10956

概率与数理統計

（下 册）

数 学 教 研 室

北 京 化 工 学 院

第六章　大数定律和中心极限定理

概率论的数学规律是把大量具体的随机现象共有的规律加以抽象而得到的，这些共同的规律与现象的大量性有关。为了研究大量的随机现象，常常采用极限形式，这就导致人们对极限定理的研究。极限定理的内容很广泛，其中最重要的有两种：大数定律和中心极限定律。

§1　大数定律

概率论中论证大量随机现象平均结果的稳定性的一系列定理统称为大数定律。

在大量随机现象中，我们不仅看到随机事件的频率具有稳定性，而且还看到一般的平均结果也具有稳定性。这就是说：无论个别随机现象的结果以及它们在进行过程中的个别特征如何，它们所产生的偏差在大量的随机现象中相互抵消了，致使大量随机现象的平均结果实际上与每一个别随机现象的特征无关，几乎不再是随机的了。

契比雪夫定理：设独立随机变量

$$\xi_1、\xi_2、\cdots、\xi_n、\cdots。$$

分别具有数学期望

$$E\xi_1、E\xi_2、\cdots、E\xi_n、\cdots，$$

及方差

$$D\xi_1、D\xi_2、\cdots、D\xi_n、\cdots，$$

并且方差是一致有上界的，即存在某一常数K，使得

$$D\xi_i < K，\quad i=1、2、\cdots n、\cdots 。$$

则对于任何正数 ε，恒有

$$\lim_{n\to\infty} P\{|\frac{1}{n}\sum_{i=1}^{n}\xi_i - \frac{1}{n}\sum_{i=1}^{n}E\xi_i|<\varepsilon\}=1 。$$

<证> 我们有

$$E\left(\frac{1}{n}\sum_{i=1}^{n}\xi_i\right)=\frac{1}{n}E\left(\sum_{i=1}^{n}\xi_i\right)=\frac{1}{n}\sum_{i=1}^{n}E\xi_i$$

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

对于随机变量$\frac{1}{n}\sum_{i=1}^{n}\xi_i$应用契比雪夫不等式得

$$P\left\{\left|\frac{1}{n}\sum_{i=1}^{n}\xi_i-\frac{1}{n}\sum_{i=1}^{n}E\xi_i\right|<\varepsilon\right\}\geq 1-\frac{\sum_{i=1}^{n}D\xi_i}{n^2\varepsilon^2}$$

因为方差是一致有上界的，所以

$$\sum_{i=1}^{n}D\xi_i< nK$$

由此得

$$P\left\{\left|\frac{1}{n}\sum_{i=1}^{n}\xi_i-\frac{1}{n}\sum_{i=1}^{n}E\xi_i\right|<\varepsilon\right\}>1-\frac{K}{n\varepsilon^2}$$

当$n\to\infty$时，得到

$$\lim_{n\to\infty}P\left\{\left|\frac{1}{n}\sum_{i=1}^{n}\xi_i-\frac{1}{n}\sum_{i=1}^{n}E\xi_i\right|<\varepsilon\right\}\geq 1$$

但概率不能大于1，所以有

$$\lim_{n\to\infty}P\left\{\left|\frac{1}{n}\sum_{i=1}^{n}\xi_i-\frac{1}{n}\sum_{i=1}^{n}E\xi_i\right|<\varepsilon\right\}=1$$

契比雪夫定理的意义[illegible]以作如下的解释：

由于独立随机变量ξ_1、ξ_2、…、ξ_n的算术平均值

$$\bar{\xi}_n=\frac{1}{n}\sum_{i=1}^{n}\xi_i$$

的数学期望$E\bar{\xi}_n$等于数学期望$E\xi_1$、$E\xi_2$、…、$E\xi_n$的算术平均值，而方差$D\bar{\xi}_n$等于方差$D\xi_1$、$D\xi_2$、…、$D\xi_n$的算术平均值的$\frac{1}{n}$。所以如果方差是一致有上界的，则当n无限增大时，$D\bar{\xi}_n$将是一个无穷小量。由此可见当n充分大时，随机变量$\bar{\xi}_n$的值将紧密地聚集在它的数学期望$E\bar{\xi}_n$附近，契比雪夫定理为此给出了精确的数学公式。

定义：设$\{a_n\}$为一常数序列，$\{\xi_n\}$为一随机变量序列。如果对于任一$\varepsilon>0$有

$$\lim_{n\to\infty}P\{|\xi_n-a_n|<\varepsilon\}=1$$

则称随机变量序列$\{\xi_n\}$当$n\to\infty$时，依概率收敛于常数序列$\{a_n\}$。

利用依概率收敛的概念，契比雪夫定理也可以这样说：

设独立随机变量ξ_1、ξ_2、…、ξ_n、…，有数学期望及方差，且方差是一致有上界的，则ξ_1、ξ_2、…、ξ_n的算术平均值依概率收敛于它们的数学期望的算术平均值。

作为契比雪夫定理的特例，有下面的推论：

推论：设独立随机变量ξ_1、ξ_2、…、ξ_n、…服从同一分布，数学期望为a，方差为σ^2，则ξ_1、ξ_2、…、ξ_n的算术平均值当$n\to\infty$时，依概率收敛于数学期望a，即对于任何正数ε恒有

$$\lim_{n\to\infty}P\{|\frac{1}{n}\sum_{i=1}^{n}\xi_i-a|<\varepsilon\}=1$$

这一推论使我们关于算术平均值的法则有了理论依据。假设我们要测量某一物理量a，在不变条件下重复测量n次，得到的结果x_1，x_2，…、x_n是不完全相同的，这些结果可以看作n个独立随机变量ξ_1、ξ_2、…、ξ_n（显然它们服从同一分布，并且数学期望为a）的实验数值。于是按大数定理当n充分大时，我们取n次测量结果的算术平均值作为a的近似值：

$$a \approx \frac{1}{n}(x_1+x_2+\cdots+x_n)$$

所发生的误差是很小的。

贝努里定理：设随机变量

$$\eta_n \sim B(n, p), \qquad n=1, 2, \cdots. \quad 1>p>0$$

那末对于任何$\varepsilon>0$有

$$\lim_{n\to\infty} P\{|\frac{\eta_n}{n}-p|<\varepsilon\}=1$$

<证> 由于$\eta_n \sim B(n,p)$，所以$E\eta_n=np$，

$D\eta_n=np(1-p)$。因此

$$E(\frac{\eta_n}{n})=\frac{np}{n}=p \quad .$$

$$D(\frac{\eta_n}{n})=\frac{D\eta_n}{n^2}=\frac{1}{n}p(1-p).$$

按契比雪夫不等式，对于任一$\varepsilon>0$有

$$P\{|\frac{\eta_n}{n}-p|<\varepsilon\}\geq 1-\frac{D(\frac{\eta_n}{n})}{\varepsilon^2}=1-\frac{p(1-p)}{n\varepsilon^2} \quad .$$

当$n\to\infty$时，便有

$$\lim_{n\to\infty} P\{|\frac{\eta_n}{n}-p|<\varepsilon\}=1 。$$

由于$\frac{\eta_n}{n}$是n次独立重复试验中事件A出现的频率，而p是每次试验中事件A出现的概率。贝努里定理说明：当n足够大时，事件A出现的频率在它的概率附近摆动，所以用频率来代替概率所产生的误差是很小的。

§2　中心极限定理

概率论中有关论证随机变量和的极限分布是正态分布的一般定理称为中心极限定理。

在介绍两个定理之前再复述一下标准化的随机变量。

设随机变量ξ的数学期望为a，方差为σ^2，则随机变量

$$\eta=\frac{\xi-a}{\sigma}$$

的数学期望$E\eta=0$，方差$D\eta=1$，称η为ξ的标准化的随机变量。

设相互独立的随机变量ξ_1、ξ_2、…、ξ_n的数学期望为$E\xi_1$、$E\xi_2$、…、$E\xi_n$，方差为$D\xi_1$、$D\xi_2$、…、$D\xi_n$。则随机变量

$$\eta_n=\sum_{i=1}^{n}\frac{\xi_i-E\xi_i}{B_n}\text{，其中}\quad B_n^2=\sum_{i=1}^{n}D\xi_i\ .$$

的数学期望$E\eta_n=0$，方差$D\eta_n=1$。称η_n为随机变量ξ_1、ξ_2、…、ξ_n和的标准化的随机变量。

同分布的中心极限定理：设相互独立的随机变量ξ_1、ξ_2、…、ξ_n、…，具有相同的分布，并且

$$0<D\xi_i<+\infty\ .\qquad i=1,2,\cdots,n,\cdots\ .$$

则当$n\to\infty$时，对x均匀地有

$$P\{\sum_{i=1}^{n}\frac{\xi_i-E\xi_i}{B_n}<x\}\to\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x}e^{-\frac{t^2}{2}}dt\ 。$$

棣莫佛——拉普拉斯中心极限定理：设随机变量 $\eta_n \sim B(n, p)$，$n = 1、2、\cdots$ $0<p<1$。则标准化的随机变量

$$\frac{\eta_n - np}{\sqrt{np(1-p)}}$$

落在任一区间 (a, b) 内的概率有

$$\lim_{n\to\infty} P\{a \le \frac{\eta_n - np}{\sqrt{np(1-p)}} < b\} = \int_a^b \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt \text{ 。}$$

例1 一加法器同时收到20个噪声电压 ξ_k ($K=1.2.\cdots.20$)，设它们是相互独立的随机变量，且都在区间 $(0, 10)$ 上服从均匀分布，记

$$\eta = \sum_{k=1}^{20} \xi_k$$

求 $P\{\eta > 105\}$ 。

<解>按 ξ_k 在区间 $(0, 10)$ 上服从均匀分布可知

$$E\xi_k = 5 \text{ ,} \quad D\xi_k = \frac{100}{12} \text{ ,} \quad K = 1, 2, \cdots 20 \text{ 。}$$

由同分布的中心极限定理知标准化的随机变量

$$\sum_{k=1}^{20} \frac{\xi_k - 5}{\sqrt{20 \times \frac{100}{12}}} = \frac{\eta - 5\times 20}{\sqrt{20 \times \frac{100}{12}}}$$

近似地服从正态分布 $N(0, 1)$ 。于是

$$P\{\eta > 105\} = P\{\frac{\eta - 5\times 20}{\sqrt{20\times\frac{100}{12}}} > \frac{105 - 5\times 20}{\sqrt{20\times\frac{100}{12}}}\}$$

$$=P\{\frac{\eta-100}{\sqrt{20\times\frac{100}{12}}}>0.3873\}$$

$$=1-P\{\frac{\eta-100}{\sqrt{20\times\frac{100}{12}}}\leq 0.3873\}$$

$$\approx 1-F_{0,1}(0.3873)$$

$$=1-0.6517=0.3483\text{。}$$

例2　射击不断地进行，设每次射击的命中率为$\frac{1}{10}$。①求500次射击中，命中次数在区间〔49，55）之中的概率；②问至少要射击多少次，才能使命中的次数超过50次的概率大于80%　？

＜解＞　设500次射击中命中次数为η

$$\eta\sim B(500,\frac{1}{10})$$

由隶莫佛——拉普斯中心极限定理知标准化的随机变量

$$\frac{\eta-500\times\frac{1}{10}}{\sqrt{500\times\frac{1}{10}\times(1-\frac{1}{10})}}=\frac{\eta-50}{\sqrt{45}}$$

近似地服从正态分布$N(0,1)$，于是在500次射击中命中次数η落在区间〔49，55）之中的概率

$$P\{49\leq\eta<55\}=P\{\frac{49-50}{\sqrt{45}}\leq\frac{\eta-50}{\sqrt{45}}<\frac{55-50}{\sqrt{45}}\}$$

$$=P\{-0.149\leq\frac{\eta-50}{\sqrt{45}}<0.745\}$$

$$\approx F_{0,1}(0.745)-F_{0,1}(-0.149)$$

$$=0.7708-0.4400=0.3308\text{。}$$

设至少要射击n次，才能使命中的次数η_n超过50次的概率大于0.80。即

$$P\{\eta_n>50\}>0.80$$

由隶莫佛——拉普拉斯中心极限定理知标准化的随机变量

$$\frac{\eta_n-n\times\frac{1}{10}}{\sqrt{n\times\frac{1}{10}\times\frac{9}{10}}}=\frac{\eta_n-\frac{n}{10}}{\frac{3}{10}\sqrt{n}}$$

近似地服从正态分布N(0,1)，于是

$$P\{\eta_n>50\}=P\left\{\frac{\eta_n-\frac{n}{10}}{\frac{3}{10}\sqrt{n}}>\frac{50-\frac{n}{10}}{\frac{3}{10}\sqrt{n}}\right\}$$

$$=P\left\{\frac{\eta_n-\frac{n}{10}}{\frac{3}{10}\sqrt{n}}>\frac{500-n}{3\sqrt{n}}\right\}$$

$$=1-P\left\{\frac{\eta_n-\frac{n}{10}}{\frac{3}{10}\sqrt{n}}\leq\frac{500-n}{3\sqrt{n}}\right\}$$

$$\approx 1-F_{0,1}\left(\frac{500-n}{3\sqrt{n}}\right)$$

解不等式

$$1-F_{0,1}\left(\frac{500-n}{3\sqrt{n}}\right)>0.80$$

即

$$F_{0,1}\left(\frac{500-n}{3\sqrt{n}}\right)<0.20$$

反查标准正态表得

$$\frac{500-n}{3\sqrt{n}} < -0.845$$

即

$$n-2.535\sqrt{n}-500>0\ .$$

分解因式

$$(\sqrt{n}+21.129)(\sqrt{n}-23.664)>0\ .$$

取

$$\sqrt{n}>23.664\ ,\quad n>559.985\ .$$

取最小的正整数 n 为560，即当射击次数超过560 次时，才能使命中次数超过50 次的概率大于80%。

习 题 六

1. 卡车装运水泥，设每袋水泥重量 ξ 是一个随机变量，且 $\xi \sim N(50, 2.5^2)$。问装多少袋水泥才能使总重量超过20000公斤的概率为0.05？

2. 计算机在进行加法时，对每个加数取整（取为最接近于它的整数），设所有的取整误差是相互独立的，且它们都在区间 $(-0.5, 0.5)$ 上服从均匀分布。

① 若将1500个数相加，问误差总和的绝对值超过15 的概率是多少？

② 多少个数加在一起可使得误差总和的绝对值小于10 的概率为0.90？

3. 设有30个电子器件D_1，D_2，…，D_{30}，它们的使用情况如下：D_1损坏，D_2立即使用，D_2损坏D_3立即使用，…。设器件D_i的寿命ξ_i是服从参数为$\lambda=0.1$(小时)$^{-1}$的指数分布的随机变量。令ζ为30个器件使用的总时间，求ζ超过350小时的概率。

4. 将一枚硬币连掷100次。求出现正面的次数大于60的概率。

5. 一个复杂的系统，由100个相互独立起作用的部件所组成，在整个运行期间每个部件损坏的概率为0.10，为了使整个系统起作用，至少必需有85个部件工作，求整个系统工作的概率。

6. 一个复杂的系统，由n个相互独立起作用的部件所组成，在整个运行期间，每个部件的可靠性（即部件工作的概率）为0.90，且至少必须有80%的部件工作才能使整个系统工作。问n至少为多少才能使系统的可靠性为0.95。

7. 某个单位设置一电话总机，共有200架电话分机。设每个分机有5%的时间要使用外线通话。假设每个分机是否使用外线通话是相互独立的，问总机要按多少条外线才能以90%的概率保证在某一时刻每个分机要使用外线时可供使用？

第七章　数理统计予备知识

数理统计是研究大量随机现象的规律性的一个数学分支。它是从实际观察得到的资料来对随机变量的数字特征，分布函数等进行估计、分析和推断，来对随机变量的取值进行予报和控制等等。本章先介绍一些基本概念。

§1　总体和标本

数理统计中把所研究的对象的全体称为总体，组成总体的每一个单元称为个体。

如一批灯泡中各个灯泡的寿命就是总体，其中一个灯泡的寿命就是个体；二月三日的平均气温是总体，一九八二年二月三日的平均气温是个体；一个棉球中各根棉纤维的长度是总体，其中一根棉纤维的长度是个体。

任何一个总体都可以用一个随机变量来代表它。以后凡提到总体就是指一个随机变量，提到随机变量就是指一个总体，说总体的概率分布就是指随机变量的概率分布。

从一个总体ξ中随机地轴取n个个体ξ_1、ξ_2、…，ξ_n称$(\xi_1, \xi_2, \cdots, \xi_n)$为总体$\xi$的一个容量为n的样本。由于每个$\xi_i(i=1, 2, \cdots, n)$是从总体$\xi$中随机取得，所以每个$\xi_i$都是一个随机变量，而样本$(\xi_1, \xi_2, \cdots, \xi_n)$则是一个n维随机变量。

一次抽取的结果是n个具体数值$(x_1, x_2, \cdots, x_n)$，称为样本$(\xi_1, \xi_2, \cdots, \xi_n)$的一个观察值，简称样本值。两次不同抽取得到的样本值一般是不相同的。

在数理统计中是用样本观察值$(x_1, x_2, \cdots, x_n)$来对总体的某些特征进行估计、推断，这就需要对样本提出一些要求：

首先样本的选取必须是独立的，也就是ξ_1、ξ_2、…ξ_n必须

是相互独立的随机变量。

其次样本必须具有代表性，即要求每个ξ_i必须与总体ξ有相同的概率分布。只有这样才可能根据样本来对总体的某些特征进行估计和判断。

满足相互独立和与总体同分布这两个条件的样本称为简单随机样本，以后所提及的样本都是简单随机样本。

综上所述，所谓总体就是一个随机变量ξ，所谓样本就是相互独立且与总体同分布的随机变量ξ_i所组成的n维随机变量$(\xi_1, \xi_2, \cdots, \xi_n)$。每次抽取得到的数据是这个n维随机变量的值。因此利用样本进行统计推断，完全建立在相互独立同分布的随机变量的概率理论的基础上。如果总体ξ的概率密度为$\varphi(x)$，分布函数为$F(x)$，则样本$(\xi_1, \xi_2, \cdots, \xi_n)$的概率密度与分布函数分别为

$$\prod_{i=1}^{n} \varphi(x_i) \qquad\qquad \prod_{i=1}^{n} F(x_i)$$

其中$\varphi(x_i)$与$F(x_i)$分别是$\xi_i(i=1, 2, \cdots, n)$的概率密度和分布函数。

§2 统计量

样本$(\xi_1, \xi_2, \cdots, \xi_n)$的函数$\Phi(\xi_1, \xi_2, \cdots, \xi_n)$称为统计量。因为样本是随机变量，所以统计量也是随机变量，它们也有概率分布。下面介绍几个常用的统计量。

一、样本均值

设$(\xi_1, \xi_2, \cdots, \xi_n)$为取自正态总体$N(\mu, \sigma^2)$的样本，则样本均值

$$\bar{\xi} = \frac{1}{n} \sum_{i=1}^{n} \xi_i$$

这个统计量是服从正态分布$N(\mu,\frac{\sigma^2}{n})$的随机变量。

二、χ^2变量

设$(\xi_1,\xi_2,\cdots,\xi_n)$为取自正态总体$N(0,1)$的样本，称统计量

$$\sum_{i=1}^{n}\xi_i^2$$

是自由度为n的χ^2变量。

当$(\xi_1,\xi_2,\cdots,\xi_n)$为取自正态总体$N(\mu,\sigma^2)$的样本，则统计量

$$\frac{1}{\sigma^2}\sum_{i=1}^{n}(\xi_i-\mu)^2$$

是自由度为n的χ^2变量。而统计量

$$\frac{1}{\sigma^2}\sum_{i=1}^{n}(\xi_i-\bar{\xi})^2$$

是自由度为$n-1$的χ^2变量。

自由度可粗略的解释为平方和式中独立变量的个数。

$$\sum_{i=1}^{n}\xi_i^2 \quad 与 \quad \frac{1}{\sigma^2}\sum_{j=1}^{n}(\xi_i-\mu)^2$$

中独立变量个数都为n，故自由度为n，而

$$\frac{1}{\sigma^2}\sum_{i=1}^{n}(\xi_i-\bar{\xi})^2$$

中因为有一个约束$\bar{\varepsilon}=\frac{1}{n}\sum_{j=1}^{n}\varepsilon_i$，所以平方和式中独立变量的个数减少1，故自由度为n－1。

χ^2变量是一个随机变量，自由度为m的χ^2变量记作$\chi^2(m)$，其概率密度为

$$\varphi(x)=\begin{cases}\dfrac{1}{2^{\frac{m}{2}}\Gamma\left(\frac{m}{2}\right)}x^{\frac{m}{2}-1}\cdot e^{-\frac{x}{2}} & x>0, \\ 0 & x\leq 0\end{cases}$$

下图画出了m取不同值时χ^2变量的概率密度曲线。

三、t变量

设$(\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n)$为取自正态总体$N(\mu,\sigma^2)$的样本，称统计量

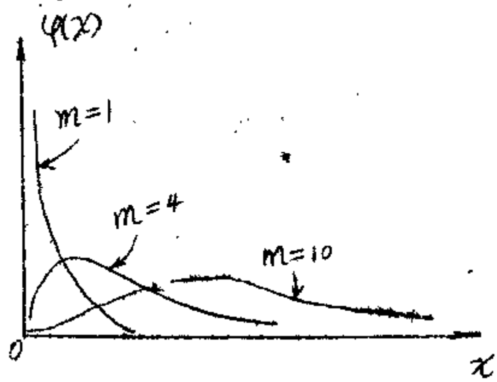

图22

$$\frac{\bar{\varepsilon}-\mu}{\sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(\varepsilon_i-\bar{\varepsilon})^2\cdot\frac{1}{n}}}$$

是自由度为n－1的t变量。

设$(\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n)$为取自正态总体$N(\mu_1,\sigma^2)$的样本，$(\eta_1,\eta_2,\cdots,\eta_m)$为取自正态总体$N(\mu_2,\sigma^2)$的样本。这两个样本相互独立，则统计量

$$\frac{(\bar{\xi}-\bar{\eta})-(\mu_1-\mu_2)}{\sqrt{\frac{1}{n+m-2}\left[\sum_{i=1}^{n}(\xi_i-\bar{\xi})^2+\sum_{j=1}^{m}(\eta_j-\bar{\eta})^2\right]\left(\frac{1}{n}+\frac{1}{m}\right)}}$$

是自由度为$n+m-2$的t变量。

t变量是随机变量，自由度为K的t变量记为t(K)，其概率密度为

$$\varphi(x)=\frac{\Gamma\left(\frac{K+1}{2}\right)}{\sqrt{K\pi}\;\Gamma\left(\frac{K}{2}\right)}\cdot\left(1+\frac{x^2}{K}\right)^{-\frac{K+1}{2}}\qquad -\infty<x<+\infty$$

$\varphi(x)$的图形关于$x=0$对称，形状类似于正态变量的概率密度曲线。

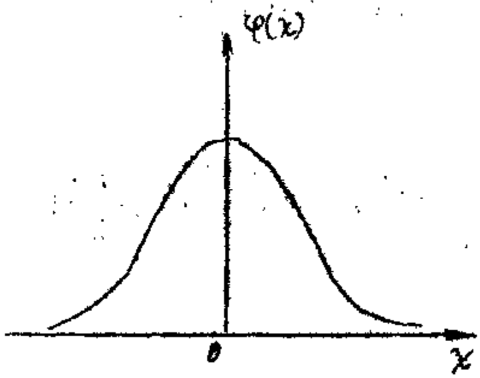

图 28

当$n\to\infty$时，利用Γ函数的斯特林公式可得

$$\lim_{n\to\infty}\varphi(x)=\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}$$

故当n很大时，t分布近似于N（0，1），但对于小的n，t分布与N（0，1）相差就很大。

四、F变量

设（ξ_1，ξ_2，…，ξ_n）为取自正态总体N（μ_1，σ_1^2）的样本，（η_1，η_2，…，η_m）为取自正态总体N（μ_2，σ_2^2）的样本，且两个样相互独立，称统计量

$$\frac{\dfrac{1}{n-1}\sum\limits_{i=1}^{n}(\xi_i-\bar{\xi})^2/\sigma_1^2}{\dfrac{1}{m-1}\sum\limits_{j=1}^{m}(\eta_j-\bar{\eta})^2/\sigma_2^2}$$

是自由度为（n－1，m－1）的F变量，其中n－1称为第一个自由度，m－1称为第二个自由度。

F变量是随机变量，自由度为（a，b）的F变量记作F（a，b），其概率密度为

$$\varphi(x)=\begin{cases}\dfrac{\Gamma\left(\dfrac{a+b}{2}\right)}{\Gamma\left(\dfrac{a}{2}\right)\cdot\Gamma\left(\dfrac{b}{2}\right)}\cdot\left(\dfrac{a}{b}\right)^{\frac{a}{2}}\cdot x^{\frac{a}{2}-1}\cdot\left(1+\dfrac{a}{b}x\right)^{-\frac{a+b}{2}}, & x>0\\ 0 & x\leq 0\end{cases}$$