随机事件与概率

- 常用公式
    - 加法

$$P(A \cup B) = P(A) + P(B) - P(AB)$$

$$P(A \cup B \cup C) = P(A) + P(B) + P(C) - P(AB) - P(AC) - P(BC) + P(ABC)$$

    - 减法

$$P(A - B) = P(A) - P(AB)$$

- 条件概率

$$P(A|B) = \frac{P(AB)}{P(B)}$$

- 全概率公式
  若 $\sum_{i=1}^{n} B_i = \Omega$，则

$$P(A) = \sum_{i=1}^{n} P(A|B_i)P(B_i)$$

- 贝叶斯公式
  若 $\sum_{i=1}^{n} B_i = \Omega$，则

$$P(B_i|A) = \frac{P(A|B_i)P(B_i)}{P(A)} = \frac{P(A|B_i)P(B_i)}{\sum_{j=1}^{n} P(A|B_j)P(B_j)}$$

  简化为

$$P(B|A) = \frac{P(A|B)P(B)}{P(A)}$$

- 独立事件
  若 A 和 B 独立，则

$$P(AB) = P(A)P(B)$$
$$P(A|B) = P(A)$$
$$P(B|A) = P(B)$$

> 此时 $\overline{A}, B$ 、 $\overline{B}, A$ 和 $\overline{B}, \overline{A}$ 也相互独立

古典概型略

一维随机变量及其分布函数

离散型随机变量及其分布

- 两点分布

$$P(X = 0) = 1 - p, \quad P(X = 1) = p$$

- 二项分布 $B(n, p)$

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

- 松柏分布 $P(\lambda)$

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

- 超几何分布 $H(n,m,N)$

$$P(X=k)=\frac{C_m^k C_{N-m}^{n-k}}{C_N^n},\quad k=0,1,\dots,n$$

- 松柏定理

  当 $n\to\infty$ 时，令 $\lambda=np$，则二项分布 $B(n,p)$ 近似于泊松分布 $P(\lambda)$。

$$\lim_{n\to\infty} B(n,p)=P(np)$$

离散型随机变量的分布函数为

$$F(x)=P(X\le x)=\sum_{k\le x}P(X=k)$$

可以得到

$$P(a<X\le b)=F(b)-F(a)$$

连续型随机变量及其分布

若存在非负可积函数 $f(x)$，使得随机变量 X 取值于任意区间 $[a,b]$ 上的概率为

$$P(a\le X\le b)=\int_a^b f(x)dx$$

则称 $f(x)$ 为随机变量 X 的概率密度函数，简称密度函数。

规范性条件:

1. $f(x)\ge 0,\forall x\in\mathbb{R}$
2. $\int_{-\infty}^{+\infty} f(x)dx=1$

- 均匀分布 $U(a,b)$

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

- 指数分布 $E(\lambda)$

$$f(x)=\begin{cases}\lambda e^{-\lambda x}, & x\ge 0\\ 0, & \text{otherwise}\end{cases}$$

- 正态分布 $N(\mu,\sigma^2)$

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

  标准正态分布 $N(0,1)$ 的密度函数为

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

连续型随机变量的分布函数为

$$F(x) = P(X < x) = \int_{-\infty}^{x} f(t)dt$$

同样有

$$P(a < X < b) = F(b) - F(a)$$

> 由于连续性随机变量单点概率为0的特点，因此这里的比较符号可任意使用 $\leq$ 或 $<$。

二维随机变量及其分布函数

二维随机变量的分布函数通常又称为**联合分布函数**或**联合密度函数**

$$F(x, y) = P(X \leq x, Y \leq y)$$

- 联合分布函数的性质
    1. $F(x, y)$ 是非递减函数
    2. $\lim_{x \to -\infty} F(x, y) = 0, \forall y$
    3. $\lim_{y \to -\infty} F(x, y) = 0, \forall x$
    4. $\lim_{x \to +\infty} F(x, y) = P(Y \leq y), \forall y$
    5. $\lim_{y \to +\infty} F(x, y) = P(X \leq x), \forall x$
    6. $\lim_{x \to +\infty, y \to +\infty} F(x, y) = 1$
    7. 对于四个实数 $a \leq b, c \leq d$，有

$$F(b, d) - F(a, d) - F(b, c) + F(a, c) = P(a \leq X \leq b, c \leq Y \leq d) \geq 0$$

边缘分布函数

对于二维随机变量 (X, Y)，其边缘分布函数为

$$F_X(x) = P(X \leq x) = \lim_{y \to +\infty} F(x, y)$$
$$F_Y(y) = P(Y \leq y) = \lim_{x \to +\infty} F(x, y)$$

条件分布函数

对于二维随机变量 (X, Y)，其条件密度函数为

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

其条件分布函数为

$$F_{Y|X}(y|x) = \frac{F(x, y)}{F_X(x)}$$
$$F_{X|Y}(x|y) = \frac{F(x, y)}{F_Y(y)}$$

二维离散型随机变量

- 对于二维离散型随机变量 (X,Y)，其分布律为

$$P(X=x_i,Y=y_j)=P(X=x_i\cap Y=y_j)=p_{ij}$$

- 其联合分布函数为

$$F(x,y)=P(X\le x,Y\le y)=\sum_{x_i\le x}\sum_{y_j\le y}p_{ij}$$

- 其边缘分布律为

$$F_X(x)=F(x,+\infty)=\sum_{y_j}p_{X,Y}(x,y_j)$$
$$F_Y(y)=F(+\infty,y)=\sum_{x_i}p_{X,Y}(x_i,y)$$

则有

$$P(X=x_i)=p_X(x_i)=\sum_{y_j}p_{X,Y}(x_i,y_j)$$
$$P(Y=y_j)=p_Y(y_j)=\sum_{x_i}p_{X,Y}(x_i,y_j)$$

- 离散型随机变量独立性判别条件

若 X,Y 独立，则

$$f(x,y)=f_X(x)f_Y(y) \tag{1}$$

- 条件分布律为

$$P(Y=y_j|X=x_i)=\frac{P(X=x_i,Y=y_j)}{P(X=x_i)}=\frac{p_{ij}}{p_X(x_i)}$$
$$P(X=x_i|Y=y_j)=\frac{P(X=x_i,Y=y_j)}{P(Y=y_j)}=\frac{p_{ij}}{p_Y(y_j)}$$

二维连续型随机变量

- 对于二维连续型随机变量 (X,Y)，其联合分布函数为

$$F(x,y)=P(X\le x,Y\le y)=\int_{-\infty}^{x}\int_{-\infty}^{y}f(u,v)dvdu$$
$$\frac{\partial^2F(x,y)}{\partial x\partial y}=f(x,y)$$

$f(x,y)$ 为其联合密度函数，满足

$$f(x,y)\ge 0,\forall x,y\in\mathbb{R}$$
$$\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}f(x,y)dxdy=1$$
$$P\{(X,Y)\in D\}=\int_D f(x,y)dxdy$$

- 其边缘分布函数为

$$F_X(x)=P(X\le x)=F(x,+\infty)=\int_{-\infty}^{x}\left(\int_{-\infty}^{+\infty}f(u,y)dy\right)du$$

$$F_Y(y)=P(Y\le y)=F(+\infty,y)=\int_{-\infty}^{y}\left(\int_{-\infty}^{+\infty}f(x,v)dx\right)dy$$

若 X,Y 为随机变量，则其概率密度函数为

$$f_X(x)=\int_{-\infty}^{+\infty}f(x,y)dy$$

$$f_Y(y)=\int_{-\infty}^{+\infty}f(x,y)dx$$

- 连续性随机变量的独立性判别条件

同式(1)

- 条件概率密度函数为

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

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

则条件分布函数为

$$F_{Y|X}(y|x)=\frac{F(x,y)}{F_X(x)}=\int_{-\infty}^{y}f_{Y|X}(v|x)dv$$

$$F_{X|Y}(x|y)=\frac{F(x,y)}{F_Y(y)}=\int_{-\infty}^{x}f_{X|Y}(u|y)du$$

连续型随机变量函数的分布

1. $Z=X+Y$

其分布函数为

$$F_Z(z)=P(Z\le z)=\iint\limits_{x+y\le z}f(u,v)\mathrm{d}u\mathrm{d}v$$

化简该二重积分，得到

$$F_Z(z)=\int_{-\infty}^{+\infty}\left[\int_{-\infty}^{z}f(t-y,y)\mathrm{d}y\right]\mathrm{d}t$$

$$f_Z(z)=F_Z'(z)=\int_{-\infty}^{+\infty}f(z-y,y)\mathrm{d}y=\int_{-\infty}^{+\infty}f(x,z-x)\mathrm{d}x \tag{2}$$

当 X,Y 独立时，式（2）可化为卷积公式，即

$$f_Z(z)=\int_{-\infty}^{+\infty}f_X(x)f_Y(z-x)\mathrm{d}x$$

2. $M=\max(X,Y)$ 及 $N=\min(X,Y)$

其分布函数为

$$F_M(z)=P(M\le z)=P(X\le z,Y\le z)=P(X\le z)P(Y\le z)=F_X(z)F_Y(z)$$

$$\begin{aligned}F_N(z)&=P(N\le z)=1-P(N>z)=1-P(X>z,Y>z)\\&=1-P(X>z)P(Y>z)=1-[1-F_X(z)][1-F_Y(z)]\end{aligned}$$

对于多维连续随机变量 $X_1,X_2,X_3...X_i$ ，其满足

$$F_M(z)=\prod_{i=1}^{n}F_{X_i}(z)$$

$$F_N(z)=1-\prod_{i=1}^{n}[1-F_{X_i}(z)]$$

随机变量的数字特征

数学期望

- 离散型随机变量

  对于离散型随机变量 X，其数学期望为

  $$E(X)=\sum_{x_i}x_iP(X=x_i)$$

- 连续型随机变量

  对于连续型随机变量 X，其数学期望为

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

随机变量函数的数学期望

- 离散型随机变量

  对于离散型随机变量 X，其函数 $g(X)$ 的数学期望为

  $$E(g(X))=\sum_{x_i}g(x_i)P(X=x_i)$$

  对于二维离散型随机变量 (X,Y)，其函数 $g(X,Y)$ 的数学期望为

  $$E(g(X,Y))=\sum_{x_i}\sum_{y_j}g(x_i,y_j)P(X=x_i,Y=y_j)$$

- 连续型随机变量

  对于连续型随机变量 X，其函数 $g(X)$ 的数学期望为

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

  对于二维连续型随机变量 (X,Y)，其函数 $g(X,Y)$ 的数学期望为

  $$E(g(X,Y))=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}g(x,y)f(x,y)dxdy$$

数学期望的性质

1. 设 C 为常数，则

$$E(C) = C$$

2. 设 X 为随机变量，C 为常数，则

$$E(CX) = CE(X)$$

3. 设 X, Y 为随机变量，则

$$E(X+Y) = E(X) + E(Y)$$

4. 设 X, Y 为独立随机变量，则

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

方差

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

方差的性质

1. 设 C 为常数，则

$$D(C) = 0$$

2. $D(X) \le 0$, $D(X) = 0$ 的充要条件为存在常数 C ，使得 $P(X = C) = 1$ 成立
3. 设 X 为随机变量，C 为常数，则

$$D(CX) = C^2 D(X)$$

4. 设 X, Y 为独立随机变量，则

$$D(X+Y) = D(X) + D(Y)$$

切比雪夫不等式

对于任意常数 $\epsilon > 0$，有

$$P(|X - E(X)| \ge \epsilon) \le \frac{D(X)}{\epsilon^2}$$

常见分布的期望和方差

	(0,1)分布	二项分布	泊松分布	均匀分布	指数分布	正态分布
$E(X)$	p	np	λ	$\frac{a+b}{2}$	$\frac{1}{\lambda}$	μ
$D(X)$	$p(1-p)$	$np(1-p)$	λ	$\frac{(b-a)^2}{12}$	$\frac{1}{\lambda^2}$	σ^2

协方差、相关系数和矩

- 协方差

  对于二维随机变量 (X, Y)，其协方差为

$$Cov(X,Y)=E[(X-E(X))(Y-E(Y))]=E(XY)-E(X)E(Y)$$

基于协方差，对任意的随机变量 X,Y，有

$$D(X+Y)=D(X)+D(Y)+2Cov(X,Y)$$
$$D(X-Y)=D(X)+D(Y)-2Cov(X,Y)$$

- 协方差的性质
  - $Cov(X,X)=D(X)$
  - $Cov(X,C)=0$，其中 C 为常数
  - $Cov(X,Y)=Cov(Y,X)$
  - 对于任意常数 a,b，有 $Cov(aX,bY)=abCov(X,Y)$
  - $Cov(X+Y,Z)=Cov(X,Z)+Cov(Y,Z)$
  - 若 X,Y 独立，则 $Cov(X,Y)=0$
- 相关系数

  相关系数 ρ_{XY} 定义为

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

- 相关系数的性质
  - $|\rho_{XY}|\leq 1$
  - $\rho_{XY}\pm 1$ 的充要条件为 X 和 Y 完全正相关或负相关，即存在常数 a,b 使得 $Y=aX+b$ 成立。
  - 对于二维正态分布，X,Y 相互独立的充分必要条件是 $\rho_{XY}=0$。

大数定律和中心极限定理

大数定律

- 依概率收敛

  对于一列独立同分布的随机变量 $X_1,X_2,\ldots,X_n$，若存在常数 a，使得对任意 $\epsilon>0$，有

$$\lim_{n\to\infty}P\left(|X_n-a|<\epsilon\right)=1$$

  则称 X_n 依概率收敛于 a。

- 切比雪夫大数定律

  设 $X_1,X_2,\ldots,X_n$ 为一列两两不相关的随机变量，且期望存在，方差有界，则对于任意 $\epsilon>0$，有

$$\lim_{n\to\infty}P\left(\left|\frac{1}{n}\sum_{i=1}^{n}(X_i-E(X_i))\right|<\epsilon\right)=0$$

- 伯努利大数定律

  设 μ_n 为 n 重伯努利试验中事件 A 出现的次数，又设 p 为事件 A 的成功概率，则对于任意 $\epsilon>0$，有

$$\lim_{n\to\infty}P\left(\left|\frac{\mu_n}{n}-p\right|<\epsilon\right)=1$$

- 辛钦大数定律

  设 $X_1, X_2, \ldots, X_n$ 为一列独立同分布的随机变量，且 $E(X_i) = \mu$，则对于任意 $\epsilon > 0$，有

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

中心极限定理

设 $X_1, X_2, \ldots, X_n$ 为一列独立同分布的随机变量，且其期望和方差均存在,则

$$\sum_{i=1}^{n} X_i \to N(E(\sum_{i=1}^{n} X_i), D(\sum_{i=1}^{n} X_i))$$

- 独立同分布中心极限定理

  设 $X_1, X_2, \ldots, X_n$ 为一列独立同分布的随机变量，且其期望和方差均存在，记

$$Y_n = \frac{\sum_{i=1}^{n} X_i - \sum_{i=1}^{n} E(X)}{\sqrt{\sum_{i=1}^{n} D(X_i)}} \to N(0,1)$$

  则 Y_n 的分布函数 $F_n(x)$ 满足

$$\lim_{n\to\infty} F_n(x) = \Phi(x)$$

  其中 $\Phi(x)$ 为标准正态分布的分布函数。

- 德莫佛-拉普拉斯中心极限定理

  设 μ_n 为 n 重伯努利试验中事件 A 出现的次数，又设 p 为事件 A 的成功概率，则对于任意 $\epsilon > 0$，有

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

数理统计

随机样本

从总体 X 中随机地抽取 n 个样本 $X_1, X_2, \ldots, X_n$ 组成一个整体，称为总体 X 的**样本**，组成样本的个体称为样本的分量，一个样本中所含分量的个数称为**样本容量。**

- 由于 n 次试验是在完全相同的条件下进行的，每个个体被抽到的机会相等，所以可认为每个分量 X_i 与总体 X 具有相同的分布，且各个分量之间相互独立。

设总体 X 的分布函数为 $F(x)$，则样本 $X_1, X_2, \ldots, X_n$ 的联合分布函数为

$$F(x_1, x_2, \ldots, x_n) = P(X_1 \le x_1, X_2 \le x_2, \ldots, X_n \le x_n) = \prod_{i=1}^{n} F(x_i) \tag{3}$$

如果总体 X 为离散型随机变量，则式(3)可视为

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

如果总体 X 为连续型随机变量，设其概率密度函数为 $f(x)$，则式(3)可视为

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

统计量

- 统计量是样本的函数，通常用来估计总体的某些特征。

常用统计量

- 样本均值

  对于样本 $X_1,X_2,\ldots,X_n$，其样本均值为

$$\bar{X}=\frac{1}{n}\sum_{i=1}^{n}X_i$$

- 样本方差

  对于样本 $X_1,X_2,\ldots,X_n$，其样本方差为

$$S^2=\frac{1}{n-1}\sum_{i=1}^{n}(X_i-\bar{X})^2=\frac{1}{n-1}(\sum_{i=1}^{n}X_i^2-n\bar{X}^2)$$

- 样本标准差

  对于样本 $X_1,X_2,\ldots,X_n$，其样本标准差为

$$S=\sqrt{S^2}$$

- 样本 k 阶极点矩
  对于样本 $X_1,X_2,\ldots,X_n$，其样本 k 阶极点矩为

$$A_k=\frac{1}{n}\sum_{i=1}^{n}X_i^k$$

- 样本 k 阶中心矩
  对于样本 $X_1,X_2,\ldots,X_n$，其样本 k 阶中心矩为

$$B_k=\frac{1}{n}\sum_{i=1}^{n}(X_i-\bar{X})^k$$

抽样分布

1. χ^2 分布

   设 $X_1,X_2,\ldots,X_n$ 是来自正态总体 $N(0,1)$ 的随机样本，则称统计量

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

   服从 χ^2 分布，其自由度为 n，记作 $\chi^2\sim\chi^2(n)$。自由度是指包含的独立随机变量的数量。

   - χ^2 分布具有可加性

   如果 $X_1,X_2,\ldots,X_n$ 是来自正态总体 $N(0,1)$ 的随机样本，且 $Y_1,Y_2,\ldots,Y_m$ 是来自正态总体 $N(0,1)$ 的随机样本，则

$$\chi^2=\sum_{i=1}^{n}X_i^2+\sum_{j=1}^{m}Y_j^2$$

服从 χ^2 分布，其自由度为 $n+m$，记作 $\chi^2 \sim \chi^2(n+m)$。

- χ^2 分布的期望和方差

对于自由度为 n 的 χ^2 分布，其期望为 $E(\chi^2)=n$，方差为 $D(\chi^2)=2n$。

- χ^2 分布的上 α 分位点

对于给定的 $0<\alpha<1$，则称满足条件

$$P(\chi^2 > \chi^2_\alpha(n)) = \alpha$$

的点 $\chi^2_\alpha(n)$ 为 χ^2 分布的上 α 分位点。

2. t 分布

设 $X_1, X_2, ..., X_n$ 是来自正态总体 $N(0,1)$ 的随机样本，$Y \sim \chi^2(n)$，并且 X 和 Y 独立，则称统计量

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

服从 t 分布，其自由度为 n，记作 $T \sim t(n)$。

- t 分布的期望和方差

对于自由度为 n 的 t 分布，其期望为 $E(t)=0$，方差为 $D(t)=\frac{n}{n-2}$（当 $n>2$ 时）。

- t 分布的上 α 分位点

对于给定的 $0<\alpha<1$，则称满足条件

$$P(T > t_\alpha(n)) = \alpha$$

的点 $t_\alpha(n)$ 为 t 分布的上 α 分位点。

其满足

$$t_{1-\alpha}(n) = -t_\alpha(n)$$

3. F 分布

设 $X \sim \chi^2(n_1)$， $Y \sim \chi^2(n_2)$，并且 X 和 Y 独立，则称统计量

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

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 分布的性质

对于第一自由度为 n_1，第二自由度为 n_2 的 F 分布，其期望为

$$E(F) = \frac{n_2}{n_2-2}$$

当 $n_2>4$ 时，方差为

$$D(F) = \frac{2n_2^2(n_1+n_2-2)}{n_1(n_2-2)^2(n_2-4)}$$

若 $F \sim F(n_1, n_2)$，则

$$\frac{1}{F} \sim F(n_2, n_1)$$

- F 分布的上 α 分位点

对于给定的 $0<\alpha<1$，则称满足条件

$$P(F>F_\alpha(n_1,n_2))=\alpha$$

的点 $F_\alpha(n_1,n_2)$ 为 F 分布的上 α 分位点。

其满足

$$F_{1-\alpha}(n_1,n_2)=\frac{1}{F_\alpha(n_2,n_1)}$$

正态总体的样本均值和样本方差的分布

1. 设 $X_1,X_2,\ldots,X_n$ 为来自正态总体 $N(\mu,\sigma^2)$ 的随机样本，则样本均值 $\bar{X}$ 的分布为

$$\bar{X}\sim N\left(\mu,\frac{\sigma^2}{n}\right)$$

2. 设 $X_1,X_2,\ldots,X_n$ 为来自正态总体 $N(\mu,\sigma^2)$ 的随机样本，$\bar{X}S^2$
分别为样本均值和样本方差，则

$$\frac{(n-1)S^2}{\sigma^2}\sim\chi^2(n-1)$$

且 $\bar{X}$ 与 S^2 独立。

3. 设 $X_1,X_2,\ldots,X_n$ 为来自正态总体 $N(\mu,\sigma^2)$ 的随机样本，$\bar{X}S^2$
分别为样本均值和样本方差，则

$$\frac{\bar{X}-\mu}{S/\sqrt{n}}\sim t(n-1)$$

参数估计

点估计

矩估计

大数定理：样本的矩依概率收敛于总体的矩。

极大似然估计

设总体 X 的概率密度函数为 $f(x;\theta)$，其中 θ 为未知参数。对于样本 $X_1,X_2,\ldots,X_n$，其似然函数为

$$L(\theta)=\prod_{i=1}^n f(X_i;\theta)$$

若某个统计量 $\hat{\theta}$ 使得 $L(\hat{\theta})$ 最大，则称 $\hat{\theta}$ 为 θ 的极大似然估计。

极大似然估计的不变原理：如果 $\hat{\theta}$ 是 θ 的极大似然估计，则对于任何可逆变换 $g(\cdot)$，$g(\hat{\theta})$ 也是 $g(\theta)$ 的极大似然估计。

例题：设总体 X 的概率密度函数为

$$f(x;\theta)=\begin{cases}\theta x^{\theta-1}, & 0<x<1,\theta>0\\ 0, & \text{elsewhere}\end{cases}$$

$X_1, X_2, X_3, \ldots, X_n$ 为来自总体 X 的随机样本

1. 求 θ 的一个矩估计
2. 若 $x_1, x_2, \ldots, x_n$ 为样本数据，求 θ 的极大似然估计
3. 求 $U = e^{\frac{1}{\theta}}$ 的极大似然估计

解:

(1) 在此求样本的一阶原点矩

$$E(X) = \int_0^1 x\theta x^{\theta-1}\mathrm{d}x = \frac{\theta}{\theta+1} = \frac{1}{n}\sum_{i=1}^{n} X_i = \bar{X}$$

则 θ 的一个矩估计为

$$\theta = \frac{\bar{X}}{1-\bar{X}}$$

(2) 设似然函数为

$$L(\theta) = \prod_{i=1}^{n} f(X_i;\theta) = \theta^n \cdot (\prod_{i=1}^{n} x_i)^{\theta-1}$$

取对数有

$$\ln L(\theta) = n\ln\theta + (\theta-1)\sum_{i=1}^{n}\ln x_i$$

两边对 θ 求导并令其为0，得到

$$\frac{n}{\theta} + \sum_{i=1}^{n}\ln x_i = 0$$

解得

$$\hat{\theta} = \frac{n}{\sum_{i=1}^{n}\ln x_i}$$

(3) 由极大似然估计不变性可得

$$\hat{U} = e^{\frac{1}{\hat{\theta}}} = e^{\frac{\sum_{i=1}^{n}\ln x_i}{n}}$$

评估量的评判标准

- 无偏性

  设 $\hat{\theta}$ 为 θ 的估计量，则 $\hat{\theta}$ 是无偏的，当且仅当

  $$E(\hat{\theta}) = \theta$$

- 有效性

  设 $\hat{\theta}_1$ 和 $\hat{\theta}_2$ 是 θ 的两个无偏估计量，则 $\hat{\theta}_1$ 比 $\hat{\theta}_2$ 更有效，当且仅当

  $$D(\hat{\theta}_1) < D(\hat{\theta}_2)$$

区间估计

设 θ 是总体 X 的未知参数，$X_1, X_2, ..., X_n$ 为来自总体 X 的随机样本。对于一个给定的置信水平 $1-\alpha$，若有两个统计量 $\hat{\theta_L}$ 和 $\hat{\theta_U}$，使得

$$P(\hat{\theta_L} < \theta < \hat{\theta_U}) = 1 - \alpha$$

则称区间 $[\hat{\theta_L}, \hat{\theta_U}]$ 为 θ 的 $1-\alpha$ 置信区间。

$\hat{\theta_L}$ 和 $\hat{\theta_U}$ 分别称为 θ 的双侧置信下限和上限

正态总体参数的区间估计

1. σ 已知

   设 $X_1, X_2, ..., X_n$ 为来自正态总体 $N(\mu, \sigma^2)$ 的随机样本，且 σ 已知，则 μ 的 $1-\alpha$ 置信区间为

   $$\left[\bar{X} - \mu_{\alpha/2}\frac{\sigma}{\sqrt{n}}, \bar{X} + \mu_{\alpha/2}\frac{\sigma}{\sqrt{n}}\right]$$

   其中 $\mu_{\alpha/2}$ 为标准正态分布的上 $\alpha/2$ 分位点。

2. σ 未知

   设 $X_1, X_2, ..., X_n$ 为来自正态总体 $N(\mu, \sigma^2)$ 的随机样本，且 σ 未知，则 μ 的 $1-\alpha$ 置信区间为

   $$\left[\bar{X} - t_{\alpha/2}(n-1)\frac{S}{\sqrt{n}}, \bar{X} + t_{\alpha/2}(n-1)\frac{S}{\sqrt{n}}\right]$$

   其中 $t_{\alpha/2}(n-1)$ 为自由度为 $n-1$ 的 t 分布的上 $\alpha/2$ 分位点。