第一章

统计学的基本概念(2)

§1.4　统计量

统计学的基本任务之一是利用样本所提供的信息对总体分布中未知的量进行推断．但是，样本观测值常常表现为一大堆数字，很难直接用来解决我们所要研究的具体问题．因此，我们需要对数据进行加工，使其成为一些简单明了的数字特征，这些加工后的数字特征就称为是统计量的观测值.

定义　　不含有未知参数的样本的函数 $g(X_1, X_2, \cdots, X_n)$ 统称为统计量.

用途

1.估计问题：用统计量的观测值去估计总体分布中的未知参数的数值;

2.检验问题：在假设检验问题中用作检验统计量.

例12 假设总体 $X \sim N(\mu,\sigma^2)$, $(X_1,X_2,\cdots,X_n)$为取自该总体的一个样本, 其中参数 μ 未知, 而 σ^2已知. 设 $n=4$, 试判别下列各函数表达式是否为统计量?

(1) $\sum_{i=1}^{4} X_i^2$; (2) $\frac{1}{4}\sum_{i=1}^{4}(X_i-\mu)^2$; (3) $\frac{1}{\sigma^2}\sum_{i=1}^{4}(X_i-\bar{X})^2$.

其中 $\bar{X}=\frac{1}{n}\sum_{i=1}^{n} X_i$.

结论: (1)(3)是统计量, (2)不是.

常用统计量

(1)样本均值 $\bar{X}=\frac{1}{n}\sum_{i=1}^{n}X_i$，样本均值观测值 $\bar{x}=\frac{1}{n}\sum_{i=1}^{n}x_i$；

(2)样本方差 $S^2=\frac{1}{n}\sum_{i=1}^{n}(X_i-\bar{X})^2$，样本方差的观测值

$$s^2=\frac{1}{n}\sum_{i=1}^{n}(x_i-\bar{x})^2;$$

(3)样本标准差 $S=\sqrt{\frac{1}{n}\sum_{i=1}^{n}(X_i-\bar{X})^2}$，样本标准差的观测值

$$s=\sqrt{\frac{1}{n}\sum_{i=1}^{n}(x_i-\bar{x})^2};$$

(4)修正的样本方差 $S^{*2}=\frac{1}{n-1}\sum_{i=1}^{n}(X_i-\bar{X})^2$，修正的样本方差观测值 $s^*=\sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(x_i-\bar{x})^2}$ ；

显然有 $(n-1)S^{*2}=nS^2=\sum_{i=1}^{n}(X_i-\bar{X})^2$ ，

且 $\sum_{i=1}^{n}(X_i-\bar{X})^2=\sum_{i=1}^{n}(X_i^2-2X_i\bar{X}+\bar{X}^2)=\sum_{i=1}^{n}X_i^2-n\bar{X}^2$ ；

故有 $S^{*2}=\frac{1}{n-1}(\sum_{i=1}^{n}X_i^2-n\bar{X}^2), S^2=\frac{1}{n}\sum_{i=1}^{n}X_i^2-\bar{X}^2$.

例13 设有样本观测值 0, 1, 1, 1, 2, 2, 3, 4, 4，试计算样本均值与修正的样本方差的观测值 $\overline{x}$, s^{*2}.

解 $$\overline{x}=\frac{1}{9}\times(0+1\times3+2\times2+3+4\times2)=2;$$

$$\sum_{i=1}^{9}x_i^2=1^2\times3+2^2\times2+3^2+4^2\times2=52,$$

所以 $$s^{*2}=\frac{1}{8}(52-9\times2^2)=2.$$

常用统计量（续）

(5)样本的 k 阶原点矩 $A_k = \frac{1}{n}\sum_{i=1}^{n} X_i^k$ （可知 $A_1 = \bar{X}$）

(6)样本的 k 阶中心矩 $M_k = \frac{1}{n}\sum_{i=1}^{n} (X_i - \bar{X})^k$ $(M_2 = S^2)$

其中 k 是任意正整数，且每个统计量也都有相应的观测值.

定理1.2 设总体 X 的均值 $E(X)=\mu$, 方差 $D(X)=\sigma^2$, $(X_1,X_2,\cdots,X_n)$ 为取自该总体的一个样本, 则

(1) $E(\bar{X})=E(X)=\mu$, $D(\bar{X})=\dfrac{D(X)}{n}=\dfrac{\sigma^2}{n}$;

(2) $E(S^2)=\dfrac{n-1}{n}\cdot D(X)=\dfrac{n-1}{n}\sigma^2$, $n\geq 2$;

(3) $E(S^{*2})=D(X)=\sigma^2$.

证第(2)条：$E(S^2)=\dfrac{n-1}{n}\sigma^2$.

$$nE(S^2)=E(\sum_{i=1}^{n}(X_i-\bar{X})^2)=E(\sum_{i=1}^{n}X_i^2-n\bar{X}^2)$$

$$=\sum_{i=1}^{n}E(X_i^2)-nE(\bar{X}^2)$$

$$=n\{D(X_1)+E^2(X_1)-D(\bar{X})-E^2(\bar{X})\}$$

$$=n(\sigma^2-\frac{\sigma^2}{n})=(n-1)\sigma^2$$

$$\Rightarrow E(S^2)=\frac{n-1}{n}\sigma^2$$

	$B(1,p)$	$P(\lambda)$	$R(a,b)$	$E(\lambda)$	$N(\mu,\sigma^2)$
$E(X)$	p	λ	$\frac{a+b}{2}$	$\frac{1}{\lambda}$	μ
$D(X)$	$p(1-p)$	λ	$\frac{(b-a)^2}{12}$	$\frac{1}{\lambda^2}$	σ^2
$E(\bar{X})$	p	λ	$\frac{a+b}{2}$	$\frac{1}{\lambda}$	μ
$D(\bar{X})$	$\frac{p(1-p)}{n}$	$\frac{\lambda}{n}$	$\frac{(b-a)^2}{12n}$	$\frac{1}{n\lambda^2}$	$\frac{\sigma^2}{n}$
$E(S^{*2})$	$p(1-p)$	λ	$\frac{(b-a)^2}{12}$	$\frac{1}{\lambda^2}$	σ^2

例14　设总体 $X \sim R(-1,3)$, $(X_1, X_2, \cdots, X_{10})$ 为来自该总体的一个样本, 试写出 $E(\bar{X})$, $D(\bar{X})$, $E(S^{*2})$.

解　由题设条件知 $E(X)=1, D(X)=\dfrac{16}{12}=\dfrac{4}{3}, n=10,$

由定理1.2即得 $E(\bar{X})=1, D(\bar{X})=\dfrac{1}{10}\cdot\dfrac{4}{3}=\dfrac{2}{15}, E(S^{*2})=\dfrac{4}{3}.$

例15 设总体 $X \sim N(1,4)$, $(X_1, X_2, \cdots, X_9)$ 为来自该总体的一个样本, 试写出 $E(\bar{X})$, $D(\bar{X})$, $E(S^{*2})$.

解 由题设条件知 $E(X)=1, D(X)=4, n=9,$

由定理1.2 $E(\bar{X})=1, D(\bar{X})=\dfrac{1}{9}\times 4=\dfrac{4}{9}, E(S^{*2})=4.$

(7)次序统计量

设$(X_1, X_2, \cdots, X_n)$为取自总体 X 的样本，其中取值最大的记为 $X_{(n)}$，称为最大次序统计量；而取值最小的记为 $X_{(1)}$，称为最小次序统计量；$X_{(i)}$ 称为第 i 个次序统计量，$i = 1, 2, \cdots, n$；即次序统计量 $X_{(1)}, X_{(2)}, \cdots, X_{(n)}$ 总是满足

$$X_{(1)} \le X_{(2)} \le \cdots \le X_{(n)}.$$

相应地，次序统计量的观测值满足 $x_{(1)} \le x_{(2)} \le \cdots \le x_{(n)}$.

(8)按大小排列，处于中间位置的次序统计量称为样本中位数，记为 $\tilde{X}$，也即规定

$$\tilde{X}=\begin{cases} X_{(\frac{n+1}{2})}, & \text{当}n\text{为奇数,} \\ X_{(\frac{n}{2}+1)}, & \text{当}n\text{为偶数.} \end{cases}$$

(9)称 $R \triangleq X_{(n)}-X_{(1)}$ 为极差；

(10)称 $\frac{1}{2}(X_{(1)}+X_{(n)})$ 为极差中值.

§1.5　三个常用分布

1.卡方分布

2.T分布

3.F分布

要求掌握：1.定义；2.常见性质；3.分位数概念.

1.卡方分布

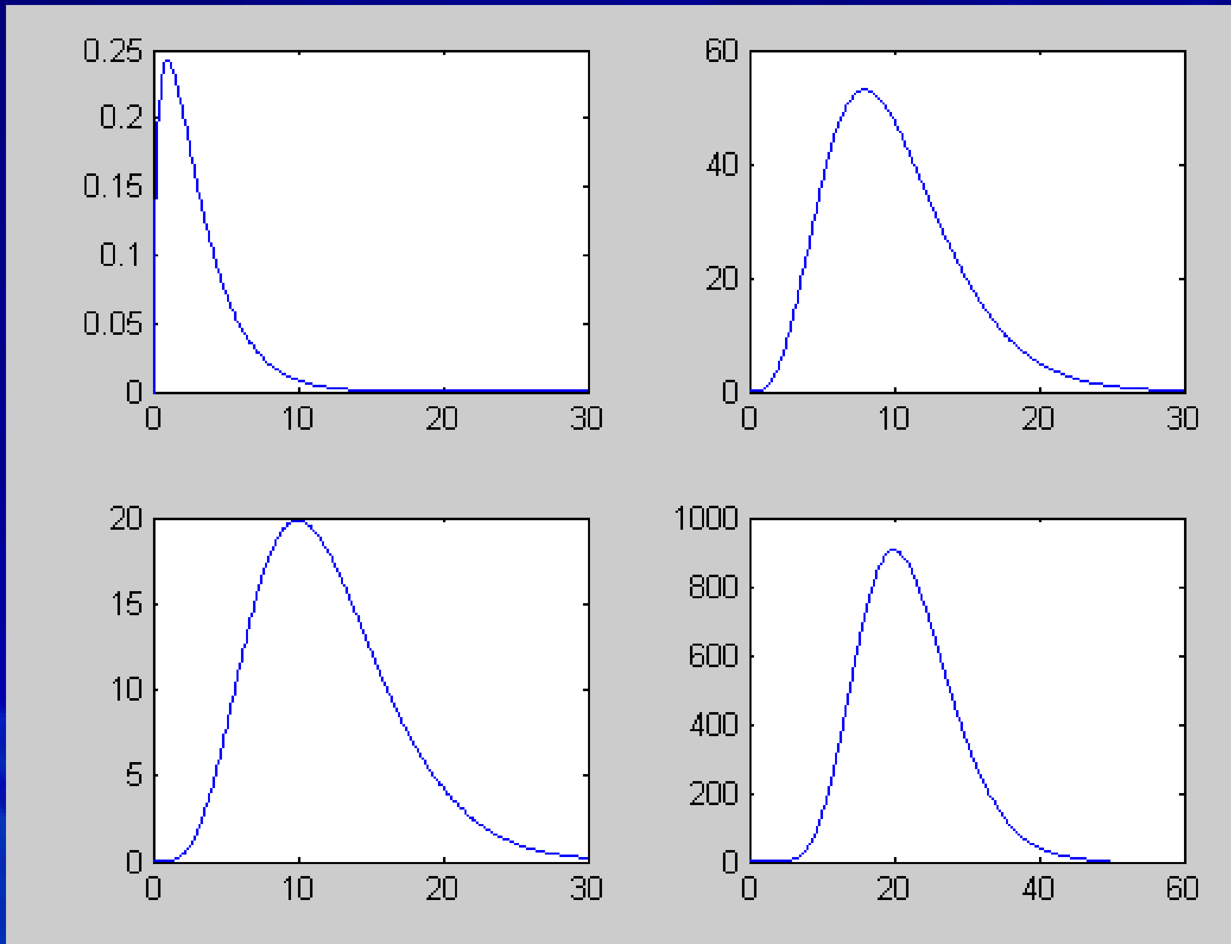
(1)定义

设 $X_1,\cdots,X_n$ 是独立同分布的随机变量，且都服从标准正态分布 $N(0,1)$，称随机变量 $Y \triangleq \sum_{i=1}^{n} X_i^2$ 所服从的分布为**自由度为 n 的 χ^2 分布**，记为 $Y \sim \chi^2(n)$.

例16 设 X_1, X_2 相互独立，都服从 $N(0,1)$，由定义可知

$$X_1^2 \sim \chi^2(1),\ X_2^2 \sim \chi^2(1),\ X_1^2 + X_2^2 \sim \chi^2(2).$$

下图分别是当 $n = 1, 4, 10, 20$ 时服从卡方分布的随机变量的概率密度函数图形.


0.25
0.2
0.15
0.1
0.05
0
0 10 20 30
60
40
20
0
0 10 20 30
20
15
10
5
0
0 10 20 30
1000
800
600
400
200
0
0 20 40 60

定理1.5 设随机变量 $X \sim \chi^2(n)$，则其概率密度函数为

$$f_n(x) = \begin{cases} \dfrac{1}{2^{\frac{n}{2}}\Gamma(\frac{n}{2})} \cdot x^{\frac{n}{2}-1} \cdot \mathrm{e}^{-\frac{x}{2}}, & x > 0, \\ 0, & \text{其余}. \end{cases}$$

利用卷积公式，用数学归纳法证明. 步骤如下：

(一)证 $n=1$ 时，结论成立，即 $f_1(x) = \dfrac{1}{2^{\frac{1}{2}}\Gamma(\frac{1}{2})} \cdot x^{\frac{1}{2}-1} \cdot \mathrm{e}^{-\frac{x}{2}}, \quad x > 0$

证明所需知识:

(1)伽玛函数 $\Gamma(\alpha) \triangleq \int_0^{\infty} x^{\alpha-1} \mathrm{e}^{-x} \mathrm{d}x$ ， $\alpha > 0$ ，且有性质

① $\Gamma(\alpha+1) = \alpha\Gamma(\alpha)$ ； ② $\Gamma(\frac{1}{2}) = \sqrt{\pi}$ ；

③ α 为正整数时， $\Gamma(\alpha) = (\alpha-1)!$.

(2)贝塔函数 $B(p,q) \triangleq \int_0^1 x^{p-1}(1-x)^{q-1} \mathrm{d}x$， $p,q > 0$ ，

且有 $B(p,q) = \dfrac{\Gamma(p)\Gamma(q)}{\Gamma(p+q)}$.

(3)当 X,Y 相互独立时, $X+Y \sim \int_{-\infty}^{\infty} f_X(x) f_Y(z-x) \mathrm{d}x$.

第一步即要证：若 $X \sim N(0,1)$，则 $Y = X^2 \sim \chi^2(1)$.

证：当 $y>0$ 时，注意到 $F_Y(y) = P(Y \le y) = P(X^2 \le y)$

$= P(-\sqrt{y} \le X \le \sqrt{y}) = 2\Phi(\sqrt{y}) - 1$，对自变量 y 求导

即得 $f_Y(y) = 2\varphi(\sqrt{y}) \cdot \dfrac{1}{2\sqrt{y}} = \dfrac{1}{\sqrt{2\pi} \cdot \sqrt{y}} \cdot \mathrm{e}^{-\frac{y}{2}}$

$$= \frac{1}{2^{\frac{1}{2}}\Gamma(\frac{1}{2})} \cdot y^{\frac{1}{2}-1} \cdot \mathrm{e}^{-\frac{y}{2}},\quad y>0.$$ 即证.

㈡归纳假定：设 $n=k$ 时，结论成立，即

$$f_k(x)=\frac{1}{2^{\frac{k}{2}}\Gamma(\frac{k}{2})}\cdot x^{\frac{k}{2}-1}\cdot \mathrm{e}^{-\frac{x}{2}},\quad x>0$$

㈢再证 $n=k+1$ 时，结论也成立，即

$$f_{k+1}(x)=\frac{1}{2^{\frac{k+1}{2}}\Gamma(\frac{k+1}{2})}\cdot x^{\frac{k+1}{2}-1}\cdot \mathrm{e}^{-\frac{x}{2}},\quad x>0$$

即要证明若 $X \sim \chi^2(k), Y \sim \chi^2(1)$ ，且相互独立，则

$$Z = X + Y \sim \chi^2(k+1)$$

利用卷积公式，当 X, Y 相互独立时，

$$Z \sim f_Z(z) = \int_{-\infty}^{\infty} f_X(x) f_Y(z-x)\mathrm{d}x$$

将归纳假定和第一步证明结果代入知，当 $z > 0$ 时

$$f_Z(z) = \int_0^z f_k(x) f_1(z-x)\mathrm{d}x$$

$$= \int_0^z \frac{1}{2^{\frac{k}{2}}\Gamma(\frac{k}{2})} x^{\frac{k}{2}-1} \mathrm{e}^{-\frac{x}{2}} \cdot \frac{1}{2^{\frac{1}{2}}\Gamma(\frac{1}{2})} \cdot (z-x)^{\frac{1}{2}-1} \cdot \mathrm{e}^{-\frac{z-x}{2}} \mathrm{d}x$$

$$=\frac{1}{2^{\frac{k+1}{2}}\Gamma(\frac{k}{2})\cdot\Gamma(\frac{1}{2})}\int_0^z x^{\frac{k}{2}-1}(z-x)^{\frac{1}{2}-1}\cdot e^{-\frac{z}{2}}dx$$

$$\triangleq C(k)\cdot e^{-\frac{z}{2}}\cdot\int_0^z x^{\frac{k}{2}-1}(z-x)^{\frac{1}{2}-1}dx \qquad \text{令 } x=tz$$

$$=C(k)\cdot e^{-\frac{z}{2}}\cdot\int_0^1 (tz)^{\frac{k}{2}-1}(z-tz)^{\frac{1}{2}-1}(zdt)$$

$$=C(k)\cdot z^{\frac{k-1}{2}}\cdot e^{-\frac{z}{2}}\int_0^1 t^{\frac{k}{2}-1}(1-t)^{\frac{1}{2}-1}dt$$

$$=C(k)z^{\frac{k-1}{2}}e^{-\frac{z}{2}}B\left(\frac{k}{2},\frac{1}{2}\right)$$

$$B(\frac{k}{2},\frac{1}{2})=\frac{\Gamma(\frac{k}{2})\Gamma(\frac{1}{2})}{\Gamma(\frac{k+1}{2})}$$

$$=\frac{1}{2^{\frac{k+1}{2}}\Gamma(\frac{k+1}{2})}\cdot z^{\frac{k-1}{2}}\cdot \mathrm{e}^{-\frac{z}{2}}$$

即 $f_Z(z)=\dfrac{1}{2^{\frac{k+1}{2}}\Gamma(\frac{k+1}{2})}\cdot z^{\frac{k+1}{2}-1}\cdot \mathrm{e}^{-\frac{z}{2}},\quad z>0$.

(2)性质

定理1.6 (χ^2分布的性质)

①当 $Y \sim \chi^2(n)$ 时，$E(Y) = n, D(Y) = 2n$；

② (χ^2分布的可加性) 设 $X \sim \chi^2(m), Y \sim \chi^2(n)$，且 X 与 Y 相互独立, 则 $X + Y \sim \chi^2(m+n)$.

证①：$E(Y)=n,\ D(Y)=2n$

由卡方分布定义即得

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

$$=n\{D(X_1)+E^2(X_1)\}=n$$

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

$$=n\{E(X_1^4)-E^2(X_1^2)\}=n(3-1)=2n$$

用卷积公式类似可证明②.

(3)分位数

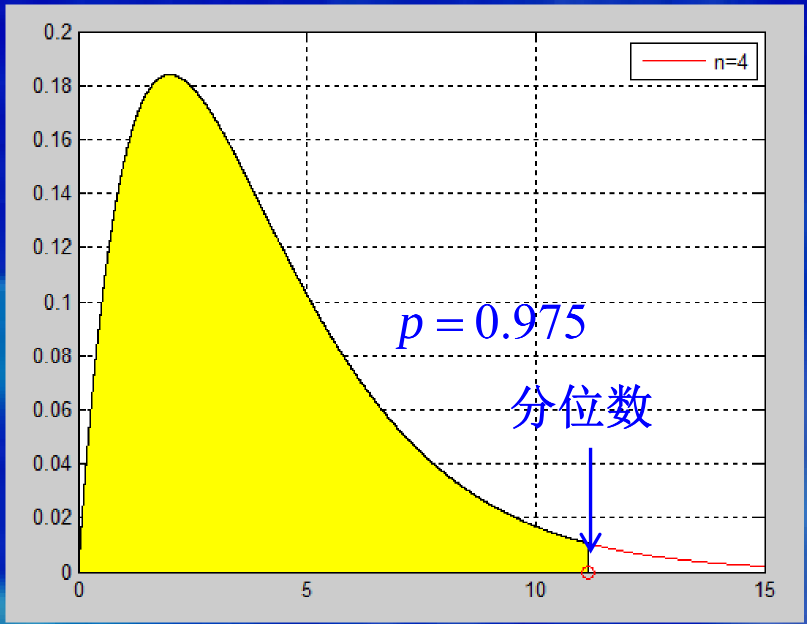
设 $X \sim \chi^2(n)$，记它的 p 分位数为 $\chi_p^2(n)$，即 $\chi_p^2(n)$ 满足

$P(X \le \chi_p^2(n)) = p$. 分位数结果可查课本**296**页附表二**.**

例 $\chi_{0.95}^2(4) = 9.488$， $\chi_{0.1}^2(14) = 7.79$.

当 $n > 45$ 时，有下列近似计算

公式 $\chi_p^2(n) \approx \dfrac{1}{2}(u_p + \sqrt{2n-1})^2$.



例17　设$(X_1,X_2,\cdots,X_n)$是来自总体 $\chi^2(n)$ 的一个样本,

样本均值 $\bar{X}=\frac{1}{n}\sum_{i=1}^{n}X_i$ ，试求 $E(\bar{X}),D(\bar{X})$.

解　由卡方分布性质知 $E(n)=n, D(X)=2n$,

再由定理1.2即得 $E(\bar{X})=n, D(\bar{X})=\frac{1}{n}\times 2n=2$.

2.T分布

(1)定义

设 $X \sim N(0,1), Y \sim \chi^2(n)$，且 X 与 Y 相互独立，称

$T \triangleq \dfrac{X}{\sqrt{Y/n}}$ 所服从的分布为自由度为 n 的 t 分布或学生氏分布，记为 $T \sim t(n)$.

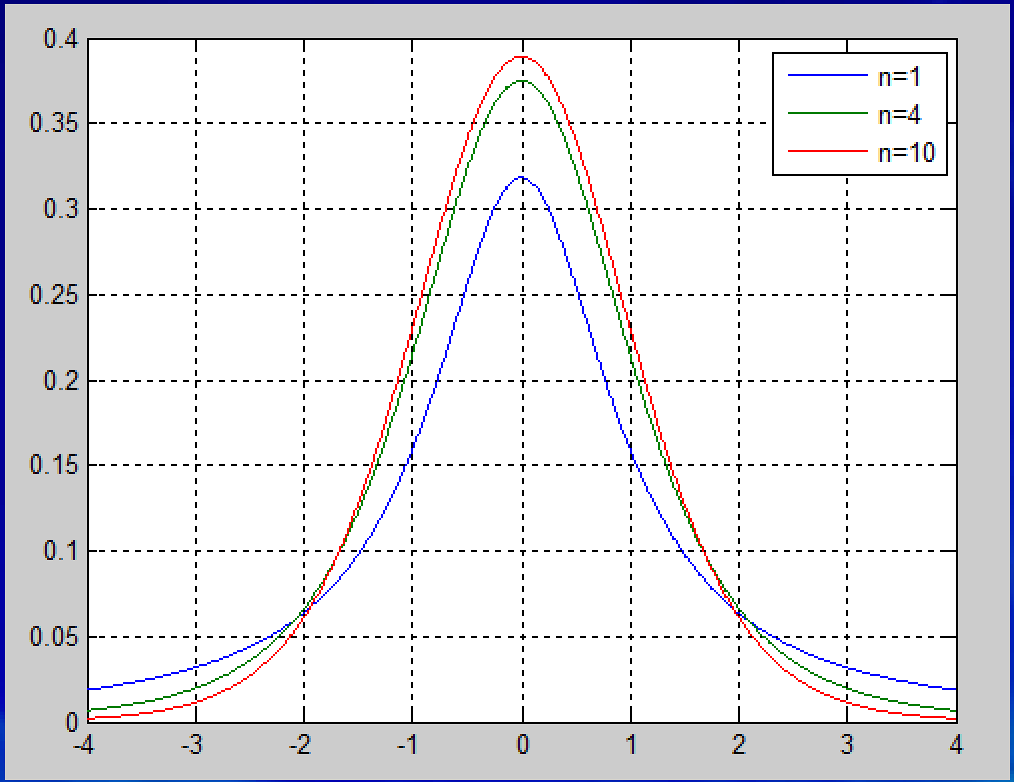
例18 设 X_1, X_2, X_3 独立同分布，都服从 $N(0,1)$，则有

$$\frac{\sqrt{2}X_1}{\sqrt{X_2^2+X_3^2}} \sim t(2).$$

证明：由定义可知 $X_1 \sim N(0,1)$，$X_2^2+X_3^2 \sim \chi^2(2)$，

且 X_1 与 $X_2^2+X_3^2$ 相互独立，故

$$\frac{X_1}{\sqrt{\dfrac{X_2^2+X_3^2}{2}}} = \frac{\sqrt{2}X_1}{\sqrt{X_2^2+X_3^2}} \sim t(2).$$



$t(n)$ 的密度函数图形

定理1.7　设随机变量 $T \sim t(n)$，则其概率密度函数为

$$f(x) = \frac{\Gamma(\frac{n+1}{2})}{\sqrt{n\pi} \cdot \Gamma(\frac{n}{2})} \cdot (1 + \frac{x^2}{n})^{-\frac{n+1}{2}}, \qquad x \in R$$

(2)分位数

设 $X \sim t(n)$，记它的 p 分位数为 $t_p(n)$，即 $t_p(n)$ 满足

$$P(X \le t_p(n)) = p.$$

由 t 分布密度函数的对称性，知其有性质 $t_p(n) = -t_{1-p}(n)$.

(该性质类似于正态分布的分位数性质)

当 $n > 45$ 时，有下列近似公式 $t_p(n) \approx u_p$.

分位数结果可查课本**299**页附表三，如 $t_{0.975}(7) = 2.3646$,

$t_{0.01}(9) = -t_{0.99}(9) = -2.8214.$

例19 设 $X_1, X_2, \cdots, X_5$ 是独立同分布的随机变量，且每个 X_i 都服从 $N(0,1)$，试给出常数 d，使得

$$d\frac{X_1+X_2}{\sqrt{X_3^2+X_4^2+X_5^2}}$$

服从 t 分布，并指出它的自由度.

解　由题意知 $X_3^2+X_4^2+X_5^2 \sim \chi^2(3), X_1+X_2 \sim N(0,2)$

且两者相互独立，由 **t-** 分布定义知

$$\frac{\dfrac{X_1+X_2}{\sqrt{2}}}{\sqrt{\dfrac{X_3^2+X_4^2+X_5^2}{3}}} \sim t(3) \Rightarrow \frac{\sqrt{3}}{\sqrt{2}} \frac{X_1+X_2}{\sqrt{X_3^2+X_4^2+X_5^2}} \sim t(3)$$

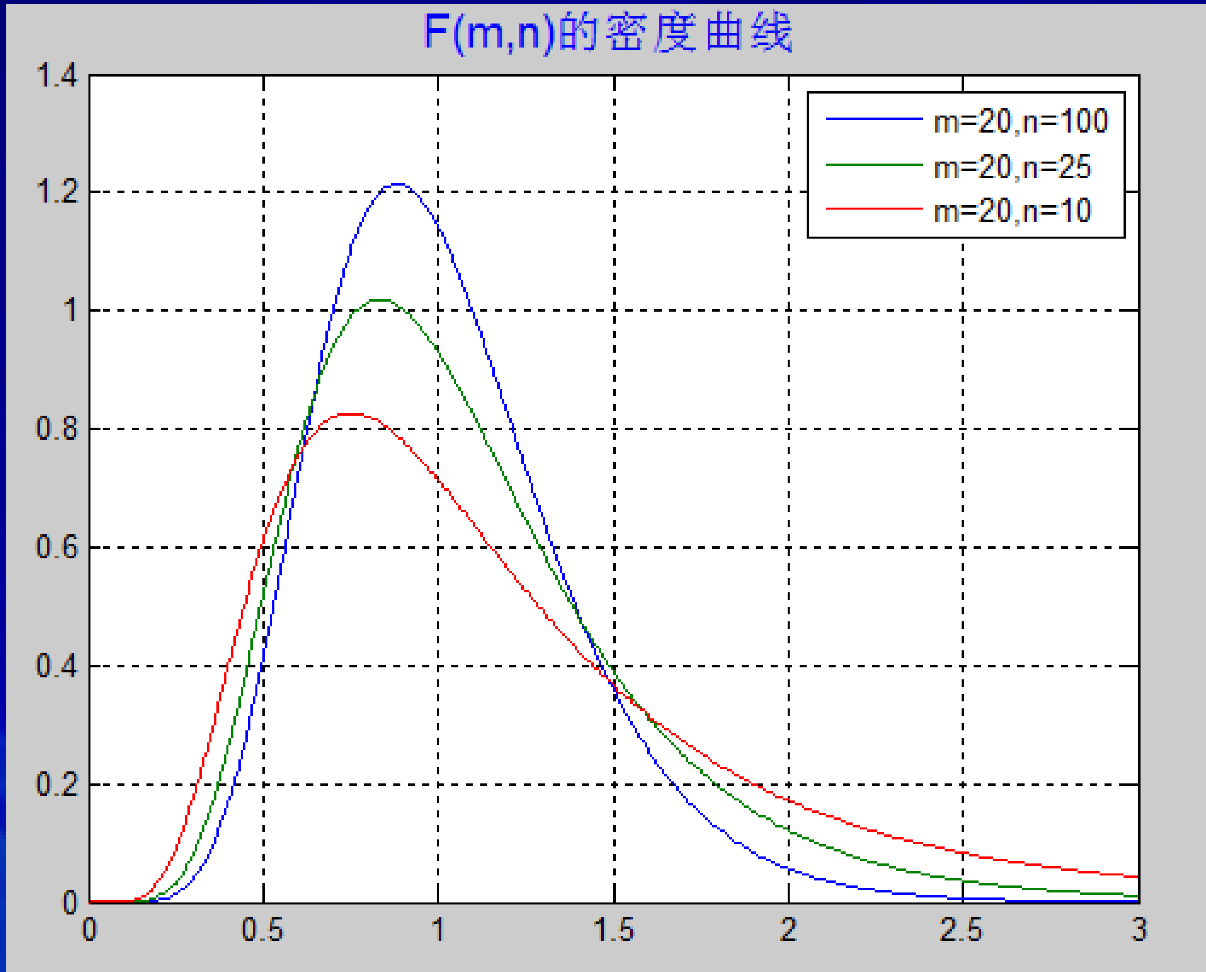
取 $d=\dfrac{\sqrt{3}}{\sqrt{2}}=\dfrac{\sqrt{6}}{2}$ 即得结论，且自由度为**3.**

3. F分布

(1)定义

设随机变量 X 与 Y 相互独立, 且 $X \sim \chi^2(m), Y \sim \chi^2(n)$,

称随机变量 $F \triangleq \dfrac{X/m}{Y/n}$ 所服从的分布为自由度为 (m,n) 的 $F-$分布, 记为 $F \sim F(m,n)$.


F(m,n)的密度曲线
m=20,n=100
m=20,n=25
m=20,n=10
1.4
1.2
1
0.8
0.6
0.4
0.2
0
0
0.5
1
1.5
2
2.5
3

定理1.8 设随机变量 $F \sim F(m,n)$，则其密度函数为

$$f(x)=\begin{cases}\dfrac{\Gamma(\frac{m+n}{2})}{\Gamma(\frac{m}{2})\cdot\Gamma(\frac{n}{2})}\cdot\left(\dfrac{m}{n}\right)^{\frac{m}{2}}\cdot x^{\frac{m}{2}-1}\cdot\left(1+\dfrac{m}{n}x\right)^{-\frac{m+n}{2}}, & x>0,\\ 0, & \text{其余}.\end{cases}$$

(2)分位数

设 $X \sim F(m,n)$，记它的 p 分位数为 $F_p(m,n)$，即 $F_p(m,n)$ 满足 $P(X \le F_p(m,n)) = p$.

分位数满足下列性质：$F_p(m,n) = \dfrac{1}{F_{1-p}(n,m)}$.

证明:　设 $X \sim F(m,n)$，则 $\dfrac{1}{X} \sim F(n,m)$，而

$$P(X \le F_p(m,n)) = p \Leftrightarrow P\left(\frac{1}{X} \ge \frac{1}{F_p(m,n)}\right) = p$$

$$\Leftrightarrow P\left(\frac{1}{X} \le \frac{1}{F_p(m,n)}\right) = 1 - p \Rightarrow \frac{1}{F_p(m,n)} = F_{1-p}(n,m)$$

此即 $F_p(m,n) = \dfrac{1}{F_{1-p}(n,m)}$.

分位数结果可查课本**300**页附表四．例 $F_{0.95}(4,10)=3.48,$

$$F_{0.05}(12,3)=\frac{1}{F_{0.95}(3,12)}=\frac{1}{3.49}.$$

例20　设 $X_1, X_2, \cdots, X_n, X_{n+1}, \cdots, X_{n+m}$ 是来自正态总体 $N(0,\sigma^2)$ 且容量为 $m+n$ 的样本，试求下列统计量的抽样分布.

(1) $Y = \dfrac{1}{\sigma^2}\sum_{i=1}^{m+n} X_i^2$;

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

例20　设 $X_1, X_2, \cdots, X_n, X_{n+1}, \cdots, X_{n+m}$ 是来自正态总体 $N(0,\sigma^2)$ 且容量为 $m+n$ 的样本，求下列统计量的抽样分布.

(2) $Z = \dfrac{\sqrt{m}\sum\limits_{i=1}^{n} X_i}{\sqrt{n}\sqrt{\sum\limits_{i=n+1}^{m+n} X_i^2}}$；

解　由题意知 $\sum\limits_{i=1}^{n} X_i \sim N(0, n\sigma^2)$，

$\sum\limits_{i=n+1}^{n+m} \left(\dfrac{X_i}{\sigma}\right)^2 \sim \chi^2(m)$，且两者独立，故

$$\frac{\sum\limits_{i=1}^{n} X_i / \sqrt{n}\sigma}{\sqrt{\dfrac{\sum\limits_{i=n+1}^{n+m}\left(\dfrac{X_i}{\sigma}\right)^2}{m}}} \sim t(m) \Rightarrow \frac{\sqrt{m}}{\sqrt{n}} \frac{\sum\limits_{i=1}^{n} X_i}{\sqrt{\sum\limits_{i=n+1}^{n+m} X_i^2}} \sim t(m)$$

例20 设 $X_1, X_2, \cdots, X_n, X_{n+1}, \cdots, X_{n+m}$ 是来自正态总体 $N(0, \sigma^2)$ 且容量为 $m+n$ 的样本，求下列统计量的抽样分布.

(3) $F = \dfrac{m\sum_{i=1}^{n} X_i^2}{n\sum_{i=n+1}^{m+n} X_i^2}$.

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

$\sum_{i=n+1}^{n+m}\left(\dfrac{X_i}{\sigma}\right)^2 \sim \chi^2(m)$，且相互独立，故

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

例21　设 (X_1, X_2, X_3, X_4) 是取自总体 $N(0,\sigma^2)$ 的一个简单随机样本，试确定常数 c，使得

$$P\left(\frac{(X_1+X_2)^2}{(X_1+X_2)^2+(X_3-X_4)^2}>c\right)=0.9.$$

解　由 $X_1+X_2 \sim N(0,2\sigma^2)$ 可知，$\dfrac{X_1+X_2}{\sqrt{2}\sigma} \sim N(0,1)$，

故 $\left(\dfrac{X_1+X_2}{\sqrt{2}\sigma}\right)^2 \sim \chi^2(1)$；　同理 $\left(\dfrac{X_3-X_4}{\sqrt{2}\sigma}\right)^2 \sim \chi^2(1)$.

又 X_1+X_2 与 X_3-X_4 相互独立，故

$$\left(\frac{X_3-X_4}{\sqrt{2}\sigma}\right)^2 \Big/ \left(\frac{X_1+X_2}{\sqrt{2}\sigma}\right)^2 \sim F(1,1)\text{，即 }\frac{(X_3-X_4)^2}{(X_1+X_2)^2} \sim F(1,1);$$

$$P\left(\frac{(X_1+X_2)^2}{(X_1+X_2)^2+(X_3-X_4)^2}>c\right)=0.9,\ \frac{(X_3-X_4)^2}{(X_1+X_2)^2}\sim F(1,1)$$

所以 $0.9=P\left(\frac{(X_1+X_2)^2}{(X_1+X_2)^2+(X_3-X_4)^2}>c\right)$

$$=P\left(\frac{(X_1+X_2)^2+(X_3-X_4)^2}{(X_1+X_2)^2}<\frac{1}{c}\right)$$

$$=P\left(1+\frac{(X_3-X_4)^2}{(X_1+X_2)^2}<\frac{1}{c}\right)=P\left(\frac{(X_3-X_4)^2}{(X_1+X_2)^2}\le\frac{1}{c}-1\right)=0.9$$

即 $\frac{1}{c}-1=F_{0.9}(1,1)=39.86\Rightarrow c=0.02447.$