环境数理统计学

上册

卢崇飞 编

北京大学环境科学中心

1983. 9.

环境数理统计学（上册）目录

序

大量偶然现象的统计规律性

人类对大量偶然现象存在统计规律性的认识、始于中华。远在公元前两千年，已发现在大量新生婴儿中，两性存在稳定比例，约各占1／2。至十八世纪，英国人统计了伦敦1629—1710年婴儿出生的材料，也发现了"男女数目精确地维持平衡"。但却以此作为神之存在的论据。甚至说"男孩的比例略超过女孩，是因上帝为了男人可能遭到一些危险（如战争、航海等）而制订的平衡补充"。对某一家庭，生男或生女带偶然性，（当然，也可从生理学上去探究其必然原因。）但是大量的家庭却生出了不以人的意志为转移的男女各半的必然性。男女各半的必然性是通过每次或男或女的偶然性表现出来的。这就是大量偶然现象的统计规律性。这种由统计规律性所表现的必然性，已不是单个家庭的属性。而是在另一个层次上的即人类的属性。

又例如，在氧气瓶中，氧分子不停地运动、互相碰撞并冲击器壁。某分子在某时刻碰壁是偶然的，但大量分子的同时存在却表现出压力的稳定性。压力及其稳定性已不是单个分子的属性，而是在另一个层次上的即大量分子总体运动所具有的统计规律性。

再例如，在一定生产条件下，抽查一个产品，它是正品还是废品带偶然性，但是大量的抽查则会发现废品数具有稳定的百分比，废品率的稳定性已不是单个产品的属性，而是大量产品所具有的统计规律性，这是在另一个层次上的即整个生产水平的表现。

上述"另一个层次"，实际上是"更上一层楼"，是在高一级层次上体现必然与偶然的辩证关系。偶然性根源于宇宙间一切事物

相互联系相互制约的无限多样性。某时某地某点的SO_2浓度，为什么是这样而不是那样，其全部原因是不可穷尽的。一个SO_2分子从污染源出发，经大气搬运、湍流扩散，途中不知经过了多少不同尺度的涡旋的携带，也不知碰撞多少亿万万次，它是否正好落入取样器，是偶然的。但它在运动过程中的每一瞬间的每一步，可以说又都符合必然性的力学规律。这些无穷无尽的必然性的交叉，就构成了事先无法予断的偶然性。跟踪追究每个分子在每个瞬时的必然运动规律是无法实现的繁琐哲学。拉普拉斯说过："假如有某一个无所不知的大天才，能够写出适合于宇宙在给定时刻状态的所有方程，并且把它积分出来，那么他就能够完全准确地预见宙宇在无穷时间过程中全部演化。"这等于说，偶然性是因为无知。正如恩格斯所说的："他是用根本否认偶然性的办法来对付偶然性。"属于机械论的自然观，以为了解了每个局部也就等于了解了整体，而没有认识到在高一级的层次上有新质。统计规律性就是这种新质正如人并不等于人身上所有原子的总和，人有思想，思想已不是各种单个原子的属性。难怪拉普拉斯，这位古典概率的奠基人，由于时代的局限性，他研究了统计规律性，却不认识其重大意义，因为古典概率论形成的时代正是经典力学走向全盛的时代，强调初始条件决定了整个运动。当时整个自然科学界正被机械论所统治。到十九、二十世纪，机械论的自然观被打破了。恩格斯说达尔文的进行论"是从最广泛地存在着的偶然性的基础出发的"，"偶然性推翻了至今人们所理解的必然性，必然性原有的概念失效了"。动摇机械论自然观最有力的科学进展，发生在这一时期的物理学中，玻尔兹曼和吉布斯把统计规律性引入了物理学，将气体分子运动论与热

力学规律联系起来，建立了统计力学。玻尔兹曼从理论上证明了不可逆方向的熵函数和状态概率的自然对数成正比，揭露了热力学运动和机械运动的本质差别。"层次"的概念引进了物理学。这是在一门学科领域内突破机械决定论的束缚，运用了整体的辩证的认识方法，它揭示出，对于组成系统的大量分子中任何一个的行为作出详尽的推断是不可能的，但大量分子的同时存在却表现出系统总体性质的十分确定的规律，对此，传统方法已无能为力，必须引进概率统计方法。因此，很多自然科学家称之为"自然定律中新质方向的第一例"，"一种可贵的进步倾向"。控制论的创始人维纳甚至认为"必须把二十世纪物理学的第一次大革命归功于吉布斯，而不是归功于爱因斯坦、海森伯或普朗克"，维纳的提法，未免有过誉之嫌，但他对概率统计在科学认识史上的地位之评价，还是中肯的，他说"概率就不仅作为物理学的数学工具，而且作为物理学的部分经纬，被人们接受来了。"

概率论与数理统计，作为一门数学，就是从量的方面去研究必然与偶然的关系，研究大量偶然现象的统计规律性。经十七世纪中叶起步，到现在它的诸分支极其迅速发展，可以说它不仅作为一门数学工具，而且作为部分经纬广泛渗透自然科学、社会科学和工农业生产的众多领域。特别对于新兴的环境科学，更是前途广阔，概率统计大有用武之地。这是因为环境科学本身就是多种学科综合的新的"层次"。恩格斯说："人类对自然界的每次征服，自然界都反过来报复了我们"。由人类干预物质循环；特别是大规模现代工业的发展，引起物质分布的改变，造成有害物质的积聚即污染，导致原有生态系统的急剧变化甚至崩溃。为此，人类又发展环境科学

来进行反报复。这是一个作用、反作用、反反作用……极其复杂的过程，不仅涉及自然科学的众多领域，而且也是社会科学问题。因此环境科学所研究的不仅是化学、物理、生物等等等等多种学科的特定主题，而且是新的“层次”上的大系统。在这个系统中，例如当我们研究污染之来源、形成、扩散、分布、变化转换与归宿，进行预测预报与环境质量评价，制定监测与控制措施，检验治理效果等等，经数量角度看，就是要研究环境变量之间的关系，寻求其规律，据此施加人之干预。而所谓环境变量，诸如：大气中之SO_2、NO_x、O_x、CO、TCH、飘尘等等的浓度；水土中甲基汞与镉的含量；虫、鱼生物之数量；噪声分贝数；汽车通行量与尾气量；职工发病数；以及气象因子等等等等，多因时因地而异，受多种因素的影响，其必然规律是通过偶然来表现的、数学上称之为随机变量。环境科学是涉及多种随机变量的、有输入有输出有反馈的复杂系统。要想攀登这个领域的高峰，除去各门专业知识的研究探讨外，还需要一双防滑的登山鞋，这便是研究随机变量（自身及相互关系）统计规律性的数学分支——《概率论与数学统计》，它与环境科学的结合，估且命名为：《环境统计学》。

恩格斯说：“被断定为必然的东西是由纯粹的偶然构成的，而所谓偶然的东西，是一种有必然性隐藏在里面的形式。”

预备知识（Ⅰ）　矩阵

§1　矩阵的运算

将多个数字或变量排成矩形队伍，称为矩阵。例如

$$\begin{pmatrix} 8 & 7 & 3 \\ 2 & 4 & 6 \end{pmatrix},\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix},\begin{pmatrix} 1 & x & x^2 \\ 1 & y & y^2 \\ 1 & z & z^2 \end{pmatrix}$$

概括为一般的形式如下：

$$\begin{pmatrix} a_{11} & a_{12} & \cdots\cdots & a_{1m} \\ a_{21} & a_{22} & \cdots\cdots & a_{2m} \\ \cdots\cdots & & & \\ a_{n1} & a_{n2} & \cdots\cdots & a_{nm} \end{pmatrix}$$

简记为A或$A_{n\times m}$或$(a_{ij})_{n\times m}$。其中每个元素有两个足码，第一足码表示横行数。第二足码表示纵列数，即a_{ij}表示第i行第j列的元素。A矩阵共有$n\times m$个元素。若$n=m$则称A为方阵。若A的各元素皆为零，则记$A=0$。若A为方阵，且对角线上各元素皆为1，其余皆为0，则称A为单位阵，记为$A=I_{n\times m}$，例如

$$I_{3\times 3}=\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

（一）矩阵加法

设

$$A=\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1m} \\ a_{21} & a_{22} & \cdots & a_{2m} \\ & \cdots\cdots\cdots\cdots & & \\ a_{n1} & a_{n2} & & a_{nm} \end{pmatrix} \quad B=\begin{pmatrix} b_{11} & b_{12} & \cdots & b_{1m} \\ b_{21} & b_{22} & \cdots & b_{2m} \\ & \cdots\cdots\cdots\cdots & & \\ b_{n1} & b_{n2} & & b_{nm} \end{pmatrix}$$

定义

$$A+B \begin{pmatrix} a_{11}+b_{11} & a_{12}+b_{12} & \cdots & a_{2m}+b_{2m} \\ a_{21}+b_{21} & a_{22}+b_{22} & \cdots & a_{2m}+b_{2m} \\ & \cdots\cdots\cdots\cdots & & \\ a_{n1}+b_{n1} & a_{n2}+b_{n2} & \cdots & a_{nm}+b_{nm} \end{pmatrix}$$

即两矩阵的对应元素相加组成新矩阵。显然，两矩阵的行数与列数必须相同才能相加。易证矩阵加法满足

结合律：　　$(A+B)+C=A+(B+C)$

交换律：　　$A+B=B+A$

（二）矩阵乘法

设

$$A=\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1m} \\ a_{21} & a_{22} & \cdots & a_{2m} \\ & \cdots\cdots & & \\ a_{n1} & a_{n2} & & a_{nm} \end{pmatrix}, \quad B=\begin{pmatrix} b_{11} & b_{12} & \cdots & b_{1p} \\ b_{21} & b_{22} & \cdots & b_{2p} \\ & \cdots\cdots & & \\ b_{m1} & b_{m2} & \cdots & b_{mp} \end{pmatrix}$$

定义

$$A+B=C=\begin{pmatrix} c_{11} & c_{12} & \cdots & c_{1P} \\ c_{21} & c_{22} & \cdots & c_{2P} \\ \cdots & \cdots & \cdots & \cdots \\ c_{n1} & c_{n2} & \cdots & c_{nP} \end{pmatrix}$$

其中 $c_{ij}=\sum_{k=1}^{m} a_{ik} b_{kj}$，$i=1,2,\cdots n; j=1,2,\cdots P$

即 乘积矩阵C的第i行第j列的元素等于矩阵A的第i行与矩阵B的第j列（皆为m个元素)依次对应的元素相乘再相加。

例1

$$\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix}\begin{pmatrix} 1 & 2 & 3 \\ 2 & 4 & 2 \\ 3 & 1 & 2 \end{pmatrix}=\begin{pmatrix} 14 & 12 & 13 \\ 32 & 34 & 34 \end{pmatrix}$$

例2

$$\begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}\begin{pmatrix} 4 & 5 & 6 \end{pmatrix}=\begin{pmatrix} 4 & 5 & 6 \\ 8 & 10 & 12 \\ 12 & 15 & 18 \end{pmatrix}$$

例3

$$\begin{pmatrix} 1 & 2 & 3 \end{pmatrix}\begin{pmatrix} 4 \\ 5 \\ 6 \end{pmatrix}=32$$

显然，前一矩阵的列数必须等于第二矩阵的行数，否则无法相乘，由例2例3可见，矩阵乘法次序不可交换，即一般情况下，$AB \neq BA$。

易证矩阵乘法满足结合律，即矩阵A，B，C相乘有：

$$(AB)C=A(BC)$$

由矩阵乘法可定义方矩阵的幂

$$A^k = \underbrace{AAA\cdots\cdots A}_{k\text{个}A\text{相乘}}$$

由结合律得

$$A^k A^\theta = A^{k+\theta} \quad ; \quad (A^k)^\theta = A^{k\theta}$$

单位阵在矩阵乘法中的地位相当于数1在数相乘中的地位，易证：

$$\underset{m\times n}{A}\ I_n = \underset{m\times n}{A}\ , \quad I_n\ \underset{n\times m}{B} = \underset{n\times m}{B}$$

（三）数量乘法

数k乘矩阵A定义为A的每一元素皆乘以数k所组成的新矩阵，称为矩阵的数量乘法。据矩阵加法与数量乘法可定义矩阵的减法如下：

$$A-B=A+(-1)B$$

显然，设k，θ是数，A，B是矩阵，则有

$$(k+\theta)A=kA+\theta A$$

$$k(A+B)=kA+kB$$

$$k(\theta A)=(k\theta)A$$

$$k(AB)=(kA)B=A(kB)$$

（四） 矩阵的转置

设

$$A=\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ & \cdots\cdots & & \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}$$

定义 A的转置阵（记为A′）

$$A'=\begin{pmatrix} a_{11} & a_{21} & \cdots & a_{m1} \\ a_{12} & a_{22} & \cdots & a_{m2} \\ & \cdots\cdots & & \\ a_{1n} & a_{2n} & \cdots & a_{mn} \end{pmatrix}$$

例如

$$\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix}' = \begin{pmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{pmatrix}$$

易证矩阵转置运算适合以下规律

$$(A')'=A$$

$$(A+B)'=A'+B'$$

$$(AB)'=B'A'$$

上式中第三式最重要，相乘后转置等于先交换次序再转置后相乘。

〔例〕 统计中的样本协方差矩阵

对P维随机向量

$$x=\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_P \end{pmatrix}$$

独立观测N次，得N个向量

$$x^{\alpha}=\begin{pmatrix} x_{\alpha 1} \\ x_{\alpha 2} \\ \vdots \\ x_{\alpha P} \end{pmatrix} \quad \alpha=1, 2, \cdots, N$$

它们的平均向量（即每个分量皆取平均数）

$$\bar{x}=\begin{pmatrix} \bar{x}_1 \\ \bar{x}_2 \\ \vdots \\ \bar{x}_P \end{pmatrix}=\begin{pmatrix} \frac{1}{N}\sum_{\alpha=1}^{N} x_{\alpha 1} \\ \frac{1}{N}\sum_{\alpha=1}^{N} x_{\alpha 2} \\ \vdots \\ \frac{1}{N}\sum_{\alpha=1}^{N} x_{\alpha P} \end{pmatrix}$$

则样本协方差矩阵

$$S = \frac{1}{N-1} \sum_{\alpha=1}^{N} (x^{\alpha} - \bar{x})(x^{\alpha} - \bar{x})'$$

$$= \frac{1}{N-1} \sum_{\alpha=1}^{N} \begin{pmatrix} x_{\alpha 1} - \bar{x}_1 \\ x_{\alpha 2} - \bar{x}_2 \\ \vdots \\ x_{\alpha P} - \bar{x}_P \end{pmatrix} (x_{\alpha 1} - \bar{x}_1, x_{\alpha 2} - \bar{x}_2, \cdots x_{\alpha P} - \bar{x}_P)$$

$$= \left(\begin{array}{l} \frac{1}{N-1} \sum_{\alpha=1}^{N} (x_{\alpha 1} - \bar{x}_1)(x_{\alpha 1} - \bar{x}_1) \\ \frac{1}{N-1} \sum_{\alpha=1}^{N} (x_{\alpha 2} - \bar{x}\ )(x_{\alpha 1} - \bar{x}_1) \\ \cdots\cdots\cdots\cdots\cdots\cdots \\ \frac{1}{N-1} \sum_{\alpha=1}^{N} (x_{\alpha P} - \bar{x}_P)(x_{\alpha 1} - \bar{x}_1) \end{array} \right.$$

$$\frac{1}{N-1} \sum_{\alpha=1}^{N} (x_{\alpha 1} - \bar{x}_1)(x_{\alpha 2} - \bar{x}_2) \cdots\cdots$$

$$\frac{1}{N-1} \sum_{\alpha=1}^{N} (x_{\alpha 2} - \bar{x}_2)(x_{\alpha 2} - \bar{x}_2) \cdots\cdots$$

$$\cdots\cdots\cdots$$

$$\frac{1}{N-1} \sum_{\alpha=1}^{N} (x_{\alpha P} - \bar{x}_P)(x_{\alpha 2} - \bar{x}_2) \cdots\cdots$$

$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$

$$\left.\begin{matrix} \frac{1}{N-1}\sum\limits_{\alpha=1}^{N}(x_{\alpha 1}-\bar{x}_1)(x_{\alpha p}-\bar{x}_p) \\ \frac{1}{N-1}\sum\limits_{\alpha=1}^{N}(x_{\alpha 2}-\bar{x}_2)(x_{\alpha p}-\bar{x}_p) \\ \cdots\cdots\cdots\cdots \\ \frac{1}{N-1}\sum\limits_{\alpha=1}^{N}(x_{\alpha p}-\bar{x}_p)(x_{\alpha p}-\bar{x}_p) \end{matrix}\right)$$

$$\overset{记}{=\!=}\begin{pmatrix} S_{11} & S_{12} & \cdots & S_{1p} \\ S_{21} & S_{22} & \cdots & S_{2p} \\ \cdots & \cdots & \cdots & \cdots \\ S_{p1} & S_{p2} & \cdots & S_{pp} \end{pmatrix}$$

〔例〕 线性方程组

$$\begin{matrix} a_{11}x_1+a_{12}x_2+\cdots+a_{1m}x_m=b_1 \\ a_{21}x_1+a_{22}x_2+\cdots+a_{2m}x_m=b_2 \\ \cdots\cdots\cdots\cdots \\ a_{m1}x_1+a_{m2}x_2+\cdots+a_{mm}x_m=b_m \end{matrix}$$

可用矩阵乘法的形式表示如下：

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1m} \\ a_{21} & a_{22} & \cdots & a_{2m} \\ \cdots & \cdots & \cdots & \cdots \\ a_{m1} & a_{m2} & \cdots & a_{mm} \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{pmatrix}=\begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}$$