§6.2 常用统计分布

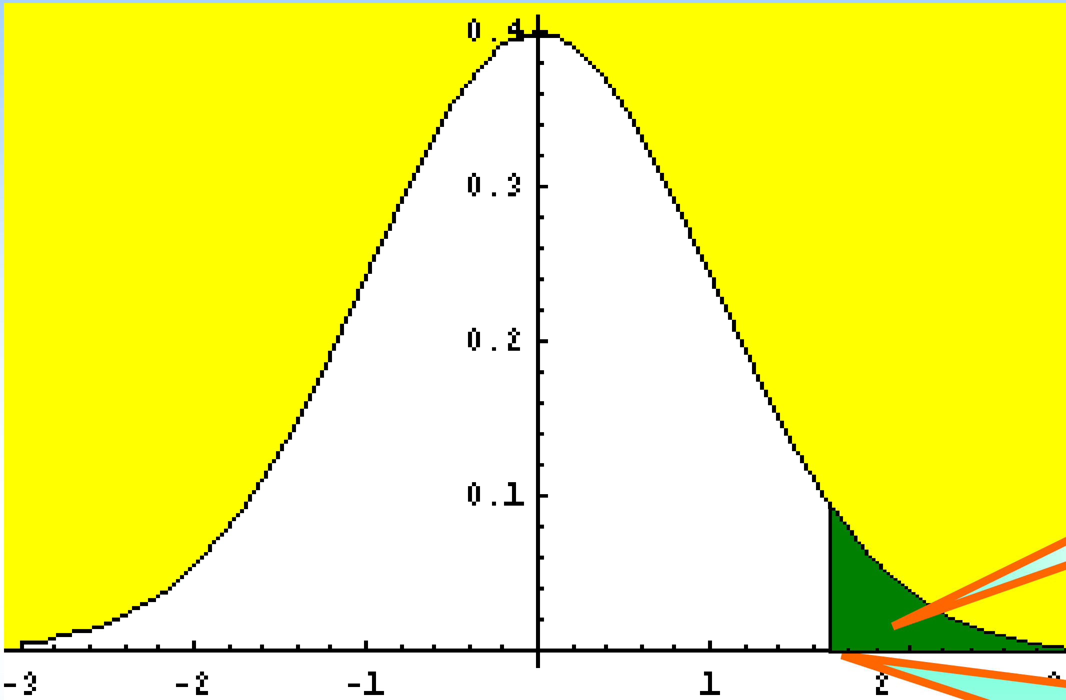
一、四种常用统计分布

1. 标准正态分布

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

上侧分位数u_α(0< α<1）满足

$$P\{X>u_\alpha\}=\int_{u_\alpha}^{+\infty}f(x)dx=\alpha$$



阴影部分面积为α

上侧分位数u_α

对于正态分布有: $\boxed{\Phi(u_\alpha)=1-\alpha}$

$$\because P\{X>u_\alpha\}=1-P\{X\le u_\alpha\}=1-\Phi(u_\alpha)=\alpha$$

查表　　如 $\alpha = 0.025$ 时， $u_\alpha = ?$

$$\Phi(u_{0.025}) = 1 - 0.025 = 0.975$$

$$u_{0.025} = 1.96$$

α分位数例子

2. χ^2 (卡方)分布

$$f_{\chi^2}(x) = \begin{cases} \dfrac{1}{2\Gamma(\frac{n}{2})}\left(\dfrac{x}{2}\right)^{\frac{n}{2}-1} e^{-\frac{x}{2}}, & x > 0 \\ 0, & x \le 0 \end{cases},$$

其中$\Gamma(\alpha)$为Gama函数，**称随机变量X服从自由度为n的χ^2分布，记为** $\chi^2 \sim \chi^2(n)$

Gama函数为

$$\Gamma(\alpha)=\int_0^{+\infty} x^{\alpha-1}e^{-x}dx, \quad (\alpha>0)$$

主要性质:

$$\Gamma(1)=1, \quad \Gamma\left(\frac{1}{2}\right)=\sqrt{\pi}, \quad \Gamma(\alpha)=(\alpha-1)\Gamma(\alpha-1)$$

例1 若$X \sim N(0, 1)$， 则 $Y = X^2$ 的概率密度为

$$f_Y(y) = \begin{cases} \dfrac{1}{\sqrt{2\pi}} y^{-\frac{1}{2}} e^{-\frac{y}{2}}, & y > 0; \\ 0, & y \le 0. \end{cases}$$

教材
例3.4.5

$$= \begin{cases} \dfrac{1}{2\Gamma(\frac{1}{2})} (\frac{y}{2})^{\frac{1}{2}-1} e^{-\frac{y}{2}}, & y > 0; \\ 0, & y \le 0. \end{cases}$$

即$Y = X^2$ 服从$\chi^2(1)$分布.

例2 若X_1，X_2相互独立, $X_1 \sim N(0, 1)$, $X_2 \sim N(0,1)$, 则$Y = X_1^2 + X_2^2$ 有

$$f_Y(y) = \begin{cases} \frac{1}{2} e^{-\frac{y}{2}}, & y > 0; \\ 0, & y \le 0. \end{cases}$$

$$= \begin{cases} \frac{1}{2\Gamma(1)} \left(\frac{y}{2}\right)^{\frac{2}{2}-1} e^{-\frac{y}{2}}, & y > 0; \\ 0, & y \le 0. \end{cases}$$

即$Y = X_1^2 + X_2^2 \sim \chi^2(2)$ 分布.

定理6.2.1 设 $X_1, X_2, \ldots, X_n$ 相互独立且都服从标准正态分布，则

标准正态随机变量的独立平方和

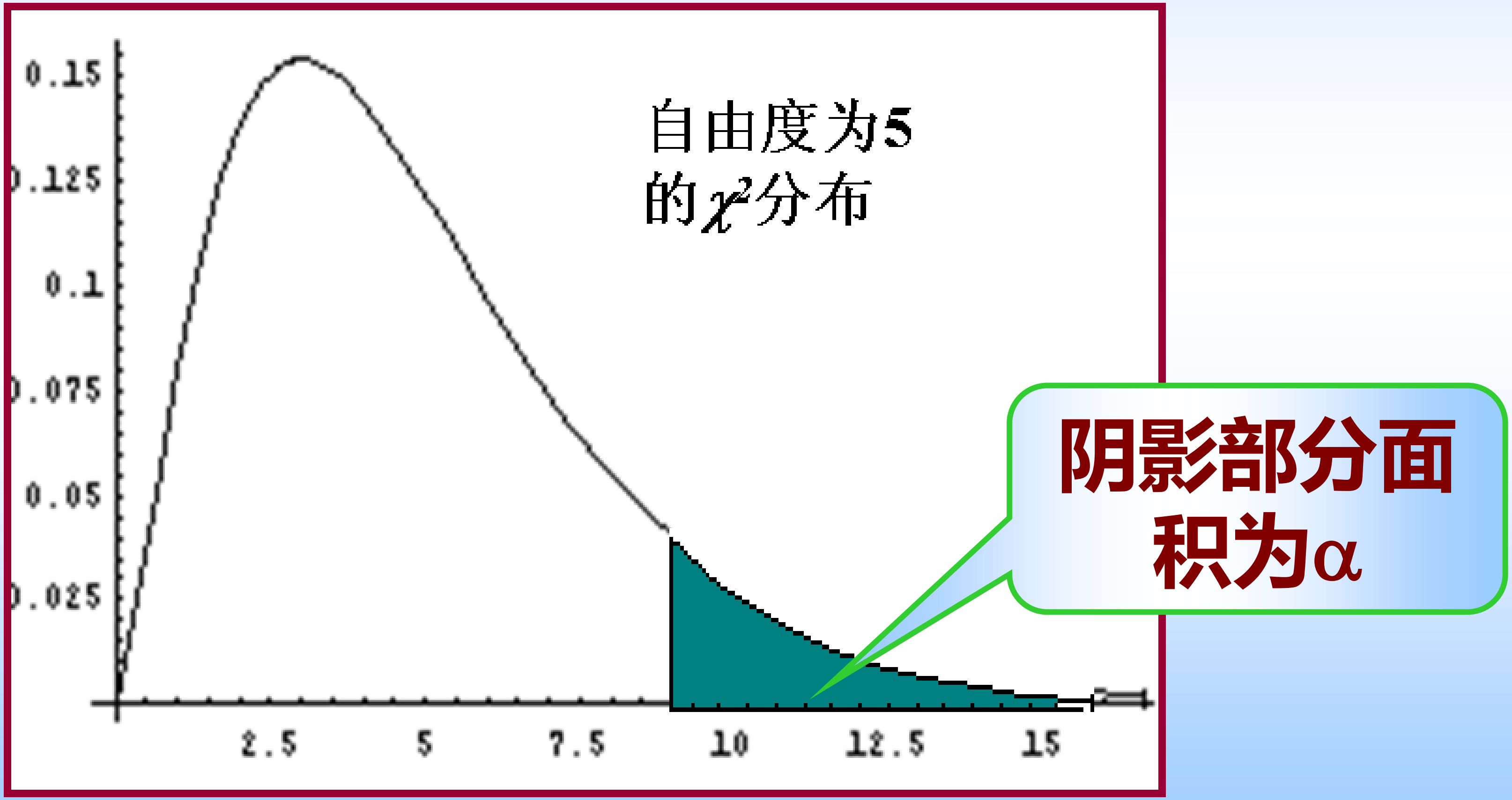
$$\chi^2 = \sum_{i=1}^{n} X_i^2 \sim \chi^2(n)$$

即随机变量 χ^2 服从自由度为 n 的卡方分布.

统计量的分布 (之一)

$\chi^2(n)$ 的上侧分位数($0< \alpha<1$):

$$P\{\chi^2 > \chi^2_\alpha(n)\} = \int_{\chi^2_\alpha(n)}^{+\infty} f_{\chi^2}(x)dx = \alpha$$



χ^2分布的三条性质：

性质1. (数字特征) 设 $\chi^2 \sim \chi^2(n)$，则有

$$E(\chi^2) = n, \qquad D(\chi^2) = 2n$$

证明

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

且 X_1，X_2，$\ldots$，X_n相互独立，$X_i \sim N(0,1)$，

$$E(\chi^2) = \sum_{i=1}^{n} E(X_i^2) = \sum_{i=1}^{n} D(X_i) = n,$$

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

$$=\sum_{i=1}^{n}\{E(X_i^4)-[E(X_i^2)]^2\}=2n.$$

性质2(可加性) 设Y_1，Y_2相互独立, 且$Y_1\sim\chi^2(n_1)$, $Y_2\sim\chi^2(n_2)$，则$Y_1+Y_2\sim\chi^2(n_1+n_2)$.

证明 记

$$Y_1=\sum_{i=1}^{n_1}X_i^2,\qquad Y_2=\sum_{i=n_1+1}^{n_1+n_2}X_i^2$$

则 $$Y_1+Y_2=\sum_{i=1}^{n_1+n_2} X_i^2$$

且X_i，$i=1,2,\ldots,n_1+n_2$相互独立，$X_i \sim N(0,1)$,

从而 $Y_1+Y_2 \sim \chi^2(n_1+n_2)$.

性质3.(大样本分位数 $\chi_\alpha^2(n)$) 当n 足够大(如$n>45$)时，有

$$\chi_\alpha^2(n) \approx n+u_\alpha\sqrt{2n}$$

证明

其中u_α满足$\Phi(u_\alpha)=1-\alpha$.

TIPS

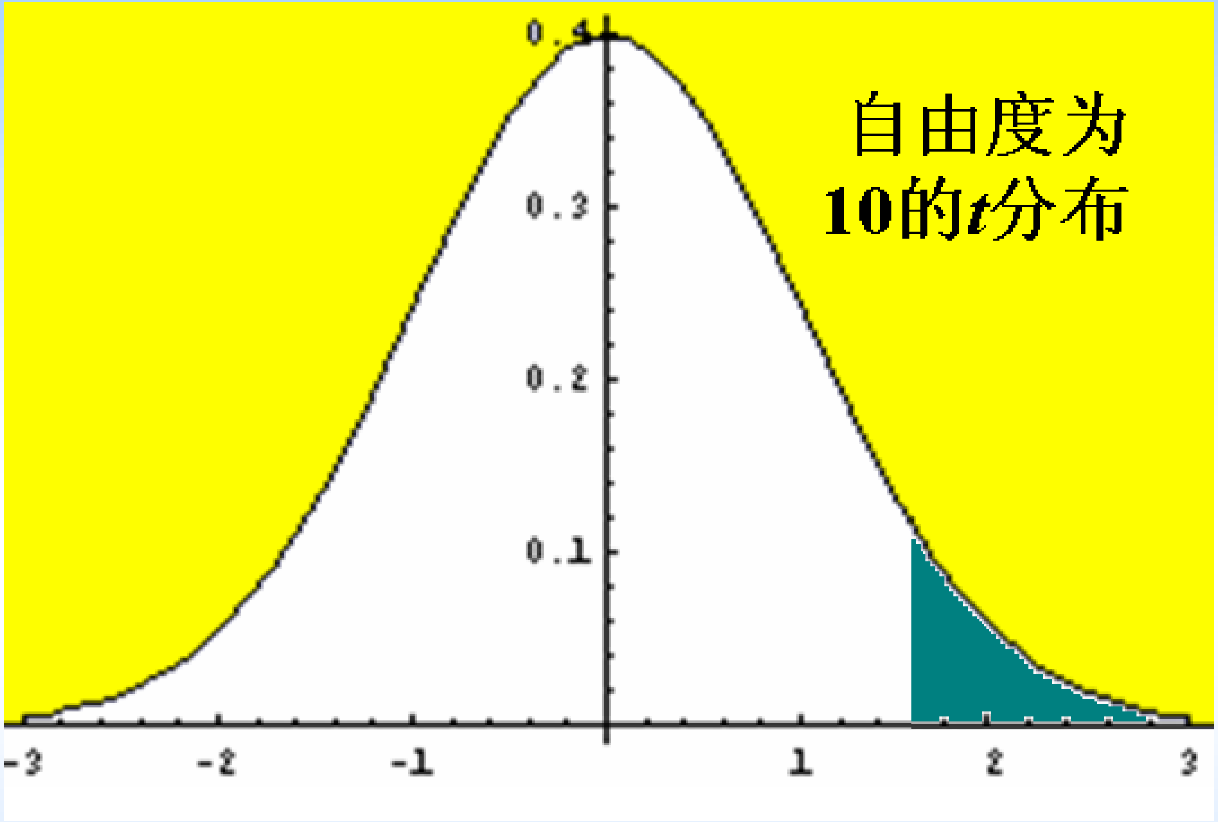
查表计算

3.自由度为 n的 t 分布 $T \sim t(n)$

又称学生氏分布--第一个研究者以*Student*作笔名发表文章.

$$f_T(x) = \frac{\Gamma(\frac{n+1}{2})}{\sqrt{n\pi}\,\Gamma(\frac{n}{2})}\left(1+\frac{x^2}{n}\right)^{-\frac{n+1}{2}}, \quad x \in R$$

$t(n)$ 的上侧分位数 $t_\alpha(n)$ （$0<\alpha<1$）:



$$P\{T > t_\alpha(n)\} = \int_{t_\alpha(n)}^{+\infty} f_T(x)dx = \alpha$$

定理6.2.2 设随机变量X, Y相互独立, $X\sim N(0,1)$，$Y\sim \chi^2(n)$，则

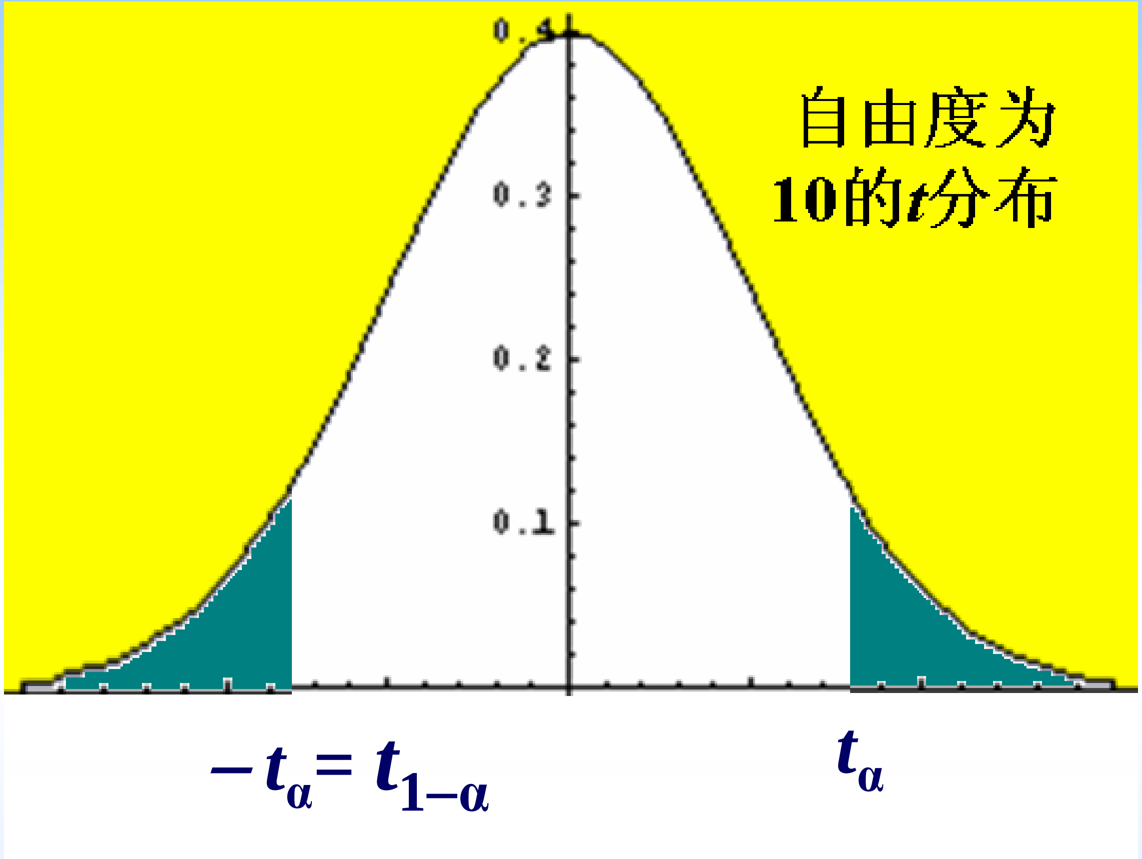
$$T=\frac{X}{\sqrt{Y/n}}\sim t(n)$$

即随机变量 T 服从自由度为 n 的 t 分布.

T 分布的特点:

1.关于纵轴对称:

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



因 $\alpha = P\{T > t_\alpha\} = P\{T \le -t_\alpha\} = 1 - P\{T > -t_\alpha\}$

故 $P\{T > -t_\alpha\} = 1 - \alpha.$ 即 $t_\alpha = -t_{1-\alpha}$

2. n 较大时，$\lim\limits_{n\to\infty} f_T(x) = \varphi(x)$.

$\Rightarrow \quad t_\alpha(n) \approx u_\alpha \quad (n > 30)$

例 查表计算：

$$t_{0.95}(20) = ? \qquad t_{0.95}(80) = ?$$

解 $t_{0.95}(20) = t_{1-0.05}(20) = -t_{0.05}(20) = -1.7247$

$$t_{0.95}(80) = -t_{0.05}(80) \approx -u_{0.05} = -1.645$$

4. F 分布　$F \sim F(n_1, n_2)$

$$f_F(x)=\begin{cases} n_1^{\frac{n_1}{2}} n_2^{\frac{n_2}{2}} \dfrac{\Gamma(\frac{n_1+n_2}{2})}{\Gamma(\frac{n_1}{2})\Gamma(\frac{n_2}{2})} x^{\frac{n_1}{2}-1}(n_1x+n_2)^{-\frac{n_1+n_2}{2}}, & x>0 \\ 0, & x\le 0 \end{cases}$$

称X 服从第一自由度为n_1,第二自由度为n_2的F分布.

定理6.2.3 设随机变量X，Y 相互独立，$X \sim \chi^2(n_1)$，$Y \sim \chi^2(n_2)$，则

$$F = \frac{X/n_1}{Y/n_2} \sim F(n_1, n_2)$$

即随机变量 F 服从第一自由度为n_1，第二自由度为n_2的F分布.

上侧分位数$F_\alpha(n_1, n_2)$ （$0 < \alpha < 1$）:

$$P\{F > F_\alpha(n_1, n_2)\} = \int_{F_\alpha(n_1,n_2)}^{+\infty} f_F(x)dx = \alpha$$

推论1:

$$\text{若}\quad F \sim F(n_1,n_2) \Rightarrow \frac{1}{F} \sim F(n_2,n_1)$$

推论2:

$$\text{若}\quad F \sim F(n_1,n_2) \Rightarrow F_{1-\alpha}(n_1,n_2) = \frac{1}{F_{\alpha}(n_2,n_1)}$$

证

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

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

又因 $\dfrac{1}{F} \sim F(n_2,n_1) \Rightarrow \dfrac{1}{F_{1-\alpha}(n_1,n_2)} = F_{\alpha}(n_2,n_1)$

TIPS

例 统计量的分布(之二)

二、抽样分布定理

定理6.2.4

设 $X_1, X_2, \ldots, X_n$ 是正态总体 $X \sim N(\mu, \sigma^2)$ 的样本，$\overline{X}, S^2$ 分别是样本均值和样本方差，则

(1) $\overline{X}$ 与 S^2 相互独立;　　(2) $\dfrac{\overline{X}-\mu}{\sigma/\sqrt{n}} \sim N(0,1)$;

(3) $\dfrac{n-1}{\sigma^2}S^2 \sim \chi^2(n-1)$;　　(4) $\dfrac{\overline{X}-\mu}{S/\sqrt{n}} \sim t(n-1)$

证明(4)：由(2) $\Rightarrow U = \dfrac{\overline{X}-\mu}{\sigma/\sqrt{n}} \sim N(0,1)$

由(3) $\Rightarrow V = \dfrac{(n-1)S^2}{\sigma^2} \sim \chi^2(n-1)$

由(1)可知U和V是相互独立的,再由定理6.2.2可得

$$\frac{U}{\sqrt{V/(n-1)}} = \frac{\overline{X}-\mu}{\sigma/\sqrt{n}} \Bigg/ \sqrt{\frac{(n-1)S^2}{\sigma^2}\cdot\frac{1}{n-1}}$$

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

应用例子

定理6.2.5 设正态总体 X 与 Y 相互独立，

$X \sim N(\mu_1, \sigma_1^2)$, 样本为$X_1$， X_2， … X_{n_1},样本均值和样本方差为 $\bar{X}$, S_1^2;

$Y \sim N(\mu_2, \sigma_2^2)$ 样本为Y_1， Y_2， … Y_{n_2},样本均值和样本方差为$\bar{Y}$, S_2^2.

有

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

(2)当$\sigma_1^2 = \sigma_2^2$时，

$$T=\frac{(\overline{X}-\overline{Y})-(\mu_1-\mu_2)}{S_w\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}}\sim t(n_1+n_2-2)$$

其中，$S_w=\sqrt{\dfrac{(n_1-1)S_1^2+(n_2-1)S_2^2}{n_1+n_2-2}}$

分析 $\overline{X}-\overline{Y}$ **服从正态分布，$S_w{}^2$可化为χ^2分布，二者组合而成的统计量应服从 t 分布.**

证明： (2) 因 $\sigma_1^2=\sigma_2^2=\sigma^2$

故 $\overline{X}-\overline{Y} \sim N(\mu_1-\mu_2,(\frac{1}{n_1}+\frac{1}{n_2})\sigma^2)$

令 $$U=\frac{(\overline{X}-\overline{Y})-(\mu_1-\mu_2)}{\sigma\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}} \sim N(0,1)$$

由χ^2分布的可加性，

$$V=\frac{(n_1-1)S_1^2+(n_2-1)S_2^2}{\sigma^2} \sim \chi^2(n_1+n_2-2)$$

因 $\bar{X}, S_1^2$ $\bar{Y}, S_2^2$ 相互独立，故U与 V也相互独立，从而

$$T=\frac{U}{\sqrt{V/(n_1+n_2-2)}}$$

$$=\frac{(\bar{X}-\bar{Y})-(\mu_1-\mu_2)}{S_w\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}}\sim t(n_1+n_2-2)$$

总体、个体

简单随机样本

正态总体的
2个抽样定理

统计量

- 样本均值
- 样本方差
- 样本矩（样本相关系数）

统计量的分布

- χ^2分布
- t 分布
- F分布

分位数

结构定理

例6.2.1 设随机变量X服从正态分布$N(0,1)$，对给定的$\alpha(0<\alpha<1)$，数u_α满足

$$P\{X > u_\alpha\} = \alpha$$

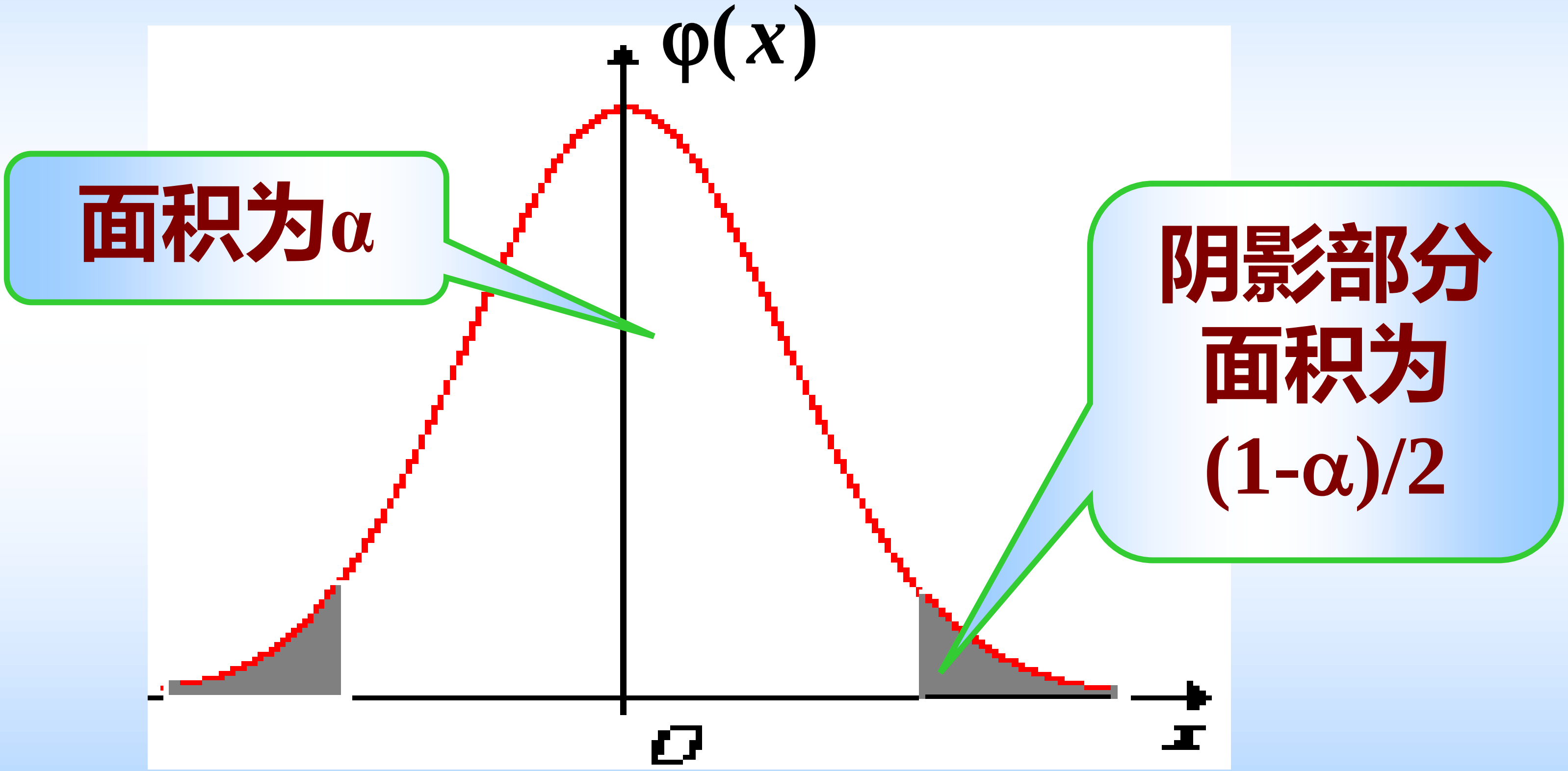
若$P\{|X| < x\} = \alpha$，则x等于 C

(A) $u_{\alpha/2}$; (B) $u_{1-\frac{\alpha}{2}}$; (C) $u_{\frac{1-\alpha}{2}}$; (D) $u_{1-\alpha}$.

解 $$P\{X > u_\alpha\} = \alpha$$

$\Rightarrow$ u_α是上侧α分位数.

若$P\{|X| < x\} = \alpha$, $\Rightarrow$ $x = u_{\frac{1-\alpha}{2}}$.



若$P\{|X|<x\}=\alpha$， $\Rightarrow$ $P\{|X|\geq x\}=1-\alpha$

$$\Rightarrow 1-\alpha = P\{X\geq x\}+P\{X\leq -x\}$$

因X的概率曲线关于y轴对称，故

$$P\{X\geq x\}=P\{X\leq x\}\Rightarrow P\{X\geq x\}=\frac{1-\alpha}{2}$$

$$\Rightarrow x=u_{\frac{1-\alpha}{2}}$$

例6.2.2 统计量的分布(之一)

设样本$X_1, X_2, \ldots, X_n$是来自正态总体$X \sim N(\mu,\sigma^2)$，容量为 n, 求下列统计量的概率分布:

1. $\overline{X}=\frac{1}{n}\sum_{i=1}^{n}X_i$,　　2. $Y=\frac{1}{\sigma^2}\sum_{i=1}^{n}(X_i-\mu)^2$

解 1. $X_i \sim N(\mu,\sigma^2) \Rightarrow \sum_{i=1}^{n}X_i \sim N(n\mu, n\sigma^2)$,

故 $$\frac{1}{n}\sum_{i=1}^{n}X_i \sim N(\mu,\frac{\sigma^2}{n}).$$

2. $$\frac{X_i-\mu}{\sigma}\sim N(0,1)\Rightarrow\sum_{i=1}^{n}\left(\frac{X_i-\mu}{\sigma}\right)^2\sim\chi^2(n)$$

故 $$Y=\frac{1}{\sigma^2}\sum_{i=1}^{n}(X_i-\mu)^2\sim\chi^2(n).$$

#

例6.2.3 查表计算概率

1. $X \sim N(0,1)$, $P\{-1.58 \le X \le 1.96\} = ?$

2. $\chi^2 \sim \chi^2(15)$, $\mathrm{P}\{6.262 \le \chi^2 \le 24.996\} = ?$

1. $\because X \sim N(0,1)$,

$\therefore P\{-1.58 \le X \le 1.96\} = P\{X \le 1.96\} - P\{X \le -1.58\}$

$= P\{X \le 1.96\} - \left[1 - P\{X \le 1.58\}\right]$

$= \Phi(1.96) - [1 - \Phi(1.58)] \quad = 0.975 - (1 - 0.943)$

$= \quad 0.918$

$$\chi^2_{0.975}(15) = 6.262$$
$$\chi^2_{0.05}(15) = 24.996$$

$$2. \because \chi^2 \sim \chi^2(15)$$

$$\therefore P\{6.262 \le \chi^2 \le 24.996\}$$

$$= P\{\chi^2 \ge 6.262\} - P\{\chi^2 \ge 24.996\}$$

$$= 0.975 - 0.05 \quad = 0.925$$

注意 应注意分布表的定义与查法!

#

例6.2.4 统计量的分布(之二)

设 $X_1, X_2, \ldots, X_{n+m}$ 是来自正态总体 $X \sim N(0, \sigma^2)$ 的样本，求下列统计量的概率分布

$$1.\ Y = \frac{1}{\sigma^2}\sum_{i=1}^{n+m} X_i^2, \quad 2.\ Z = \sqrt{\frac{m}{n}}\sum_{i=1}^{n} X_i \Big/ \sqrt{\sum_{i=n+1}^{n+m} X_i^2}, \quad 3.\ \frac{1}{Z^2}.$$

解

1. $\dfrac{X_i}{\sigma} \sim N(0,1)$, 且所有 $\dfrac{X_i}{\sigma}$ 相互独立

$$(i = 1,2,\ldots,n+m)$$

故 $Y=\frac{1}{\sigma^2}\sum_{i=1}^{n+m}X_i^2=\sum_{i=1}^{n+m}\left(\frac{X_i}{\sigma}\right)^2\sim\chi^2(n+m)$

2.因 $\sum_{i=1}^{n}X_i\sim N(0,n\sigma^2)\Rightarrow U=\sum_{i=1}^{n}X_i\Big/\sqrt{n\sigma^2}\sim N(0,1)$

同时 $V=\sum_{i=n+1}^{n+m}\frac{X_i^2}{\sigma^2}\sim\chi^2(m)$， U与V相互独立

由 t 分布结构定理知

$$Z=\sqrt{\frac{m}{n}}\sum_{i=1}^{n}X_i\Bigg/\sqrt{\sum_{i=n+1}^{n+m}X_i^2}=\frac{U}{\sqrt{V/m}}\sim t(m)$$

3. 因$Z \sim t(m)$, 根据 t 分布结构定理，有

$$Z = \frac{U}{\sqrt{V/m}},$$

$V \sim \chi^2(m)$, U与V相互独立

而且 $U \sim N(0,1) \Rightarrow U^2 \sim \chi^2(1)$,

故 $$\frac{1}{Z^2} = \frac{V/m}{U^2/1} \sim F(m,1)$$

#

例6.2.5 设$X_1, X_2, \ldots\ldots, X_9$是来自正态总体$X$的随机样本，令

$$Y_1 = \frac{1}{6}(X_1 + \cdots + X_6), \quad Y_2 = \frac{1}{3}(X_7 + X_8 + X_9),$$

$$S^2 = \frac{1}{2}\sum_{i=7}^{9}(X_i - Y_2)^2, \quad Z = \frac{\sqrt{2}(Y_1 - Y_2)}{S},$$

证明统计量Z 服从自由度为2 的t 分布.

证 记 $D(X)=\sigma^2$(未知)，有

$$E(Y_1) = E(Y_2), \quad D(Y_1) = \frac{\sigma^2}{6}, D(Y_2) = \frac{\sigma^2}{3},$$

由于Y_1, Y_2相互独立，故

$$E(Y_1 - Y_2) = 0, \quad D(Y_1 - Y_2) = \frac{\sigma^2}{6} + \frac{\sigma^2}{3} = \frac{\sigma^2}{2},$$

从而

$$U = \frac{Y_1 - Y_2}{\sigma/\sqrt{2}} \sim N(0,1)$$

由抽样分布定理知

$$\chi^2 = \frac{2S^2}{\sigma^2} \sim \chi^2(2),$$

因Y_1和Y_2相互独立，Y_1和S^2 相互独立，而且Y_2和S^2 相互独立，故Y_1 - Y_2和S^2 相互独立.

根据t分布结构定理知

$$Z=\frac{U}{\sqrt{\chi^2/2}}=\frac{\sqrt{2}(Y_1-Y_2)}{S}$$

服从自由度为2的t分布.

#

证明 $\chi_{\alpha}^{2}(n) \approx n + u_{\alpha}\sqrt{2n}$

证 因 $\chi^{2} = \sum_{i=1}^{n} X_{i}^{2}$

且 X_1，X_2，…，X_n相互独立，$X_i \sim N(0,1)$，

$\Rightarrow$ $X_1^2, X_2^2, \cdots, X_n^2$ 也相互独立同分布，

根据独立同分布中心极限定理有

$$\chi^{2} \sim N(n, 2n)$$

近似成立，故

$$1-\alpha = P\{\chi^2 < \chi^2_\alpha(n)\} = P\{\frac{\chi^2 - n}{\sqrt{2n}} < \frac{\chi^2_\alpha(n) - n}{\sqrt{2n}}\}$$

$$\approx \Phi(\frac{\chi^2_\alpha(n) - n}{\sqrt{2n}})$$

$$\Rightarrow \quad \frac{\chi^2_\alpha(n) - n}{\sqrt{2n}} \approx u_\alpha \Rightarrow \chi^2_\alpha(n) \approx n + u_\alpha\sqrt{2n}$$

#