目录

1 随机变量的数字特征

r.v. 的分布函数/密度函数是对它的概率性质最完整的刻画
有时更关注 r.v. 的某些特征 (也是由分布决定)
如果是 value, 就称为数字特征

1. 度量中心: 期望/中位数
2. 度量散布程度: 方差/绝对偏差/极差
3. 分布形状: 偏度系数/峰度系数
4. 相关程度: 相关系数 (这是对于两个随机变量而言的)

1.1 数学期望与中位数

1. **数学期望**
   定义
   (a) 离散型:
   $$EX = \sum_{n=1}^{N} x_i P(X = x_i)$$
   (b) 连续型:
   $$EX = \int x f(x)$$
   (c) 条件期望:(以连续型为例)
   $$E(X|Y = y) = \int x f(x|Y = y)$$
   但期望存在有要求: 绝对收敛/绝对可积, 否则称期望不存在
   e.g. Cauchy 分布
   $$f(x) = \frac{1}{\pi(1+x^2)}$$
   $$\int_{-\infty}^{+\infty} \left|\frac{x}{\pi(1+x^2)}\right| dx = +\infty$$
   故不存在期望
   常见期望
   (a) 二项分布 $X \sim B(n,p)$
   $$EX = np$$
   $$\begin{aligned} EX &= \sum_{i=0}^{n} i \frac{n!}{i!(n-i)!} p^i (1-p)^{n-i} \\ &= \sum_{i=1}^{n} np \frac{(n-1)!}{(i-1)!((n-1)-(i-1))!} p^{i-1} (1-p)^{(n-1)-i-1} \\ &= np \sum_{j=0}^{n-1} \frac{(n-1)!}{j!(n-1-j)!} p^j (1-p)^{n-1-j} \\ &= np \end{aligned}$$

凑配

$$\begin{aligned} E[X(X-1)] &= \sum_{j=1}^{n} j(j-1)p^j q^{n-j} C_n^j \\ &= p^2 \frac{\mathrm{d}^2}{\mathrm{d}x^2}\left(\sum_{j=1}^{n} x^j q^{n-j}\right)\Big|_{x=p} \\ &= p^2 \frac{\mathrm{d}^2}{\mathrm{d}x^2}(x+p)^2\Big|_{x=p} \\ &= \boxed{n(n-1)p^2} \end{aligned}$$

(b) 负二项分布 $X \sim NB(r,p)$

$$EX = \boxed{\frac{r}{p}}$$

(c) Poisson 分布 $X \sim P(\lambda)$

$$EX = \lambda$$

$$\begin{aligned} EX &= \sum_{i=0}^{+\infty} i\frac{\lambda^i}{i!}e^{-\lambda} \\ &= \sum_{i=1}^{+\infty} \lambda\frac{\lambda^{i-1}}{(i-1)!}e^{-\lambda} \\ &= \lambda\sum_{j=0}^{+\infty} \frac{\lambda^j}{j!}e^{-\lambda} \\ &= \lambda \end{aligned}$$

(d) 指数分布 $X \sim Exp(\lambda)$

$$EX = \frac{1}{\lambda}$$

$$\begin{aligned} EX &= \int_0^{+\infty} x\lambda e^{-\lambda x}dx \\ &= -xe^{-\lambda x}\Big|_0^{+\infty} + \int_0^{+\infty} e^{-\lambda x}dx \\ &= 0 - \frac{1}{\lambda}e^{-\lambda x}\Big|_0^{+\infty} \\ &= \frac{1}{\lambda} \end{aligned}$$

(e) 卡方分布 $X = \sum_{i=1}^{n} X_i^2 \sim \chi_n^2$

其中 $X_i \sim N(0,1)$

$$EX = n$$

(f) t 分布 $X = \frac{X_1}{\sqrt{X_2/n}} \sim t_n$

其中 X_1, X_2 独立,$X_1 \sim N(0,1), X_2 \sim \chi_n^2$

$$EX = 0$$

(g) F 分布 $X = \frac{X_1/m}{X_2/n} \sim F_{m,n}$

其中 $X_1 \sim \chi_m^2, X_2 \sim \chi_n^2$

$$EX = \frac{n}{n-2}$$

计算 $E(X^2)$:

$$E(X^2+Y^2)=\iint_{R^2}(x^2+y^2)f(x,y)dxdy \quad (1)$$
$$=\int_0^{+\infty} r^3 e^{-\frac{r^2}{2}}dr=2 \quad (2)$$

由 $E(X^2+Y^2)=2$, 以及 iid.

得到

$$E(X^2)=1 \quad (3)$$

(4)

卡方分布可轻易得到该结论

数学期望的性质

以下均假设 r.v. 期望存在

(a) r.v. 线性组合之期望

$$E(\sum_{i=1}^{n} c_i X_i)=\sum_{i=1}^{n} c_i E(X_i)$$

这里 c_i 为常数

神奇之处: 等式始终成立, 不要求 X_i 相互独立!

(b) r.v. 乘积之期望

$$E(\prod_{i=1}^{n} X_i)=\prod_{i=1}^{n} E(X_i)$$

这里要求 X_i 相互独立!

(c) 全期望公式

$$EX=E[E(X|Y)]$$

这里 $E(X|Y)$ 是随机变量 Y 的函数, 故也是随机变量, 具有分布与期望

(d)

$$EX=\int_0^{+\infty} P(X>x)\,dx \quad (5)$$

证明

$$EX=\int_0^{+\infty}\int_x^{+\infty} f(t)dtdx$$
$$=\int_0^{+\infty} f(t)\int_0^{t} dxdt$$
$$=\int_0^{+\infty} tf(t)dt$$

2. **中位数/p 分位数**

定义

(a) 连续型中位数

m 为 X 的中位数 $\Leftrightarrow P(X\le m)\ge \frac{1}{2}, P(X\ge m)\ge \frac{1}{2}$

(b) 连续型 p 分位数
μ_p 为 X 的 p 分位数 $\Leftrightarrow P(X \le \mu_p) \ge p, P(X \ge \mu_p) \ge 1-p$

中位数特点
优点:

(a) 中位数总存在
(b) 中位数受极端值影响小. 如收入, 收入差距大时中位数比均值更有效

缺点:(这使得数学期望重要性超过中位数)

(a) 期望有 $E(\sum_{i=1}^n c_i X_i) = \sum_{i=1}^n c_i E(X_i), EX = E[E(X|Y)]$ 这些优良性质, 而中位数没有这些简单的关系
(b) 中位数不唯一, 对于离散型 r.v. 不易定义

1.2 方差与协方差

1. **方差 (Variance)**

定义

给定随机变量 X, 平方可积

$$Var(X) = E(X-EX)^2 = \sigma^2$$

为方差,$\sqrt{Var(X)} = \sigma$ 为标准差 (为了与原 r.v. 保持量纲一致而引入)
(若上述积分不收敛, 则称不存在方差或方差为正无穷)

性质

(a) $$Var(X) = E(X^2) - (EX)^2$$

(b)
$$0 \le Var(X) = E(X^2) - (EX)^2$$
$$\Rightarrow E(X^2) \le (EX)^2$$

(c) $$Var(cX) = c^2 Var(X)$$

(d) X 退化到常数 $c \Leftrightarrow Var(X) = 0, E(X) = c$

(e) $$Var(X) \le E((X-c)^2)$$
, 等号成立但且仅当 $c = EX, E(X-c) = 0$

(f) 若 X_1, X_2 相互独立,

$$Var(X_1 + X_2) = Var(X_1) + Var(X_2)$$

证明:

$$\begin{aligned}
Var(X_1+X_2) &= E((X_1+X_2)^2) - (E(X_1+X_2))^2 \\
&= E(X_1^2 + X_2^2 + 2X_1X_2) - (E(X_1) + E(X_2))^2 \\
&= E(X_1^2) + E(X_2^2) + 2EX_1X_2 - (EX_1)^2 - (EX_2)^2 - 2EX_1EX_2 \\
&= Var(X_1) + Var(X_2) + 2EX_1EX_2 - 2EX_1EX_2 \\
&= Var(X_1) + Var(X_2)
\end{aligned}$$

内积不等式

let

$$g(t) = E(tX-Y)^2 = t^2EX^2 - 2tEXY + EY^2 \qquad \text{for } (tX-Y)\geq 0, g(t)\geq 0$$

$$\Rightarrow \Delta = B^2 - 4AC \leq 0$$

then we have

$$\boxed{|EXY|^2 \leq EX^2EY^2}$$

常见方差

(a) 二项分布 $X \sim B(n,p)$

$$Var(X) = np(1-p)$$

$$\begin{aligned} Var(X) &= \sum_{i=0}^{n} i^2 \frac{n!}{i!(n-i)!} p^i (1-p)^{n-i} - (EX)^2 \\ &= \sum_{i=1}^{n} np(i-1+1)\frac{(n-1)!}{(i-1)!((n-1)-(i-1))!} p^{i-1}(1-p)^{(n-1)-i-1} - (np)^2 \\ &= np + np\sum_{j=0}^{n-1} j \frac{(n-1)!}{j!(n-1-j)!} p^j (1-p)^{n-1-j} - (np)^2 \\ &= np + n(n-1)p^2 \sum_{k=0}^{n-2} \frac{(n-2)!}{k!(n-2-k)!} p^k (1-p)^{n-2-k} - (np)^2 \\ &= np + n^2p^2 - np^2 - n^2p^2 \\ &= np(1-p) \end{aligned}$$

(b) Poisson 分布 $X \sim P(\lambda)$

$$Var(X) = \lambda$$

$$\begin{aligned} Var(X) &= \sum_{i=0}^{+\infty} i^2 \frac{\lambda^i}{i!} e^{-\lambda} \\ &= \sum_{i=1}^{+\infty} \lambda((i-1)+1)\frac{\lambda^{i-1}}{(i-1)!} e^{-\lambda} - (EX)^2 \\ &= \lambda \sum_{j=0}^{+\infty} \frac{\lambda^j}{j!} e^{-\lambda} + \lambda \sum_{j=0}^{+\infty} j \frac{\lambda^j}{j!} e^{-\lambda} - \lambda^2 \\ &= \lambda + \lambda^2 \sum_{k=0}^{+\infty} \frac{\lambda^k}{k!} e^{-\lambda} - \lambda^2 \\ &= \lambda \end{aligned}$$

(c) 均匀分布 $X \sim U(a,b)$

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

$$
\begin{aligned}
Var(X) &= \int_a^b (x-\frac{a+b}{2})^2 \frac{1}{b-a} dx \\
&= \frac{1}{3(b-a)}(x-\frac{a+b}{2})^3\Big|_a^b \\
&= \frac{1}{3(b-a)}(\frac{b-a}{2})^3 - \frac{1}{3}(\frac{a-b}{2})^3 \\
&= \frac{(b-a)^2}{12}
\end{aligned}
$$

(d) 指数分布 $X \sim Exp(\lambda)$

$$Var(X) = \frac{1}{\lambda^2}$$

$$
\begin{aligned}
Var(X) &= \int_0^{+\infty} x^2\lambda e^{-\lambda x} dx - (EX)^2 \\
&= -x^2 e^{-\lambda x}\Big|_0^{+\infty} + \int_0^{+\infty} 2xe^{-\lambda x} dx - \frac{1}{\lambda^2} \\
&= 0 - \frac{2x}{\lambda} e^{-\lambda x}\Big|_0^{+\infty} + \int_0^{+\infty} \frac{2}{\lambda} e^{-\lambda x} dx - \frac{1}{\lambda^2} \\
&= -\frac{2}{\lambda^2} e^(-\lambda x)\Big|_0^{+\infty} - \frac{1}{\lambda^2} \\
&= \frac{1}{\lambda^2}
\end{aligned}
$$

(e) 卡方分布 $X = \sum_{i=1}^n X_i^2 \sim \chi_n^2$

其中 $X_i \sim N(0,1)$

$$Var(X) = 2n$$

(f) t 分布 $X = \frac{X_1}{\sqrt{X_2/n}} \sim t_n$

其中 X_1, X_2 独立,$X_1 \sim N(0,1), X_2 \sim \chi_n^2$

$$Var(X) = \frac{n}{n-2}$$

(g) F 分布 $X = \frac{X_1/m}{X_2/n} \sim F_{m,n}$

其中 $X_1 \sim \chi_m^2, X_2 \sim \chi_n^2$

$$Var(X) = \frac{2n^2(m+n-2)}{m(n-2)^2(n-4)}$$

标准化随机变量

随机变量 X 的标准化随机变量定义为

$$X^* = \frac{X-EX}{\sqrt{Var(X)}}$$

特点:$EX^* = 0, Var(X^*) = 1$, 且两者分布图相同. 可以用于在比较不同随机变量时消除单位的影响而比较分布图

e.g. 标准正态分布

$X \sim N(\mu, \sigma^2), Y = \frac{X-\mu}{\sigma} \sim N(0,1)$

2. **矩 (Moments)**

定义:

随机变量 X,c 为常数,r 为正整数, 则 $E[(X-c)^r]$ 称为 X 关于 c 点的 r 阶矩

依据 c 的不同, 主要分为两种:

(a) $c=0$, 记作 $a_r=EX^r$, 称为 X 的 r 阶原点矩

(b) $c=EX$, 记作 $\mu_r=E[(X-EX)^r]$, 称为 X 的 r 阶中心矩

偏度系数

$$\gamma_1=E[(\frac{X-\mu}{sigma})^3]=\frac{\mu_3}{sigma^3}$$

峰度系数

$$\gamma_2=E[(\frac{X-\mu}{\sigma})^4]=\frac{\mu_4}{\sigma^4}$$

3. **协方差 (Covariance)**

定义

$$\begin{aligned}Var(X+Y)&=E((X+Y)-E(X+Y))^2\\&=E((X-EX)+(Y-EY))^2\\&=E(X-EX)^2+E(Y-EY)^2+2E((X-EX)(Y-EY))\\&=Var(X)+Var(Y)+2Cov(X,Y)\\Cov(X,Y)&=E((X-EX)(Y-EY))\\&=E(XY)-(EX)(EY)\end{aligned}$$

协方差有正有负有零这里 XY 可以认为是随机向量 (X,Y) 的函数, 也可以利用全期望公式求得

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

性质

(a) $$Cov(X,Y)=E(XY)-(EX)(EY)$$

若 X,Y 相互独立,$Cov(X,Y)=0$

(b) $$Cov(X_1+X_2,Y)=Cov(X_1,Y)+Cov(X_2,Y)$$

证明:

$$\begin{aligned}Cov(X_1+X_2,Y)&=E((X_1+X_2)Y)-E(X_1+X_2)(EY)\\&=E(X_1Y+X_2Y)-(EX_1+EX_2)(EY)\\&=E(X_1Y)+E(X_2Y)-(EX_1)(EY)-(EX_2)(EY)\\Cov(X_1,Y)+Cov(X_2,Y)&=E(X_1Y)-(EX_1)(EY)+E(X_2Y)-(EX_2)(EY)\end{aligned}$$

(c) $$Cov(\sum_{i=1}^m a_iX_i,\sum_{j=1}^n b_jY_j)=\sum_{0\le i\le m,0\le j\le n}a_ib_jCov(X_i,Y_j)$$

借助上面的结论可证

协方差矩阵

若 $\xi_1 \ldots \xi_n$ 为定义在同一概率空间下的随机变量，且都平方可积 (方差存在). 记 $\boldsymbol{\xi} = (\xi_1, \ldots, \xi_n)$, 称矩阵

$$
\begin{aligned}
\Sigma &= E[(\boldsymbol{\xi} - \bar{\boldsymbol{\xi}})(\boldsymbol{\xi} - \bar{\boldsymbol{\xi}})^\top] \\
&= [b_{ij}] = [Cov(\xi_i, \xi_j)] \\
&= \begin{bmatrix} Var(\xi_1) & Cov(\xi_1, \xi_2) & \cdots & Cov(\xi_1, \xi_n) \\ Cov(\xi_2, \xi_1) & Var(\xi_2) & \cdots & Cov(\xi_2, \xi_n) \\ \vdots & \vdots & \ddots & \vdots \\ Cov(\xi_n, \xi_1) & Cov(\xi_n, \xi_2) & \cdots & Var(\xi_n) \end{bmatrix}
\end{aligned}
$$

为 $\boldsymbol{\xi} = (\xi_1, \ldots, \xi_n)$ 的协方差矩阵 **性质**

(a) $\Sigma \geq 0$

$$
\begin{aligned}
&\forall \boldsymbol{u} \in R^n, \text{令} c =< \boldsymbol{u}, \boldsymbol{\xi} > \\
&\quad \boldsymbol{u}^\top \Sigma \boldsymbol{u} = \boldsymbol{u}^\top \boldsymbol{\xi} \boldsymbol{\xi}^\top \boldsymbol{u} \\
&\quad\quad = c^2 \geq 0
\end{aligned}
$$

(b) 设 $\boldsymbol{Z} = (X, Y) \sim N(a, b, \sigma_X^2, \sigma_Y^2, \rho)$, 则关于 $\boldsymbol{Z} = (X, Y)$ 的协方差矩阵有

$$
\begin{aligned}
\Sigma &= \begin{bmatrix} \sigma_X^2 & \rho\sigma_X\sigma_Y \\ \rho\sigma_X\sigma_Y & \sigma_Y^2 \end{bmatrix} \\
f_Z(x, y) &= \frac{1}{2\pi\sigma_X\sigma_Y\sqrt{1-\rho^2}} e^{-\frac{1}{2(1-\rho^2)}[\frac{(x-a)^2}{\sigma_X^2} - 2\rho\frac{(x-a)(y-b)}{\sigma_X\sigma_Y} + \frac{(y-b)^2}{\sigma_Y^2}]} \\
&= \frac{1}{(2\pi)^{k/2}\sqrt{|\Sigma|}} \exp -\frac{1}{2}\boldsymbol{z}^\top \Sigma^{-1} \boldsymbol{z} \\
\boldsymbol{z} &= (x-a, y-b)^\top, k = 2
\end{aligned}
$$

2 大数定理与中心极限定理

符号	意义	
a.s.	almost sure	几乎处处收敛
iid	Independent Identically Distribution	独立同分布
$\xrightarrow{d}$		依分布收敛

样本均值趋于总体均值

2.1 强大数定理

$\overline{X} \xrightarrow{a.s.} \mu$

2.2 (弱) 大数定理

$\overline{X} \xrightarrow{P} \mu$

题 1： 如何理解“弱”

用随机变量 U 表示一位专职司机在一个工作日内因交通事故造成的损失，U 取值越大说明损失越大，$U=0$ 表示没有造成实际损失. 对于正数 ε，用 $U \geq \varepsilon$ 表示造成较大的损失. 对于一位优秀的老司机来讲，假设他的 U 已经很小. 为方便，假设他的 $U=0$.

设甲某是一位新的专职司机，用 U_n 表示他在第 n 个工作日的交通事故造成的损失. 因为他的开车经验在不断提高，所以随着时间的推移，他的 U_n 会向**老司机**的 $U=0$ 收敛.

如果 $U_n \xrightarrow{P} U$, 则 $n \to \infty$ 时，我们只能得到

$$P(U) = P(|U_n - U| \geq \varepsilon) \to 0. \tag{6}$$

所以对任意大的 n，都不能保证 $P(U_n \geq \varepsilon) = 0$.

也就是说，无论有多长的开车经验，这位新司机因交通事故造成较大损失的概率都是正数，从而都有可能造成较大的损失.

用 u_n 表示 U_n 的观测值，如果 $U_n \xrightarrow{a.s.} U$

说明存在 n_0, 使得 $n \geq n_0, u_n < \varepsilon$. 也就是说，从某天开始

这位新司机就再也不会发生有较大损失的交通事故了.

2.2.1 切比雪夫不等式

$$\lim_{n\to\infty} P\left(\left|\overline{X} - \mu\right| \geq \varepsilon\right) \leq \frac{1}{\varepsilon^2} Var(X) \tag{7}$$

$$\begin{aligned} &X_1, \cdots, X_n\ i.i.d. \sim X, EX = \mu, \forall \varepsilon > 0 \\ &\qquad \lim_{n\to\infty} P\left(\left|\overline{X} - \mu\right| \geq \varepsilon\right) = 0 \end{aligned} \tag{8}$$

不需要方差存在, 期望存在即可

一般样本会是多个独立同分布的随机变量, 可以应用大数定理

2.3 中心极限定理

部分和的分布会趋向标准正态分布，依分布收敛

1. 一般中心极限定理

$$X_1, \cdots, X_n\ i.i.d. \sim X, EX = \mu, Var X = \sigma^2$$

$$\lim_{x\to\infty} P\left(\sqrt{n}\frac{\overline{X} - \mu}{\sigma} \leq x\right) = \Phi(x)$$

也记作

$$\sqrt{n}\frac{\overline{X} - \mu}{\sigma} \xrightarrow{d} N(0,1)$$

2. 针对两点分布的中心极限定理

$$X_1,\cdots,X_n\ i.i.d. \sim B(1,p)$$
$$\lim_{x\to\infty} P\left(\sqrt{n}\frac{\overline{X}-p}{\sqrt{p(1-p)}} \le x\right) = \Phi(x)$$

3. 针对二项分布的中心极限定理

$$\begin{aligned}
X &\sim B(n,p)\\
\lim_{x\to\infty}\frac{X-EX}{\sqrt{VarX}} &\sim N(0,1)\\
P(t_1 < X \le t_2) &= P(\frac{t_1-EX}{\sqrt{VarX}} < \frac{X-EX}{\sqrt{VarX}} \le \frac{t_2-EX}{\sqrt{VarX}})\\
&\approx \Phi(y_2)-\Phi(y_1),\\
y_2 &= \frac{t_2-EX}{\sqrt{VarX}},\\
y_1 &= \frac{t_1-EX}{\sqrt{VarX}}
\end{aligned}$$

4. 修正

$$\begin{aligned}
y_2 &= \frac{t_2+1/2-EX}{VarX}\\
y_1 &= \frac{t_1-1/2-EX}{VarX}
\end{aligned}$$

5. 针对泊松分布的中心极限定理

$$\begin{aligned}
X &\sim P(\lambda(n))\\
\lim_{n\to\infty}\frac{X-\lambda(n)}{\sqrt{\lambda(n)}} &\sim N(0,1)\\
\lambda(n) &\sim n
\end{aligned}$$

经验分布函数

$$\begin{aligned}
F_n(x) &= \sum_{k=1}^{n}\frac{I_{(X_i\le x)}}{n}\\
\lim_{n\to\infty}F_n(X) &= \lim_{n\to\infty}\sum_{k=1}^{n}\frac{I_{(X_i\le x)}}{n}\\
&= EI_{(X_i\le x)}\\
&= F(x)
\end{aligned}$$

2.4 三大分布

1. 卡方分布 (chi-square distribution)

$\boxed{\chi^2 = X_1^2+X_2^2+\cdots+X_n^2}$

$X_1^2,X_2^2,\cdots,X_n^2$ 是来自标准正态总体 $N(0,1)$ 的样本主要性质:

(a) 可加性;

(b) 若 $\chi^2 \sim \chi^2(n)$, 则 $E(\chi^2) = n, D(\chi^2) = 2n$;

(c) **上侧分位点**

设 $X \sim \chi^2(n)$ 对给定的正数 $\alpha(0 < \alpha < 1)$

称满足 $P\left(X > \chi^2_\alpha(n)\right) = \alpha$ 的点 $\chi^2_\alpha(n)$ 为 χ^2 分布的 上α分位点

例如 $\chi^2_{0.1}(25) = 34.382$

2. t 分布 (student' s dtstribution)

$\boxed{T = \dfrac{X}{\sqrt{Y/n}}}$ 记作 $T \sim t(n)$

其中 $X \sim N(0,1), Y \sim \chi^2(n)$.

(a) 概率密度函数 $f(t)$ 是偶函数, 关于纵轴对称分布;

(b) $n > 45$ 时,$T \xrightarrow{\text{近似}} N(0,1)$, 因此 $t_\alpha(n) \approx z_\alpha$;

(c) 若 $t \sim t(n)$, 则 $E(t) = 0, D(t) = \frac{n}{n-2}(n > 2)$ $n = 1$ 是柯西分布

(d) **上侧分位点**

设 $t \sim t(n)$ 对给定的正数 $\alpha(0 < \alpha < 1)$

称满足 $P(X > t_\alpha(n)) = \alpha$ 的点 $t_\alpha(n)$ 为 t 分布的 上α分位点

由对称性，有 $\boxed{t_{1-\alpha}(n) = -t_\alpha(n)}$

3. F 分布 (F-dtstribution)

$\boxed{F = \dfrac{U/m}{Y/n}}$ 记作 $T \sim F(m,n)$

其中 $X \sim \chi^2(m), Y \sim \chi^2(n)$.

(a) $\frac{1}{F} \sim F(n,m)$;

(b) $t^2 \sim F(1,n)$;

(c) **上侧分位点**

设 $F \sim F(m,n)$ 对给定的正数 $\alpha(0 < \alpha < 1)$

称满足 $P(X > F_\alpha(m,n)) = \alpha$ 的点 $F_\alpha(m,n)$ 为 F 分布的 上α分位点

有 $\boxed{F_{1-\alpha}(m,n) = \dfrac{1}{F_\alpha(n,m)}}$

3 估计

3.1 点估计

1. 矩估计

适合非参数分布

能用低阶矩不用高阶矩

2. 极大似然估计

只适合参数分布

有 n 个参数时有 n 个方程, 用已有估计量代替未知量

3. 柯西分布使用中位数估计

4. 点估计的优良性准则

   (a) 特性: 无偏性
   对于不同的参数, 估计量的期望总等于参数真值

$$
\begin{aligned}
ES^2 &= \sigma^2 \\
ES &\neq \sigma \\
\sigma^2 &= E(S^2) \\
&= Var(s) + (ES)^2 \\
\therefore E(c_n S) &= c_n ES = \sigma \\
c_n &> 1
\end{aligned}
$$

   对于正态总体

$$
c_n = \sqrt{\frac{n-1}{2}}\Gamma(\frac{n-1}{2})/\Gamma(\frac{n}{2})
$$

   (b) 数量性指标: 有效性: 均方误差小为佳
   例子:
   在期望 μ 已知的条件下, 作为对于 σ^2 的估计,
   $\sum_{k=1}^{n}\frac{(X_k-\mu)}{n}$ 比 $\sum_{k=1}^{n}\frac{(X_k-\bar{X})}{n-1}$ 更优

3.2 区间估计

给定一个很小的数 $\alpha > 0$,
n 个样本 $X_1, \cdots, X_n$,
得到两个统计量 $\hat{\theta}_1(X_1,\cdots,X_n), \hat{\theta}_2(X_1,\cdots,X_n)$,
$P_\theta(\hat{\theta}_1(X_1,\cdots,X_n) < \theta \leq \hat{\theta}_2(X_1,\cdots,X_n)) = 1-\alpha$,
且区间长度尽可能小
则得到了对于 θ 的一个区间估计,
此区间估计的置信系数/置信水平为 $1-\alpha$

枢轴变量法

常用分布

1. 卡方分布 $X = \sum_{i=1}^{n} X_i^2 \sim \chi_n^2$
   其中 $X_i \sim N(0,1)$

2. t 分布 $X = \frac{X_1}{\sqrt{X_2/n}} \sim t_n$
   其中 X_1, X_2 独立, $X_1 \sim N(0,1), X_2 \sim \chi_n^2$

3. F 分布 $X = \frac{X_1/m}{X_2/n} \sim F_{m,n}$
   其中 $X_1 \sim \chi_m^2, X_2 \sim \chi_n^2$

上分位数

1. 正态

$$u_\beta, \Phi(u_\beta) = 1-\beta$$

2. 卡方分布

$$\chi^2_n(\beta)$$

3. t 分布

$$t_n(\beta)$$

4. F 分布

$$F_{n,m}(\beta)$$

常用枢轴变量

1. 样本来自一个正态总体

$$\text{samples } X_1,\cdots,X_n \ i.i.d. \sim N(\mu,\sigma^2)$$

× 表示未知,√ 表示已知

μ ×,σ √, 估计 μ

$$\sqrt{n}\frac{\bar{X}-\mu}{\sigma} \sim N(0,1)$$

μ ×,σ ×, 估计 μ

$$T = \frac{\sqrt{n}(\bar{X})-\mu}{S} \sim t_{n-1}$$

μ √,σ ×, 估计 σ

$$\sum_{k=1}^{n}(\frac{X_i-\mu}{\sigma})^2 \sim \chi^2_n$$

μ ×,σ ×, 估计 σ

$$\frac{(n-1)S^2}{\sigma^2} = \sum_{k=1}^{n}(\frac{X_i-\bar{X}}{\sigma})^2 \sim \chi^2_{n-1}$$

2. 样本来自两个正态总体

$$\text{samples } X_1,\cdots,X_n \ i.i.d. \sim N(\mu_X,\sigma^2_X)$$
$$Y_1,\cdots,Y_m \ i.i.d. \sim N(\mu_Y,\sigma^2_Y)$$

μ_X, μ_Y 都已知, 方差 $\sigma_X^2 = \sigma_Y^2 = \sigma^2$ 相等且未知, 估计 $\mu_X - \mu_Y$

$$\bar{X}_n \sim N(\mu_1, \frac{\sigma_1^2}{n}), \bar{Y}_m \sim N(\mu_2, \frac{\sigma_2^2}{m})$$
$$\bar{X}_n - \bar{Y}_m \sim N(\mu_1 - \mu_2, \frac{\sigma_1^2}{n} + \frac{\sigma_2^2}{n})$$
$$Z = \frac{(\bar{X}_n - \bar{Y}_m) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n} + \frac{\sigma_2^2}{n}}} \sim N(0, 1.)$$

μ_X, μ_Y 都未知, 方差 $\sigma_X^2 = \sigma_Y^2 = \sigma^2$ 相等且未知, 估计 $\mu_X - \mu_Y$

$$S^2 = \frac{\sum_{i=1}^n (X_i - \bar{X}) + \sum_{j=1}^m (Y_j - \bar{Y})^2}{\sigma^2} \sim \chi^2_{n+m-2}$$
$$Z = \frac{(\bar{X} - \bar{Y}) - (\mu_X - \mu_Y)}{\sigma\sqrt{\frac{1}{n} + \frac{1}{m}}} \sim N(0, 1)$$
$$\therefore T = \frac{Z}{\sqrt{S^2/(n+m-2)}} \sim t_{n+m-2}$$
$$= \frac{(\bar{X} - \bar{Y}) - (\mu_X - \mu_Y)}{\sqrt{(\sum_{i=1}^n (X_i - \bar{X})^2 + \sum_{j=1}^m (Y_j - \bar{Y})^2)(n+m-2)}} \sqrt{\frac{nm}{n+m}} \sim t_{n+m-2}$$

$\mu_X, \mu_Y, \sigma_X^2, \sigma_Y^2$ 都未知, 方差不一定相等, 估计 $\frac{\sigma_X^2}{\sigma_Y^2}$

$$\frac{(n-1)S_X^2}{\sigma_X^2} \sim \chi^2_{n-1}$$
$$\frac{(m-1)S_Y^2}{\sigma_Y^2} \sim \chi^2_{m-1}$$
$$\therefore F = \frac{\frac{(n-1)S_X^2}{\sigma_X^2}/(n-1)}{\frac{(m-1)S_Y^2}{\sigma_Y^2}/(m-1)} \sim F_{n-1,m-1}$$
$$\frac{S_X^2}{S_Y^2}\frac{\sigma_Y^2}{\sigma_X^2} \sim F_{n-1,m-1}$$

3. 样本来自指数分布总体

$$\text{samples } X_1, \cdots, X_n \ i.i.d. \sim \exp(\lambda)$$
$$f(x, \lambda) = \lambda e^{-\lambda x} I_{[x>0]}$$
$$2\lambda X_i \sim \chi_2^2$$
$$\therefore 2\lambda n\bar{X} = 2\lambda \sum_{i=1}^n X_i \sim \chi^2_{2n}$$

大样本法 对于很多上面没有列举的总体, 特别是离散型总体, 则可以考虑极限分布, 利用中心极限定理建立枢轴变量. 下面举一些例子

1. 两点分布, 估计 p

$$\text{samples } X_1,\cdots,X_n\ i.i.d. \sim Ber(p)$$
$$\lim_{n\to\infty}\sqrt{n}\frac{\bar{X}-p}{\sqrt{p(1-p)}}\sim N(0,1)$$

出于计算方便, 可用 $\hat{p}=\bar{X}$ 代替分母中的 p

$$\lim_{n\to\infty}\sqrt{n}\frac{\bar{X}-p}{\sqrt{\hat{p}(1-\hat{p})}}\sim N(0,1)$$

2. 泊松分布, 估计 λ

$$\text{samples } X_1,\cdots,X_n\ i.i.d. \sim P(\lambda)$$
$$\lim_{n\to\infty}\sqrt{n}\frac{\bar{X}-\lambda}{\sqrt{\lambda}}\sim N(0,1)$$

3. 总体分布图未知, 期望 θ、方差 σ^2 存在但未知, 做期望 θ 的区间估计

$$\lim_{n\to\infty}\sqrt{n}\frac{\bar{X}-\theta}{\sigma}\sim N(0,1)$$

由于 σ 未知, 无法直接作为枢轴变量. 由于样本均方差 S 是 σ 的相合估计,$n\to\infty$ 时可近似用 S 代替 σ

$$\lim_{n\to\infty}\sqrt{n}\frac{\bar{X}-\theta}{S}\sim N(0,1)$$

4. 贝伦斯-费歇尔问题

$$\begin{aligned}\text{samples } X_1,\cdots,X_n\ i.i.d. &\sim N(\mu_X,\sigma_X^2)\\ Y_1,\cdots,Y_m\ i.i.d. &\sim N(\mu_Y,\sigma_Y^2)\end{aligned}$$

$\mu_X,\mu_Y,\sigma_X^2\sigma_Y^2$ 未知, 方差不一定相等, 估计 $\mu_X-\mu_Y$

$$\frac{(\bar{X}-\bar{Y})-(\mu_X-\mu_Y)}{\sqrt{\sigma_X^2/n+\sigma_Y^2/m}}\sim N(0,1)$$

次式严格成立, 但同样由于 $\sigma_X^2\sigma_Y^2$ 未知而无法作为枢轴变量. 由于样本方差 S^2 是 σ^2 的无偏估计,$n\to\infty$ 时可近似用 S^2 代替 σ^2

$$\frac{(\bar{X}-\bar{Y})-(\mu_X-\mu_Y)}{\sqrt{S_X^2/n+S_Y^2/m}}\sim N(0,1)$$

4 假设检验

4.1 问题提法和基本概念

原假设与对立假设
一般来说, 待检验命题作为对立假设, 才可以以较大的把握肯定
但对于形如" 什么什么相等"" 什么等于多少" 的命题 (如方差检验), 则将" 什么等于什么" 作为原假设 (出于可行性)

功效函数 区分不同的检验法的优劣
一个给定的检验法, 在不同的未知参数下, 都会给出一个原假设被拒绝的概率 (源于样本是随机变量)
此时该概率为未知参数的函数, 即为功效函数
当未知参数处于原假设时我们希望该概率小以避免第一类错误, 并且一个满足要求的检验法应确保该概率在此时小于等于检验的水平 α
当未知参数处于对立假设时, 我们则希望该概率尽可能得高, 以避免第二类错误

重要参数检验 假设检验临界值, 借助枢轴变量构造功效函数 (含检验临界值与未知参数, 不含统计量), 类比区间估计, 得到包含已知枢轴变量上分位数、已知原假设临界值和待定检验临界值的方程, 解出检验临界值即可.
所依赖枢轴变量与区间估计相同

拟合优度检验 检验随机变量是否服从假设的分布
只提零假设不提对立假设

离散总体情形

1. 理论总体不含未知参数 (所有参数都有假设值) 时
   X 为离散型随机变量, 零假设:
   $$H_0:\ P(X=a_i)=p_i\ \ ,i=1,\cdots,k$$
   经过观测, 得到 n 个样本, 其中各个取值观测频数 $n_i,\cdots,n_k$, 检验统计量:
   $$T=\sum_{i=1}^{k} np_i(\frac{n_i-np_i}{np_i})^2=\sum_{i=1}^{k}\frac{(n_i-np_i)^2}{np_i}=\sum\frac{(O-E)^2}{E}\sim\chi^2_{k-1}$$
   其中 O 代表观测值,E 代表预测值
   拒绝域
   $$T>\chi^2(\alpha)$$
2. 理论总体分布含未知参数 (有的参数需要你估) 时
   修正
   $$T\sim\chi^2_{k-1-r}$$
   其中 r 为需要你估计的独立参数的个数

列联表的独立性和齐一性检验

1. 独立性
   原假设: 两属性独立.
   基于此估计参数、根据独立性乘得预测值、检验原假设即可.

2. 齐一性
   原假设: 不同条件 (某一属性的不同水平) 下另一属性条件分布相同.
   方法与独立性检验无异

3. 两者比较
   两种检验区别在于两个属性的地位相同与否, 但检验方法无异

连续总体情形 离散化总体分布, 化为离散总体情形, 进行检验