

第四章 随机变量的数字特征

本章导读

前几章用分布完整描述随机变量, 但很多问题不需要保存全部分布信息. 例如, 比较两种方案时可能只关心长期平均收益、波动大小以及两个指标是否同向变化. 数字特征就是把分布压缩成少数有解释力的数:

中心位置: 期望, 波动程度: 方差, 共同变化: 协方差与相关系数.

本章始终沿着同一条计算链工作:

已知分布 $\longrightarrow E(X), E(X^2), E(XY) \longrightarrow D(X), \text{Cov}(X, Y), \rho_{XY}$.

其中期望是根. 方差、协方差和相关系数最终都要回到期望计算.

4.1 数学期望

4.1.1 定义、存在性与直观含义

定义 4.1 (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

$$P(X = x_k) = p_k, \quad k = 1, 2, \dots$$

若级数 $\sum_k x_k p_k$ 绝对收敛, 则定义

$$E(X) = \sum_k x_k p_k.$$

否则称 X 的数学期望不存在.

定义 4.2 (连续型随机变量的数学期望)

设连续型随机变量 X 的密度为 $f(x)$. 若

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

则定义

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

数学期望是按概率加权的平均位置. 它不一定是随机变量可能取到的值, 也不保证每个随机变量都存在期望. 绝对收敛保证求和或积分次序不影响唯一结果.

注: [对称不能代替存在性检查] 标准柯西密度

$$f(x) = \frac{1}{\pi(1+x^2)}, \quad x \in \mathbb{R},$$

关于原点对称, 但

$$\int_{-\infty}^{+\infty} |x|f(x) dx = +\infty.$$

因此其期望不存在. 把左右两端对称抵消得到的主值为 0, 不是概率论定义下的期望.

例题 4.3 (离散型期望与函数期望)

设

x	-2	0	2
$P(X=x)$	$\frac{4}{10}$	$\frac{3}{10}$	$\frac{3}{10}$

求 $E(X)$ 与 $E(3X^2 + 5)$.

直接按“取值乘概率”求和:

$$E(X) = (-2)\frac{4}{10} + 0 \cdot \frac{3}{10} + 2\frac{3}{10} = -\frac{1}{5}.$$

先求

$$E(X^2) = 4\frac{4}{10} + 0 + 4\frac{3}{10} = \frac{14}{5},$$

再利用期望的线性性:

$$E(3X^2 + 5) = 3E(X^2) + 5 = \frac{67}{5}.$$

方法选择: 无需先求 $3X^2 + 5$ 的分布律; 直接计算函数期望更短.

4.1.2 随机变量函数的期望

定理 4.4 (一维随机变量函数的期望)

若 $Y = g(X)$ 且相应绝对收敛条件成立, 则:

- 离散型:

$$E[g(X)] = \sum_k g(x_k) p_k;$$

- 连续型:

$$E[g(X)] = \int_{-\infty}^{+\infty} g(x) f_X(x) dx.$$

这条公式常称为“无须先求函数分布的期望公式”. 口诀是: 求谁的期望, 就把谁乘在原分布前面求和或积分.

例题 4.5 (分段函数的期望)

设 $X \sim E(1)$, 令 $Z = \max(X, 1)$, 求 $E(Z)$.

Z 在 $0 < X < 1$ 时等于 1, 在 $X \geq 1$ 时等于 X , 故

$$\begin{aligned} E(Z) &= \int_0^1 1 \cdot e^{-x} dx + \int_1^{+\infty} x e^{-x} dx \\ &= (1 - e^{-1}) + 2e^{-1} = 1 + e^{-1}. \end{aligned}$$

关键转折: 先判断 Z 在不同区间的表达式, 再把正确的函数乘密度; 不能把 $\max(X, 1)$ 误当成连续型随机变量处处都有普通变换密度后才计算.

4.1.3 二维随机变量函数的期望

定理 4.6 (二维函数期望公式)

设 $Z = g(X, Y)$, 相应绝对收敛条件成立.

- 若 (X, Y) 离散, 联合分布律为 p_{ij} , 则

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

- 若 (X, Y) 连续, 联合密度为 $f(x, y)$, 则

$$E[g(X, Y)] = \iint_{\mathbb{R}^2} g(x, y)f(x, y) dx dy.$$

例题 4.7 (二维离散型函数期望)

设联合分布律为

	$X = 1$	$X = 2$	$X = 3$
$Y = -1$	0.2	0.1	0
$Y = 0$	0.1	0	0.3
$Y = 1$	0.1	0.1	0.1

若 $Z = (Y - X)^2$, 则

$$E(Z) = \sum_{x,y} (y - x)^2 P(X = x, Y = y).$$

逐格代入即可, 无需先把相同的 Z 值合并成新分布律.

例题 4.8 (相约问题的平均等待时间)

甲乙约在 8-12 点见面, 二人到达时刻 X, Y 独立且均服从 $U(0, 4)$. 先到者等待时间为 $|X - Y|$, 求平均等待时间.

联合密度在正方形 $(0, 4)^2$ 上为 $1/16$. 由对称性只积上三角形再乘 2:

$$\begin{aligned} E|X - Y| &= 2 \int_0^4 \int_0^y (y - x) \frac{1}{16} dx dy \\ &= \frac{4}{3}. \end{aligned}$$

所以平均等待时间为 $4/3$ 小时, 即 80 分钟.

结果检查: 等待时间在 $[0, 4]$ 内, 平均值 $4/3$ 合理; 对称性只减少计算量, 不改变联合密度.

4.1.4 期望的运算性质

定理 4.9 (期望的常用性质)

在相关期望存在时:

1. $E(C) = C$;
2. $E(aX + bY + c) = aE(X) + bE(Y) + c$;
3. 更一般地, $E(\sum_i a_i X_i + c) = \sum_i a_i E(X_i) + c$;
4. 若 X, Y 独立, 则 $E(XY) = E(X)E(Y)$.

注: [线性无条件, 乘积拆分有条件] 期望的加减可以无条件拆开, 不要求独立; 乘积期望拆成期望乘积则需要独立, 或至少需要后文所说的“不相关”. 这是本章最常见的条件混淆.

注: [证明等级] 期望线性性质在离散、连续情形下逐条验证属于 C 级, 本阶段直接熟练使用; 但遇到具体题目要求证明某个期望等于指定值时, 必须从分布、函数期望公式或线性性质完整写出计算链.

小结

期望是按概率加权的平均位置. 使用前先检查绝对收敛; 求函数期望通常直接把函数乘原分布, 无需先求新分布. 二维情形把求和推广为双重求和、把积分推广为二重积分. 期望线性无须独立, 只有乘积期望的拆分需要额外条件.

4.2 方差与标准差

4.2.1 定义与计算公式

定义 4.10 (方差与标准差)

若 $E[(X - E(X))^2]$ 存在, 则定义

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

为 X 的方差, 定义

$$\sigma_X = \sqrt{D(X)}$$

为标准差.

方差衡量随机变量围绕期望的波动程度. 方差单位是原单位的平方; 标准差与 X 同量纲, 更便于直观解释.

定理 4.11 (方差的易计算公式)

若 $E(X^2)$ 存在, 则

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

注: [公式为什么成立 (B 级)] 把 $E(X)$ 视为常数, 展开

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

即可得到结论. 真正起作用的是期望的线性性.

证明: **证明目标:** 根据定义验证给定随机变量的方差.

出发点: $D(X) = E(X^2) - [E(X)]^2$.

关键转换: 先从分布分别计算一阶矩和二阶矩.

正式推导: 若离散型 X 的分布律为 p_k , 则

$$E(X) = \sum_k x_k p_k, \quad E(X^2) = \sum_k x_k^2 p_k,$$

所以

$$D(X) = \sum_k x_k^2 p_k - \left(\sum_k x_k p_k \right)^2.$$

连续型时把求和替换为对密度的积分即可.

证明小结: 验证方差的固定路线是“一阶矩、二阶矩、相减”, 最后检查结果非负. ■

例题 4.12 (由函数分布求方差)

设 $X \sim U(-3, 2)$, 令

$$Y = \begin{cases} 1, & X \geq 0, \\ -1, & X < 0. \end{cases}$$

求 $D(Y)$.

区间长度为 5, 故

$$P(Y = 1) = P(X \geq 0) = \frac{2}{5}, \quad P(Y = -1) = \frac{3}{5}.$$

于是

$$E(Y) = \frac{2}{5} - \frac{3}{5} = -\frac{1}{5}, \quad E(Y^2) = 1.$$

所以

$$D(Y) = 1 - \left(-\frac{1}{5}\right)^2 = \frac{24}{25}.$$

检查: Y 只取 ± 1 , 方差不超过 1, 结果合理.

4.2.2 方差的运算性质

定理 4.13 (方差的常用性质)

1. $D(C) = 0$;
2. $D(X + C) = D(X)$;
3. $D(aX) = a^2 D(X)$;
4. 若 X, Y 独立, 则

$$D(X \pm Y) = D(X) + D(Y).$$

更一般的非独立情形将在协方差一节补全.

注: [和差方差的三个易错点] 常数平移在方差中消失; 系数提出要平方; 独立时无论是 $X + Y$ 还是 $X - Y$, 两个方差之间都是加号. 若未给独立或不相关, 不能擅自拆开.

定理 4.14 (方差为零的含义)

$D(X) = 0$ 当且仅当存在常数 c , 使 $P(X = c) = 1$. 该严格证明属于 C 级, 本阶段准确记住结论并会在题目中调用.

小结

方差是到期望的平方偏差的平均, 标准差与原变量同量纲. 计算时优先用“二阶矩减一阶矩平方”. 平移不改变方差, 缩放系数要平方; 和差方差只有在独立或不相关时才能直接相加.

4.3 常见分布的期望与方差

4.3.1 常用结论表

分布	参数与记号	$E(X)$	$D(X)$
0-1 分布	$P(X = 1) = p$	p	$p(1 - p)$
二项分布	$B(n, p)$	np	$np(1 - p)$
泊松分布	$P(\lambda)$	λ	λ
几何分布	$Ge(p)$, 取值从 1 起	$1/p$	$(1 - p)/p^2$
均匀分布	$U(a, b)$	$(a + b)/2$	$(b - a)^2/12$
指数分布	$E(\lambda)$	$1/\lambda$	$1/\lambda^2$
正态分布	$N(\mu, \sigma^2)$	μ	σ^2
卡方分布	$\chi^2(n)$	n	$2n$

必须连同参数口径一起记. 例如泊松分布的期望和方差等于 λ ; 指数分布的期望与方差却分别是 $1/\lambda$ 与 $1/\lambda^2$.

例题 4.15 (由期望和方差反求二项参数)

已知 $X \sim B(n, p)$, 且

$$E(X) = 2.4, \quad D(X) = 1.44.$$

求 n, p .

由

$$np = 2.4, \quad np(1 - p) = 1.44,$$

两式相除得

$$1 - p = \frac{1.44}{2.4} = 0.6, \quad p = 0.4.$$

进而 $n = 2.4/0.4 = 6$.

检查: n 是正整数、 $0 < p < 1$, 且代回方差得到 $6 \times 0.4 \times 0.6 = 1.44$.

例题 4.16 (由方差反求二阶矩)

若 $X \sim P(\lambda)$ 且 $E(X^2) = 6$, 则

$$E(X^2) = D(X) + [E(X)]^2 = \lambda + \lambda^2 = 6.$$

由 $\lambda > 0$ 得 $\lambda = 2$. 当期望、方差有现成结论时, 不必重新对分布律求和.

4.3.2 独立正态的线性组合

若 $X_i \sim N(\mu_i, \sigma_i^2)$ 相互独立, 则

$$Z = \sum_{i=1}^n a_i X_i + b \sim N\left(\sum_{i=1}^n a_i \mu_i + b, \sum_{i=1}^n a_i^2 \sigma_i^2\right).$$

例题 4.17 (现场确定正态参数)

设 $X \sim N(-3, 1)$ 、 $Y \sim N(2, 1)$, 且相互独立. 令 $Z = X - 2Y + 7$, 则

$$E(Z) = -3 - 2 \times 2 + 7 = 0,$$

$$D(Z) = 1 + (-2)^2 \times 1 = 5.$$

因此 $Z \sim N(0, 5)$.

4.3.3 最大值与最小值的期望技巧

对任意实数 x, y , 有

$$\max(x, y) + \min(x, y) = x + y, \quad \max(x, y) \min(x, y) = xy,$$

以及

$$\max(x, y) = \frac{x + y + |x - y|}{2}, \quad \min(x, y) = \frac{x + y - |x - y|}{2}.$$

这些恒等式常比先求极值分布再积分更短.

例题 4.18 (两个独立指数变量的极值期望)

设 X, Y 独立且均服从 $E(1)$. 令

$$U = \max(X, Y), \quad V = \min(X, Y).$$

求 $E(U), E(V), E(UV)$.

独立指数变量的最小值仍为指数分布, 参数相加, 因此

$$V \sim E(2), \quad E(V) = \frac{1}{2}.$$

又由 $U + V = X + Y$,

$$E(U) = E(X) + E(Y) - E(V) = 1 + 1 - \frac{1}{2} = \frac{3}{2}.$$

由 $UV = XY$ 且 X, Y 独立,

$$E(UV) = E(XY) = E(X)E(Y) = 1.$$

易错点: U, V 一般不独立, 不能把 $E(UV)$ 直接拆成 $E(U)E(V)$; 这里是先用 $UV = XY$ 换回独立的 X, Y .

小结

常见分布的期望、方差应连同参数口径熟记. 反求参数时联立期望和方差; 已知期望、方差时可反求二阶矩. 独立正态线性组合直接由期望和方差确定参数; 极值期望优先检查和、积及绝对值恒等式, 避免不必要地先求完整分布.

4.4 切比雪夫不等式

定理 4.19 (切比雪夫不等式)

若 $E(X) = \mu$ 、 $D(X) = \sigma^2$ 存在, 则对任意 $\varepsilon > 0$,

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

等价地,

$$P(|X - \mu| < \varepsilon) \geq 1 - \frac{\sigma^2}{\varepsilon^2}.$$

注: [为什么成立 (B 级证明骨架)] 在事件 $|X - \mu| \geq \varepsilon$ 上, $(X - \mu)^2 \geq \varepsilon^2$. 因此该事件的概率质量乘以 ε^2 不超过全部平方偏差的期望 $D(X)$. 离散型用求和、连续型用积分即可写成严格放缩. 本阶段应能复述这条思路并准确调用结论.

例题 4.20 (直接估计偏离期望的概率)

若 $D(X) = 2$, 则

$$P(|X - E(X)| \geq 2) \leq \frac{2}{2^2} = \frac{1}{2}.$$

切比雪夫不等式只利用期望和方差, 不需要知道 X 的具体分布, 因此给出的是普适上界, 不一定精确.

例题 4.21 (先构造中心化变量再使用不等式)

已知

$$E(X) = E(Y) = 2, \quad D(X) = 1, \quad D(Y) = 4, \quad \rho_{XY} = 0.5.$$

估计 $P(|X - Y| \geq 6)$ 的上界.

令 $Z = X - Y$. 先检查中心:

$$E(Z) = E(X) - E(Y) = 0,$$

所以 $|X - Y| = |Z - E(Z)|$. 由

$$\text{Cov}(X, Y) = \rho_{XY} \sqrt{D(X)D(Y)} = 0.5 \times 1 \times 2 = 1,$$

得

$$D(Z) = D(X) + D(Y) - 2 \text{Cov}(X, Y) = 1 + 4 - 2 = 3.$$

因此

$$P(|X - Y| \geq 6) \leq \frac{3}{36} = \frac{1}{12}.$$

方法选择: 切比雪夫形式必须围绕期望; 若题面没有显式出现期望, 应先构造新变量并计算其期望、方差.

小结

切比雪夫不等式用方差控制偏离期望的概率, 不依赖具体分布. 使用时先确认期望、方差存在, 并把目标事件写成 $|Z - E(Z)| \geq \varepsilon$; 右端一定是“方差除以距离平方”.

4.5 协方差

4.5.1 定义、计算与性质

定义 4.22 (协方差)

若相应期望存在, 定义

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

为 X, Y 的协方差.

协方差为正, 表示两个变量相对各自期望的偏差倾向同号; 协方差为负, 表示偏差倾向异号. 但它受单位和尺度影响, 不能直接跨量纲比较强弱.

定理 4.23 (协方差的易计算公式)

$$\text{Cov}(X, Y) = E(XY) - E(X)E(Y).$$

注: [展开思路 (B 级)] 把定义中的乘积展开并使用期望线性性, 四项中有三项合并, 得到“乘积期望减期望乘积”. 因此从联合分布计算协方差时, 固定求 $E(XY), E(X), E(Y)$ 三个量.

定理 4.24 (协方差的运算性质)

1. $\text{Cov}(X, Y) = \text{Cov}(Y, X)$;
2. $\text{Cov}(X, C) = 0$;
3. $\text{Cov}(aX, bY) = ab \text{Cov}(X, Y)$;
4. $\text{Cov}(X, X) = D(X)$;
5. 协方差对每个变量都具有线性性;
6. 若 X, Y 独立, 则 $\text{Cov}(X, Y) = 0$;
- 7.

$$D(X \pm Y) = D(X) + D(Y) \pm 2 \text{Cov}(X, Y).$$

证明: 证明目标: 推导一般和差方差公式, 并指出独立情形为何可简化.

出发点: 方差与协方差的定义.

关键转换: 把中心化后的和差平方展开.

正式推导:

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

若 X, Y 独立, 则 $E(XY) = E(X)E(Y)$, 故协方差为零, 和差方差均化为方差之和.

证明小结: 方差不能无条件拆开, 缺少的“小尾巴”正是二倍协方差. 独立只是一种让协方差归零的充分条件. ■

例题 4.25 (由联合分布计算协方差)

设

	$Y = -1$	$Y = 0$	$Y = 1$
$X = 0$	0.07	0.18	0.15
$X = 1$	0.08	0.32	0.20

求 $\text{Cov}(X^2, Y^2)$.

由于 $X \in \{0, 1\}$, 有 $X^2 = X$; 由于 $Y \in \{-1, 0, 1\}$, Y^2 是 $Y \neq 0$ 的指示变量.

$$E(X^2) = P(X = 1) = 0.60,$$

$$E(Y^2) = P(|Y| = 1) = 0.07 + 0.15 + 0.08 + 0.20 = 0.50,$$

$$E(X^2 Y^2) = P(X = 1, |Y| = 1) = 0.08 + 0.20 = 0.28.$$

所以

$$\text{Cov}(X^2, Y^2) = 0.28 - 0.60 \times 0.50 = -0.02.$$

关键转折: 先利用取值简化平方, 再把乘积期望识别成一个事件概率.

小结

协方差衡量两个变量中心化偏差的共同变化, 计算时固定求三个期望. 协方差对加法线性、常数消失, 且补全了非独立情形的和差方差公式. 独立推出协方差为零, 但逆命题一般不成立.

4.6 相关系数

协方差能够判断共同变化的方向, 但它会随单位和尺度改变. 例如把“米”换成“厘米”, 协方差会被放大. 为了得到一个无量纲、便于比较的指标, 需要把协方差标准化.

定义 4.26 (相关系数)

设随机变量 X, Y 的方差存在且

$$D(X) > 0, \quad D(Y) > 0.$$

称

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

为 X, Y 的相关系数.

注: 定义前必须检查两个方差都严格大于零. 若某个变量几乎处处为常数, 它没有随机波动, 分母为零, 相关系数便没有定义. 相关系数描述的是**线性相关程度**, 不是一般意义下所有依赖关系的强弱.

定理 4.27 (相关系数的基本性质)

若 $D(X) > 0, D(Y) > 0$, 则:

1. $\rho_{XY} = \rho_{YX}$;
2. $-1 \leq \rho_{XY} \leq 1$;
3. 对 $a, c \neq 0$,

$$\rho_{aX+b, cY+d} = \text{sgn}(ac)\rho_{XY};$$

4. $|\rho_{XY}| = 1$ 当且仅当存在常数 $a \neq 0, b$, 使

$$Y = aX + b \quad \text{几乎处处成立.}$$

此时 $\rho_{XY} = 1$ 对应 $a > 0$, $\rho_{XY} = -1$ 对应 $a < 0$.

注: [为什么相关系数的绝对值不超过 1] 令

$$U = \frac{X - E(X)}{\sqrt{D(X)}}, \quad V = \frac{Y - E(Y)}{\sqrt{D(Y)}}.$$

则 $E(U) = E(V) = 0, E(U^2) = E(V^2) = 1$, 并且 $\rho_{XY} = E(UV)$. 由柯西-施瓦茨不等式,

$$|E(UV)|^2 \leq E(U^2)E(V^2) = 1.$$

所以 $|\rho_{XY}| \leq 1$. 等号成立恰好对应 U, V 几乎处处成比例, 从而 X, Y 存在几乎处处成立的线性关系.

这里使用了较高层次的不等式及其等号条件. 现阶段应掌握标准化思路和证明骨架; 严格的不等式理论细节本阶段不要求复现.

例题 4.28 (完全负线性相关)

在长度为 1 的木棍上随机选一点截断, 令左段长度为 X , 右段长度为 Y . 求 ρ_{XY} .

由几何关系, $Y = 1 - X$. 只要截点不是固定的, 就有 $D(X) > 0$, 且

$$\text{Cov}(X, Y) = \text{Cov}(X, 1 - X) = -D(X),$$

$$D(Y) = D(1 - X) = D(X).$$

因此

$$\rho_{XY} = \frac{-D(X)}{\sqrt{D(X)}\sqrt{D(X)}} = -1.$$

方法选择: 看到确定的线性关系, 直接使用 $|\rho| = 1$ 的性质, 比先求联合分布更简洁.

结果检查: 左段变长时右段必等量变短, 完全反向线性变化, 故结果应为 -1 .

例题 4.29 (尺度变化不改变相关强弱)

若 $D(X), D(Y) > 0$ 且 $\rho_{XY} = 0.8$, 令 $U = 100X + 3$, $V = -2Y + 5$, 则

$$\rho_{UV} = \text{sgn}(100 \times (-2))\rho_{XY} = -0.8.$$

平移不影响相关系数; 正比例缩放不改变符号, 负比例缩放会反转符号.

协方差与相关系数的辨析

比较项	协方差	相关系数
定义	$\text{Cov}(X, Y)$	协方差除以两个标准差
量纲	有量纲, 随单位改变	无量纲
取值范围	没有统一范围	$[-1, 1]$
主要作用	判断共同变化方向, 补全和差方差	衡量线性相关的方向与标准化程度
使用条件	相关二阶矩存在	还要求两个方差严格大于零
常见误判	数值大就认为关系强	$\rho = 0$ 就认为完全没有关系

小结

相关系数是标准化后的协方差, 取值在 $[-1, 1]$, 只衡量线性相关. 使用前先检查两个方差是否严格大于零; $|\rho| = 1$ 表示几乎处处存在确定的线性关系, 而 $\rho = 0$ 只表示没有线性相关, 不能直接推出独立.

4.7 不相关与独立

这一节解决本章最重要的逻辑辨析: 两个变量的协方差为零, 究竟能说明什么? 它与独立到底相差多远?

定义 4.30 (不相关)

若 X, Y 的二阶矩存在, 且

$$\text{Cov}(X, Y) = 0,$$

则称 X, Y 不相关. 当 $D(X) > 0, D(Y) > 0$ 时, 这也等价于 $\rho_{XY} = 0$.

定理 4.31 (不相关的等价刻画)

设 X, Y 的二阶矩存在, 则下列命题等价:

1. $\text{Cov}(X, Y) = 0$;
2. $E(XY) = E(X)E(Y)$;
3. $D(X + Y) = D(X) + D(Y)$;
4. $D(X - Y) = D(X) + D(Y)$.

若再有 $D(X) > 0, D(Y) > 0$, 它们还等价于 $\rho_{XY} = 0$.

证明: 证明目标: 证明“不相关”的几种常见说法可以互相替换.

出发点: 协方差计算公式、和差方差公式以及相关系数定义.

关键转换: 每个等式与标准形式比较, 分离出其中的协方差项.

正式推导: 由

$$\text{Cov}(X, Y) = E(XY) - E(X)E(Y),$$

立即得到命题 1 与命题 2 等价. 又因为

$$D(X + Y) - D(X) - D(Y) = 2 \text{Cov}(X, Y),$$

$$D(X - Y) - D(X) - D(Y) = -2 \text{Cov}(X, Y),$$

所以命题 1 分别与命题 3、命题 4 等价. 若两个方差严格大于零, 则相关系数的分母为正, 故

$$\rho_{XY} = 0 \iff \text{Cov}(X, Y) = 0.$$

证明小结: 这些等价关系都来自同一个量 $\text{Cov}(X, Y)$. 使用时可选择题目中最容易计算的一种形式, 但必须先确认相关矩存在; 涉及相关系数时还要检查方差不为零. ■

定理 4.32 (独立与不相关的关系)

若 X, Y 相互独立且二阶矩存在, 则 X, Y 不相关. 反过来一般不成立.

证明: 证明目标: 由独立性推出协方差为零.

出发点: 独立随机变量的函数乘积期望可拆分.

关键转换: 把 $E(XY)$ 拆为 $E(X)E(Y)$.

正式推导: 由于 X, Y 独立,

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

因此

$$\text{Cov}(X, Y) = E(XY) - E(X)E(Y) = 0.$$

证明小结: 独立控制的是联合分布的整体结构, 所以能推出不相关; 不相关只控制一个混合二阶矩, 信息远少于独立, 因而通常不能反推. ■

例题 4.33 (不相关但不独立的反例)

设 $X \sim U(-1, 1)$, 令 $Y = X^2$. 判断 X, Y 是否相关、是否独立.

由于 X 关于原点对称,

$$E(X) = 0, \quad E(XY) = E(X^3) = 0.$$

于是

$$\text{Cov}(X, Y) = E(XY) - E(X)E(Y) = 0,$$

所以 X, Y 不相关.

但 $Y = X^2$ 是由 X 完全决定的. 例如一旦知道 X , 便确定知道 Y , 二者显然不是独立的.

关键转折: 奇函数在对称区间上的期望为零, 只能说明线性共同变化被抵消, 不能消除确定的非线性关系.

易错提醒: “相关系数为零”应读作“无线性相关”, 不能读作“没有任何关系”.

定理 4.34 (二维正态分布的特殊结论)

若 (X, Y) 服从二维正态分布, 则

$$X, Y \text{ 相互独立} \iff \text{Cov}(X, Y) = 0 \iff \rho_{XY} = 0,$$

其中相关系数形式要求两个方差严格大于零.

注: 这是二维正态分布的特殊结构, 不能推广到任意联合分布. 做题时只有明确给出“联合正态”或“二维正态”, 才能由不相关反推独立; 仅知道两个边缘分布各自为正态仍不够.

独立与不相关的辨析

比较项	独立	不相关
本质	联合分布可由边缘分布分解	协方差为零
控制范围	所有适当函数的联合关系	只控制一个二阶线性指标
常用判定	联合分布函数或密度分解、事件乘法	$E(XY) = E(X)E(Y)$
一般关系	有矩条件下推出不相关	一般不能推出独立
特殊情形	二维正态中与不相关等价	需先确认联合正态

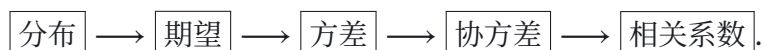
小结

不相关等价于协方差为零, 也等价于乘积期望可拆分以及和差方差可直接相加. 独立在矩存在时一定推出不相关, 但不相关通常不能推出独立; 只有在二维正态等特殊结构中才能反推. 判断时先问“题目给的是一般联合分布, 还是联合正态”.

4.8 本章总结

知识结构

本章可以按一条主线组织:



期望描述中心; 方差描述单个变量围绕中心的波动; 协方差描述两个变量共同变化; 相关系数把协方差标准化. 独立性属于联合分布层面的性质, 不相关只是二阶矩层面的性质.

核心公式与适用条件

对象	核心公式	使用前检查
离散型期望	$E(X) = \sum_k x_k p_k$	绝对收敛
连续型期望	$E(X) = \int_{-\infty}^{+\infty} x f(x) dx$	绝对可积
函数期望	$E[g(X)] = \sum g(x_k) p_k$ 或 $\int g(x) f(x) dx$	原变量分布已知, 期望存在
方差	$D(X) = E(X^2) - [E(X)]^2$	二阶矩存在
和差方差	$D(X \pm Y) = D(X) + D(Y) \pm 2 \text{Cov}(X, Y)$	二阶矩存在; 不能漏协方差
协方差	$\text{Cov}(X, Y) = E(XY) - E(X)E(Y)$	相关矩存在
相关系数	$\rho = \text{Cov}(X, Y) / \sqrt{D(X)D(Y)}$	两个方差严格大于零
切比雪夫	$P(X - E(X) \geq \varepsilon) \leq D(X) / \varepsilon^2$	$\varepsilon > 0$, 期望和方差存在

常见题型与方法选择

1. **求函数的期望:** 已知原变量分布时, 优先直接计算 $E[g(X)]$, 不要先求新变量分布.
2. **求方差:** 通常先求 $E(X)$ 、 $E(X^2)$, 再用快捷公式; 系数提出时必须平方.
3. **参数反求:** 把题目给出的期望、方差代入对应分布公式, 再检查参数范围.
4. **求协方差:** 固定计算 $E(XY)$, $E(X)$, $E(Y)$; 若存在简单线性关系, 优先用运算性质.
5. **求线性组合方差:** 先写完整协方差公式, 只有确认独立或不相关后才能删去交叉项.
6. **概率粗界:** 只知道期望和方差时, 把事件整理为偏离期望的形式, 再使用切比雪夫不等式.
7. **判断独立:** 一般回到联合分布; 只有联合正态等明确特殊条件下, 才可用零相关反推独立.

重点辨析

容易混淆的说法	正确结论	常见误判
期望线性性	$E(aX + bY) = aE(X) + bE(Y)$, 不要求独立	看到两个变量就额外要求独立
乘积期望	独立可推出 $E(XY) = E(X)E(Y)$	无条件拆分乘积
方差线性性	缩放系数平方, 和差还含协方差	把方差当作期望直接拆
协方差与相关系数	前者有量纲, 后者标准化且有界	用协方差绝对值直接比较强弱
零相关与独立	独立推出零相关, 逆命题一般不成立	看见 $\rho = 0$ 就判独立
二维正态	联合正态下零相关与独立等价	两个边缘正态就套用该结论

易错点清单

- 不检查期望的绝对收敛或绝对可积, 误把发散的正负部分抵消成有限值;
- 求 $E[g(X)]$ 时误把密度写成 $g(f(x))$, 忘记“函数乘原密度”;
- 把 $D(aX + b)$ 写成 $aD(X) + b$;
- 未知独立性时擅自写 $D(X + Y) = D(X) + D(Y)$;
- 切比雪夫不等式没有围绕 $E(X)$ 整理事件, 或忘记分母平方;
- 相关系数分母为零时仍继续计算;
- 把“无相关”理解成“无任何依赖关系”;
- 忘记二维正态结论要求的是联合正态, 而非仅边缘正态.

证明掌握清单

- **A 级:** 能由定义计算并验证期望、方差、协方差和相关系数; 能完整推出和差方差公式, 并逐项说明条件.
- **B 级:** 能复述切比雪夫不等式的证明骨架; 理解方差快捷公式、线性变换、协方差性质、 $|\rho| \leq 1$ 及等号条件的核心转换.
- **C 级:** 知道期望线性性的离散与连续严格证明、 $D(X) = 0$ 的严格等价证明, 以及相关系数有界性的高层不等式细节, 本阶段不要求完整复现.

自测题

练习 4.35

设离散型随机变量 X 的分布律为

$$P(X = -1) = \frac{1}{4}, \quad P(X = 0) = \frac{1}{2}, \quad P(X = 2) = \frac{1}{4}.$$

求 $E(X), E(X^2), D(X)$, 并用 $D(3X - 2) = 9D(X)$ 检查 $3X - 2$ 的方差.

练习 4.36

设 $E(X) = 2, D(X) = 3$. 不假设具体分布, 利用切比雪夫不等式给出

$$P(-1 < X < 5)$$

的下界, 并说明为什么这只是一个普适下界.

练习 4.37

设 $D(X) = 4, D(Y) = 9, \text{Cov}(X, Y) = -3$. 求 $D(2X - Y)$ 与 ρ_{XY} , 并判断二者线性变化方向.

练习 4.38

构造或复述一个“不相关但不独立”的例子, 并分别指出“不相关”和“不独立”的判定依据. 若再补充 (X, Y) 联合正态, 结论如何变化?

注: [章末自测清单] 合上讲义后, 检查自己能否做到:

1. 看到分布就能选择求期望和方差的公式, 并说明存在条件;
2. 求函数期望时不必先求新变量分布;
3. 写线性组合方差时不遗漏协方差项;
4. 把切比雪夫事件整理到以期望为中心;
5. 解释协方差与相关系数的区别;

6. 用反例说明“不相关不等于独立”；
7. 说清二维正态中零相关与独立等价的前提.

若其中任一项不能独立完成, 应回到对应小节重做例题, 而不是只背结论.