一、总体与样本

1、总体：全体　10万

有限总体，无限总体

不一定是无限

X：总体分布

样本：抽样 $(X_1, X_2, \cdots, X_n)$　$(x_1, x_2, \cdots, x_n)$

变量　　观测值

简单随机抽样

(1) 同分布　(2) 独立

2、样本的分布 $(X_1, X_2, \cdots, X_n)$

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

$P(X_1 = x_1 \cdots)$

$f(x_1, \cdots)$

二、统计量

1、Def：不含任何未知参数的样本的函数

构造

$X_1 + X_2 + X_3$ ✓

$\frac{X_1 + X_2 + \cdots + X_n}{n}$ ✓

2、常见统计量

(1) 样本均值：$\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$

(2) 未修正的样本方差：$S_0^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})^2$

$S^2 = \frac{n}{n-1} S_0^2$

(3) (修正的) 样本方差：$S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2$

(4) 标准差：　$S = \sqrt{S^2}$

(5) K阶原点矩：$A_k=\frac{1}{n}\sum_{i=1}^{n}X_i^k$ $\quad A_1=\overline{X}$

(6) K阶中心矩：$B_k=\frac{1}{n}\sum_{i=1}^{n}(X_i-\overline{X})^k$ $\quad B_2=S_0^2$

3、样本均值和样本方差的性质

设总体X的均值$EX=\mu$，$DX=\sigma^2$，样本$(X_1,\cdots,X_n)$来自X，

则 (1) $E\overline{X}=\mu$ (2) $D\overline{X}=\frac{1}{n}\sigma^2$ (3) $ES^2=\sigma^2$

证(3)：

$$\sum_{i=1}^{n}(X_i-\overline{X})^2=\sum_{i=1}^{n}[(X_i-\mu)-(\overline{X}-\mu)]^2$$
$$=\sum_{i=1}^{n}[(X_i-\mu)^2-2(X_i-\mu)(\overline{X}-\mu)+(\overline{X}-\mu)^2]$$
$$=\sum_{i=1}^{n}(X_i-\mu)^2-2(\overline{X}-\mu)\sum_{i=1}^{n}(X_i-\mu)+\sum_{i=1}^{n}(\overline{X}-\mu)^2$$
$$=\sum_{i=1}^{n}(X_i-\mu)^2-2(\overline{X}-\mu)(\sum_{i=1}^{n}X_i-\sum_{i=1}^{n}\mu)+n(\overline{X}-\mu)^2$$
$$=\sum_{i=1}^{n}(X_i-\mu)^2-2(\overline{X}-\mu)(n\overline{X}-n\mu)+n(\overline{X}-\mu)^2$$
$$=\sum_{i=1}^{n}(X_i-\mu)^2-2n(\overline{X}-\mu)^2+n(\overline{X}-\mu)^2$$
$$=\sum_{i=1}^{n}(X_i-\mu)^2-n(\overline{X}-\mu)^2$$

$$\sum_{i=1}^{n}(X_i-\overline{X})^2=\sum_{i=1}^{n}(X_i-\mu)^2-n(\overline{X}-\mu)^2$$
$$ES^2=E[\frac{1}{n-1}\sum_{i=1}^{n}(X_i-\overline{X})^2]$$
$$=\frac{1}{n-1}E[\sum_{i=1}^{n}(X_i-\mu)^2-n(\overline{X}-\mu)^2]$$
$$=\frac{1}{n-1}\{E[\sum_{i=1}^{n}(X_i-\mu)^2]-nE(\overline{X}-\mu)^2\}$$
$$=\frac{1}{n-1}[\sum_{i=1}^{n}E(X_i-\mu)^2-nD\overline{X}]$$
$$=\frac{1}{n-1}[\sum_{i=1}^{n}DX_i-nD\overline{X}]$$
$$=\frac{1}{n-1}(n\sigma^2-n\frac{1}{n}\sigma^2)=\sigma^2$$

三、抽样分布

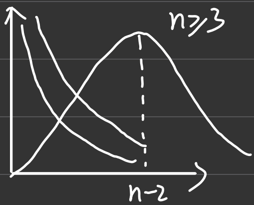
1、χ^2分布（卡方分布）

自由度

(1) def: $X_1, X_2, \cdots, X_n$ 独立，$X_i \sim N(0,1)$，则 $\sum_{i=1}^{n} X_i^2 = \chi^2(n)$

$E\chi^2 = n \quad D\chi^2 = 2n$

(2) 图像：



解释：$\chi^2(2) \sim Exp(\frac{1}{2})$

n很大时，可用正态分布近似

(3) 正态近似

由中心极限定理，$X \sim \chi^2(n)$，n充分大时，$\frac{X-n}{\sqrt{2n}} \sim N(0,1)$

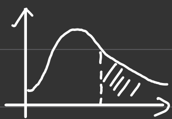
(4) 可加性

① $X \sim \chi^2(n)$，$Y \sim \chi^2(m)$，X, Y独立，则 $X+Y \sim \chi^2(m+n)$

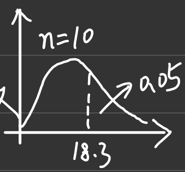
② $X_i \sim \chi^2(m_i)$，独立，$\sum_{i=1}^{n} X_i \sim \chi^2(\sum_{i=1}^{n} m_i)$

(5) 上α分位数

$P(\chi^2 > \underbrace{\chi^2_\alpha(n)}_{点}) = \underbrace{\alpha}_{面积}$



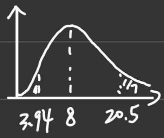
例1: $\underbrace{\chi^2_{0.05}(10) = 18.3}_{要求的数}$



例2：$X\sim\chi^2(10)$，$P(X>a)=0.025$，$P(X<b)=0.05$，求a和b

解：$n=10$，$\alpha_a=0.025$，查表 $\chi^2_{0.025}(10)=20.5=a$

$\alpha_b=0.95$，查表 $\chi^2_{0.95}(10)=3.94=b$



例3：$X_1,\cdots,X_6\sim N(\overset{\mu}{0},\overset{\sigma^2}{2^2})$，求 $P(\sum_{i=1}^{6}X_i^2>6.54)$

解：化成标准正态：$\sum_{i=1}^{6}\left(\frac{X_i}{2}\right)^2\sim\chi^2(6)$

$$P\left(\sum_{i=1}^{6}\left(\frac{X_i}{2}\right)^2>\frac{6.54}{4}\right)=0.95$$

2、t分布　$X\sim t(n)$

(1) def:

$X\sim N(0,1)$，$Y\sim\chi^2(n)$，X,Y独立，则 $\frac{X}{\sqrt{Y/n}}\sim t(n)$

(2) 上α分位数　$P(T>t_\alpha(n))=\alpha$

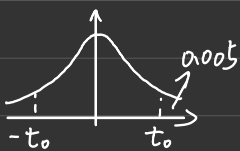
由于对称性　$t_{1-\alpha}(n)=-t_\alpha(n)$

例1：$X\sim N(2,1)$，$Y_1,\cdots Y_4\sim N(0,4)$，独立，令 $T=\frac{4(X-2)}{\sqrt{\sum_{i=1}^{4}Y_i^2}}$

求T的分布和 $P(|T|>t_0)=0.01$ 的 t_0.

解：$\frac{X-2}{1}\sim N(0,1)$，$\sum_{i=1}^{4}\left(\frac{Y_i}{2}\right)^2\sim\chi^2(4)$

$$T=\frac{\frac{X-2}{1}}{\sqrt{\sum_{i=1}^{4}\left(\frac{Y_i}{2}\right)^2/4}}\sim t(4)$$



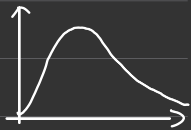
$t_{0.005}(4)=4.604$

3. F分布

(1) def.

$X \sim \chi^2(n_1)$, $Y \sim \chi^2(n_2)$, 独立, 则 $F \sim \dfrac{X/n_1}{Y/n_2} \sim F(n_1, n_2)$

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



(2) 上α分位数

$P(F > F_\alpha(n_1, n_2)) = \alpha$ $\quad F_{1-\alpha}(n_1, n_2) = \dfrac{1}{F_\alpha(n_2, n_1)}$

证: $F \sim F(n_1, n_2)$, $\dfrac{1}{F} \sim F(n_2, n_1)$

看成新变量

$$1-\alpha = P(F > F_{1-\alpha}(n_1, n_2)) = P\left(\frac{1}{F} < \frac{1}{F_{1-\alpha}(n_1, n_2)}\right) = 1 - \underbrace{P\left(\frac{1}{F} > \frac{1}{F_{1-\alpha}(n_1, n_2)}\right)}_{=\alpha}$$

3、正态分布下的抽样分布

总体是正态分布，抽样本，构造统计量

(1) $X \sim N(\mu, \sigma^2)$ $\quad \{X_1, X_2, \cdots, X_n\}$

① $\bar{X} \sim N(\mu, \frac{\sigma^2}{n})$, $\quad \dfrac{\bar{X} - \mu}{\sigma}\sqrt{n} \sim N(0, 1)$

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

$\bar{X} = \frac{1}{n}(X_1 + X_2 + \cdots + X_n)$ 多一个约束

③ $\bar{X}$ 与 S^2 独立

④ $\dfrac{1}{\sigma^2}\sum_{i=1}^{n}(X_i - \mu)^2 = \sum_{i=1}^{n}\left(\dfrac{X_i - \mu}{\sigma}\right)^2 \sim \chi^2(n)$

⑤ $\dfrac{\bar{X} - \mu}{S}\sqrt{n} = \dfrac{\frac{\bar{X} - \mu}{\sigma}\sqrt{n}}{\sqrt{\frac{(n-1)S^2}{\sigma^2}/n-1}} \sim t(n-1)$

(2) 两个正态总体

$X \sim N(\mu_1, \sigma_1^2)$ $\qquad\qquad$ $Y \sim N(\mu_2, \sigma_2^2)$

$\{X_1, \cdots, X_{n_1}\}$ $\qquad\qquad$ $\{Y_1, \cdots, Y_{n_2}\}$

① $$\frac{(\bar{X}-\bar{Y})-(\mu_1-\mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}} \sim N(0,1)$$

② $$\frac{S_1^2/\sigma_1^2}{S_2^2/\sigma_2^2} = \frac{\frac{(n_1-1)S_1^2}{\sigma_1^2}\Big/(n_1-1)}{\frac{(n_2-1)S_2^2}{\sigma_2^2}\Big/(n_2-1)} \sim F(n_1-1, n_2-1)$$

③ $\sigma_1^2=\sigma_2^2=\sigma^2$ 时 (两个正态总体方差相等)

$$T = \frac{(\bar{X}-\bar{Y})-(\mu_1-\mu_2)}{S_w\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}} \sim t(n_1+n_2-2), \quad S_w^2 = \frac{(n_1-1)S_1^2+(n_2-1)S_2^2}{n_1+n_2-2}$$