概率论与数理统计

[1]随机事件及其概率

随机试验（E）：

1. **重复性**：在相同条件下，实验可重复进行
2. **明确性**：实验所有可能的结果先均已知
3. **随机性**：每次实验的具体结果在实验前无法预知

样本点（ω）：随机试验每一个可能出现的结果

样本空间（Ω）:随机试验所有**样本点的全体**

随机事件：具有某种特征的**样本点的集合**

- 基本事件：由一个样本点构成的**单点**
- 必然事件（Ω）：必然发生的事件
- 不可能事件（$\varnothing$）：不可能发生的事件

概率的性质：

- 非负性：$0 \leq P(A) \leq 1$
- 规范性：$P(\Omega)=1,\ P(\varnothing)=0$
- 可加性：如果事件 $A_1,\ A_2,\ \cdots,\ A_n$ **两两互不相容**，则 $P(A_1 \cup A_2 \cup \cdots \cup A_n)=P(A_1)+P(A_2)+\cdots+P(A_n)$
- 差事件概率计算公式：$P(A-B)=P(A)-P(AB)$
- 并事件概率计算公式：

$$P(A_1 \cup A_2 \cup \cdots \cup A_n)=\sum_{i=1}^{n} P(A_i)-\sum_{1 \leq i<j \leq n} P(A_i A_j)+\sum_{1 \leq i<j<k \leq n} P(A_i A_j A_k)+\cdots+(-1)^{n-1} P(A_1 A_2 \cdots A_n)$$

条件概率：已知事件 A 发生的条件下，事件 B 发生的概率 $P(B \mid A)=\frac{P(AB)}{P(B)}$

乘法公式：$P(A_1 A_2 \cdots A_n)=P(A_1)P(A_2|A_1)P(A_3|A_1A_2)\cdots P(A_n|A_1A_2\cdots A_{n-1})$

完备事件组（完备组）：事件组 $A_!,\ A_2,\ \cdots,\ A_i,\ \cdots$ 两两互不相容，且 $A_1 \cup A_2 \cup \cdots \cup A_i \cup \cdots=\Omega$

全概率公式：

$$P(B)=\sum P(A_i)\,P(B|A_i)$$
$$P(C|B)=\sum P(A_i\,|B)\,P(C\,|A_iB)$$

贝叶斯公式：完备事件组 A_i

$$P(A_i\,|B)=\frac{P(A_iB)}{P(B)}=\frac{P(B\,|A_i)\,P(A_i)}{\sum P(A_i)\,P(B\,|A_i)}$$

- 先验概率：$P(A_i)$
- 后验概率：$P(A_i\,|B)$

相互独立：

- $P(AB) = P(A)\,P(B)$
- $P(B) = P(B\,|\,A) = P(B\,|\,\bar{A})$

两两独立： $n\,(n \geq 2)$ 个随机事件 $A_1,\, A_2,\, \cdots,\, A_n$ ，对其中任意 $k\,(k = 2,\, 3,\, \cdots,\, n)$ 个随机事件，均有

$$P(A_{i_1}A_{i_2}\cdots A_{i_n}) = P(A_{i_1})\,P(A_{i_2})\cdots P(A_{i_k})$$

古典概型： 随机试验只有**有限个**样本点且各**基本事件**出现的**概率相等**

几何概型： 随机试验的样本空间为**一几何区域**且各**基本事件**出现的**机会均等**

伯努利概型： 随机试验 E 重复进行 n 次，若每次实验

- **样本空间**相同
- 结果**相互独立**
- 仅考虑两种结果 A 或 $\bar{A}$

[2]一维随机变量及其分布

随机变量： 与没一个样本点 ω 唯一确定对应的**实数** X

分布函数： 设 X 为一随机变量，对于任意实数 x ，称函数 $P(X \leq x)$ 为 X 的分布函数，记为 $F(X)$.
即 $F(x) = P(X \leq x\,|\,-\infty < x < +\infty)$

- $P\{< X \leq b\} = F(b) - F(a)$
- $P\{X = x_0\} = F(x_0) - F(x_0 - 0)$
- $0 \leq F(x) \leq 1$
- $F(-\infty) = 0,\ F(+\infty) = 1$
- $F(x)$ 单调不减，且处处**右连续**，图像**左闭右开**

分布律：

- 离散型随机变量：X 的取值**有限个**或**无穷多个**，记为

X	x_1	x_2	$\cdots$	x_i	$\cdots$
P	p_1	p_2	$\cdots$	p_i	$\cdots$

或 $X \sim \begin{pmatrix} x_1 & x_2 & \cdots & x_i & \cdots \\ p_1 & p_2 & \cdots & p_i & \cdots \end{pmatrix}$

- 连续型随机变量：对**任意** x 有 $F(x) = \int_{\infty}^{x} f(x)\mathrm{d}x$
  - $F(x)$ 连续，最多有有限的可导点
  - 在 $F(x)$ 的可导处，$f(x) = F'(x)$
  - 在 $F(x)$ 的不可导处，$f(x)$为**任意**非负实数
  - $P\{X = x_0\} = 0$
  - 不可能事件未必是空集，全概率事件未必是全集
  - 连续型变量所在某区间的概率与区间端点是否取到无关
  - 在有限点改变 $f(x)$ 的取值，不影响 X 的整体分布，即 $f(x)$ 的表达式不唯一
- 性质：$\sum p_i = \int_{-\infty}^{-\infty} f(x)\mathrm{d}x = 1$

离散型随机变量分布：

- **两点分布：** $X \sim B(1,\, p)$

$$P\{X = k\} = p^k(1-p)^{1-k}$$

- **二项分布：** $X \sim B(n,\, p)$

$$P\{X = k\} = \mathrm{C}_n^k p^k(1-p)^{n-k}$$

-

- **泊松分布：** $X \sim P(\lambda)$

$$P\{X=k\}=\frac{\lambda^k}{k!}e^{-\lambda}$$

- $\lambda=\lim_{n\to\infty} np$
- n 充分大，p 充分小，λ 较适中

$$\mathrm{C}_n^k p^k(1-p)^{n-k}\approx\frac{\lambda^k}{k!}e^{-\lambda}$$

- **几何分布**：$X\sim G(p)$

$$P\{X=k\}=(1-p)^{k-1}p$$

  事件首次发生时的概率

- **超几何分布**：

$$P\{X=k\}=\frac{\mathrm{C}_{N-M}^{n-k}\mathrm{C}_M^k}{\mathrm{C}_N^n}$$

  从 N 个球里取 n 个，其中有 k 个球的颜色是 M 的概率

连续型随机变量分布：

- **均匀分布**：$X\sim U[a,\, b]$

$$f(x)=\begin{cases}\frac{1}{b-a} & a\le x\le b\\ 0 & \text{其他}\end{cases}$$

$$F(x)=\begin{cases}0 & x<a\\ \frac{x-a}{b-a} & a\le a\le b\\ 1 & b\le x\end{cases}$$

- **指数分布**：$X\sim E(\lambda)$

$$f(x)=\begin{cases}0 & x\le 0\\ \lambda e^{-\lambda x} & x>0\end{cases}$$

$$F(x)=\begin{cases}0 & x<0\\ 1-e^{-\lambda x} & x\ge 0\end{cases}$$

  - 当 $s>0,\, t>0$ 时

$$P\{\ X>s+t\mid X>s\ \}=P\{\ X>t\ \}$$

  - 指数分布与泊松分布的联系：
    - 泊松分布指事件 t 内事件发生次数 $N(t)$ 的概率
    - 指数分布指随机事件发生的时间间隔 $T(x)$ 的概率
    - $\{X>t\}\iff\{N(t)=0\}$
- **正态分布**：$X\sim N(\mu,\, \sigma^2)$

$$f(x)=\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

  - 对称轴位置：$x=\mu$
  - 离散程度：σ 大图形矮胖，σ 小图形高瘦
  - 最大值 $f(\mu)=\frac{1}{\sqrt{2\pi}\sigma}$
  - $(-\infty,\, \mu)$ 内单调递增，$(\mu,\, +\infty)$ 内单调递减
  - 密度函数的**水平渐近线**：x 轴
  - 密度函数的**拐点**（凹凸性改变）：$x=\mu\pm\sigma$
  - 标准正态分布：$X\sim N(0,\, 1)$

$$\varphi(x)=\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}$$
$$\Phi(x)=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x}e^{-\frac{t^2}{2}}\,\mathrm{d}x$$
$$F(x)=\Phi(\frac{x-\mu}{\sigma})$$

- 3σ 原则：X 的取值基本上都落在区间 $(\mu-3\sigma,\,u+3\sigma)$ 内

$$P\{|x-\mu|<\sigma\}=0.6826$$
$$P\{|x-\mu|<2\sigma\}=0.9544$$
$$P\{|x-\mu|<3\sigma\}=0.9974$$

- 令 $Y=\frac{X-\mu}{\sigma}$，则 $Y\sim N(0,\,1)$

$Y=g(X)$的分布：

若

- 连续型随机变量 X 的密度函数为 $f_Y(x)$
- $y=g(x)$ 是单调函数
- 其反函数 $x=h(y)$ 具有**连续一阶导数**

则

- Y 是连续型变量随机变量
- $f_Y(y)=\begin{cases} f_X(h(y))\,|h'(y)|, & \alpha<y<\beta \\ 0, & 其他 \end{cases}$
- $\alpha=\min\{g(-\infty),\,g(+\infty)\},\,\beta=\max\{g(-\infty),\,g(+\infty)\}$

推论

- $Y=aX+b\sim N\left(a\mu+b,\,a^2\sigma^2\right)$，$a\neq 0$
- $Y=\frac{X-\mu}{\sigma}\sim N(0,\,1)$

重要公式：

$$\int_0^{+\infty}e^{-x^2}=\frac{\sqrt{\pi}}{2}$$
$$\sum_{n=0}^{\infty}\frac{x^n}{n!}=e^x$$

[3]多维随机变量及其分布

联合分布函数：设 $(X,\,Y)$ 是二维随机变量，称

$$F(x,\,y)=P\{\,X\leq x,\,Y\leq y\,\},\quad -\infty<x<+\infty,\quad -\infty<y<+\infty$$

为 $(X,\,Y)$ 的分布函数

- $F(x,\,y)$ 分别关于变量 x 和 y 处处右连续
- $P\{\,x_1\leq X\leq x_2,\,y_1\leq Y\leq y_2\,\}=F(x_2,\,y_2)-F(x_1,\,y_1)-F(x_2,\,y_1)+F(x_1,\,y_1)$

二维随机变量：设二维随机变量 $(X,\,Y)$ 的分布函数为 $F(x,\,y)$，密度函数为 $f(x,\,y)$，在 $f(x,\,y)$ 的连续点处有 $\frac{\partial^2F(x,\,y)}{\partial x\partial y}=f(x,\,y)$

常见二维随机变量分布：

- **均匀分布**：$(X,\,Y)\sim U(D)$

$$f(x,\,y)=\begin{cases}\frac{1}{S_D} & (x,\,y)\in D\\ 0 & (x,\,y)\notin D\end{cases}$$

- **正态分布：** $(X,\,Y)\sim N(\mu_1,\,\mu_2,\,\sigma_1^2,\,\sigma_2^2,\,\rho)$

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

  - $\mu_1,\,\mu_2\in(-\infty,+\infty)$
  - $\sigma_1,\,\sigma_2>0$
  - $\rho\in(-1,\,1)$
  - $(X,\,Y)\sim N(\mu_1,\,\mu_2,\,\sigma_1^2,\,\sigma_2^2,\,\rho)\implies X\sim N(\mu_1,\,\sigma_1^2),\;Y\sim N(\mu_2,\,\sigma_2^2)$

边缘分布函数：

$$F_X(x)=F(x,\,+\infty)$$
$$F_Y(y)=F(+\infty,\,y)$$

- 边缘分布律不能唯一确定联合分布律
-

Y	X					$p_{\cdot j}$
	x_1	x_2	$\cdots$	x_i	$\cdots$	
y_1	p_{11}	p_{21}	$\cdots$	p_{i1}	$\cdots$	$p_{\cdot 1}$
y_2	p_{12}	p_{22}	$\cdots$	p_{i2}	$\cdots$	$p_{\cdot 2}$
$\vdots$			$\ddots$			$\vdots$
y_i	p_{1i}	p_{2i}	$\cdots$	p_{ii}	$\cdots$	$p_{\cdot i}$
$\vdots$			$\ddots$			$\vdots$
$p_{i\cdot}$	$p_{1\cdot}$	$p_{2\cdot}$	$\cdots$	$p_{i\cdot}$	$\cdots$	1

边缘密度函数：

$$f_X(x)=\int_{\infty}^{+\infty}f(x,\,y)\,\mathrm{d}x$$
$$f_Y(y)=\int_{\infty}^{+\infty}f(x,\,y)\,\mathrm{d}y$$

条件分布函数：

$$F_{X\,|\,Y}(x\,|\,y)=P\{\,X\le x\,|\,Y=y\,\}$$
$$F_{Y\,|\,X}(y\,|\,x)=P\{\,Y\le y\,|\,X=x\,\}$$

条件密度函数：

$$f_{X\,|\,Y}(x\,|\,y)=\frac{f(x,\,y)}{f_Y(y)}$$
$$f_{Y\,|\,X}(y\,|\,x)=\frac{f(x,\,y)}{f_X(x)}$$

相互独立： $(X,\,Y)$ 相互独立的充要条件为

- 离散型：$P_{ij}=p_{i\cdot}p_{\cdot j}$，$i=1,\,2,\,\cdots$，$j=1,\,2,\,\cdots$
- 连续型：对平面上几乎所有点有 $f(x,\,y)=f_X(x)f_Y(y)$
- 正态分布：$\rho=0$
- 推论：X 和 Y 相互独立
  - $X\sim P(\lambda_1)$，$Y\sim P(\lambda_2)\implies X+Y\sim P(\lambda_1+\lambda_2)$

- $X \sim B(n, p)$，$Y \sim B(m, p) \implies X + Y \sim (n + m, p)$

$Z = g(X, Y)$ 的分布：

- $F_Z(z) = P\{\ Z \le z\ \} = \iint_{g(x,y) \le z} f(x, y)\,\mathrm{d}x\mathrm{d}y$
- $f_Z(z) = \begin{cases} \int_{-\infty}^{+\infty} f(x, z - x)\,\mathrm{d}x = \int_{-\infty}^{+\infty} f(z - y, y)\,\mathrm{d}y & Z = X + Y \\ \int_{-\infty}^{\infty} f(yz, y)|y|\,\mathrm{d}y & Z = \frac{X}{Y} \end{cases}$
- $M = \max\{X, Y\}$, $N = \min\{X, Y\}$
  $F_M = F_X(x)F_Y(y)$, $F_N(x) = 1 - [1 - F_X(x)][1 - F_Y(x)]$

[4]随机变量的数字特征

随机变量的数学期望：

- 设随机变量 X 的分布律为

$$P\{\ X = x_i\ \} = p_i,\quad i = 1, 2, \cdots,\quad \text{或} \; X \sim \begin{pmatrix} x_1 & x_2 & \cdots & x_i & \cdots \\ p_1 & p_2 & \cdots & p_i & \cdots \end{pmatrix}$$

  如果无穷级数 $\sum_{i=1}^{\infty} x_i p_i$ 绝对收敛，就称之为 X 的数学期望值或者均值，记为 EX 或 $E(X)$，即

$$EX = \sum_{i=1}^{\infty} x_i p_i$$

- 设随机变量 X 的密度函数为 $f(x)$，如果反常积分 $\int_{-\infty}^{+\infty} x f(x)\,\mathrm{d}x$ 绝对收敛，则

$$EX = \int_{-\infty}^{+\infty} x f(x)\,\mathrm{d}x$$

随机变量函数的数学期望：

- 设离散型随机变量 $X \sim \begin{pmatrix} x_1 & x_2 & \cdots & x_i & \cdots \\ p_1 & p_2 & \cdots & p_i & \cdots \end{pmatrix}$，$Y = g(x)$，且无穷级数 $\sum_{i=1}^{\infty} g(x_i)p_i$ 绝对收敛，则

$$EY = E[\,g(X)\,] = \sum_{i=1}^{\infty} g(x_i)p_i$$

- 设连续型随机变量 X 的密度函数为 $f(x)$，$Y = g(x)$，且广义积分 $\int_{-\infty}^{+\infty} g(x)f(x)\,\mathrm{d}x$ 绝对收敛，则

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

- 设二维离散型随机变量 (X, Y) 的分布律为 $P = \{\ X = x_i, Y = y_i\ \} = p_{ij}$，$i = 1, 2, \cdots$，$j = 1, 2, \cdots$，$Z = g(X, Y)$，且 $\sum_{i=1}^{\infty}\sum_{j=1}^{\infty} g(x_i, y_i)p_{ij}$ 绝对收敛，则

$$EZ = E[\,g(X, Y)\,] = \sum_{i=1}^{\infty}\sum_{j=1}^{\infty} g(x_i, y_i)p_{ij}$$

- 设二维随机变量 (X, Y) 的密度函数为 $f(x, y)$，$Z = g(X, Y)$，且 $\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} g(x, y)f(x, y)\,\mathrm{d}x\mathrm{d}y$ 绝对收敛，则

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

- $E[\,g(X)\,] = \sum_{i=1}^{\infty}\sum_{j=1}^{\infty} g(x_i)p_{ij} = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} g(x)f(x, y)\,\mathrm{d}x\mathrm{d}y$

数学期望的性质：

- $Ec = c$
- $E(kX) = kE(X)$
- $E(X \pm Y) = EX \pm EY$

- 若随机变量 X 和 Y 相互独立，则 $EXY = EXEY$，**反之不成立**
- 若随机变量 $X \geq a$ （或 $X \leq a$），则 $EX \geq a$ （或 $X \leq a$）
- $[E(XY)]^2 \leq E(X^2)E(Y^2)$

方差：设 X 为随机变量，如果 $E\left[(X - EX)^2\right]$ 存在，就称之为 X 的方差，记为 DX 或 $D(X)$，即

$$DX = E\left[(X - EX)^2\right] = E(X^2) - (EX)^2$$

并称 $\sqrt{DX}$ 为 X 的**标准差**

- $DX = \sum_{i=1}^{\infty}(x_i - EX)^2 p_i = \sum_{i=1}^{\infty} X_i^2 p_i - \left(\sum_{i=1}^{\infty} x_i p_i\right)^2$
- $DX = \int_{-\infty}^{+\infty}(x - EX)^2 f(x)\,\mathrm{d}x = \int_{-\infty}^{+\infty} x^2 f(x)\,\mathrm{d}x - \left(\int_{-\infty}^{+\infty} x f(x)\,\mathrm{d}x\right)^2$

方差的性质：

- $Dc = 0$
- $DX \geq 0$
- $DX = 0 \Longleftrightarrow P\{X = EX = 1\}$
- $D(kX + c) = K^2 DX$
- 若随机变量 X 和 Y 相互独立，则 $D(X \pm Y) = DX + DY$
- **标准化随机变量**：
  - $X^* = \frac{X - EX}{\sqrt{DX}}$
  - $EX^* = 0$
  - $DX^* = 1$
- 若随机变量 $X_1, X_2, \cdots, X_n$ 相互独立，且 $EX_i = \mu$，$DX_i = \sigma^2$，$i = 1, 2, \cdots, n$，记 $\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$
  - $E\overline{X} = \mu$
  - $D\overline{X} = \frac{\sigma^2}{n}$

常见分布的数学期望和方差：

序号	分布	分布律或密度函数	EX	DX
1	$0-1$ 两点分布 $X \sim B(1, p)$	$P\{X = k\} = p^k(1-p)^{1-k}$ $k = 0, 1$	p	$1 - p$
2	二项分布 $X \sim B(n, p)$	$P\{X = k\} = C_n^k p^k (1-p)^{n-k}$ $k = 0, 1, \cdots, n$	np	$np(1-p)$
3	泊松分布 $x \sim P(\lambda)$	$P\{X = k\} = \frac{\lambda^k}{k!} e^{-\lambda}$ $k = 0, 1, 2, \cdots$	λ	λ
4	几何分布 $X \sim G(p)$	$P\{X = k\} = (1-p)^{k-1} p$ $k = 1, 2, \cdots$	$\frac{1}{p}$	$\frac{1}{p}$
5	超几何分布 $X \sim H(M, N, n)$	$P\{X = k\} = \frac{C_{N-M}^{n-k} C_M^k}{C_N^n}$	$\frac{nM}{N}$	$\frac{nM(N-n)(N-M)}{N^2(N-1)}$
6	均匀分布 $X \sim U[a, b]$	$f(x) = \begin{cases} \frac{1}{b-a} & a \leq x \leq b \\ 0 & \text{其他} \end{cases}$	$\frac{a+b}{2}$	$\frac{(b-a)^2}{12}$
7	指数分布 $X \sim E(\lambda)$	$f(x) = \begin{cases} \lambda e^{-\lambda x} & x \geq 0 \\ 0 & \text{其他} \end{cases}$	$\frac{1}{\lambda}$	$\frac{1}{\lambda^2}$
8	正态分布 $X \sim N(\mu, \sigma^2)$	$f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$ $-\infty < x < +\infty$	μ	σ^2

协方差：设 (X, Y) 为二维随机变量，如果有 $E[(X-EX)(Y-EY)]$ 存在，就称之为 X 和 Y 的协方差，记为$\text{Cov}(X, Y)$，即

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

协方差的性质：

- $\text{Cov}(X, X) = DX$
- $\text{Cov}(X, Y) = \text{Cov}(Y, X)$
- $\text{Cov}(X, c) = 0$
- $\text{Cov}(aX, bY) = ab\text{Cov}(X, Y)$
- $\text{Cov}(X_1 \pm X_2, Y) = \text{Cov}(X_1, Y) \pm \text{Cov}(X_2, Y)$
- $D(X \pm Y) = DX + DY \pm 2\,\text{Cov}(X, Y)$
- $[\text{Cov}(X, Y)] \leq DXDY$

相关系数：设 (X, Y) 为二维随机变量，如果 $DX > 0$，$DY > 0$，记随机变量 X 与 Y 的相关系数为 ρ_{XY}，且

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

- 设 $X^* = \frac{X-EX}{\sqrt{DX}}$，$Y^* = \frac{Y-DY}{\sqrt{DY}}$，有 $E(X^*Y^*) = \rho_{XY}$
- $|\rho_{XY}| \leq 1$
- $|\rho_{XY}| = 1 \Leftrightarrow Y = aX + b$
- X 与 Y 不相关 $\Leftrightarrow cov(X, Y) = 0$
- X 与 Y 相互独立 $\implies \rho_{XY} = 0$

矩：对于正整数 k

- 若$E(X^k)$ 存在，则称为 X 的 k 阶原点矩
- 若 $E[(X-EX)^k]$ 存在，则称为X 的 k 阶中心矩
- 若 $E[X^kY^l]$ 存在，则称为 X 和 Y 的 $k+l$ 阶混合原点矩
- 若 $E[(X-EX)^k(T-EY)^l]$，则称为 X 和 Y 的 $k+l$ 阶混合中心距

$[5]$大数定律和中心极限定理

切比雪夫不等式：设随机变量的数学期望 $EX = \mu$，方差 $DX = \sigma^2$，则对任意的 $\epsilon > 0$，有

$$P\{|X-\mu| \geq \varepsilon\} \leq \frac{\sigma^2}{\varepsilon^2} \quad \text{或} \quad P\{|X-\mu| < \varepsilon\} \geq 1 - \frac{\sigma^2}{\varepsilon^2}$$

依概率收敛：设有随机变量序列 $X_1, X_2, \cdots, X_n, \cdots$，如果存在常数 a，使得对任意的 $\varepsilon > 0$，有

$$\lim_{n\to\infty} P\{|X_n - a| < \varepsilon\} = 1$$

就称序列 X_n 依概率收敛于 a，记为

$$\lim_{n\to\infty} X_n \overset{P}{=} a$$

切比雪夫大数定律：>设 $X_1, X_2, \cdots, X_n, \cdots$ 是**相互独立**的随机变量序列，如果数学期望 EX_i 和方差 DX_i 均存在，且存在常数 c，使得 $DX_i \leq c$，则对任意的 $\varepsilon > 0$，有

$$\lim_{n\to\infty} P\left\{\left|\frac{1}{n}\sum_{i=1}^{n} X_i - \frac{1}{n}\sum_{i=1}^{n} EX_i\right| < \varepsilon\right\} = 1$$

伯努利大数定律：设 n_A 是 n 重独立重复试验中事件 A 发生的次数，p 是事件 A 在每次实验中发生的概率，则对任意的 $\varepsilon > 0$，有

$$\lim_{n\to\infty} P\{\,|\frac{n_A}{n}-p|<\varepsilon\,\}=1$$

辛钦大数定律：>设随机变量 $X_1, X_2, \cdots, X_n, \cdots$ 相互独立**同分布**，若果 $EX_i=\mu$，则对任意的 $\varepsilon>0$，有

$$\lim_{n\to\infty} P\{\left|\frac{1}{n}\sum_{i=1}^{n}X_i-\mu\right|<\varepsilon\,\}=1$$

列维-林德伯格中心极限定理：设随机变量序列 $X_1, X_2, \cdots, X_n, \cdots$ 独立**同分布**，且 $EX_i=\mu$，$DX_i=\sigma^2$，令 Y_n 为 $\sum_{i=1}^{n}X_i$ 的标准化随机变量，即 $Y_n=\frac{\sum_{i=1}^{n}X_i-n\mu}{\sqrt{n}\sigma}$，记 Y_n 的分布函数为 $F_{Y_n}(x)$，有

$$\lim_{n\to\infty} F_{Y_n}(x)=\Phi(x)=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x}e^{-\frac{t^2}{2}}\,\mathrm{d}t$$

理解为，当 n 充分大时，

$$\sum_{i=1}^{n}X_i\overset{近似于}{\sim}N(n\mu,n\sigma^2)$$

棣莫弗-拉普拉斯中心极限定理：设随机变量 $X_n\sim B(n,\,p)$，则

$$\lim_{n\to\infty} P\{\left|\frac{X_n-np}{\sqrt{np(1-p)}}\right|<x\,\}=\Phi(x)$$

理解为，当 n 充分大时，

$$X_n\overset{近似于}{\sim}N(\,np,\,np(1-p)\,)$$

重要推论：

- $\overline{X}=\frac{1}{n}\sum X_i$
- $\sum(X_i-\overline{X})^2=\sum X_i^2-n\overline{X}^2$
- $P\{\,a<\sum_{i=1}^{n}X_i\le b\,\}=\Phi(\frac{b-n\mu}{\sqrt{n}\sigma})-\Phi(\frac{a-n\mu}{\sqrt{n}\sigma})$

[6]数理统计基础

基本概念：

- 总体：研究问题中所有被考察对象的全体. 分为有限总体和无限总体
- 个体：总体中的每一个成员
- 样本（简单随机样本）：总体 X 中，按一定的规则任意抽取的部分个体 $X_1, x_2, \cdots, X_n$. n 称为样本容量
  - 代表性：每个 X_i 与总体 X 同分布
  - 独立性：$X_1, X_2, \cdots, X_n$ 相互独立
- 离散型样本的分布律

$$P\{\,X_1=x_1,\,X_2=x_2,\cdots,\,X_n=x_n\,\}=\prod_{i=1}^{n}p(x_i)$$

- 连续型样本的密度函数

$$f(x_1,\,x_2,\cdots,\,x_n)=\prod_{i=1}^{n}f_{X_i}(x_i)$$

- 统计量：$g(x_1, x_2, \cdots, x_n)$ 不依赖于总体 X 中的任何**未知参数**，称随机变量 $g(X_1, X_2, \cdots, X_n)$ 为一个统计量
- 观察值：$(x_1, x_2, \cdots, x_n)$
- 抽样分布：统计量的分布

常见统计量：

- 均值：$\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$
- 方差：$S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{X})^2 = \frac{1}{n-1}(\sum_{i=1}^{n} X_i^2 - n\overline{X}^2)$
- 标准差：$S = \sqrt{S^2}$
- k 阶原点矩：$A_k = \frac{1}{n}\sum_{i=1}^{n} X_i^k$
- k 阶中心矩：$B_k = \frac{1}{n}\sum_{i=1}^{n}(X_i - \overline{X})^k$
- 2 阶中心矩：$S_n^2 = B_2$
- 设总体 X 数学期望 $EX = \mu$，方差 $DX = \sigma^2$，有
  - $E(\overline{X}) = \mu$
  - $D(\overline{X}) = \frac{\sigma^2}{\mu}$
  - $E(S^2) = \sigma^2$

χ^2 **分布**：设 $(X_1, X_2, \cdots, X_n)$ 为来自总体 $X \sim N(0, 1)$ 的一个样本，称统计量

$$\chi^2 = \sum_{i=1}^{n} X_i^2$$

服从**自由度为** n **的** χ^2 **分布**，记作 $\chi^2 \sim \chi^2(n)$

- $E(\chi^2) = n$
- $D(\chi^2) = 2n$
- $X \sim N(0, 1) \implies \chi^2 \sim \chi^2(1)$
- $\chi_1^2 \sim \chi_1^2(n_1)$ 与 $\chi_2^2 \sim \chi_2^2(n_2)$ 相互独立 $\implies \chi_1^2 + \chi_2^2 \sim \chi^2(n_1 + n_2)$
- $\chi^2 = \frac{(n-1)S^2}{\sigma^2} \sim \chi^2(n-1)$
- $n \to \infty$，$\chi^2 \overset{近似于}{\sim} N(n, 2n)$

t **分布**：设随机变量 $X \sim N(0, 1)$，$Y \sim \chi^2(n)$，且 X 和 Y 相互独立，称随机变量

$$T = \frac{X}{\sqrt{Y/n}}$$

服从自由度为 n 的 t 分布，记作 $T \sim t(n)$

- $ET = 0$
- $DT = \frac{n}{n-2}$
- $n \to \infty$，$T \overset{近似于}{\sim} N(0, 1)$
- $T \sim t(n) \implies T^2 \sim F(1, n)$
- $T = \frac{\overline{X}-\mu}{S/\sqrt{n}} \sim (n-1)$

F **分布**：设随机变量 $X \sim \chi^2(n_1)$，$Y \sim \chi^2(n_2)$，且 X 和 Y 相互独立，则称随机变量

$$F = \frac{X/n_1}{Y/n_2}$$

服从第一自由度为 n_1，第二自由度为 $n-2$ 的 F 分布，记作 $F \sim F(n_1, n_2)$

- $F = \frac{S_X^2/\sigma_1^2}{S_Y^2/\sigma_2^2} \sim F(n_1-1,\, n_2-1)$
- $F \sim F(n_1,\, n_2) \implies \frac{1}{F} \sim F(n_2,\, x_1)$

上侧（右侧）分位点： x_p

- $0 < p < q$
- $P\{\ X \geq x_p\ \} \geq p$
- $P\{\ X \leq x_p\ \} \geq 1-p$
- $t_{1-\alpha}(n) = -t_\alpha(n)$
- $F_{1-\alpha}(n_1,\, n_2) = \frac{1}{F_\alpha(n_2,\, n_1)}$

[7]参数估计

点估计： 使用样本 $(X_1,\, X_2,\, \cdots,\, X_n)$，构造一个合适的统计量 $\hat{\theta} = \hat{\theta}(X_1,\, X_2,\, \cdots, X_n)$ 作为未知参数 θ 的估计.

- 估计量：$\hat{\theta}$
- 估计值：$\hat{\theta}(x_1,\, x_2,\, \cdots,\, x_n)$

矩估计法：

$$\frac{1}{n}\sum_{i=1}^{n} X_i^r \approx E(X^r)$$

最大似然估计法：

- 似然函数：$L(\theta) = L(X_1,\, X_2,\, \cdots,\, X_n;\, \theta)$
- 最大似然估计量：$L(X_1,\, X_2, \cdots,\, X_n;\, \hat{\theta}) = \max L(X_1,\, X_2,\, \cdots,\, X_n;\, \theta)$

步骤

1. 写出似然函数

$$L(\theta_1,\, \theta_2,\, \cdots,\, \theta_m) = \begin{cases} \prod_{i=1}^{n} p(x_i;\, \theta_1,\, \cdots,\, \theta_m) & X \text{ 为离散型} \\ \prod_{i=1}^{n} f(x_i;\, \theta_1,\, \cdots,\, \theta_m) & X \text{ 为连续型} \end{cases}$$

2. 两边取对数
3. 求导数

$$\frac{\partial \ln L}{\partial \theta_i} = 0, \quad i = 1,\, 2,\, \cdots,\, m$$

估计量的评价标准： 设 $\hat{\theta} = \hat{\theta}(X_1,\, X_2,\, \cdots,\, X_n)$ 为 θ 的估计量

- **均方误差：**

$$MSE(\hat{\theta}) = E[\,(\hat{\theta}-\theta)^2\,] = D\,\hat{\theta} + (\,\theta - E\,\hat{\theta}\,)^2$$

- **无偏性：**
  - 无偏估计：$E\,\hat{\theta} = \theta$
  - 渐进无偏估计：$\lim_{n\to\infty} E\,\hat{\theta} = \theta$
  - 若估计量 $\theta_1,\, \theta_2,\, \cdots,\, \theta_m$ 均为 θ 的无偏估计，且 $\sum_{i=1}^{m} c_i = 1$，则 $\sum_{i=1}^{m} c_i\theta_i$ 也为 θ 的无偏估计
- **有效性：** 若 $\hat{\theta}_1$ 和 $\hat{\theta}_2$ 均为 θ 的无偏估计，且 $D\,\hat{\theta}_1 < D\,\hat{\theta}_2$，则称 $\hat{\theta}_1$ 比 $\hat{\theta}_2$ 有效
- **相合性（一致性）：** 若对任意的 $\varepsilon > 0$，有 $\lim_{n\to\infty} P\{\,|\,\hat{\theta}-\theta\,| < \varepsilon\,\} = 1$，则称 $\hat{\theta}$ 为 θ 的相合估计

置信度： $P\{\ \theta \in (\,\theta_1,\, \theta_2\,)\ \} = P\{\ \theta_1 < \theta < \theta_2\ \} = 1-\alpha$

- **置信系数（置信水平）：** $1-\alpha$
- **置信区间：** (θ_1, θ_2) 为 θ 的置信度为 $1-\alpha$ 的置信区间
- 置信上下限：θ_1，θ_2

区间估计：

- 概率$P\{\theta \in (\theta_1, \theta_2)\}$ 反映置信度
- 区间长度 $\theta_2-\theta_1$ 反映精度
- 在确保置信度为 $1-\alpha$ 的前提下，构造统计量 θ_1 和 θ_2，使得 $\theta_2-\theta_1$ 尽可能小

置信区间估计表：

待估参数	其他条件	中间统计量及其分布	置信区间
μ	σ^2 已知	$U=\frac{\overline{X}-\mu}{\sigma/\sqrt{n}} \sim N(0,1)$	$\left(\overline{X}-u_{\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}}, \overline{X}+u_{\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}}\right)$
	σ^2 未知	$T=\frac{\overline{X}-\mu}{S/\sqrt{n}} \sim t(n-1)$	$\left(\overline{X}-t_{\frac{\alpha}{2}}(n-1)\frac{S}{\sqrt{n}}, \overline{X}+t_{\frac{\alpha}{2}}(n-1)\frac{S}{\sqrt{n}}\right)$
σ^2	μ已知	$\chi^2=\frac{1}{\sigma^2}\sum_{i=1}^n(X_i-\mu)^2 \sim \chi^2(n)$	$\left(\frac{\sum_{i=1}^n(X_i-\mu)^2}{\chi^2_{\frac{\alpha}{2}}(n)}, \frac{\sum_{i=1}^n(X_i-\mu)^2}{\chi^2_{1-\frac{\alpha}{2}}(n)}\right)$
	μ未知	$\chi^2=\frac{(n-1)S^2}{\sigma^2} \sim \chi^2(n-1)$	$\left(\frac{(n-1)S^2}{\chi^2_{\frac{\alpha}{2}}(n-1)}, \frac{(n-1)S^2}{\chi^2_{1-\frac{\alpha}{2}}(n-1)}\right)$

[8]假设检验

原假设 H_0	备择假设 H_1	均为正态总体	检验统计量及在 H_0 成立时的分布	拒绝域
$\mu=\mu_0$	$\mu\neq\mu_0$	σ^2 已知	$Z=\frac{\overline{X}-\mu}{\frac{\sigma}{\sqrt{n}}} \sim N(0,1)$	$\lvert z\rvert \geq z_{\frac{\alpha}{2}}$
$\mu\leq\mu_0$	$\mu>\mu_0$			$z\geq z_\alpha$
$\mu\geq\mu_0$	$\mu<\mu_0$			$z\leq -z_\alpha$
$\mu=\mu_0$	$\mu\neq\mu_0$	σ^2 未知	$T=\frac{\overline{X}-\mu}{\frac{s}{\sqrt{n}}} \sim t(n-1)$	$\lvert t\rvert \geq t_{\frac{\alpha}{2}}(n-1)$
$\mu\leq\mu_0$	$\mu>\mu_0$			$t\geq t_\alpha(n-1)$
$\mu\geq\mu_0$	$\mu<\mu_0$			$t\leq -t_\alpha(n-1)$
$\sigma^2=\sigma_0^2$	$\sigma^2\neq\sigma_0^2$	μ 未知	$\chi^2=\frac{(n-1)S^2}{\sigma^2} \sim \chi^2(n-1)$	$\chi^2\leq\chi^2_{1-\frac{\alpha}{2}}(n-1)$ 或 $\chi^2\geq\chi^2_{\frac{\alpha}{2}}(n-1)$
$\sigma^2\leq\sigma_0^2$	$\sigma^2>\sigma_0^2$			$\chi^2\geq\chi^2_\alpha(n-1)$
$\sigma^2\geq\sigma_0^2$	$\sigma^2<\sigma_0^2$			$\chi^2\leq\chi^2_{1-\alpha}(n-1)$