概　　率　　论

Probability

已学知识点

- 第一章 事件与概率
  - 随机现象与统计规律性

    ① 概率的频率解释依然是当今最通行的解释.

    ② 描述频率趋近于概率的大数定律总是概率论的第一大数定律.

    ③ 实际当中用频率作为概率的估计是十分自然的.
  - 样本空间与事件

符号	集合论含义	概率论含义
Ω	空间或全集	样本空间或必然事件
Φ	空集	不可能事件
ω	元素	样本点
A	子集	随机事件
$\omega \in A$	ω 是 A 的元素	事件 A 包含样本点 ω
$A \subset B$	A 是 B 的子集	A 发生则 B 发生
$AB = \Phi$	A, B 不相交	A, B 不可能同时发生
$A \cup B$	并集	A, B 至少有一个发生
$A \cap B$	交集	A, B 同时发生
$A - B$	差集	A 发生而 B 不发生
$\bar{A}$	余集	A 不发生

已学知识点

- 第一章 事件与概率
  - 古典概型 (等可能概率模型)：(1) 样本空间样本点有限；(2) 每个样本点等可能出现.
    - 计数方法：排列组合.
    - 三个基本性质：非负性、规范性、有限可加性.
  - 几何概率：以等可能性定义概率，处理无限场合，概率是几何体的测度之比.
    - 三个基本性质：非负性、规范性、可列可加性.
  - 概率空间：$\left(\Omega, \mathscr{F}, P\right)$
    - 难点和要点：事件域 $\mathscr{F}$ 的选择，太小不能满足需要，太大难以定义概率.
    - 选择包含我们关注的所有事件的 σ 域，保证事件对交、并、逆、差作可列次运算的封闭性.
    - 在这种 σ 域上，能定义满足非负、规范和可列可加性的概率测度.

已学知识点

- 第二章 条件概率与统计独立性
  - 条件概率：$P(A\,|\,B)=\dfrac{P(AB)}{P(B)}.$
    - 乘法公式：$P\left(A_1A_2\cdots A_n\right)=P\left(A_1\right)\cdot P\left(A_2\middle|\,A_1\right)\cdot P\left(A_3\middle|\,A_1A_2\right)\cdot\cdots\cdot P\left(A_n\middle|\,A_1A_2\cdots A_{n-1}\right)$
    - 全概率公式：$P(B)=\sum\limits_{i=1}^{\infty}P\left(A_i\right)\cdot P\left(B\,\middle|A_i\right)$
    - Bayes 公式：$P\left(A_k\,\middle|B\right)=\dfrac{P\left(A_k\right)\cdot P\left(B\,\middle|A_k\right)}{\sum\limits_{i=1}^{\infty}P\left(A_i\right)\cdot P\left(B\,\middle|A_i\right)}$

$$\left.\begin{array}{l} A_i\cap A_j=\Phi\quad(i\neq j)\\ \bigcup\limits_{i=1}^{\infty}A_i=\Omega\end{array}\right.$$

已学知识点

- 第二章 条件概率与统计独立性
  - 事件独立性：两个事件 $P(AB) = P(A) \cdot P(B)$. 三个事件 $\begin{cases} P(AB) = P(A) \cdot P(B) \\ P(AC) = P(A) \cdot P(C) \\ P(BC) = P(B) \cdot P(C) \\ P(ABC) = P(A) \cdot P(B) \cdot P(C) \end{cases}$
  - 试验独立性：一个试验的结果对其它各试验的可能结果的概率都无影响.
  - Bernoulli 试验 E：概率空间 $\left(\Omega, \mathscr{F}, P\right)$，其中

$$A \subset \Omega, \quad \mathscr{F} = \left\{\Phi, A, \bar{A}, \Omega\right\}, \quad P(A) = p, P\left(\bar{A}\right) = q, \quad (p > 0, q > 0, p + q = 1)$$

- Bernoulli 分布
- 二项分布
- 几何 分布
- Pascal 分布
- 多项分布
- Poisson 分布

已学知识点

- 第三章 随机变量与分布函数

  - 随机变量 (r.v.) 与分布函数 (c.d.f.)：

    - 随机变量 ξ 是概率空间 $(\Omega, \mathscr{F}, P)$ 中定义域为 Ω、取值于 $\mathbb{R}$ 的一个单值 Borel 函数.

    - 分布函数 $F(x) = P\left\{\xi(\omega) < x\right\}$ ，$-\infty < x < \infty$ 是单调不降、取值 $[0, 1]$ 的左连续函数. 它完整描述了随机变量，是研究的主要对象.

    - 随机变量依取值 $\begin{cases} \text{离散型：} & \text{分布律 (mass function)、分布列} \\ \text{连续型：} & \text{概率密度 (p.d.f.)} \end{cases}$

    - 主要分布：$\begin{cases} \text{离散型：} & \text{Bernoulli, binomial, Poisson, hyper-geometric, geometric} \\ \text{连续型：} & \text{uniform, exponential, normal, } \Gamma \end{cases}$

    - 概率计算：$\begin{cases} \text{离散型：} & P\{x \in D\} = \sum\limits_{x_i \in D} p_i, \quad P\left\{(x, y) \in D\right\} = \sum\limits_{(x_i, y_j) \in D} p_{ij} \\ \text{连续型：} & P\{x \in D\} = \int\limits_D p(x)\mathrm{d}x, \quad P\left\{(x, y) \in D\right\} = \iint\limits_D p(x, y)\mathrm{d}x\mathrm{d}y \end{cases}$

已学知识点

- 第三章 随机变量与分布函数

▶ 随机向量，随机变量的独立性：

- 随机向量即多元随机变量
  - 联合分布： 联合分布函数、联合分布律、联合密度
  - 边际分布： 边际分布函数、边际分布律、边际密度
  - 条件分布： 条件分布函数、条件分布律、条件密度
  - 独立性： 与事件独立性几乎完全相同

- 主要分布：
  - 离散型： 多项分布、多元超几何分布
  - 连续型： 二元均匀分布、二元正态分布

$$
\begin{aligned}
(\xi,\eta) &\sim N\left(\mu_1,\mu_2,\sigma_1^2,\sigma_2^2,\rho\right)\\
&\sim N\left(\begin{pmatrix}\mu_1\\ \mu_2\end{pmatrix},\begin{pmatrix}\sigma_1^2 & \sigma_1\sigma_2\rho\\ \sigma_1\sigma_2\rho & \sigma_2^2\end{pmatrix}\right)
\end{aligned}
\Longrightarrow
\begin{cases}
\xi\sim N\left(\mu_1,\sigma_1^2\right), \qquad \eta\sim N\left(\mu_2,\sigma_2^2\right)\\
\left(\eta\,\middle|\,\xi=x\right)\sim N\left(\mu_2+\rho\dfrac{\sigma_2}{\sigma_1}(x-\mu_1),\ \sigma_2^2\left(1-\rho^2\right)\right)\\
\left(\xi\,\middle|\,\eta=y\right)\sim N\left(\mu_1+\rho\dfrac{\sigma_1}{\sigma_2}(y-\mu_2),\ \sigma_1^2\left(1-\rho^2\right)\right)\\
\xi \text{ 与 } \eta \text{ 相互独立} \Longleftrightarrow \rho=0
\end{cases}
$$

已学知识点

- 第三章 随机变量与分布函数

随机变量的函数及其分布：

- 随机变量的函数什么情况下还是随机变量
- 离散型易：$\begin{cases} \eta = g(\xi): & \text{对应法} \\ \zeta = g(\xi, \eta): & \text{表上作业法，独立情形和的卷积公式} \end{cases}$
- 连续型难：

$$\begin{cases} \eta = g(\xi): \text{直接法 } F_\eta(y) = P\{\eta < y\} = \displaystyle\int_{g(x)<y} p_\xi(x)\mathrm{d}x \\ \zeta = g(\xi, \eta): \text{直接法 } F_\zeta(z) = P\{\zeta < z\} = \displaystyle\iint_{g(x,\, y)<z} p(x, y)\mathrm{d}x\mathrm{d}y \\ \qquad\qquad \text{和 (卷积公式)、差、商、积、 max , min 的公式} \\ \begin{cases} \zeta_1 = g_1(\xi, \eta) \\ \zeta_2 = g_2(\xi, \eta) \end{cases}: \text{变换法} \begin{cases} u = g_1(x, y) \\ v = g_2(x, y) \end{cases} \Longrightarrow \begin{cases} x = h_1^{-1}(u, v) \\ y = h_2^{-1}(u, v) \end{cases} \Longrightarrow J = \begin{vmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} \end{vmatrix} \end{cases}$$

$$q_{\zeta_1,\, \zeta_2}(u, v) = p\left(h_1^{-1}(u, v), h_2^{-1}(u, v)\right) \cdot |J|$$

已学知识点

- 第四章 数字特征与特征函数
  - 数学期望：
    - 定义 $\begin{cases} \text{离散型：} E(\xi) = \sum\limits_i x_i \cdot p_i \\ \text{连续型：} E(\xi) = \int\limits_{-\infty}^{\infty} x \cdot p_\xi(x)\,\mathrm{d}x \end{cases}$ 绝对收敛！
    - 随机变量函数的数学期望 $\begin{cases} \text{离散型：} E\left[g(\xi)\right] = \sum\limits_i g\left(x_i\right) \cdot p_i \\ \text{连续型：} E\left[g(\xi)\right] = \int\limits_{-\infty}^{\infty} g(x) \cdot p_\xi(x)\,\mathrm{d}x \end{cases}$

已学知识点

- 第四章 数字特征与特征函数

  - 常见分布的数学期望与方差：

分布	符号表示	数学期望
二项分布	$\xi \sim b(n, p)$	$E(\xi) = np$
Bernoulli 分布	$\xi \sim b(1, p)$	$E(\xi) = p$
Poisson 分布	$\xi \sim P(\lambda)$	$E(\xi) = \lambda$
均匀分布	$\xi \sim U[a, b]$	$E(\xi) = \frac{a+b}{2}$
Gamma 分布	$X \sim \Gamma(r, \lambda)$	$E(\xi) = \frac{r}{\lambda}$
指数分布	$\xi \sim \mathrm{Exp}(\lambda)$	$E(\xi) = \frac{1}{\lambda}$
χ^2 分布	$X \sim \chi_n^2$	$E(\xi) = n$
正态分布	$\xi \sim N(\mu, \sigma^2)$	$E(\xi) = \mu$

第四章 数字特征与特征函数

4.1 数学期望

一、数学期望的概念
二、离散型场合
三、应用实例
四、连续型场合
五、一般场合
六、随机变量函数的数学期望
七、多维场合
八、数学期望的基本性质

七、多维场合

- 随机向量函数的期望

  - 设随机向量 $(\xi_1, \xi_2, \ldots, \xi_n)$ 的联合分布函数为 $F(x_1, x_2, \ldots, x_n)$，$g(x_1, x_2, \ldots, x_n)$ 为 n 元 Borel 函数，则 $\eta = g(\xi_1, \xi_2, \ldots, \xi_n)$ 为一元随机变量，其数学期望为 (以 $n = 2$ 为例)

$$E(\eta) = E\left[g(\xi_1, \xi_2)\right] = \begin{cases} \text{离散型：} \sum\limits_i \sum\limits_j g(x_i, y_j) \cdot p_{ij} \\ \text{连续型：} \int\limits_{-\infty}^{\infty} \int\limits_{-\infty}^{\infty} g(x, y) \cdot p_{(\xi,\eta)}(x, y)\,\mathrm{d}x\,\mathrm{d}y \end{cases}$$

定义：定义随机向量 $(\xi_1, \xi_2, \ldots, \xi_n)$ 的数学期望为 $\left(E(\xi_1), E(\xi_2), \ldots, E(\xi_n)\right)$.

七、多维场合

$$E(\eta)=E\left[g(\xi_1,\xi_2)\right]=\begin{cases}\text{离散型：} & \sum_i\sum_j g(x_i,y_j)\cdot p_{ij}\\ \text{连续型：} & \int_{-\infty}^{\infty} g(x,y)\cdot p_{(\xi,\eta)}(x,y)\,\mathrm{d}x\,\mathrm{d}y\end{cases}$$

- 例：(ξ,η) 的联合分布律为

$\eta \backslash \xi$	0	1	2
0	0.10	0.15	0.15
1	0.25	0.20	0.15

求 $\zeta=\sin\dfrac{\pi(\xi+\eta)}{2}$ 的数学期望 $E(\zeta)$.

$$E(\zeta)=E\left\{\sin\frac{\pi(\xi+\eta)}{2}\right\}$$

$$=0.10\times\sin\frac{\pi(0+0)}{2}+0.25\times\sin\frac{\pi(0+1)}{2}+0.15\times\sin\frac{\pi(1+0)}{2}+0.20\times\sin\frac{\pi(1+1)}{2}$$

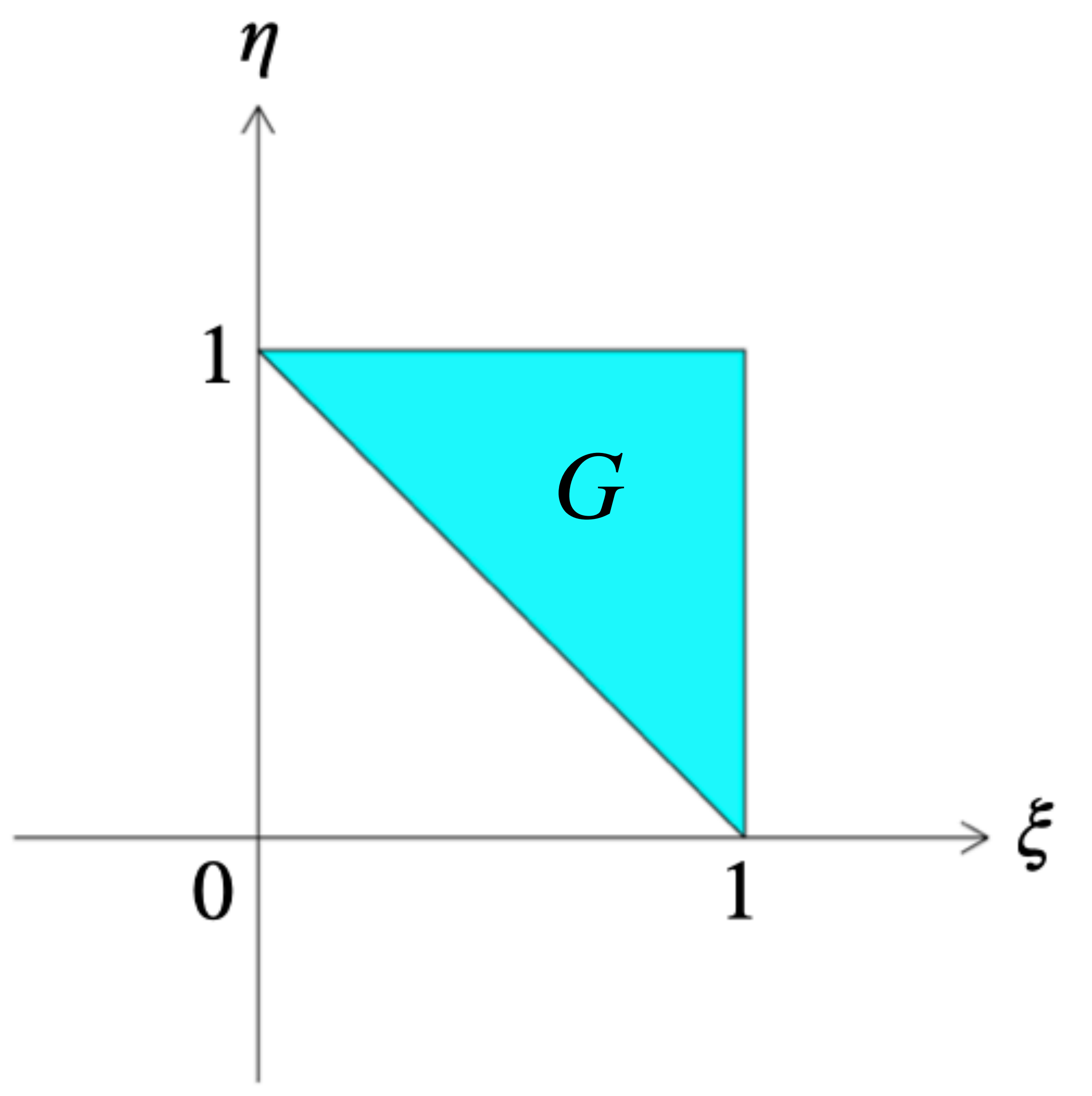
$$+0.15\times\sin\frac{\pi(2+0)}{2}+0.15\times\sin\frac{\pi(2+1)}{2}$$

$$=0.25$$

$$E(\eta)=E\left[g(\xi_1,\xi_2)\right]=\begin{cases}\text{离散型：} & \sum_i\sum_j g(x_i,y_j)\cdot p_{ij}\\ \text{连续型：} & \int_{-\infty}^{\infty} g(x,y)\cdot p_{(\xi,\eta)}(x,y)\,\mathrm{d}x\,\mathrm{d}y\end{cases}$$

七、多维场合

- 例：(ξ,η) 服从以 $(0,1)$，$(1,0)$ 和 $(1,1)$ 为顶点的三角形区域上的均匀分布，求 $\zeta=\xi+\eta$ 的数学期望 $E(\zeta)$.



(ξ,η) 的联合密度： $p(x,y)=\begin{cases}2, & (x,y)\in G\\ 0, & (x,y)\notin G\end{cases}$

$$E(\zeta)=E\left(\xi+\eta\right)=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}(x+y)\,p(x,y)\,\mathrm{d}x\,\mathrm{d}y$$

$$=\iint_G 2\,(x+y)\,\mathrm{d}x\,\mathrm{d}y=2\int_0^1\mathrm{d}y\int_{1-y}^1(x+y)\,\mathrm{d}x$$

$$=\frac{4}{3}$$

八、数学期望的基本性质

- **性质 1** 若 $a \le \xi \le b$，则 $a \le E(\xi) \le b$．特别，若 c 为常数，则 $E(c) = c$.

- **性质 2 (单调性)** 若 $\xi \le \eta$ 几乎处处 (a.e.) 成立，则 $E(\xi) \le E(\eta)$.

- **性质 3 (线性性)** 对任意常数 c_i , $i = 1, 2, \ldots, n$ 以及 b，都有

$$E\left(\sum_{i=1}^{n} c_i\,\xi_i + b\right) = \sum_{i=1}^{n} c_i\,E(\xi_i) + b$$

- **性质 4 (和的期望等于期望的和)** 对任意 n 个随机变量 $\xi_1, \xi_2, \ldots, \xi_n$，非负可测，都有

$$E(\xi_1 + \xi_2 + \cdots + \xi_n) = E(\xi_1) + E(\xi_2) + \cdots + E(\xi_n)$$

- **性质 5 (独立随机变量乘积的期望等于期望的乘积)** 若随机变量 $\xi_1, \xi_2, \ldots, \xi_n$ 相互独立，则有

$$E(\xi_1 \cdot \xi_2 \cdot \cdots \cdot \xi_n) = E(\xi_1) \cdot E(\xi_2) \cdot \cdots \cdot E(\xi_n)$$

八、数学期望的基本性质

- 例：计算正态分布 $\xi \sim N\left(\mu,\ \sigma^2\right)$ 的均值.

$$\xi_0 \sim N(0,\ 1) \implies E\left(\xi_0\right) = \int_{-\infty}^{+\infty} x \cdot \frac{1}{\sqrt{2\pi}} \mathrm{e}^{-\frac{x^2}{2}} \mathrm{d}x = 0$$

$$\xrightarrow{\text{令}} \xi = \mu + \sigma \cdot \xi_0 \implies \xi \sim N\left(\mu,\ \sigma^2\right) \implies E\left(\xi\right) = \mu + \sigma \cdot E\left(\xi_0\right) = \mu$$

八、数学期望的基本性质

- 例：计算二项分布及超几何分布的均值.

N 件产品中有 M 件次品，定义

$$\xi_i = \begin{cases} 1, & \text{第 } i \text{ 次取到次品} \\ 0, & \text{第 } i \text{ 次取到正品} \end{cases}, \qquad i = 1, 2, \dots, n$$

ξ_i	1	0
$P\{\xi_i = k\}$	$\frac{M}{N}$	$1-\frac{M}{N}$

$$\Longrightarrow \quad E(\xi_i) = \frac{M}{N} \triangleq p$$

▶ 有放回抽样时：$\xi_1, \xi_2, \dots, \xi_n$ 相互独立,

$$\xi \triangleq \xi_1 + \xi_2 + \cdots + \xi_n \sim b(n, p) \Longrightarrow E(\xi) = E(\xi_1) + E(\xi_2) + \cdots + E(\xi_n) = np$$

▶ 不放回抽样时：$\xi_1, \xi_2, \dots, \xi_n$ 不相互独立,

$$\xi \triangleq \xi_1 + \xi_2 + \cdots + \xi_n \sim \mathrm{HG}(n, N, M) \Longrightarrow E(\xi) = E(\xi_1) + E(\xi_2) + \cdots + E(\xi_n) = np = n\frac{M}{N}$$

4.2 方差，相关系数，矩

一、方差

二、切比雪夫 (Chebyshev) 不等式

三、相关系数

四、矩

五、条件数学期望

一、方差

- 例：两个射击选手，技术表现如下

射手甲

ξ 命中环数	8	9	10
概率	0.1	0.8	0.1

射手乙

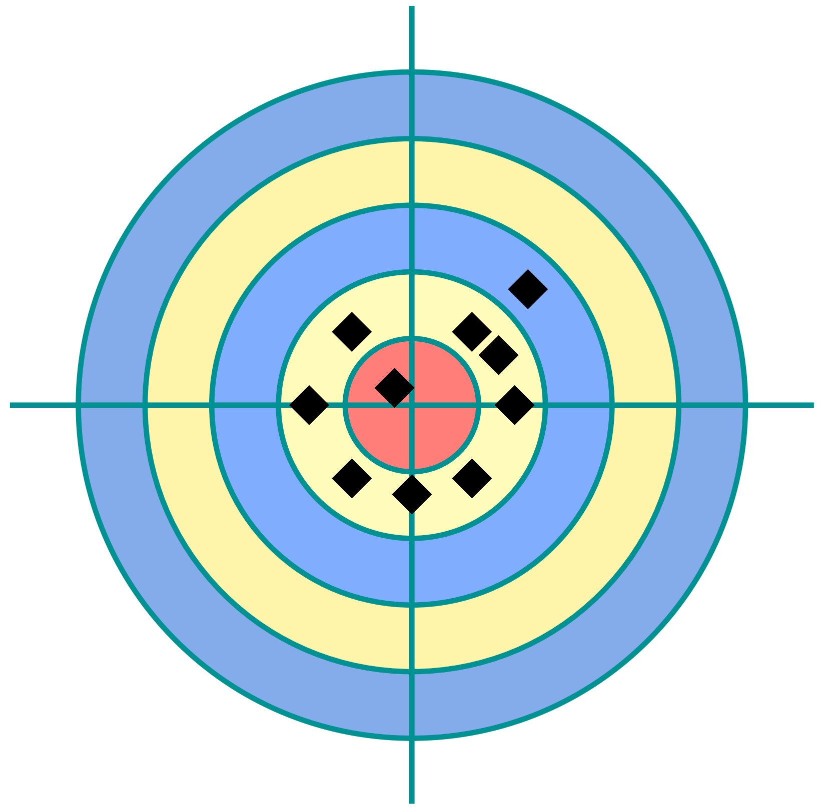
η 命中环数	8	9	10
概率	0.3	0.4	0.3

$\Longrightarrow$ $E(\xi) = 8\times 0.1 + 9\times 0.8 + 10\times 0.1 = 9$

$\Longrightarrow$ $E(\eta) = 8\times 0.3 + 9\times 0.4 + 10\times 0.3 = 9$

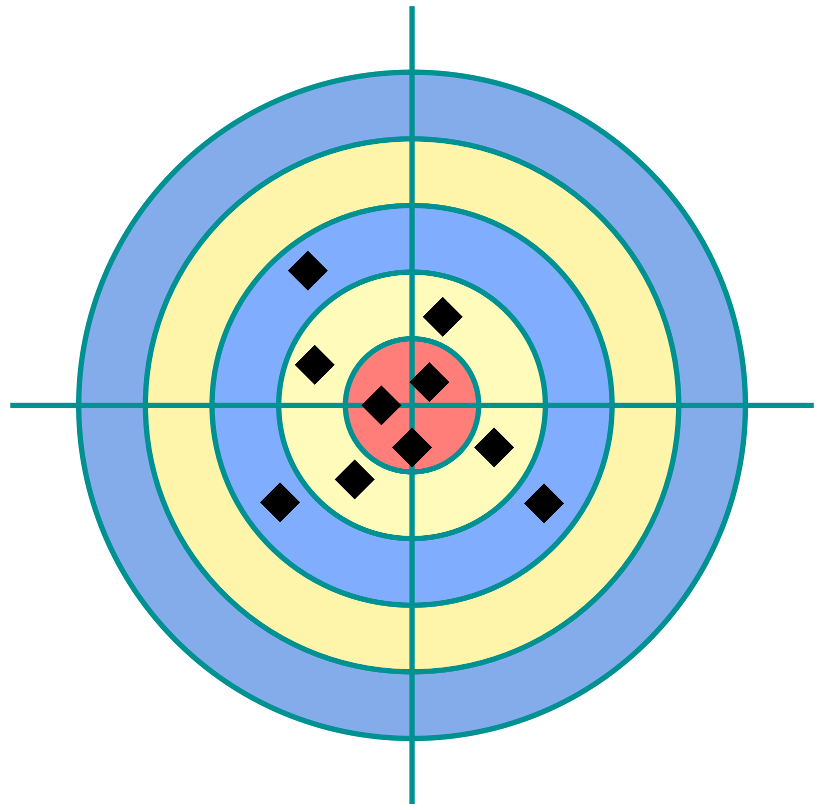
▶ 从均值的角度而言，两个选手的水平相当.

$x_1 = 9$， $x_6 = 9$
$x_2 = 9$， $x_7 = 8$
$x_3 = 9$， $x_8 = 9$
$x_4 = 9$， $x_9 = 9$
$x_5 = 10$， $x_{10} = 9$



$\sum_{i=1}^{n}\left|(x_i-9)^2\right|$

$\sum_{i=1}^{n}\left|(y_i-9)^2\right|$



$y_1 = 8$， $y_6 = 9$
$y_2 = 10$， $y_7 = 9$
$y_3 = 10$， $y_8 = 8$
$y_4 = 9$， $y_9 = 10$
$y_5 = 8$， $y_{10} = 9$

▶ 从稳定性而言，甲选手的命中环数更集中在均值 9 附近，而乙选手则散布度较大.

一、方差

定义：设 ξ 是一个随机变量，若

$$E\left\{\left[\xi-E(\xi)\right]^2\right\}$$

存在，则称之为随机变量 ξ 的方差 (variance)，记为

$$D(\xi)=E\left\{\left[\xi-E(\xi)\right]^2\right\}$$

称其算术平方根 $\sqrt{D(\xi)}$ 为均方差或标准差 (standard deviation).

- 方差、标准差描述了随机变量关于其数学期望的偏离程度 (dispersion)，在概率论与数理统计中有着十分重要的作用. 标准差的量纲与 ξ 的量纲相同.

一、方差

- 常用计算公式：$D(\xi) = E(\xi^2) - \left[E(\xi)\right]^2$

$$D(\xi) = E\left\{\left[\xi - E(\xi)\right]^2\right\} = E\left\{\xi^2 - 2\xi \cdot E(\xi) + \left[E(\xi)\right]^2\right\}$$

$$= E(\xi^2) - 2E(\xi) \cdot E(\xi) + \left[E(\xi)\right]^2$$

$$= E(\xi^2) - \left[E(\xi)\right]^2$$

▶ 推论：(1) $E(\xi^2) = D(\xi) + \left[E(\xi)\right]^2$

(2) $E(\xi^2) \geq \left[E(\xi)\right]^2$

一、方差

- 例：两个射击选手，技术表现如下

射手甲

ξ 命中环数	8	9	10
概率	0.1	0.8	0.1

射手乙

η 命中环数	8	9	10
概率	0.3	0.4	0.3

$$\Longrightarrow \quad E(\xi) = E(\eta) = 9$$

教练欲从中选一人参加比赛，问哪位射手更合适?

$$\Longrightarrow \quad D(\xi) = E(\xi^2) - \left[E(\xi)\right]^2$$

$$= 8^2 \times 0.1 + 9^2 \times 0.8 + 10^2 \times 0.1 - 9^2$$

$$= 0.2$$

$$\Longrightarrow \quad D(\eta) = E(\eta^2) - \left[E(\eta)\right]^2$$

$$= 8^2 \times 0.3 + 9^2 \times 0.4 + 10^2 \times 0.3 - 9^2$$

$$= 0.6$$

结论：选择射手甲参加比赛更合适，因为其方差小，发挥更稳定一些.

一、方差

- 常见分布的方差

① Bernoulli 分布

X	0	1
p_i	$1-p$	p

$\Longrightarrow E(X)=p$

$$\Longrightarrow D(X)=E\left(X^2\right)-\left[E(X)\right]^2=p-p^2=p(1-p)$$

② 二项分布：$X\sim b(n,p)$ $P\{X=k\}=C_n^k p^k(1-p)^{n-k}$ ， $k=0,1,2,\ldots,n$ $\Longrightarrow E(X)=np$

$$\Longrightarrow D(X)=E\left(X^2\right)-\left[E(X)\right]^2=\sum_{k=0}^{n}k^2C_n^kp^k(1-p)^{n-k}-(np)^2=\sum_{k=1}^{n}k^2\frac{n!}{k!(n-k)!}p^k(1-p)^{n-k}-(np)^2$$

$$\xrightarrow{i=k-1,\ m=n-1}=np\sum_{k=1}^{n}k\frac{(n-1)!}{(k-1)!(n-k)!}p^{k-1}(1-p)^{n-k}-n^2p^2=np\sum_{i=0}^{m}(i+1)\frac{m!}{i!(m-i)!}p^i(1-p)^{m-i}-n^2p^2$$

$$\xrightarrow{j=i-1,\ t=m-1}=np+np\sum_{i=1}^{m}i\frac{m!}{i!(m-i)!}p^i(1-p)^{m-i}-n^2p^2=np+np\cdot mp\sum_{j=0}^{t}\frac{t!}{j!(t-j)!}p^j(1-p)^{t-j}-n^2p^2$$

$$=np+(n-1)np^2-n^2p^2=np-np^2=np(1-p)$$

一、方差

- 常见分布的方差

③ Poisson 分布：$X \sim P(\lambda)$ $\quad P\{X=k\}=\dfrac{\lambda^k}{k!}\mathrm{e}^{-\lambda},\quad k=0,1,2,\ldots \quad \Longrightarrow \quad E(X)=\lambda$

$$\Longrightarrow \quad D(X)=E\left(X^2\right)-[E(X)]^2=\sum_{k=0}^{\infty}\left(k^2\frac{\lambda^k}{k!}\mathrm{e}^{-\lambda}\right)-\lambda^2=\lambda\sum_{k=1}^{\infty}\left(\frac{k\lambda^{k-1}}{(k-1)!}\mathrm{e}^{-\lambda}\right)-\lambda^2$$

$$\xrightarrow{i=k-1} \quad =\lambda\sum_{i=0}^{\infty}\left(\frac{(i+1)\lambda^i}{i!}\mathrm{e}^{-\lambda}\right)-\lambda^2=\lambda\underbrace{\sum_{i=0}^{\infty}\left(i\cdot\frac{\lambda^i}{i!}\mathrm{e}^{-\lambda}\right)}_{E(X)=\lambda}+\lambda\underbrace{\sum_{i=0}^{\infty}\left(\frac{\lambda^i}{i!}\mathrm{e}^{-\lambda}\right)}_{=1}-\lambda^2$$

$$=\lambda$$

一、方差

- 常见分布的方差

④ 均匀分布：$X \sim U[a\ ,\ b]$

$$p_X(x) = \begin{cases} \dfrac{1}{b-a}\ , & a \le x \le b \\ 0\ , & \text{其它} \end{cases} \Longrightarrow E(X) = \frac{a+b}{2}$$

$$\Longrightarrow D(X) = E(X^2) - [E(X)]^2 = \int_{-\infty}^{+\infty} x^2 p_X(x)\,\mathrm{d}x - \left(\frac{a+b}{2}\right)^2$$

$$= \int_a^b \frac{x^2}{b-a}\mathrm{d}x - \left(\frac{a+b}{2}\right)^2 = \frac{x^3}{3(b-a)}\bigg|_a^b - \left(\frac{a+b}{2}\right)^2$$

$$= \frac{a^2+ab+b^2}{3} - \frac{(a+b)^2}{4}$$

$$= \frac{(b-a)^2}{12}$$

一、方差

- 常见分布的方差

⑤ 指数分布：$X \sim \mathrm{Exp}(\lambda)$

$$p_X(x) = \begin{cases} \lambda \mathrm{e}^{-\lambda x}, & x > 0 \\ 0, & x \le 0 \end{cases} \Longrightarrow E(X) = \frac{1}{\lambda}$$

$$\Longrightarrow D(X) = E(X^2) - [E(X)]^2 = \int_{-\infty}^{+\infty} x^2 \cdot p_X(x)\,\mathrm{d}x - \left(\frac{1}{\lambda}\right)^2$$

$$= \int_0^{+\infty} x^2 \cdot \lambda \mathrm{e}^{-\lambda x}\,\mathrm{d}x - \left(\frac{1}{\lambda}\right)^2 = -\int_0^{+\infty} x^2\,\mathrm{d}\left(\mathrm{e}^{-\lambda x}\right) - \left(\frac{1}{\lambda}\right)^2$$

$$= -\left(x^2 \mathrm{e}^{-\lambda x}\right)\Big|_0^{\infty} + \int_0^{\infty} \mathrm{e}^{-\lambda x}\,\mathrm{d}\left(x^2\right) - \frac{1}{\lambda^2} = 2\int_0^{\infty} x\mathrm{e}^{-\lambda x}\,\mathrm{d}x - \frac{1}{\lambda^2}$$

$$= -\frac{2}{\lambda}\int_0^{\infty} x\,\mathrm{d}\left(\mathrm{e}^{-\lambda x}\right) - \frac{1}{\lambda^2} = \frac{2}{\lambda^2} - \frac{1}{\lambda^2} = \frac{1}{\lambda^2}$$

一、方差

- 常见分布的方差

⑥ 正态分布：$X \sim N\left(\mu\ ,\ \sigma^2\right)$ $\quad p_X(x) = \dfrac{1}{\sqrt{2\pi}\ \sigma}\mathrm{e}^{-\frac{(x-\mu)^2}{2\sigma^2}}\ ,\quad -\infty < x < +\infty \quad \Longrightarrow \quad E(X) = \mu$

$$\Longrightarrow \quad D(X) = E\left\{\left[X - E(X)\right]^2\right\} = \int_{-\infty}^{+\infty} (x-\mu)^2 \cdot p_X(x)\,\mathrm{d}x$$

$$= \frac{1}{\sqrt{2\pi}\ \sigma}\int_{-\infty}^{+\infty}(x-\mu)^2 \cdot \mathrm{e}^{-\frac{(x-\mu)^2}{2\sigma^2}}\mathrm{d}x = \frac{\sigma^2}{\sqrt{2\pi}}\int_{-\infty}^{+\infty} t^2 \cdot \mathrm{e}^{-\frac{t^2}{2}}\mathrm{d}t \quad \longleftarrow \quad t = \frac{x-\mu}{\sigma}\ ,\ \mathrm{d}x = \sigma\mathrm{d}t$$

$$= -\frac{\sigma^2}{\sqrt{2\pi}}\int_{-\infty}^{+\infty} t\ \mathrm{d}\left(\mathrm{e}^{-\frac{t^2}{2}}\right) = -\frac{\sigma^2}{\sqrt{2\pi}}\underbrace{\left(t\mathrm{e}^{-\frac{t^2}{2}}\right)\Bigg|_{-\infty}^{+\infty}}_{=0} + \sigma^2 \underbrace{\int_{-\infty}^{+\infty}\frac{1}{\sqrt{2\pi}}\mathrm{e}^{-\frac{t^2}{2}}\mathrm{d}t}_{=1}$$

$$= \sigma^2$$

▶ 特别：$X \sim N(0\ ,\ 1) \Longrightarrow \quad E(X) = 0\ ,\ D(X) = 1.$

- 常见分布的数学期望与方差：

分布	符号表示	数学期望	方差
二项分布	$\xi \sim b(n, p)$	$E(\xi) = np$	$D(\xi) = np(1-p)$
Bernoulli 分布	$\xi \sim b(1, p)$	$E(\xi) = p$	$D(\xi) = p(1-p)$
Poisson 分布	$\xi \sim P(\lambda)$	$E(\xi) = \lambda$	$D(\xi) = \lambda$
均匀分布	$\xi \sim U[a, b]$	$E(\xi) = \dfrac{a+b}{2}$	$D(\xi) = \dfrac{(b-a)^2}{12}$
Gamma 分布	$X \sim \Gamma(r, \lambda)$	$E(\xi) = \dfrac{r}{\lambda}$	$D(\xi) = \dfrac{r}{\lambda^2}$
指数分布	$\xi \sim \mathrm{Exp}(\lambda)$	$E(\xi) = \dfrac{1}{\lambda}$	$D(\xi) = \dfrac{1}{\lambda^2}$
χ^2 分布	$X \sim \chi_n^2$	$E(\xi) = n$	$D(\xi) = 2n$
正态分布	$\xi \sim N(\mu, \sigma^2)$	$E(\xi) = \mu$	$D(\xi) = \sigma^2$

一、方差

- 方差的基本性质

① 设 C 是常数，则 $D(C)=0$.

② 设 ξ 是随机变量，c 为常数，则 $D(\xi+c)=D(\xi)$.

③ 设 ξ 是随机变量，c 为常数，则 $D(c\xi)=c^2D(\xi)$.

④ 若 $c\neq E(\xi)$，则 $D(\xi)<E\left[(\xi-c)^2\right]$.

期望的一个重要的极值性质. $E\left[(\xi-c)^2\right]$ 当 $c=E(\xi)$ 时达到极小，这说明了在 $D(\xi)$ 定义中取 $c=E(\xi)$ 的合理性.

$$D(\xi)=E\left\{\left[\xi-E(\xi)\right]^2\right\}=E\left\{\left[\xi-c-\left(E(\xi)-c\right)\right]^2\right\}$$

$$=E\left\{(\xi-c)^2-2(\xi-c)\left(E(\xi)-c\right)+\left(E(\xi)-c\right)^2\right\}$$

$$=E\left[(\xi-c)^2\right]-\left[E(\xi)-c\right]^2<E\left[(\xi-c)^2\right]$$

一、方差

- 方差的基本性质

① 设 C 是常数，则 $D(C)=0.$

② 设 ξ 是随机变量，c 为常数，则 $D(\xi+c)=D(\xi).$

③ 设 ξ 是随机变量，c 为常数，则 $D(c\xi)=c^2D(\xi).$

④ 若 $c\neq E(\xi)$，则 $D(\xi)<E\left[(\xi-c)^2\right].$

⑤ $D(\xi)=0 \iff P\{\xi=C\}=1.$

⑥ 若 ξ 与 η 独立，则 $D(\xi\pm\eta)=D(\xi)+D(\eta).$

$$D(\xi+\eta)=E\left[(\xi+\eta)^2\right]-\left[E(\xi+\eta)\right]^2=E\left[\xi^2+2\xi\eta+\eta^2\right]-\left[E(\xi)+E(\eta)\right]^2$$

$$=E(\xi^2)-\left[E(\xi)\right]^2+2E(\xi\eta)-2E(\xi)\cdot E(\eta)+E(\eta^2)-\left[E(\eta)\right]^2=D(\xi)+D(\eta)$$

一、方差

- 方差的基本性质

① 设 C 是常数，则 $E(C)=0$.

② 设 ξ 是随机变量，c 为常数，则 $D(\xi+c)=D(\xi)$.

③ 设 ξ 是随机变量，c 为常数，则 $D(c\xi)=c^2D(\xi)$.

④ 若 $c\neq E(\xi)$，则 $D(\xi)<E\left[(\xi-c)^2\right]$.

⑤ $D(\xi)=0 \iff P\{\xi=C\}=1$.

⑥ 若 ξ 与 η 独立，则 $D(\xi\pm\eta)=D(\xi)+D(\eta)$.

▶ 推广：若 $\xi_1, \xi_2, \ldots, \xi_n$ 相互独立，则

$$D\left(c_1\xi_1+c_2\xi_2+\cdots+c_n\xi_n\right)=c_1^2D\left(\xi_1\right)+c_2^2D\left(\xi_2\right)+\cdots+c_n^2D\left(\xi_n\right)$$

一、方差

- **例**：设 $X \sim b(n, p)$，求 $D(X)$.

X_i	0	1
p_i	$1-p$	p

$, i = 1, 2, \ldots, n$ 且相互独立

$E(X_i) = p\ ,\quad D(X_i) = p(1-p)$

$$X \triangleq X_1 + X_2 + \cdots + X_n \ \sim\ b(n, p)$$

$$\Longrightarrow \quad D(X) = D(X_1 + X_2 + \cdots + X_n) = D(X_1) + D(X_2) + \cdots + D(X_n) = np(1-p)$$

- **随机变量的标准化**

设随机变量 ξ 的均值 $E(\xi)$ 与方差 $D(\xi)$ 都存在，且 $D(\xi) > 0$，称

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

为随机变量 ξ 的标准化随机变量.

▶ 标准化的目的：通过线性变换将随机变量的均值变换为 0、方差变换为 1.

二、切比雪夫 (Chebyshev) 不等式

Chebyshev 不等式：设 ξ 为具有有限均值 $E(\xi)$ 与有限方差 $D(\xi)$ 的随机变量，则 $\forall\, \varepsilon > 0$ 有

$$P\left\{\left|\xi - E(\xi)\right| \geq \varepsilon\right\} \leq \frac{D(\xi)}{\varepsilon^2}$$

$F_\xi(x)$ 是 ξ 的分布函数

$$D(\xi) = \int_{-\infty}^{+\infty} \left[x - E(\xi)\right]^2 \mathrm{d}F_\xi(x)$$

$$= \int_{-\infty}^{E(\xi)-\varepsilon} \left[x - E(\xi)\right]^2 \mathrm{d}F_\xi(x) + \underbrace{\int_{E(\xi)-\varepsilon}^{E(\xi)+\varepsilon} \left[x - E(\xi)\right]^2 \mathrm{d}F_\xi(x)}_{\geq 0} + \int_{E(\xi)+\varepsilon}^{+\infty} \left[x - E(\xi)\right]^2 \mathrm{d}F_\xi(x)$$

$$\geq \int_{\left|x-E(\xi)\right|\geq\varepsilon} \left[x - E(\xi)\right]^2 \mathrm{d}F_\xi(x) \geq \int_{\left|x-E(\xi)\right|\geq\varepsilon} \varepsilon^2 \,\mathrm{d}F_\xi(x) = \varepsilon^2 P\left\{\left|\xi - E(\xi)\right| \geq \varepsilon\right\}$$

二、切比雪夫 (Chebyshev) 不等式

Chebyshev 不等式：设 ξ 为具有有限均值 $E(\xi)$ 与有限方差 $D(\xi)$ 的随机变量，则 $\forall\ \varepsilon > 0$ 有

$$P\left\{\left|\xi - E(\xi)\right| \geq \varepsilon\right\} \leq \frac{D(\xi)}{\varepsilon^2}$$

- Chebyshev 不等式的等价形式

$$P\left\{\left|\xi - E(\xi)\right| < \varepsilon\right\} \geq 1 - \frac{D(\xi)}{\varepsilon^2} \quad 或 \quad P\left\{\left|\frac{\xi - E(\xi)}{\sqrt{D(\xi)}}\right| \geq \delta\right\} \leq \frac{1}{\delta^2}$$

- Chebyshev 不等式的意义：利用随机变量 ξ 的数学期望以及方差对 ξ 的概率分布进行估计.

它断言不管 ξ 的分布是什么，ξ 落在 $\left(E(\xi) - \varepsilon\ ,\ E(\xi) + \varepsilon\right)$ 中的概率均不小于 $1 - \dfrac{D(\xi)}{\varepsilon^2}$.

二、切比雪夫 (Chebyshev) 不等式

Chebyshev 不等式：设 ξ 为具有有限均值 $E(\xi)$ 与有限方差 $D(\xi)$ 的随机变量，则 $\forall\ \varepsilon > 0$ 有

$$P\left\{\left|\xi - E(\xi)\right| \geq \varepsilon\right\} \leq \frac{D(\xi)}{\varepsilon^2}$$

- 从 Chebyshev 不等式还可以看出，当方差 $D(\xi)$ 愈小时，事件

$$\left\{\left|\xi - E(\xi)\right| \geq \varepsilon\right\}$$

的概率也愈小，从这里可以看出：方差是描述随机变量与其期望值离散程度的一个量.

- 特别：若 $D(\xi) = 0$，则 $\forall\ \varepsilon > 0$，恒有

$$P\left\{\left|\xi - E(\xi)\right| \geq \varepsilon\right\} = 0 \implies P\left\{\xi \neq E(\xi)\right\} = 0 \implies P\left\{\xi = E(\xi)\right\} = 1$$

所以：方差为零的随机变量是常数.

二、切比雪夫 (Chebyshev) 不等式

Chebyshev 不等式：设 ξ 为具有有限均值 $E(\xi)$ 与有限方差 $D(\xi)$ 的随机变量，则 $\forall\, \varepsilon > 0$ 有

$$P\left\{\left|\xi - E(\xi)\right| \geq \varepsilon\right\} \leq \frac{D(\xi)}{\varepsilon^2}$$

- 3σ 法则：$\sigma^2 \triangleq D(\xi)$，$\sigma = \sqrt{D(\xi)}$，$\mu \triangleq E(\xi)$

$$P\left\{\left|\xi - \mu\right| < \varepsilon\right\} \geq 1 - \frac{\sigma^2}{\varepsilon^2} \Longrightarrow \begin{cases} P\left\{\left|\xi-\mu\right| \leq \sigma\right\} \geq 0\ , & \varepsilon = \sigma \\ P\left\{\left|\xi-\mu\right| \leq 2\sigma\right\} \geq 0.75\ , & \varepsilon = 2\sigma \\ P\left\{\left|\xi-\mu\right| \leq 3\sigma\right\} \geq 0.8889\ , & \varepsilon = 3\sigma \end{cases}$$

与正态分布的比较： $X \sim N(\mu\ ,\sigma^2)$ $\Longrightarrow \begin{cases} P\left\{\left|\xi-\mu\right| \leq \sigma\right\} = 0.6826 \\ P\left\{\left|\xi-\mu\right| \leq 2\sigma\right\} = 0.9544 \\ P\left\{\left|\xi-\mu\right| \leq 3\sigma\right\} = 0.9974 \end{cases}$