概率论与数理统计

第二节　常用统计分布

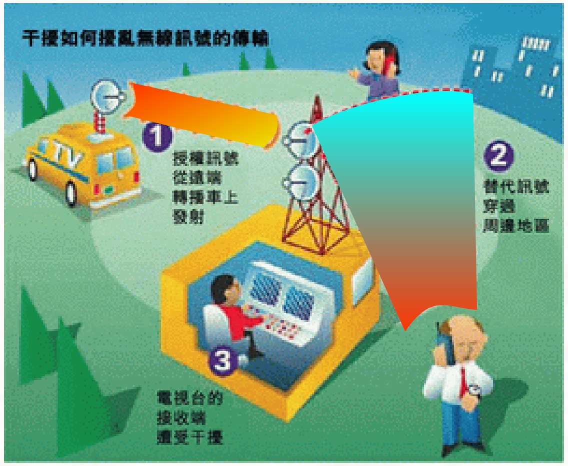
- 一、常见分布
- 二、概率分布的分位数

一、常见分布

在实际中我们往往会遇到这样的问题,要求有关随机变量的函数的概率分布.

例如在无线电接收中，某时刻接收到的信号是一个随机变量X，若我们把这个信号通过平方示波器，则输出的信号为

$$Y = X^2$$



通常需要求出Y的概率分布.

本节介绍一些最常见的统计分布.

1. χ^2 分布

正态分布是自然界中最常见的一类概率分布，例如测量的误差；人的生理尺寸：身高，体重等都近似服从正态分布.常见的问题是关于这些正态随机变量的平方以及平方和的概率分布问题.

例如在统计物理中，若气体分子速度是随机向量 $V=(X,Y,Z)$ 各分量相互独立，且均服从 $N(0,1.5)$，要求该分子运动动能

$$S=\frac{1}{2}m(X^2+Y^2+Z^2)$$

的分布规律.

要求S的分布,自然首先就要知道S中的随机变量

$$X^2+Y^2+Z^2$$

的概率分布.

对于这种在实际中经常碰到的随机变量平方和问题，我们自然希望能够对其加以总结，卡方分布就是在类似的实际背景下提出的.

(1) 定义

定义5.6 设$X_1, X_2, \cdots, X_n$是来自总体$N(0,1)$的样本，则称统计量$\chi_n^2 = X_1^2 + X_2^2 + \cdots + X_n^2$服从自由度为$n$的$\chi^2$分布.

自由度:

指 $\chi_n^2 = X_1^2 + X_2^2 + \cdots + X_n^2$ 中右端包含独立变量的个数.

(2) χ_n^2分布的概率分布

定理5.4 χ_n^2分布的概率密度：

$$p(x)=\begin{cases}\dfrac{1}{2^{\frac{n}{2}}\Gamma(\frac{n}{2})}x^{\frac{n}{2}-1}e^{-\frac{x}{2}} & x>0\\ 0 & \text{其它}\end{cases}$$

证 因为 $\chi^2(1)$ 分布即为$\Gamma\left(\dfrac{1}{2},\dfrac{1}{2}\right)$分布,

又因为 $X_i \sim \Gamma$

即 $X_i^2 \sim \Gamma\left(\dfrac{1}{2},\right.$

定义设随机变量X的分布密度函数为

$$f(x;\alpha,\beta)=\begin{cases}\dfrac{\beta^\alpha}{\Gamma(\alpha)}x^{\alpha-1}\mathrm{e}^{-\beta x}, & x>0,\\ 0, & x\le 0,\end{cases}$$

则称X服从Γ分布，记为$X\sim\Gamma(\alpha,\beta)$,其中$\alpha>0,\beta>0$为参数，$\Gamma$分布族常记为$\{\Gamma(\alpha,\beta):\alpha>0,\beta>0\}$.

因为$X_1, X_2, \cdots, X_n$相互独立,

所以 $X_1^2, X_2^2, \cdots, X_n^2$也相互独立,

根据Γ分布的可加性知 $\chi_n^2 = \sum_{i=1}^{n} X_i^2 \sim \Gamma\left(\frac{n}{2}, \frac{1}{2}\right).$

性质2（可加性） 若$X_j \sim \Gamma(\alpha_j, \beta), j = 1, 2, \cdots, n,$

而且X_j间相互独立，则

$$\sum_{j=1}^{n} X_j \sim \Gamma(\sum_{j=1}^{n} \alpha_j, \beta),$$

(3) χ^2 分布的性质

性质1 (χ^2 分布的可加性)

设 $Y_1 \sim \chi^2(n_1)$, $Y_2 \sim \chi^2(n_2)$, 并且 Y_1, Y_2 独立, 则 $Y_1 + Y_2 \sim \chi^2(n_1 + n_2)$.

(此性质可以推广到多个随机变量的情形)

设 $Y_i \sim \chi^2(n_i)$, 并且 Y_i $(i = 1, 2, \cdots, m)$ 相互独立, 则 $\sum\limits_{i=1}^{m} Y_i \sim \chi^2(n_1 + n_2 + \cdots + n_m)$.

性质2（χ^2分布的数学期望和方差）

若 $\chi_n^2 \sim \chi^2(n)$, 则 $E(\chi_n^2) = n,\quad D(\chi_n^2) = 2n.$

证 因为 $X_i \sim N(0,1)$, 所以

$$E(X_i^2) = D(X_i) + [E(X_i)]^2 = 1,$$

$$E(X_i^4) = \int_{-\infty}^{+\infty} x^4 \cdot \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \mathrm{d}x = \frac{-2}{\sqrt{2\pi}} \int_0^{+\infty} x^3 \cdot \mathrm{d}e^{-\frac{x^2}{2}}$$

$$= \frac{-2}{\sqrt{2\pi}} \left[x^3 e^{-\frac{x^2}{2}} \Big|_0^{+\infty} - \int_0^{+\infty} 3x^2 \cdot e^{-\frac{x^2}{2}} \mathrm{d}x \right] = 3$$

$$\therefore \quad D(X_i^{\,2}) = E(X_i^{\,4}) - [E(X_i^{\,2})]^2$$

$$= 3 - 1 = 2, \qquad (i = 1, 2, \cdots, n)$$

故 $$E(\chi_n^2) = E\left(\sum_{i=1}^{n} X_i^{\,2}\right)$$

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

$$D(\chi_n^2) = D\left(\sum_{i=1}^{n} X_i^{\,2}\right) = \sum_{i=1}^{n} D(X_i^{\,2}) = 2n.$$

性质3 设$\chi_n^2 \sim \chi^2(n)$,则对任意x,有

$$\lim_{n\to\infty} P\left\{\frac{\chi_n^2 - n}{\sqrt{2n}} \le x\right\} = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt.$$

证 由假设和定义5.6,$\chi_n^2 = \sum_{i=1}^{n} X_i^2$,其中$X_1, X_2, \cdots, X_n$

独立且每个$X_i \sim N(0,1)$,因而$X_1^2, X_2^2, \cdots, X_n^2$独立同分布,且

$$E(X_i^2) = 1, \qquad D(X_i^2) = 2 \qquad (i = 1, 2, \cdots, n)$$

由中心极限定理得

$$\lim_{n\to\infty} P\{\frac{\chi_n^2 - n}{\sqrt{2n}} < x\}$$

$$= \lim_{n\to\infty} P\{\frac{\sum_{i=1}^{n} X_i^2 - n\mu}{\sqrt{n}\sigma} \le x\} = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt$$

即χ^2分布的极限分布是正态分布，也即当n

很大时，$\frac{\chi_n^2 - n}{\sqrt{2n}}$服从$N(0,1)$，进而$\chi_n^2 \to N(n,2n)$.

例1 设$X \sim N(0,4), Y \sim \chi^2(2)$，且$X,Y$相互独立，试求解$\dfrac{X^2}{4}+Y$的概率分布.

解 因为$X \sim N(0,4)$且X,Y相互独立，所以

$$\frac{X}{2} \sim N(0,1)$$

且$\dfrac{X^2}{4}$与Y相互独立

又因为$\dfrac{X^2}{4} \sim \chi^2(1)$，由可加性得

得
$$\frac{X^2}{4}+Y \sim \chi^2(3).$$

例2 设$X_1, X_2, \cdots, X_6$为来自正态总体$N(0,1)$的一组样本,求C_1, C_2使得

$$Y = C_1(X_1 + X_2)^2 + C_2(X_3 + X_4 + X_5 + X_6)^2$$

服从χ^2分布.

解 $X_1 + X_2 \sim N(0,2)$, 则 $Y_1 = \dfrac{X_1 + X_2}{\sqrt{2}} \sim N(0,1)$

同理 $X_3 + X_4 + X_5 + X_6 \sim N(0,4)$,

则 $Y_2 = \dfrac{X_3 + X_4 + X_5 + X_6}{\sqrt{4}} \sim N(0,1)$

又$\because$ $Y_1=\dfrac{X_1+X_2}{\sqrt{2}}$ 与 $Y_2=\dfrac{X_3+X_4+X_5+X_6}{\sqrt{4}}$

相互独立.

所以 $(\dfrac{X_1+X_2}{\sqrt{2}})^2+(\dfrac{X_3+X_4+X_5+X_6}{\sqrt{4}})^2$

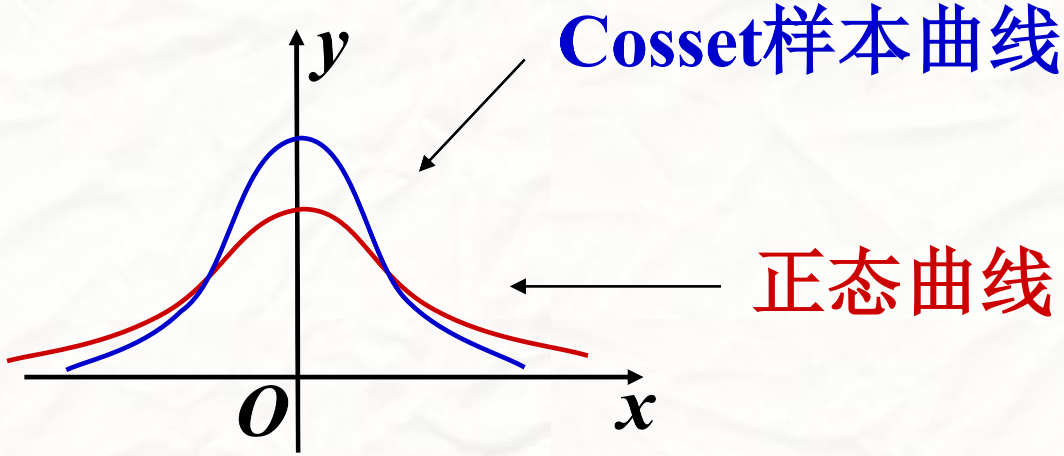
$=Y_1^2+Y_2^2 \sim \chi^2(2)$

则$C_1=1/2, C_2=1/4$.

2. t 分布

历史上，正态分布由于其广泛的应用背景和良好的性质，曾一度被看作是“万能分布”，在这样的背景下，十九世纪初英国一位年轻的酿酒化学技师Cosset. WS, 他在酒厂从事试验数据分析工作，对数据误差有着大量感性的认识，我们知道在总体均值和方差已知情况下，样本均值的分布将随样本量增大而接近正态分布,

但是Cosset在实验中遇到的样本容量仅有5~6个，在其中他发现实际数据的分布情况与正态分布有着较大的差异.



于是Cosset怀疑存在一个不属于正态的其他分布，通过学习终于得到了新的密度曲线，并在1908年以"Student"笔名发表了此项结果，后人称此分布为"*t* 分布"或"学生氏"分布.

(1) 定义

定义5.7　设 $X \sim N(0,1)$, $Y \sim \chi^2(n)$, 且 X, Y 独立,则称随机变量

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

服从自由度为 n 的 t 分布,　记为 $T \sim t(n)$.

t 分布又称**学生氏**

(Student)**分布**.

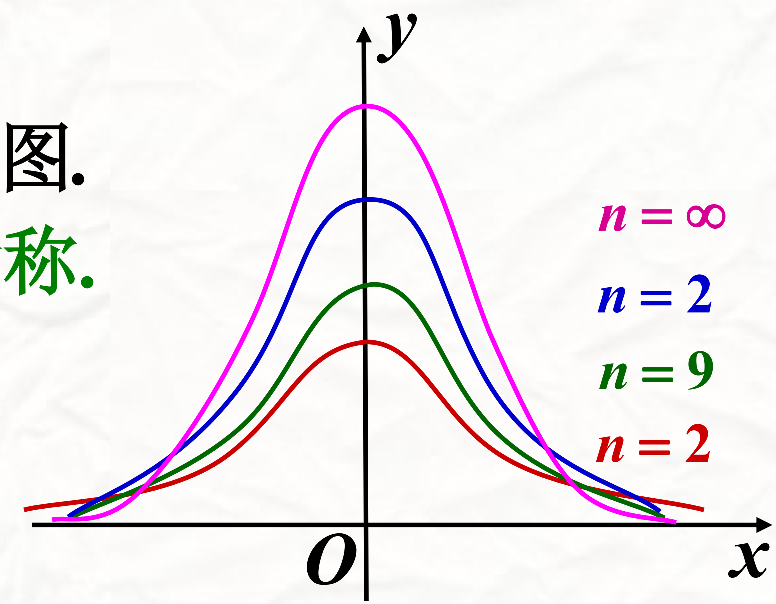
(2) $t(n)$ 分布的概率密度函数为

$$h(t)=\frac{\Gamma\left(\frac{n+1}{2}\right)}{\sqrt{\pi n}\,\Gamma\left(\frac{n}{2}\right)}\left(1+\frac{t^2}{n}\right)^{-\frac{n+1}{2}},\quad -\infty<t<+\infty$$

t分布的概率密度曲线如图.

显然图形是关于 $t=0$ 对称.

当n充分大时，其图形类似于标准正态变量概率密度的图形.



利用Stirling公式可以证明

$$\lim_{n\to\infty}\frac{\Gamma\left(\frac{n+1}{2}\right)}{\sqrt{n}\,\Gamma\left(\frac{n}{2}\right)}=\frac{1}{\sqrt{2}}$$

因为$\lim\limits_{n\to\infty} h(t)=\frac{1}{\sqrt{2\pi}}e^{-\frac{t^2}{2}}$,

所以当n足够大时t分布近似于$N(0,1)$分布,

但对于较小的n，t分布与$N(0,1)$分布相差很大.

(3) T的数字特征

$$E(T)=0,$$

$$D(T)=\frac{n}{n-2}\qquad (n>2).$$

目 上 下 回 停 续

例3 设总体X和Y相互独立,且都服从N（0,9）$X_1,X_2,\cdots,X_9$和$Y_1,Y_2,\cdots,Y_9$来自总体X,Y的样本,求统计量T的分布，其中

$$T=\sum_{i=1}^{9}X_i\Big/\sqrt{\sum_{i=1}^{9}Y_i^2}.$$

解 从抽样分布知$\overline{X}\sim N(0,1)$

而 $Y_i\sim N(0,9)$,故$Y_i/3\sim N(0,1)$,

从而 $(\frac{Y_i}{3})^2\sim\chi^2(1), i=1,2,\cdots,9.$

目 上 下 例 停 续

由可加性知 $\sum_{i=1}^{9}\left(\frac{Y_i}{3}\right)^2 \sim \chi^2(9)$

于是由t 的定义有 $\dfrac{\overline{X}}{\sqrt{\frac{1}{9}\sum_{i=1}^{9} Y_i^{2}/9}} = \dfrac{9\overline{X}}{\sqrt{\sum_{i=1}^{9} Y_i^{2}}} \sim t(9)$

即 $$T = \frac{\sum_{i=1}^{9} X_i}{\sqrt{\sum_{i=1}^{9} Y_i^{2}}} \sim t(9).$$

3. F分布

(1) 定义

定义5.8 设$X \sim \chi^2(n_1)$, $Y \sim \chi^2(n_2)$,且X,Y相互独立,

则称随机变量

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

服从自由度为(n_1, n_2)的F分布，记为

$$F \sim F(n_1, n_2).$$

(2) $F(n_1, n_2)$分布的概率密度为

$$p(y)=\begin{cases}\dfrac{\Gamma\left(\dfrac{n_1+n_2}{2}\right)\left(\dfrac{n_1}{n_2}\right)^{\frac{n_1}{2}} y^{\frac{n_1}{2}-1}}{\Gamma\left(\dfrac{n_1}{2}\right)\Gamma\left(\dfrac{n_2}{2}\right)\left[1+\left(\dfrac{n_1 y}{n_2}\right)\right]^{\frac{n_1+n_2}{2}}}, & y>0\\ 0, & \text{其它}\end{cases}$$

(3) F分布有以下性质

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

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

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

3) 设$F \sim F(n_1, n_2)$,则当$n_2 > 4$时,对任意x有

$$\lim_{n_1 \to \infty} P\left\{\frac{F - E(F)}{\sqrt{D(F)}} \le x\right\} = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt$$

这说明F分布极限分布也是正态分布.

例4 已知 $T \sim t(n)$, 试证 $T^2 \sim F(1,n)$.

证 因为 $T \sim t(n)$, 由定义5.7有

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

其中 $X \sim N(0,1), Y \sim \chi^2(n)$, 且 X, Y 独立,

从而 $X^2 \sim \chi^2(1)$, 且 X^2 与 Y 独立,

$\therefore$ 由定义5.8, 有 $T^2 = \dfrac{X^2}{Y/n} \sim F(1,n)$.

例5 设$X \sim F(n,m)(n>4)$,试求EX^{-1}, DX^{-1}.

解 因为 $X \sim F(n,m)$,所以

由F分布的性质知

$$X^{-1} \sim F(m,n)$$

所以得

$$EX^{-1} = \frac{n}{n-2},$$

$$DX^{-1} = \frac{n^2(2m+2n-4)}{m(n-2)^2(n-4)}.$$

二、概率分布的分位数

1. 定义

定义5.9 对于总体X和给定的$\alpha(0<\alpha<1)$,

若存在x_α,使

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

则称x_α为X的分布的上侧α分位数.

2. 常用分布的上侧分位数记号

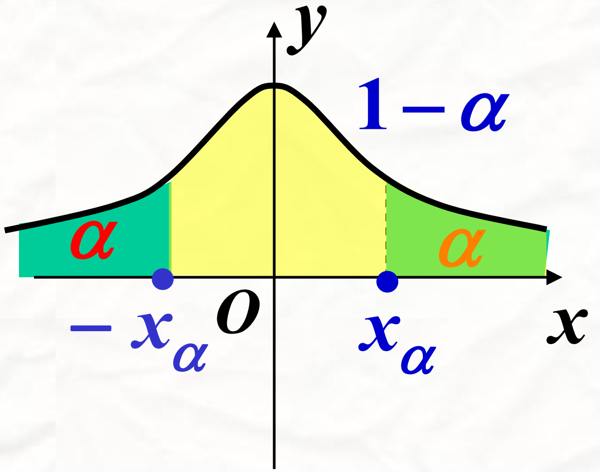
分布	$N(0,1)$	$\chi^2(n)$	$t(n)$	$F(n_1,n_2)$
记号	u_α	$\chi^2_\alpha(n)$	$t_\alpha(n)$	$F_\alpha(n_1,n_2)$

目 上 下 回 停 续

3. 查表法

(1) 若X的分布密度关于y轴对称，则

$$x_{1-\alpha} = -x_{\alpha}$$



特例：

1) $N(0,1)$: $u_{1-\alpha} = -u_{\alpha}$

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

1) 正态分布的上侧分位数u_α:

设X服从标准正态分布$N(0,1)$,则其上侧分位数u_α满足

$$P\{X > u_\alpha\} = \frac{1}{\sqrt{2\pi}}\int_{u_\alpha}^{+\infty} e^{-\frac{x^2}{2}}\,\mathrm{d}x$$

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

即 $\Phi(u_\alpha) = 1 - \alpha$

给定α,由附表2可查得u_α的值.

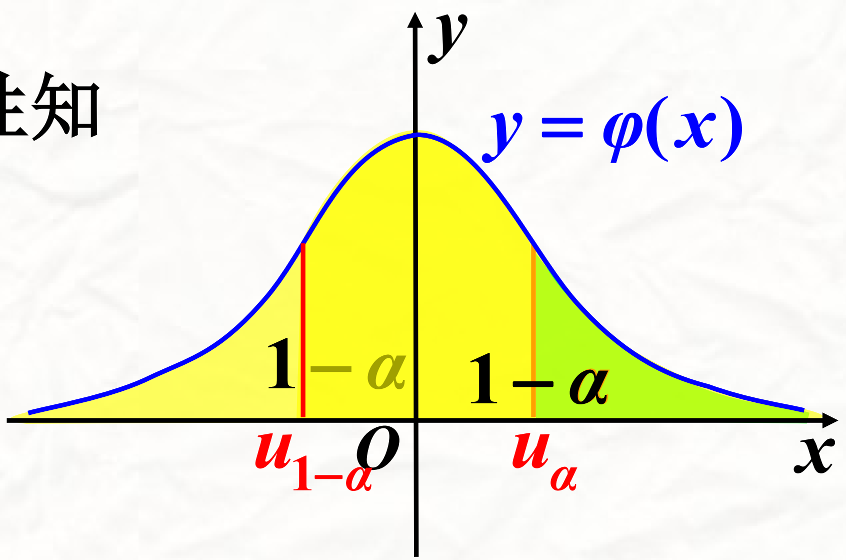
$$\boxed{\Phi(u_\alpha) = 1 - \alpha}$$

$u_{0.05} = 1.645,$ 0.95 $(\alpha = 0.05)$

$u_{0.025} = 1.96,$ 0.975 $(\alpha = 0.025)$

根据正态分布的对称性知

$$u_{1-\alpha} = -u_\alpha.$$



2) t分布的上分位 $t_\alpha(n)$:

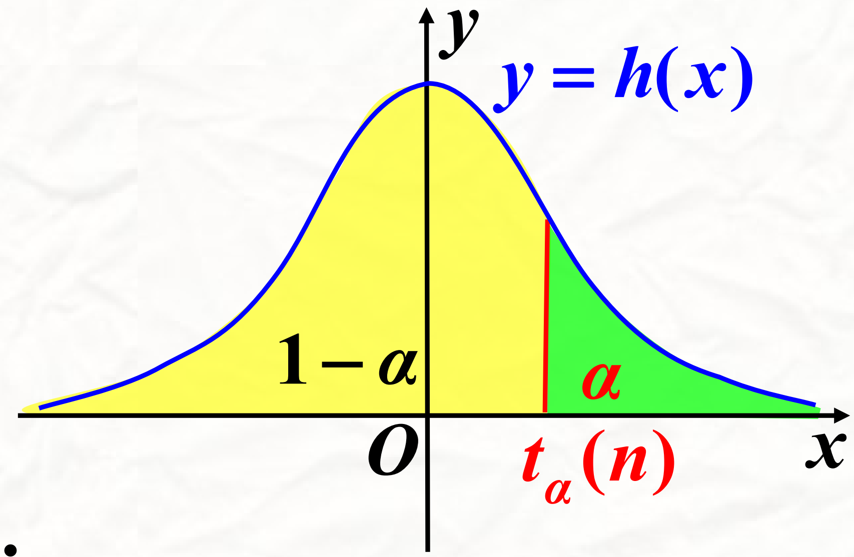
对于给定的α, $0<\alpha<1$, 称满足条件

$$P\{t>t_\alpha(n)\}=\int_{t_\alpha(n)}^{\infty}h(t)\mathrm{d}t=\alpha$$

的点$t_\alpha(n)$为$t(n)$分布的上α分位点.

可以通过查表求

得上α分位点的值.



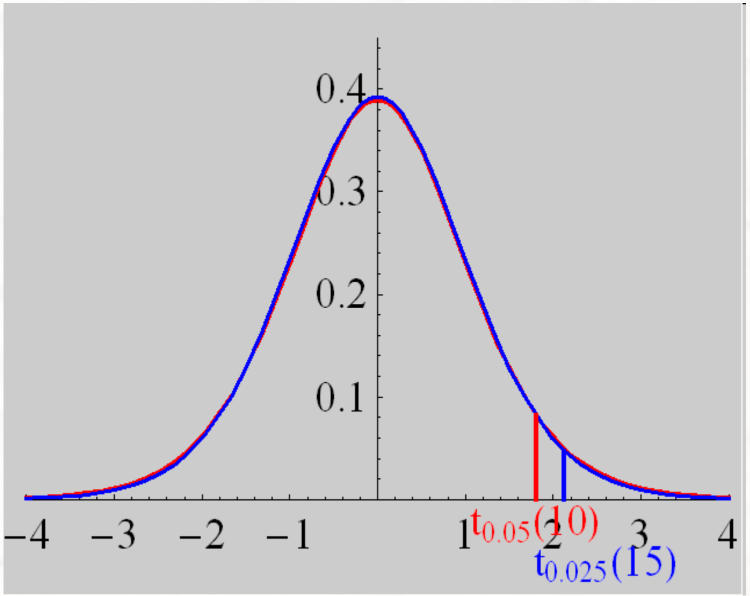
由分布的对称性知

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

当$n>45$时, $t_\alpha(n)\approx u_\alpha$.

$$t_{0.05}(10) = 1.8125,$$

$$t_{0.025}(15) = 2.1315.$$



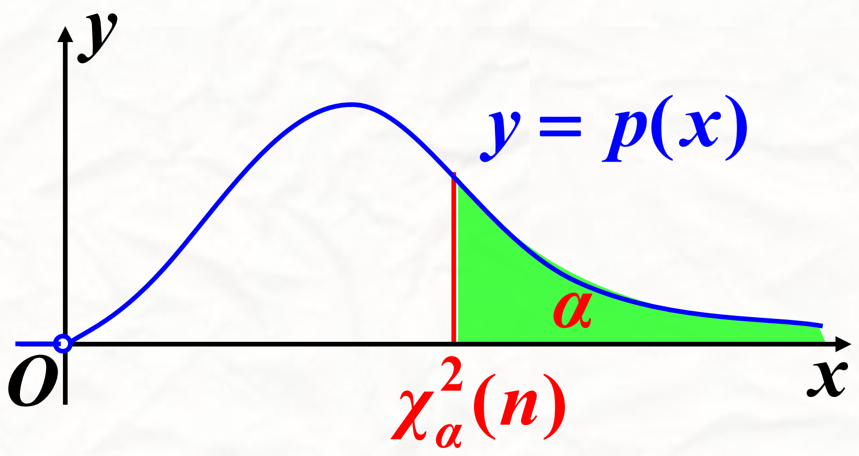
目 上 下 回 停 续

(2) X的分布密度无对称性的情形

1) $\chi_\alpha^2(n)$**:** 对于给定的正数α, $0<\alpha<1$, 称满足

$$P\{\chi^2 > \chi_\alpha^2(n)\} = \int_{\chi_\alpha^2(n)}^{\infty} p(y)\,\mathrm{d}y = \alpha$$

的点 $\chi_\alpha^2(n)$ 为 $\chi^2(n)$ 分布的上侧分位数.

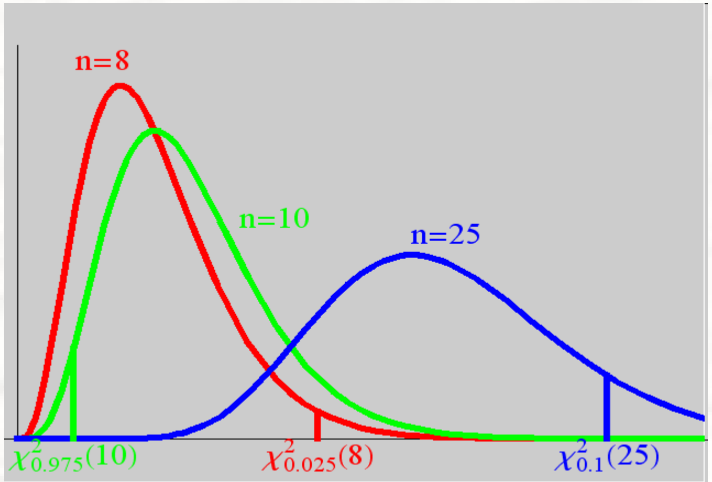


当$n \le 60$时，可查表4(表4只详列到 n=60 为止).

$$\chi^2_{0.025}(8) = 17.535,$$

$$\chi^2_{0.975}(10) = 3.247,$$

$$\chi^2_{0.1}(25) = 34.382.$$



当$n > 60$时，$\chi^2_\alpha(n) \approx n + \sqrt{2n}\, u_\alpha$.

费歇(R.A.Fisher)公式：

费歇资料

当 n 充分大时, $\chi^2_\alpha(n) \approx n + \sqrt{2n} u_\alpha$.

其中 u_α 是标准正态分布的上 α 分位点.

例如：$\chi^2_{0.05}(120) \approx 120 + \sqrt{2 \times 120} \times u_{0.05}$

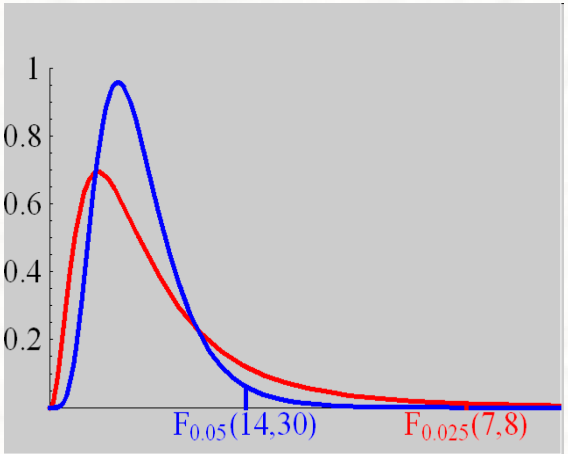
$$= 120 + \sqrt{240} \times 1.64$$

$$= 145.5.$$

2) $F_\alpha(n_1,n_2)$: 对于 $\alpha = 0.01, \ 0.025, \ 0.05, \ 0.1$ 等，可直接查表5 ~ 8.

$$F_{0.05}(14,30) = 2.04 .$$

$$F_{0.025}(7,8) = 4.53 ,$$



此外，还可利用关系

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

由F_α求得$F_{1-\alpha}$.

如： $$F_{0.95}(12,9) = \frac{1}{F_{0.05}(9,12)} = \frac{1}{2.8} = 0.357 .$$

目 上 下 回 停 续

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

证 因为$F \sim F(n_1, n_2)$,

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

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

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

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

因为$\dfrac{1}{F} \sim F(n_2, n_1)$,

所以$P\left\{\dfrac{1}{F} > F_{\alpha}(n_2, n_1)\right\} = \alpha$,

比较后得$\dfrac{1}{F_{1-\alpha}(n_1, n_2)} = F_{\alpha}(n_2, n_1)$,

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

内容小结

1.三大抽样分布:

$$\chi^2\text{ 分布},\ t\text{ 分布},\ F\text{ 分布}$$

的定义,性质.

2.概率分布的分位数概念.

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

目 上 下 回 停 续


再见

备用题

例1-1

设$(X_1,\cdots X_n)$为来自于正态总体$N(\mu,\sigma^2)$的样本，

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

解 $\frac{X_i-\mu}{\sigma}\sim N(0,1), i=1,\cdots,n,$ 且它们独立.

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

例1-2 设$X_1, X_2, \cdots, X_n$来自正态分布N（$0,\sigma^2$）的样本，

试求 $Y=\sum_{i=1}^{n} X_i^2$的分布函数.

解 令 $T=\dfrac{Y}{\sigma^2}$,则由χ^2分布性质知 $T \sim \chi_n^2$

其中χ_n^2表自由度为n的χ^2分布函数.

所以Y的分布函数为

$$F(y)=P(Y \le y)=P(T \le \frac{y}{\sigma^2})$$

$$=\chi_n^2(\frac{y}{\sigma^2}), y \ge 0.$$

目 上 下 回 停 续

相应的由公式法可得，密度函数为

密度变换公式

$$p(y)=\frac{1}{2\Gamma(n/2)}\left(\frac{y}{\sigma^2}\right)^{\frac{n}{2}-1}e^{-\frac{y}{2\sigma^2}}\cdot\frac{1}{\sigma^2}$$

$$=\frac{1}{2\sigma^n\Gamma(n/2)}y^{\frac{n}{2}-1}e^{-\frac{y}{2\sigma^2}},\ y\geq 0.$$

$X\sim p(x), Y=f(x)$且单调连续，则Y的

密度函数为 $p[f^{-1}(y)]\cdot\left|[f^{-1}(y)]'\right|.$

例2-1 设总体为标准正态分布，从中抽取n个样本,$\overline{X}, S^2$分别为样本均值与方差，则

$(A)\overline{X} \sim N(0,1)$　　$(B) n\overline{X} \sim N(0,1)$

$(C)\sum_{i=1}^{n} X_i^2 \sim \chi^2(n)$　　$(D)\overline{X}/S^2 \sim t(n-1)$

解 因$\overline{X}$为样本均值，由独立正态变量的线性组合仍为正态随机变量

所以 $\overline{X} \sim N(0,\frac{1}{n})$ 相应的 $n\overline{X} \sim N(0,n)$

因为$X_i \sim N(0,1)$,由卡方分布的定义有

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

因为$\sqrt{n}\overline{X} \sim N(0,1)$且$nS^2 \sim \chi^2(n-1)$

所以，由$\sqrt{n}\overline{X}$与nS^2的独立性有

$$\frac{\sqrt{n-1}\overline{X}}{S} = \frac{\sqrt{n}\overline{X}}{\sqrt{nS^2/(n-1)}} \sim t(n-1)$$

综上可得,正确答案为C.

例3-1 设$X \sim N(\mu,\sigma^2)$, $\dfrac{Y}{\sigma^2} \sim \chi^2(n)$,且$X,Y$相互独立,

试求 $T = \dfrac{X-\mu}{\sqrt{Y/n}}$

的概率分布.

解 $\because \quad X \sim N(\mu,\sigma^2), \quad \therefore \quad \dfrac{X-\mu}{\sigma} \sim N(0,1)$

又$\dfrac{Y}{\sigma^2} \sim \chi^2(n)$,且$X,Y$独立,则$\dfrac{X-\mu}{\sigma}$与$\dfrac{Y}{\sigma^2}$独立,

由定义5.7, $T = \dfrac{X-\mu}{\sqrt{Y/n}} = \dfrac{(X-\mu)/\sigma}{\sqrt{(Y/\sigma^2)/n}} \sim t(n).$

例3-2 设X_1，$X_2,\cdots$，X_n是来自正态分布$N(\mu,\sigma^2)$的样本，$\overline{X}$和S_n^2是样本均值和方差，又设X_{n+1}服从$N(\mu,\sigma^2)$,分布，且与$X_1,\cdots,X_n$相互独立，试求

$$T=\frac{X_{n+1}-\overline{X}}{S_n}\sqrt{\frac{n-1}{n+1}}$$

的概率分布.

解 因为$X_{n+1}-\overline{X}\sim N(0,\dfrac{n+1}{n}\sigma^2)$

所以 $\dfrac{X_{n+1}-\overline{X}}{\sqrt{\dfrac{n+1}{n}\sigma^2}}\sim N$（0,1）

又 $\dfrac{nS_n^2}{\sigma^2} \sim \chi^2(n-1)$

且 $X_{n+1} - \dfrac{\overline{X}}{\sigma\sqrt{\dfrac{n+1}{n}}}$ 与 $\dfrac{nS_n^2}{\sigma^2}$ 相互独立

故 $$\frac{X_{n+1} - \overline{X}}{S_n^2}\sqrt{\frac{n-1}{n+1}} = \frac{\sigma\sqrt{\dfrac{n+1}{n}}}{\sqrt{\dfrac{nS_n^2}{\sigma^2(n-1)}}} \quad \sim t(n-1).$$

例3-3 设$X \sim N(0,1), Y \sim \chi^2(4)$,令$T = \dfrac{X}{\sqrt{Y}}$

则$ET = \underline{\ 0\ }, DT = \underline{1/2}$

解 因为 $2T = \dfrac{X}{\sqrt{Y/4}} \sim t(4)$,所以

$$ET = \frac{1}{2}E2T = 0$$

$$DT = \frac{1}{4}D(2T) = \frac{1}{4} \cdot \frac{4}{4-2} = \frac{1}{2}.$$

例3-4 设$X_1,\cdots,X_m$和$Y_1,\cdots,Y_n$分别来自正态总体$N(\mu_1,\sigma^2)$和$N(\mu_2,\sigma^2)$且相互独立，试求

$$T=\frac{\alpha(\overline{X}-\mu_1)-\beta(\overline{Y}-\mu_2)}{\sqrt{\dfrac{mS_1^2+nS_2^2}{m+n-2}}\sqrt{\dfrac{\alpha^2}{m}+\dfrac{\beta^2}{n}}},\alpha,\beta\text{为实数}$$

的概率分布.

解 由于$\overline{X},\overline{Y}$服从正态分布且独立,

所以$\overline{X}\sim N(\mu_1,\dfrac{\sigma^2}{m}),\overline{Y}\sim N(\mu_2,\dfrac{\sigma^2}{n})$

因此$\overline{X}-\mu_1\sim N(0,\dfrac{\sigma^2}{m}),\overline{Y}-\mu_2\sim N(0,\dfrac{\sigma^2}{n})$

又因为 $E(\overline{X}-\overline{Y})=E[\alpha(\overline{X}-\mu_1)+\beta(\overline{Y}-\mu_2)]=0$

$$D(\overline{X}-\overline{Y})=D[\alpha(\overline{X}-\mu_1)+\beta(\overline{Y}-\mu_2)]$$

$$=(\frac{\alpha^2}{m}+\frac{\beta^2}{n})\sigma^2$$

由于独立正态变量的线性组合仍是正态变量

故 $\alpha(\overline{X}-\mu_1)+\beta(\overline{Y}-\mu_2)\sim N(0,(\frac{\alpha^2}{m}+\frac{\beta^2}{n})\sigma^2)$

整理得

$$U=\frac{1}{\sqrt{\alpha^2/m+\beta^2/n}}\cdot\frac{\alpha(\overline{X}-\mu_1)+\beta(\overline{Y}-\mu_2)}{\sigma}\sim N(0,1)$$

由卡方分布的定义知

我们有 $$\frac{mS_1^2}{\sigma^2} \sim \chi^2(m-1),$$

$$\frac{nS_2^2}{\sigma^2} \sim \chi^2(n-1)$$

且它们相互独立, 再利用伽玛分布的可加性知

$$V = \frac{mS_1^2}{\sigma^2} + \frac{nS_2^2}{\sigma^2} \sim \chi^2(m+n-2)$$

从而，由t分布的定义有

$$T=\frac{\alpha(\bar{X}-\mu_1)-\beta(\bar{Y}-\mu_2)}{\sqrt{\dfrac{mS_1^2+nS_2^2}{m+n-2}}\sqrt{\dfrac{\alpha^2}{m}+\dfrac{\beta^2}{n}}}$$

$$=\frac{U}{\sqrt{\dfrac{V}{m+n-2}}}\sim t(m+n-2).$$

注　本例要求两个正态总体的方差相同！

例3-5 设$X,Y \sim N(0,1)$且相互独立，试求统计量T的分布函数，其中$T=\dfrac{X}{Y}$.

解 因为 $Y \sim N(0,1)$且与X独立，

所以 $Y^2 \sim \chi^2(1)$，且X与Y^2独立

故由t的定义有

$$T=\frac{X}{Y}=\frac{X}{\sqrt{Y^2/1}} \sim t(1)$$

因而T的分布密度为

$$p(t)=\frac{1}{\sqrt{n\pi}}(1+t^2)^{-1}, -\infty < t < +\infty$$

例4-1 设$X_1,\cdots,X_n,X_{n+1},\cdots,\ X_{n+m}$是来自正态总体$N(0,\sigma^2)$的容量为$m+n$的样本,试求$F$的分布，其中

$$F=\frac{m\sum_{i=1}^{n}X_i^2}{n\sum_{i=n+1}^{m+n}X_i^2}.$$

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

且 $\sum_{i=1}^{n}\left(\frac{X_i}{\sigma}\right)^2$与$\sum_{i=n+1}^{m+n}\left(\frac{X_i}{\sigma}\right)^2$相互独立,

所以

$$\frac{\dfrac{m\sum_{i=1}^{n}X_i^2}{m+n}}{\dfrac{n\sum_{i=n+1}^{m+n}X_i^2}{}} = \frac{\dfrac{\sum_{i=1}^{n}X_i^2}{n\sigma^2}}{\dfrac{\sum_{i=n+1}^{m+n}X_i^2}{m\sigma^2}} \sim F(n,m).$$

例4-2 设X和$Y_1,\cdots,Y_n$分别来自正态总体$N(0,\sigma_1^2)$ $N(\mu_2,\sigma_2^2)$且相互独立的样本，试求

$$F=\frac{X^2\sigma_2^2}{S_n^{*2}\sigma_1^2}\text{，其中}S_n^{*2}=\frac{1}{n-1}\sum_{i=1}^{n}(Y_i-\overline{Y})^2$$

的概率分布.

解 由卡方分布的定义有

$$\frac{X^2}{\sigma_1^2}\sim\chi^2(1)$$

$$\frac{(n-1)S_n^{*2}}{\sigma_2^2}=\sum_{i-1}^{n}\left(\frac{Y_i-\overline{Y}}{\sigma_2}\right)^2\sim\chi^2(n-1)$$

又因为X^2与$\dfrac{(n-1)S_n^{*2}}{\sigma_2^2}$相互独立,

由F分布性质知

$$F=\frac{X^2\sigma_2^2}{S_n^{*2}\sigma_1^2}=\frac{\dfrac{X^2}{\sigma_1^2}/1}{\dfrac{(n-1)S_n^{*2}}{\sigma_2^2}/(n-1)}\sim F(1,n-1).$$