

第一节 随机样本

一、总体与个体

二、随机样本的定义

三、小结



一、总体与个体

1. 总体

试验的全部可能的观察值称为总体。

2. 个体 总体中的每个可能观察值称为个体。

实例1 在研究2000名学生的年龄时, 这些学生的年龄的全体就构成一个总体, 每个学生的年龄就是个体。



3. 有限总体和无限总体

实例2 某工厂10月份生产的灯泡寿命所组成的总体中, 个体的总数就是10月份生产的灯泡数, 这是一个有限总体; 而该工厂生产的所有灯泡寿命所组成的总体是一个无限总体, 它包括以往生产和今后生产的灯泡寿命.

当有限总体包含的个体的总数很大时, 可近似地把它看成是无限总体.



4. 总体分布

实例3 在2000名大学一年级学生的年龄中, 年龄指标值为“15”, “16”, “17”, “18”, “19”, “20”的依次有9, 21, 132, 1207, 588, 43 名, 它们在总体中所占比率依次为

$$\frac{9}{2000}, \frac{21}{2000}, \frac{132}{2000}, \frac{1207}{2000}, \frac{588}{2000}, \frac{43}{2000},$$

即学生年龄的取值有一定的分布.



一般地, 我们所研究的总体, 即研究对象的某项数量指标 X , 其取值在客观上有一定的分布, X 是一个随机变量.

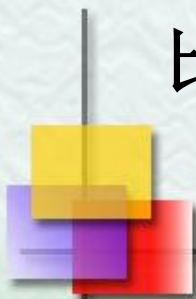
总体分布的定义

我们把数量指标取不同数值的比率叫做总体分布.

如实例3中, 总体就是数集 $\{15, 16, 17, 18, 19, 20\}$.

总体分布为

年龄	15	16	17	18	19	20
比率	$\frac{9}{2000}$	$\frac{21}{2000}$	$\frac{132}{2000}$	$\frac{1207}{2000}$	$\frac{588}{2000}$	$\frac{43}{2000}$



二、随机样本的定义

1. 样本的定义

设 X 是具有分布函数 F 的随机变量, 若 $X_1, X_2, \dots, X_n$ 是具有同一分布函数 F 、相互独立的随机变量, 则称 $X_1, X_2, \dots, X_n$ 为从分布函数 F (或总体 F 、或总体 X) 得到的容量为 n 的简单随机样本, 简称样本.

它们的观察值 $x_1, x_2, \dots, x_n$ 称为样本值, 又称为 X 的 n 个独立的观察值.



2. 简单随机抽样的定义

获得简单随机样本的抽样方法称为简单随机抽样。

根据定义得：若 $X_1, X_2, \dots, X_n$ 为 F 的一个样本，

则 $X_1, X_2, \dots, X_n$ 的联合分布函数为

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

又若 X 具有概率密度 f ，

则 $X_1, X_2, \dots, X_n$ 的联合概率密度为

$$f^*(x_1, x_2, \dots, x_n) = \prod_{i=1}^n f(x_i).$$



例4 设总体 X 服从参数为 λ ($\lambda > 0$) 的指数分布, $(X_1, X_2, \dots, X_n)$ 是来自总体的样本, 求样本 $(X_1, X_2, \dots, X_n)$ 的概率密度.

解 总体 X 的概率密度为 $f(x) = \begin{cases} \lambda e^{-\lambda x}, & x > 0, \\ 0, & x \leq 0, \end{cases}$

因为 $X_1, X_2, \dots, X_n$ 相互独立, 且与 X 有相同的分布, 所以 $(X_1, X_2, \dots, X_n)$ 的概率密度为

$$f_n(x_1, x_2, \dots, x_n) = \prod_{i=1}^n f(x_i) = \begin{cases} \lambda^n e^{-\lambda \sum_{i=1}^n x_i}, & x_i > 0, \\ 0, & \text{其他.} \end{cases}$$



例5 设总体 X 服从两点分布 $B(1, p)$, 其中 $0 < p < 1$, $(X_1, X_2, \dots, X_n)$ 是来自总体的样本, 求样本 $(X_1, X_2, \dots, X_n)$ 的分布律.

解 总体 X 的分布律为

$$P\{X = i\} = p^i (1 - p)^{1-i} \quad (i = 0, 1)$$

因为 $X_1, X_2, \dots, X_n$ 相互独立,

且与 X 有相同的分布,

所以 $(X_1, X_2, \dots, X_n)$ 的分布律为



$$\begin{aligned} & P\{X_1 = x_1, X_2 = x_2, \cdots, X_n = x_n\} \\ &= P\{X_1 = x_1\}P\{X_2 = x_2\} \cdots P\{X_n = x_n\} \\ &= p^{\sum_{i=1}^n x_i} (1-p)^{n-\sum_{i=1}^n x_i} \end{aligned}$$

其中 $x_1, x_2, \cdots, x_n$ 在集合 $\{0,1\}$ 中取值.



三、小结

基本概念：个体 总体 $\left\{ \begin{array}{l} \text{有限总体} \\ \text{无限总体} \end{array} \right.$ 随机样本

说明1 一个总体对应一个随机变量 X , 我们将不区分总体和相应的随机变量, 统称为总体 X .

说明2 在实际中遇到的总体往往是有限总体, 它对应一个离散型随机变量; 当总体中包含的个体的个数很大时, 在理论上可认为它是一个无限总体.



第二节 抽样分布

一、基本概念

二、常见分布

三、小结



一、基本概念

1. 统计量的定义

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

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



实例1 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?

$$T_1 = X_1,$$

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

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

是

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

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

不是

2. 几个常用统计量的定义

设 $X_1, X_2, \dots, X_n$ 是来自总体的一个样本，
 $x_1, x_2, \dots, x_n$ 是这一样本的观察值。

(1) 样本平均值 $\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i;$

其观察值 $\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i.$

(2) 样本方差

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



其观察值

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

(3) 样本标准差

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

其观察值

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



(4) 样本 k 阶(原点)矩 $A_k = \frac{1}{n} \sum_{i=1}^n X_i^k, k = 1, 2, \dots;$

其观察值 $\alpha_k = \frac{1}{n} \sum_{i=1}^n x_i^k, k = 1, 2, \dots.$

(5) 样本 k 阶中心矩

$$B_k = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^k, k = 2, 3, \dots;$$

其观察值 $b_k = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^k, k = 2, 3, \dots.$



由以上定义得下述**结论**:

若总体 X 的 k 阶矩 $E(X^k)$ 记成 μ_k 存在,
则当 $n \rightarrow \infty$ 时, $A_k \xrightarrow{P} \mu_k, k = 1, 2, \dots$.
证明 因为 $X_1, X_2, \dots, X_n$ 独立且与 X 同分布,
所以 $X_1^k, X_2^k, \dots, X_n^k$ 独立且与 X^k 同分布,
故有 $E(X_1^k) = E(X_2^k) = \dots = E(X_n^k) = \mu_k$.

再根据第五章**辛钦定理**知

辛钦定理



$$\frac{1}{n} \sum_{i=1}^n X_i^k \xrightarrow{P} \mu_k, \quad k = 1, 2, \dots;$$

由第五章关于依概率收敛的序列的性质知

$$g(A_1, A_2, \dots, A_k) \xrightarrow{P} g(\mu_1, \mu_2, \dots, \mu_k),$$

其中 g 是连续函数.

以上结论是下一章所要介绍的矩估计法的理论根据.



3. 经验分布函数

总体分布函数 $F(x)$ 相应的统计量称为经验分布函数.

经验分布函数的做法如下:

设 $X_1, X_2, \dots, X_n$ 是总体 F 的一个样本,

用 $S(x) (-\infty < x < +\infty)$ 表示 $X_1, X_2, \dots, X_n$ 中不大于 x 的随机变量的个数,

定义经验分布函数 $F_n(x)$ 为

$$F_n(x) = \frac{1}{n} S(x), \quad (-\infty < x < +\infty)$$



对于一个样本值, $F_n(x)$ 的观察值容易求得.
($F_n(x)$ 的观察值仍以 $F_n(x)$ 表示.)

实例2 设总体 F 具有一个样本值 1, 2, 3,

则经验分布函数
 $F_3(x)$ 的观察值为

$$F_3(x) = \begin{cases} 0, & x < 1, \\ \frac{1}{3}, & 1 \leq x < 2, \\ \frac{2}{3}, & 2 \leq x < 3, \\ 1, & x \geq 3. \end{cases}$$



实例3 设总体 F 具有一个样本值 $1, 1, 2$,

则经验分布函数 $F_3(x)$ 的观察值为

$$F_3(x) = \begin{cases} 0, & x < 1, \\ \frac{2}{3}, & 1 \leq x < 2, \\ 1, & x \geq 2. \end{cases}$$



一般地,

设 $x_1, x_2, \dots, x_n$ 是总体 F 的一个容量为 n 样本值,

先将 $x_1, x_2, \dots, x_n$ 按自小到大的次序排列,

并重新编号, $x_{(1)} \leq x_{(2)} \leq \dots \leq x_{(n)},$

则经验分布函数 $F_n(x)$ 的观察值为

$$F_n(x) = \begin{cases} 0, & x < x_{(1)}, \\ \frac{k}{n}, & x_{(k)} \leq x < x_{(k+1)}, \\ 1, & x \geq x_{(n)}. \end{cases}$$



格里汶科定理

格里汶科

对于任一实数 x , 当 $n \rightarrow \infty$ 时, $F_n(x)$ 以概率 1 一致收敛于分布函数 $F(x)$, 即

$$P\left\{\lim_{n \rightarrow \infty} \sup_{-\infty < x < +\infty} |F_n(x) - F(x)| = 0\right\} = 1.$$

对于任一实数 x 当 n 充分大时, 经验分布函数的任一个观察值 $F_n(x)$ 与总体分布函数 $F(x)$ 只有微小的差别, 从而在实际上可当作 $F(x)$ 来使用.



二、常见分布

统计量的分布称为抽样分布.

1. χ^2 分布

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

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

随机数演示

分布函数与密度函数演示



$\chi^2(n)$ 分布的概率密度为

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

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

又因为 $X_i \sim N(0, 1)$, 由定义 $X_i^2 \sim \chi^2(1)$,

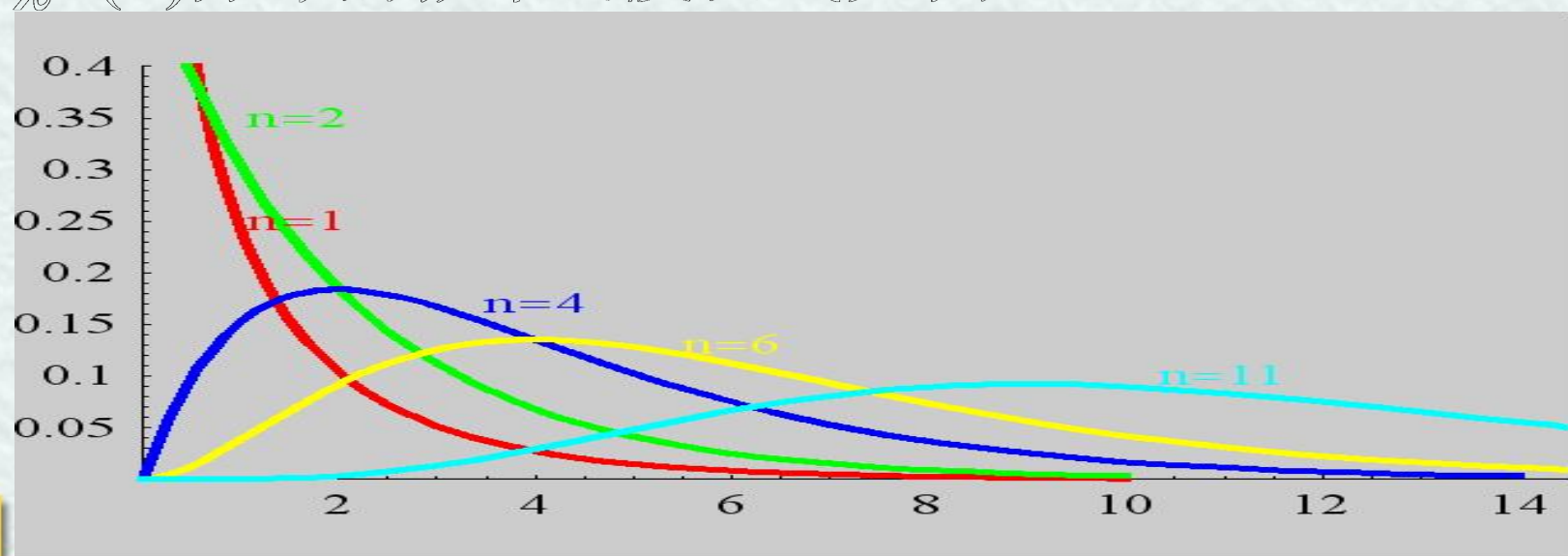
即 $X_i^2 \sim \Gamma\left(\frac{1}{2}, 2\right)$, $i = 1, 2, \dots, n$.



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

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

$\chi^2(n)$ 分布的概率密度曲线如图.



χ^2 分布的性质

性质1 (χ^2 分布的可加性)

设 $\chi_1^2 \sim \chi^2(n_1)$, $\chi_2^2 \sim \chi^2(n_2)$, 并且 χ_1^2 , χ_2^2 独立, 则 $\chi_1^2 + \chi_2^2 \sim \chi^2(n_1 + n_2)$.

(此性质可以推广到多个随机变量的情形.)

设 $\chi_i^2 \sim \chi^2(n_i)$, 并且 χ_i^2 ($i = 1, 2, \dots, m$) 相互独立, 则 $\sum_{i=1}^m \chi_i^2 \sim \chi^2(n_1 + n_2 + \dots + n_m)$.



性质2 (χ^2 分布的数学期望和方差)

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

证明 因为 $X_i \sim N(0, 1)$, 所以 $E(X_i^2) = D(X_i) = 1$,
 $D(X_i^2) = E(X_i^4) - [E(X_i^2)]^2 = 3 - 2 = 1, i = 1, 2, \dots, n$.

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

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



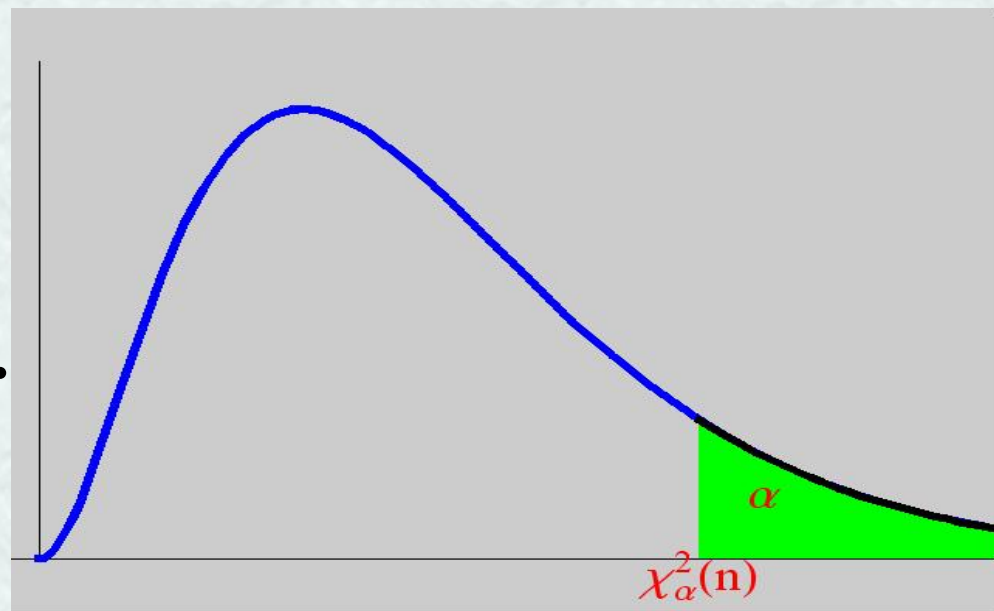
χ^2 分布的分位点

对于给定的正数 α , $0 < \alpha < 1$, 称满足条件

$$P\{\chi^2 > \chi_{\alpha}^2(n)\} = \int_{\chi_{\alpha}^2(n)}^{\infty} f(y)dy = \alpha$$

的点 $\chi_{\alpha}^2(n)$ 为 $\chi^2(n)$ 分布的上 α 分位点.

对于不同的 α, n ,
可以通过查表求
得上 α 分位点的值.



例1 设 X 服从标准正态分布 $N(0,1)$, $N(0,1)$ 的上

$$\alpha \text{ 分位点 } z_{\alpha} \text{ 满足 } P\{X > z_{\alpha}\} = \frac{1}{\sqrt{2\pi}} \int_{z_{\alpha}}^{+\infty} e^{-\frac{x^2}{2}} dx = \alpha,$$

求 z_{α} 的值, 可通过查表完成.

$$z_{0.05} = 1.645,$$

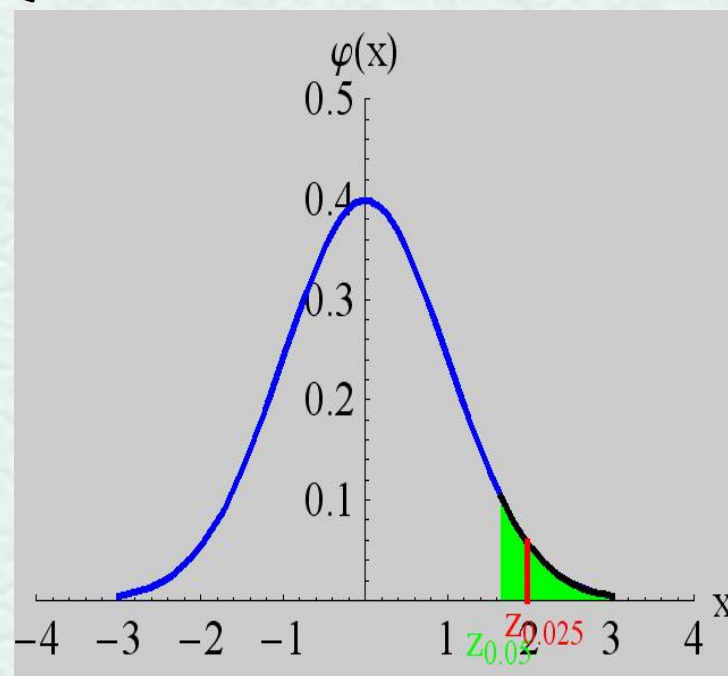
附表2-1

$$z_{0.025} = 1.96,$$

附表2-2

根据正态分布的对称性知

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



例2 设 $Z \sim \chi^2(n)$, $\chi^2(n)$ 的上 α 分位点满足

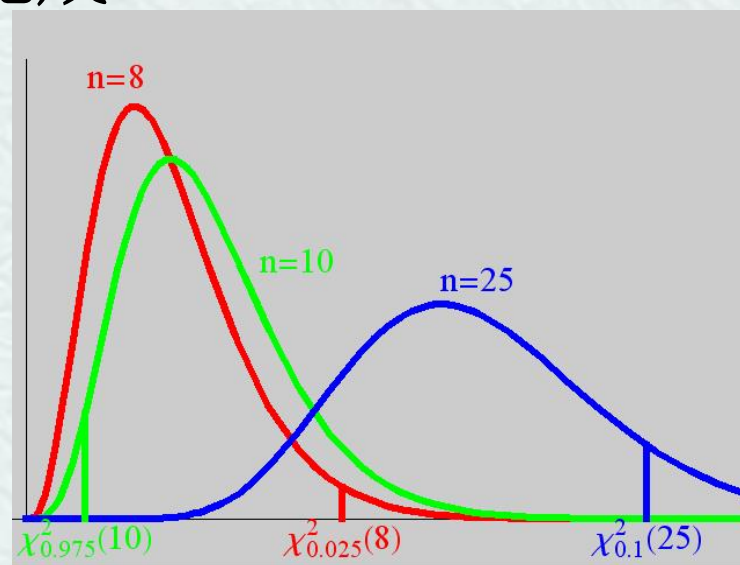
$$P\{Z > \chi_{\alpha}^2(n)\} = \int_{\chi_{\alpha}^2(n)}^{+\infty} \chi^2(y; n) dy = \alpha,$$

求 $\chi_{\alpha}^2(n)$ 的值, 可通过查表完成.

$$\chi_{0.025}^2(8) = 17.535, \quad \text{附表4-1}$$

$$\chi_{0.975}^2(10) = 3.247, \quad \text{附表4-2}$$

$$\chi_{0.1}^2(25) = 34.382. \quad \text{附表4-3}$$



附表4只详列到 $n=45$ 为止.

在Matlab中求解



费舍尔(R.A.Fisher)证明:

费舍尔资料

$$\text{当 } n \text{ 充分大时, } \chi_{\alpha}^2(n) \approx \frac{1}{2}(z_{\alpha} + \sqrt{2n-1})^2.$$

其中 z_{α} 是标准正态分布的上 α 分位点.

利用上面公式,

可以求得 $n > 45$ 时, 上 α 分位点的近似值.

例如 $\chi_{0.05}^2(50) \approx \frac{1}{2}(1.645 + \sqrt{99})^2 = 67.221.$

而查详表可得 $\chi_{0.05}^2(50) = 67.505.$



2. t 分布

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 $t \sim t(n)$.

学生氏资料

t 分布又称**学生氏(Student)分布**.

随机数演示

$t(n)$ 分布的概率密度函数为

分布函数与密度函数演示

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

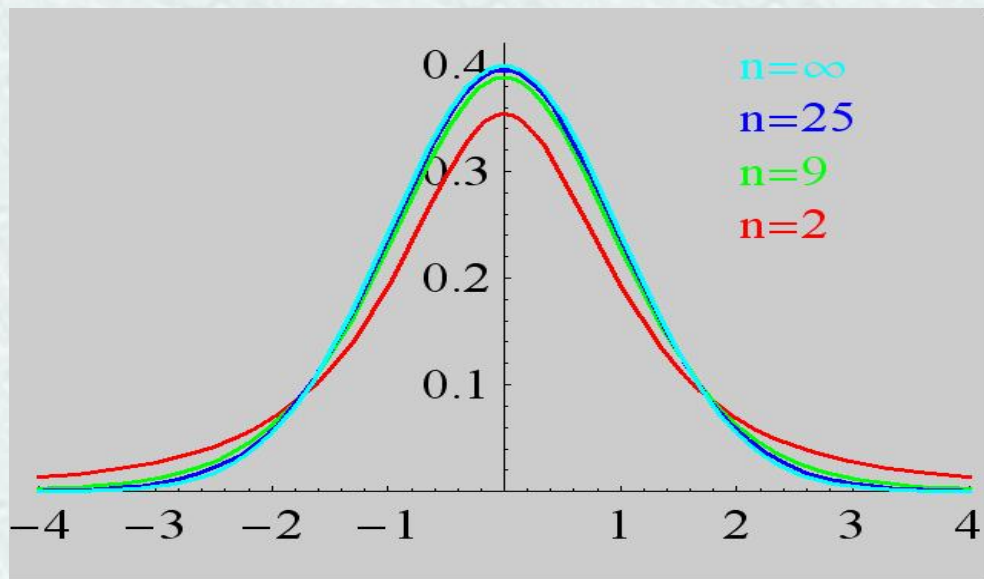


t 分布的概率密度曲线如图

显然图形是关于

$t = 0$ 对称的。

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。



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

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。



t 分布的分位点

对于给定的 α , $0 < \alpha < 1$, 称满足条件

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

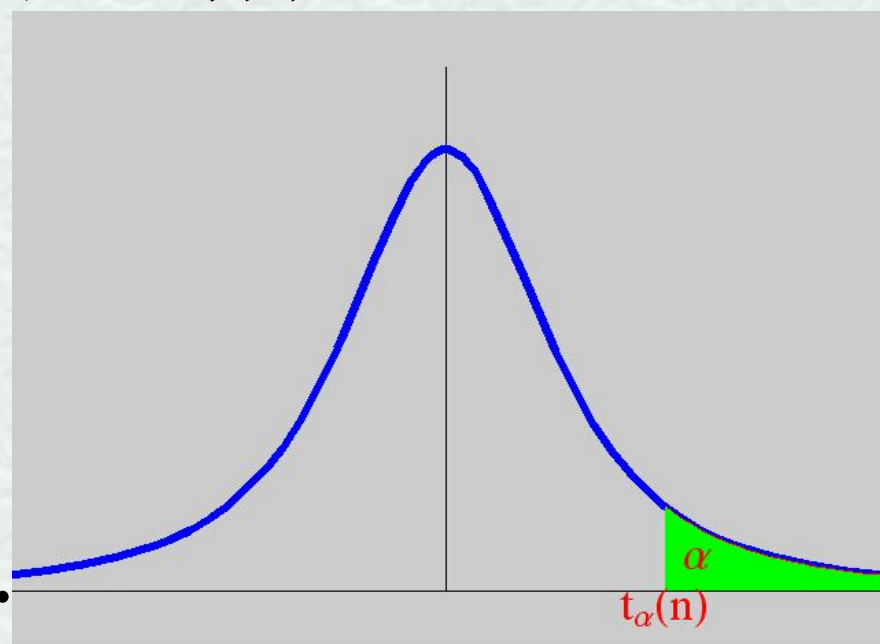
的点 $t_{\alpha}(n)$ 为 $t(n)$ 分布的上 α 分位点.

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, $t_{\alpha}(n) \approx z_{\alpha}$.



例3 设 $T \sim t(n)$, $t(n)$ 的上 α 分位点满足

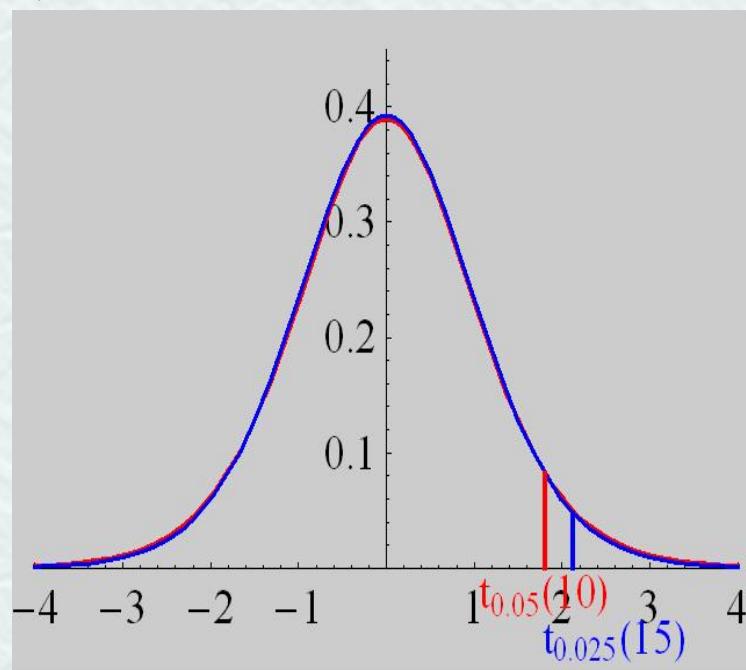
$$P\{T > t_{\alpha}(n)\} = \int_{t_{\alpha}(n)}^{+\infty} t(y; n) dy = \alpha,$$

求 $t_{\alpha}(n)$ 的值, 可通过查表完成.

$$t_{0.05}(10) = 1.8125, \quad \text{附表3-1}$$

$$t_{0.025}(15) = 2.1315. \quad \text{附表3-2}$$

在Matlab中求解



3. F 分布

设 $U \sim \chi^2(n_1)$, $V \sim \chi^2(n_2)$, 且 U, V 独立, 则称随机变量 $F = \frac{U/n_1}{V/n_2}$ 服从自由度为 (n_1, n_2) 的 F 分布, 记为 $F \sim F(n_1, n_2)$.

随机数演示

分布函数与密度函数演示



$F(n_1, n_2)$ 分布的概率密度为

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



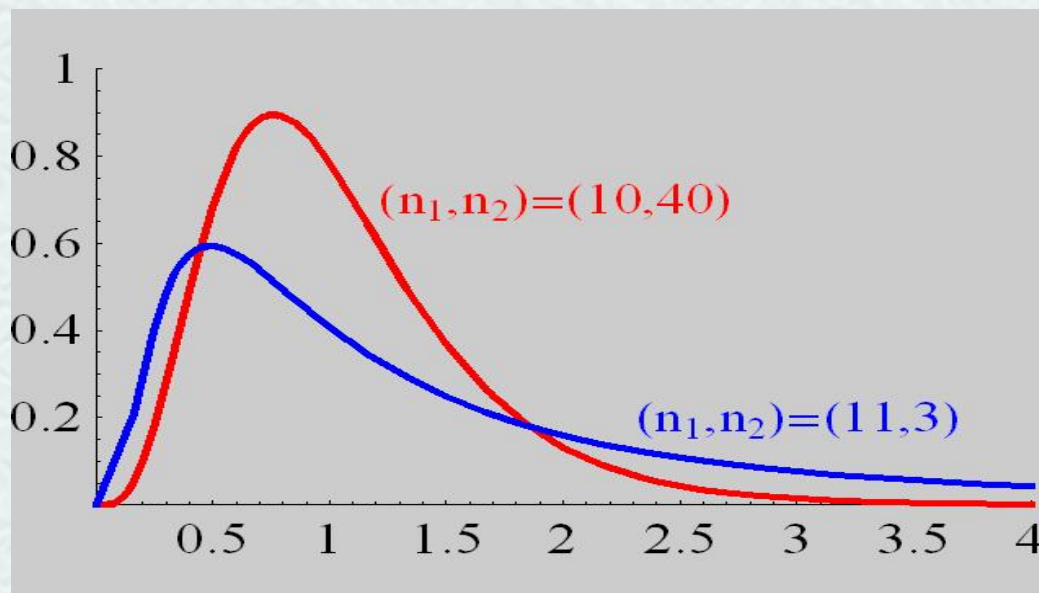
F 分布的概率密度曲线如图

根据定义可知,

若 $F \sim F(n_1, n_2)$,

则 $\frac{1}{F} \sim F(n_2, n_1)$.

F 分布的分位点



对于给定的 α , $0 < \alpha < 1$, 称满足条件

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

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



例4 设 $F(n_1, n_2)$ 分布的上 α 分位点满足

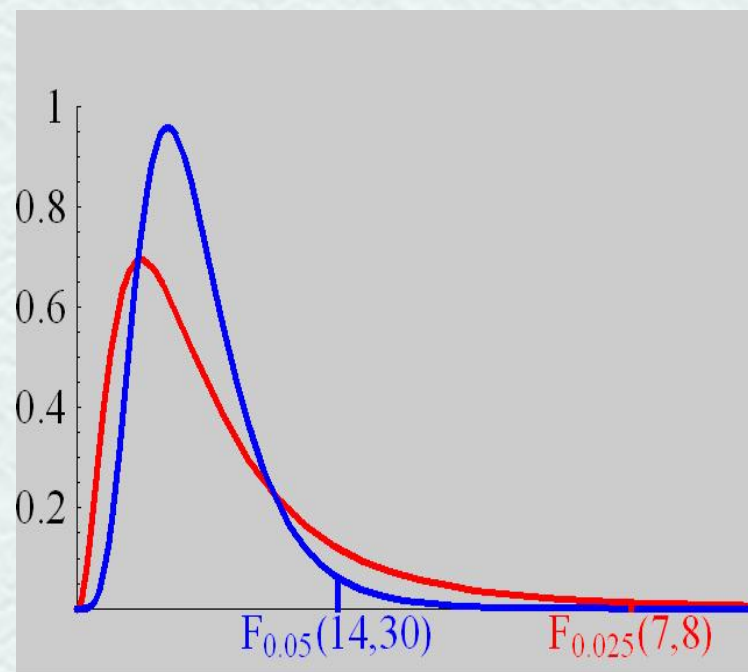
$$P\{F > F_\alpha(n_1, n_2)\} = \int_{F_\alpha(n_1, n_2)}^{+\infty} \psi(y) dy = \alpha,$$

求 $F_\alpha(n_1, n_2)$ 的值, 可通过查表完成.

$F_{0.025}(7, 8) = 4.90$, 附表5-1

$F_{0.05}(14, 30) = 2.31$. 附表5-2

在Matlab中求解



F 分布的上 α 分位点具有如下性质：

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

证明 因为 $F \sim F(n_1, n_2)$,

$$\begin{aligned} \text{所以 } 1-\alpha &= P\{F > F_{1-\alpha}(n_1, n_2)\} \\ &= P\left\{\frac{1}{F} < \frac{1}{F_{1-\alpha}(n_1, n_2)}\right\} = 1 - P\left\{\frac{1}{F} \geq \frac{1}{F_{1-\alpha}(n_1, n_2)}\right\} \\ &= 1 - P\left\{\frac{1}{F} > \frac{1}{F_{1-\alpha}(n_1, n_2)}\right\}, \end{aligned}$$

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



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

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

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

用来求分布表中未列出的一些上 α 分位点。

例 $F_{0.95}(12, 9) = \frac{1}{F_{0.05}(9, 12)} = \frac{1}{0.28} = 0.357$.



4. 正态总体的样本均值与样本方差的分布

定理一

设 $X_1, X_2, \dots, X_n$ 是来自正态总体 $N(\mu, \sigma^2)$ 的样本, $\bar{X}$ 是样本均值, 则有 $\bar{X} \sim N(\mu, \sigma^2 / n)$.

正态总体 $N(\mu, \sigma^2)$ 的样本均值和样本方差有以下两个重要定理.



定理二

设 $X_1, X_2, \dots, X_n$ 是总体 $N(\mu, \sigma^2)$ 的样本, $\bar{X}, S^2$ 分别是样本均值和样本方差, 则有

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

(2) $\bar{X}$ 与 S^2 独立.



定理三 设 $X_1, X_2, \dots, X_n$ 是总体 $N(\mu, \sigma^2)$ 的样本, $\bar{X}, S^2$ 分别是样本均值和样本方差, 则有

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

证明 因为 $\frac{\bar{X} - \mu}{\sigma / \sqrt{n}} \sim N(0, 1), \frac{(n-1)S^2}{\sigma^2} \sim \chi^2(n-1),$

且两者独立, 由 t 分布的定义知

$$\frac{\bar{X} - \mu}{\sigma / \sqrt{n}} \bigg/ \sqrt{\frac{(n-1)S^2}{\sigma^2(n-1)}} \sim t(n-1).$$



定理四 设 $X_1, X_2, \dots, X_{n_1}$ 与 $Y_1, Y_2, \dots, Y_{n_2}$ 分别是具有相同方差的两正态总体 $N(\mu_1, \sigma^2), N(\mu_2, \sigma^2)$ 的样本, 且这两个样本互相独立, 设 $\bar{X} = \frac{1}{n_1} \sum_{i=1}^{n_1} X_i$,

$\bar{Y} = \frac{1}{n_2} \sum_{i=1}^{n_2} Y_i$ 分别是这两个样本的均值,

$$S_1^2 = \frac{1}{n_1 - 1} \sum_{i=1}^{n_1} (X_i - \bar{X})^2, \quad S_2^2 = \frac{1}{n_2 - 1} \sum_{i=1}^{n_2} (Y_i - \bar{Y})^2$$

分别是这两个样本的方差, 则有



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

(2) 当 $\sigma_1^2 = \sigma_2^2 = \sigma^2$ 时,

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

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



证明 (1) 由定理二

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

由假设 S_1^2, S_2^2 独立, 则由 F 分布的定义知

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

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



$$(2) \quad \text{因为 } \bar{X} - \bar{Y} \sim N\left(\mu_1 - \mu_2, \frac{\sigma^2}{n_1} + \frac{\sigma^2}{n_2}\right)$$

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

$$\text{由 } \frac{(n_1 - 1)S_1^2}{\sigma^2} \sim \chi^2(n_1 - 1), \quad \frac{(n_2 - 1)S_2^2}{\sigma^2} \sim \chi^2(n_2 - 1),$$

且它们相互独立, 故由 χ^2 分布的可加性知



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

由于 U 与 V 相互独立, 按 t 分布的定义.

$$\begin{aligned} & \frac{U}{\sqrt{V/(n_1 + n_2 - 2)}} \\ &= \frac{(\bar{X} - \bar{Y}) - (\mu_1 - \mu_2)}{S_w \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \sim t(n_1 + n_2 - 2). \end{aligned}$$



三、小结

两个最重要的统计量:

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

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

三个来自正态分布的抽样分布:

χ^2 分布, t 分布, F 分布。



辛钦定理

设随机变量 $X_1, X_2, \dots, X_m, \dots$ 相互独立, 服从同一分布, 且具有数学期望 $E(X_k) = \mu$ ($k = 1, 2, \dots$),

则对于任意正数 ε , 有 $\lim_{m \rightarrow \infty} P \left\{ \left| \frac{1}{m} \sum_{k=1}^m X_k - \mu \right| < \varepsilon \right\} = 1$.



附表2-1

标准正态分布表

z	0	1	2	3	4	5	6	7	8	9
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.6951	0.6987	0.7023	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.7	0.7580	0.7611	0.7642	0.7673	0.7703	0.7734	0.7764	0.7794	0.7823	0.7852
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
1.0	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810	0.8830
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997	0.9015
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9278	0.9292	0.9306	0.9319
1.5	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394	0.9406	0.9418	0.9430	0.9441
1.6	0.9452	0.9463	0.9474	0.9484	0.9495	0.9505	0.9515	0.9525	0.9535	0.9545

1.645



附表2-2

标准正态分布表

z	0	1	2	3	4	5	6	7	8	9
1.6	0.9452	0.9463	0.9474	0.9484	0.9495	0.9505	0.9515	0.9525	0.9535	0.9545
1.7	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	0.9608	0.9616	0.9625	0.9633
1.8	0.9641	0.9648	0.9656	0.9664	0.9671	0.9678	0.9686	0.9693	0.9700	0.9706
1.9	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.9750	0.9756	0.9762	0.9767
2.0	0.9772	0.9778	0.9783	0.9788	0.9793	0.9798	0.9803	0.9808	0.9812	0.9817
2.1	0.9821	0.9826	0.9831	0.9836	0.9842	0.9846	0.9846	0.9850	0.9854	0.9853
2.2	0.9861	0.9864	0.9868	0.9871	0.9878	0.9881	0.9881	0.9884	0.9887	0.9890
2.3	0.9893	0.9896	0.9898	0.9901	0.9906	0.9909	0.9909	0.9911	0.9913	0.9916
2.4	0.9918	0.9920	0.9922	0.9925	0.9927	0.9929	0.9931	0.9932	0.9934	0.9936
2.5	0.9938	0.9940	0.9941	0.9943	0.9945	0.9946	0.9948	0.9949	0.9951	0.9952
2.6	0.9953	0.9955	0.9956	0.9957	0.9959	0.9960	0.9961	0.9962	0.9963	0.9964
2.7	0.9965	0.9966	0.9967	0.9968	0.9969	0.9970	0.9971	0.9972	0.9973	0.9974
2.8	0.9974	0.9975	0.9976	0.9977	0.9977	0.9978	0.9979	0.9979	0.9980	0.9981
2.9	0.9981	0.9982	0.9982	0.9983	0.9984	0.9984	0.9985	0.9985	0.9986	0.9986
3.0	0.9987	0.9990	0.9993	0.9995	0.9997	0.9998	0.9998	0.9999	0.9999	1.0000

1.96



附表4-1

χ^2 分布表

n	$\alpha=0.25$	0.10	0.05	0.025	0.01	0.005
1	1.323	2.706	3.841	5.024	6.635	7.879
2	2.773	4.605	5.991	7.378	9.210	10.597
3	4.108	6.251	7.815	9.348	11.345	12.838
4	5.385	7.779	9.488	11.143	13.277	14.860
5	6.626	9.236	11.070	12.833	15.086	16.750
6	7.841	10.597	12.592	14.449	16.812	18.548
7	9.037	12.017	14.067	16.013	18.475	20.278
8	10.219	13.362	15.507	17.535	20.090	21.955
9	11.389	14.684	16.919	19.023	21.666	23.589
10	12.549	15.987	18.307	20.483	23.209	25.188
11	13.701	17.275	19.675	21.920	24.725	26.757
12	14.845	18.549	21.026	23.337	26.217	28.299
13	15.984	19.812	22.362	24.736	27.688	29.891
14	17.117	20.064	23.685	26.119	29.141	31.319
15	18.245	22.307	24.996	27.488	30.578	32.801
16	19.369	23.542	26.296	28.845	32.000	34.267

17.535

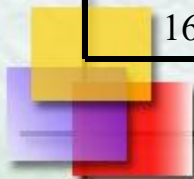


附表4-2

χ^2 分布表

n	$\alpha = 0.995$	0.99	0.975	0.95	0.90	0.75
1	—	—	0.001	0.004	0.016	0.102
2	0.010	0.020	0.051	0.103	0.211	0.575
3	0.072	0.115	0.216	0.352	0.584	1.213
4	0.207	0.297	0.484	0.711	1.064	1.923
5	0.412	0.554	0.831	1.145	1.610	2.675
6	0.675	0.872	1.237	1.635	2.204	3.455
7	0.975	1.239	1.690	2.167	2.833	4.255
8	1.312	1.649	2.180	2.733	3.490	5.071
9	1.735	2.088	2.700	3.325	4.168	5.899
10	2.156	2.558	3.247	3.940	4.865	6.737
11	2.603	3.053	3.816	4.575	5.578	7.584
12	3.074	3.571	4.404	5.226	6.304	8.438
13	3.565	4.107	5.009	5.892	7.042	9.299
14	4.075	4.660	5.629	6.571	7.790	10.165
15	4.601	5.229	6.262	7.261	8.547	11.037
16	5.142	5.812	6.908	7.962	9.312	11.912

3.247



附表4-3

χ^2 分布表

n	$\alpha=0.25$	0.10	0.05	0.025	0.01	0.005
17	20.489	24.769	27.587	30.191	33.409	35.718
18	21.605	25.989	28.869	31.526	34.805	37.156
19	22.718	27.204	30.144	32.852	36.191	38.582
20	23.828	28.412	31.410	34.170	37.566	39.997
21	24.935	29.615	32.671	35.479	38.932	41.401
22	26.039	30.813	33.924	36.781	40.289	42.796
23	27.141	32.007	35.178	38.076	41.638	44.181
24	28.241	33.196	36.415	39.364	42.980	45.559
25	29.339	34.382	37.652	40.646	44.314	46.928
26	30.435	35.563	38.885	41.923	45.642	48.290
27	31.528	36.741	40.113	43.194	46.963	49.645
28	32.620	37.916	41.337	44.461	48.278	50.993
29	33.711	39.087	42.557	45.712	49.588	52.336
30	34.800	40.256	43.773	46.979	50.892	53.672
31	35.887	41.422	44.985	48.232	52.191	55.003
32	36.973	42.585	46.194	49.480	53.486	56.328

34.382

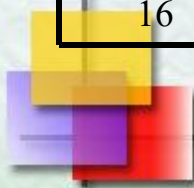


附表3-1

t 分布表

n	$\alpha = 0.25$	0.10	0.05	0.025	0.01	0.005
1	1.0000	3.0777	6.3138	12.7062	31.8207	63.6574
2	0.8165	1.8856	2.9200	4.3027	6.9646	9.9248
3	0.7649	1.6377	2.3534	3.1824	4.5407	5.8409
4	0.7407	1.5332	2.1318	2.7764	3.7469	4.6041
5	0.7267	1.4759	2.0150	2.5706	3.3649	4.0322
6	0.7176	1.4398	1.9432	2.4477	3.1427	3.7074
7	0.7111	1.4149	1.8946	2.3646	2.9980	3.4995
8	0.7064	1.3968	1.8595	2.3060	2.8965	3.3554
9	0.7027	1.3830	1.8331	2.2622	2.8214	3.2498
10	0.6998	1.3722	1.8125	2.2281	2.7638	3.1693
11	0.6974	1.3634	1.7959	2.2010	2.7181	3.1058
12	0.6955	1.3562	1.7823	2.1788	2.6810	3.0545
13	0.6938	1.3502	1.7709	2.1604	2.6503	3.0123
14	0.6924	1.3450	1.7613	2.1448	2.6245	2.9768
15	0.6912	1.3406	1.7531	2.1315	2.6025	2.9467
16	0.6901	1.3368	1.7459	2.1199	2.5835	2.9208

1.8125

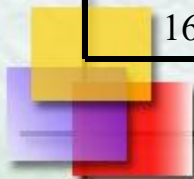


附表3-2

t 分布表

n	$\alpha=0.25$	0.10	0.05	0.025	0.01	0.005
1	1.0000	3.0777	6.3138	12.7062	31.8207	63.6574
2	0.8165	1.8856	2.9200	4.3027	6.9646	9.9248
3	0.7649	1.6377	2.3534	3.1824	4.5407	5.8409
4	0.7407	1.5332	2.1318	2.7764	3.7469	4.6041
5	0.7267	1.4759	2.0150	2.5706	3.3649	4.0322
6	0.7176	1.4398	1.9432	2.4469	3.1427	3.7074
7	0.7111	1.4149	1.8946	2.3646	2.9980	3.4995
8	0.7064	1.3968	1.8595	2.3060	2.8965	3.3554
9	0.7027	1.3830	1.8331	2.2622	2.8214	3.2498
10	0.6998	1.3707	1.8125	2.2281	2.7638	3.1693
11	0.6974	1.3601	1.7953	2.2010	2.7181	3.1058
12	0.6955	1.3502	1.7799	2.1788	2.6810	3.0545
13	0.6938	1.3450	1.7661	2.1604	2.6503	3.0123
14	0.6924	1.3406	1.7531	2.1448	2.6245	2.9768
15	0.6912	1.3368	1.7415	2.1315	2.6025	2.9467
16	0.6901	1.3336	1.7307	2.1199	2.5835	2.9208

2.1315



附表5-1

F分布表

$$\alpha = 0.025$$

$n_1 \backslash n_2$	1	2	3	4	5	6	7	8	9	10	12	15	20	24	30	40	120	∞
1	647.8	799.5	864.2	899.6	921.8	937.1	948.2	956.7	963.3	968.6	976.7	984.9	993.1	997.2	1001	1006	1014	1018
2	38.51	39.00	39.17	39.25	39.30	39.33	39.36	39.37	39.39	39.40	39.41	39.43	39.45	39.46	39.46	39.47	39.49	39.50
3	17.44	16.04	15.44	15.10	14.88	14.73	14.62	14.54	14.47	14.42	14.34	14.25	14.17	14.12	14.08	14.04	13.95	13.90
4	12.22	10.65	9.98	9.60	9.36	9.20	9.07	8.98	8.90	8.84	8.75	8.66	8.56	8.51	8.46	8.41	8.31	8.26
5	10.01	8.43	7.76	7.39	7.15	6.98	6.85	6.76	6.68	6.62	6.52	6.43	6.33	6.28	6.23	6.18	6.07	6.02
6	8.81	7.26	6.60	6.23	5.99	5.82	5.70	5.60	5.52	5.46	5.37	5.27	5.17	5.12	5.07	5.01	4.90	4.85
7	8.07	6.54	5.89	5.52	5.29	5.12	4.99	4.90	4.82	4.76	4.67	4.57	4.47	4.42	4.36	4.31	4.20	4.14
8	7.57	6.06	5.42	5.50	4.82	4.65	4.53	4.44	4.36	4.30	4.20	4.10	4.00	3.59	3.89	3.84	3.73	3.67
9	7.21	5.71	5.08	4.72	4.48	4.23	4.20	4.11	4.03	3.96	3.87	3.77	3.67	3.61	3.56	3.51	3.39	3.33
10	6.94	5.46	4.83	4.47	4.24	4.07	3.95	3.86	3.78	3.72	3.62	3.52	3.42	3.37	3.31	3.26	3.14	3.08
11	6.72	5.26	4.63	4.28	4.04	3.88	3.76	3.67	3.59	3.53	3.43	3.33	3.23	3.17	3.12	3.06	2.94	2.88
12	6.55	5.10	4.47	4.12	3.89	3.73	3.61	3.52	3.44	3.38	3.28	3.18	3.07	3.02	2.96	2.91	2.79	2.72
13	6.41	4.97	4.35	4.00	3.77	3.60	3.48	3.39	3.31	3.25	3.15	3.05	2.95	2.89	2.84	2.78	2.66	2.60
14	6.30	4.86	4.24	3.89	3.66	3.50	3.38	3.29	3.21	3.15	3.05	2.95	2.84	2.79	2.73	2.67	2.55	2.49
15	6.20	4.77	4.15	3.80	3.58	3.41	3.29	3.20	3.12	3.06	2.96	2.86	2.76	2.70	2.64	2.59	2.46	2.40
16	6.12	4.69	4.08	3.73	3.50	3.34	3.22	3.13	3.05	2.99	2.89	2.79	2.68	2.63	2.57	2.51	2.38	2.32
17	6.04	4.62	4.01	3.66	3.44	3.28	3.16	3.07	2.99	2.93	2.83	2.72	2.62	2.56	2.50	2.44	2.32	2.25
18	5.95	4.56	3.95	3.61	3.38	3.22	3.10	3.01	2.93	2.87	2.77	2.67	2.56	2.50	2.44	2.38	2.26	2.19
19	5.92	4.51	3.90	3.56	3.33	3.17	3.05	2.96	2.88	2.82	2.72	2.62	2.51	2.45	2.39	2.33	2.20	2.13

4.90



附表5-2

F分布表

$$\alpha = 0.05$$

$n_1 \backslash n_2$	1	2	3	4	5	6	7	8	9	10	15	20	24	30	40	60	120	∞
1	161.4	199.5	215.7	224.6	230.2	234.0	236.8	238.9	24.05	241.9	245.9	248.0	249.1	250.1	151.1	252.2	253.3	254.3
2	18.51	19.00	19.16	19.25	19.30	19.33	19.35	19.37	19.38	19.40	19.43	19.45	19.45	19.46	19.47	19.48	19.49	19.50
3	10.13	9.55	9.28	9.12	9.01	8.94	8.89	8.85	8.81	8.79	8.70	8.66	8.64	8.62	8.59	8.57	8.55	8.53
4	7.71	6.94	6.59	6.39	6.26	6.16	6.09	6.04	6.00	5.96	5.86	5.80	5.77	5.75	5.72	5.69	5.66	5.63
5	6.61	5.79	5.41	5.19	5.05	4.95	4.88	4.82	4.77	4.74	4.62	4.56	4.53	4.50	4.46	4.43	4.40	4.36
6	5.99	5.14	4.76	4.53	4.39	4.28	4.21	4.15	4.10	4.06	3.94	3.87	3.84	3.81	3.77	3.74	3.70	3.67
7	5.59	4.74	4.35	4.12	3.97	3.87	3.79	3.73	3.68	3.64	3.51	3.44	3.41	3.38	3.34	3.30	3.27	3.23
8	5.32	4.46	4.07	3.84	3.69	3.58	3.50	3.44	3.39	3.35	3.22	3.15	3.12	3.08	3.04	3.01	2.97	2.93
9	3.12	4.26	3.81	3.63	3.48	3.37	3.29	3.23	3.18	3.14	3.01	2.94	2.90	2.86	2.83	2.79	2.75	2.71
10	4.96	4.10	3.71	3.48	3.33	3.22	3.14	3.07	3.02	2.98	2.85	2.77	2.74	2.70	2.66	2.62	2.58	2.54
11	4.84	3.98	3.59	3.36	3.20	3.09	3.01	2.95	2.89	2.85	2.72	2.65	2.61	2.57	2.53	2.49	2.45	2.40
12	4.75	3.89	3.49	3.26	3.11	3.00	2.91	2.85	2.79	2.75	2.62	2.55	2.51	2.47	2.43	2.38	2.34	2.30
13	4.67	3.81	3.41	3.18	3.03	2.92	2.83	2.77	2.71	2.67	2.54	2.47	2.42	2.38	2.34	2.30	2.25	2.21
14	4.60	3.74	3.34	3.11	2.96	2.85	2.76	2.70	2.64	2.60	2.47	2.40	2.35	2.31	2.27	2.22	2.18	2.13
15	4.54	3.68	3.29	3.06	2.90	2.79	2.71	2.65	2.59	2.55	2.42	2.35	2.29	2.25	2.20	2.16	2.11	2.07
16	4.49	3.63	3.24	3.01	2.85	2.74	2.66	2.59	2.54	2.49	2.35	2.28	2.24	2.19	2.15	2.11	2.06	2.01
17	4.45	3.59	3.20	2.96	2.81	2.70	2.61	2.55	2.49	2.45	2.31	2.23	2.19	2.15	2.10	2.06	2.01	1.96
18	4.41	3.55	3.16	2.93	2.77	2.66	2.58	2.51	2.46	2.41	2.27	2.19	2.15	2.11	2.06	2.02	1.97	1.92
19	4.38	3.52	3.13	2.90	2.74	2.63	2.54	2.48	2.42	2.38	2.23	2.16	2.11	2.07	2.03	1.98	1.93	1.88

2.31



费舍尔资料



Ronald Aylmer Fisher

**Born: 17 Feb 1890 in
London, England**

**Died: 29 Jul. 1962 in
Adelaide, Australia**



学生氏资料



William Sealey Gosset

**Born: 13 Jun. 1876 in
Canterbury, England**

**Died: 16 Oct. 1937 in
Beaconsfield, England**



格里汶科资料

Boris Vladimirovich Gnedenko



**Born: 1 Jan. 1912 in
Simbirsk (now
Ulyanovskaya), Russia**

**Died: 27 Dec. 1995 in
Moscow, Russia**

