概率论与数理统计

范正平
fanzhp@mail.sysu.edu.cn

中山大学 数据科学与计算机学院

Chapter 6
样本及抽样分布

数理统计的基本概念

数理统计的分类

描述统计学——

对随机现象进行观测、试验，以取得有代表性的观测值

推断统计学——

对已取得的观测值进行整理、分析，作出推断、决策，从而找出所研究的对象的规律性

推断统计学

- 推断统计学
  - 参数估计
  - 假设检验
  - 方差分析
  - 回归分析

基本概念

一个统计问题总有它明确的研究对象.

- 总体 研究对象的全体

研究某批灯泡的质量

该批灯泡寿命的全体就是总体

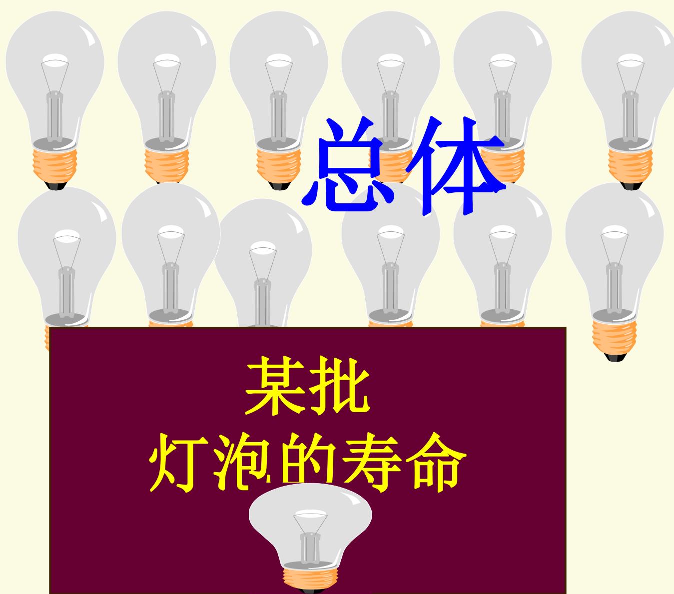
考察国产 轿车的油耗

所有国产轿车每公里耗油量的全体就是总体

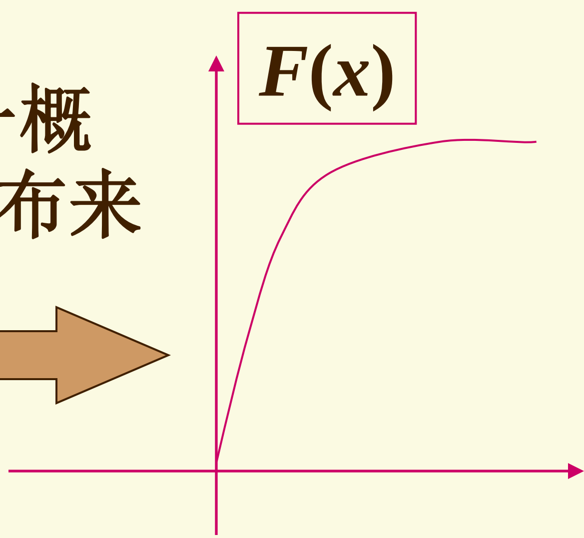
基本概念(Cont.)

- 总体可以用一个随机变量 X 或其分布来描述

如:研究某批灯泡的寿命时，我们关心的就是寿命，那么，寿命这个总体就可以用随机变量X表示，或用其分布函数$F(x)$表示.



寿命可用一概率(指数)分布来刻划



基本概念(Cont.)

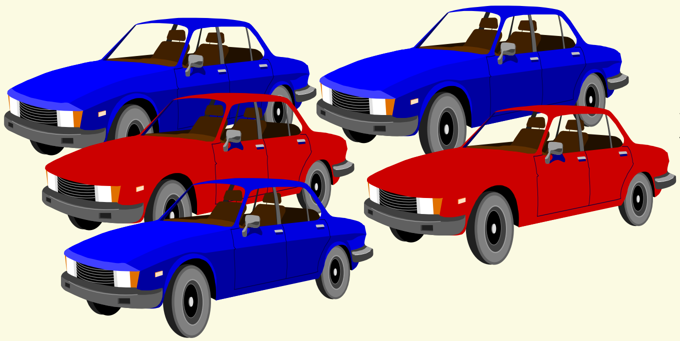
- 个体

总体中每个对象称为个体.

- 样本

为推断总体分布及各种特征,按一定规则从总体中抽取若干个体进行观察试验以获得有关总体的信息. 所抽取的部分个体称为样本. 样本中所包含的个体数目称为样本容量.

基本概念(Cont.)



从国产轿车中抽5
辆进行耗油量试验

样本容量为 5

抽到哪 5 辆是随机的！

样本是随机变量

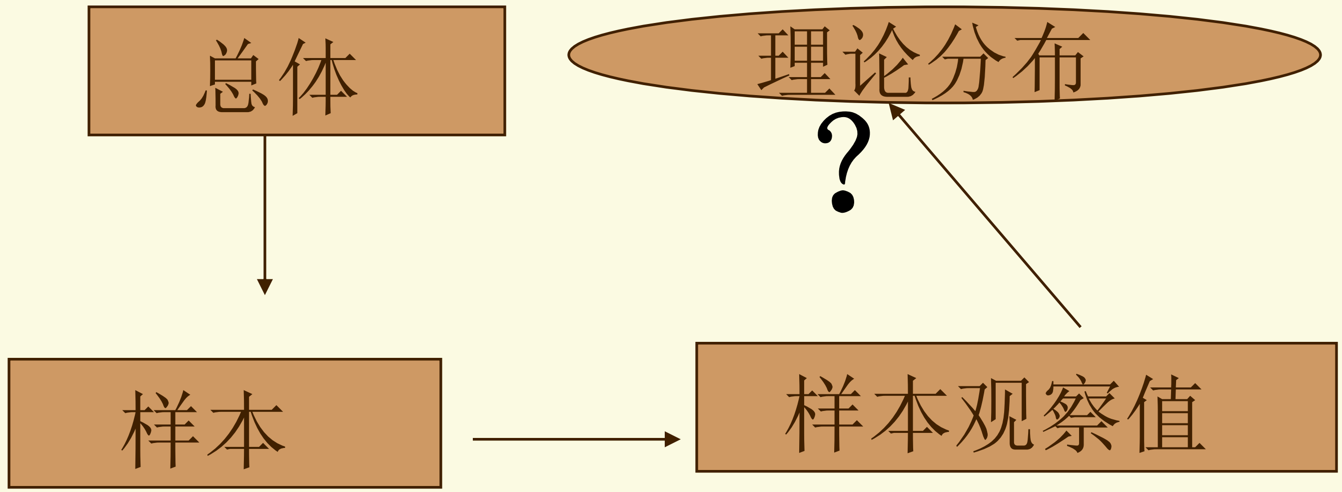
容量为 n 的样本可以看作n 维随机变量$(X_1, X_2, \cdots, X_n)$.

一旦取定一组样本，得到的是 n 个具体的数 $x_1, x_2, \cdots, x_n$，称为样本$(X1, X2, \ldots, Xn)$的一组观测值，简称样本值.

总体、样本、样本值的关系

事实上我们抽样后得到的资料都是具体的、确定的值. 如我们从某班大学生中抽取10人测量身高,得到10个数，它们是样本取到的值而不是样本. 我们只能观察到随机变量取的值而见不到随机变量.

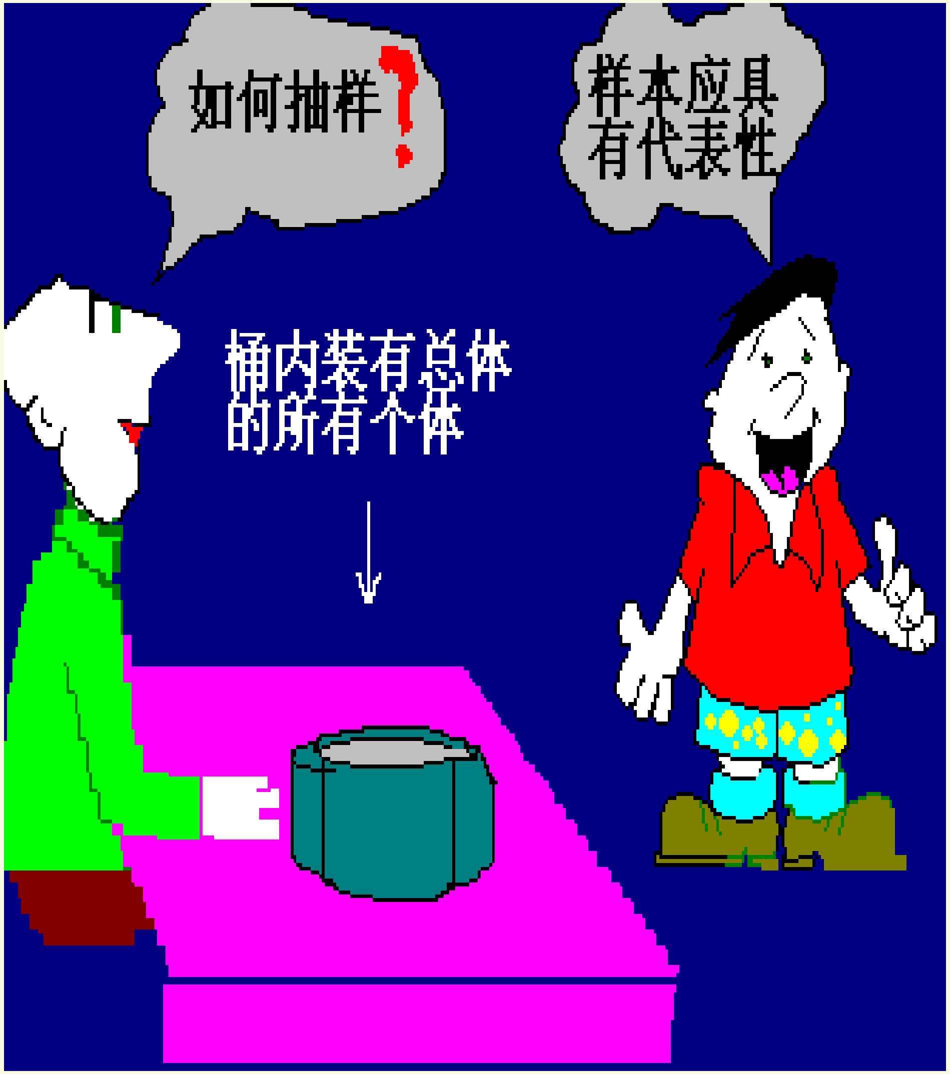
总体、样本、样本观察值的关系



样本空间 —— 样本所有可能取值的集合

简单随机样本

抽取样本的目的是为了利用样本对总体进行统计推断，这就要求样本能很好的反映总体的特性且便于处理。因此：



简单随机样本(Cont.)

抽取的样本$X_1, X_2, \cdots, X_n$ 满足下面两点:

1.独立性： $X_1, X_2, \cdots, X_n$ 是相互独立的随机变量；

2.代表性： X_i（$i=1,2,\cdots,n$）与所考察的总体 X 具有相同的分布.

简单随机样本是应用中最常见的情形,今后，说到"$X_1, \cdots, X_n$是来自某总体的样本"时，若不特别说明，就指简单随机样本.

总体与样本

总体：随机变量X，或分布函数$F(x)$

样本：n 维随机变量$(X1, X2, \ldots, Xn)$.

$x_1, x_2, \cdots, x_n$，为样本值.

- 若总体 X 的分布函数为$F(x)$，则其样本的联合分布函数为

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

- 若总体 X 的概率密度为 $p(x)$,则其样本的联合概率密度为

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

直方图与箱线图

一、直方图

实际统计工作中首先涉及到的是一系列数据，数据的变异性系统地表现了数据的分布。常常在工厂、学校等部门表现为图表。下面讨论直方图。

例1：观察新生女婴儿的体重X，统计数据如下：

2880,　2440,　2700,　3500,　3500

3600,　3080,　3860,　3200，　3100

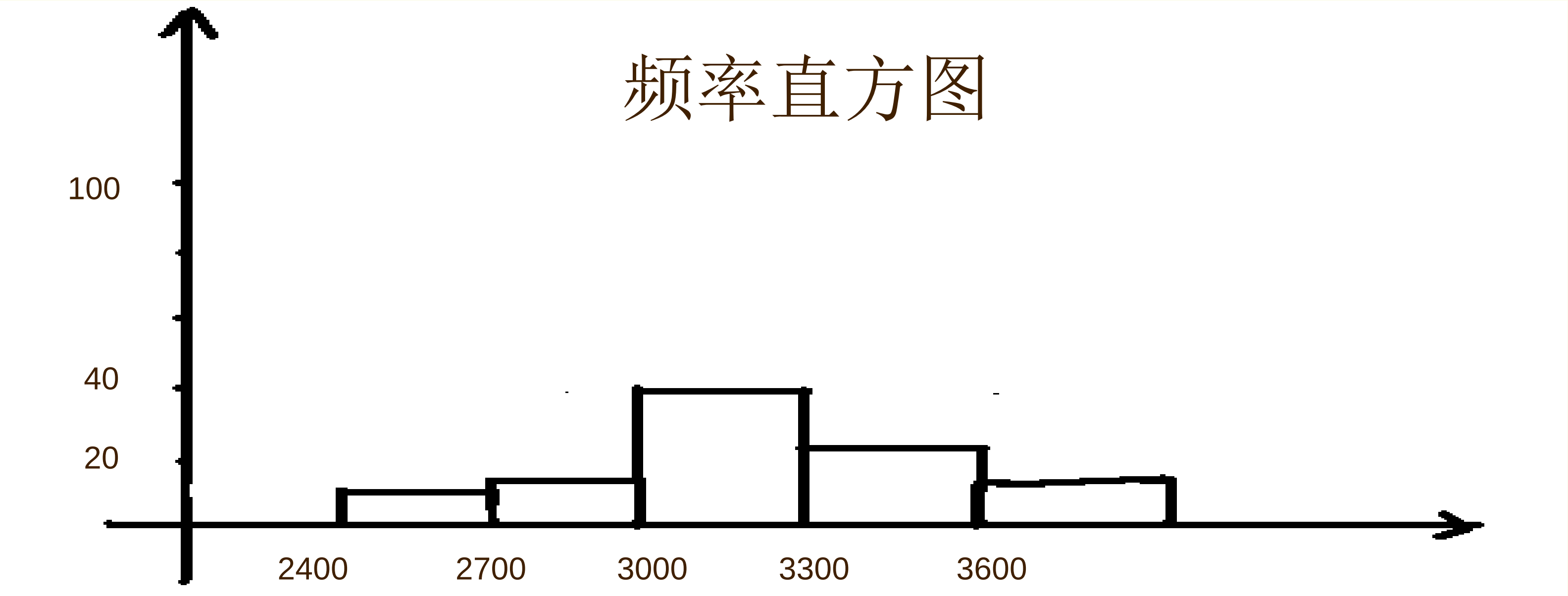
3180, 3200, 3300, 3020, 3040

3420, 2900, 3440, 3000, 2620

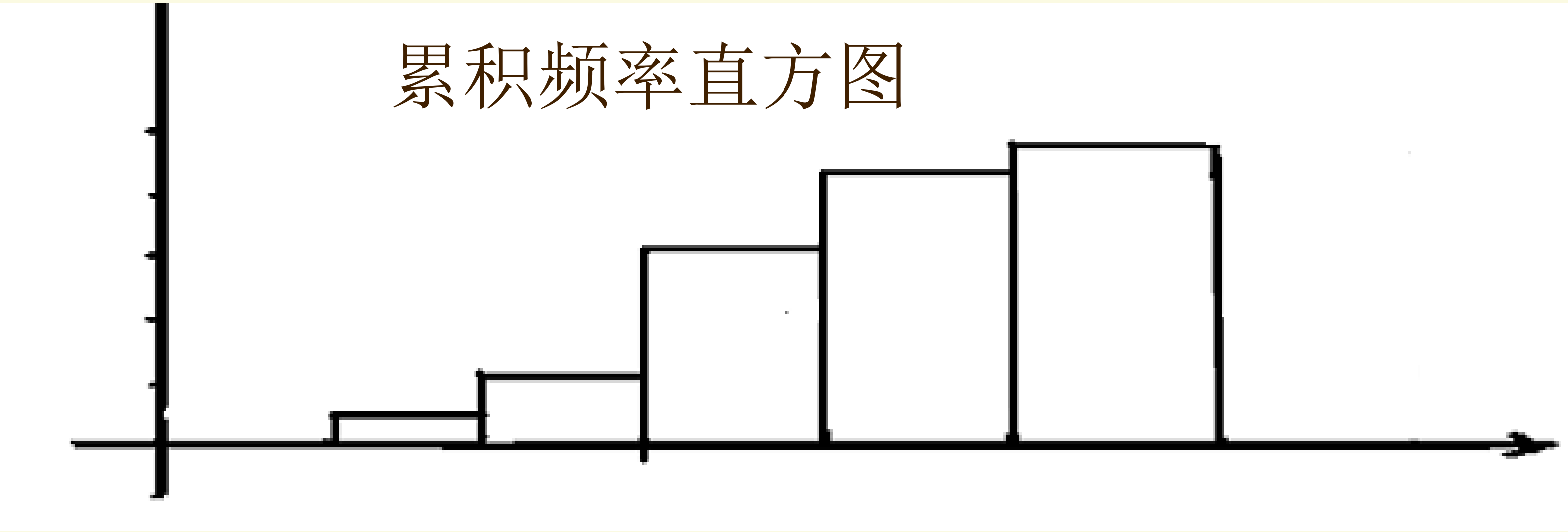
考察数据的分布？

解　共20个数据，对数据进行分组, 列表, 统计.

分组	1	2	3	4	5
组限	[2400，2700）	[2700，3000）	[3000，3300）	[3300，3600）	[3600，3900）
组中值	2550	2850	3150	3450	3750
组频数	2	3	8	5	2
组频率	0.10	0.15	0.40	0.25	0.10
累积频率	0.10	0.25	0.65	0.90	1

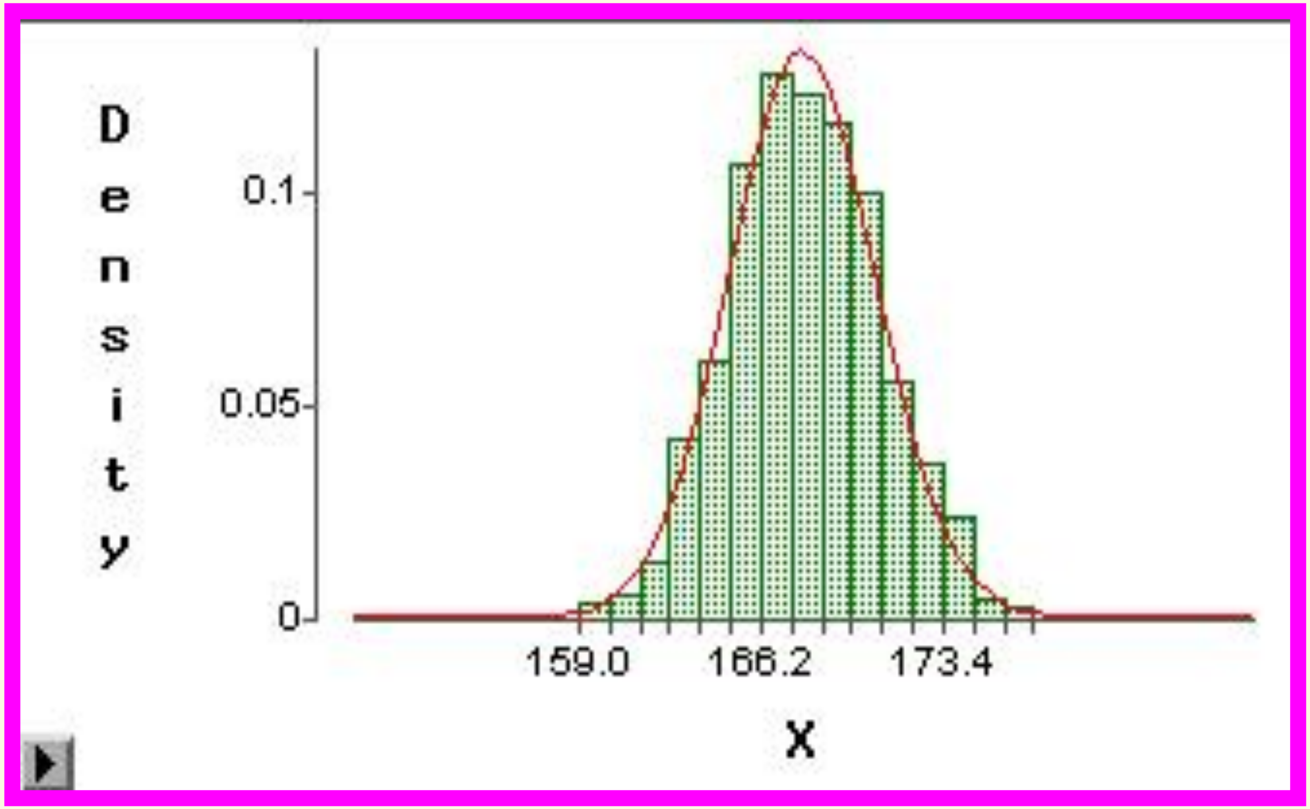


频率直方图
100
40
20
2400
2700
3000
3300
3600



累积频率直方图

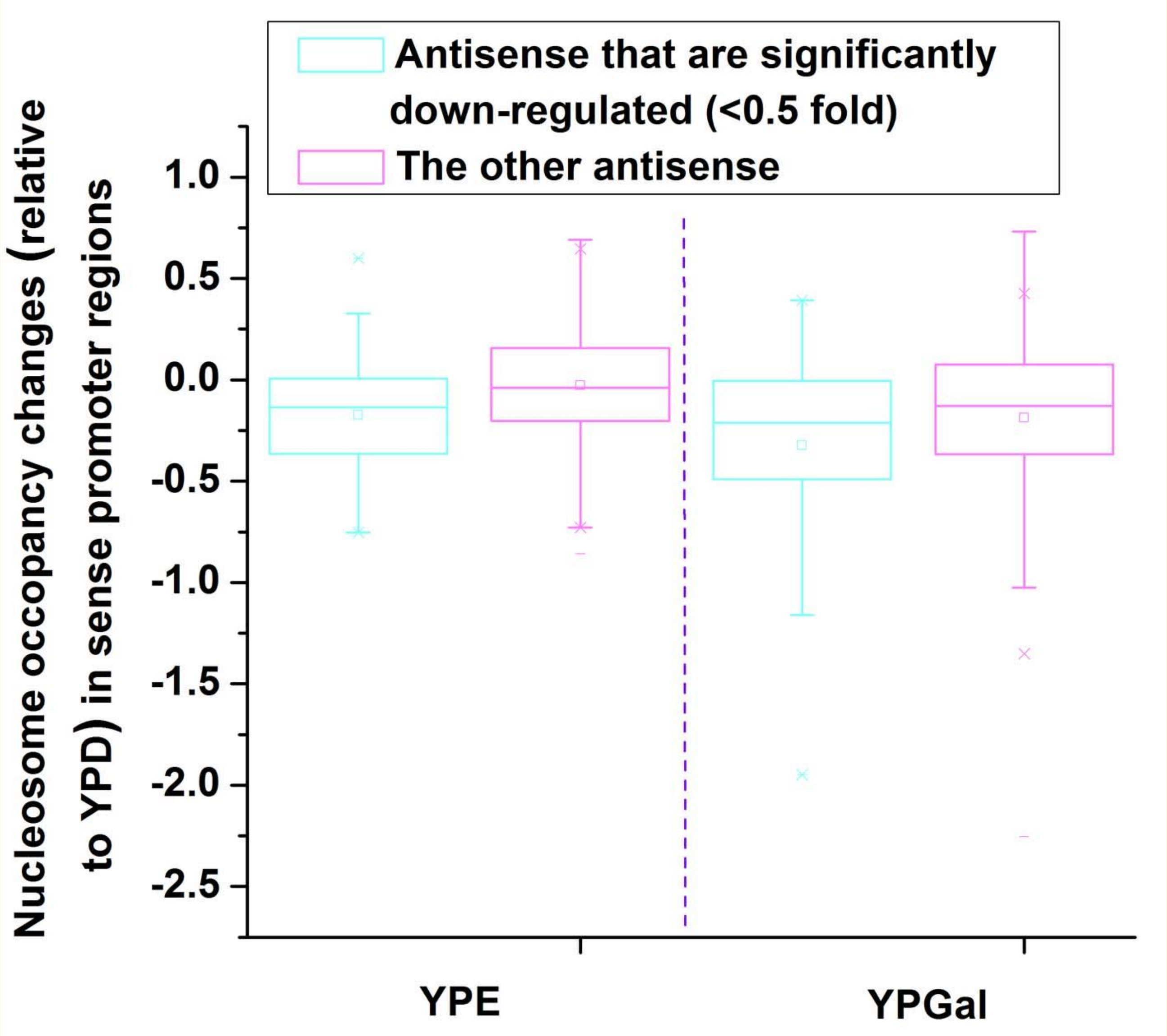
下面是我们用某大学大学生的身高的数据画出的频率直方图。



红线是拟合的正态密度曲线

箱线图

最大值、
最小值、
中位数、
四分位数
(第一、第三）



统计量

由样本推断总体特征, 需要对样本值进行 "加工", "提炼". 这就需要构造一些样本的函数, 它把样本中所含的信息集中起来.

统计量的定义

设 $X_1, X_2, \cdots, X_n$ 是来自总体 X 的一个样本, $g(X_1, X_2, \cdots, X_n)$ 是 $X_1, X_2, \cdots, X_n$ 的函数, 若 g 中不含未知参数, 则称 $g(X_1, X_2, \cdots, X_n)$ 是一个统计量.

设 $x_1, x_2, \cdots, x_n$ 是相应于样本 $X_1, X_2, \cdots, X_n$ 的样本值, 则称 $g(x_1, x_2, \cdots, x_n)$ 是 $g(X_1, X_2, \cdots, X_n)$ 的观察值.

统计量(Cont.)

实例 设X_1, X_2, X_3是来自总体$N(\mu, \sigma^2)$的一个样本, 其中μ为已知, σ^2为未知, 判断下列各式哪些是统计量, 哪些不是?

$$T_1 = X_1, \qquad T_2 = X_1 + X_2 e^{X_3},$$

$$T_3 = \frac{1}{3}(X_1 + X_2 + X_3),$$

$$T_4 = \max(X_1, X_2, X_3), \qquad T_5 = X_1 + X_2 - 2\mu,$$

是

$$T_6 = \frac{1}{\sigma^2}(X_1^2 + X_2^2 + X_3^2).$$

不是

常见统计量

1. 样本均值 $\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$

2. 样本方差

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

3. 样本k阶原点矩

$$A_k = \frac{1}{n}\sum_{i=1}^{n} X_i^k$$

4. 样本k阶中心矩

$$B_k = \frac{1}{n}\sum_{i=1}^{n} (X_i - \overline{X})^k$$

$X_1, X_2, \cdots, X_n$为来自X的样本 样本数字特征(随机变量)：	$x_1, x_2, \cdots, x_n$为样本值 样本数字特征观察值：
1. 样本均值 $\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$	$\overline{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$
2. 样本方差 $S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{X})^2$	$s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \overline{x})^2$
3. 样本k阶原点矩 $A_k = \frac{1}{n}\sum_{i=1}^{n} X_i^k$	$a_k = \frac{1}{n}\sum_{i=1}^{n} x_i^k$
4. 样本k阶中心矩 $B_k = \frac{1}{n}\sum_{i=1}^{n}(X_i - \overline{X})^k$	$b_k = \frac{1}{n}\sum_{i=1}^{n}(x_i - \overline{x})^k$

例1 从一批钢筋中随机抽取10条，测得其直径（单位：mm）为： 24.2, 25.4, 24, 24, 25, 25, 24.4, 24.6, 25.2, 25.2.

(1)写出总体、样本、样本值、样本容量；

(2)求样本观测值的均值、方差及二阶原点矩(保留二位).

解 (1)总体为该批钢筋的直径；样本为$X_1, X_2, \dots, X_{10}$

样本值：24.2, 25.4, 24, 24, 25, 25, 24.4, 24.6, 25.2, 25.2.

样本容量：n=10;

(2)样本均值 $$\frac{1}{n}\sum_{i=1}^{n} x_i = \frac{1}{10}(24.2+25.2+\dots+25.2) = 24.68mm$$

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

$$= \frac{1}{9}\begin{bmatrix}(-0.48)^2+(0.72)^2+(-0.68)^2+(-0.68)^2+(0.32)^2 \\ +(0.32)^2+(-0.28)^2+(-0.08)^2+(-0.52)^2+(0.52)^2\end{bmatrix} \approx 0.278$$

二阶原点矩 $$A_2 = \frac{1}{n}\sum_{i=1}^{n} x_i^{\,2} = \frac{1}{10}\left[24.8^2+25.4^2+\cdots+25.2^2\right] = 610.34.$$

常用统计量的分布

统计学上的三大分布：

χ^2分布、t分布和F分布。

一、正态分布

定理1. 若 $X_1, X_2, \cdots, X_n$ 相互独立,

$X_i \sim N(\mu_i, \sigma_i^2)$ ，则

$$\sum_{i=1}^{n} a_i X_i \sim N\left(\sum_{i=1}^{n} a_i \mu_i, \sum_{i=1}^{n} a_i^2 \sigma_i^2\right)$$

特别地, 若 $X_1, X_2, \cdots, X_n \overset{i.i.d}{\sim} N(\mu, \sigma^2)$

则 $\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i \sim N\left(\mu, \frac{\sigma^2}{n}\right)$

$$U = \frac{\sqrt{n}\left(\overline{X} - \mu\right)}{\sigma} \sim N(0,1)$$

在已知总体μ，σ^2时，
可用本定理计算样
本均值$\overline{X}$.

标准正态分布的 α 分位数

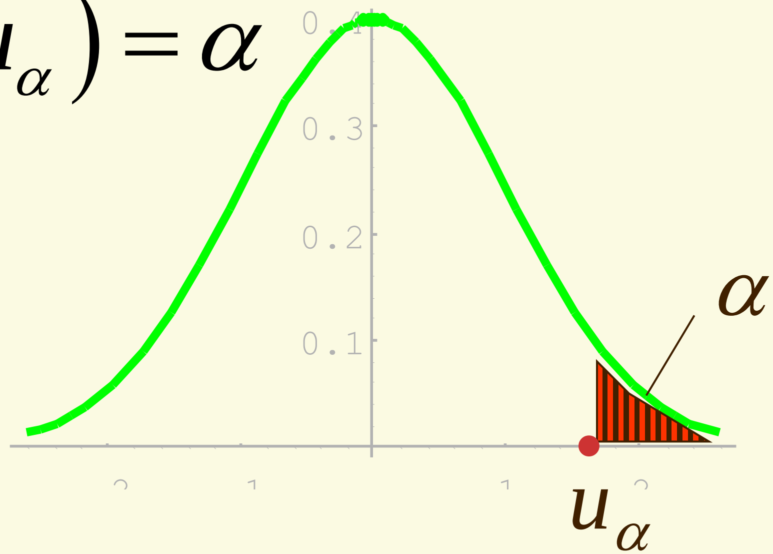
定义

若 $P(X > u_\alpha) = \alpha$ 则称u_α为标准正态分布的上α 分位数.

若 $P\left(|X| > u_{\frac{\alpha}{2}}\right) = \alpha$ 则称 $u_{\frac{\alpha}{2}}$ 为标准正态分布的双侧α分位数.

标准正态分布的α 分位数图形

$$P\left(X > u_\alpha\right) = \alpha$$

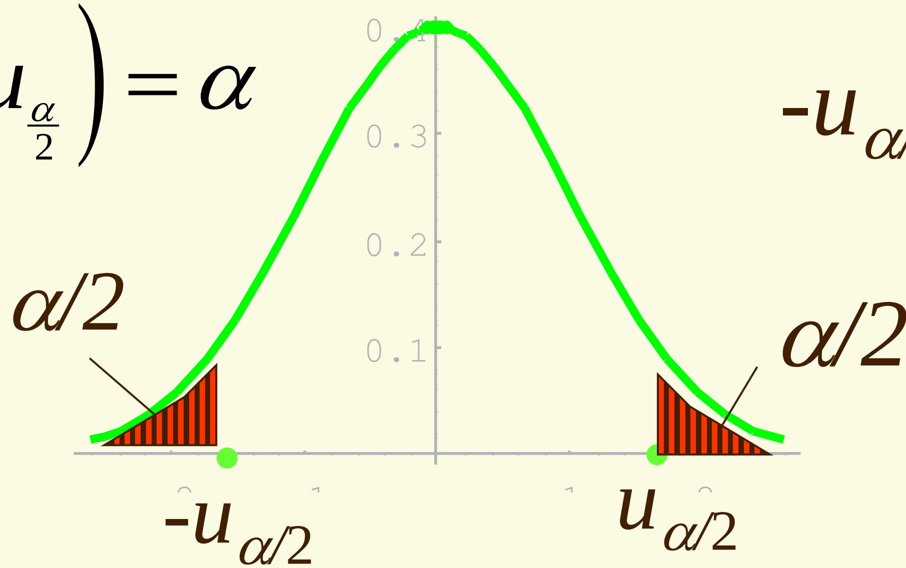


$$\left.\begin{aligned} u_{0.05} &= 1.645 \\ u_{0.025} &= 1.96 \end{aligned}\right\} \boxed{\text{常用数字}}$$

$$u_{0.005} = 2.575$$

$$P\left(|\,X\,| > u_{\frac{\alpha}{2}}\right) = \alpha$$

$$-u_{\alpha/2} = u_{1-\alpha/2}$$



二、 $\chi^2(n)$ 分布（ n为自由度 ）（卡方分布）

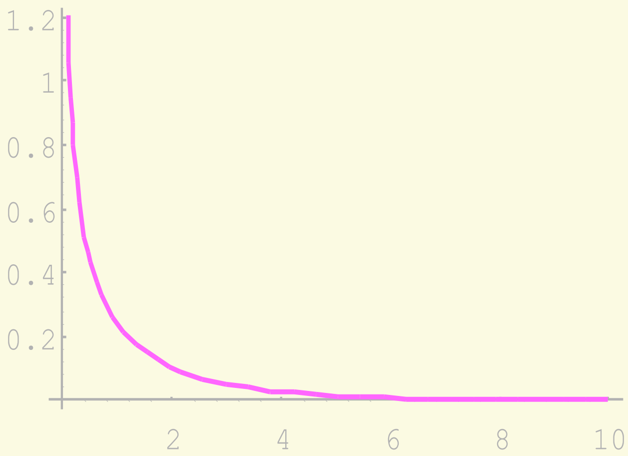
定义 设 $X_1, X_2, \cdots, X_n$相互独立，

且都服从标准正态分布$N(0, 1)$, 则

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

注1 $n = 1$ 时, 其密度函数为

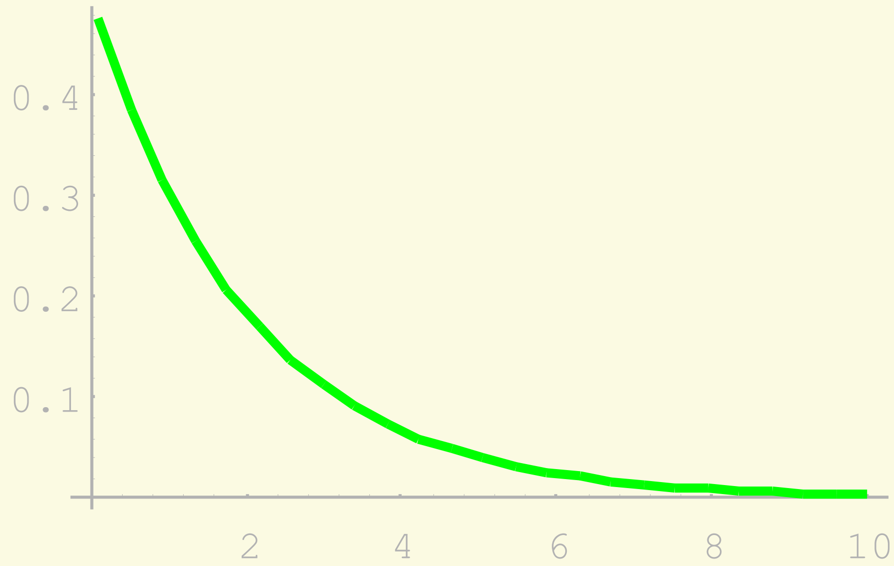
$$f(x) = \begin{cases} \dfrac{1}{\sqrt{2\pi}} x^{-\frac{1}{2}} e^{-\frac{x}{2}}, & x > 0 \\ 0, & x \le 0 \end{cases}$$



注2 $n=2$ 时, 其密度函数为

$$f(x)=\begin{cases}\dfrac{1}{2}e^{-\frac{x}{2}}, & x>0\\ 0, & x\le 0\end{cases}$$

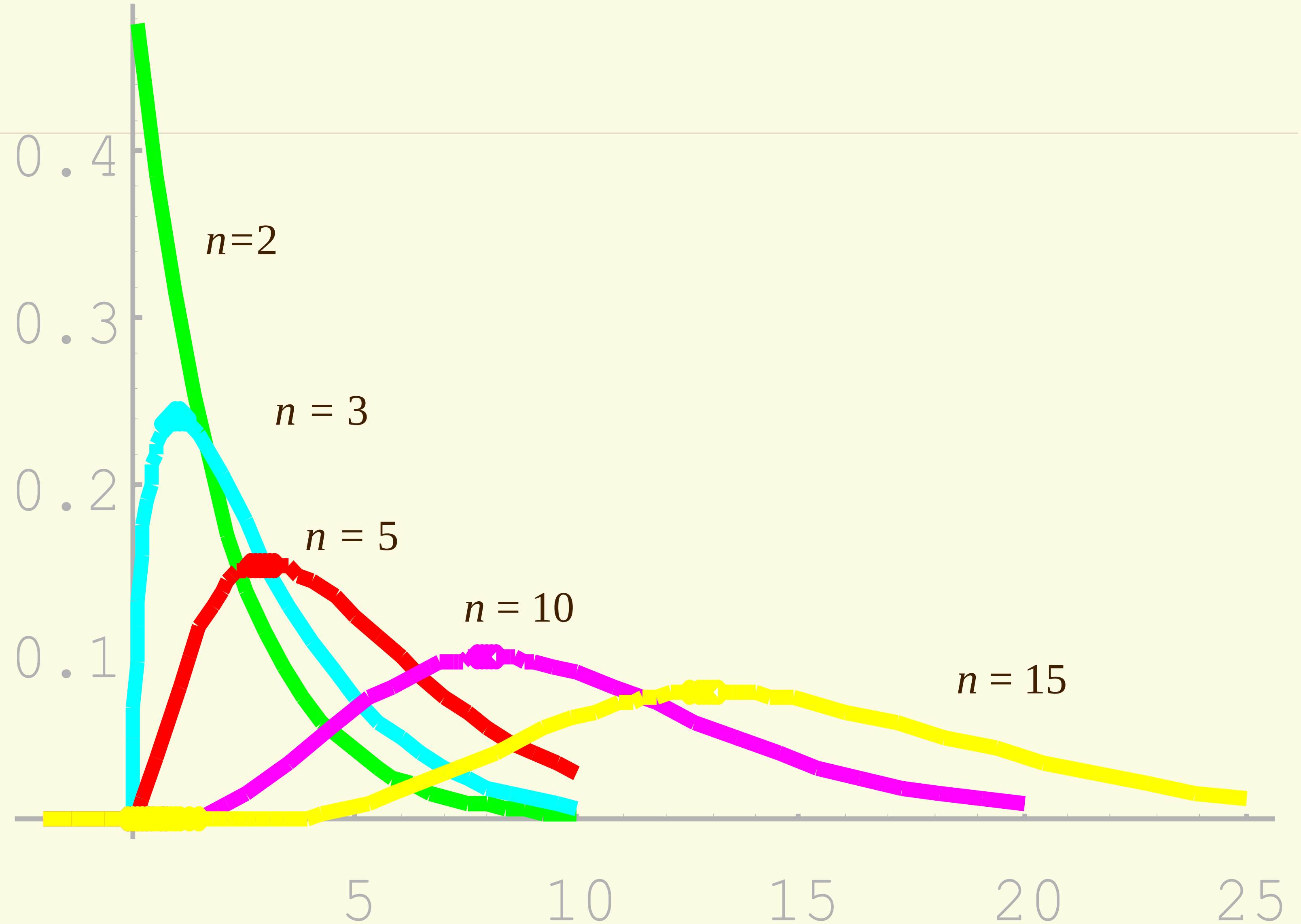
为参数为1/2的指数分布.



一般 自由度为 n 的 $\chi^2(n)$ 的密度函数为

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

其中 $\Gamma(\alpha)=\int_0^{+\infty}t^{\alpha-1}e^{-t}dt$



0.4
0.3
0.2
0.1
n=2
n = 3
n = 5
n = 10
n = 15
5
10
15
20
25

$\chi^2(n)$ 分布的性质

1° $E\left(\chi^2(n)\right)=n,\ D\left(\chi^2(n)\right)=2n$

2° 若$X_1=\chi^2(n_1), X_2=\chi^2(n_2), X_1, X_2$相互独立，

则 $X_1+X_2\sim\chi^2(n_1+n_2)$

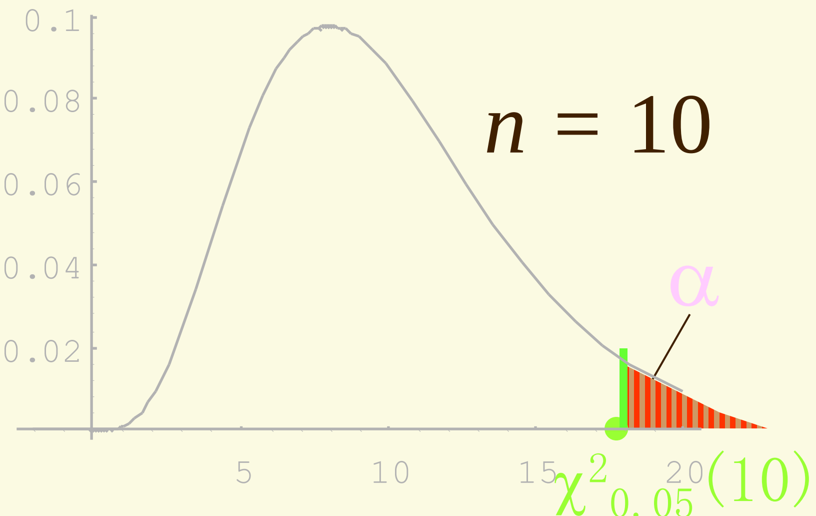
3° $n\to\infty$时，$\chi^2(n)\to$正态分布

4° $\chi^2(n)$分布的上α分位数有表可查

例如

$\chi^2_{0.05}(10)=18.307$

$P\left(\chi^2(10)>18.307\right)=0.05$



证 1°设 $\chi^2(n)=\sum_{i=1}^{n} X_i^2 \quad X_i \sim N(0,1)\ \ i=1,2,\cdots,n$

$X_1, X_2, \cdots, X_n$ 相互独立,

则 $E(X_i)=0,\ D(X_i)=1,\ E(X_i^2)=1$

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

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

$$D(X_i^2)=E(X_i^4)-E^2(X_i^2)=2$$

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

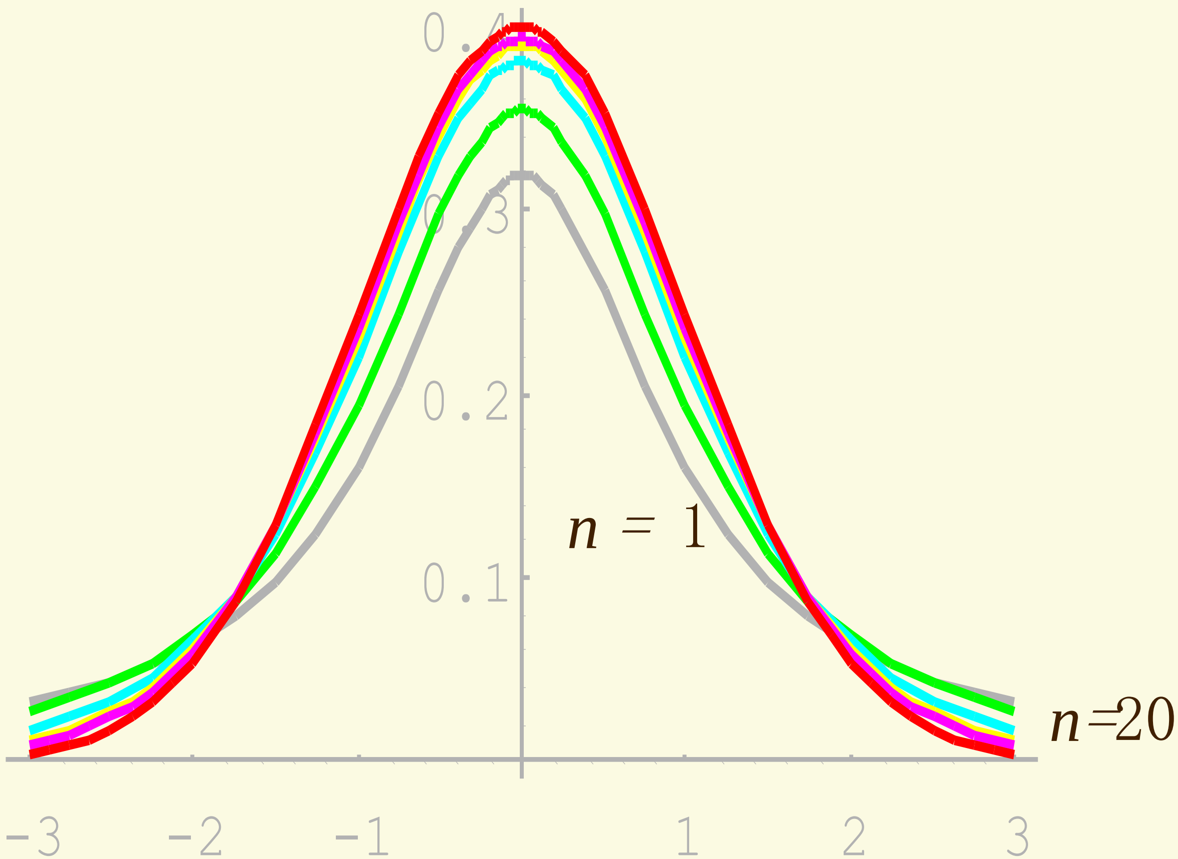
三、 t 分布（Student 分布）

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

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

则称T服从自由度为 n 的T分布, 记为$T \sim t(n)$.
其密度函数为

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



t 分布的图形(红色的是标准正态分布)

t分布的性质：

1. 具有自由度为n的t分布$t \sim t(n)$,其数学期望与方差为：$E(t)=0, D(t)=n/(n-2) \quad (n>2)$

2. t分布的密度函数关于$t=0$对称.当n充分大时，其图形近似于标准正态 分布概率密度的图形，

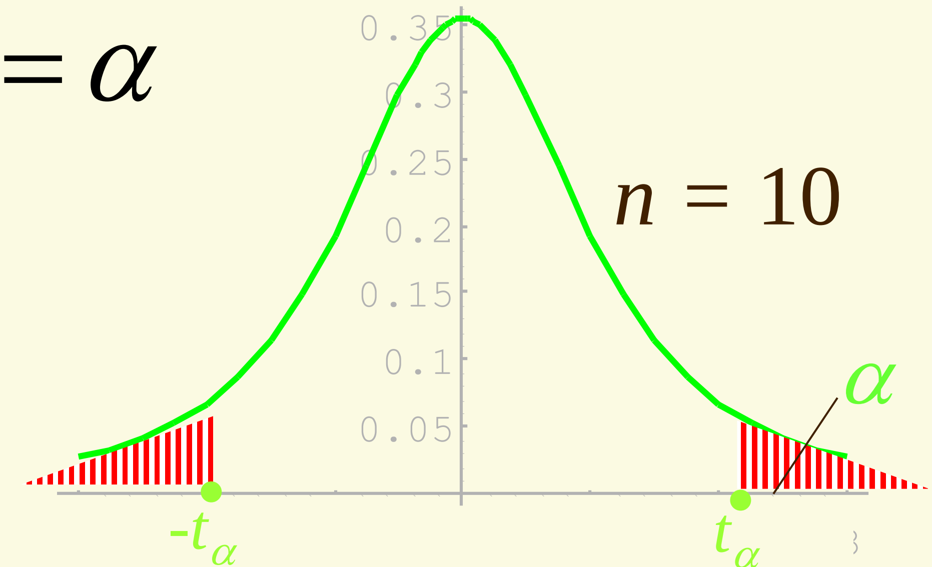
再由Γ函数的性质有 $\lim\limits_{n\to\infty} h(t)=\dfrac{1}{\sqrt{2\pi}}e^{-t^2/2}$.

即当n足够大时，$t \overset{\text{近似}}{\sim} N(0,1)$.

3 T 分布的上 α 分位数 t_{α} 与双测 α 分位数 $t_{\alpha/2}$ 均有表可查.

$$P(T > t_\alpha) = \alpha$$

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



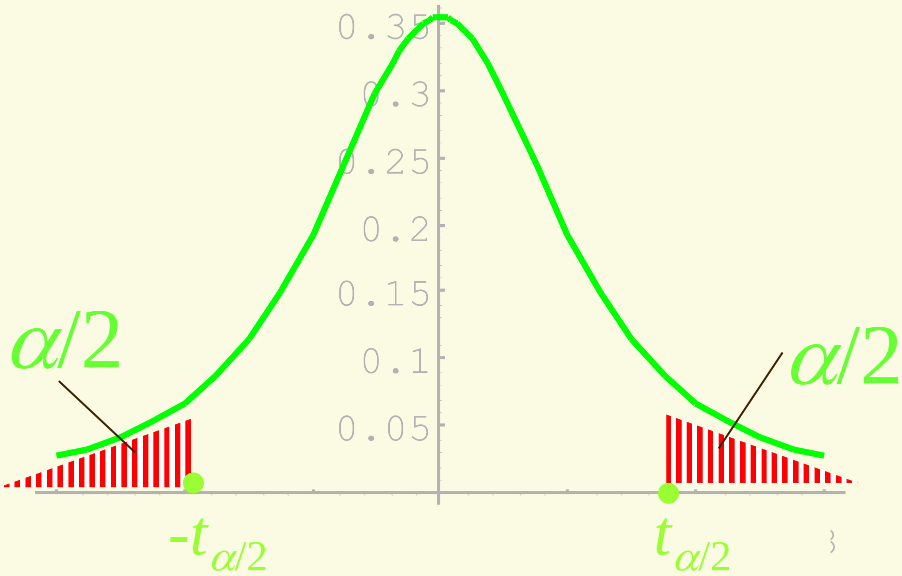
$$P(T > 1.8125) = 0.05 \Rightarrow t_{0.05}(10) = 1.8125$$

$$P(T < -1.8125) = 0.05, \quad P(T > -1.8125) = 0.95$$

$$\Rightarrow t_{0.95}(10) = -1.8125$$

$$P(T > t_{\alpha/2}) = \frac{\alpha}{2}$$

$$P\left(|T| > t_{\alpha/2}\right) = \alpha$$



$$P\left(T > 2.2281\right) = 0.025$$

$$P\left(|T| > 2.2281\right) = 0.05$$

$$\Rightarrow t_{0.025}(10) = 2.2281$$

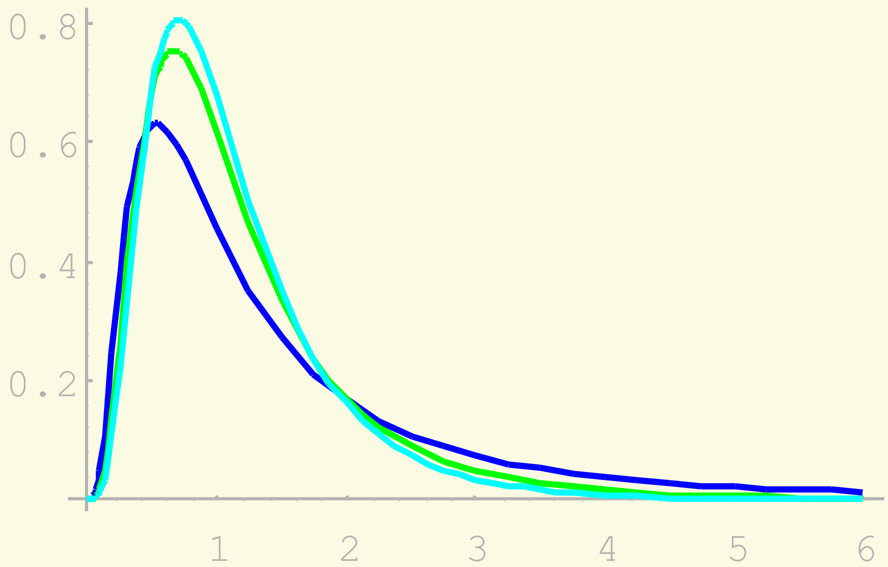
四、 F 分布

定义 设 $X \sim \chi^2(n), Y \sim \chi^2(m), X, Y$ 相互独立，

令 $$F = \frac{X/n}{Y/m}$$

则称 F 服从为第一自由度为n，第二自由度为 m 的F 分布. 其密度函数为

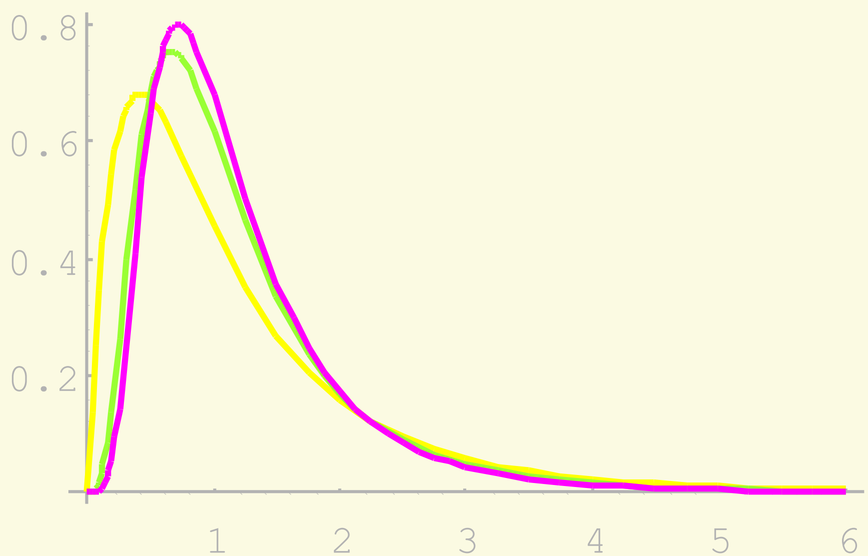
$$f(t,n,m) = \begin{cases} \dfrac{\Gamma\left(\dfrac{n+m}{2}\right)}{\Gamma\left(\dfrac{n}{2}\right)\Gamma\left(\dfrac{m}{2}\right)}\left(\dfrac{n}{m}\right)^{\frac{n}{2}} t^{\frac{n}{2}-1}\left(1+\dfrac{n}{m}t\right)^{-\frac{n+m}{2}} & t > 0 \\ 0, & t \le 0 \end{cases}$$



m = 10, n = 4

m = 10, n = 10

m = 10, n = 15



m = 4, n =10

m = 10, n = 10

m = 15, n = 10

F 分布的性质

1° 若$F \sim F(n,m)$，则 $1/F \sim F(m,n)$

2° $F(n,m)$的上α分位数$F_\alpha(n,m)$有表可查：

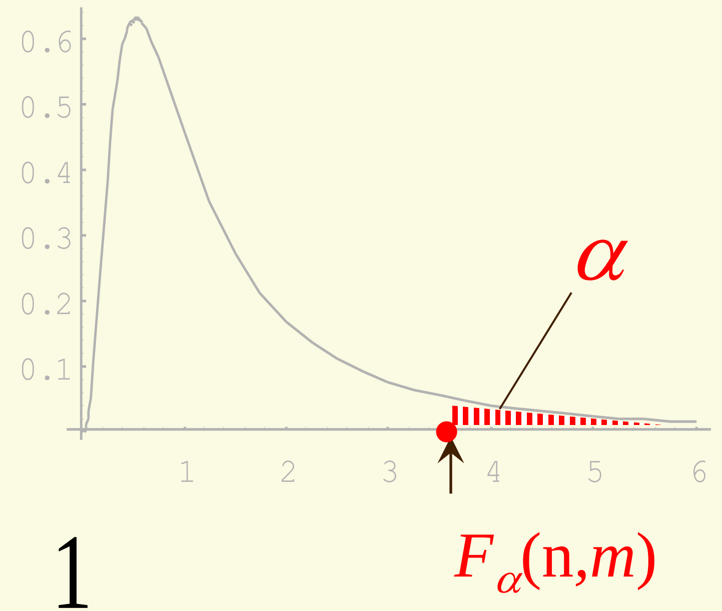
$$P(F > F_\alpha(n,m)) = \alpha$$

例如 $F_{0.05}(4,5) = 5.19$

求 $F_{0.95}(5,4) = ?$

事实上，$F_{1-\alpha}(n,m) = \dfrac{1}{F_\alpha(m,n)}$

故 $F_{0.95}(5,4) = \dfrac{1}{F_{0.05}(4,5)} = \dfrac{1}{5.19}$



F分布的性质

1.F分布的数学期望为:

$$E(F)=\frac{n_2}{n_2-2}$$ 若$n_2>2$

即它的数学期望并不依赖于第一自由度n_1.

当总体为正态分布时，给出几个重要的抽样分布定理.

定理 1 (样本均值的分布)

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

$$\overline{X} \sim N(\mu, \frac{\sigma^2}{n})$$

即 $$\frac{\overline{X}-\mu}{\sigma/\sqrt{n}} \sim N(0,1)$$

定理 2 (样本方差的分布)

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

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

(2) $\overline{X}$与S^2独立.

样本方差

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

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

定理 3 (样本均值的分布)

设$X_1,X_2,\ldots,X_n$是取自正态总体 $N(\mu,\sigma^2)$

的样本， $\overline{X}$和S^2 分别为样本均值和样本方差，

则有

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

证明: $\because \quad \overline{X} \sim N(\mu,\frac{\sigma^2}{n})$, $\quad \therefore \quad \frac{\overline{X}-\mu}{\sigma/\sqrt{n}} \sim N(0,1)$

又 $\frac{(n-1)}{\sigma^2}S^2 \sim \chi^2(n-1)$ 且与 $\frac{\overline{X}-\mu}{\sigma/\sqrt{n}}$ 相互独立，

故

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

定理 4 两个正态总体

设 $X_1, X_2, \cdots, X_n$ 是来自正态总体 $X \sim N(\mu_1, \sigma_1^2)$ 的一个简单随机样本

$Y_1, Y_2, \cdots, Y_m$ 是来自正态总体 $Y \sim N(\mu_2, \sigma_2^2)$ 的一个简单随机样本

它们相互独立.

令 $\overline{X} = \dfrac{1}{n}\sum_{i=1}^{n} X_i$ $\qquad \overline{Y} = \dfrac{1}{m}\sum_{j=1}^{m} Y_j$

$$S_1^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{X})^2 \qquad S_2^2 = \frac{1}{m-1}\sum_{j=1}^{m}(Y_j - \overline{Y})^2$$

则 **(i)**

$$\frac{S_1^2 / \sigma_1^2}{S_2^2 / \sigma_2^2} \sim F(n-1, m-1)$$

若 $\sigma_1 = \sigma_2$ 则 $$\frac{S_1^2}{S_2^2} \sim F(n-1, m-1)$$

(ii) 有

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

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

(iii) 有

$$\frac{(n-1)S_1^2}{\sigma^2} \sim \chi^2(n-1), \quad \frac{(m-1)S_2^2}{\sigma^2} \sim \chi^2(m-1)$$

$$\longrightarrow \quad \frac{(n-1)S_1^2}{\sigma^2} + \frac{(m-1)S_2^2}{\sigma^2} \sim \chi^2(n+m-2)$$

$\overline{X}-\overline{Y}$ 与 $\dfrac{(n-1)S_1^2}{\sigma^2} + \dfrac{(m-1)S_2^2}{\sigma^2}$ 相互独立

$$\frac{\dfrac{(\overline{X}-\overline{Y})-(\mu_1-\mu_2)}{\sqrt{\dfrac{\sigma^2}{n}+\dfrac{\sigma^2}{m}}}}{\sqrt{\dfrac{\dfrac{(n-1)S_1^2}{\sigma^2}+\dfrac{(m-1)S_2^2}{\sigma^2}}{n+m-2}}}$$

$$=\boxed{\frac{(\overline{X}-\overline{Y})-(\mu_1-\mu_2)}{\sqrt{\dfrac{1}{n}+\dfrac{1}{m}}\sqrt{\dfrac{(n-1)S_1^2+(m-1)S_2^2}{n+m-2}}}\sim t(n+m-2)}$$

例： $X \sim N(72, 100)$，为使样本均值大于70的概率不小于90%，则样本容量至少取多少？

解　设样本容量为 n，则　$\overline{X} \sim N(72, \frac{100}{n})$

故 $P(\overline{X} > 70) = 1 - P(\overline{X} \le 70)$

$$= 1 - \Phi\left(\frac{70 - 72}{\frac{10}{\sqrt{n}}}\right) = \Phi(0.2\sqrt{n})$$

令　$\Phi(0.2\sqrt{n}) \ge 0.9$　　得　$0.2\sqrt{n} \ge 1.29$

即　$n \ge 41.6025$　　所以取 $n = 42$

例 设 $(X_1, X_2, \cdots, X_{10})$ 是来自总体 $N(0, 0.3^2)$ 的样本, 求 $P\left(\sum_{i=1}^{10} X_i^2 > 1.44\right)$.

解:将 $X_i, i=1,2,\cdots,10$ 标准化得 $\frac{X_i - 0}{0.3} \sim N(0,1), i=1,2,\cdots,10$.

由 χ^2 分布的构造知, $\sum_{i=1}^{10}\left(\frac{X_i}{0.3}\right)^2 \sim \chi^2(10)$. 因此有

$$P\left(\sum_{i=1}^{10} X_i^2 > 1.44\right) = P\left\{\sum_{i=1}^{10}\left(\frac{X_i}{0.3}\right)^2 > \frac{1.44}{0.09}\right\}$$

$$= P\left\{\sum_{i=1}^{10}\left(\frac{X_i}{0.3}\right)^2 > 16\right\} = 0.1$$

例 设 $X_1, X_2, \cdots, X_{10}$ 是来自总体 $X \sim N(\mu, 4)$ 的样本, 求样本方差 S^2 大于 2. 622 的概率.

解:由于 $\dfrac{(10-1)S^2}{4} \sim \chi^2(9)$,

$$P(S^2 > 2.622) = P\left(\frac{9}{4}S^2 > 5.8995\right)$$

查表得 $x_{0.75}^2(9) = 5.899$,因此, $P(S^2 > 2.622) \approx 0.75$.