

第三章、随机变量的数字特征

- 分布的完整刻画:
 - 离散型的分布列(PMF)、
 - 连续型的密度函数(PDF)、
 - 一般类型的分布函数(CDF).
- 实际问题中, 难以完整刻画/描述/确定分布.
- 数字特征容易计算, 是研究工具. 如,
 - 中心位置特征(期望);
 - 分散程度特征(方差).

§3.1 离散型随机变量的数学期望

例. 设某种彩票发行了10万张, 每张售1元, 其中10张有奖, 各奖励5000元, 其它无奖. 问: 买一张彩票, 平均盈利多少?

- 将“盈利”视为随机变量, 记为 X . 则 X 可取值5000 - 1 或 -1. 对应的概率为 $\frac{10}{100000}$ 和 $1 - \frac{10}{100000}$, (空间百分比).
- 平均盈利为

$$EX = 4999 \cdot \frac{10}{100000} - 1 \cdot \left(1 - \frac{10}{100000}\right) = -0.5(\text{元}).$$

- 注: “计算4999 与 -1 的算术平均, 得2499 元” 是荒谬的.
- 注: 期望是可能值的加权平均.

例. 放射性粒子求平均粒子数.

数据如右表所示:

- $\hat{p}_k =$ 取 k 的频率/(时间)百分比.

- 泊松分布列:

$$p_k, k = 0, 1, 2, \dots$$

- 平均粒子数:

$$\begin{aligned} & \frac{1}{2608} (0 \times 57 + 1 \times 203 + \\ & 2 \times 383 + \dots + 10 \times 16) \\ &= \sum_{k=0}^{\infty} k \hat{p}_k \approx \sum_{k=0}^{\infty} k p_k. \end{aligned}$$

- 注: 期望 \approx 数据的算术平均.

X	频数	频率	p_k
0	57	0.022	0.021
1	203	0.078	0.081
2	383	0.147	0.156
3	525	0.201	0.202
4	532	0.204	0.195
5	408	0.156	0.151
6	273	0.105	0.097
7	139	0.053	0.054
8	45	0.017	0.026
9	27	0.010	0.011
≥ 10	16	0.006	0.007
总计	2608	1.000	1.000

- 分布特征中最重要的特征之一: 中心位置特征(期望).
- 例. 一批数据 a_1, a_2, \dots, a_n 的平均值:

$$\bar{a} = \frac{1}{n} \sum_{i=1}^n a_i.$$

统计频数/频率: 可能值 x_k 出现 n_k 次, 频率为 $\frac{n_k}{n}$. 于是

$$\bar{a} = \frac{1}{n} \sum_k x_k n_k = \sum_k x_k \cdot \frac{n_k}{n} \approx \sum_k x_k p_k.$$

- 定义1.1. 设 X 为离散型随机变量, 概率分布如下:

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

称 $\sum_k x_k p_k$ 为 X 的(数学)期望或均值, 记为 EX 或 $E(X)$.

- 注: 也称为该分布的期望/均值.
- 注: 假设级数绝对收敛. 否则期望不存在.

两点分布 $B(1, p)$ 的期望

• X 的分布:

X	1	0
概率	p	$1-p$

• X 的期望:

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

二项分布 $B(n, p)$ 的期望

- X 的分布: 记 $q = 1 - p$, 则

$$P(X = k) = C_n^k p^k q^{n-k}, \quad k = 0, 1, \dots, n.$$

- X 的期望:

$$\begin{aligned} EX &= \sum_{k=0}^n k \cdot C_n^k p^k q^{n-k} = \sum_{k=1}^n k \cdot \frac{n!}{k!(n-k)!} p^k q^{n-k} \\ &= \sum_{k=1}^n \frac{n!}{(k-1)!(n-k)!} p^k q^{n-k} \quad (k' = k - 1) \\ &= np \sum_{k'=0}^{n-1} \frac{n!}{k'!(n-k')!} p^{k'} q^{n-k'} = np, \quad (n' = n - 1). \end{aligned}$$

泊松分布Poisson(λ) 的期望

- X 的分布:

$$P(X = k) = \frac{\lambda^k}{k!} e^{-\lambda}, \quad k = 0, 1, 2, \dots$$

- X 的期望:

$$\begin{aligned} EX &= \sum_{k=0}^{\infty} k \cdot \frac{\lambda^k}{k!} e^{-\lambda} = \sum_{k=1}^{\infty} \frac{\lambda^k}{(k-1)!} e^{-\lambda} \quad (k' = k - 1) \\ &= \lambda \sum_{k'=0}^{\infty} \frac{\lambda^{k'}}{k'!} e^{-\lambda} = \lambda. \end{aligned}$$

- 注: $k \cdot p_k = \lambda \cdot p_{k-1}$, $k = 1, 2, \dots$

超几何分布(参数为 N, M, n)的期望

- X 的分布:

$$P(X = m) = \frac{C_M^m C_{N-M}^{n-m}}{C_N^n}, \quad m = 0, 1, 2, \dots, n.$$

- X 的期望: 记 $x' = x - 1$. 则

$$\begin{aligned} EX &= \sum_{m=0}^n m \cdot \frac{M!}{m!(M-m)!} \cdot \frac{C_{N-M}^{n-m}}{C_N^n} \\ &= \sum_{m=1}^n \frac{M!}{m!(M-m)!} \cdot \frac{C_{N-M}^{n-m}}{C_N^n} = \sum_{m=1}^n \frac{M \cdot M'}{m!(M'-m')!} \cdot \frac{C_{N-M}^{n-m}}{C_N^n} \\ &= M \sum_{m'=0}^{n'} C_{M'}^{m'} \cdot \frac{C_{N-M'}^{n-m'}}{C_N^n} = M \cdot \frac{C_{N'}^{n'}}{C_N^n} \sum_{m'=0}^{n'} C_{M'}^{m'} \cdot \frac{C_{N'-M'}^{n-m'}}{C_{N'}^{n'}} \\ &= M \cdot \frac{N'!}{n!(N'-n')!} \cdot \frac{n!(N-n)!}{N!} = \frac{Mn}{N} = n \cdot \frac{M}{N}. \end{aligned}$$

§3.2 连续型随机变量的期望

- 定义2.1. 设 X 为连续型随机变量, 密度为 $p(x)$, 称

$$\int_{-\infty}^{\infty} xp(x)dx \quad (2.1)$$

为 X 的(数学)期望/均值, 记作 $E(X)$, 简记为 EX .

- 注: 也称为该分布的期望/均值.
- 注: 假设积分收敛, 否则期望不存在.
- 注: $p(x)$ 不是概率, $p(x)dx$ 才对应着概率.
- 直观: 把 X 的取值范围划分为小区间 $(x_i, x_{i+1}]$, $i = 1, 2, \dots$, 设 $p(x)$ 连续, 则近似地

$$\begin{aligned} EX &\approx \sum_i x_i P(x_i < X \leq x_{i+1}) \\ &\approx \sum_i \int_{x_i}^{x_{i+1}} xp(x)dx \approx \int_{-\infty}^{\infty} xp(x)dx. \end{aligned}$$

均匀分布 $U[a, b]$ 的期望

- X 的密度:

$$p(x) = \begin{cases} \frac{1}{b-a}, & x \in [a, b], \\ 0, & \text{其它}. \end{cases}$$

- X 的期望:

$$\begin{aligned} EX &= \int_{-\infty}^{\infty} xp(x)dx = \int_a^b x \cdot \frac{1}{b-a} dx \\ &= \frac{1}{b-a} \cdot \frac{x^2}{2} \Big|_a^b = \frac{1}{2} \cdot \frac{b^2 - a^2}{b-a} = \frac{a+b}{2}. \end{aligned}$$

- 注: 期望是权重的对称中心. (前提: 存在).

指数分布 $E(\lambda)$ 的期望

- X 的密度:

$$p(x) = \begin{cases} \lambda e^{-\lambda x}, & x \geq 0, \\ 0, & \text{其它.} \end{cases}$$

- X 的期望:

$$\begin{aligned} EX &= \int_{-\infty}^{\infty} xp(x)dx = \int_0^{\infty} x\lambda e^{-\lambda x}dx \\ &= \frac{1}{\lambda} \int_0^{\infty} te^{-t}dt \quad (\text{令 } t = \lambda x) \\ &= \frac{1}{\lambda} \left((-te^{-t}) \Big|_0^{\infty} + \int_0^{\infty} e^{-t}dt \right) = \frac{1}{\lambda}. \end{aligned}$$

正态分布 $N(\mu, \sigma^2)$ 的期望

- X 的密度:

$$p(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left\{-\frac{(x-\mu)^2}{2\sigma^2}\right\}.$$

- X 的期望:

$$\begin{aligned} EX &= \int_{-\infty}^{\infty} xp(x)dx \\ &= \int_{-\infty}^{\infty} x \cdot \frac{1}{\sqrt{2\pi}\sigma} \exp\left\{-\frac{(x-\mu)^2}{2\sigma^2}\right\} dx \\ &= \int_{-\infty}^{\infty} (\mu + \sigma y) \cdot \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{y^2}{2}\right\} dy \quad \left(y = \frac{x-\mu}{\sigma}\right) \\ &= \mu \int_{-\infty}^{\infty} \phi(y) dy + \sigma \int_{-\infty}^{\infty} y\phi(y) dy = \mu. \end{aligned}$$

- 注: μ 是权重的对称中心.

伽玛分布 $\Gamma(\alpha, \beta)$ 的期望

- X 的密度:

$$p(x) = \begin{cases} \frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\beta x}, & x > 0, \\ 0, & x \leq 0. \end{cases} \quad (\alpha > 0, \beta > 0)$$

- X 的期望:

$$\begin{aligned} EX &= \int_{-\infty}^{\infty} xp(x)dx = \frac{\beta^\alpha}{\Gamma(\alpha)} \int_0^{\infty} x^\alpha e^{-\beta x} dx \\ &= \frac{\beta^\alpha}{\Gamma(\alpha)} \int_0^{\infty} x^{\alpha'-1} e^{-\beta x} dx \quad (\alpha' = \alpha + 1) \\ &= \frac{\beta^\alpha}{\Gamma(\alpha)} \cdot \frac{\Gamma(\alpha')}{\beta^{\alpha'}} = \frac{\beta^\alpha}{\Gamma(\alpha)} \cdot \frac{\Gamma(\alpha + 1)}{\beta^{\alpha+1}} = \frac{\alpha}{\beta}. \end{aligned}$$

§3.3 期望的简单性质及随机变量函数的期望公式

- 设 X 为随机变量, c, k, b 为常数. 则

$$(1) E(c) = c;$$

$$(2) E(kX) = kE(X);$$

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

$$(4) \text{线性: } E(kX + b) = kE(X) + b.$$

(3.1)

- 证: (1) $P(X = c) = 1$. 按定义知(1) 成立.
由(2), (3) 知(4) 成立.

(2) $E(kX) = kE(X)$ 的证明:

- 若 $k = 0$, 则两边均为0, \checkmark . 下设 $k \neq 0$.
- 离散型: 设 X 的分布为 $P(X = x_i) = p_i, i = 1, 2, \dots$
则 $Y = kX$ 的分布为 $P(Y = kx_i) = p_i, i = 1, 2, \dots$
按定义, $E(Y) = \sum_i kx_i \cdot p_i = k \sum_i x_i p_i = kE(X)$.
- 连续型: 设 X 的密度为 $p(x)$,
则 $Y = kX$ 的密度为 $p_Y(y) = \frac{1}{|k|} p\left(\frac{y}{k}\right)$. (参见第二章, 例4.8).
按定义

$$\begin{aligned} E(Y) &= \int_{-\infty}^{\infty} y \cdot p_Y(y) dy = \frac{1}{|k|} \int_{-\infty}^{\infty} y \cdot p\left(\frac{y}{k}\right) dy \\ &= k \int_{-\infty}^{\infty} x \cdot p(x) dx = kE(X). \quad (x = \frac{y}{k}). \end{aligned}$$

(3) $E(X + b) = E(X) + b$ 的证明:

- 离散性与(2) 类似. 下设 X 为连续型, 密度为 $p(x)$.
- $p_Y(y) = p(y - b)$. (参见第二章, 例4.8)

按定义,

$$\begin{aligned} E(Y) &= \int_{-\infty}^{\infty} y \cdot p_Y(y) dy = \int_{-\infty}^{\infty} y \cdot p(y - b) dy \\ &= \int_{-\infty}^{\infty} (x + b) p(x) dx \quad (x = y - b) \\ &= \int_{-\infty}^{\infty} x p(x) dx + b \int_{-\infty}^{\infty} p(x) dx = E(X) + b. \end{aligned}$$

随机变量函数的期望公式

- 离散型: 设 X 的分布列为 $P(X = x_i) = p_i, i = 1, 2, \dots$.
若下式右边绝对收敛, 则

$$Ef(X) = \sum_i f(x_i)p_i. \quad (3.3)$$

- 连续型: 设 X 的密度为 $p(x)$. 若下式右边绝对收敛, 则

$$Ef(X) = \int_{-\infty}^{\infty} f(x)p(x)dx. \quad (3.2)$$

- 注: 免去了求 $Y = f(X)$ 的分布列或密度的过程.
- 注: 符号约定. 例 EX^k 表示 $E(X^k)$, 不表示 $(EX)^k$.

例3.1. 设 $X \sim N(0, 1)$. 求 EX^2 .

- 解法1.(更简洁). 用公式(3.2).

$$\begin{aligned} EX^2 &= \int_{-\infty}^{\infty} x^2 \cdot \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx = - \int_{-\infty}^{\infty} x d \left(\frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \right) \\ &= - \left(x \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \right) \Big|_{-\infty}^{\infty} + \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx = 0 + 1 = 1. \end{aligned}$$

- 解法2. 先求 X^2 的密度(参见§2.4 例). $X^2 \sim \Gamma(\frac{1}{2}, \frac{1}{2})$.

$$\begin{aligned} p_{X^2}(y) &= \frac{1}{\sqrt{2\pi}} y^{-\frac{1}{2}} e^{-\frac{1}{2}y}, \quad (y > 0); \quad p_{X^2}(y) = 0, \quad (y \leq 0). \\ \Rightarrow EX^2 &= \int_0^{\infty} y \cdot p_{X^2}(y) dy = \int_0^{\infty} \frac{\sqrt{y}}{\sqrt{2\pi}} e^{-\frac{y}{2}} dy \\ &= \int_0^{\infty} \frac{x}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} 2x dx \quad (x = \sqrt{y}) \\ &= - \int_0^{\infty} \frac{1}{\sqrt{2\pi}} 2x d(e^{-\frac{x^2}{2}}) = 0 + 2 \int_0^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx = 1. \end{aligned}$$

例3.2 $X \sim U[0, 2\pi]$, 求 $E \sin X$.

• 解:

$$\begin{aligned} E \sin X &= \int_{-\infty}^{\infty} \sin(x) \cdot p_X(x) dx \\ &= \int_0^{2\pi} \sin(x) \cdot \frac{1}{2\pi} dx = 0. \end{aligned}$$

• 注: 直观上, 由对称性知 $E \sin X = 0$.

§3.4 方差及其简单性质

- 分布特征中最重要的特征之二：分散程度/宽窄特征(方差).
- 例. 甲、乙两个女生小合唱队的身高如下, 平均都是1.60.
甲队: 1.60, 1.62, 1.59, 1.60, 1.59; 整齐/方差小.
乙队: 1.80, 1.60, 1.50, 1.50, 1.60. 参差不齐/方差大.
- 例. 一批数据 a_1, a_2, \dots, a_n 的波动程度.
如, 产品的某种特性(如强度)波动大, 说明生产不稳定.
如, 生物的某种特性(如血压)波动大, 表示病态.
- 一批数据的分散程度/波动程度:

$$\frac{1}{n-1} \left((a_1 - \bar{a})^2 + (a_2 - \bar{a})^2 + \dots + (a_n - \bar{a})^2 \right). \quad (4.1)$$

- 数据来自随机变量 X , 则 $\bar{a} \approx EX \triangleq \mu$.
上式约为 $Y = (X - \mu)^2$ 对应的数据的平均值, 故 $\approx EY$.

- 定义: 设 EX 存在, 称

$$E(X - EX)^2 \quad (4.3)$$

为 X 的方差, 记为 $D(X)$ 或 $\text{Var}(X)$. 也称为其分布的方差.

- 定义4.1(离散型) & 4.2(连续型)

设 X 的分布列如下, 或密度为 $p(x)$.

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

则 $D(X)$ 有如下表达式:

$$\sum_k (x_k - EX)^2 p_k, \quad \int_{-\infty}^{\infty} (x - EX)^2 p(x) dx \quad (4.2, 4.2')$$

- 注: 若级数/积分收敛, 则方差存在; 否则方差发散.
- 方差是非负的: $D(X) \geq 0$.

- 方差的恒等式/计算公式:

$$D(X) = EX^2 - (EX)^2. \quad (4.4)$$

- 需要用第四章的定理: $E(X + Y) = E(X) + E(Y)$.
- (4.4) 的证明: 记 $\mu = EX$.

$$\begin{aligned} D(X) &= E(X - \mu)^2 = E(X^2 - 2\mu \cdot X + \mu^2) \\ &= EX^2 - 2\mu \cdot EX + \mu^2 = EX^2 - \mu^2. \end{aligned}$$

- 另证: 如, $X \sim p(x)$, 记 $\mu = EX$. 则

$$\begin{aligned} D(X) &= \int_{-\infty}^{\infty} (x - \mu)^2 p(x) dx = \int_{-\infty}^{\infty} (x^2 - 2\mu x + \mu^2) p(x) dx \\ &= \int_{-\infty}^{\infty} x^2 p(x) dx - 2\mu \int_{-\infty}^{\infty} x p(x) dx + \mu^2 \cdot 1 \\ &= EX^2 - 2\mu \cdot \mu + \mu^2 = EX^2 - \mu^2. \end{aligned}$$

两点分布 $B(1, p)$ 的方差

- X 的期望: $EX = p$.
- X 的方差: 按定义,

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

- X 的方差: 按计算公式,

$$EX^2 = 1^2 \cdot p + 0^2 \cdot (1 - p) = p$$

$$\Rightarrow D(X) = EX^2 - (EX)^2 = p - p^2 = p(1 - p).$$

二项分布 $B(n, p)$ 的方差

- X 的期望: $EX = np$.
- X 的方差: 按定义, ……
- X 的方差: 记 $q = 1 - p$, 按计算公式,

$$\begin{aligned} EX^2 &= \sum_{k=0}^n k^2 \cdot \frac{n!}{k!(n-k)!} p^k q^{n-k} \\ &= \sum_{k=1}^n (k(k-1) + k) \frac{n!}{k!(n-k)!} p^k q^{n-k} \\ &= n(n-1)p^2 \sum_{k'=0}^{n'} \frac{n!}{k'!(n-k')!} p^{k'} q^{n-k'} + EX \quad (x' = x - 2) \\ &= n(n-1)p^2 + np = n^2p^2 - np^2 + np \\ \Rightarrow D(X) &= EX^2 - (EX)^2 = np(1-p). \end{aligned}$$

泊松分布Poisson(λ) 的方差

- X 的期望: $EX = \lambda$.
- X 的方差: $k' = k - 1, k'' = k - 2$, 按计算公式,

$$\begin{aligned} EX^2 &= \sum_{k=1}^{\infty} k^2 \cdot \frac{\lambda^k}{k!} e^{-\lambda} = \sum_{k=1}^{\infty} k \cdot \frac{\lambda^k}{(k-1)!} e^{-\lambda} \\ &= \sum_{k=1}^{\infty} ((k-1) + 1) \cdot \frac{\lambda^k}{(k-1)!} e^{-\lambda} \\ &= \sum_{k=2}^{\infty} \frac{\lambda^k}{(k-2)!} e^{-\lambda} + \sum_{k=1}^{\infty} \frac{\lambda^k}{(k-1)!} e^{-\lambda} \\ &= \lambda^2 \sum_{k''=0}^{\infty} \frac{\lambda^{k''}}{k''!} e^{-\lambda} + \lambda \sum_{k'=0}^{\infty} \frac{\lambda^{k'}}{k'!} e^{-\lambda} = \lambda^2 + \lambda \\ \Rightarrow D(X) &= EX^2 - (EX)^2 = \lambda. \end{aligned}$$

均匀分布 $U[a, b]$ 的方差

- X 的期望: $EX = \frac{a+b}{2}$.
- X 的方差: 按计算公式,

$$\begin{aligned} EX^2 &= \int_a^b x^2 \cdot \frac{1}{b-a} dx = \frac{1}{b-a} \cdot \frac{x^3}{3} \Big|_a^b = \frac{1}{3}(a^2 + ab + b^2) \\ \Rightarrow D(X) &= EX^2 - (EX)^2 \\ &= \frac{1}{3}(a^2 + ab + b^2) - \frac{1}{4}(a^2 + 2ab + b^2) = \frac{1}{12}(b-a)^2. \end{aligned}$$

- X 的方差: 按定义,

$$\begin{aligned} D(X) &= E(X - EX)^2 = \int_a^b \left(x - \frac{a+b}{2}\right)^2 \cdot \frac{1}{b-a} dx \\ &= \frac{1}{b-a} \cdot \frac{\left(x - \frac{a+b}{2}\right)^3}{3} \Big|_a^b = \frac{1}{3(b-a)} \cdot 2\left(\frac{b-a}{2}\right)^3 = \frac{1}{12}(b-a)^2. \end{aligned}$$

指数分布 $\text{Exp}(\lambda)$ 的方差

- X 的期望: $EX = \frac{1}{\lambda}$.
- X 的方差: 按计算公式,

$$\begin{aligned} EX^2 &= \int_0^{\infty} x^2 \cdot \lambda e^{-\lambda x} dx = \frac{1}{\lambda^2} \int_0^{\infty} t^2 e^{-t} dt \quad (t = \lambda x) \\ &= \frac{1}{\lambda^2} \Gamma(3) = \frac{1}{\lambda^2} \cdot 2! = \frac{2}{\lambda^2} \\ \Rightarrow D(X) &= EX^2 - (EX)^2 = \frac{1}{\lambda^2}. \end{aligned}$$

正态分布 $N(\mu, \sigma^2)$ 的方差

- X 的期望: $EX = \mu$.
- X 的方差: 按定义,

$$\begin{aligned} D(X) &= E(X - \mu)^2 \\ &= \int_{-\infty}^{\infty} (x - \mu)^2 \cdot \frac{1}{\sqrt{2\pi}\sigma} \exp\left\{-\frac{1}{2} \frac{(x - \mu)^2}{\sigma^2}\right\} dx \\ &= \sigma^2 \int_{-\infty}^{\infty} t^2 \cdot \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}t^2} dt \quad (\text{令 } t = \frac{x - \mu}{\sigma}) \\ &= \sigma^2. \quad (\text{见例3.1}) \end{aligned}$$

- 注: 两个参数 μ, σ^2 分别是期望和方差.

伽玛分布 $\Gamma(\alpha, \beta)$ 的方差

- X 的密度: 固定 β ,

$$p_{\alpha}(x) = \frac{\beta^{\alpha}}{\Gamma(\alpha)} x^{\alpha-1} e^{-\beta x}, \quad (x > 0).$$

- X 的期望: $EX = \frac{\alpha}{\beta}$.
- X 的方差: $\alpha' = \alpha + 2$, 按计算公式,

$$\begin{aligned} EX^2 &= \int_0^{\infty} x^2 \frac{\beta^{\alpha}}{\Gamma(\alpha)} x^{\alpha-1} e^{-\beta x} dx \\ &= \frac{\beta^{\alpha}}{\Gamma(\alpha)} \cdot \frac{\Gamma(\alpha')}{\beta^{\alpha'}} \int_0^{\infty} \frac{\beta^{\alpha'}}{\Gamma(\alpha')} x^{\alpha'-1} e^{-\beta x} dx \\ &= \frac{\Gamma(\alpha+2)}{\beta^2 \Gamma(\alpha)} = \frac{(\alpha+1)\alpha}{\beta^2} \\ \Rightarrow D(X) &= \frac{(\alpha+1)\alpha}{\beta^2} - \frac{\alpha^2}{\beta^2} = \frac{\alpha}{\beta^2}. \end{aligned}$$

方差的简单性质

- 设 X 为随机变量, c, k, b 为常数. 则

$$(1) \quad D(c) = 0;$$

$$(2) \quad D(kX) = k^2 D(X);$$

$$(3) \quad D(X + b) = D(X);$$

$$(4) \quad D(kX + b) = k^2 D(X).$$

(4.5)

- 证: (1) $E(c) = c$. 按定义, $D(X) = E(c - c)^2 = 0$.
由(2), (3) 知(4) 成立.

(2) $D(kX) = k^2D(X)$ 的证明:

- 按计算公式,

$$\begin{aligned} E(kX) &= kE(X), \quad E[(kX)^2] = E(k^2X^2) = k^2EX^2 \\ \Rightarrow D(kX) &= k^2EX^2 - (kEX)^2 \\ &= k^2(EX^2 - (EX)^2) = k^2D(X). \end{aligned}$$

(3) $D(X + b) = D(X)$ 的证明:

- 按定义,

$$\begin{aligned} E(X + b) &= E(X) + b \\ \Rightarrow D(X + b) &= E((X + b) - E(X + b))^2 \\ &= E(X - EX)^2 = D(X). \end{aligned}$$

切比雪夫不等式

§3.5 其它

定理5.1 设 $\mu = EX$ 与 $\sigma^2 = D(X)$ 存在. 则对任意 $\varepsilon > 0$,

$$P(|X - \mu| \geq \varepsilon) \leq \frac{\sigma^2}{\varepsilon^2}. \quad (5.1)$$

- 连续型情形的证明:

$$\begin{aligned} D(X) &= \int_{-\infty}^{\infty} (x - \mu)^2 p(x) dx \geq \int_{x \leq \mu - \varepsilon \text{ 或 } x \geq \mu + \varepsilon} (x - \mu)^2 p(x) dx \\ &\geq \int_{-\infty}^{\mu - \varepsilon} \varepsilon^2 p(x) dx + \int_{\mu + \varepsilon}^{\infty} \varepsilon^2 p(x) dx \\ &\geq \varepsilon^2 \left(\int_{-\infty}^{\mu - \varepsilon} p(x) dx + \int_{\mu + \varepsilon}^{\infty} p(x) dx \right) \\ &= \varepsilon^2 P(X \leq \mu - \varepsilon \text{ 或 } X \geq \mu + \varepsilon) = \varepsilon^2 P(|X - \mu| \geq \varepsilon). \end{aligned}$$

- 注: σ 反映了 X 分布的权重的分散程度.

定理5.1 设 $\mu = EX$ 与 $\sigma^2 = D(X)$ 存在. 则对任意 $\varepsilon > 0$,

$$P(|X - \mu| \geq \varepsilon) \leq \frac{\sigma^2}{\varepsilon^2}. \quad (5.1)$$

- 称 $\sigma = \sqrt{D(X)}$ 为 X 的标准差.
- 取 $\varepsilon = k\sigma$, 则

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

- 特别地, 取 $k = 3$, 则推出

$$P(|X - \mu| \geq 3\sigma) \leq \frac{1}{9}.$$

- 注: 正态分布的经验规则 $P(|X - \mu| \geq 3\sigma) \leq 0.0027$. (§2.3)

原点矩与中心矩

- 分别称 EX^k , $E(X - EX)^k$ 为 X 的 k 阶 **原点矩**, **中心矩**, 分别记为 ν_k , μ_k , ($k = 1, 2, \dots$).
- 注: 均值为 ν_1 , 方差为 μ_2 .
- 注: k 也可以不是正整数.

分位数与中位数

- 设 X 的分布函数 $F(\cdot) : x \mapsto p = F(x)$ 连续且严格单调上升. 则存在反函数 $p \mapsto x = x_p, p \in (0, 1)$, 即存在唯一的 x_p 使得

$$F(x_p) = p, \quad \forall p \in (0, 1).$$

- 称 x_p 为 X 的 p 分位数. 称 $p \mapsto x_p$ 为 X 的分位数函数.
- 一般情形. 若下式成立, 则称 x_p 是 X 的 p 分位数.

$$P(X < x_p) \leq p \leq P(X \leq x_p). \quad (5.4)$$

$$\text{等价地, } P(X \leq x_p) \geq p, \quad P(X \geq x_p) \geq 1 - p. \quad (5.4')$$

- 令 $p = \frac{1}{2}$. 称 $x_{\frac{1}{2}}$ 为中位数.
- 注: 分位数存在(如, $x_p = \inf\{x : F(X) \geq p\}$), 但不一定唯一.

例. 两点分布 $B(1, p_0)$ 的分位数.

- 记 $q_0 = 1 - p_0$. 分布函数:

$$F(x) = \begin{cases} 0, & x < 0; \\ q_0, & 0 \leq x < 1; \\ 1, & x \geq 1. \end{cases}$$

- 分位数:

$$x_p : \begin{cases} 0, & p \in (0, q_0); \\ [0, 1], & p = q_0; \\ 1, & p \in (q_0, 1). \end{cases}$$