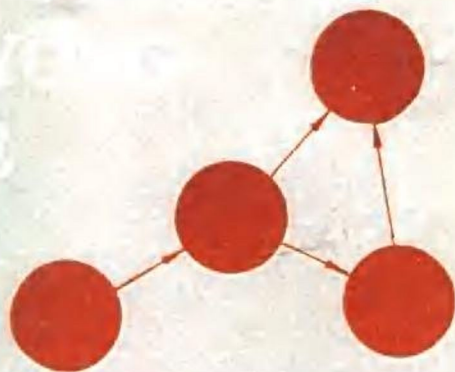


吴天滨 编

GONGCHANG
ZHILIANGLIANLI
CHANGYONG
SHUXUE



3

工厂质量管理常用数学

山东科学技术出版社

F406.3
49

工厂质量管理常用数学

吴天滨 编

100701

山东科学技术出版社

一九八三年·济南

B 036275

责任编辑 原式溶

工厂质量管理常用数学

吴天滨 编

*

山东科学技术出版社出版

山东省新华书店发行

山东新华印刷厂潍坊厂印刷

*

787×1092毫米32开本 10.625印张 224千字

1983年1月第1版 1983年1月第1次印刷

印数: 1—6,000

书号 15195·111 定价 0.90 元

前 言

在工业各部门的质量管理及有关科学试验中,数理统计等数学方法有广泛的应用。系统学习数理统计等理论知识,需要较广较深的数学基础,这是大多数搞实际工作的同志所难以做到的。另一方面,实践已证明,完全避开理论,只学习具体的做法,往往收效甚微。针对以上情况,本书对数理统计等有关理论知识,力求作深入浅出的介绍,通过实例着重叙述若干常用数学方法的具体应用,注意避免繁难的数学推导,内容简明、实用,可供广大工业科技人员和技术工人学习参考。书中缺点错误敬请读者指正。

吴 天 滨

一九八二年五月

目 录

第一章 随机抽样和数据整理	1
第一节 随机抽样	1
第二节 频率、概率和统计方法	5
第三节 大子样的频率分布和直方图	11
第四节 子样特征数	13
第五节 随机变量和概率分布	23
第二章 图解方法	38
第一节 直方图的应用	38
第二节 排队图的作法和应用	46
第三节 鱼骨图的作法和应用	50
第四节 控制图的作法和应用	52
第三章 统计推断	62
第一节 预备知识	62
第二节 u 检验法及区间估计	66
第三节 t 检验法及区间估计	78
第四节 χ^2 检验法及区间估计	84
第五节 F 检验法	91
第六节 柯氏检验法和斯氏检验法	94
第七节 符号检验法和秩和检验法	99
第四章 优选法和正交试验法	106
第一节 单因素优选法	106
第二节 正交试验法(最简单情形)	115
第三节 交互作用及其直观分析	122
第四节 水平数不同的试验	130

第五节	试验安排中的常用技巧和注意事项	133
第六节	逐次选优试验法	139
第五章	方差分析	147
第一节	单因素方差分析	147
第二节	二因素方差分析	161
第三节	正交试验的方差分析	170
第四节	逐次选优试验的方差分析	183
第六章	回归分析	188
第一节	一元线性回归问题	188
第二节	一元线性回归问题的解法	191
第三节	多元线性回归和正交表回归	202
第四节	子样相关系数及其应用	215
第五节	非线性回归	220
第六节	正交多项式	232
第七章	抽样验收法	243
第一节	概述	243
第二节	计件的一次抽样方案	245
第三节	计件的序贯抽样方案	248
第四节	计件的二次抽样方案	254
第五节	接收概率曲线	256
第六节	验收制度与平均出厂次品率	258
第八章	统筹法	261
第一节	统筹法与箭头图	261
第二节	箭头图的作法	263
第三节	统筹调整的方法	266
第四节	非肯定型(随机统筹)问题	272
第五节	利用正态分布解非肯定型问题	275
附表		282
主要参考资料		332

第一章 随机抽样和数据整理

数理统计方法的一个主要特点是由大量对象的一少部分来推断总体，这一少部分就是子样。本章介绍抽取随机子样和对子样（数据）进行初步整理的基本方法。

第一节 随 机 抽 样

一、母体、个体和子样

母体（或称总体）是指在某一次统计分析工作中所要研究的对象的全体，组成母体的每一个单位叫做一个个体。从一个总体中取出一部分个体，就构成一个子样（或称样本）。子样中含有个体的个数，叫做子样的大小（或称容量）。

【例 1】 某糖厂用自动打包机包装砂糖，每包标准重量为 100 斤，某日开工后测得 9 包重量（斤）如下：

99.3, 98.7, 100.5, 101.2, 98.3,

99.7, 99.5, 102.1, 100.5

在本例中，某日该糖厂按开工后的生产条件用自动打包机打的全部糖包（包括已打出的和将要打出的）的单包重量是一母体，每包重量是一个个体，而测得的 9 包重量是一个大小为 9 的子样。

【例 2】 某厂所产的铆钉，直径在 13.20~13.60 毫米范围内为合格。从该厂产的铆钉中随机抽取 200 个，测量其直

径得到 200 个数据，见表 1—1。

表 1—1

单位：毫米

13.39	13.43	13.54	13.64	13.40	13.55	13.40	13.26
13.42	13.50	13.32	13.31	13.28	13.52	13.45	13.63
13.38	13.44	13.52	13.53	13.37	13.33	13.24	13.13
13.53	13.53	13.39	13.57	13.51	13.34	13.39	13.47
13.51	13.48	13.62	13.58	13.37	13.33	13.51	13.40
13.30	13.48	13.40	13.57	13.51	13.40	13.52	13.56
13.40	13.34	13.23	13.37	13.48	13.48	13.62	13.35
13.28	13.59	13.47	13.46	13.62	13.54	13.20	13.38
13.43	13.35	13.56	13.51	13.47	13.40	13.29	13.20
13.46	13.44	13.42	13.29	13.41	13.39	13.50	13.48
13.53	13.34	13.45	13.42	13.29	13.38	13.45	13.50
13.55	13.33	13.32	13.69	13.46	13.32	13.32	13.48
13.29	13.25	13.44	13.60	13.43	13.51	13.43	13.38
13.24	13.58	13.31	13.31	13.31	13.45	13.43	13.44
13.34	13.49	13.50	13.38	13.48	13.43	13.37	13.29
13.54	13.33	13.36	13.46	13.23	13.44	13.38	13.27
13.66	13.26	13.40	13.52	13.59	13.48	13.46	13.40
13.43	13.26	13.50	13.38	13.43	13.34	13.41	13.24
13.42	13.55	13.37	13.41	13.38	13.14	13.42	13.52
13.38	13.54	13.30	13.18	13.32	13.46	13.39	13.35
13.34	13.37	13.50	13.61	13.42	13.32	13.35	13.40
13.57	13.31	13.40	13.36	13.28	13.58	13.58	13.38
13.26	13.37	13.28	13.39	13.32	13.20	13.43	13.34
13.33	13.33	13.31	13.45	13.39	13.45	13.41	13.45
13.40	13.36	13.45	13.48	13.29	13.58	13.44	13.56

在此例中，某厂所产全部该种规格的铆钉的直径是一母体，每个铆钉的直径是一个个体，我们得到的是一个大小为 200 的子样。

在实际应用中，通常把大小在 50 以下的子样叫做小子

样；把大小超过 50 的子样叫做大子样。例 1 是一个小子样，例 2 是一个大子样。一般小子样的大小多取在 10 左右，也就是包含 10 来个数据。取小子样比较节省人力、物力和时间，但反映总体情况不如大子样全面。在实用中，大子样主要用于分析已有的资料，小子样广泛应用于抽样检验和质量控制等在生产现场进行的统计分析。

应注意，在同一母体中可能有两个以上个体取同一数值，这时仍认为是两个以上个体。例如，在例 2 中抽得的子样中有 6 个铆钉直径为 13.31 毫米，这里应作为 6 个个体来看待。

还应注意，统计学中的个体，一般不是笼统地指某个具体事物，而是指该具体事物的某一数量指标，如糖包的重量、铆钉的直径、单块砖的抗折强度、单个灯泡的耐用时间等。

二、随机抽样法和随机子样

抽取子样的方法和过程叫做抽样。抽样的目的是通过子样来判断母体，因此要求抽得的子样具有代表性。也就是说，抽样时应不加任何主观偏向，使所有的个体处于平等地位，每一个体被抽到的可能性相同。符合上述要求的抽样法，叫做随机抽样法；用随机抽样法得到的子样叫做随机子样。下面所说的抽样，总是指随机抽样；所说的子样，也都是指随机子样。

当母体不太大而且每个个体尺寸也较小时，可以直接进行随机抽样。但在实际问题中，母体一般都是比较大的，为了保证抽样的随机性，最好利用随机数表进行抽样。

三、利用随机数表抽样

随机数表就是由随机出现的数字组成的表。常用的随机数表包括 5000 或 10000 个数字，其中每个数字都是从 0~9

这 10 个数中随机取得的。随机数表可以自己制作：把 0~9 这 10 个数分别写在 10 张同样的小纸片上，放在盒里或袋里，随机取一张，记下数码，放回去；再抽一张，记下数码，放回去……。如此进行 1000 次，顺次记下的 1000 个数字，就组成一个随机数表。附表 1 是一个包含 5000 个数字的随机数表。

下面介绍利用随机数表抽样的标准做法：先把母体中的全部个体编号，如要从 50000 件产品中随机抽取 20 件，可把产品编为 00000~49999 号。然后，在随机数表中随机决定页码与起点。例如，附表 1 共 4 页，可取 4 张纸片写上 1, 2, 3, 4，随机抽其一，若抽到 3，就用数表的第 3 页。再在第 3 页上随机投一根针，取离针尖最近的数字作为起点。例如，离针尖最近的是第 5 行第 2 个数字(1)，就从这个数字开始，依次取数码作成五位数，但碰到首位是 5, 6, 7, 8, 9 时，跳过不用。如此，从表上得 20 个五位数如下：

17538, 17831, 31089, 42983, 40230,
24617, 09609, 15253, 29772, 02019,
05653, 18760, 42774, 48445, 03875,
08363, 40649, 00429, 12139, 37914

然后，按这 20 个编号取 20 件产品。

随机数表可以灵活运用，如可以按由右向左，由下向上的顺序，也可以每隔一个数字取一个数字等等。

四、不放回抽样和放回抽样

抽样时如果抽出第 1 个不放回母体中去；再抽第 2 个也不放回去……。一直抽出 n 个，得到一个大小为 n 的子样，这种抽样叫做不放回抽样。

另一种抽样法是抽出第 1 个记下数据后放回到母体中；再抽第 2 个记下数据后放回去……。最后，抽出第 n 个，得到一个大小为 n 的子样，这种抽样叫做放回抽样。

两种抽样的区别在于：在不放回抽样中，每个个体至多被抽到 1 次，各次抽样时母体有变化(每次减少一个个体，当母体很大时该变化可以忽略不计)；在放回抽样中，每个个体有可能被重复抽到，各次抽样时母体不变。

今后若不特别声明，我们所说的抽样总是指不放回抽样。

五、数据记录

正确地测得和记录数据，是搞好统计分析工作的基础。数据记录最好表格化。一组数据就是一个子样，要把抽样的条件（来自什么母体）记录清楚。同一组数据应准确到小数点后同样的位数。例如，本节例 2 的 200 个数据都记到小数点后第二位，即准确到小数点后第一位，第二位是由四舍五入得到的，数据的误差在 $(-0.005, 0.005)$ 范围内，即误差不超过 0.005。

第二节 频率、概率和统计方法

一、随机现象、随机试验和随机事件

在自然界、生产实践、科学试验以及社会生活中，处处存在着随机现象(或称偶然现象)，随机现象可以用随机试验来描述。

我们对试验作广义的理解：所谓试验就是在某种确定条件下观察发生了什么现象。

【例 1】 使悬挂着的两个带正电的小球互相靠近（条

件),我们观察到的是两个小球互相排斥(现象)。

【例 2】 随手向桌面上掷一枚硬币(条件), 我们观察到的是有时正面向上, 有时反面向上(现象)。

象例 1 这种有唯一确定试验结果的试验, 叫做决定性试验, 它描述的是决定性现象。这种在确定条件下必定发生的事件, 叫做必然事件。

象例 2 那样试验结果不唯一的, 叫做随机试验, 它描述的是随机现象。这种在确定条件下可能发生也可能不发生的事件叫随机事件。

必然事件的反面是不可能事件。在例 1 中两小球不互相排斥就是不可能事件。容易看出, 在决定性试验中, 只有必然事件及其反面——不可能事件。而在随机试验中, 有随机事件, 同时也有必然事件和不可能事件。例 2 中或出正面或出反面就是必然事件, 不出正面也不出反面则是不可能事件。随机事件、必然事件和不可能事件, 三者统称事件。

由上述可见, 一个事件是随机的、必然的或不可能的, 是与试验条件分不开的。

还应注意, 一个试验是决定性的还是随机的, 不但要看试验条件如何, 还要看观察哪些现象。一般地说, 在同样条件下, 观察较笼统粗糙的是决定性试验, 而观察较深入细致的是随机试验。比如在例 1 中, 若观察两小球是否互相排斥, 这是决定性试验; 若更细致地观察两小球互相排斥达成的距离, 则是一个随机试验。由于许多偶然因素的影响, 重复试验几次就可发现两小球达成的距离并不完全确定, 有时大些, 有时小些。再如在例 2 中, 若观察落下的硬币是正面向上还是反面向上, 这是随机试验; 若笼统地观察硬币是否落下来,

则是一个决定性试验了(硬币必定会落下来)。

对客观事物观察的细致深入,正是科学技术精确化和深刻化的表征。因此,随着生产和科学技术的发展,必须研究随机现象。研究随机现象数量规律性的数学分支,称为概率论与数理统计,其理论和方法现在已渗透到一切自然科学和技术科学领域中,得到广泛应用。

二、频率和频率稳定性

研究随机现象的数量规律性,从考虑随机事件发生的可能性大小开始。一个熟练工人在生产中出次品是随机事件,一个初步掌握技术的工人在同样生产条件下出次品也是随机事件。这两个随机事件发生的可能性是不同的,前者小而后者大。这里“可能性”大小的确切意思是:当多次重复生产此种产品时,熟练工人出次品的百分率较少。例如,同是重复生产 1000 件产品,熟练工人出了 5 件次品,次品百分率为 0.5%;初步掌握技术的工人出了 100 件次品,次品百分率为 10%。这里的百分率在统计学中称为频率。

一般以大写字母 $A, B, C, \dots$ 表示某随机试验下的随机事件。将此试验重复进行 k 次,以 l 表示 A 在 k 次试验中的发生次数,则 l 称为 A 在 k 次试验中的频数, $\frac{l}{k}$ 称为 A 在 k 次试验中的频率。

如上所述,频率 $\frac{l}{k}$ 可以衡量随机事件发生的可能性的 大小,但 $\frac{l}{k}$ 并不是确定的值,它是随试验的具体结果而变的。同一个熟练工人,在这次生产 1000 件产品可能出 5 件次品(频率为 0.5%),在另一次生产 1000 件产品时可能出 4

件或 6 件次品(频率为 0.4% 或 0.6%)。经验证明, 当 k 较大时, 频率 $\frac{l}{k}$ 虽不能唯一确定, 但其数值是相当稳定的, 这叫做频率稳定性。举例如下。

【例 3】 随手向桌上掷一枚硬币, 以 A 表示落下后正面向上。将此试验重复 k 次, 以 l 表示其中 A 发生的次数, 当 k 充分大时, 频率稳定在 $\frac{1}{2}$ 附近。历史上有许多人认真进行过这种掷硬币试验, 表 1—2 是几个实际试验结果:

表 1—2

试 验 次 数 k	出 正 面 次 数 l	频 率 $\frac{l}{k}$
4040	2048	0.5069
12000	6019	0.5016
24000	12012	0.5005

【例 4】 对某工厂生产的大批同类产品进行合格率检查, 抽取 5 件、10 件、60 件、150 件、600 件、900 件、1200 件、1800 件, 结果见表 1—3。这里每抽一件产品是一次随机试验, 而抽到合格品是随机事件。

表 1—3

抽取件数 k	5	10	60	150	600	900	1200	1800
合格品数 l	5	7	53	131	542	820	1091	1631
合格频率 $\frac{l}{k}$	1	0.7	0.883	0.873	0.913	0.911	0.909	0.906

类似的例子很多，这种频率稳定性说明随机事件在大量试验中存在一种客观规律性，这种规律性来自大量统计，所以叫做统计规律性。

三、概率和实际推断原理

一个随机事件 A 在大量重复试验中的频率 $\frac{l}{k}$ 呈现稳定性，表明客观上存在一个数值作为频率的稳定中心，此数值叫做 A 的概率。 A 的概率通常记为 $P(A)$ 。当试验次数 k 充分大时，就取频率 $\frac{l}{k}$ 作为概率的近似值。

由上述可见，概率概念反映大量随机现象的数量规律性，而对个别随机试验来说，概率可以表明随机事件发生的可能性大小。例如，某工厂出次品的概率为 1%，意思是在该厂大量产品中次品约占 1%。而若从该厂产品中随机取 1 件，则取到次品的可能性很小，可以预言几乎一定会取到正品。

在生产及其他实践活动中，人们常把可能性很小（概率接近于 0）的随机事件看作实际上不可能发生，而把可能性很大（概率接近于 1）的随机事件看作实际上必然发生。这种处理问题的方法，叫做实际推断原理。例如，人们常把枪法高明的射击者（在一定条件下打中目标的概率接近于 1，但不会等于 1），叫做神枪手，认为他百发百中。又如通常乘车船而遇险的概率很小，人们在乘车船时一般是把遇险看作实际上不会发生的。

四、统计方法

研究一个母体可以有两种方法：一种是对母体中各个体无遗漏地全面研究，叫做全面统计；另一种是从母体中抽取子样，通过研究子样来对母体作出判断，叫做抽样统计。抽

样统计的方法简称统计方法。顾名思义地理解，似乎全面统计也应算统计方法，但在概率统计及有关的科技领域中，统计方法专指抽样统计的方法。

在工业中，当母体很大（如有的工业产品数以万计，十万计）或需要进行破坏性检验统计（如检验灯泡耐用时间，砖块抗折强度等）时，进行全面统计是行不通的，这时必须采用抽样统计方法。统计方法既然是由子样推断母体，那就不可能以百分之百的把握保证结论的正确性，它只能以大的概率（如99%、95%）保证结论正确。按照实际推断原理，可以用这种结论来指导行动，在绝大多数情况下（约占99%或95%），结论是正确的；在很少数情况下（约占1%或5%），结论会出错。通常，应用这种方法能节省大量人力物力，而少数错误结论（如判断少数正品为次品，判断错产品的等级等）造成的损失较小，因此统计方法在工业生产中有广泛的应用。

五、频率和概率的直接应用

对管理生产的同志来说，最要紧的莫过于直接应用频率和概率的简单概念于生产实际。心中有数意思不仅指心中要有绝对数字，而且指心中要有百分数，有频率和概率概念。例如研究一种生产措施的效果如何，至少应该统计一下采取这种措施的各单位中效果好的占百分之多少，效果一般的占百分之多少，效果差的占百分之多少。有一种不科学的方法是举例法，这种方法很容易被诡辩者利用。一种好的生产措施用于很多单位时，可能因某些原因在少数单位效果不好，诡辩者可以举这少数单位的例子否定这个措施。一种坏的生产措施用于很多单位时，在个别单位也可能因某些特殊原因取得较好效果，诡辩者也可以以此为例来肯定这种措施。由此

可见，为了实行实事求是的科学管理，应该提倡统计方法，首先是最简单的统计百分数，即频率的方法。

第三节 大子样的频率分布和直方图

本节以第一节例2中200个铆钉的直径为例，介绍对大子样进行整理分析的基本方法。

一、数据的分组

对于一个大子样 $x_1, x_2, \dots, x_n (n \geq 50)$ ，首先找出其中最小的数据（记为 x_1^* ）和最大的数据（记为 x_n^* ）。例如，200个铆钉直径的 $x_1^* = 13.13$ ， $x_n^* = 13.69$ 。取 a 略小于 x_1^* ，取 b 略大于 x_n^* ，则区间 $\Delta = (a, b)$ 就可以包含全部 n 个数据。

适当选取一个包含全部数据的区间 Δ ，把 Δ 等分为 k 个小区间 $\Delta_1, \Delta_2, \dots, \Delta_k$ ，由此相应地把数据分成 k 组。在铆钉直径的例子中，取 $\Delta = (13.095, 13.695)$ ， $k = 12$ ， $\Delta_1 = (13.095, 13.145)$ ， $\Delta_2 = [13.145, 13.195)$ ， $\dots$ ， $\Delta_{11} = [13.595, 13.645)$ ， $\Delta_{12} = [13.645, 13.695)$ ，则200个数据全都包含于 Δ 中。其中，包含于 Δ_1 中的为第1组，包含于 Δ_2 中的为第2组， $\dots$ 包含于 Δ_{12} 中的为第12组。

在取 Δ 和把 Δ 等分为 $\Delta_1, \Delta_2, \dots, \Delta_k$ 时，可以灵活掌握，一般 k 取10左右为宜。小区间的记法 $[\dots)$ ，表明左闭右开，即包括左端点而不包括右端点。

每个组的小区间长度 d ，叫做组距，左端点叫做组下限，右端点叫做组上限，小区间的中点值叫做组中值。

二、计算各组频数和频率

子样分组后，每组含有数据的个数，叫做子样在该组的