

CHAPTER 3 数学特征与特征函数

ZHEYU XIE¹

[1]

1. 数学期望

1.1. 定义.

Definition 1 (离散型随机变量的数学期望). 设离散型随机变量 ξ 的概率分布为 $p_i = P\{\xi = x_i\}$, $i = 1, 2, \dots$, 若 $\sum_i |x_i|p_i < +\infty$, 则称

$$(1) \quad E(\xi) = \sum_i x_i p_i$$

为随机变量 ξ 的数学期望。

Definition 2 (连续型随机变量的数学期望). 设连续型随机变量 ξ 的概率密度为 $p(x)$, 若 $\int_{-\infty}^{+\infty} |x|p(x)dx < +\infty$, 则称

$$(2) \quad E(\xi) = \int_{-\infty}^{+\infty} xp(x)dx$$

为随机变量 ξ 的数学期望。

Definition 3 (数学期望的统一写法). 设 ξ 为随机变量, 则定义

$$(3) \quad E(\xi) = \int_{-\infty}^{+\infty} x dF(x)$$

为随机变量 ξ 的数学期望。

1.2. 基本性质.

Proposition 1 (数学期望的性质). 设 ξ, η 为随机变量, 且都有有限的数学期望, 则有

$$(a) \quad E(c) = c$$

$$(b) \quad E(c\xi) = cE(\xi)$$

$$(c) \quad E(\xi + \eta) = E(\xi) + E(\eta)$$

$$(d) \quad \text{若 } \xi \geq 0, \text{ 则 } E(\xi) \geq 0$$

$$(e) \quad \text{若 } \xi \geq \eta, \text{ 则 } E(\xi) \geq E(\eta)$$

¹ DEPARTMENT OF MATHEMATICS, TSINGHUA UNIVERSITY, BEIJING, CHINA.

E-mail address: xie.zeyu20@gmail.com.

Date: 2024 年 4 月 16 日.

Proposition 2 (Borel 函数下的数学期望). 设 ξ 为随机变量, $f(x)$ 为 Borel 函数, 则有

$$(4) \quad E[f(\xi)] = \int_{-\infty}^{+\infty} f(x) dF(x)$$

此性质在多元随机变量的情况下也成立: 设随机变量 $(\xi_1, \xi_2, \dots, \xi_n)$ 有联合分布函数 $F(x_1, x_2, \dots, x_n)$, f 为 n 元 Borel 函数, 则有

$$(5) \quad E[f(\xi_1, \xi_2, \dots, \xi_n)] = \int_{-\infty}^{+\infty} \cdots \int_{-\infty}^{+\infty} f(x_1, x_2, \dots, x_n) dF(x_1, x_2, \dots, x_n)$$

注: 以上性质说明, 可直接用 $(\xi_1, \xi_2, \dots, \xi_n)$ 的联合分布计算 $\eta = f(\xi_1, \xi_2, \dots, \xi_n)$ 的数学期望, 而不必先求出 η 的分布。

1.3. 独立随机变量的性质.

Proposition 3 (独立随机变量数学期望的性质). 设 ξ, η 为独立随机变量, 且 ξ 与 η 均可积, 则乘积 $\xi\eta$ 也可积, 且有

$$(6) \quad E(\xi\eta) = E(\xi)E(\eta)$$

Proposition 4 (独立随机变量数学期望的等价条件). 概率空间 (Ω, \mathcal{F}, P) 中的随机变量 ξ, η 独立的充要条件是: 对任意使得 $f(\xi)$ 和 $g(\eta)$ 可积的 Borel 函数 f, g , 有

$$(7) \quad E[f(\xi)g(\eta)] = E[f(\xi)]E[g(\eta)]$$

1.4. 极限性质. 我们考虑

$$(8) \quad \lim_{n \rightarrow \infty} E(\xi_n) = E(\lim_{n \rightarrow \infty} \xi_n)$$

成立的条件。

Proposition 5 (单调收敛定理). 设随机变量序列 $\{\xi_n\}$ 满足条件

$$(9) \quad 0 \leq \xi_1(\omega) \leq \xi_2(\omega) \leq \cdots \leq \xi_n(\omega) \uparrow \xi(\omega), \quad \forall \omega \in \Omega$$

则 (8) 成立。

Proposition 6 (Fatou 引理). 设 $\{\xi_n\}$ 是一随机变量序列

(a) 若存在可积随机变量 σ , 使得 $\xi_n \geq \sigma$, 则有

$$(10) \quad E(\liminf_{n \rightarrow \infty} \xi_n) \leq \liminf_{n \rightarrow \infty} E(\xi_n)$$

(b) 若存在可积随机变量 τ , 使得 $\xi_n \leq \tau$, 则有

$$(11) \quad E(\limsup_{n \rightarrow \infty} \xi_n) \geq \limsup_{n \rightarrow \infty} E(\xi_n)$$

Proposition 7 (Lebesgue 控制收敛定理). 设 $\{\xi_n\}$ 是一随机变量序列, 若存在可积随机变量 η 使得 $|\xi_n| \leq \eta$, 且 $\lim_{n \rightarrow \infty} \xi_n = \xi$, 则 8 成立, 即

$$(12) \quad \lim_{n \rightarrow \infty} E(\xi_n) = E(\xi)$$

1.5. 常见分布的期望.

Proposition 8 (Bernoulli 分布的数学期望). 设随机变量 ξ 服从参数为 p 的 *Bernoulli* 分布, 则有

$$(13) \quad E(\xi) = p$$

Proposition 9 (二项分布的数学期望). 设随机变量 ξ 服从参数为 (n, p) 的二项分布, 则有

$$(14) \quad E(\xi) = np$$

Proposition 10 (Poisson 分布的数学期望). 设随机变量 ξ 服从参数为 λ 的 *Poisson* 分布, 则有

$$(15) \quad E(\xi) = \lambda$$

证明: 直接计算

Proposition 11 (几何分布的数学期望). 设随机变量 ξ 服从参数为 p 的几何分布, 则有

$$(16) \quad E(\xi) = \frac{1}{p}$$

Proposition 12 (均匀分布的数学期望). 设随机变量 ξ 服从参数为 (a, b) 的均匀分布, 则有

$$(17) \quad E(\xi) = \frac{a+b}{2}$$

Proposition 13 (正态分布的数学期望). 设随机变量 ξ 服从参数为 (μ, σ^2) 的正态分布, 则有

$$(18) \quad E(\xi) = \mu$$

Proposition 14 (χ^2 分布的数学期望). 设随机变量 ξ 服从参数为 n 的 χ^2 分布, 则有

$$(19) \quad E(\xi) = n$$

Proposition 15 (Cauchy 分布的数学期望). 设随机变量 ξ 有密度函数

$$(20) \quad p(x) = \frac{1}{\pi(1+x^2)}$$

注意到

$$(21) \quad \int_{-\infty}^{+\infty} |x|p(x)dx = \int_{-\infty}^{+\infty} \frac{|x|}{\pi(1+x^2)}dx = +\infty$$

故 Cauchy 分布的数学期望不存在。

注：常见的分布中，数学期望不存在的仅有 Cauchy 分布。

1.6. 两个引理.

Proposition 16. $E(\xi^2) = 0$ 的充分必要条件为 $\xi = 0 \quad a.s.$

Proposition 17 (Cauchy-Schwarz 不等式). 设 $\xi, \eta \in L^2$, 则

$$(22) \quad [E(\xi\eta)]^2 \leq E(\xi^2)E(\eta^2)$$

2. 方差

Definition 4 (方差). 随机变量 $\xi \in L^2$ 的数学期望为 $E(\xi)$, 则称

$$(23) \quad D(\xi) = E[(\xi - E(\xi))^2]$$

为随机变量 ξ 的方差。

Definition 5 (标准差). 随机变量 $\xi \in L^2$ 的方差为 $D(\xi)$, 则称

$$(24) \quad \sigma(\xi) = \sqrt{D(\xi)}$$

为随机变量 ξ 的标准差。

Proposition 18 (方差的基本性质). 设 $\xi \in L^2$, c 为常数, 则有

(a) $D(\xi) \geq 0$, 等号成立当且仅当 $\xi = E(\xi) \quad a.s.$

(b) $D(c\xi) = E(\xi^2) - [E(\xi)]^2$

(c) $D(c\xi) = c^2 D(\xi)$

(d) $f(c) = E(\xi - c)^2$ 在 $c = E(\xi)$ 处取得最小值 $D(\xi)$

Example 1: 求 Poisson 分布的方差

Solution 1: 设随机变量 ξ 服从参数为 λ 的 Poisson 分布 $P(\lambda)$, 则有

$$(25) \quad E(\xi) = \lambda, \quad E(\xi^2) = \lambda^2 + \lambda$$

故 Poisson 分布的方差为

$$(26) \quad D(\xi) = E(\xi^2) - [E(\xi)]^2 = \lambda$$

Proposition 19 (常见分布的方差). (a) *Possion* 分布: 设 $\xi \sim P(\lambda)$, 则 $D(\xi) = \lambda$

(b) 正态分布: 设 $\xi \sim N(\mu, \sigma^2)$, 则 $D(\xi) = \sigma^2$

(c) χ^2 分布: 设 $\xi \sim \chi^2(n)$, 则 $D(\xi) = 2n$

(d) *Cauchy* 分布: 方差不存在 (因为数学期望不存在)

(e) 二项分布: 设 $\xi \sim B(n, p)$, 则 $D(\xi) = np(1-p)$

Definition 6 (标准化随机变量). 设随机变量 ξ 的数学期望为 $E(\xi)$, 方差为 $D(\xi)$, 若 $D(\xi) > 0$, 即 ξ 不为常数, 则称

$$(27) \quad \xi^* = \frac{\xi - E(\xi)}{\sqrt{D(\xi)}}$$

为随机变量 ξ 的标准化随机变量。 ξ^* 具有以下性质:

(a) $E(\xi^*) = 0$

(b) $D(\xi^*) = 1$

3. 协方差阵

Definition 7 (协方差). 设随机变量 $\xi, \eta \in L^2$, 则称

$$(28) \quad cov(\xi, \eta) = E[(\xi - E(\xi))(\eta - E(\eta))]$$

为随机变量 ξ, η 的协方差, 协方差反映了 ξ, η 之间的相依程度, 是一个有单位的量。

Definition 8 (协方差阵). 设随机向量 $(\xi_1, \xi_2, \dots, \xi_n)$, 则称

$$(29) \quad B = [b_{ij}] = \begin{pmatrix} cov(\xi_1, \xi_1) & cov(\xi_1, \xi_2) & \cdots & cov(\xi_1, \xi_n) \\ cov(\xi_2, \xi_1) & cov(\xi_2, \xi_2) & \cdots & cov(\xi_2, \xi_n) \\ \vdots & \vdots & \ddots & \vdots \\ cov(\xi_n, \xi_1) & cov(\xi_n, \xi_2) & \cdots & cov(\xi_n, \xi_n) \end{pmatrix}$$

为随机向量 $(\xi_1, \xi_2, \dots, \xi_n)$ 的协方差阵。

Proposition 20 (协方差阵的性质). 设随机向量 $(\xi_1, \xi_2, \dots, \xi_n)$, 则协方差阵 B 具有以下性质:

(a) $cov(\xi_i, \xi_j) = E(\xi_i \xi_j) - E(\xi_i)E(\xi_j)$

(b) B 为对称半正定阵

(c) $D(\xi_i + \xi_j) = D(\xi_i) + D(\xi_j) + 2cov(\xi_i, \xi_j)$

Proposition 21 (二维正态分布的协方差阵). 设二维随机向量 (ξ, η) 服从参数为 $(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho)$ 的二维正态分布, 则有

$$(30) \quad B = \begin{pmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{pmatrix}$$

4. 相关系数

Definition 9 (相关系数). 设随机变量 $\xi, \eta \in L^2$, 则称

$$(31) \quad \rho(\xi, \eta) = \frac{\text{cov}(\xi, \eta)}{\sqrt{D(\xi)}\sqrt{D(\eta)}} = E \left[\frac{\xi - E(\xi)}{\sqrt{D(\xi)}} \cdot \frac{\eta - E(\eta)}{\sqrt{D(\eta)}} \right]$$

为随机变量 ξ, η 的相关系数, 相关系数是一个无单位的量。

当 $\rho(\xi, \eta) = 0$ 时, 称 ξ, η 不相关。

Proposition 22 (相关系数的范围). 对空间 L^2 中的非退化随机变量 ξ, η 而言, 其相关系数 $\rho \in [-1, 1]$ 。

$\rho = 1$ 当且仅当 $\eta^* = \xi^* \quad a.s.$

$\rho = -1$ 当且仅当 $\eta^* = -\xi^* \quad a.s.$

Proposition 23. 对于空间 L^2 中的非退化随机变量, 如下四个命题等价:

(a) ξ 和 η 不相关

(b) $\text{cov}(\xi, \eta) = 0$

(c) $E(\xi\eta) = E(\xi)E(\eta)$

(d) $D(\xi + \eta) = D(\xi) + D(\eta)$

Proposition 24 (独立性和相关性的关系). ξ 和 η 独立 $\Rightarrow \xi$ 和 η 不相关, 反之不成立。

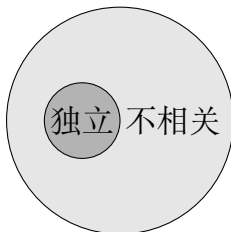


图 1. 独立性和相关性的关系

5. 条件数学期望

Definition 10 (条件分布函数). 对任何有正概率的事件 B , 定义随机变量 ξ 在事件 B 发生的条件下的条件分布函数为

$$(32) \quad F(x|B) = P\{\xi \leq x|B\} = \frac{P\{\xi \leq x, B\}}{P\{B\}}$$

注意到对固定的 B , $F(x|B)$ 也是一个分布函数, 因此可以考虑对它的 Lebesgue-Stieltjes 积分

Definition 11 (条件数学期望). 如果相应的积分绝对收敛, 则称

$$(33) \quad E(\xi|B) = \int_{-\infty}^{+\infty} x dF(x|B)$$

为随机变量 ξ 在事件 B 发生的条件下的条件数学期望。

特别地, 当 $B = \{\eta = y\}$ 时, 称 $E(\xi|\eta = y)$ 为给定 $\eta = y$ 时 ξ 的条件数学期望。

Definition 12 (给定 $\eta = y$ 的条件期望). (a) 若 η 为离散型随机变量, 则有

$$(34) \quad E(\xi|\eta = y_j) = \sum_i x_i P\{\xi = x_i | \eta = y_j\}$$

(b) 若 η 为连续型随机变量, 则有

$$(35) \quad E(\xi|\eta = y) = \int_{-\infty}^{+\infty} x p(x|\eta = y) dx$$

当 ξ 和 η 独立时, 条件分布等于无条件分布, 即 $E(\xi|\eta = y) = E(\xi)$ 。

若 g 为 Borel 函数, 则有

$$(36) \quad E[g(\xi)|\eta = y] = \int_{-\infty}^{+\infty} g(x) dF(x|\eta = y)$$

特别地, 当 ξ 平方可积时, 取 $g(\xi) = [\xi - E(\xi|\eta = y)]^2$, 则得到给定 $\eta = y$ 后 ξ 的条件方差

Definition 13 (条件方差).

$$(37) \quad D(\xi|\eta = y) = E\{(\xi - E[\xi|\eta = y])^2 | \eta = y\}$$

深入理解: 条件数学期望是一个随机变量, 它是一个关于 η 的函数, 当 η 取定值 y 时, 条件数学期望 $E(\xi|\eta = y)$ 是一个确定的数, 它是 ξ 在给定 $\eta = y$ 的条件下的数学期望。

Proposition 25 (全期望公式). 设 (ξ, η) 为联合分布的随机变量, $g(\xi)$ 为可积随机变量, 则有

$$(38) \quad E\{E[g(\xi)|\eta]\} = E[g(\xi)]$$

其含义为: 条件期望 (作为一个随机变量) 的期望等于无条件期望。

Example 1: [Baghdad 窃贼问题] 一个窃贼被关在有三个门的地牢中, 其中

- 一号门通向自由, 出这个门后走三个小时便返回地面
- 二号门通向一个地道, 在其中走五个小时后返回地牢
- 三号门通向一个地道, 在其中走七个小时后返回地牢

如果窃贼每次选择每个门的概率相等，求他为获得自由而走的平均时间。

Solution 1: 设窃贼走出地牢的时间为 ξ ，则由全期望公式有

$$(39) \quad E(\xi) = \frac{1}{3}E(\xi|1) + \frac{1}{3}E(\xi|2) + \frac{1}{3}E(\xi|3)$$

其中

$$(40) \quad E(\xi|1) = 3, \quad E(\xi|2) = 5 + E(\xi), \quad E(\xi|3) = 7 + E(\xi)$$

故有

$$(41) \quad E(\xi) = \frac{1}{3} \cdot 3 + \frac{1}{3} \cdot (5 + E(\xi)) + \frac{1}{3} \cdot (7 + E(\xi))$$

解得 $E(\xi) = 15$ ，即窃贼走出地牢的平均时间为 15 小时。

Proposition 26. 设 ξ 和 η 是联合分布着的平方可积随机变量。如果以 $f(\eta)$ 作为 ξ 的预测，其中 f 是 Borel 函数，则使预测方差 $D(\xi - f(\eta))$ 最小的函数 f 是 $f(\eta) = E(\xi|\eta)$ 。

6. 矩

Definition 14 (矩). 如果 $E|\xi|^k < +\infty$ ，则称 $m_k = E(\xi^k)$ 为随机变量 ξ (及其分布) 的 k 阶原点矩，称 $c_k = E[(\xi - E(\xi))^k]$ 为随机变量 ξ (及其分布) 的 k 阶中心矩。

Proposition 27 (低阶原点矩). (a) 数学期望是一阶原点矩 m_1

(b) 方差是二阶中心矩 c_2

Proposition 28. 根据

$$(42) \quad |\xi|^k \leq 1 + |\xi|^{k+1}$$

可得：若随机变量 ξ 的高阶矩有限，则其低阶矩也有限。

Proposition 29 (原点矩和中心矩的关系). (a) $c_k = \sum_{i=0}^k \binom{k}{i} (-m_1)^{k-i} m_i$

(b) $m_k = \sum_{i=0}^k \binom{k}{i} c_{k-i} m_1^i$

REFERENCES

[1] 杨振明. 概率论 (第二版). 北京: 科学出版社, 2007.