(Y-b_0X) + 0 = \mathsf{Var}(Y) + b_0^2\mathsf{Var}(X) - 2b_0\mathsf{Cov}(X,\ Y) \\ &= \mathsf{Var}(Y) - \frac{[\mathsf{Cov}(X,\ Y)]^2}{\mathsf{Var}(X)} = (1-\rho_{XY}^2)\mathsf{Var}(Y) \end{aligned}$$

(1) 由 $e(a_0,b_0) \geq 0$ 可得 $1-\rho_{XY}^2 \geq 0$，即 $|\rho_{XY}| \leq 1$.

(2)

$$\begin{aligned} |\rho_{XY}| = 1 &\Leftrightarrow \mathsf{E}\{[Y-(a_0+b_0X)]^2\} = 0 \\ &\Leftrightarrow \mathsf{Var}[Y-(a_0+b_0X)] = 0 \text{ 且} \mathsf{E}[Y-(a_0+b_0X)] = 0 \\ &\Leftrightarrow \mathsf{P}\{Y-(a_0+b_0X) = 0\} = 1. \end{aligned}$$

当 $\rho_{XY} > 0$ 时，$\mathsf{Cov}(X,\ Y) > 0,\ b_0 > 0$；当 $\rho_{XY} < 0$ 时，$\mathsf{Cov}(X,\ Y) < 0,\ b_0 < 0$.

相关系数 ρ_{XY}是一个用来表征 X, Y 之间线性关系紧密程度的数字特征.

- 当 $|\rho_{XY}|$ 较大时, $e(a_0, b_0)$ 较小, 表明 X, Y 线性关系的程度较强;
- 当 $|\rho_{XY}| = 1$ 时, $e(a_0, b_0) = 0$, 表明 X, Y 之间以概率 1 存在线性关系;
- 当 $|\rho_{XY}|$ 较小时, $e(a_0, b_0)$ 较大, 表明 X, Y 线性关系的程度较弱.

Definition 8 (正相关、负相关)

当 $\rho_{XY} > 0$ 时, 称 X 与 Y 为正相关, 当 $\rho_{XY} < 0$ 时, 称 X 与 Y 为负相关.

Definition 9 (不相关)

当 $\rho_{XY}=0$ 时, 称 X 与 Y 为不相关或零相关.

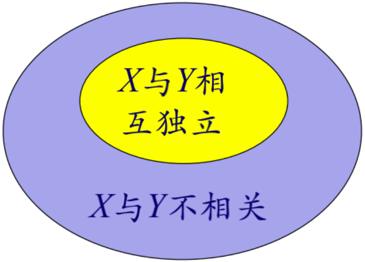
随机变量 X 与 Y 不相关, 即 $\rho_{XY}=0$ 的等价条件:

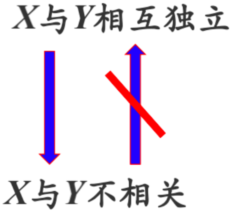
1. $\mathrm{Cov}(X, Y)=0$;
2. $\mathrm{E}(XY)=\mathrm{E}(X)\mathrm{E}(Y)$; (较常用)
3. $\mathrm{Var}(X+Y)=\mathrm{Var}(X)+\mathrm{Var}(Y)$.

相关性与独立性的关系

X 与 Y 不相关，仅针对于二者的线性关系而言；
而 X 与 Y 相互独立，是就一般关系而言的.

- X 与 Y 相互独立 $\Rightarrow$ X 与 Y 不相关
- X 与 Y 不相关 $\nRightarrow$ X 与 Y 相互独立





Example 41

设 X 的分布律如下所示，问 X 与 X^2 是否相关？是否独立？

X	-1	0	1
P	1/4	1/2	1/4

Solution

注意到 $\mathsf{E}(X) = \mathsf{E}(X^3) = 0$，因此有 $\mathsf{E}(X)\mathsf{E}(X^2) = \mathsf{E}(X \cdot X^2)$，所以 X 与 X^2 不相关.
而 X 与 X^2 是有关系的，是不独立的. 事实上

$$\mathsf{P}(X = -1, X^2 = 0) = 0 \neq \mathsf{P}(X = -1)\mathsf{P}(X^2 = 0) = 1/8.$$

在本例中也可以看出，"相关性" 度量的是随机变量之间的线性关系.

Example 42

设 X, Y 服从同一分布，其分布律为

X	-1	0	1
P	1/4	1/2	1/4

已知 $\mathrm{P}\{|X| = |Y|\} = 0$，判断 X 和 Y 是否相关？是否独立？

Solution

由题意可得 X, Y 的联合分布律

$X \backslash Y$	-1	0	1	$\mathrm{P}(X=i)$
-1	0	1/4	0	1/4
0	1/4	0	1/4	1/2
1	0	1/4	0	1/4
$\mathrm{P}(Y=j)$	1/4	1/2	1/4	1

因此 $\mathrm{E}(XY) = 0 \times (-1) \times 1/4 + (-1) \times 0 \times 1/4 + 1 \times 1 \times 1/4 + 0 \times 1 \times 1/4 = 0$, 而 $\mathrm{E}(X) = \mathrm{E}(Y) = 0$, 所以 $\mathrm{E}(XY) = \mathrm{E}(X)\mathrm{E}(Y)$, 即 X 与 Y 不相关.

此外, 注意到 $\mathrm{P}(X=-1, Y=-1) = 0 \neq \mathrm{P}(X=-1)\mathrm{P}(Y=-1) = 1/16$, 所以 X 与 Y 不独立.

Example 43

设 (X, Y) 服从二维正态分布 $N(\mu_1, \mu_2; \sigma_1^2, \sigma_2^2; \rho)$，它的联合概率密度函数为

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

求 X 和 Y 的相关系数，并证明 X 与 Y 相互独立 $\Leftrightarrow$ X 与 Y 不相关.

Solution

由前面对二维正态分布的分析可知

$$\mathsf{E}(X) = \mu_1, \mathsf{Var}(X) = \sigma_1^2; \quad \mathsf{E}(Y) = \mu_2, \mathsf{Var}(Y) = \sigma_2^2.$$

下面计算 X 与 Y 的协方差 $\mathsf{Cov}(X, Y)$.

Solution (Cont.)

$$
\begin{aligned}
\mathrm{Cov}(X,Y) &= \mathrm{E}\big[(X-\mu_1)(Y-\mu_2)\big] \\
&= \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}(x-\mu_1)(y-\mu_2)f(x,y)\,\mathrm{d}x\mathrm{d}y \\
&= \int_{-\infty}^{+\infty}\frac{x-\mu_1}{\sqrt{2\pi}\sigma_1}\mathrm{e}^{-\frac{(x-\mu_1)^2}{2\sigma_1^2}}\mathrm{d}x\int_{-\infty}^{+\infty}\frac{y-\mu_2}{\sqrt{2\pi}\sigma_2\sqrt{1-\rho^2}}\mathrm{e}^{-\frac{\left[y-\left(\mu_2+\frac{\rho\sigma_2}{\sigma_1}(x-\mu_1)\right)\right]^2}{2(1-\rho^2)\sigma_2^2}}\mathrm{d}y \\
&= \int_{-\infty}^{+\infty}\frac{x-\mu_1}{\sqrt{2\pi}\sigma_1}\mathrm{e}^{-\frac{(x-\mu_1)^2}{2\sigma_1^2}}\left[\mu_2+\frac{\rho\sigma_2}{\sigma_1}(x-\mu_1)-\mu_2\right]\mathrm{d}x \\
&= \frac{\rho\sigma_2}{\sigma_1}\int_{-\infty}^{+\infty}\frac{(x-\mu_1)^2}{\sqrt{2\pi}\sigma_1}\mathrm{e}^{-\frac{(x-\mu_1)^2}{2\sigma_1^2}} = \frac{\rho\sigma_2}{\sigma_1}\cdot\sigma_1^2 = \rho\sigma_1\sigma_2.
\end{aligned}
$$

Solution (Cont.)

因此

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

即二维正态变量 (X, Y) 中的参数 ρ 就是 X, Y 的相关系数, 因而二维正态变量的分布完全可由 X, Y 各自的均值、方差以及它们的相关系数所确定.

之前我们已经证明, 若 (X, Y) 服从二维正态分布, 那么 X 和 Y 相互独立 $\Leftrightarrow \rho = 0$. 因此, 对于二维正态变量 (X, Y) 来说, X 与 Y 相互独立 $\Leftrightarrow$ X 与 Y 不相关.

Example 44

设 X, Y 相互独立服从同一分布，方差存在，记 $U = X - Y, V = X + Y$，则随机变量 U 与 V 是否一定不相关？是否一定独立？

Solution

先求 U, V 的协方差

$$\mathrm{Cov}(U, V) = \mathrm{Cov}(X - Y, X + Y) = \mathrm{Var}(X) - \mathrm{Var}(Y) = 0,$$

所以 U 与 V 一定不相关.

Solution (Cont.)

但 U 与 V 不一定独立, 也不一定不独立. 举例如下:

1. 设 X 与 Y 独立服从正态分布, 则 (U,V) 也服从正态分布, 而对于二维正态分布, 独立与不相关等价, 从而 U 与 V 独立.
2. 设 X 与 Y 独立服从两点分布 $B(1,0.5)$(或写为 $0-1(0.5)$), 则

$$\begin{aligned}
\mathsf{P}(U=1,V=0)&=\mathsf{P}(X-Y=1,X+Y=0)=0,\\
\mathsf{P}(U=1)&=\mathsf{P}(X-Y=1)=\mathsf{P}(X=1,Y=0)=0.25,\\
\mathsf{P}(V=0)&=\mathsf{P}(X+Y=0)=\mathsf{P}(X=0,Y=0)=0.25,
\end{aligned}$$

$\mathsf{P}(U=1,V=0)\neq\mathsf{P}(U=1)\mathsf{P}(V=0)$, 因此 U 与 V 不独立.

§4.4 其它数字特征

矩 (Moment)

Definition 10 (矩)

设 X, Y 为随机变量,

- 若 $E(X^k), k = 1, 2, \ldots$ 存在, 则称它为 X 的k 阶 (原点) 矩.
- 若 $E\{[X - E(X)]^k\}, k = 2, 3, \ldots$ 存在, 则称它为 X 的k 阶中心矩.
- 若 $E(X^kY^l), k, l = 1, 2, \ldots$ 存在, 则称它为 X 和 Y 的 $k + l$ 阶混合 (原点) 矩.
- 若 $E\{[X - E(X)]^k[Y - E(Y)]^l\}, k, l = 1, 2, \ldots$ 存在, 则称它为 X 和 Y 的 $k + l$ 阶混合中心矩.

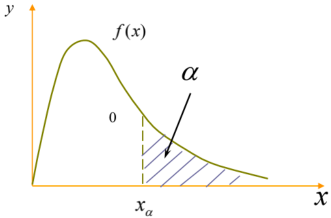
分位数 (Quantile)

Definition 11 (分位数)

设连续型随机变量 X 的分布函数和密度函数分别为 $F(x)$ 与 $f(x)$，称满足条件

$$P(X > x_\alpha) = 1 - F(x_\alpha) = \int_{x_\alpha}^{+\infty} f(x)\mathrm{d}x = \alpha$$

的实数 x_α 为随机变量 X (或此分布) 的上 (侧) α 分位数$(0 < \alpha < 1)$.



分布的上α分位数

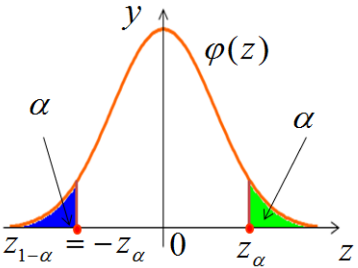
特别地，当 $\alpha = 1/2$ 时，$x_{1/2}$ 称为 X 的中位数;
当 $\alpha = 1/4$ 时，$x_{1/4}$ 称为 X 的上 1/4 分位数;
当 $\alpha = 3/4$ 时，$x_{3/4}$ 称为 X 的上 3/4 分位数.

设 $X \sim N(0,1)$, 若 z_α 满足条件

$$\mathrm{P}(X > z_\alpha) = \alpha, \quad 0 < \alpha < 1,$$

则称点 z_α 为标准正态分布的上 α 分位数. (Excel 实现: NORMSINV)

- $z_{1/2} = 0$.
- $z_{1-\alpha} = -z_\alpha$.



众数 (Mode)

Definition 12 (众数)

随机变量 X 最可能取的数值, 称为众数 (mode).

对于离散型随机变量, 众数是其分布律中最大概率所对应的那个可能取值.

§4.5 多维随机变量的数字特征

Definition 13 (数学期望 (向量))

设 n 维随机变量 $\boldsymbol{X} = (X_1, X_2, \ldots, X_n)^T$, 若其每一分量的数学期望都存在, 则称

$$E(\boldsymbol{X}) = (E(X_1), E(X_2), \ldots, E(X_n))^T$$

为 n 维随机变量 $\boldsymbol{X}$ 的数学期望 (向量).

Definition 14 (协方差矩阵)

设二维随机变量 (X_1, X_2) 的四个二阶中心矩存在, 则称

$$\begin{pmatrix} Var(X_1) & Cov(X_1, X_2) \\ Cov(X_2, X_1) & Var(X_2) \end{pmatrix},$$

为 (X_1, X_2) 的协方差矩阵.

Definition 15 (协方差矩阵)

设 n 维随机变量 $(X_1, X_2, \ldots, X_n)$ 的任意二者协方差 $Cov(X_i, X_j)$ 存在, $i, j = 1, 2, \ldots, n$, 称矩阵

$$(Cov(X_i, X_j))_{n\times n} = \begin{pmatrix} Var(X_1) & Cov(X_1, X_2) & \cdots & Cov(X_1, X_n) \\ Cov(X_2, X_1) & Var(X_2) & \cdots & Cov(X_2, X_n) \\ \vdots & \vdots & \ddots & \vdots \\ Cov(X_n, X_1) & Cov(X_n, X_2) & \cdots & Var(X_n) \end{pmatrix}$$

为 n 维随机变量 $(X_1, X_2, \ldots, X_n)$ 的协方差矩阵.

协方差矩阵是一个对称的非负定矩阵.

利用协方差矩阵, 可将二维正态变量的概率密度推广至 n 维正态变量的概率密度.

已知 $\boldsymbol{X}=(X_1,X_2)$ 服从二维正态分布, 其概率密度为

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

引入列向量 $\boldsymbol{x}=\begin{pmatrix}x_1\\x_2\end{pmatrix}$, $\boldsymbol{\mu}=\begin{pmatrix}\mu_1\\\mu_2\end{pmatrix}$, $\boldsymbol{X}$ 的协方差矩阵 $\boldsymbol{C}=\begin{pmatrix}\sigma_1^2 & \rho\sigma_1\sigma_2\\ \rho\sigma_1\sigma_2 & \sigma_2^2\end{pmatrix}$,

它的行列式为 $|\boldsymbol{C}|=\sigma_1^2\sigma_2^2(1-\rho^2)$, 逆矩阵为

$$\boldsymbol{C}^{-1}=\frac{1}{|\boldsymbol{C}|}\begin{pmatrix}\sigma_2^2 & -\rho\sigma_1\sigma_2\\ -\rho\sigma_1\sigma_2 & \sigma_1^2\end{pmatrix}=\frac{1}{1-\rho^2}\begin{pmatrix}\dfrac{1}{\sigma_1^2} & -\dfrac{\rho}{\sigma_1\sigma_2}\\ -\dfrac{\rho}{\sigma_1\sigma_2} & \dfrac{1}{\sigma_2^2}\end{pmatrix}$$

经计算得

$$(\boldsymbol{x}-\boldsymbol{\mu})^{\mathsf{T}}\boldsymbol{C}^{-1}(\boldsymbol{x}-\boldsymbol{\mu})=\frac{1}{1-\rho^2}\left[\frac{(x_1-\mu_1)^2}{\sigma_1^2}-2\rho\frac{(x_1-\mu_1)(x_2-\mu_2)}{\sigma_1\sigma_2}+\frac{(x_2-\mu_2)^2}{\sigma_2^2}\right]$$

于是 $\boldsymbol{X}$ 的概率密度可写成

$$f(\boldsymbol{x})=\frac{1}{(2\pi)^{\frac{2}{2}}|\boldsymbol{C}|^{\frac{1}{2}}}\exp\left\{-\frac{1}{2}(\boldsymbol{x}-\boldsymbol{\mu})^{\mathsf{T}}\boldsymbol{C}^{-1}(\boldsymbol{x}-\boldsymbol{\mu})\right\}$$

由此, 我们可以推广到 n 维正态变量的情况.

设 $\boldsymbol{X}=(X_1,X_2,\ldots,X_n)^{\mathsf{T}}$ 服从 n 维正态分布，其数学期望为

$$\mathsf{E}(\boldsymbol{X})=(\mathsf{E}(X_1),\mathsf{E}(X_2),\ldots,\mathsf{E}(X_n))^{\mathsf{T}}=(\mu_1,\mu_2,\ldots,\mu_n)\xlongequal{\text{记为}}\boldsymbol{\mu},$$

协方差矩阵为

$$\boldsymbol{\Sigma}=\mathsf{E}[(\boldsymbol{X}-\boldsymbol{\mu})(\boldsymbol{X}-\boldsymbol{\mu})^{\mathsf{T}}].$$

引入列向量 $\boldsymbol{x}=(x_1,x_2,\ldots,x_n)^{\mathsf{T}}$，则 $\boldsymbol{X}$ 的概率密度为

$$f(\boldsymbol{x})=\frac{1}{(2\pi)^{\frac{n}{2}}|\boldsymbol{\Sigma}|^{\frac{1}{2}}}\exp\left\{-\frac{1}{2}(\boldsymbol{x}-\boldsymbol{\mu})^{\mathsf{T}}\boldsymbol{\Sigma}^{-1}(\boldsymbol{x}-\boldsymbol{\mu})\right\}.$$

n 维正态变量具有以下四条重要性质

1. n 维正态变量 $(X_1, X_2, \ldots, X_n)^{\mathsf{T}}$ 中的任意子向量 $(X_{i_1}, X_{i_2}, \ldots, X_{i_k})^{\mathsf{T}}(1 \leq k \leq n)$ 也服从 k 维正态分布. 特别地, 每一个分量 $X_i, i = 1, 2, \ldots, n$ 都是正态变量; 反之, 若 $X_1, X_2, \ldots, X_n$ 都是正态变量, 且相互独立, 则 $(X_1, X_2, \ldots, X_n)^{\mathsf{T}}$ 是 n 维正态变量.
2. n 维随机变量 $(X_1, X_2, \ldots, X_n)^{\mathsf{T}}$ 服从 n 元正态分布 $\Leftrightarrow X_1, X_2, \ldots, X_n$ 的任意线性组合 $l_1X_1 + l_2X_2 + \cdots + l_nX_n$ 服从一元正态分布, 其中 $l_1, l_2, \ldots, l_n$ 不全为零.
3. 若 $(X_1, X_2, \ldots, X_n)^{\mathsf{T}}$ 服从 n 元正态分布, 设 $Y_1, Y_2, \ldots, Y_k$ 是 $X_j(j = 1, 2, \ldots, n)$ 的线性函数, 则 $(Y_1, Y_2, \ldots, Y_k)^{\mathsf{T}}$ 服从 k 维正态分布.
   这一性质称为正态变量的线性变换不变性.
4. 设 $(X_1, X_2, \ldots, X_n)^{\mathsf{T}}$ 服从 n 元正态分布, 则
   $X_1, \ldots, X_n$ 相互独立 $\Leftrightarrow$ $X_1, \ldots, X_n$ 两两不相关 $\Leftrightarrow$ 协方差矩阵为对角矩阵.

Example 45

设二元随机变量 (X, Y) 服从二元正态分布, $X \sim N(0,1)$, $Y \sim N(1,4)$, X 与 Y 的相关系数 $\rho = -\frac{1}{2}$. 求:

1. $\mathrm{Var}(2X - Y)$;
2. $\mathrm{P}(2X > Y)$;
3. $(Z_1, Z_2) = (X + Y, X - Y)$ 的分布.

Solution

1. 利用相关系数定义, 可知协方差

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

Solution (Cont.)

1 于是

$$\begin{aligned}\mathsf{Var}(2X-Y) &= \mathsf{Var}(2X)+\mathsf{Var}(-Y)+2\mathsf{Cov}(2X,-Y)\\ &= 2^2\mathsf{Var}(X)+(-1)^2\mathsf{Var}(Y)-4\mathsf{Cov}(X,Y)\\ &= 4\times 1+1\times 4-4\times(-1)=12\end{aligned}$$

2 有人利用 $\mathsf{P}(2X>Y)=\iint\limits_{2x>y} f(x,y)\mathrm{d}x\mathrm{d}y$ (其中 $f(x,y)$ 为 $N(0,1;1,4;-0.5)$ 的密度函数), 算不出来!
注意到多元正态变量的线性组合仍服从正态分布, 因而有 $2X-Y\sim N(-1,12)$. 所以

$$\mathsf{P}(2X>Y)=\mathsf{P}(2X-Y>0)=1-\Phi\left(\frac{0-(-1)}{\sqrt{12}}\right)=1-\Phi\left(\frac{1}{2\sqrt{3}}\right).$$

Solution (Cont.)

3 根据正态变量的线性变换不变性, (Z_1, Z_2) 也服从二维正态分布.

$$
\begin{aligned}
\mathsf{E}(Z_1) &= \mathsf{E}(X+Y) = \mathsf{E}(X) + \mathsf{E}(Y) = 1,\\
\mathsf{E}(Z_2) &= \mathsf{E}(X-Y) = \mathsf{E}(X) - \mathsf{E}(Y) = -1;\\
\mathsf{Var}(Z_1) &= \mathsf{Var}(X) + \mathsf{Var}(Y) + 2\mathsf{Cov}(X, Y) = 3,\\
\mathsf{Var}(Z_2) &= \mathsf{Var}(X) + \mathsf{Var}(Y) - 2\mathsf{Cov}(X, Y) = 7;\\
\mathsf{Cov}(Z_1, Z_2) &= \mathsf{Cov}(X+Y, X-Y) = \mathsf{Var}(X) - \mathsf{Var}(Y) = -3;
\end{aligned}
$$

$$
\Rightarrow \qquad \rho_{Z_1 Z_2} = \frac{\mathsf{Cov}(Z_1, Z_2)}{\sqrt{Z_1}\sqrt{Z_2}} = \frac{-3}{\sqrt{3}\cdot\sqrt{7}} = -\sqrt{\frac{3}{7}}.
$$

所以 $(Z_1, Z_2) \sim N\left(1, -1; 3, 7; -\sqrt{\frac{3}{7}}\right)$.

Example 46

设二维随机变量 (X, Y) 服从二维正态分布, $X \sim N(1,1)$, $Y \sim N(2,4)$, X 与 Y 的相关系数 ρ, 即 $(X, Y) \sim N(1,2;1,4;\rho)$.

1. 若 $\rho = 0$, 求 $Z = 3X - 2Y$ 的密度函数, $\mathrm{Cov}(2X - Y, X - 2Y)$ 及 $\mathrm{E}(XY)$;
2. 若 $\rho = 1/3$, 求 $\mathrm{Var}(2X - Y)$, $\mathrm{E}(XY)$ 及 $\mathrm{P}(3Y > 4 + 2X)$.

Solution

1. 注意到 Z 为二维正态随机变量分量的线性组合, 故 Z 服从正态分布,而

$$\mathrm{E}(Z) = 3\mathrm{E}(X) - 2\mathrm{E}(Y) = -1, \quad \mathrm{Var}(Z) \xlongequal{\rho=0} 9\mathrm{Var}(X) + 4\mathrm{Var}(Y) = 25.$$

所以 $Z \sim N(-1,25)$, 其密度函数为 $f(z) = \dfrac{1}{5\sqrt{2\pi}} \mathrm{e}^{-\frac{(z+1)^2}{50}}$.

Solution (Cont.)

1
$$\begin{aligned}\mathrm{Cov}(2X-Y, X-2Y) &= 2\mathrm{Var}(X)+2\mathrm{Var}(Y)-5\mathrm{Cov}(X,Y)\\ &= 2\times 1+2\times 4-5\times 0=10,\end{aligned}$$

$$\mathrm{E}(XY) \xlongequal{\rho=0} \mathrm{E}(X)\mathrm{E}(Y)=1\times 2=2.$$

2 由于 $\mathrm{Cov}(X,Y)=\rho\cdot\sqrt{\mathrm{Var}(X)\cdot\mathrm{Var}(Y)}=\frac{1}{3}\times\sqrt{1\times 4}=\frac{2}{3}$, 故

$$\begin{aligned}\mathrm{Var}(2X-Y) &= 4\mathrm{Var}(X)+\mathrm{Var}(Y)-4\mathrm{Cov}(X,Y)=4\times 1+1\times 4-4\times\frac{2}{3}=\frac{16}{3},\\ \mathrm{E}(XY) &= \mathrm{Cov}(X,Y)+\mathrm{E}(X)\mathrm{E}(Y)=1\times 2\times\frac{1}{3}+1\times 2=\frac{8}{3}.\end{aligned}$$

Solution (Cont.)

2 同样地，注意到 $\xi = 3Y - 2X$ 服从正态分布，而

$$\mathrm{E}(\xi) = 3\mathrm{E}(Y) - 2\mathrm{E}(X) = 4,$$

$$\mathrm{Var}(\xi) = 9\mathrm{Var}(Y) + 4\mathrm{Var}(X) + 2\mathrm{Cov}(3Y, -2X) = 36 + 4 - 12 \times 1 \times 2 \times \frac{1}{3} = 32.$$

所以 $\xi \sim N(4, 32)$. 所以

$$\mathrm{P}(3Y > 4 + 2X) = \mathrm{P}(3Y - 2X > 4) = \mathrm{P}(\xi > 4) = 1 - \Phi\left(\frac{4 - \mathrm{E}(\xi)}{\sqrt{\mathrm{Var}(\xi)}}\right) = 1 - \Phi(0) = 0.5.$$

复习思考题

1. 叙述 $\mathrm{E}(X)$ 和 $\mathrm{Var}(X)$ 的定义.
2. 设有一批数据 $x_1, x_2, \ldots, x_n$, 记 $\overline{x} = \dfrac{1}{n}\sum_{i=1}^{n} x_i$, 则 $\sum_{i=1}^{n}(x_i - \overline{x}) = 0$, 对吗?
3. 已知随机变量 X 具有概率密度 $f(x) = \begin{cases} \frac{3(2x-x^2)}{4}, & 0 < x < 2, \\ 0, & \text{其他}. \end{cases}$ 求 $\mathrm{E}(X)$, 试问下列哪种解法是正确的?

解法 1: $\mathrm{E}(X) = \int_{-\infty}^{+\infty} xf(x)\mathrm{d}x = \int_0^2 \frac{3x(2x-x^2)}{4}\mathrm{d}x = 1.$

解法 2: $\mathrm{E}(X) = \begin{cases} \int_0^2 \frac{3x(2x-x^2)}{4}\mathrm{d}x = 1, & 0 < x < 2, \\ \int_{-\infty}^0 0 \cdot x\mathrm{d}x + \int_2^{+\infty} 0 \cdot x\mathrm{d}x = 0, & \text{其他}. \end{cases}$

复习思考题 (Cont.)

4 试述计算随机变量 X 的函数 $g(X)$ 的数学期望 $\mathrm{E}[g(X)]$ 的两种方法.

5 设 $X \sim N(\mu, \sigma^2)$, 用如下两种方法求 $\mathrm{E}(X^2)$:

- $\mathrm{E}(X^2) = \mathrm{Var}(X) + [\mathrm{E}(X)]^2 = \sigma^2 + \mu^2$;
- $\mathrm{E}(X^2) = \mathrm{E}(X \cdot X) = \mathrm{E}(X) \cdot \mathrm{E}(X) = \mu^2$;

两种结果不一样, 问哪一种错? 为什么?

6 设 X 和 Y 为两随机变量, 且已知 $\mathrm{Var}(X) = 6$, $\mathrm{Var}(Y) = 7$, 则 $\mathrm{Var}(X - Y) = \mathrm{Var}(X) - \mathrm{Var}(Y) = 6 - 7 = -1 < 0$, 这与任意一个随机变量的方差都不小于零相矛盾, 为什么?

7 试问 $\mathrm{Var}(X - Y) = \mathrm{Var}(X) + \mathrm{Var}(Y) - 2\mathrm{Cov}(X, Y)$ 对吗?

8 叙述独立性与不相关性的关系.

9 设 X 服从 $U(1, 3)$, 则 $\mathrm{E}(1/X) = 1/\mathrm{E}(X) = 1/2$, 对吗?

复习思考题 (Cont.)

10 某企业生产的袋装水泥，每袋重量服从 $N(50, 2.5^2)$ (单位：斤)．现考虑随机抽取的 100 包水泥的总重量 Y，用以下两种方式表示：

- 设第 i 袋水泥的重量为 $X_i, i=1,2,\ldots,100$，由题意知 $X_i \sim N(50, 2.5^2)$，则 $Y=\sum_{i=1}^{100} X_i$，因此 $Y \sim N(100\times 50, 100\times 2.5^2)$.
- 设一包水泥的重量为 X，由题意知 $X \sim N(50, 2.5^2)$．若将 100 包水泥的总重量看成是 1 包水泥的 100 倍，即 $Y=100X$ 是 X 的线性函数，则

$$\mathrm{E}(Y)=100\mathrm{E}(X)=100\times 50,\quad \mathrm{Var}(Y)=100^2\mathrm{Var}(X)=100^2\times 2.5^2.$$

这两种方法得到的总重量的分布不一样 (因为方差不同，后者方差是前者的 100 倍)，试问哪一种正确？

求：(i) $\mathrm{Var}(X_1-X_2)$; (ii) Y 与 X_1 的相关系数;
(iii) $\mathrm{P}(Y>\mathrm{E}(Y))$; (iv) $\mathrm{P}(Y=\mathrm{E}(Y))$.

谢谢!