第一章 抽样和抽样分布

§2 常用抽样分布

理学院

- 本次课教学目的：
  - 掌握三大抽样分布的构造性定义并熟悉一些重要结论
- 重点难点：
  - 三大抽样分布的构造及其抽样分布
  - 抽样分布和中心极限定理的理解
  - 查表求上α分位点的熟练掌握

§2　常用抽样分布

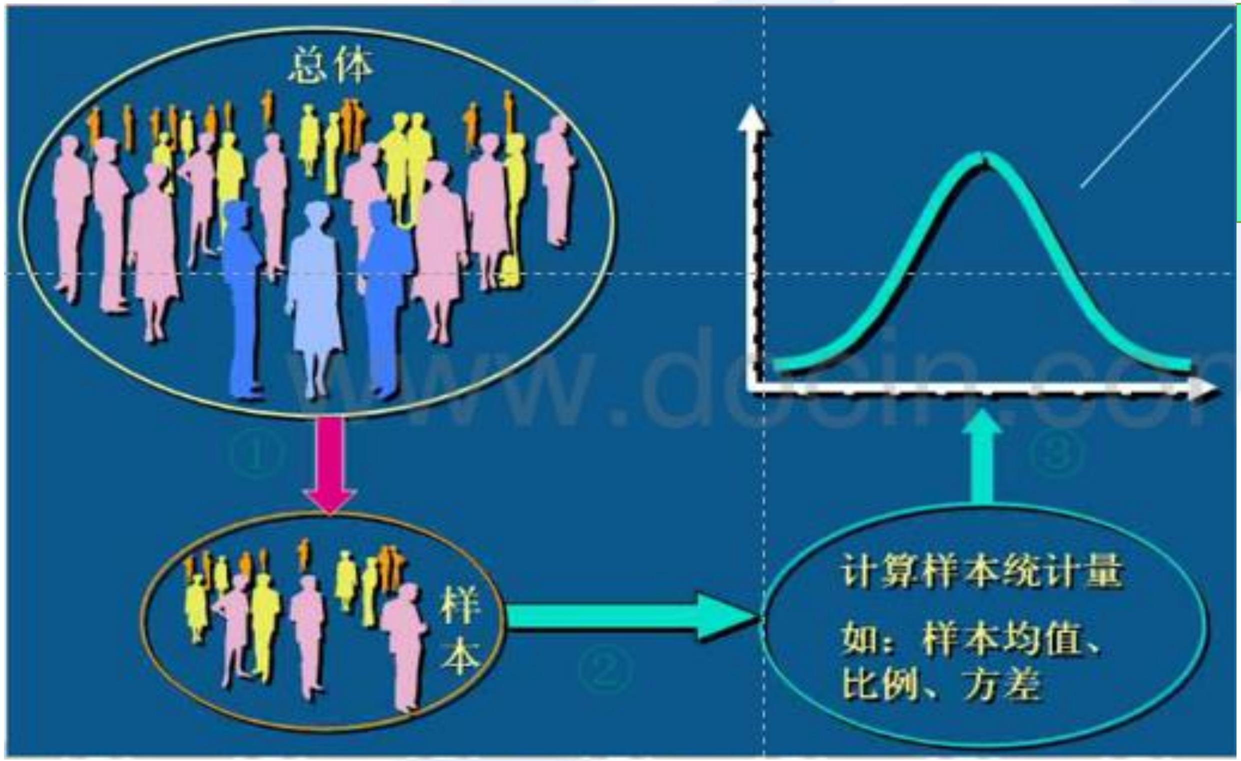
统计量既然是依赖于样本的，而后者又是随机变量，故统计量也是随机变量，因而就有一定的分布，这个分布叫做统计量的“抽样分布”　.

抽样分布
- 精确抽样分布　(小样本问题中使用)
- 渐近分布　(大样本问题中使用)

在小样本问题中，样本容量固定,如果能得到有关统计量或样本函数的精确分布,就能较精确和较满意地讨论和分析各种统计问题。在大样本问题中，用统计量的极限分布作为近似分布。

抽样分布是由一个统计量(随机变量函数)确定的分布.研究统计量的性质和评价一个统计推断的优良性，完全取决于其抽样分布的性质.

构造抽样分布的过程



样本统计量全部可能的数值对应的频数分布，即抽样分布

2.1 正态总体的样本均值与样本方差的分布

由实际问题与中心极限定理可知，讨论正态总体的样本统计量的分布非常必要。

定理2: 设 $X_1, X_2, \cdots X_n$ 是来自正态总体 $N(\mu,\sigma^2)$ 的样本，且是独立同分布的随机变量，$\overline{X}$ 为其样本均值，则总有

$$\boxed{E(\overline{X}) = \mu, \quad D(\overline{X}) = \frac{\sigma^2}{n}, \quad E(S^2) = \sigma^2.}$$

推导:

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

$$D(\overline{X}) = D(\frac{1}{n}\sum_{i=1}^{n} X_i) = \frac{1}{n^2}\sum_{i=1}^{n} D(X_i) = \frac{D(X)}{n} = \frac{\sigma^2}{n}$$

$$E(S^2) = E\left[\frac{1}{n-1}\left(\sum_{i=1}^{n} X_i^2 - n\overline{X}^2\right)\right] = \frac{1}{n-1}\left[\sum_{i=1}^{n} E(X_i^2) - nE(\overline{X}^2)\right]$$

$$= \frac{1}{n-1}\left[\sum_{i=1}^{n}(\sigma^2+\mu^2) - n(\frac{\sigma^2}{n}+\mu^2)\right] = \sigma^2$$

2.1 正态总体的样本均值与样本方差的分布

由实际问题与中心极限定理可知，讨论正态总体的样本统计量的分布非常必要。

定义1.2.1 设X是随机变量，则称满足

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

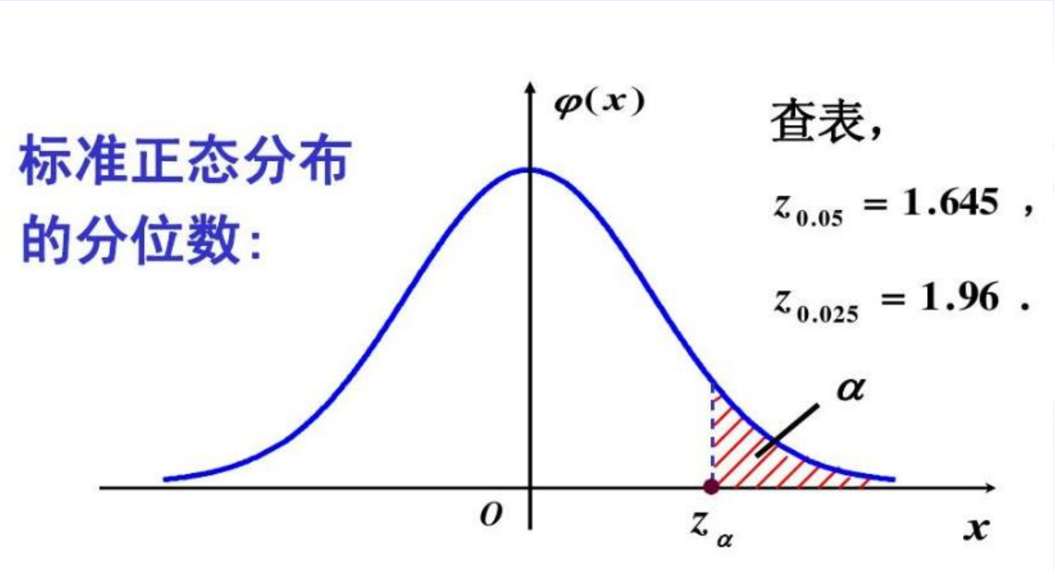
的u_α为X的上侧α分位数，简称为(上侧)分位数。

标准正态分布：

随机变量$x \sim \boldsymbol{N}(\mathbf{0}, \mathbf{1})$,其密度函数为

$$\varphi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \quad (-\infty < x < +\infty)$$

它的图形关于y轴对称.



对于数$\alpha(0<\alpha<1)$, 通过查表，可求出满足等式

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

的上侧分位数u_α（显然$u_{1-\alpha} = 1 - u_\alpha$）。当$\alpha = 0.025$时，可查得$u_{0.025} = 1.96$.

定理2: 设 $X_1, X_2, \cdots X_n$ 是来自正态总体 $N(\mu,\sigma^2)$ 的样本，且是独立同分布的随机变量，$\overline{X}$ 为其样本均值，则总有

$$\boxed{E(\overline{X})=\mu,\quad D(\overline{X})=\frac{\sigma^2}{n},\quad E(S^2)=\sigma^2.}$$

推导:

$$E(\overline{X})=E\left(\frac{1}{n}\sum_{i=1}^{n}X_i\right)=\frac{1}{n}\sum_{i=1}^{n}E(X_i)=\frac{1}{n}\sum_{i=1}^{n}E(X)=\mu$$

$$D(\overline{X})=D\left(\frac{1}{n}\sum_{i=1}^{n}X_i\right)=\frac{1}{n^2}\sum_{i=1}^{n}D(X_i)=\frac{D(X)}{n}=\frac{\sigma^2}{n}$$

$$E(S^2)=E\left[\frac{1}{n-1}\left(\sum_{i=1}^{n}X_i^2-n\overline{X}^2\right)\right]=\frac{1}{n-1}\left[\sum_{i=1}^{n}E(X_i^2)-nE(\overline{X}^2)\right]$$

$$=\frac{1}{n-1}\left[\sum_{i=1}^{n}(\sigma^2+\mu^2)-n\left(\frac{\sigma^2}{n}+\mu^2\right)\right]=\sigma^2$$

- 有许多**统计推断**是基于正态分布的假设的，以**标准正态变量为基石而构造的三个著名统计量**在实际中有广泛的应用，这是因为这三个统计量不仅有明确背景，而且其抽样分布的密度函数有明显表达式，他们被称为统计中的“**三大抽样分布**”.
- 若设 $x_1, x_2, \cdots x_n, y_1, y_2, \cdots, y_m$ 是来自标准正态分布的两个相互独立的样本，则此三个统计量的构造及其抽样分布如下表所示.

三个著名统计量的构造及其抽样分布

统计量的构造	抽样分布的密度函数	期望	方差
$\chi^2 = x_1^2 + x_2^2 + \ldots + x_n^2$	$p(y) = \dfrac{1}{\Gamma(\frac{n}{2})2^{n/2}} y^{\frac{n}{2}-1} e^{-\frac{y}{2}} (y>0)$	n	2n
$F = \dfrac{(y_1^2+\ldots+y_m^2)/m}{(x_1^2+\ldots+x_n^2)/n}$	$p(y) = \dfrac{\Gamma\left(\frac{m+n}{2}\right)\left(\frac{m}{n}\right)^{m/2}}{\Gamma\left(\frac{m}{2}\right)\Gamma\left(\frac{n}{2}\right)} y^{\frac{m}{2}-1} \cdot \left(1+\frac{m}{n}y\right)^{-\frac{m+n}{2}}$	$\dfrac{2}{n-2}$ $(n>2)$	$\dfrac{2n^2(m+n-2)}{m(n-2)^2(n-4)}$ $(n>4)$
$t = \dfrac{y_1}{\sqrt{(x_1^2+\ldots x_n^2)/n}}$	$p(y) = \dfrac{\Gamma\left(\frac{n+1}{2}\right)}{\sqrt{n\pi}\,\Gamma\left(\frac{n}{2}\right)}\left(1+\frac{y^2}{n}\right)^{-\frac{n+1}{2}}$	0 $(n>1)$	$\dfrac{n}{n-2}$ $(n>2)$

$$\Gamma(s) = \int_0^{+\infty} x^{s-1} e^{-x} dx \qquad (s>0)$$

2.2 χ^2分布

定理3：设 $X_1, X_2, \ldots, X_n \sim N(0,1)$且相互独立，则称随机变量

$$\chi^2 = X_1^2 + X_2^2 + \cdots + X_n^2$$

由正态分布派生出来的一种分布

服从自由度为n的χ^2分布，记为 $\chi^2 \sim \chi^2(n)$.

这里自由度n表示相互独立的随机变量的个数，也称为χ^2变量.

➢ 问题：如何确定χ^2的分布？

→ 思路：通过推导分布密度函数$f(y)$来求解！

注：1. $X_1, X_2, \ldots, X_n$独立同分布且 $X_i \sim N(0,1)$，则

$$X_1^2 + X_2^2 + \cdots + X_n^2 \sim \chi^2(n)$$

2. $X \sim N(0,1)$，则 $X^2 \sim \chi^2(1)$.

☞$f(y)$的推导：由前例知， $\chi^2(1)=\Gamma(\frac{1}{2},\frac{1}{2})$

而$X_i \sim N(0,1)$，由定义$X_i^2 \sim \chi^2(1)$，即

$$X_i^2 \sim \Gamma\left(\frac{1}{2},\frac{1}{2}\right)$$

再由$X_1, X_2, \ldots, X_n$的独立性及Γ分布的可加性

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

(1) $X \sim N(0,1)$, $X^2 \sim \Gamma(1/2, 1/2)$;

(2) $X_1 \sim \Gamma(\alpha_1,\beta), X_2 \sim \Gamma(\alpha_2,\beta)$

$\xrightarrow{\text{独立}} X_1+X_2 \sim \Gamma(\alpha_1+\alpha_2,\beta)$

$\chi^2(n)$分布的概率密度为

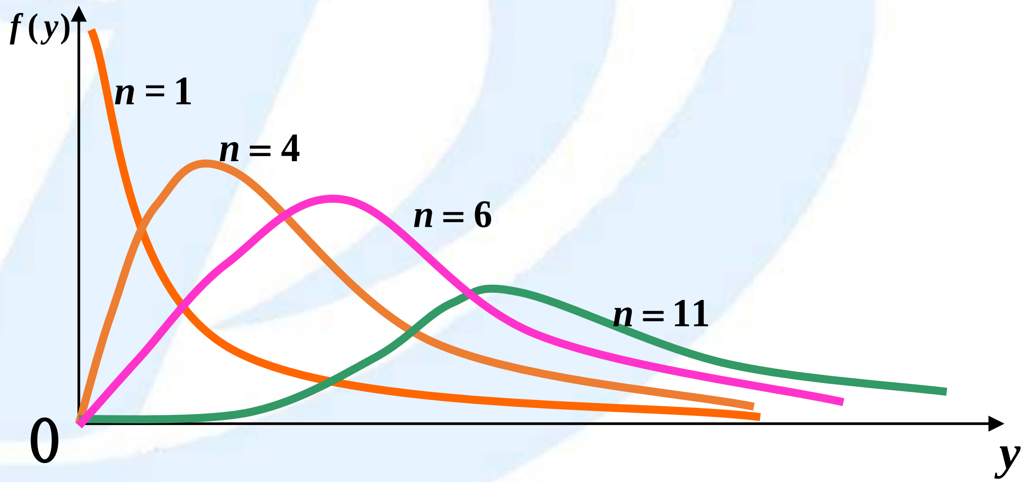
$$f(y)=\begin{cases}\dfrac{1}{2^{n/2}\Gamma(n/2)}y^{\frac{n}{2}-1}e^{-\frac{y}{2}}, & y>0,\\ 0, & \text{其他}.\end{cases}$$

其中

$$\Gamma\left(\frac{n}{2}\right)=\int_0^{+\infty}e^{-x}x^{\frac{n}{2}-1}dx.$$

☞ $f(y)$的图形

（与n有关）：



☞ $f(y)$的图形特点：

密度函数的图像是一个只取非负值的偏态分布。

χ^2变量的性质

- 性质1：设$X_1,\ldots,X_n\sim N(\mu,\sigma^2)$且相互独立

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

事实上 $$Y_i=\frac{1}{\sigma}(X_i-\mu)\sim N(0,1)$$

- 性质2：设$X\sim\chi^2(n)$，则 $E(X)=n, D(X)=2n.$

$$X^2=\sum_{I=1}^{N}X_i^2$$

由定义知 $E(X_i)=0,\ D(X_i)=1$

$$EX=E\left(\sum_{i=1}^{n}X_i^2\right)=\sum_{i=1}^{n}EX_i^2=\sum_{i=1}^{n}\left(DX_i+(EX_i)^2\right)=n$$

$$DX=D\left(\sum_{i=1}^{n}X_i^2\right)=\sum_{i=1}^{n}DX_i^2=\sum_{i=1}^{n}\left(EX_i^4-\left(EX_i^2\right)^2\right)$$

$$=\sum_{i=1}^{n}\left(\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty}xe^{-\frac{x^2}{2}}dx-\left(DX_i+(EX_i)\right)^2\right)$$

$$=\sum_{i=1}^{n}(3-1)=2n$$

➢ 性质3：χ^2 变量的可加性

若$X_1 \sim \chi^2(n_1), X_2 \sim \chi^2(n_2)$且相互独立，则

$$X_1 + X_2 \sim \chi^2(n_1 + n_2)$$

要用到独立随机变量和的卷积公式和$\Gamma(x)$ 的性质。

➢ 性质4：极限分布 ⟶ 应用Lindeberg中心极限定理

若$\chi_n^2 \sim \chi_{(n)}^2$，则$\forall x \in (-\infty, +\infty)$有

$$\lim_{n\to+\infty} P\left(\frac{\chi_n^2 - n}{\sqrt{2n}} \le x\right) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{x^2}{2}} dx$$

$$\frac{\chi_n^2 - n}{\sqrt{2n}} \xrightarrow{\text{极限分布}} N(0,1) \qquad \chi_n^2 \xrightarrow{\text{极限分布}} N(n, 2n)$$

$$\lim_{n\to+\infty} P\left(\frac{\chi_n^2 - n}{\sqrt{2n}} \le x\right) = \lim_{n\to+\infty} P\left(\frac{\sum_{i=1}^{n} X_i^2 - n\cdot\mu}{\sqrt{n}\cdot\sigma} \le x\right)$$

其中 $X_i^2 \sim \chi_{(1)}^2, EX_i^2 = 1, DX_i^2 = 2$

例6 设$X_1, X_2, \ldots, X_6$是来自总体$X\sim N(0,1)$, 又设

$$Y=(X_1+X_2+X_3)^2+(X_4+X_5+X_6)^2$$

试求常数C, 使CY服从χ^2分布.

解: 因为$X_1+X_2+X_3\sim N(0,3)$, $X_4+X_5+X_6\sim N(0,3)$,

所以 $\dfrac{X_1+X_2+X_3}{\sqrt{3}}\sim N(0,1)$ $\dfrac{X_4+X_5+X_6}{\sqrt{3}}\sim N(0,1)$

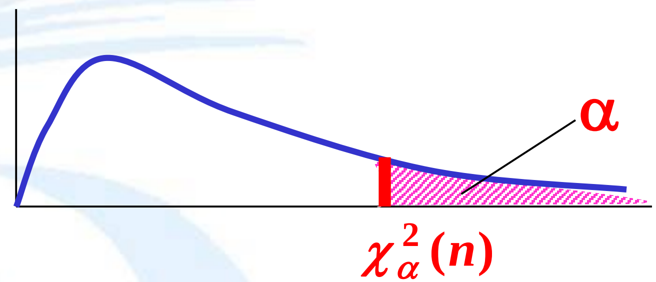
且它们相互独立.于是由χ^2分布的可加性得:

$$\left(\frac{X_1+X_2+X_3}{\sqrt{3}}\right)^2+\left(\frac{X_4+X_5+X_6}{\sqrt{3}}\right)^2\sim\chi^2(2)$$

故应取常数 $C=\dfrac{1}{3}$, 于是$\dfrac{1}{3}Y\sim\chi^2(2)$.

☞ χ^2分布的分位点：当n充分大(>45)时，对给定α (0<α<1)，有

$$P\{X \ge \chi^2_\alpha(n)\} = \alpha$$



因而

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

令$U = \dfrac{X-n}{\sqrt{2n}}$，近似服从标准正态分布 $N(0,1)$，则有

$$P\{U \ge u_\alpha\} \approx \alpha$$

标准正态分布的上α分位点

从而有

$$\boxed{\chi^2_\alpha(n) \approx n + \sqrt{2n}u_\alpha}$$

例5： 试求$\chi^2_{0.05}(120)$的数值.

解： 因为$\alpha = 0.05$，查表得u_α=1.645，根据分位点定义

$$\chi^2_{0.05}(120) \approx 120 + \sqrt{2 \times 120} \times 1.645 = 145.5$$

2.3 t 分布

定理4: 设$X\sim N(0,1)$, $Y\sim \chi^2(n)$,且X与Y相互独立,则称随机变量

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

服从自由度是n的t 分布(Student分布), 也称为t变量,记作$t\sim t(n)$.

➢ 问题—— 如何确定t 的分布?

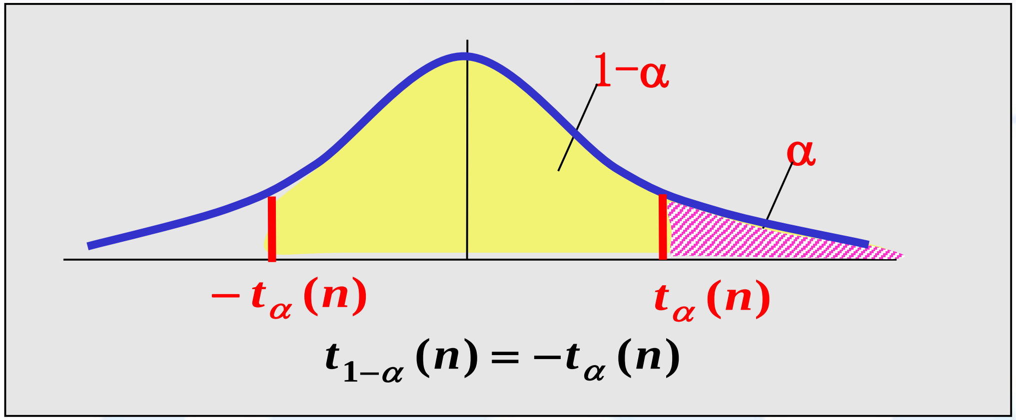
➡ 思路: 通过推导分布密度函数 $f(y)$ 来求解!

$t(n)$分布的概率密度函数为:

$$f(t)=\frac{\Gamma[(n+1)/2]}{\sqrt{\pi n}\,\Gamma(n/2)}\left(1+\frac{t^2}{n}\right)^{-(n+1)/2}, -\infty<t<\infty$$

$$\Gamma(s)=\int_0^{+\infty}x^{s-1}e^{-x}dx \qquad (s>0)$$

☞$f(t)$图形：(关于$t=0$对称，其形状与n有关)



$$\lim_{n\to\infty} f(t) = \frac{1}{\sqrt{2\pi}} e^{-t^2/2}$$

☞$f(y)$的图形特点：

密度函数的图像是是一个关于y轴对称的分布.

☞t分布的分位点：对给定α $(0<\alpha<1)$，称满足

$$P\{t > t_\alpha(n)\} = \int_{t_\alpha(n)}^{\infty} f(t)dt = \alpha$$

的点$t_\alpha(n)$为$t(n)$分布的上侧分位点.

t分布的性质：

1^0. $t \sim t(n)$为具有自由度为n的t分布的随机变量，则 t的数字特征具有如下性质：

当 $n=1$时， $t \sim t(n)$实际上是柯西分布，任何阶矩均不存在；

当$n>2$时， $E(t)=0$; $D(t)=n/(n-2)$.

2^0. $X \sim N(\mu,\sigma^2), Y/\sigma^2 \sim \chi^2_{(n)}$且相互独立，则

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

事实上 $\frac{X-\mu}{\sigma} \sim N(0,1) \longrightarrow \frac{X-\mu}{\sqrt{Y/n}}=\frac{(X-\mu)/\sigma}{\sqrt{(Y/\sigma^2)/n}}$

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

3^0. 极限分布

t分布的密度函数关于$x=0$对称，是偶函数，且

$$\lim_{|x|\to\infty} f(x;n)=0$$

当n充分大时，其图形类似于标准正态分布密度函数的图形.

当n充分大时，t 分布近似$N(0,1)$分布．但对于较小的n，t分布与$N(0,1)$分布相差很大.

$$f(t)=\frac{\Gamma[(n+1)/2]}{\sqrt{\pi n}\Gamma(n/2)}\left(1+\frac{t^2}{n}\right)^{-(n+1)/2}, -\infty<t<\infty$$

$$\Gamma(s)=\int_0^{+\infty} x^{s-1}e^{-x}dx \qquad (s>0)$$

例7 设$X\sim N(2,1)$, $Y_1, Y_2,\ldots, Y_4$ 均服从$N(0,4)$，且都相互独立，令

$$T=\frac{4(X-2)}{\sqrt{\sum_{i=1}^{4}Y_i^2}}$$

试求T的分布，并确定 t_0 的值，使 $P\{|T|>t_0\}=0.01.$

解： 由t分布的定义知，$X-2\sim N(0,1)$，$Y_i/2\sim N(0,1)$， $i=1,2,3,4.$

故

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

由 $P\{|T|>t_0\}=\alpha=0.01.$ 查表得：

$$t_0=t_{\alpha/2}(4)=t_{0.005}(4)=4.6041$$

2.4 F分布

定理5：设$X\sim\chi^2(n_1)$, $Y\sim\chi^2(n_2)$,且U与V相互独立， 则称

$$F=\frac{X_1/n_1}{X_2/n_2}$$

服从自由度为(n_1, n_2)的F分布, 也称为F变量，记为$F\sim F(n_1,n_2)$.

- 问题：如何确定F的分布？

$\longrightarrow$ 思路：通过两步推导分布密度函数来求解！

- 首先，我们导出 $Z=\dfrac{X_1}{X_2}$的密度函数
- 第二步，我们导出$F=\dfrac{n_2}{n_1}Z$ 的密度函数

首先，我们导出 $Z=\frac{X_1}{X_2}$ 的密度函数:

$p_1(x)$ 和 $p_2(x)$ 分别记为 $\chi^2(n_1)$ 和 $\chi^2(n_2)$ 的密度函数

$$X_1 \sim \chi^2(n_1) \Rightarrow p_1(x_1)=\frac{\left(\frac{1}{2}\right)^{\frac{n_1}{2}}}{\Gamma\left(\frac{n_1}{2}\right)} x_1^{\frac{n_1}{2}-1} e^{-\frac{x_1}{2}}, \quad x_1>0$$

$$X_2 \sim \chi^2(n_2) \Rightarrow p_2(x_2)=\frac{\left(\frac{1}{2}\right)^{\frac{n_2}{2}}}{\Gamma\left(\frac{n_2}{2}\right)} x_2^{\frac{n_2}{2}-1} e^{-\frac{x_2}{2}}, \quad x_2>0$$

z的密度函数为

$$p_Z(z)=\int_0^{+\infty} x_2 p_1(zx_2)p_2(x_2)dx_2=\frac{z^{\frac{n_1}{2}-1}}{\Gamma\left(\frac{n_1}{2}\right)\Gamma\left(\frac{n_2}{2}\right)2^{\frac{n_1+n_2}{2}}}\int_0^{+\infty} x_2^{\frac{n_1+n_2}{2}-1} e^{-\frac{x_2}{2}(1+z)}dx_2$$

（做变换令$u=\dfrac{x_2}{2}(1+z)$）

于是 $$p_z(z)=\frac{z^{\frac{n_1}{2}-1}(1+z)^{-\frac{n_1+n_2}{2}}}{\Gamma\left(\frac{n_1}{2}\right)\Gamma\left(\frac{n_2}{2}\right)}\int_0^{\infty} u^{\frac{n_1+n_2}{2}-1}e^{-u}du$$

$$\Gamma\left(\frac{n_1+n_2}{2}\right)$$

$$=\frac{\Gamma\left(\frac{n_1+n_2}{2}\right)}{\Gamma\left(\frac{n_1}{2}\right)\Gamma\left(\frac{n_2}{2}\right)}z^{\frac{n_1}{2}-1}(1+z)^{-\frac{n_1+n_2}{2}}\qquad z>0$$

$$\Gamma(s)=\int_0^{+\infty} x^{s-1}e^{-x}dx\qquad (s>0)$$

第二步，我们导出 $F=\frac{n_2}{n_1}Z$ 的密度函数

$$p_F(y)=p_Z\left(\frac{n_1}{n_2}y\right)\bullet\frac{n_1}{n_2}$$

$$=\frac{\Gamma\left(\frac{n_1+n_2}{2}\right)}{\Gamma\left(\frac{n_1}{2}\right)\Gamma\left(\frac{n_2}{2}\right)}\left(\frac{n_1}{n_2}y\right)^{\frac{n_1}{2}-1}\left(1+\frac{n_1}{n_2}y\right)^{-\frac{n_1+n_2}{2}}\bullet\frac{n_1}{n_2}$$

$$=\frac{\Gamma\left(\frac{n_1+n_2}{2}\right)\left(\frac{n_1}{n_2}\right)^{\frac{n_1}{2}}}{\Gamma\left(\frac{n_1}{2}\right)\Gamma\left(\frac{n_2}{2}\right)}y^{\frac{n_1}{2}-1}\left(1+\frac{n_1}{n_2}y\right)^{-\frac{n_1+n_2}{2}}$$

➢从而，$F\sim F(n_1,n_2)$ 分布的概率密度函数证毕！

F变量的性质

1^0. 若$X\sim F(n_1,n_2)$，X的数学特征:

EX不依赖于第一自由度n_1.

若$n_2>2$ $$E(X)=\frac{n_2}{n_2-2}$$

若$n_2>4$ $$D(X)=\frac{2n_2^2(n_1+n_2-2)}{n_1(n_2-2)^2(n_2-4)},$$

2^0. 若$n_1=1$时，$F\sim F(1,n_2)=t^2(n_2)$.

$$F=\frac{X/n_1}{Y/n_2}=\frac{X/1}{Y/n_2}=\left(\sqrt{\frac{X/1}{Y/n_2}}\right)^2$$

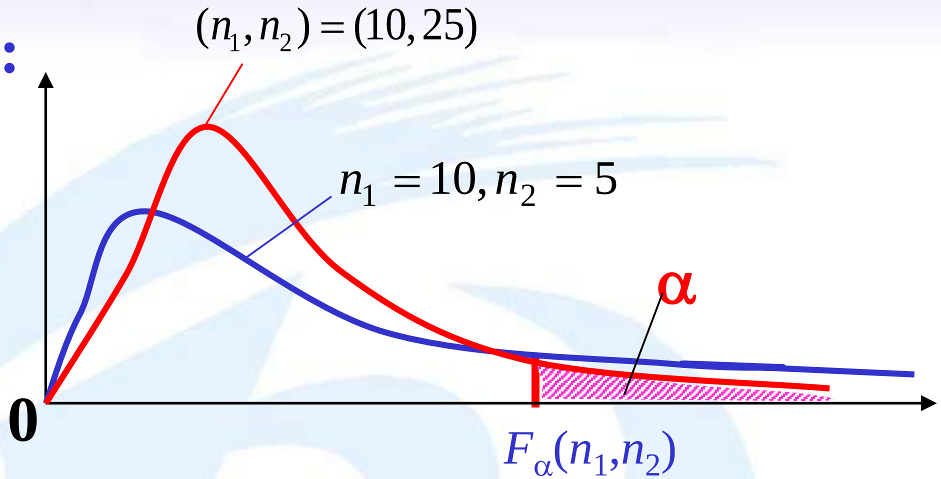
3^0. 极限分布

若$X\sim F(n_1,n_2)$， $n_2>4$，则 $\forall x\in(-\infty,+\infty)$ 有

$$\lim_{n\to+\infty}P\left(\frac{X-E(X)}{\sqrt{D(X)}}\le x\right)=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x}e^{-\frac{x^2}{2}}dx$$

➢ $F(n_1,n_2)$的图形特点:

非对称分布



➢ F分布的分位点:

对给定α (0<α<1),称满足

$$P\{F>F_\alpha(n_1,n_2)\}=\int_{F_\alpha(n_1,n_2)}^{\infty} p_F(y)dy=\alpha$$

$$P\{F\geq F_{1-\alpha}(n_1,n_1)\}=1-\alpha$$

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

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

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

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

$$=1-P\left\{\frac{1}{F}\geqslant F_\alpha(n_2,n_1)\right\}$$

$$=1-\alpha$$

例8：若取n_1=8,n_2=15，α=0.05,求其F分布.

提示：本题主要考查对F分布的基本性质的掌握.

解：由F分布的定义可知，从附表上可查得

$$F_{0.05}(8,15)=2.64$$

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

$$F_{0.95}(15,8)=\frac{1}{F_{0.05}(8,15)}=\frac{1}{2.64}=0.379$$

抽样分布定理

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

- (1) $\overline{X} \sim N(\mu, \frac{\sigma^2}{n})$ 定理1.3.1(P15) $\Rightarrow$ $\frac{\overline{X}-\mu}{\sigma/\sqrt{n}} \sim N(0,1)$
- (2) $\overline{X}$与S^2独立. 定理1.3.3(P15)
- (3) $\sum_{i=1}^{n} \frac{(X_i-\overline{X})^2}{\sigma^2} = \frac{(n-1)S^2}{\sigma^2} \sim \chi^2(n-1)$; 定理1.3.3(P15)
- (4) $\frac{\overline{X}-\mu}{S/\sqrt{n}} \sim t(n-1)$. 定理1.3.4(P16)

定理6: 设$X_1, X_2, \ldots, X_{n_1}$与$Y_1, Y_2, \ldots, Y_{n_2}$分别是来自正态总体$N(\mu_1, \sigma_1^2)$和$N(\mu_2, \sigma_2^2)$的样本,且这两个样本相互独立. 两个样本的均值和方差分别为$\overline{X}, \overline{Y}, S_1^2, S_2^2$，则有

1° $$\boxed{\frac{S_1^2 / S_2^2}{\sigma_1^2 / \sigma_2^2} \sim F(n_1-1, n_2-1);}$$ 定理1.3.7(P18)

2° 当$\sigma_1^2 = \sigma_2^2 = \sigma^2$时，

$$\boxed{\frac{(\overline{X}-\overline{Y})-(\mu_1-\mu_2)}{S_\varpi \sqrt{\dfrac{1}{n_1}+\dfrac{1}{n_2}}} \sim t(n_1+n_2-2),}$$ 定理1.3.6(P17)

其中 $$S_\varpi^2 = \frac{(n_1-1)S_1^2+(n_2-1)S_2^2}{n_1+n_2-2}, \quad S_\varpi = \sqrt{S_\varpi^2}$$

证明：1° 由定理1.3.3 知

$$\frac{(n_1-1)S_1^{\,2}}{\sigma_1^{\,2}}\sim\chi^2(n_1-1),\qquad \frac{(n_2-1)S_2^{\,2}}{\sigma_2^{\,2}}\sim\chi^2(n_2-1);$$

两者相互独立，由*F*分布定义可知

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

化简后即得1°．

2° 由$\chi^2(n)$的可加性：$\dfrac{(n_1-1)S_1^{\,2}}{\sigma_1^{\,2}}+\dfrac{(n_2-1)S_2^{\,2}}{\sigma_2^{\,2}}\sim\chi^2(n_1+n_2-2),$

$$\bar{X}-\bar{Y}\sim N\left(\mu_1-\mu_2,\frac{\sigma^2}{n_1}+\frac{\sigma^2}{n_2}\right),\quad\Longrightarrow\quad U=\frac{(\bar{X}-\bar{Y})-(\mu_1-\mu_2)}{\sigma\sqrt{\dfrac{1}{n_1}+\dfrac{1}{n_2}}}\sim N(0,1),$$

t 分布定义 $\Longrightarrow$

$$\frac{(\bar{X}-\bar{Y})-(\mu_1-\mu_2)}{\sigma\sqrt{\dfrac{1}{n_1}+\dfrac{1}{n_2}}}\Bigg/\sqrt{\frac{(n_1-1)S_1^2+(n_2-1)S_2^2}{\sigma^2(n_1+n_2-2)}}\sim t(n_1+n_2-2)$$

化简后即得2°．

Thank you!