第 4 章 随机变量的数字特征

安徽财经大学

统计与应用数学学院

目录

有甲乙两名射击选手打靶，其命中的环数分别记为 X 和 Y，显然 X 和 Y 都是随机变量．设 X 和 Y 的分布列分别为：

X	8	9	10
P	0.3	0.1	0.6

Y	8	9	10
P	0.2	0.4	0.4

现在问，甲和乙哪个射击技术较好？

有甲乙两名射击选手打靶，其命中的环数分别记为 X 和 Y，显然 X 和 Y 都是随机变量．设 X 和 Y 的分布列分别为：

X	8	9	10
P	0.3	0.1	0.6

Y	8	9	10
P	0.2	0.4	0.4

现在问，甲和乙哪个射击技术较好？

由于某射手射中某环的概率大小即反映该射手击中该环的比率，因而上面两个分布列详细刻画了甲、乙两人的射击技术．但要从分布上来比较两人的射击水平的高低并不明显．

因此我们希望能够构造出一些其他的量，它们能够更集中更明显的反映两个选手射击水平的高低．此时一个自然的想法就是考虑他们平均命中的环数．

假定两射击选手各射 N 枪, 则其命中的总环数应当为:

$$甲: 8 \times 0.3N + 9 \times 0.1N + 10 \times 0.6N = 9.3N,$$

$$乙: \quad 8 \times 0.2N + 9 \times 0.4N + 10 \times 0.4N = 9.2N.$$

平均起来甲每枪平均命中 9.3 环, 乙每枪平均命中 9.2 环, 因此有理由判定甲的射击技术要好些.
与此同时可以看到这种反映随机变量取值 “平均” 的数值, 恰好是 X 的可能取值与其对应的概率的乘积之和.

定义 (4.1.1)

设 X 是离散型随机变量，其分布列为

$$P\{X=x_k\}=p_k,\ \ k=1,2,\cdots,$$

若级数 $\sum\limits_{k=1}^{\infty} x_k p_k$ 绝对收敛，记

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

称 $E(X)$ 为随机变量 X 的**数学期望**(Mathematical expectation).

- 定义中对级数要求绝对收敛是必要的. 因为从直观上看, $E(X)$ 的取值应不受 X 值顺序改变的影响, 因此定义中应当允许任意改变求和顺序而不影响级数的收敛性与值.
- 从定义可以知道, 随机变量的数学期望可能不存在, 但当它只取有限个可能值时数学期望一定存在.
- 特别注意的是: 尽管随机变量的取值是随机的, 但其数学期望是一个数, 不再具有随机性. 数学期望有时简称为期望或者均值.

例 (4.1.1)

设 X 为一离散型随机变量，其分布列为

$$P\left(X = (-1)^k \frac{2^k}{k}\right) = \frac{1}{2^k}, \quad k = 1, 2, \cdots$$

试说明 X 的数学期望不存在.

例 (4.1.1)

设 X 为一离散型随机变量，其分布列为

$$P\left(X = (-1)^k \frac{2^k}{k}\right) = \frac{1}{2^k}, \quad k = 1, 2, \cdots$$

试说明 X 的数学期望不存在.

解

考虑级数

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

显然该级数是发散的，于是可以断定 X 的数学期望不存在.

定义 (4.1.2)

设 X 为连续型随机变量，其概率密度为 $f(x)$，若积分 $\int_{-\infty}^{+\infty}|x|f(x)\mathrm{d}x$ 收敛，记

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

称 $E(X)$ 为随机变量 X 的**数学期望**.

常见分布的数学期望

(1) 单点分布的数学期望

设随机变量分布列为 $P(X=c)=1$，则可得其数学期望

$$E(X)=1\times c=c.$$

(2) $0-1$ 分布的数学期望

设随机变量 X 服从参数为 p 的 $0-1$ 分布，则可得其数学期望为:

$$E(X)=1\times p+0\times(1-p)=p.$$

(3) 二项分布 $b(n,p)$ 的数学期望

$$E(X)=\sum_{k=0}^{n}kC_n^kp^k(1-p)^{n-k}=\sum_{k=0}^{n}k\frac{n!}{k!(n-k)!}p^k(1-p)^{n-k}$$
$$=np\sum_{k=1}^{n}\frac{(n-1)!}{(k-1)![(n-1)-(k-1)]!}p^{k-1}(1-p)^{[(n-1)-(k-1)]},$$

令 $k-1=t$, 则

$$E(X)=np\sum_{t=0}^{n-1}\frac{(n-1)!}{t![(n-1)-t]!}p^t(1-p)^{[(n-1)-t]}$$
$$=np[p+(1-p)]^{n-1}=np.$$

(4) 泊松分布 $P(\lambda)$ 的数学期望

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

(5) 几何分布 $Ge(p)$ 的数学期望

$$E(X)=\sum_{k=1}^{+\infty}kq^{k-1}p=p\sum_{k=1}^{+\infty}\frac{\mathrm{d}q^k}{\mathrm{d}q}=p\frac{\mathrm{d}}{\mathrm{d}q}\left(\sum_{k=1}^{+\infty}q^k\right)=p\frac{\mathrm{d}}{\mathrm{d}q}\left(\frac{q}{1-q}\right)=\frac{1}{p}.$$

注意: $q=1-p$.

(6) 均匀分布 $U(a, b)$ 的数学期望

$$E(X) = \int_{-\infty}^{+\infty} xf(x)\mathrm{d}x = \int_a^b x\frac{1}{b-a}\mathrm{d}x = \frac{a+b}{2}.$$

(7) 指数分布 $Exp(\lambda)$ 的数学期望

$$E(X) = \int_0^{+\infty} x\lambda e^{-\lambda x}\mathrm{d}x = -\int_0^{+\infty} x\mathrm{d}e^{-\lambda x} = \int_0^{+\infty} e^{-\lambda x}\mathrm{d}x = \frac{1}{\lambda}.$$

(8) 正态分布 $N(\mu,\sigma^2)$ 的数学期望

$$
\begin{aligned}
E(X) &= \int_{-\infty}^{+\infty} x\frac{1}{\sqrt{2\pi}\sigma}\exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)\mathrm{d}x \\
&= \int_{-\infty}^{+\infty} (x-\mu)\frac{1}{\sqrt{2\pi}\sigma}\exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)\mathrm{d}x \\
&\quad + \int_{-\infty}^{+\infty} \mu\frac{1}{\sqrt{2\pi}\sigma}\exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)\mathrm{d}x \\
&= \int_{-\infty}^{+\infty} t\frac{1}{\sqrt{2\pi}\sigma}\exp\left(-\frac{t^2}{2\sigma^2}\right)\mathrm{d}t \\
&\quad + \mu\int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi}\sigma}\exp\left(-\frac{t^2}{2\sigma^2}\right)\mathrm{d}t \ (t=x-\mu) \\
&= \mu.
\end{aligned}
$$

分布	分布律或密度函数	期望
$0-1$ 分布 $b(1,p)$	$P(X=k)=p^k(1-p)^{1-k},\ k=0,1$	p
二项分布 $b(n,p)$	$P(X=k)=C_n^k p^k(1-p)^{n-k},\ k=0,1,\cdots,n$	np
泊松分布 $P(\lambda)$	$P(X=k)=\dfrac{\lambda^k}{k!}\mathrm{e}^{-\lambda},\ k=0,1,2,\cdots,\ \lambda>0$	λ
几何分布 $Ge(p)$	$P(X=k)=p(1-p)^{k-1},\ k=1,2,\cdots$	$\dfrac{1}{p}$
均匀分布 $U(a,b)$	$f(x)=\dfrac{1}{b-a},\ a<x<b;\ f(x)=0,$ 其他	$\dfrac{a+b}{2}$
指数分布 $Exp(\lambda)$	$f(x)=\lambda\mathrm{e}^{-\lambda x},\ x>0;\ f(x)=0,\ x\leqslant 0$	$\dfrac{1}{\lambda}$
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

例 (4.1.2)

袋中有 k 号球 k 个 $(k=1,\cdots,n)$, 现从中任取一球, 求所得号码的平均值.

例 (4.1.2)

袋中有 k 号球 k 个 $(k=1,\cdots,n)$，现从中任取一球，求所得号码的平均值.

解

显然所得号码是一个随机变量，不妨设为 X，题目要求的是 $E(X)$. 注意到由古典概型知 X 的分布列为

$$P(X=k)=\frac{k}{1+2+\cdots+n}=\frac{2k}{n(n+1)},\quad k=1,2,\cdots,n.$$

所以

$$E(X)=\sum_{k=1}^{n}kP(X=k)=\sum_{k=1}^{n}k\frac{2k}{n(n+1)}=\frac{2}{n(n+1)}\sum_{k=1}^{n}k^2=\frac{2n+1}{3}.$$

定理 (4.1.1)

设 Y 是随机变量 X 的函数 $Y=g(X)$ (g 是连续函数).
(1) X 是离散型随机变量, 它的分布列为 $P(X=x_k)=p_k,\ k=1,2,\cdots$,
若 $\sum\limits_{k=1}^{\infty} g(x_k)p_k$ 绝对收敛, 则有

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

(2) 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

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

- 定理 4.1.1 的重要意义在于当我们求 $E[g(X)]$ 时, 不必知道 $Y = g(X)$ 的分布而只需知道 X 的分布就可以了.
- 当然, 我们也可以由已知的 X 的分布, 先求出其函数 $g(X)$ 的分布, 再根据数学期望的定义去求 $E[g(X)]$, 然而, 求 $Y = g(X)$ 的分布是不容易的, 所以一般不采用后一种方法.
- 上述定理还可以推广到二个或二个以上随机变量的函数情形.

定理 (4.1.2)

设 Z 是随机变量 X, Y 的函数，$Z = g(X, Y)$ (g 是连续函数)，那么 Z 也是一个随机变量，当 (X, Y) 是二维离散型随机变量，其分布律为 $P\{X = x_i, Y = y_j\} = p_{ij}, \quad (i, j = 1, 2, \cdots)$ 时，若 $\sum\limits_i \sum\limits_j g(x_i, y_j)p_{ij}$ 绝对收敛，则有

$$E(Z) = E[g(X, Y)] = \sum_i \sum_j g(x_i, y_j)p_{ij}.$$

当 (X, Y) 是二维连续型随机变量，其概率密度为 $f(x, y)$ 时，若 $\int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} g(x, y)f(x, y)\mathrm{d}x\mathrm{d}y$ 绝对收敛，则有

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

定理 (4.1.2)

特别地有

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

例 (4.1.3)

设随机变量 X 的分布律为:

X	-1	0	1	2
P	0.1	0.2	0.3	0.4

求 $E(1-X^2)$.

例 (4.1.3)

设随机变量 X 的分布律为:

X	-1	0	1	2
P	0.1	0.2	0.3	0.4

求 $E(1-X^2)$.

解

由定理 4.1.1 得

$$\begin{aligned} E(1-X^2) =&[(1-(-1)^2] \times 0.1 + (1-0^2) \times 0.2 \\ &+ (1-1^2) \times 0.3 + (1-2^2) \times 0.4 \\ =& -1. \end{aligned}$$

例 (4.1.4)

设随机变量 (X, Y) 的联合概率密度为

$$f(x, y) = \begin{cases} 4xy, & 0 < x < 1, 0 < y < 1, \\ 0, & \text{其他}. \end{cases}$$

试求 $E(X + Y)$.

例 (4.1.4)

设随机变量 (X, Y) 的联合概率密度为

$$f(x, y) = \begin{cases} 4xy, & 0 < x < 1, 0 < y < 1, \\ 0, & \text{其他}. \end{cases}$$

试求 $E(X+Y)$.

解

由定理 4.1.2 得

$$E(X+Y) = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}(x+y)f(x, y)\mathrm{d}x\mathrm{d}y = \int_0^1\int_0^1 4xy(x+y)\mathrm{d}x\mathrm{d}y = \frac{4}{3}.$$

例 (4.1.5)

某商品销售量 (单位)$X \sim U(10,30)$. 若每销售一个单位可获利 500 元, 若供大于求每积压一单位将损失 100 元, 若供不应求, 可按需调剂, 每单位可以获利 400 元. 问进货量为多少时平均利润最大?

例 (4.1.5)

某商品销售量 (单位)$X \sim U(10,30)$. 若每销售一个单位可获利 500 元，若供大于求每积压一单位将损失 100 元，若供不应求，可按需调剂，每单位可以获利 400 元. 问进货量为多少时平均利润最大?

解

设进货量为 n 个单位，显然进货量应在 $10 \sim 30$ 之间. 用 Y 表示进货量为 n 时的利润. 显然有 $Y = \begin{cases} 500X - 100(n-X), & X \leqslant n, \\ 500n + 400(X-n), & X > n. \end{cases}$

则平均利润为

$$\begin{aligned} E(Y) &= \int_{10}^{n} \frac{1}{20}[500x - 100(n-x)]\mathrm{d}x + \int_{n}^{30} \frac{1}{20}[500n + 400(x-n)]\mathrm{d}x \\ &= -5n^2 + 200n + 7500 = -5(n-20)^2 + 9500. \end{aligned}$$

得 $n = 20$，即进货量为 20 单位，可使平均利润最大.

数学期望的性质

定理 (4.1.3)

设随机变量 X, Y 的数学期望 $E(X), E(Y)$ 存在.

(1) $E(c) = c$, 其中 c 是常数.

(2) $E(cX) = cE(X)$.

(3) $E(X+Y) = E(X) + E(Y)$, $E(aX+bY+c) = aE(X) + bE(Y) + c$.

(4) 若 X, Y 是相互独立的, 则有

$$E(XY) = E(X)E(Y).$$

(5) 若随机变量 $X \geqslant 0$, 则 $E(X) \geqslant 0$, 若随机变量 $X \geqslant Y$, 则 $E(X) \geqslant E(Y)$, 若 $b \geqslant X \geqslant a$, 则 $b \geqslant E(X) \geqslant a$ (a, b 为常数).

仅就连续型的情况我们来证明性质 (3), (4).

证明.

(3) 设二维随机变量 (X, Y) 的概率密度为 $f(x, y)$, 其边际概率密度为 $f_X(x)$, $f_Y(y)$, 则

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

□

性质 (3) 可推广到任意有限个随机变量之和的情形.

证明.

(4) 若 X 和 Y 相互独立, 此时

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

故

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

□

性质 (4) 可推广到任意有限个相互独立的随机变量之积的情形.

例 (4.1.6)

已知 $X \sim N(3, 2^2)$, $Y \sim P(2)$, 试求 $E(2X - Y + 3)$.

例 (4.1.6)

已知 $X \sim N(3, 2^2)$，$Y \sim P(2)$，试求 $E(2X - Y + 3)$.

解

由数学期望的性质得

$$E(2X - Y + 3) = 2E(X) - E(Y) + 3 = 2 \times 3 - 2 + 3 = 7.$$

将一个随机变量分解为几个随机变量的和，然后利用数学期望的性质去进行计算，可以使复杂的计算变得简单.

将一个随机变量分解为几个随机变量的和，然后利用数学期望的性质去进行计算，可以使复杂的计算变得简单.

利用数学期望的性质求二项分布的期望

设 $X_i \sim b(1,p)$，$i=1,2,\cdots,n$，且相互独立，$X \sim b(n,p)$，则 X 可以表示为

$$X = X_1 + X_2 + \cdots + X_n.$$

由数学期望的性质，

$$E(X) = E(X_1) + E(X_2) + \cdots + E(X_n) = np.$$

例 (4.1.7)

一颗均匀的骰子, 连掷 6 次, 用 X 表示出现的点数之和, 求 $E(X)$.

例 (4.1.7)

一颗均匀的骰子，连掷 6 次，用 X 表示出现的点数之和，求 $E(X)$.

解

下面通过将 X 分解成为较为简单的随机变量的和来解决这个问题．为此设 $X_k=$“第 k 次出现的点数”，则有

$$P(X_k=i)=\frac{1}{6},\ \ i=1,2,\cdots,6$$

于是有 $X=X_1+X_2+\cdots+X_6$. 可得

$$E(X)=E(X_1)+E(X_2)+\cdots+E(X_6)=6E(X_1).$$

注意到

$$E(X_1)=\frac{1}{6}(1+2+\cdots+6)=\frac{21}{6}.$$

所以 $E(X)=6\times\dfrac{21}{6}=21.$

例 (4.1.8)

设电流强度 I 服从 $(0,1)$ 上的均匀分布, 电阻 R 的密度函数为

$$f_R(r)=\begin{cases}2r, & 0<r<1,\\ 0, & \text{其他}.\end{cases}$$

I 与 R 相互独立, 试求该电阻上的电压 $V=RI$ 的数学期望.

例 (4.1.8)

设电流强度 I 服从 $(0,1)$ 上的均匀分布, 电阻 R 的密度函数为

$$f_R(r) = \begin{cases} 2r, & 0 < r < 1, \\ 0, & \text{其他}. \end{cases}$$

I 与 R 相互独立, 试求该电阻上的电压 $V = RI$ 的数学期望.

解

由题给条件知

$$E(I) = \frac{1}{2}, \quad E(R) = \int_0^1 r \cdot 2r \mathrm{d}r = \frac{2}{3}.$$

注意到 I, R 的独立性, 有

$$E(V) = E(IR) = E(I)E(R) = \frac{1}{2} \times \frac{2}{3} = \frac{1}{3}.$$

前一节介绍的数学期望虽然较好地刻画随机变量取值的平均状况，但有时候仅知道这个值是不够的．如甲乙两射手各打了 6 发子弹，子弹命中的环数分别为

$$\text{甲}: 10, 7, 9, 8, 10, 6; \;\; \text{乙}: 8, 7, 10, 9, 8, 8$$

问哪个选手的技术较好?

显然甲、乙平均环数都是 8.3 环．但他们的情况不完全相同．谁的稳定性好，则他的技术就好．下面将介绍描述随机变量稳定性的一个量——方差．

1. 随机变量的数学期望
2. 方差
   - 方差的定义
   - 方差的性质
   - 常用分布的方差
3. 协方差与相关系数
4. 矩、协方差矩阵
5. 条件数学期望

方差

定义 (4.2.1)

设 X 是一个随机变量，若 $E[X-E(X)]^2$ 存在，则称

$$D(X) = E[X-E(X)]^2.$$

为 X 的方差(Variance)，有时也记为 $Var(X)$. 称 $\sqrt{D(X)}$ 为随机变量 X 的标准差(Standard deviation) 或均方差(Mean square deviation)，记为 $\sigma(X)$ 或者 σ_X.

方差

定义 (4.2.1)

设 X 是一个随机变量，若 $E[X-E(X)]^2$ 存在，则称

$$D(X)=E[X-E(X)]^2.$$

为 X 的**方差**(Variance)，有时也记为 $Var(X)$. 称 $\sqrt{D(X)}$ 为随机变量 X 的**标准差**(Standard deviation) 或**均方差**(Mean square deviation)，记为 $\sigma(X)$ 或者 σ_X.

由方差定义的数学表达式可以看出，随机变量 X 的方差实际上反映了随机变量的取值与其数学期望的偏离程度.
方差越大，则随机变量的取值越分散，方差越小，则随机变量的取值越集中在其期望附近. 若 X 取值比较集中，则 $D(X)$ 较小，反之，若 X 取值比较分散，则 $D(X)$ 较大.

利用定理 4.1.1 可以得到:
(1) 若离散型随机变量 X 的分布列为 $P\{X=x_k\}=p_k,\ k=1,2,\cdots$, 则

$$D(X)=\sum_{k=1}^{\infty}[x_k-E(X)]^2 p_k.$$

(2) 若连续型随机变量 X 的概率密度为 $f(x)$, 则

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

利用定理 4.1.1 可以得到:

(1) 若离散型随机变量 X 的分布列为 $P\{X=x_k\}=p_k$, $k=1,2,\cdots$, 则

$$D(X)=\sum_{k=1}^{\infty}[x_k-E(X)]^2p_k.$$

(2) 若连续型随机变量 X 的概率密度为 $f(x)$, 则

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

由此可见, 方差 $D(X)$ 是一个常数, 它由随机变量的分布唯一确定, 不再具有随机性. 根据数学期望的性质可以得到计算方差的一个简便公式.

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

例 (4.2.1)

设有甲, 乙两种棉花, 用 X, Y 分别表示甲, 乙两种棉花的纤维的长度 (单位: 毫米), 现从中各抽取等量的样品进行检验, 结果如下表所示.

X	28	29	30	31	32
P	0.10	0.15	0.50	0.15	0.10

Y	28	29	30	31	32
P	0.13	0.17	0.40	0.17	0.13

试求 $D(X)$ 与 $D(Y)$, 并比较它们的质量.

解 (由于)

$$E(X)=28\times 0.1+29\times 0.15+30\times 0.5+31\times 0.15+32\times 0.1=30,$$
$$E(Y)=28\times 0.13+29\times 0.17+30\times 0.4+31\times 0.17+32\times 0.13=30,$$

故得

$$\begin{aligned}
D(X)=&(28-30)^2\times 0.1+(29-30)^2\times 0.15+(30-30)^2\times 0.5\\
&+(31-30)^2\times 0.15+(32-30)^2\times 0.1\\
=&4\times 0.1+1\times 0.15+0\times 0.5+1\times 0.15+4\times 0.1=1.1,\\
D(Y)=&(28-30)^2\times 0.13+(29-30)^2\times 0.17+(30-30)^2\times 0.4\\
&+(31-30)^2\times 0.17+(32-30)^2\times 0.13\\
=&4\times 0.13+1\times 0.17+0\times 0.4+1\times 0.17+4\times 0.13=1.38.
\end{aligned}$$

因 $D(X)<D(Y)$, 所以甲种棉花纤维长度的方差小些, 说明其纤维比较均匀, 故甲种棉花质量较好.

例 (4.2.2)

设 X 的密度函数为 $f(x)=\begin{cases} 2x, & 0\leqslant x\leqslant 1, \\ 0, & \text{其他}. \end{cases}$，试求 $D(X)$.

例 (4.2.2)

设 X 的密度函数为 $f(x) = \begin{cases} 2x, & 0 \leqslant x \leqslant 1, \\ 0, & \text{其他}. \end{cases}$，试求 $D(X)$.

解

由于

$$E(X) = \int_0^1 x \cdot 2x \,\mathrm{d}x = \frac{2}{3} x^3 \Big|_0^1 = \frac{2}{3},$$

$$E(X^2) = \int_0^1 x^2 \cdot 2x \,\mathrm{d}x = \frac{2}{4} x^3 \Big|_0^1 = \frac{1}{2},$$

例 (4.2.2)

设 X 的密度函数为 $f(x)=\begin{cases} 2x, & 0\leqslant x\leqslant 1, \\ 0, & \text{其他}. \end{cases}$，试求 $D(X)$.

解

由于

$$E(X)=\int_0^1 x\cdot 2x\,\mathrm{d}x=\frac{2}{3}x^3\Big|_0^1=\frac{2}{3},$$

$$E(X^2)=\int_0^1 x^2\cdot 2x\,\mathrm{d}x=\frac{2}{4}x^3\Big|_0^1=\frac{1}{2},$$

于是

$$D(X)=E(X^2)-E^2(X)=\frac{1}{2}-\left(\frac{2}{3}\right)^2=\frac{1}{18}.$$

方差的性质

性质

设随机变量 X 与 Y 的方差存在，则

(1) 设 a 为常数，则 $D(a)=0$.

(2) 设 a 为常数，则 $D(aX)=a^2D(X)$.

(3) $D(X\pm Y)=D(X)+D(Y)\pm 2E[(X-E(X))(Y-E(Y))]$.

(4) 若 X, Y **相互独立**，则 $D(X\pm Y)=D(X)+D(Y)$.

方差的性质

性质

设随机变量 X 与 Y 的方差存在, 则

(1) 设 a 为常数, 则 $D(a) = 0$.

(2) 设 a 为常数, 则 $D(aX) = a^2 D(X)$.

(3) $D(X \pm Y) = D(X) + D(Y) \pm 2E[(X - E(X))(Y - E(Y))]$.

(4) 若 X, Y 相互独立, 则 $D(X \pm Y) = D(X) + D(Y)$.

证明.

(1) 因为 $E(a) = a$, 所以 $D(a) = E[a - E(a)]^2 = 0$.

(2) 因为 $E(aX) = aE(X)$, 所以

$$D(aX) = E[aX - E(aX)]^2 = E[a(X - E(X))]^2 = a^2 E[X - E(X)]^2 = a^2 D(X).$$

$\square$

证明.

(3)

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

证明.

(3)

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

(4) 注意到: 当 X 与 Y 相互独立时, $X - E(X)$ 与 $Y - E(Y)$ 也相互独立, 由数学期望的性质有

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

因此有 $D(X \pm Y) = D(X) + D(Y)$, 性质 (4) 得证. □

性质 (4) 可以推广到任意有限多个相互独立的随机变量之和的情况.

例 (4.2.3)

设 X, Y **相互独立，已知** $E(X)=0$，$D(X)=1$，$E(Y)=2$，$D(Y)=3$，**试求** $E(X-2Y)$；$D(X-2Y)$；$E[(X+Y)]^2$；$E[(X-2Y)]^2$.

例 (4.2.3)

设 X, Y **相互独立，已知** $E(X)=0,\ D(X)=1,\ E(Y)=2,\ D(Y)=3$，**试求** $E(X-2Y)$; $D(X-2Y)$; $E[(X+Y)]^2$; $E[(X-2Y)]^2$.

解

由数学期望和方差的性质知:

$$
\begin{aligned}
E(X-2Y) &= E(X)-2E(Y)=0-2\times 2=-4.\\
D(X-2Y) &= D(X)+(-2)^2D(Y)=1+4\times 3=13.\\
E[(X+Y)]^2 &= E(X^2+2XY+Y^2)\\
&= E(X^2)+2E(XY)+E(Y^2)\\
&= D(X)+E^2(X)+2E(X)E(Y)+D(Y)+E^2(Y)\\
&= 1+0^2+2\times 0\times 2+3+2^2=8.\\
E[(X-2Y)]^2 &= D[(X-2Y)]+E^2[(X-2Y)]\\
&= 13+(-4)^2=29.
\end{aligned}
$$

1. 随机变量的数学期望
2. 方差
   - 方差的定义
   - 方差的性质
   - 常用分布的方差
3. 协方差与相关系数
4. 矩、协方差矩阵
5. 条件数学期望

1. 单点分布

1. 单点分布

$$P(X=a)=1;\ D(X)=0.$$

2. 0 – 1 分布

2. $0-1$ 分布

设 X 服从参数为 p 的 $0-1$ 分布，其分布列为:

X	0	1
P	$1-p$	p

$$D(X)=E(X^2)-E^2(X)=1^2\times p-p^2=p(1-p).$$

3. 二项分布 $X \sim b(n,p)$: $D(X) = np(1-p)$

3. 二项分布 $X \sim b(n,p)$: $D(X) = np(1-p)$

事实上, 由上一章知: 若 X_i, $i = 1, 2, \cdots, n$ 相互独立且都服从 $0-1$ 分布 $b(1,p)$, 则有

$$X = \sum_{i=1}^{n} X_i \sim b(n,p),$$

于是

$$D(X) = \sum_{i=1}^{n} D(X_i) = np(1-p).$$

4. 泊松分布 $X \sim p(\lambda)$: $D(X) = \lambda$

4. 泊松分布 $X \sim p(\lambda)$: $D(X) = \lambda$

设 X 服从参数为 λ 的泊松分布，由上一节知 $E(X) = \lambda$，又

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

所以

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

从而有

$$D(X) = E(X^2) - [E(X)]^2 = \lambda^2 + \lambda - \lambda^2 = \lambda.$$

5. 几何分布 $X \sim Ge(p)$: $D(X) = \dfrac{1-p}{p^2}$

5. 几何分布 $X \sim Ge(p)$: $D(X) = \dfrac{1-p}{p^2}$

设 X 服从参数为 p 的几何分布，由上一节知 $E(X) = \dfrac{1}{p}$，又

$$E[X(X-1)] = \sum_{k=1}^{+\infty} k(k-1)q^{k-1}p = pq\sum_{k=1}^{+\infty} k(k-1)q^{k-2} = pq\sum_{k=1}^{+\infty} \frac{\mathrm{d}^2}{\mathrm{d}q^2}(q^k)$$

$$= pq\frac{\mathrm{d}^2}{\mathrm{d}q^2}\left(\sum_{k=1}^{+\infty} q^k\right) = pq\frac{\mathrm{d}^2}{\mathrm{d}q^2}\left(\frac{q}{1-q}\right) = \frac{2(1-p)}{p^2},$$

$$E(X^2) = E[X(X-1)+X] = E[X(X-1)] + E(X) = \frac{2(1-p)}{p^2} + \frac{1}{p} = \frac{2-p}{p^2},$$

从而有

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

6. 均匀分布 $X \sim U(a, b)$: $D(X) = \dfrac{(b-a)^2}{12}$

6. 均匀分布 $X \sim U(a, b)$: $D(X) = \dfrac{(b-a)^2}{12}$

设 X 服从 (a, b) 上的均匀分布, 由上一节知 $E(X) = \dfrac{a+b}{2}$, 又

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

所以

$$D(X) = E(X^2) - [E(X)]^2 = \frac{1}{3}(a^2 + ab + b^2) - \frac{1}{4}(a+b)^2 = \frac{(b-a)^2}{12}.$$

7. 指数分布 $X \sim Exp(\lambda)$: $D(X) = \dfrac{1}{\lambda^2}$

7. 指数分布 $X \sim Exp(\lambda)$: $D(X) = \dfrac{1}{\lambda^2}$

设 X 服从参数为 λ 的指数分布，由上一节知 $E(X) = \dfrac{1}{\lambda}$，又

$$\begin{aligned} E(X^2) &= \int_0^{+\infty} x^2 \lambda e^{-\lambda x}\,\mathrm{d}x = -\int_0^{+\infty} x^2\,\mathrm{d}(e^{-\lambda x}) \\ &= \frac{2}{\lambda}\int_0^{+\infty} x\lambda e^{-\lambda x}\,\mathrm{d}x = \frac{2}{\lambda}E(X) = \frac{2}{\lambda^2}, \end{aligned}$$

所以

$$D(X) = E(X^2) - [E(X)]^2 = \frac{1}{\lambda^2}.$$

8. 正态分布 $X \sim N(\mu, \sigma^2)$: $D(X) = \sigma^2$

8. 正态分布 $X \sim N(\mu, \sigma^2)$: $D(X) = \sigma^2$

设 $X \sim N(\mu, \sigma^2)$, 由上一节知 $E(X) = \mu$, 从而

$$D(X) = \int_{-\infty}^{+\infty} [x - E(X)]^2 f(x)\mathrm{d}x = \int_{-\infty}^{+\infty} (x-\mu)^2 \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \mathrm{d}x.$$

令 $\dfrac{x-\mu}{\sigma} = t$, 则

$$\begin{aligned} D(X) &= \frac{\sigma^2}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} t^2 e^{-\frac{t^2}{2}} \mathrm{d}t = \frac{\sigma^2}{\sqrt{2\pi}} \left(-te^{-\frac{t^2}{2}} \Big|_{-\infty}^{+\infty} + \int_{-\infty}^{+\infty} e^{-\frac{t^2}{2}} \mathrm{d}t\right) \\ &= \frac{\sigma^2}{\sqrt{2\pi}} (0 + \sqrt{2\pi}) = \sigma^2. \end{aligned}$$

- 由此可知：正态分布的概率密度中的两个参数 μ 和 σ 分别是该分布的数学期望和均方差. 因而正态分布完全可由它的数学期望和方差所确定.
- 再者，由上一章例 3.4.5 知道，若 $X_i \sim N(\mu_i, \sigma_i^2)$，$i = 1, 2, \cdots, n$，且它们相互独立，则它们的线性组合 $c_1X_1 + c_2X_2 + \cdots + c_nX_n$ ($c_1, c_2, \cdots, c_n$ 是不全为零的常数) 仍然服从正态分布.

- 由此可知: 正态分布的概率密度中的两个参数 μ 和 σ 分别是该分布的数学期望和均方差. 因而正态分布完全可由它的数学期望和方差所确定.
- 再者, 由上一章例 3.4.5 知道, 若 $X_i \sim N(\mu_i, \sigma_i^2)$, $i = 1, 2, \cdots, n$, 且它们相互独立, 则它们的线性组合 $c_1X_1 + c_2X_2 + \cdots + c_nX_n$ ($c_1, c_2, \cdots, c_n$ 是不全为零的常数) 仍然服从正态分布.
- 于是由数学期望和方差的性质知道:

$$\boxed{c_1X_1 + c_2X_2 + \cdots + c_nX_n \sim N\left(\sum_{i=1}^{n} c_i\mu_i, \sum_{i=1}^{n} c_i^2\sigma_i^2\right).}$$

这是一个重要的结果.

分布	分布律或密度函数	期望	方差
$b(1,p)$	$P(X=k)=p^k(1-p)^{1-k},\ k=0,1$	p	$p(1-p)$
$b(n,p)$	$P(X=k)=C_n^k p^k(1-p)^{n-k},\ k=0,1,\cdots,n$	np	$np(1-p)$
$P(\lambda)$	$P(X=k)=\dfrac{\lambda^k}{k!}\mathrm{e}^{-\lambda},\ k=0,1,2,\cdots,\ \lambda>0$	λ	λ
$Ge(p)$	$P(X=k)=p(1-p)^{k-1},\ k=1,2,\cdots$	$\dfrac{1}{p}$	$\dfrac{1-p}{p^2}$
$U(a,b)$	$f(x)=\dfrac{1}{b-a},\ a<x<b;\ f(x)=0,$ 其他	$\dfrac{a+b}{2}$	$\dfrac{(b-a)^2}{12}$
$Exp(\lambda)$	$f(x)=\lambda\mathrm{e}^{-\lambda x},\ x>0;\ f(x)=0,\ x\leqslant 0$	$\dfrac{1}{\lambda}$	$\dfrac{1}{\lambda^2}$
$N(\mu,\sigma^2)$	$f(x)=\dfrac{1}{\sqrt{2\pi}\sigma}\mathrm{e}^{-\frac{(x-\mu)^2}{2\sigma^2}},\ -\infty<x<+\infty$	μ	σ^2

例 (4.2.4)

设 $X \sim U(1,4)$, $Y \sim N(4,1)$, X, Y 相互独立, 求 $E(X-2Y)$, $D(X-2Y)$.

例 (4.2.4)

设 $X \sim U(1,4)$, $Y \sim N(4,1)$, X, Y 相互独立, 求 $E(X-2Y)$, $D(X-2Y)$.

证明.

由期望与方差的性质知:

$$E(X-2Y) = E(X) - 2E(Y) = \frac{1+4}{2} - 2\times 4 = -\frac{11}{2}.$$

$$D(X-2Y) = D(X) + 2^2 D(Y) = \frac{(4-1)^2}{12} + 4\times 1 = \frac{19}{4}.$$

□

例 (4.2.5)

随机变量 $X \sim N(1,1)$, 设 $Y = 1 - 2X$, 试写出 Y 的密度函数.

例 (4.2.5)

随机变量 $X \sim N(1,1)$, 设 $Y = 1-2X$, 试写出 Y 的密度函数.

解

注意到 $Y = 1-2X$ 也是正态分布, 故只需要确定其期望和方差就可以写出其密度函数.

$$E(Y) = E(1-2X) = 1-2E(X) = 1-2\times 1 = -1,$$
$$D(Y) = D(1-2X) = 0+(-2)^2 D(X) = 4\times 1 = 4.$$

所以 $Y \sim N(-1,4)$. Y 的密度函数为: $f_Y(y) = \dfrac{1}{2\sqrt{2\pi}} e^{-(y+1)^2/8}$.

例 (4.2.6)

设活塞的直径 (以 cm 计) $X \sim N(22.40, 0.03^2)$, 气缸的直径 $Y \sim N(22.50, 0.04^2)$, X, Y 相互独立, 任取一只活塞, 任取一只气缸, 求活塞能装入气缸的概率.

例 (4.2.6)

设活塞的直径 (以 cm 计) $X \sim N(22.40, 0.03^2)$, 气缸的直径 $Y \sim N(22.50, 0.04^2)$, X, Y 相互独立, 任取一只活塞, 任取一只气缸, 求活塞能装入气缸的概率.

解

由题知所求概率为: $P\{X < Y\} = P\{X - Y < 0\}$.
令 $Z = X - Y$, 则:

$$E(Z) = E(X) - E(Y) = 22.40 - 22.50 = -0.10,$$
$$D(Z) = D(X) + D(Y) = 0.03^2 + 0.04^2 = 0.05^2,$$

即 $Z \sim N(-0.10, 0.05^2)$. 故有

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

1. 随机变量的数学期望
2. 方差
3. 协方差与相关系数
   - 协方差
   - 协方差的性质
   - 相关系数及其性质
4. 矩、协方差矩阵
5. 条件数学期望

定义 (4.3.1)

设 (X, Y) 为二维随机变量，称

$$\operatorname{cov}(X, Y) := E\{[X - E(X)][Y - E(Y)]\},$$

为随机变量 X, Y 的**协方差**(Covariance).

定义 (4.3.1)

设 (X, Y) 为二维随机变量，称

$$\operatorname{cov}(X, Y) := E\{[X - E(X)][Y - E(Y)]\},$$

为随机变量 X，Y 的**协方差**(Covariance).

若 (X, Y) 为二维离散型随机变量，其联合分布列为 $P\{X = x_i, Y = y_j\} = p_{ij},\ i, j = 1, 2, \cdots$，则有

$$\operatorname{cov}(X, Y) = \sum_i \sum_j [x_i - E(X)][y_j - E(Y)] p_{ij}.$$

若 (X, Y) 为二维连续型随机变量，其概率密度为 $f(x, y)$，则有

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

定义 (4.3.1)

设 (X, Y) 为二维随机变量，称

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

为随机变量 X, Y 的**协方差**(Covariance).

定义 (4.3.1)

设 (X, Y) 为二维随机变量，称

$$\operatorname{cov}(X, Y) := E\{[X - E(X)][Y - E(Y)]\},$$

为随机变量 X, Y 的**协方差**(Covariance).

利用数学期望的性质，可以得到计算协方差的一个简便公式.

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

$$\boxed{\operatorname{cov}(X, Y) = E(XY) - E(X)E(Y)}$$

例 (4.3.1)

设 (X, Y) 的分布律如下表所示.

$X \backslash Y$	0	1	$X = x_i$
0	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{2}$
1	$\frac{1}{3}$	$\frac{1}{6}$	$\frac{1}{2}$
$Y = y_j$	$\frac{7}{12}$	$\frac{5}{12}$	

求 $\mathrm{cov}(X, Y)$.

解

$$E(X) = 0 \times \frac{1}{2} + 1 \times \frac{1}{2} = \frac{1}{2},$$

$$E(Y) = 0 \times \frac{7}{12} + 1 \times \frac{5}{12} = \frac{5}{12},$$

$$E(XY) = 1 \times \frac{1}{6} = \frac{1}{6},$$

故

$$\operatorname{cov}(X, Y) = E(XY) - E(X)E(Y) = \frac{1}{6} - \frac{1}{2} \times \frac{5}{12} = -\frac{1}{24}.$$

例 (4.3.2)

设 (X, Y) 的概率密度为

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

求 $\operatorname{cov}(X, Y)$.

例 (4.3.2)

设 (X, Y) 的概率密度为

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

求 $\operatorname{cov}(X, Y)$.

解 (由联合密度不难计算出边际密度)

$$f_X(x) = \int_{-\infty}^{+\infty} f(x, y)\mathrm{d}y = \begin{cases} x + \dfrac{1}{2}, & 0 < x < 1, \\ 0, & \text{其他}, \end{cases}$$

$$f_Y(y) = \int_{-\infty}^{+\infty} f(x, y)\mathrm{d}x = \begin{cases} y + \dfrac{1}{2}, & 0 < y < 1, \\ 0, & \text{其他}, \end{cases}$$

解

于是有

$$E(X)=\int_0^1 x(x+\frac{1}{2})\mathrm{d}x=\frac{7}{12},\ E(Y)=\int_0^1 y(y+\frac{1}{2})\mathrm{d}y=\frac{7}{12},$$

$$E(XY)=\int_0^1\int_0^1 xy(x+y)\mathrm{d}x\mathrm{d}y=\int_0^1\int_0^1 x^2y\mathrm{d}x\mathrm{d}y+\int_0^1\int_0^1 xy^2\mathrm{d}x\mathrm{d}y=\frac{1}{3},$$

因此

$$\mathrm{cov}(X,Y)=E(XY)-E(X)E(Y)=\frac{1}{3}-\frac{7}{12}\times\frac{7}{12}=-\frac{1}{144}.$$

协方差的性质

性质

(1) $\text{cov}(X, X) = D(X)$;

(2) $\text{cov}(X, Y) = \text{cov}(Y, X)$;

(3) 若 X 与 Y 相互独立，则 $\text{cov}(X, Y) = 0$;

(4) 若 a, b 为常数，则 $\text{cov}(aX, bY) = ab\text{cov}(X, Y)$;

(5) $\text{cov}(X_1 + X_2, Y) = \text{cov}(X_1, Y) + \text{cov}(X_2, Y)$;

(6) 若 a 为常数，则 $\text{cov}(a, X) = 0$.

定义 (4.3.2)

设 (X, Y) 为二维随机变量，$D(X) > 0$，$D(Y) > 0$，称

$$\rho_{XY} := \frac{\mathrm{cov}(X, Y)}{\sqrt{D(X)}\sqrt{D(Y)}}$$

为随机变量 X, Y 的相关系数(Correlation coefficient). 它是一个标量，没有量纲，是一个刻画 X 与 Y 相关程度的一个数字特征.

定义 (4.3.2)

设 (X, Y) 为二维随机变量, $D(X) > 0$, $D(Y) > 0$, 称

$$\rho_{XY} := \frac{\operatorname{cov}(X, Y)}{\sqrt{D(X)}\sqrt{D(Y)}}$$

为随机变量 X, Y 的**相关系数**(Correlation coefficient). 它是一个标量, 没有量纲, 是一个刻画 X 与 Y 相关程度的一个数字特征.

注:

设 (X, Y) 为二维随机变量, $D(X) > 0$, $D(Y) > 0$, 令

$$X^* = \frac{X - E(X)}{\sqrt{D(X)}}, \qquad Y^* = \frac{Y - E(Y)}{\sqrt{D(Y)}},$$

则有 $E(X^*) = E(Y^*) = 0$, $D(X^*) = D(Y^*) = 1$, 且

$$\rho_{XY} = \operatorname{cov}(X^*, Y^*).$$

柯西 · 施瓦茨 (Cauchy-Schwarz) 不等式:

设有两个随机变量 X, Y, 且 $E(X^2)$, $E(Y^2)$ 都存在, 则有:

$$[E(XY)]^2 \leqslant E(X^2)E(Y^2),$$

等号成立当且仅当存在常数 t_0, 使得 $P\{Y = t_0 X\} = 1$.

柯西 · 施瓦茨 (Cauchy-Schwarz) 不等式:

设有两个随机变量 X, Y, 且 $E(X^2), E(Y^2)$ 都存在, 则有:

$$[E(XY)]^2 \leqslant E(X^2)E(Y^2),$$

等号成立当且仅当存在常数 t_0, 使得 $P\{Y = t_0 X\} = 1$.

证明.

对于任意的实数 t, 令 $g(t) = E[(tX - Y)^2]$, 故有

$$g(t) = t^2 E(X^2) - 2tE(XY) + E(Y^2)$$

是关于 t 的二次函数. 注意到对于任意的实数 t, $g(t) \geqslant 0$. 因而判别式

$$\Delta = 4[E(XY)]^2 - 4E(X^2)E(Y^2) \leqslant 0.$$

即 $[E(XY)]^2 \leqslant E(X^2)E(Y^2)$. 此外, $g(t) = 0$ 有一个重根 t_0 存在等价于 $[E(XY)]^2 = E(X^2)E(Y^2)$. 即 $E[(t_0X - Y)^2] = 0 \Rightarrow P\{Y = t_0X\} = 1$. □

相关系数的重要性质

性质 (4.3.1)

$|\rho_{XY}| \leqslant 1.$

相关系数的重要性质

性质 (4.3.1)

$|\rho_{XY}| \leqslant 1.$

证明.

方法一: 注意到 $X-E(X)$, $Y-E(Y)$ 也是随机变量, 对它们们应用柯西 – 许瓦兹不等式可得

$$[E(X-E(X))(Y-E(Y))]^2 \leqslant E[X-E(X)]^2 E[Y-E(Y)]^2$$

即 $\operatorname{cov}^2(X, Y) \leqslant D(X)D(Y)$. 所以

$$|\rho_{XY}|^2 = \frac{\operatorname{cov}^2(X, Y)}{D(X)D(Y)} \leqslant 1.$$

□

性质 (4.3.1)

$|\rho_{XY}| \leqslant 1.$

证明.

方法二: 因为

$$
\begin{aligned}
& E\left[\frac{X-E(X)}{\sqrt{D(X)}} \pm \frac{Y-E(Y)}{\sqrt{D(Y)}}\right]^2 \\
= & E\left[\frac{X-E(X)}{\sqrt{D(X)}}\right]^2 + E\left[\frac{Y-E(Y)}{\sqrt{D(Y)}}\right]^2 \pm 2E\left[\frac{(X-E(X))(Y-E(Y))}{\sqrt{D(X)}\sqrt{D(Y)}}\right] \\
= & \frac{E[X-E(X)]^2}{D(X)} + \frac{E[Y-E(Y)]^2}{D(Y)} \pm 2\frac{E[(X-E(X))(Y-E(Y))]}{\sqrt{D(X)}\sqrt{D(Y)}} \\
= & 1 + 1 \pm 2\rho_{XY} = 2 \pm 2\rho_{XY} \geqslant 0,
\end{aligned}
$$

所以, $|\rho_{XY}| \leqslant 1.$ $\square$

性质 (补充)

$$\rho_{XY}=1 \Leftrightarrow P\{X^*=Y^*\}=1 \Leftrightarrow P\left\{\frac{X-E(X)}{\sqrt{D(X)}}=\frac{Y-E(Y)}{\sqrt{D(Y)}}\right\}=1.$$

$$\rho_{XY}=-1 \Leftrightarrow P\{X^*=-Y^*\}=1 \Leftrightarrow P\left\{\frac{X-E(X)}{\sqrt{D(X)}}=-\frac{Y-E(Y)}{\sqrt{D(Y)}}\right\}=1.$$

性质 (补充)

$$\rho_{XY}=1 \Leftrightarrow P\{X^*=Y^*\}=1 \Leftrightarrow P\left\{\frac{X-E(X)}{\sqrt{D(X)}}=\frac{Y-E(Y)}{\sqrt{D(Y)}}\right\}=1.$$

$$\rho_{XY}=-1 \Leftrightarrow P\{X^*=-Y^*\}=1 \Leftrightarrow P\left\{\frac{X-E(X)}{\sqrt{D(X)}}=-\frac{Y-E(Y)}{\sqrt{D(Y)}}\right\}=1.$$

性质 (4.3.2)

$|\rho_{XY}|=1$ 的充分必要条件是存在常数 a, b 使

$$P\{Y=aX+b\}=1,\ (a\neq 0).$$

$\rho_{XY}=1$的充要条件是存在常数 $a>0, b$ 使 $P\{Y=aX+b\}=1.$
$\rho_{XY}=-1$的充要条件是存在常数 $a<0, b$ 使 $P\{Y=aX+b\}=1.$

定义 (4.3.3)

如果随机变量 X, Y 的相关系数 $\rho_{XY}=0$, 则称X 与 Y 不相关. 它表明 X, Y 之间不存在线性关系.
若 $\rho_{XY}>0$, 则称X 与 Y 正相关; 若 $\rho_{XY}<0$, 则称X 与 Y 负相关.
若 $\rho_{XY}=1$, 称X 与 Y 完全正相关; 若 $\rho_{XY}=-1$, 称X 与 Y 完全负相关.

定义 (4.3.3)

如果随机变量 X, Y 的相关系数 $\rho_{XY}=0$, 则称 X 与 Y 不相关. 它表明 X, Y 之间不存在线性关系.
若 $\rho_{XY}>0$, 则称 X 与 Y 正相关; 若 $\rho_{XY}<0$, 则称 X 与 Y 负相关.
若 $\rho_{XY}=1$, 称 X 与 Y 完全正相关; 若 $\rho_{XY}=-1$, 称 X 与 Y 完全负相关.

性质 (4.3.3)

若 X, Y 相互独立, 则 X, Y 一定不相关.

定义 (4.3.3)

如果随机变量 X, Y 的相关系数 $\rho_{XY}=0$, 则称 X 与 Y 不相关. 它表明 X, Y 之间不存在线性关系.
若 $\rho_{XY}>0$, 则称 X 与 Y 正相关; 若 $\rho_{XY}<0$, 则称 X 与 Y 负相关.
若 $\rho_{XY}=1$, 称 X 与 Y 完全正相关; 若 $\rho_{XY}=-1$, 称 X 与 Y 完全负相关.

性质 (4.3.3)

若 X, Y 相互独立, 则 X, Y 一定不相关.

证明.

注意到 $\operatorname{cov}(X, Y)=E(XY)-E(X)E(Y)$. 又 X, Y 独立, 所以有

$$E(XY)=E(X)E(Y),\ \operatorname{cov}(X, Y)=0,\ \rho_{XY}=0.$$

$\square$

性质 (4.3.3)

若 X, Y 相互独立, 则 X, Y 一定不相关.

定理 (对于随机变量 X, Y, 以下事实是等价的:)

(1) $\rho_{XY} = 0$;

(2) X 与 Y 不相关;

(3) $\operatorname{cov}(X, Y) = 0$;

(4) $E(XY) = E(X)E(Y)$;

(5) $D(X+Y) = D(X) + D(Y)$.

(6) $D(X-Y) = D(X) + D(Y)$.

例 (4.3.3)

设 X 服从 $[0,2\pi]$ 上均匀分布，$Y=\cos X$，$Z=\cos(X+a)$，这里 a 是常数. 求 ρ_{YZ}.

例 (4.3.3)

设 X 服从 $[0, 2\pi]$ 上均匀分布，$Y = \cos X$，$Z = \cos(X + a)$，这里 a 是常数．求 ρ_{YZ}.

解

$$E(Y) = \int_0^{2\pi} \cos x \cdot \frac{1}{2\pi} \mathrm{d}x = 0,\ E(Z) = \frac{1}{2\pi} \int_0^{2\pi} \cos(x + a) \mathrm{d}x = 0,$$

$$D(Y) = E\{[Y - E(Y)]^2\} = \frac{1}{2\pi} \int_0^{2\pi} \cos^2 x \mathrm{d}x = \frac{1}{2},$$

$$D(Z) = E\{[Z - E(Z)]^2\} = \frac{1}{2\pi} \int_0^{2\pi} \cos^2 (x + a) \mathrm{d}x = \frac{1}{2},$$

$$\mathrm{cov}(Y, Z) = E\{[Y - E(Y)][Z - E(Z)]\} = \frac{1}{2\pi} \int_0^{2\pi} \cos x \cdot \cos(x + a) \mathrm{d}x = \frac{\cos a}{2},$$

解

因此

$$\rho_{YZ}=\frac{\operatorname{cov}(Y,Z)}{\sqrt{D(Y)}\sqrt{D(Z)}}=\frac{\frac{1}{2}\cos a}{\sqrt{\frac{1}{2}}\cdot\sqrt{\frac{1}{2}}}=\cos a.$$

注意到: 当 $a=\dfrac{\pi}{2}$ 或 $a=\dfrac{3\pi}{2}$ 时, $\rho_{YZ}=0$, 这时 Y 与 Z 不相关, 但这时却有 $Y^2+Z^2=1$, 因此, Y 与 Z 不独立. 或者, 令 $a=\dfrac{3\pi}{2}$, 则 $Z=\sin X$, 设 Y,Z 的联合分布函数为 $F(y,z)$, 易证

$$F\left(\frac{\sqrt{2}}{2},\frac{\sqrt{2}}{2}\right)=\frac{1}{2}\neq F_Y\left(\frac{\sqrt{2}}{2}\right)F_Z\left(\frac{\sqrt{2}}{2}\right)=\frac{3}{4}\cdot\frac{3}{4}.$$

解

因此

$$\rho_{YZ}=\frac{\operatorname{cov}(Y,Z)}{\sqrt{D(Y)}\sqrt{D(Z)}}=\frac{\frac{1}{2}\cos a}{\sqrt{\frac{1}{2}}\cdot\sqrt{\frac{1}{2}}}=\cos a.$$

注意到: 当 $a=\dfrac{\pi}{2}$ 或 $a=\dfrac{3\pi}{2}$ 时, $\rho_{YZ}=0$, 这时 Y 与 Z 不相关, 但这时却有 $Y^2+Z^2=1$, 因此, Y 与 Z 不独立. 或者, 令 $a=\dfrac{3\pi}{2}$, 则 $Z=\sin X$, 设 Y,Z 的联合分布函数为 $F(y,z)$, 易证

$$F\left(\frac{\sqrt{2}}{2},\frac{\sqrt{2}}{2}\right)=\frac{1}{2}\neq F_Y\left(\frac{\sqrt{2}}{2}\right)F_Z\left(\frac{\sqrt{2}}{2}\right)=\frac{3}{4}\cdot\frac{3}{4}.$$

这个例子说明: 当两个随机变量不相关时, 它们并不一定相互独立, 它们之间还可能存在其他的函数关系.

例 (4.3.4)

设 (X, Y) 服从二维正态分布，它的概率密度为

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

求 $\operatorname{cov}(X, Y)$ 和 ρ_{XY}.

解

可以计算得 (X, Y) 的边际概率密度为

$$f_X(x) = \frac{1}{\sqrt{2\pi}\sigma_1} e^{-\frac{(x-\mu_1)^2}{2\sigma_1^2}}, \quad -\infty < x < +\infty,$$

$$f_Y(y) = \frac{1}{\sqrt{2\pi}\sigma_2} e^{-\frac{(y-\mu_2)^2}{2\sigma_2^2}}, \quad -\infty < y < +\infty,$$

故 $E(X) = \mu_1$, $E(Y) = \mu_2$, $D(X) = \sigma_1^2$, $D(Y) = \sigma_2^2$. 而

$$\mathrm{cov}(X, Y) = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}(x-\mu_1)(y-\mu_2)f(x, y)\mathrm{d}x\mathrm{d}y = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \times$$

$$\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}(x-\mu_1)(y-\mu_2)e^{-\frac{(x-\mu_1)^2}{2\sigma_1^2}} e^{-\frac{1}{2(1-\rho^2)}\left[\frac{y-\mu_2}{\sigma_2}-\rho\frac{x-\mu_1}{\sigma_1}\right]^2}\mathrm{d}x\mathrm{d}y.$$

解

令 $t = \dfrac{1}{\sqrt{1-\rho^2}}\left(\dfrac{y-\mu_2}{\sigma_2} - \rho\dfrac{x-\mu_1}{\sigma_1}\right)$, $u = \dfrac{x-\mu_1}{\sigma_1}$, 则

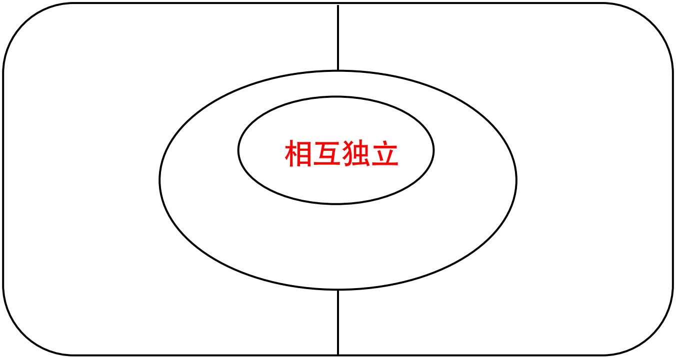
$$\begin{aligned}
\operatorname{cov}(X, Y) =& \frac{1}{2\pi}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}(\sigma_1\sigma_2\sqrt{1-\rho^2}\,tu + \rho\sigma_1\sigma_2 u^2)e^{-\frac{u^2}{2}-\frac{t^2}{2}}\,\mathrm{d}t\mathrm{d}u \\
=& \frac{\sigma_1\sigma_2\rho}{2\pi}\left(\int_{-\infty}^{+\infty} u^2 e^{-\frac{u^2}{2}}\mathrm{d}u\right)\left(\int_{-\infty}^{+\infty} e^{-\frac{t^2}{2}}\mathrm{d}t\right) \\
&+ \frac{\sigma_1\sigma_2\sqrt{1-\rho^2}}{2\pi}\left(\int_{-\infty}^{+\infty} ue^{-\frac{u^2}{2}}\mathrm{d}u\right)\left(\int_{-\infty}^{+\infty} te^{-\frac{t^2}{2}}\mathrm{d}t\right) \\
=& \frac{\rho\sigma_1\sigma_2}{2\pi}\sqrt{2\pi}\cdot\sqrt{2\pi} = \rho\sigma_1\sigma_2.
\end{aligned}$$

于是

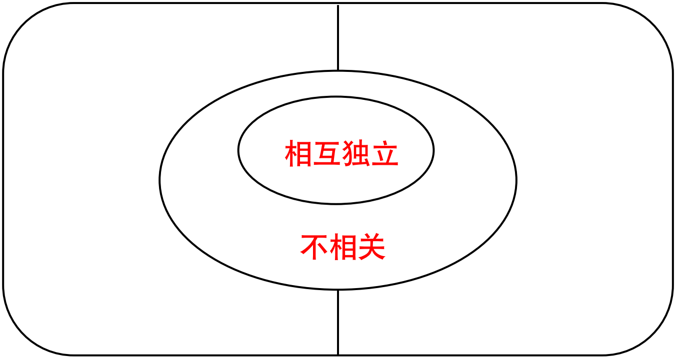
$$\rho_{XY} = \frac{\operatorname{cov}(X, Y)}{\sqrt{D(X)}\sqrt{D(Y)}} = \rho.$$

- 这说明二维正态随机变量 (X, Y) 的概率密度中的参数 ρ 就是 X 和 Y 的相关系数，从而二维正态随机变量的分布完全可由 X, Y 的各自的数学期望、方差以及它们的相关系数所确定.
- 随机变量的独立性和不相关性是两个不同的概念. 两个随机变量相互独立，是说它们在取值上互不影响. 而随机变量不相关，是说它们在取值上没有线性关系.
- 从上文我们已经看到，一般地：随机变量相互独立，则它们必不相关(性质 4.3.3)，但如果两个随机变量不相关，并不能够肯定它们独立 (例 4.3.3).
- 但若 (X, Y) 服从二维正态分布，那么 X 和 Y 相互独立的充要条件是 $\rho = 0$，即 X 与 Y 不相关.

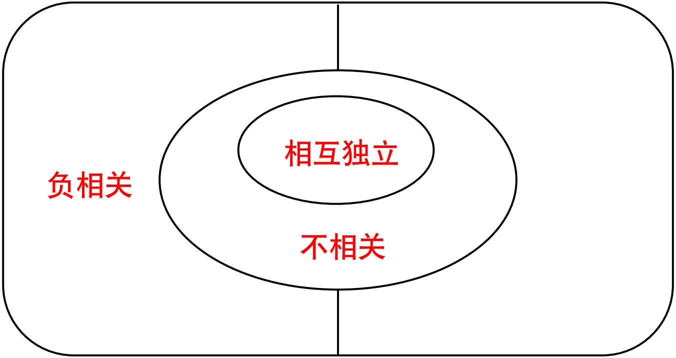
相互独立与线性无关、线性相关之间的关系



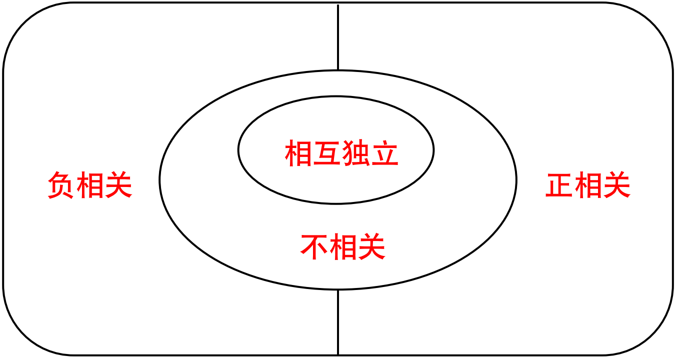
相互独立与线性无关、线性相关之间的关系



相互独立与线性无关、线性相关之间的关系



相互独立与线性无关、线性相关之间的关系



矩

定义 (4.4.1)

设 X 和 Y 是随机变量, 若 $E(X^k)$, $k=1,2,\cdots$ 存在, 称它为 X 的k 阶原点矩, 简称k 阶矩.
若 $E[X-E(X)]^k$, $k=1,2,\cdots$ 存在, 称它为 X 的k 阶中心矩.
若 $E(X^kY^l)$, $k,l=1,2,\cdots$ 存在, 称它为 X 和 Y 的$k+l$ 阶混合矩.
若 $E\{[X-E(X)]^k[Y-E(Y)]^l\}$ 存在, 称它为 X 和 Y 的$k+l$ 阶混合中心矩.

矩

定义 (4.4.1)

设 X 和 Y 是随机变量, 若 $E(X^k)$, $k=1,2,\cdots$ 存在, 称它为 X 的 k 阶原点矩, 简称 k 阶矩.

若 $E[X-E(X)]^k$, $k=1,2,\cdots$ 存在, 称它为 X 的 k 阶中心矩.

若 $E(X^kY^l)$, $k,l=1,2,\cdots$ 存在, 称它为 X 和 Y 的 $k+l$ 阶混合矩.

若 $E\{[X-E(X)]^k[Y-E(Y)]^l\}$ 存在, 称它为 X 和 Y 的 $k+l$ 阶混合中心矩.

- 数学期望 $E(X)$ 是 X 的一阶原点矩.
- 方差 $D(X)$ 是 X 的二阶中心矩.
- 协方差 $\operatorname{cov}(X,Y)$ 是 X 和 Y 的 $1+1$ 阶混合中心矩.

矩

当 X 为离散型随机变量, 其分布列为 $P\{X=x_i\}=p_i,\ i=1,2,\cdots$, 则

$$E(X^k)=\sum_{i=1}^{\infty}x_i^k p_i,\quad E[X-E(X)]^k=\sum_{i=1}^{\infty}[x_i-E(X)]^k p_i.$$

当 X 为连续型随机变量, 其概率密度为 $f(x)$, 则

$$E(X^k)=\int_{-\infty}^{+\infty}x^k f(x)\mathrm{d}x,\quad E[X-E(X)]^k=\int_{-\infty}^{+\infty}[x-E(X)]^k f(x)\mathrm{d}x.$$

协方差矩阵

定义 (4.4.2)

对于二维随机变量 (X_1, X_2)，记

$$\sigma_{ij} = \operatorname{cov}(X_i, X_j) = E\{[X_i - E(X_i)][X_j - E(X_j)]\}, i, j = 1, 2$$

并排成矩阵

$$\begin{pmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{21} & \sigma_{22} \end{pmatrix},$$

称这个矩阵为 (X_1, X_2) 的**协方差矩阵**.

定义 (4.4.3)

设 n 维随机变量 $(X_1, X_2, \cdots, X_n)$ 的 $1+1$ 阶混合中心矩

$$\sigma_{ij} = \operatorname{cov}(X_i, X_j) = E\{[X_i - E(X_i)][X_j - E(X_j)]\}, i, j = 1, 2, \cdots, n$$

都存在，则称矩阵

$$\Sigma = \begin{pmatrix} \sigma_{11} & \sigma_{12} & \cdots & \sigma_{1n} \\ \sigma_{21} & \sigma_{22} & \cdots & \sigma_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{n1} & \sigma_{n2} & \cdots & \sigma_{nn} \end{pmatrix}$$

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

由于 $\sigma_{ij} = \sigma_{ji}$, $(i, j = 1, 2, \cdots, n)$, 因此 Σ 是一个对称矩阵.

利用协方差矩阵还可以引入 n 维正态分布的概率密度. 首先用协方差矩阵重写二维正态随机变量 (X_1, X_2) 的概率密度.

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

令 $X = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$, $\mu = \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}$, (X_1, X_2) 的协方差矩阵为

$$\Sigma = \begin{pmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{21} & \sigma_{22} \end{pmatrix} = \begin{pmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{pmatrix}.$$

它的行列式 $|\Sigma| = \sigma_1^2\sigma_2^2(1-\rho^2)$，逆矩阵

$$\Sigma^{-1} = \frac{1}{|\Sigma|}\begin{pmatrix} \sigma_2^2 & -\rho\sigma_1\sigma_2 \\ -\rho\sigma_1\sigma_2 & \sigma_1^2 \end{pmatrix}.$$

由于

$$\begin{aligned}
&(X-\mu)^T\Sigma^{-1}(X-\mu) \\
=&\frac{1}{|\Sigma|}(x_1-\mu_1, x_2-\mu_2)\begin{pmatrix} \sigma_2^2 & -\rho\sigma_1\sigma_2 \\ -\rho\sigma_1\sigma_2 & \sigma_1^2 \end{pmatrix}\begin{pmatrix} x_1-\mu_1 \\ x_2-\mu_2 \end{pmatrix} \\
=&\frac{1}{1-\rho^2}\left[\frac{(x_1-\mu_1)^2}{\sigma_1^2} - 2\rho\frac{(x_1-\mu_1)(x_2-\mu_2)}{\sigma_1\sigma_2} + \frac{(x_2-\mu_2)^2}{\sigma_2^2}\right],
\end{aligned}$$

因此 (X_1, X_2) 的概率密度可写成

$$f(x_1, x_2) = \frac{1}{2\pi\sqrt{|\Sigma|}}\exp\left\{-\frac{1}{2}(X-\mu)^T\Sigma^{-1}(X-\mu)\right\}.$$

上式容易推广到 n 维的情形.

设 $(X_1, X_2, \cdots, X_n)$ 是 n 维随机变量, 令

$$X = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}, \ \mu = \begin{pmatrix} \mu_1 \\ \mu_2 \\ \vdots \\ \mu_n \end{pmatrix} = \begin{pmatrix} E(X_1) \\ E(X_2) \\ \vdots \\ E(X_n) \end{pmatrix},$$

定义 n 维正态随机变量 $(X_1, X_2, \cdots, X_n)$ 的概率密度为

$$f(x_1, x_2, \cdots, x_n) = \frac{1}{(2\pi)^{\frac{n}{2}}\sqrt{|\Sigma|}} \exp\left\{-\frac{1}{2}(X-\mu)^T\Sigma^{-1}(X-\mu)\right\}.$$

其中 Σ 是 $(X_1, X_2, \cdots, X_n)$ 的协方差矩阵.

n 维正态随机变量具有以下几条重要性质:

(1) n 维随机变量 $(X_1, X_2, \cdots, X_n)$ 服从 n 维正态分布的充要条件是 $X_1, X_2, \cdots, X_n$ 的任意的线性组合

$$l_1 X_1 + l_2 X_2 + \cdots + l_n X_n$$

服从一维正态分布. (其中 $l_1, l_2, \cdots, l_n$ 不全为零).

(2) 若 $(X_1, X_2, \cdots, X_n)$ 服从 n 维正态分布, 设 $(Y_1, Y_2, \cdots, Y_k)$ 是 $(X_1, X_2, \cdots, X_n)$ 的线性函数, 则 $(Y_1, Y_2, \cdots, Y_k)$ 服从 k 维正态分布.

(3) 设 $(X_1, X_2, \cdots, X_n)$ 服从 n 维正态分布, 则 $X_1, X_2, \cdots, X_n$ 相互独立的充要条件是 $X_1, X_2, \cdots, X_n$ 两两不相关.

变异系数 (补充)

由于方差、标准差受量纲的影响，所以在实际工作中，常用变异系数这个数字特征.

定义 (变异系数)

设随机变量 X 的数学期望 $E(X) \neq 0$, 方差 $D(X)$ 存在，那么称

$$\delta_X \triangleq \frac{\sqrt{D(X)}}{|E(X)|}$$

为随机变量 X 的**变异系数**.

变异系数无量纲，反映随机变量在单位均值上的波动程度. 例如，当 $X \sim E(\lambda)$ 时，$\delta_X = 1$. 当 $X \sim N\left(\mu, \sigma^2\right)$ 时，$\delta_X = \dfrac{\sigma}{|\mu|}(\mu \neq 0)$.

条件期望

定义 (4.5.1)

设 (X, Y) 是二维离散型随机变量，其联合分布列是

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

若级数 $\sum_j y_j P(Y = y_j \mid X = x_i)$ 绝对收敛，则称

$$E(Y \mid X = x_i) := \sum_j y_j P(Y = y_j \mid X = x_i)$$

为在 $X = x_i$ 条件下 Y 的条件期望(Conditional Expectation)；称

$$E(X \mid Y = y_j) := \sum_i x_i P(X = x_i \mid Y = y_j)$$

为在 $Y = y_j$ 条件下 X 的条件期望.

条件方差

定义 (4.5.1)

称

$$D(Y \mid X = x_i) := \sum_j [y_j - E(Y \mid X = x_i)]^2 P(Y = y_j \mid X = x_i)$$

为在 $X = x_i$ 条件下 Y 的条件方差(Conditional Variance); 称

$$D(X \mid Y = y_j) := \sum_i [x_i - E(X \mid Y = y_j)]^2 P(X = x_i \mid Y = y_j)$$

为在 $Y = y_j$ 条件下 X 的条件方差.

条件期望

定义 (4.5.2)

设 (X, Y) 是二维连续型随机变量，其联合密度函数是 $f(x, y)$. 条件密度分别为 $f_{Y\mid X}(y\mid x)$, $f_{X\mid Y}(x\mid y)$. 若 $\int_{-\infty}^{+\infty} y f_{Y\mid X}(y\mid x)\mathrm{d}y$ 绝对收敛，则称

$$E(Y\mid X=x) := \int_{-\infty}^{+\infty} y f_{Y\mid X}(y\mid x)\mathrm{d}y$$

为在 $X=x$ 条件下 Y 的条件期望; 称

$$E(X\mid Y=y) := \int_{-\infty}^{+\infty} x f_{X\mid Y}(x\mid y)\mathrm{d}x$$

为在 $Y=y$ 条件下 X 的条件期望.

条件方差

定义 (4.5.2)

设 (X, Y) 是二维连续型随机变量，其联合密度函数是 $f(x, y)$. 条件密度分别为 $f_{Y \mid X}(y \mid x)$, $f_{X \mid Y}(x \mid y)$. 若 $\int_{-\infty}^{+\infty} y f_{Y \mid X}(y \mid x) \mathrm{d} y$ 绝对收敛，则称

$$D(Y \mid X=x):=\int_{-\infty}^{+\infty}[y-E(Y \mid X=x)]^2 f_{Y \mid X}(y \mid x) \mathrm{d} y$$

为在 $X=x$ 条件下 Y 的条件方差; 称

$$D(X \mid Y=y):=\int_{-\infty}^{+\infty}[x-E(X \mid Y=y)]^2 f_{X \mid Y}(x \mid y) \mathrm{d} x$$

为在 $Y=y$ 条件下 X 的条件方差.

1. 实数域中的 Cauchy-Schwarz 不等式

设 $a_i, b_i \in \mathbb{R}(i = 1, 2, \cdots, n)$, 则

$$\left(\sum_{i=1}^{n} a_i b_i\right)^2 \leqslant \sum_{i=1}^{n} a_i^2 \cdot \sum_{i=1}^{n} b_i^2,$$

等号成立当且仅当 $\dfrac{b_1}{a_1} = \dfrac{b_2}{a_2} = \cdots = \dfrac{b_n}{a_n}$.

2. n 维欧氏空间中的 Cauchy-Schwarz 不等式

在 n 维欧氏空间中, 对任意向量 α, β 有

$$\langle \alpha, \beta \rangle^2 \leqslant \langle \alpha, \alpha \rangle \langle \beta, \beta \rangle,$$

等号成立当且仅当 α, β 线性相关.

3. 数学分析中的 Cauchy-Schwarz 不等式

设 $f(x), g(x)$ 在 $[a, b]$ 上可积，则

$$\left[\int_a^b f(x)g(x)\,\mathrm{d}x\right]^2 \leqslant \int_a^b f^2(x)\,\mathrm{d}x \cdot \int_a^b g^2(x)\,\mathrm{d}x,$$

等号成立当且仅当存在常数 t 使得 $f(x) = tg(x)$ 或 $g(x) = tf(x)$.

4. 概率空间中的 Cauchy-Schwarz 不等式

设 X, Y 为任意随机变量, 若 $E\left(X^2\right), E\left(Y^2\right)$ 都存在, 则 $E(XY)$ 也存在, 且

$$[E(XY)]^2 \leqslant E\left(X^2\right) \cdot E\left(Y^2\right),$$

等号成立当且仅当存在常数 t_0, 使得 $P\{Y = t_0X\} = 1$.

本章小结

1. **数学期望与方差**

$$E(X)=\sum_{k=1}^{+\infty}x_kp_k;$$

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

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

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

1. 数学期望与方差

性质:

(1) $E(c)=c$, 其中 c 是常数; $D(c)=0$;

(2) $E(cX)=cE(X)$; $D(cX)=c^2D(X)$;

(3) $E(X+Y)=E(X)+E(Y)$;

(4) X, Y 相互独立 $\Rightarrow$ X, Y 不相关 $\Leftrightarrow E(XY)=E(X)E(Y)$
$\Leftrightarrow D(X\pm Y)=D(X)+D(Y) \Leftrightarrow \rho=0 \Leftrightarrow \operatorname{cov}(X,Y)=0$.

实际意义: 数学期望表示的是随机变量取值的 "加权平均"; 方差是描述随机变量的取值偏离其数学期望的程度.

方差计算的常用公式: $D(X)=E(X^2)-E^2(X)$.

方差与协方差的关系式: $D(X)=\operatorname{cov}(X,X)$

$$D(X\pm Y)=D(X)+D(Y)\pm 2\operatorname{cov}(X,Y).$$

七种常见分布的数学期望与方差.

分布	分布律或密度函数	期望	方差
$b(1,p)$	$P(X=k)=p^k(1-p)^{1-k},\ k=0,1$	p	$p(1-p)$
$b(n,p)$	$P(X=k)=C_n^k p^k(1-p)^{n-k},\ k=0,1,\cdots,n$	np	$np(1-p)$
$P(\lambda)$	$P(X=k)=\dfrac{\lambda^k}{k!}\mathrm{e}^{-\lambda},\ k=0,1,2,\cdots,\ \lambda>0$	λ	λ
$Ge(p)$	$P(X=k)=p(1-p)^{k-1},\ k=1,2,\cdots$	$\dfrac{1}{p}$	$\dfrac{1-p}{p^2}$
$U(a,b)$	$f(x)=\dfrac{1}{b-a},\ a<x<b;\ f(x)=0,$ 其他	$\dfrac{a+b}{2}$	$\dfrac{(b-a)^2}{12}$
$Exp(\lambda)$	$f(x)=\lambda\mathrm{e}^{-\lambda x},\ x>0;\ f(x)=0,\ x\leqslant 0$	$\dfrac{1}{\lambda}$	$\dfrac{1}{\lambda^2}$
$N(\mu,\sigma^2)$	$f(x)=\dfrac{1}{\sqrt{2\pi}\sigma}\mathrm{e}^{-\frac{(x-\mu)^2}{2\sigma^2}},\ -\infty<x<+\infty$	μ	σ^2

2. 随机变量函数的数学期望

(1) 若 X 的分布列为 $P(X=x_k)=p_k,\ k=1,2,\cdots$，则有 $Y=g(X)$ 的期望为

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

(2) X 的概率密度为 $f(x)$，则有 $Y=g(X)$ 的期望为

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

(3) 设 (X,Y) 的联合分布列为 $P\{X=x_i,Y=y_j\}=p_{ij},\ (i,j=1,2,\cdots)$ 时，则有 $Z=g(X,Y)$ 的数学期望为

$$E(Z)=E[g(X,Y)]=\sum_i\sum_j g(x_i,x_j)p_{ij}.$$

(4) 当 (X,Y) 的概率密度为 $f(x,y)$ 时，则有 $Z=g(X,Y)$ 的数学期望为

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

3. X, Y 的协方差与相关系数

$$\operatorname{cov}(X, Y) := E\{[X-E(X)][Y-E(Y)]\};\ \rho_{XY} := \frac{\operatorname{cov}(X, Y)}{\sqrt{D(X)}\sqrt{D(Y)}}.$$

性质:

(1) $\operatorname{cov}(X, X) = D(X)$;

(2) $\operatorname{cov}(X, Y) = \operatorname{cov}(Y, X)$;

(3) 若 X 与 Y 相互独立, 则 $\operatorname{cov}(X, Y) = 0$;

(4) 若 a, b 为常数, 则 $\operatorname{cov}(aX, bY) = ab \cdot \operatorname{cov}(X, Y)$;

(5) $\operatorname{cov}(X_1 + X_2, Y) = \operatorname{cov}(X_1, Y) + \operatorname{cov}(X_2, Y)$;

(6) 若 a 为常数, 则 $\operatorname{cov}(a, X) = 0$.

3. X, Y 的协方差与相关系数

(7) $|\rho_{XY}| \leqslant 1$; $|\rho_{XY}| = 0$ 时称 X, Y 不相关;

$$|\rho_{XY}| = 1 \Leftrightarrow P\{Y = aX + b\} = 1, \ (a \neq 0).$$

实际意义: X, Y 的相关系数描述了 X, Y 之间线性关系的紧密程度.
协方差的计算公式: $\operatorname{cov}(X, Y) = E(XY) - E(X)E(Y)$;

(8) 柯西 – 许瓦兹不等式: $[E(XY)]^2 \leqslant E(X^2)E(Y^2)$.

4. 随机变量的矩及协方差矩阵

5. 条件数学期望 *