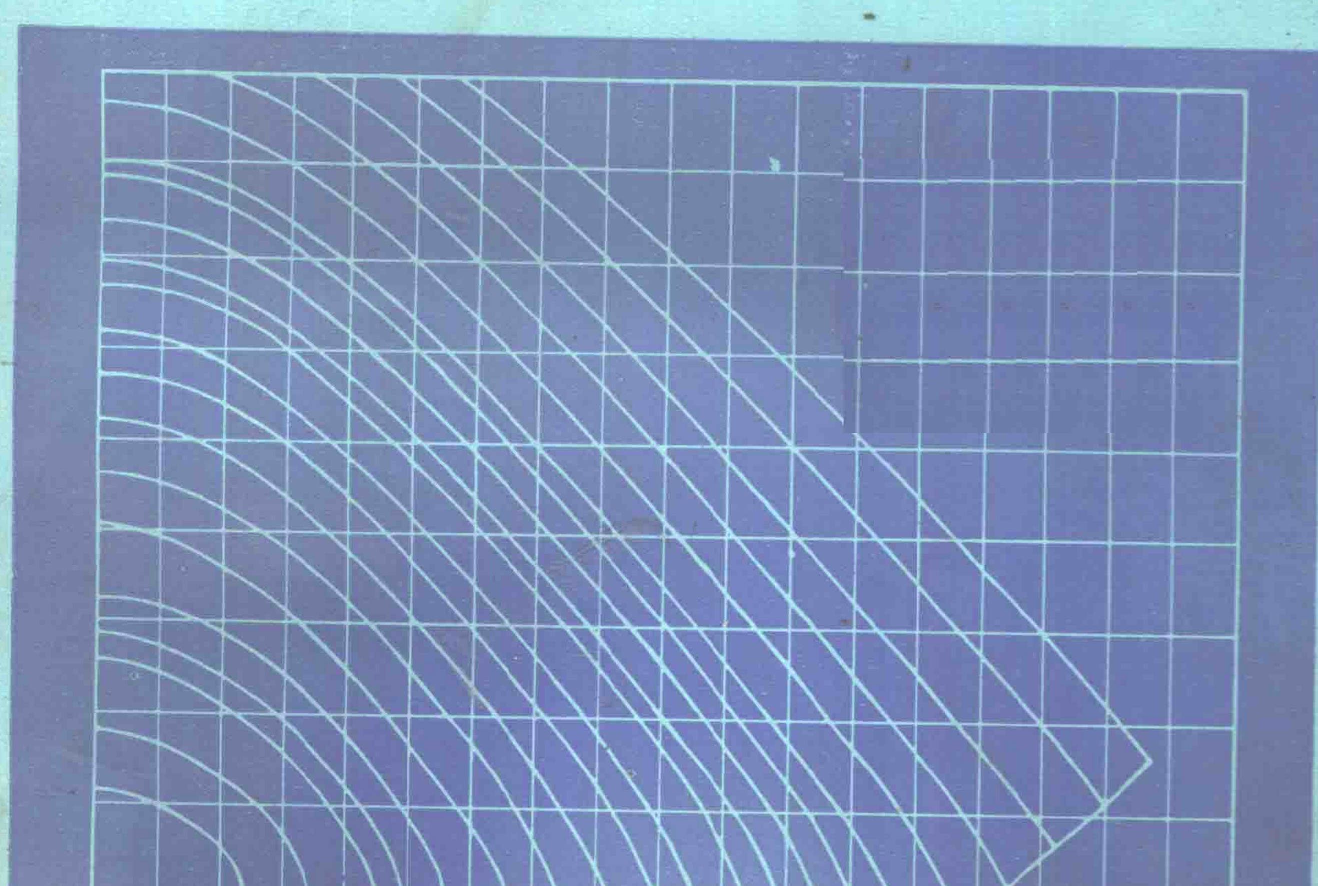
# ISO数理统计方法标准译文集

四川省环境保护科研监测所



四川科学技术出版社

# ＩＳＯ数理统计方法标准译文集

四川省环境保护科研监测所

四 川 科 学 技 术 出 版 社

一九八四年·成都

责任编辑：罗孝昌　崔泽海

封面设计：邱云松

**ISO数理统计方法标准译文集**

四川科学技术出版社出版　　成都盐道街三号
四川省新华书店发行　　四川教育学院印刷厂印刷

开本787×1092毫米1/16　印张27　插页4　字数643千
1984年5月第一版　　1984年5月第一次印刷
印数：1—2,000册

书号：15298.23　　定价：5.84元

# 前　言

ISO（The International Organization for Standardization，国际标准化组织）是一个世界性的各国标准化机构（集体会员）的联合组织，通过ISO技术委员会开展研制国际标准的工作。凡技术委员会安排的又为集体会员感兴趣的课题，每个集体会员都有权向该委员会申报。与ISO有联系的政府的或民间的国际组织，也参加了部分工作。

技术委员会接到的国际标准草案，在被ISO会议采纳作为正式国际标准之前，应先向各集体会员通报，征求意见。

TC 69（Technical Committees 69，统计方法应用技术委员会）的工作任务是：统计学术语、符号和试验及检验结果的提出与解释的标准化；工农业产品和工艺过程规范与检验中统计方法应用条件的标准化。TC 69工作的推广促进了统计方法在世界各国各领域内的应用，而随着统计方法的广泛应用又给TC 69提出了越来越多的研究课题。

国务院于1982年12月25日批转《国家标准局关于加强标准化工作报告》通知指出：随着国民经济建设和科学技术的发展，国际贸易的扩大，采用国际标准是我国的技术经济政策。在四川省城乡建设环境保护厅领导的重视下，我所针对科研课题和业务技术工作的需要，并考虑到统计方法在科学技术领域、工农业生产、环境保护和国民经济各个部门的广泛应用，决定由技术情报室负责收集和组织编译了《ISO数理统计方法标准译文集》。本译文集所收集的11个标准是由ISO/TC 69制订的。ISO/DIS 5479正态性检验是一个投票未通过的标准草案，未通过的原因不是方法本身的问题，而是这个方法是否需要作为一个标准的争论。由于正态性假设是许多统计方法应用的前提，所以我们仍将该标准草案译出供读者参考。

这11个标准是先后研制的，由不同的会员国编写；各标准的叙述程序和章节的编码有些差异。现将它们的译文汇编成册，在编辑过程中将内容相关的各标准放在一起，每个标准作为一章，章节和图表的编码作了适当改变。

本书是集体劳动的成果，参加翻译和校对工作的除我所科技人员外还邀请了部分高等院校、科研单位的教授、专家参加工作。译、校者计有：孙顺华、黄斌、潘大健、林兆升、李智果、谢清成、胡师度、李隆缘、罗孟华、黄德才、伍登祥、白苏华、梁晓禽、薛佩珠、何光宗、段梦生、刘天厚、马伯宜、刘荣模、付维周、高世雄等，并将其姓名署于每篇译文末尾。全书由林兆升工程师校对；敖硕昌教授审阅，定稿。黄时达、江伟铿、崔莲溪、詹永东描图。在此对参加这项工作的全体同志，表示感谢。

由于我们水平有限，书中难免存在不少缺点错误，请读者批评，指正。

四川省环境保护科研监测所

一九八三年八月

# 目　　录

# 一、统计学——术语和符号

**目的与应用范围**

本国际标准用英文和法文给出了统计术语的定义，这些术语在起草其他国际标准时可以使用。此外，还规定了少量术语的符号。

术语依下列大标题分成四类：

1. 概率论中使用的术语
2. 一般的统计术语
3. 与抽样法有关的一般术语
4. 与抽样检查有关的术语

## §1 概率论中使用的术语

许多术语在本节和下节“一般统计术语”里都有。分别在两节加以定义，对于提醒读者注意以下两点似乎是有益的：

a） 概率意义的术语系应用于原理与任何实际应用无关；

b） 统计意义的术语系用于与之有关的观测值，这些定义有特殊的运算特征。

**关于概率概念的注释**

可以用两种方式来介绍概率的概念，一种方式是用它表示可信程度，另一种是把它视为频率的极限。在这两种方式中，概念的介绍都要涉及国际标准的内容中不能阐述的概念，请读者参阅有关的专著。

然而，为实用计，可以认为：只要试验条件可以重复实现，事件E发生的概率Pr（E)是这样的值，事件出现的频率围绕它摆动，而且当试验的次数无限增加时，事件E出现的频率趋近于它。

**1.1 随机变量；变量：** 一个可以取一特定数集中任何值的，而且与之相应有一概率分布（见1.2）的变量。

只取孤立值的随机变量称为“离散的”。可以取有限区间或无穷区间的所有值的随机变量称为“连续的”。

**1.2 随机变量的概率分布：** 一个确定随机变量取任意给定值或属于给定数集的概率的函数。

随机变量取值于整个变域的概率为1。

**1.3分布函数**：是一个函数，对于x的每一个值，它给定随机变量X小于或等于x的概率，即

$$F(x)=Pr[X\leqslant x]$$

**1.4 连续随机变量的概率密度函数**：分布函数的导数（当它存在时）。

$$f(x)=\frac{dF(x)}{dx}$$

注：f(x)dx是"概率元素"；$f(x)dx=p_r[x<X<x+dx]$

**1.5 离散随机变量的概率**：设$x_i$是离散随机变量X所取的一个值，则概率$p_i$是：

$$p_i=Pr[X=x_i]$$

**1.6 二维分布**：这分布确定一对随机变量取任一对给定值，或属于一给定值集的概率。

**1.7 多维分布**：这分布确定同时考虑的几个随机变量取任意给定诸值或属于给定值集的概率。

**1.8 边缘分布**：在k个随机变量的概率分布情况，在（k—p）个随机变量取它们的变域区间的任何值时，p个随机变量的子集合的分布。

实例：在有三个随机变量X、Y、Z的概率分布中：有三个二维边缘分布；它们是变量对（X、Y），（X、Z），（Y、Z）的分布；

还有三个一维边缘分布：X、Y和Z的分布。

**1.9 条件分布**：在k个随机变量的概率分布情况，当(k—p)个随机变量有固定值，另p个随机变量的子集合的分布。

实例：在两个随机变量X，Y的概率分布中，有：

X的条件分布，表为"当Y=y，X的分布"的一个特殊分布。

Y的条件分布，表为"当X=x，Y的分布"的一个特殊分布。

**1.10 相关**：两个或几个变量间的相互依赖关系，在这一关系之中包含有随机部分。

注：在这种情况下，称随机变量之间存在着"随机连接"。

**1.11 概率分布的分位数（或分位点）**：设p是0和1之间的数，p分位数是随机变量的一个值，它使得分布函数在该点的值等于p或是从小于或等于p的值"跳"到大于p的值。

分布函数在随机变量的两个相邻可能值间的整个区间中有可能都等于p，这种情况下，间中的任一值都可以作为p分位数。

**1.12 中位数**：概率分布的0.50分位数。

**1.13 众数**：随机变量的值（或诸值），它使得概率（离散变量）或密度（连续变量）在该点（或诸点）有极大值。

若只有一个众数，则变量的概率分布称为"单峰的"；有多个众数时称为"多峰的"（两个众数时称为双峰的）。

**1.14 随机变量的数学期望（平均值）**：

a) 对一个离散的以概率$p_i$取值$x_i$的随机变量X，其期望定义为

$$E(X)=\sum p_i x_i$$

式中求和遍历X的所有可能取的值$x_i$。

b) 对一个连续的具有密度f(x)的随机变量X，其期望定义为

$$E(X)=\int x f(x)dx$$

式中，积分是对X的变域区间取的。

随机变量的数学期望也就是它的概率分布的数学期望。

注：1 边缘期望：随机变量的边缘分布的数学期望。
2 条件期望：随机变量的条件分布的数学期望。

**1.15 中心随机变量**：数学期望等于零的随机变量。

注：如果随机变量X的数学期望为E（X），则相应的中心随机变量是X—E（X）。

**1.16 方差**（随机变量的，或概率分布的）：中心变量平方的数学期望（见1.22）。

**1.17 标准差**（随机变量的，或概率分布的）：方差的正平方根。

**1.18 变异系数**（随机变量的，或概率分布的）：标准差对数学期望的绝对值的比。

**1.19 标准化变量**：数学期望等于零，标准差等于1的随机变量。

注：1 若随机变量X的数学期望等于E（X），标准差等于σ，则相应的标准化随机变量是随机变量

$$\frac{X-E(X)}{\sigma}$$

标准化随机变量的分布称为“标准分布”。

2 标准化随机变量的概念可以推广为“化约随机变量”的概念，化约变量是用别的位置参数和别的尺度参数来定义的。

**1.20 q 阶原点矩**：1) 在一维分布，随机变量X的q次幂的数学期望

$$E[X^q]$$

注：一阶原点矩就是变量X的数学期望。

**1.21 对点a的q阶矩**：1) 在一维分布，随机变量（X－a）的q次幂的数学期望

$$E[(X-a)^q]$$

**1.22 q 阶中心矩**：1) 在一维分布，中心随机变量〔X－E(X)〕的q次幂的数学期望

$$E\{[X-E(X)]^q\}$$

注：二阶中心矩就是随机变量X的方差。

**1.23 q阶和s阶联合原点矩**：1) 在二维分布，随机变量X的q次幂与随机变量Y的s次幂的乘积的数学期望

$$E[X^qY^s]$$

注：一阶和零阶矩的联合原点矩是X的边缘分布的数学期望。零阶和一阶矩的联合原点矩是Y的边缘分布的数学期望。

**1.24 对点（a、b）的q阶和s阶联合矩**：1) 在二维分布，随机变量（X－a）的q次幂与随机变量（Y－b）的s次幂的乘积的数学期望

$$E[(X-a)^q(Y-b)^s]$$

**1.25 q阶和s阶的联合中心矩**：1) 在二维分布，中心随机变量〔X－E(X)〕的q次幂与中心随机变量〔Y－E(Y)〕的s次幂的乘积的数学期望

$$E\{[X-E(X)]^q[Y-E(Y)]^s\}$$

注：二阶和零阶的联合中心矩是X的边缘分布的方差。零阶和二阶的联合中心矩是Y的边缘分布的方差。

---

1) 在矩的定义中，若X，X—a，Y，Y—b等量分别用它们的绝对值，即|X|，|X—a|，|Y|，|Y—b|等代替，则定义的矩称为绝对矩。

*原文中1.30条“不包含零”这一限制是不必要的—总审注

**1.26　协方差**：一阶和一阶的联合中心矩

$$E\{[X-E(X)][Y-E(Y)]\}$$

**1.27　相关系数**：两个随机变量的协方差与它们的标准差的乘积之比。

**1.28　回归曲线**：在两个随机变量的情况，Y对X＝x的数学期望，当x变动时所形成的曲线称为"Y对X的回归曲线"。

当Y对X的回归曲线是一条直线时，回归称为"线性的"。在这种情况，Y对X的线性回归的系数是回归直线方程里x的系数（斜率）。

**1.29　回归曲面**：在三个随机变量的情况下，Z对X＝x和Y＝y的数学期望，当x和y变动时所形成的曲面，称为"Z对X和Y的回归曲面"。

当回归曲面是一平面时，回归称为"线性的"。在这在种情况，Z对X的线性偏回归的系数是回归平面方程里x的系数。

注：上面的定义可以推广到三个以上的随机变量。

**1.30　一致分布；矩形分布**：在不包含零*的有限区间内概率密度函数为常数的连续随机变量的概率分布。

**1.31　正态分布；拉普拉斯—高斯分布**：连续随机变量X的概率分布，若x是任一实数，其概率密度是

$$f(x)=\frac{1}{\sigma\sqrt{2\pi}}\exp\left[-\frac{1}{2}\left(\frac{x-m}{\sigma}\right)^2\right]$$

$$-\infty<x<+\infty$$

注：m是正态分布的数学期望值，σ是正态分布的标准差。

**1.32　标准正态分布**：标准正态变量的概率分布。

以m和σ作参数的正态变量X，其标准化随机变量是

$$U=\frac{X-m}{\sigma}$$

随机变量U的概率密度是

$$f(u)=\frac{1}{\sqrt{2\pi}}\exp\left(-\frac{u^2}{2}\right)$$

$$-\infty<u<+\infty$$

**1.33　$\chi^2$分布**：独立的标准正态变量平方和的分布。

这些变量的个数ν是$\chi^2$分布变量的自由度，它是分布的一个参数。

$\chi^2$分布变量的概率密度函数是

$$f(\chi^2,\nu)=\frac{(\chi^2)^{(\nu/2)-1}}{2^{\nu/2}\Gamma\left(\frac{\nu}{2}\right)}\exp\left(-\frac{\chi^2}{2}\right)$$

$$\chi^2\geqslant 0$$

注：随机变量$\frac{\chi^2}{2}$的概率分布是参数$m=\frac{\nu}{2}$的Γ分布（见1.38）。

**1.34 t 分布；学生分布**：独立随机变量的商的分布，它的分子是标准化的正态变量、它的分母是 $\chi^2$ 分布变量与它的自由度的商的正平方根。

$\chi^2$ 的自由度就是 t 分布变量的自由度 $\nu$。t 分布变量的概率密度函数是

$$f(t,\nu)=\frac{1}{\sqrt{\pi\nu}}\ \frac{\Gamma[(\nu+1)/2]}{\Gamma(\nu/2)}\ \frac{1}{(1+t^2/\nu)^{(\nu+1)/2}}$$

$$-\infty<t<+\infty$$

**1.35 F 分布**：两个独立 $\chi^2$ 分布变量各除以它的自由度后的商的分布，分子的 $\chi^2$ 分布变量的自由度 $\nu_1$ 和分母的 $\nu_2$ 依次为 F 分布变量的第一和第二自由度数。

注：F分布变量的概率密度函数是

$$f(F,\nu_1,\nu_2)=\frac{\Gamma[(\nu_1+\nu_2)/2]}{\Gamma(\nu_1/2)\Gamma(\nu_2/2)}(\nu_1)^{\nu_1/2}(\nu_2)^{\nu_2/2}\ \frac{F^{(\nu_1/2)-1}}{(\nu_1F+\nu_2)^{(\nu_1+\nu_2)/2}}$$

$$F\geqslant 0$$

**1.36 对数正态分布**：连续随机变量X的一种概率分布，X可以取 a 到正无穷的任一值，其概率密度为

$$f(x)=\frac{1}{(x-a)\sigma\sqrt{2\pi}}\exp\left[-\frac{1}{2}\left(\frac{\log_e(x-a)-m}{\sigma}\right)^2\right]$$

$$x>a$$

m 和 $\sigma$ 分别是 $\log_e(X-a)$ 的平均值和标准差。

注： 1 随机变量 $\log_e(X-a)$ 的概率分布是正态分布；m和 $\sigma$ 分别是这变量的数学期望和标准差。

2 经常用 $\log_{10}$ 而不用 $\log_e$。在这种情况

$$f(x)=\frac{0.4343}{(x-a)\sigma\sqrt{2\pi}}\exp\left[-\frac{1}{2}\left(\frac{\log_{10}(x-a)-m}{\sigma}\right)^2\right]$$

m和 $\sigma$ 分别是 $\log_{10}(X-a)$ 的平均值和标准差。

**1.37 指数分布**：连续随机变量X的一种概率分布，X可以取 0 到 $+\infty$ 区间的任意值，其分布函数是

$$F(X)=1-e^{-\lambda x},\qquad \lambda>0$$

把 x 换成 $x-a$ $(x\geqslant a)$ 可以推广这个概率分布。

**1.38 Γ分布**：连续随机变量X的一种概率分布，X可以取 0 到 $+\infty$ 区间的任意值，其概率密度为

$$f(x)=\frac{e^{-x}x^{m-1}}{\Gamma(m)}$$

式中 $m>0$，$x\geqslant 0$，$\Gamma(m)=\int_0^\infty e^{-x}x^{m-1}dx$

m 是正常数，确定分布形式。当 m 是整数时，有 $\Gamma(m)=(m-1)!$

当 $m=1$ 时，Γ分布成为指数分布。

把x换成（$x-a$）/b，（$x\geqslant a$，$b>0$）可以推广这概率分布。

**1.39 贡贝尔(Gumbel)分布（Ⅰ型极值分布）**：连续随机变量X的一种概率分布，它的分布函数是

$$F(x)=\exp(-e^{-y})$$

式中 $y=(x-a)/b$，$b>0$，$-\infty<x<+\infty$；

**1.40 弗雷谢(Fréchet)分布（Ⅱ型极值分布）**：连续随机变量X的一种概率分布，它的分布函数是

$$F(x)=\exp(-y^{-k})$$

式中 $y=(x-a)/b$，$b>0$，$x\geqslant a$；k是一正常数，确定分布的形状。

**1.41 威布尔(Weilbull)分布（Ⅲ型极值分布）**：连续随机变量X的一种概率分布，它的分布函数是

$$F(x)=1-\exp(-y^{k})$$

式中 $y=(x-a)/b$，$b>0$，$x\geqslant a$；k是一正常数，确定分布的形状。

**1.42 二项分布**：离散随机变量的一种概率分布，若 $X\in\{0, 1, 2\cdots, n\}$，则

$$\Pr[X=x]=\frac{n!}{x!(n-x)!}p^{x}(1-p)^{n-x}$$

其中 $0<p<1$

**1.43 负二项分布**：离散随机变量X的一种概率分布，若 $X\in\{0, 1, 2\cdots\}$，则

$$\Pr[X=x]=\frac{c(c+1)\cdots(c+x-1)}{x!}p^{c}(1-p)^{x-c}$$

式中 $c>0$，$0<p<1$

注："负二项分布"这个名称是这样得来的：对x=0，1，2，…的各概率是由负指数－c的二项式，$p^{c}\{1-(1-p)\}^{-c}$，按1—p的正整次幂展开得到的。

**1.44 泊松(Poisson)分布**：离散的随机变量X的一种概率分布，若 $X\in\{0, 1, 2\cdots\}$ 则

$$\Pr[X=x]=e^{-m}\frac{m^{x}}{x!}$$

m是正参数

注：泊松分布的数学期望和方差都等于m。

**1.45 超几何分布**：若给出三个为正整数或零的N，n和d，使得下表中的数都是正整数或零。

| | | |
|---|---|---|
| N | d | N－d |
| n | x | n－x |
| N－n | d－x | N－n－d＋x |

则

$$\Pr[X=x]=\frac{n!(N-n)!d!(N-d)!}{N!x!(n-x)!(d-x)!(N-n-d+x)!}$$

**1.46 二维正态分布；二维拉普拉斯—高斯分布：** 两个连续变量X和Y的概率分布，其概率密度函数是

$$f(x,\ y)=\frac{1}{2\pi\sigma_X\sigma_Y\sqrt{1-\rho^2}}\ \exp\left\{-\frac{1}{2(1-\rho^2)}\left[\left(\frac{x-m_X}{\sigma_X}\right)^2-2\rho\left(\frac{x-m_X}{\sigma_X}\right)\left(\frac{y-m_Y}{\sigma_Y}\right)+\left(\frac{y-m_Y}{\sigma_Y}\right)^2\right]\right\}$$

$$-\infty<x<+\infty$$

$$-\infty<y<+\infty$$

注：1 $m_x$，$m_Y$是X，Y的边缘分布数学期望；$\sigma_x$，$\sigma_Y$是X，Y的边缘分布的标准差，它们是正态分布；$\rho$ 是X，Y的相关系数。

2 这概念可以推广到多于两个以上的变量的分布。

**1.47 标准化二维正态分布；标准化拉普拉斯—高斯分布：** 一对标准化正态变量的概率分布。对于参数为（$m_X, m_Y$）和（$\sigma_X, \sigma_Y$）的两个正态变量(X，Y)，相应的标准化变量是

$$U=\frac{X-m_X}{\sigma_X}\qquad\qquad V=\frac{Y-m_Y}{\sigma_Y}$$

概率密度函数是

$$f(u,\ v)=\frac{1}{2\pi\sqrt{1-\rho^2}}\ \exp\left[\frac{1}{2(1-\rho^2)}(u^2-2\rho uv+v^2)\right]$$

$$-\infty<u<+\infty,\quad -\infty<v<+\infty$$

（ρ是X，Y的相关系数，也是U，V的相关系数。）

注：这概念可以推广到两个以上的变量的分布。

**1.48 多项分布：** k个离散变量$X_1$，$X_2$，…$X_k$的一种概率分布，若$x_1$，$x_2$，…$x_k$取整数0，1，2，…n，使得它们的和$x_1+x_2+\cdots+x_k=n$

则 $$\Pr[X_1=x_1,X_2=x_2,\ldots,X_k=x_k]=\frac{n!}{x_1!x_2!\ldots x_k!}\ p_1^{x_1}p_2^{x_2}\ldots p_k^{x_k}$$

$$\rho_i\geqslant 0\ (i=1,\ 2,\ \ldots,k),\ \sum_{i=1}^{k}p_i=1$$

## §2 一般统计术语

为了避免与§1(“概率论中使用的术语”)相似的术语混淆，本节有关抽样概念一些定义可以冠以经验二字。然而，当不会含混不清时，省掉这一限定词。另外这节里提到的有关抽样的一些术语将在§3(“与抽样方法有关的一般术语”)里定义。

**2.1 单元：** a)一个实际的或约定的对象，关于这一对象可以进行一组观测。或者，b)物质

的一个确定数量，关于这个量可以进行一组观测值。或者，c）一个观测值，它可能是定性的，也可能是定量的。

注：英语术语“individual”（个体）、“unit”（单位)有时用来作为“item”（单元）的同义语。

**2.2 总体**：所考虑中的单元的全体。

总体的每一个有明确定义的部分称为子总体。在随机变量的情况，概率分布（1.2）被认为是规定了这个变量的总体。

**2.3 特征**：一种有助于区别给定总体中的单元之间的差异的性质，差异既可是定的性，也可是定量的。

**2.4 检验**：为了观测特征或将特征分类而施行的操作。

**2.5 观测值**：由观测或检验结果而确定的特征值。

**2.6 绝对差**：两个值的差的绝对值。

**2.7 极差**：定量特征的最大观测值与最小观测值的差。

**2.8 极差中值**：定量特征的最大观测值与最小观测值的算术平均值。

**2.9 组**：在定量特征的情形中，整个变域区间被分成的诸相邻的区间的每一个。

**2.10 组限**：组的上、下限。

注：应当规定两个限中的哪一个属于这个组。

**2.11 组中值**：一个组的上、下限的算术平均值。

**2.12 组距**：一个组的上、下限之间的差。

**2.13 绝对频率；容量**：总体、批、样本、组等中的单元的数目。

**2.14 累积绝对频率**：在定量特征的情形，其值小于或等于给定值，或小于等于给定的组的上限的单元的数目。

**2.15 相对频率**：一个特定值（或属于给定组内的一个值）被观测到的次数对总观测次数的比。

**2.16 累积相对频率**：在定量特征的情形，其值小于或等于给定值，或小于等于给定组的上限的单元的相对频率的和。

**2.17 频率分布**：特征值和它们的绝对频率，或相对频率之间的相互关系。

若特征值可用连续标度来度量，则分布经常用有特殊分组的表格表示。

**2.18 单变量分布**：一个变量情形下，两个方面之间的关系：一方面是特征的值、组或形式；另一方面是它们的容量或频率。

**2.19 直方图**：连续变量的频率分布的一种图形表示。在有线性标度的一条轴上以相邻区间表示组。组的绝对（或相对）频率由以区间为底的矩形来表示，矩形面积与组的绝对（或相对）频率成正比。

**2.20 条线图**：离散变量的频率分布的一种图形表示。变量的值表在有线性标度的轴上，作垂直于此轴的直线段(条线)，线段的长度与变量的绝对（或相对）频率成正比。

**2.21 累积绝对（相对）频率折线图**：连接各点所得的折线，点的横坐标是每组的上限，点的纵坐标是累积绝对频率或累积相对频率。

**2.22 二维分布**：同时考虑两个变量情况下，两个方面之间的关系：一方面是特征的值，组或形式；另一方面是它们的容量或频率。

**2.23　多维分布**：同时考虑几个变量情况下，两个方面之间的关系：一方面是特征的值，组或形式；另一方面是它们的容量或频率。

**2.24　二向表**：用来表示有两个特征的分布的一种表；在表中，行和列分别对应于第一特征的和第二特征的值或组。

在行和列交叉处出现的数或频率对应于以行列标记的组合值、组或形式。

**2.25　列联表**：在两个定性特征情形的二向表。此概念可以推广到两个特征以上的情形。

**2.26　边缘分布**：在两个变量的分布情形，忽视一个变量时，另一个变量的分布。就所考虑的变量的每一值或组，把另一变量的对应的绝对（或相对）频率加起来就可得到边缘分布。

边缘分布的概念可以推广到两个以上的变量。在这种情况，边缘分布是单元变量分布或多元变量分布，其他的变量因总计而消逝。

**2.27　条件分布**：在两个变量的分布情况，当一个变量固定时，另一个变量的分布。对于每一个变量，另一个变量有多少值或组，它就有多少条件分布。

条件分布是由指明它所涉及的变量，而另一个变量的值或组固定来表示。例如：给定$X=x$，变量Y的条件分布。

绝对频率的条件分布由在频率的二向表的行和列里直接读出。相对频率的条件分布由每一行（或列）的诸数除以该行（或列）的总数而得。

条件分布的概念可以推广到变量多于两个以上的情况。条件分布是当另一个变量或一些变量是已给定时，其它一个或多个变量的分布。

**2.28　算术平均**：诸值的总和除以它们的个数。

**2.29　（算术）加权平均**：每一个值赋以一个称为权的非负系数，每个值与它的权的乘积的和除以权的和。

**2.30　中位数**：把n个值从小到大排列，并由1到n编号。若n是奇数，这些值的中位数是第$\frac{n+1}{2}$个；若n是偶数，中位数介于第$\frac{n}{2}$个与第$\frac{n+1}{2}$个之间，不是唯一确定的，除非另有规定中位数可取为这两个值的算术平均。

**2.31　众数**：变量的，以最大绝对（或相对）频率出现的值。

分布可以具有几个局部众数，在这种情况下，分布称为多峰分布。

**2.32　平均偏差**：对某一起点的所有偏差的绝对值的算术平均值。

注：尽管取中位数作起原点时，平均偏差为极小，但是一般都选译算术平均值为起点。

**2.33　方差**：分散性的一个度量，它是以对算术平均的均方偏差为基础的。根据所考虑的情况，用偏差数目或偏差数目减1来除对算术平均的诸偏差的平方和可得方差。

因此，对于n个观测值$x_1$，$x_2$，…，$x_n$，有平均值$\bar{x}=\frac{1}{n}\Sigma x_i$，可用表达式$\frac{1}{n}\Sigma(x_i-\bar{x})^2$或$\frac{1}{n-1}\Sigma(x_i-\bar{x})^2$中的一个。

重要的是要明确采用哪一个定义。

注：一般，用$s^2$表示第二个表达式（见2.38和2.58）。

**2.34　标准差**：方差的正平方根。

**2.35　变异系数**：标准差与算术平均的绝对值之比。

注：比可以用百分数表示。

**2.36　q 阶原点矩**：在一个变量的分布，观测值的 q 次幂的算术平均

$$\frac{1}{n}\sum_i n_i x_i^q$$

这里$n_i$是$x_i$的观测次数；$n=\sum_i n_i$是观测总数。

注：一阶矩是观测值的算术平均。

**2.37　对a的q阶矩**：在一个变量的分布，观测值和值 a 的差的 q 次幂的算术平均

$$\frac{1}{n}\sum_i n_i(x_i-a)^q$$

这里$n_i$是$x_i$的观测次数；$n=\sum_i n_i$是观测总数。

**2.38　q 阶中心矩**：在一个变量的分布，观测值与其平均值$\overline{x}$的差的 q 次幂的算术平均

$$\frac{1}{n}\sum_i n_i(x_i-\overline{x})^q$$

这里$n_i$是$x_{ji}$值的观测次数，$n=\sum_i n_i$是观测的总数。

注：一阶中心矩等于零，二阶中心矩是观测值的方差。

**2.39　q 阶和 s 阶的原点联合矩**：在两个变量的分布，对构成变量X，Y的分布的所有数对($x_i$，$y_j$)，$x_i$的 q 次幂与$y_j$的 s 次幂的乘积的算术平均

$$\frac{1}{n}\sum_{ij} n_{ij}x_i^q y_j^s$$

这里 $n_{ij}$ 是每数对（$x_i$，$y_j$）的观测次数；$n=\sum_{ij} n_{ij}$是观测数对的总数。

注：1　q 阶和 S 阶联合矩是（q+S）阶矩之一。

2　一阶和零阶联合矩是X的边缘分布的平均值，零阶和一阶联合矩是Y的边缘分布的平均值。

**2.40　对（a、b）的q阶和s阶的联合矩**：在两个变量的分布，构成变量X、Y的分布的所有数对（$x_i$，$y_j$），$x_i$与 a 的差的 q 次幂和$y_j$与 b 的差的 s 次幂的乘积的算术平均

$$\frac{1}{n}\sum_{ij} n_{ij}(x_i-a)^q(y_j-b)^s$$

这里 $n_{ij}$ 是数对（$x_i$，$y_j$）的观测次数；$n=\sum_{ij} n_{ij}$是观测数对的总数。

**2.41　q阶和s阶联合中心矩**：在两个变量的分布，对构成的变量X、Y的分布的所有数对（$x_i$，$y_j$），$x_i$与平均值$\overline{x}$的差的 q 次幂与 $y_j$ 与平均值$\overline{y}$的差的 s次幂的乘积的算术平均

$$\frac{1}{n}\sum_{ij} n_{ij}(x_i-\overline{x})^q(y_j-\overline{y})^s$$