Shiyong Jishu Zhiliang GailüSheji

实用技术质量概率设计

汪觉生　陈 沉　汪民生　著

人民交通出版社

内 容 提 要

本书从实用、普及和生产需要出发，通俗地介绍了处理偏态分布的概率设计方法——瑞利法。首次完整地发表了作者多年研究和生产应用成果：瑞利分布数值表、概率纸、控制图、分布检验、异常值剔除，C_p 值计算等瑞利系列方法和工具，研究了在尺寸链计算、制订标准和质量评定等方面的应用，同时还综合介绍了其它偏态分布的研究和应用，发表了以原始正态参数 σ μ为参数的©χ ©γ分布的分布函数表。

本书所介绍的系列方法和工具，在工厂的生产，技术和质量管制实践中，特别是在形位误差的技术质量控制过程中证明是实用、简便而有效的。

本书适合技术和质量人员阅读，生产现场的中高级技工也可以通过本书掌握实用性很强的各种方法和工具的应用技巧。

实用技术质量概率设计

汪觉生　陈 沉　汪民生　著

人民交通出版社出版发行

(100013 北京和平里东街 10 号)

各地新华书店经销

印刷厂印刷

开本：787×1092 $\frac{1}{16}$ 印张：13.5 字数：346 千

1997年12月　第1版

1997年12月　第1版　第1次印刷

印数：0001—1 000 册　定价：26.00 元

ISBN 7-114- -

·

前　言

在机械制造中,通常遇到的技术设计有产品设计、工艺设计和装备设计。

一个完整的技术设计应由设计输入、设计主体图样或文件、设计输出三个部分组成。设计输入内容可以在设计任务书、计算书或设计说明资料等文件中阐明,设计输入所要阐明的主要问题是条件性和前提性的问题。设计输出内容可以在所设计完成的主体图样或文件中阐明,往往也以技术条件、技术要求,质量规范等文件作记载。

设计输出是功能性的规划,在规划功能性要求时应确定:

(1)指标(标志):技术质量特性;

(2)标准(公差):技术质量特性值的容许分散范围;

(3)观测值满足标准要求的可靠程度。

传统的技术条件或技术要求,把功能性的指标概括为技术特性,而在近代质量学、质量管理和控制中,则把功能性的指标概括为质量特性。

考虑到技术是质量的保证,质量是技术过程的输出内容,是技术设计所追求的主要目标。因此本书把传统的技术特性和近代质量学中的质量特性概括统一为技术质量特性。

当我们把功能性指标在概念上理解为技术质量特性时,那么反映这个概念的形式上的名词是什么就变得无关紧要。技术人员可以用技术特性来表达这一概念,质量人员可以用质量特性来表达这个概念,在这种情况下,我们的叙述就将技术、质量工作并入同一轨道。

近代概率概念和数理统计学逐渐渗透到各种科技学科和具体工作中,并与之相结合取得了广泛而显著的效果,从而显示其价值和应用前景。

当我们应用指标、标准和可靠程度来规划技术设计的功能性要求时,同时承认下述三个方面的事实:

(1)指标值即技术质量特性值是散布在一定范围内的,是随着机会而变的带有不确定性的变量,因此要求研究和掌握这种散布的规律。

(2)必须应用观测值即利用技术质量特性值散布的规律性和标准进行比较,从而回答观测值落在我们所关心的区间、特别是公差范围内的概率(稳定的比例或频率)问题,因为回答了概率问题就等同地知道了指标值落在有关区间内的可靠程度。

(3)若采用概率统计方法来制订标准,将更有利于在设计阶段就对观测值(或加工结果)满足标准要求的可靠程度作出估计。

在规划功能性要求时,要确定如上所述的指标、标准和可靠程度三类内容,同时考虑到以上三个方面的事实,因而提出技术质量概率设计概念就具有现实意义。

技术质量概率设计的含义对技术质量特性值以多大的概率落在所关心的区间,或落在容许的分散范围内加以规划和控制称为技术质量概率设计。

产品形成的全过程是由很多阶段组成的,最基本的过程单元称之为工序。一个产品的形成将历经很多工序,前面的工序为后面的工序服务。

如果将工艺准备看作一道工序,工艺人员对技术质量特性值加以规划和控制,那么后续的

现场生产工序就会产出一个相应的稳定质量水平。

如果车削加工工序对 5MIE 因素的生产条件进行监察和维持, 那么它就可以有保证地向后续工序提交具有稳定质量水平的车削成品, 因而有利于继续后面的磨削加工或其它工序。

稳定的质量水平, 是指服从一定分布的工序的总体质量。我们对该总体进行抽样检查, 就可以藉若干样品的观测值来估计考察期内的质量满足标准要求的程度, 或者判断质量是否进步或退化, 同时可以估计合格概率和所关心的区间概率, 掌握技术质量特性满足标准要求的可靠程度。

规划和控制可以在每个工序中进行, 只不过有时侧重规划, 有时侧重控制。因此, 不仅仅在技术设计中存在着, 而且在现场生产中也存在着技术质量的概率设计问题。

要解决技术质量概率设计问题, 就需要研究技术质量特性值的概率分布。

在形状和位置方面的技术质量特性, 由于形位公差的定义, 测量方法和形成形位误差生产条件的不同, 可能呈现各种分布规律。目前获得普遍承认的是瑞利分布和©χ©分布(绝对正态分布), 但是有关这方面的生产应用的研究资料甚少。为此, 本书力图从实用出发, 着重应用方法的通俗阐述, 系统地介绍瑞利分布的生产应用, 服务于形位误差的规划、控制和质量保证工作, 同时希望促进对瑞利分布生产应用的进一步研究和普及。

本书由中国一汽无锡柴油机厂汪觉生, 北京大学陈沉、汪民生著, 北京大学沈钧涛教授提供计算机计算和指导。本书的写作始于许金钊教授的提议和鼓励, 在写作过程中得到无锡市太湖耐热铸造厂丁荣兴厂长的支持, 在此一并致谢。

限于水平和经验, 书中错误和不足之处在所难免, 敬请读者批评指正。

作　者

目　　录

绪　论

研究统计规律的意义在于: 揭示波动的数据以及波动的质量中蕴涵着的规律性, 对工程进行质量控制和改进活动。

一、波动的数据

日常所说的操作者甲磨出零件的外圆直径尺寸和乙磨出的完全一样, 这句话是相对而言的。我们用钢皮尺测量这个零件可能是相等的; 而用游标卡尺测量就可能不相等了; 用游标卡尺测量可能是相等的, 而用外径百分尺测量就可能不相等而显示出波动性。当然, 我们用更精密的量具去测量它, 将会显示出更加细微的差别。如用微米千分尺可显示出千分之一毫米的差别, 在现代机械制造中要显示出万分之一毫米的差别也是常事。

在非常良好的条件下, 配备较高级的技师十分用心地磨制多件精密轴的外圆, 每根轴的外圆尺寸也不会是一样的, 只要采用相应精确度的量具, 就可鉴别出其中细微的差别。

因此, 数据的波动是绝对的, 在现实世界中这是一个基本事实。

古希腊朴素辩证法的奠基人赫拉克利特有一句名言:“人不能两次踏入同一条河流”, 因为河水在不停地流动, 当人们第二次踏入这条河流时, 接触到的已经不是原来的流水, 而是变化了的新流水。

对于磨削加工这个系统来说, 这“河水”是指什么?“流动”又是指什么呢?赫拉克利特认为:“世界是包括一切的整体”。世界万物是在永不停息地变化中作有规律的运动。这个整体就是原始的“系统”概念。磨削这个小世界自成一个整体, 自成一个系统。在这个系统里, 磨削输出的尺寸数据, 就象流动的水一样, 在生产中我们永远得不到两个“完全一样”的尺寸。

由于科学研究、生产技术和管理决策的需要, 系统理论有了很大的发展。现在认识到磨削功能的实现和其它加工技术一样, 是处于一定的系统运动中才能实现的。“系统”的概念逐渐为人们所接受和掌握。

所谓“系统”, 是指一个要研究和处理的对象, 是由相互联系、相互依赖、相互制约、相互作用的事物和过程所形成的统一整体。对于磨削加工来说, 我们所要研究和处理的对象是人、机器、材料、方法、测量和环境。它们是影响尺寸变化的因素。在质量控制中, 把影响质量的人(Man)、机器(Machine)、材料(Material)、方法(Methods)、测量(Measurement)和环境(Environment)称为 5M1E 因素。在这里我们把 5M1E 因素当作“系统”从整体上考虑问题。现在可以明白, 影响磨削尺寸的“物质”是 5M1E, 它们在运动中相互联系, 相互依赖, 相互制约, 相互作用。

在生产中由物质构成的生产系统的运动和变化是数据波动的根源, 在生产中数据发生动态的波动是必然的。

二、波动的质量

每一种有用的物体有许多属性, 从而可以在种种不同的方面有效用。譬如手表, 有的买主

希望走时准确,有的买主着眼外观漂亮,有的买主在意材质的价值,此时的手表就是具有三个方面属性的全体,从而在三个方面有效用。所以产品质量是使用价值的体现,是指产品质量特性满足使用要求的程度。它是产品本身所固有的,且在产品使用过程中表现出来的能满足人们需要的自然属性——质量特性所组成。

在制造过程中,符合性质量是指产品以某一特定标准进行衡量后的符合程度。这一点始终是质量概念的重要内涵。若将手表的走时准确度、漂亮程度、表壳含金量与特定标准进行衡量,显然其符合程度存在差异,这就是说,手表的质量存在差异,因而是波动的。

所谓手表的走时准确度、漂亮程度、表壳含金量就是关于使用要求方面的属性,就是质量特性。在实际生产中我们往往将质量特性与特定标准定量地相比较,从而体现出产品质量的优劣程度。

如把磨削尺寸作为产品的结构特性,当其与规定的公差带进行比较就可体现出该尺寸的优劣。尺寸的变化,用波动的数据来定量地表示,数据发生了波动,与公差比较后的符合程度也发生了相应波动,因此我们说质量是波动的。

当然,这个以衡量为目的的特定标准应与使用要求相适应,应符合使用要求。这个使用要求是销售后在用户手里使用时实现的,这就是通常所说的适用性质量。显然适用性质量也存在着波动。这个适用性质量是由标准的优劣和符合性质量的综合作用的结果。

质量是波动的,我们不能消除这种波动,但能控制这种波动,使它符合我们的要求,或是减少波动提高产品的使用价值,或是增加波动,在较大程度上降低造价,降低产品价值。后者也是一种质量的提高活动,因为我们追求的是全面的经济质量。

三、波动中的规律性

千差万别的原始数据常常惹人烦恼,甚至被杂乱无章的表面现象所迷惑而束手无策。但是建立在概率论基础上的数理统计技术可以为我们指点迷津。

例 0-1　在外圆磨床上磨　119.95±0.05mm 的外圆,将测得的原始数据记录于表 0-1。

磨外圆尺寸的数据记录表　(单位 0.001)　　表 0-1

-5	16	24	12	16	7	6	-9	14	-2
-8	-3	6	-16	4	5	11	8	-4	-6
-4	5	-1	16	-10	8	29	2	13	7
-2	-5	13	4	$-25^{(L)}$	18	-3	9	-3	-18
14	2	6	$38^{(H)}$	3	12	8	1	15	-11

仅将所录这一大堆数字资料进行直观观察,是很难发现这批数字资料的重要特征的。但是,若将数据按区间分组统计后(表 0-2),我们就可很快地看出,落入不同区间的数据个数是不同的,我们即刻掌握了数据不同区间的分布情况。

为更直观地掌握数据的分布情况,我们来制作数据球的模型。每个球代表一个数据,该球称数据球。每区间置一透明玻璃管,将数据球投入相应的玻璃管(图 0-1),此时我们可更直观而形象地看到数据在各区间的分布情况。一眼就可看出数据集中在哪个区间周围,哪一个区间数据最多,数据的分散程度、分散范围。图 0-1 中还清楚地显示数据球全部落在公差带(T_L,T_u)范围内,均值 X 稍偏离公差中值 L。凭直观可看出所磨出外圆尺寸对公差带的符合程度较高,质量较好。

数据统计表　　表 0-2

组　　号	组界(a、b]	组　中　值	频　　数	频　　率	累　积　频　率
I	- 30～- 20	- 25	1	0.02	0.02
II	- 20～- 10	- 15	4	0.08	0.10
III	- 10～0	- 5	13	0.26	0.36
IV	0～10	5	17	0.34	0.70
V	10～20	15	12	0.24	0.94
VI	20～30	25	2	0.04	0.98
VII	30～40	35	1	0.02	1.00
合　　计			50	1.00	

当用曲线勾画出数据球轮廓后，可得出一钟形曲线，该曲线就是以后将要谈到的正态分布曲线。利用数理统计方法我们可求出表示数据集中趋势的指标：平均值 X 和表示分散程度的指标：标准差 S。同时可很方便地求出数据落入各区间的概率，对总体作出描述和评价。此时，我们对数据的认识将从感性达到理性的境界，更充分地掌握蕴涵于数据中的规律性。

由此可知，经过统计处理的数据，将显示其规律性，而该规律性原先是蕴藏在波动的原始数据之中的，它需要我们去揭示。

图 0-1　数据球模型

四、工程质量与产品质量

我们明白了在生产过程中数据发生动态的变化是必然的，但某一具体数值的发生却纯属偶然。在偶然发生的数据背后，我们都可以找到它的必然规律，进而找到它的必然的原因。

譬如我们通过对磨削尺寸的数据收集、整理，了解到它的统计规律，就可以发现磨削工程系统中的 5M1E 处于什么状态。当我们改变这种状态时，统计规律将生相应的变化。这样我们就掌握了 5M1E 和数据整体的因果关系，达到驾驭系统、控制系统、改造系统，让它源源不断

地输出符合我们需要的统计规律的数据和产品(图 0-2)。

我们将影响系统输出质量的相互联系、相互依赖、相互制约的诸因素综合定义为工程, 诸因素的综合质量越高, 则工程质量越高。从整体上看工程质量越高, 系统输出质量就越高。

图 0-2　质量改进循环

加工系统的 5M1E 诸因素的综合也是一种工程, 不过是出产品的工序工程而已。这样, 我们从因果关系上看产品质量是由工程质量决定的(图 0-3)。工序的产品质量由 5M1E 工程因素的综合影响所决定(图 0-4)。在机械等工业生产中遇到的大都是工序质量。因此工程质量概念遂被工序质量所替代, 并由标准加以固定下来。

图 0-3　因果关系图

图 0-4　因果分析图

五、产品质量与统计技术

我们从质量工程学角度说明了研究数据统计规律性的目的是控制质量。其意义是揭示数据的统计规律, 通过该统计规律与工程系统的因果关系, 进而对工程系统采取控制改进措施, 使 5M1E 处于控制状态, 达到控制质量的目的。因而在生产活动中数理统计是一门行动的学问。脱离行动, 数理统计将迷失方向, 我们所追求的是结合实际, 采取措施, 取得实效, 并将措施纳入标准。若仅热衷于数据和方法而不注重措施, 则将使我们的研究毫无意义。

此外, 我们必须充分地认识到在工程中, 人作为管理者同时又是被管理者的双重作用。在工程系统中, 人始终处于中心地位, 是最活跃也是最难管理的因素, 但这是属于心理学和行为科学的讨论范畴。

当然, 数理统计的应用, 绝不仅限于质量控制和技术设计, 它在社会、生产、管理、自然各学

科中都有广泛而有效的应用。

统计技术实为工厂生产最需要的技术之一，譬如过去有相当数量的工厂，它们的成批生产设备长期生产不出合乎质量要求的产品，技术部门与生产部门互相推诿，从数理统计角度究其原因就是设备投产的时候没有应用统计技术对工序的能力进行论证，没有明确工序的控制标准和控制手段，事后又不认真地用统计技术去对工序作探讨和改进。

没有建立在统计意义上的质量评价往往是虚假的，譬如在工厂里经常听到某性能已达到一等品规定，认真追究却有很多不合格产品。作者曾经历过一件事：某零件磨削加工 4 万件无废品，但分布曲线却显示出 20% 的不合格率，因此统计技术对生产来说决不是可有可无的点缀。

在国际上，自从统计技术从英国传到印度和美国后，印度在理论上取得了进展，美国在应用上有了很大发展。战后，日本从美国引进了统计技术，创造出了结合日本国情的理论和应用体系。数理统计成为日本经济高速度发展的决定性因素之一。

在 ISO 9000 族的指南和要求标准中，设专门章节叙述了统计技术的应用。可见在国际商务和国际社会中是何等重视统计技术。

提高产品质量成为我国的重要经济决策后，10 多年来，随着质量管理的深入发展和积极推行 ISO 9000 族质量管理和质量保证标准，我们在使用统计技术时，逐步认识到不对称的偏态分布在质量控制中有广阔的应用场合和应用前景。譬如形位误差质量特性值。大部分不是呈现正态分布而是呈偏态的瑞利分布，因而原来以正态分布为理论假设前提的质量控制方法和设计程序不能适应处理形位质量特性值的需要，发生误用统计包现象也绝非个别。这已经引起我国数理统计、标准化工作、质量管理、生产技术各界的关注。由于目前国内外尚缺乏这方面专著，故结合作者多年工作经验和实践所得，本书重点将讨论瑞利分布的统计技术。希望为广大技术质量工作者提供一种简便、实用、有效的新概率设计工具。

第一章　概率论基础

第一节　概 率 初 步

一、概率的统计定义

定义　当试验条件可以重复，试验次数无限增加时，如果事件 E 发生的频率趋近于一个稳定的数值，这个数值就称为事件 E 的概率，以 P(E) 表示。

在上述定义中，把概率看成频率的稳定值，用来刻划事件 E 发生的可信程度。

掷硬币的模拟试验。约定掷 100 次硬币算做一次试验。由第 1 次试验到 n 次试验为止，其间出现正面的次数与掷硬币的次数之比，是出现正面的频率，试验结果示于表 1-1-1。

掷硬币试验结果表　　表 1-1-1

实验次数	出正面的次数	正面的(累计)频率	实验次数	出正面的次数	正面的(累计)频率
1	51	0.510	60	53	0.502
2	49	0.500	…	…	…
3	48	0.493	70	54	0.500
4	55	0.508	…	…	…
5	45	0.496	73	55	0.503
6	49	0.495	…	…	…
7	61	0.511	80	49	0.501
8	62	0.525	81	50	0.501
9	51	0.523	82	51	0.501
10	60	0.531	83	47	0.501
11	53	0.531	84	48	0.501
12	49	0.528	85	46	0.500
13	34	0.513	86	47	0.500
14	59	0.519	87	59	0.501
15	52	0.519	88	53	0.501
16	50	0.518	89	52	0.501
17	52	0.518	90	47	0.501
18	48	0.516	91	53	0.501
19	47	0.513	92	51	0.501
20	47	0.511	93	53	0.502
…	…	…	94	55	0.502
30	53	0.502	95	47	0.502
…	…	…	96	51	0.502
40	51	0.502	97	54	0.502
…	…	…	98	46	0.502
50	49	0.498	99	39	0.501
			100	56	0.501

随着试验次数的增加，频率的变化见图 1-1-1。

由图 1-1-1 可知，频率围绕稳定值 0.5 摆动，当试验次数增加时，频率变动的幅度逐渐变小。也就是说，频率趋向于稳定值 0.5，该频率的稳定值是 0.5，就是掷一次硬币出现正面这一

事件的概率。它表示在掷一次硬币时,出现正面的可信程度是 0.5,或者说是 50%。

在相同或稳定条件下,重复试验事件出现的频率,可以当作概率来使用。

在生产中如磨外圆的机器、材料、方法、测量、环境条件不变,所谓高级工磨外圆的合格品率为 98%,初级工磨外圆的合格品率为 90%,是指长期以来随着磨削次数的增加,高级工所磨制的合格品率趋近于 98%,初级工所磨制的合格品率趋近于 90%,这里的合格品率就是合格品概率。在各磨外圆少数次时高级工出合格品少,而初级工出合格品多的可能性是存在的,但高级工出合格品的可信程度是 0.98,初级工出合格品的可信程度依然是 0.90。若所磨外圆质量标准为 $50^{+0.05}_{0}$,则表示高级工磨出的外圆尺寸在 50～50.05 之间的频率稳定值即概率是 0.98,而初级工为 0.9。

图 1-1-1 掷硬币试验的频率变动

二、贝努里大数定理

设 μ为 n 次独立试验中事件 A 出现次数,P 是事件 A 在每次试验中出现的概率,则对任意的 ε 0 都有

$$\lim_{n\to\infty} P\left|\frac{\mu}{n}-P\right|<\varepsilon=1$$

贝努里定理有很大的实际意义,它阐明了事件频率与事件概率的关系。上式表明在 $n\to\infty$ 时,频率$\frac{\mu}{n}$与概率 P 之间的差,小于任意正数 ε的概率趋近于 1。它说明了当试验次数无限地增加时,事件 A 出现的频率围绕概率摆动,并将愈来愈接近概率,即$\left|\frac{\mu}{n}-P\right|<\varepsilon$的可能性愈来愈大,而成为必然事件。但这里并不是说概率 P 是频率$\frac{\mu}{n}$的极限,因为我们找不到这样足够大的数 N,使得 $n>N$ 时的一切 n 都满足不等式$\left|\frac{\mu}{n}-P\right|<\varepsilon$ 在 $n>N$ 时仍有可能产生$\left|\frac{\mu}{n}-P\right|>\varepsilon$情况,不过随着试验次数的增大,其可能性愈来愈小而已。

在掷硬币的试验中,从表 1-1-1 和图 1-1-1 可以看出,随着试验次数的增加,频率趋向于稳定值 0.5。试验次数超过 40 次(实际掷 4000 次)后,$\left|\frac{\mu}{n}-P\right|<\varepsilon$ 0.003 的可能性增大,频率没有越出 50±0.003 范围。当试验次数继续增加,这个频率的变动将会更接近于值 0.5,但是,从某个时候起将和 0.5 完全一致这种事情似乎没有。

三、概率的公理化结构

现在我们比较直观地了解了概率的统计定义,但因为它是建立在大量重复试验基础上的,在理论与实际上都存在一定困难,因而概率的统计定义是实用而近似的。20 世纪 30 年代开始

建立起严格的概率的公理化结构，用集合定义事件，用测度定义概率，用微积分等现代数学工具进行概率计算，使概率论的发展进入了新阶段。

概率的公理化结构是由前苏联数学家柯尔莫哥洛夫(A·H·ko Лmo ropob)在 1933 年完成的。

概率的公理化结构不直接给出概率是什么，而是在概率的统计定义基础上，提出一组关于随机事件 A 的概率 $P(A)$ 的公理。

公理 1　对任一事件 A，有 $0 \le P(A) \le 1$；

公理 2　对必然事件 S，有 $P(S) = 1$；

对不可能事件 ，有 $P(\ ) = 0$；

公理 3　若 $A_1, A_2, \cdots, A_n$ 两两相斥，A_i 和 A_j 不可能同时发生，即 $A_i \cdot A_j = \ \ (i \ne j)$ 则有

$$P(A_1 + A_2 + \cdots + A_n) = P(A_1) + P(A_2) + \cdots + P(A_n)$$

称满足上述公理化结构条件的一个函数 $P(A)$ 为概率。

例 1-1-1　一颗均衡的骰子，投掷一次组成一次试验，朝上的一面可以出现 6 个事件。

A_1——1 点　　A_2——2 点　　A_3——3 点

A_4——4 点　　A_5——5 点　　A_6——6 点

这是一次试验可能出现的全部结果，如图 1-1-2 所示。因所掷骰子是均衡的，各事件的出现等可能

$$P(A) = P(A_1) = P(A_2) = P(A_3) =$$

$$P(A_4) = P(A_5) = P(A_6) = \frac{1}{6}$$

不可再分割的事件是基本事件，以上 6 个基本事件及其概率值 $P(A) = \frac{1}{6}$ 是不是概率呢？

(1) 对任一事件 A_3 满足公理 1

如　　$0 \le P(A_3) \le 1$；

(2) 对任一必然事件满足公理 2

如　$P(S) = P(A_4 + A_4) = P(A_4) + P(A_4) = \frac{1}{6} + \frac{5}{6} = 1$；

对任一不可能事件满足公理 2

图 1-1-2　掷骰子的全部不可分割结果

如　　$P(\ ) = P(A_2 \cdot A_3) = 0$

(3) 全部事件概率和满足公理 3

$$P(A_1 + A_2 + \cdots + A_6) = P(A_1) + P(A_2) + \cdots + P(A_6) = 6 \times \frac{1}{6} = 1$$

$P(A_1), P(A_2), \cdots, P(A_6)$ 符合三条件，因此是概率。

例 1-1-2　一颗均衡的骰子，投掷一次，组成一次试验，朝上的一面出现二个事件 A_{I} 和 A_{II}

A_{I}——少于 4 点　　$A_{\mathrm{I}} = A_1 + A_2 + A_3$

A_{II}——不少于 4 点　　$A_{\mathrm{II}} = A_4 + A_5 + A_6$

这是一次试验可能出现的全部结果，如图 1-1-3 所示。各事件出现等可能 $P(A) = P(A_{\mathrm{I}}) = P(A_{\mathrm{II}}) = \frac{1}{2}$。

A_{I} 可再分割为 A_1, A_2, A_3，是三个基本事件的集合，A_{II} 可再分割为 A_4、A_5、A_6 也是三个基本事件的集合。

很多场合，我们对可分割事件的兴趣胜过基本事件，因此，现在要回答可分割事件 A_I 的概率值 $P(A=A_I)=\frac{1}{2}$ 是不是概率？

(1) 对任一事件如 $P(A_I)=\frac{1}{2}$ 满足：$0\leq P(A_I)=\frac{1}{2}\leq 1$

(2) 对任一必然事件如 A_I+A_I 满足

$$P(U)=P(A_I+A_I)=P(A_I)+P(A_I)=\frac{1}{2}+\frac{1}{2}=1$$

图 1-1-3　掷骰子的二个可分割结果

对任一不可能事件如 $A_I\cdot A_{II}$ 满足：

$$P(\ )=P(A_I\cdot A_{II})=0$$

(3) 全部事件概率和：

$$P(A_I+A_{II})=P(A_I)+P(A_{II})=\frac{1}{2}+\frac{1}{2}=1$$

$P(A_I)$，$P(A_{II})$ 满足概率三条件因此是概率。

概率的公理化定义可以更广泛地确切地描述随机现象，从而很好地满足科学和生产的需要。

当代各门自然科学、技术科学、管理科学、经济科学现阶段发展的一个重要特征是应用概率论。

四、事件与概率空间

1. 事件与样本空间

试验　试验就是对不同结果的观测过程，也就是确定条件的一次实现。试验可以是真实的，也可以是概念上的。

事件　试验或观测的一种结果称为一个事件。

基本事件　每次试验总可得一结果，每一个可能的单一试验结果称为基本事件，它不能再分割为另外的事件。

样本空间　每次试验的各种可能结果的集合称为样本空间(或基本空间)。

如果把样本空间看成元素集合，那么基本事件是集合中元素，基本事件与集合的一个元素且只与一个元素对应，样本空间的一个元素叫做一个样本点。

样本空间可用代号 S 表示。

在例 1-1-1，例 1-1-2 中，一颗骰子投掷一次生成的样本空间 S 分别是

6 个事件　$S=\{1, 2, 3, 4, 5, 6\}$

2 个事件　$S=\{<4, \geq 4\}$

我们还会经常遇见一种由二个基本事件 0、1 组成的双项分类的样本空间，基本事件 0 表示是什么，基本事件 1 表示不是什么，在表 1-1-2 中，列举了四个试验的双项样本空间。

双项分类的样本空间表　　表 1-1-2

试　验	基本事件 0	基本事件 1	试　验	基本事件 0	基本事件 1
掷硬币	正面朝上	反面朝上	磨外圆	合格	不合格
瞄准靶心打靶	中	不中	购买设备决策	购买	不购买

双项分类的样本空间可表达为

$$S=\{0,1\}$$

概率论开始于赌博，但其产生和发展的根源都只可能是生产、科学和社会的需要。

例 1-1-3　1654 年法国贵族德·梅吕骑士问著名的数学家和哲学家巴斯加，把两粒骰子掷 24 次，出现两个 6 的情况会有多少？掷质量均匀的两粒骰子这一确定条件的实现就是一次试验，一次试验的结果可能是：(1, 1)，(1, 2)，⋯，(2, 1)，(2, 2)，⋯，它们是基本事件共有 36 个(表 1-1-3)。

掷骰子试验结果表　　表 1-1-3

甲骰子							甲骰子						
乙骰子	1	2	3	4	5	6	乙骰子	1	2	3	4	5	6
6	(1, 6)	(2, 6)	(3, 6)	(4, 6)	(5, 6)	(6, 6)	3	(1, 3)	(2, 3)	(3, 3)	(4, 3)	(5, 3)	(6, 3)
5	(1, 5)	(2, 5)	(3, 5)	(4, 5)	(5, 5)	(6, 5)	2	(1, 2)	(2, 2)	(3, 2)	(4, 2)	(5, 2)	(6, 2)
4	(1, 4)	(2, 4)	(3, 4)	(4, 4)	(5, 4)	(6, 4)	1	(1, 1)	(2, 1)	(3, 1)	(4, 1)	(5, 1)	(6, 1)

这 36 个基本事件的集合是样本空间：

$$S=\{(1,1),(1,2),\cdots,(2,1),(2,2),\cdots,(5,6),(6,6)\}$$

或者说这 36 个元素构成掷两颗骰子试验的样本空间。

例 1-1-4　某磨床生产的 10 个成品，其中合格品 6 个，不合格品 4 个，从中任取 3 个，试述 10 中取 3 试验的样本空间。

解：某磨床磨出 10 个成品，其中合格品 6 个，不合格品 4 个，从中任取 3 个这一组条件的实现是一次试验。“3 个中恰有 0 个不合格品”是在合格品 6 中取 3，故有 $C_6^3=20$ 种可能试验结果，即有 20 个基本事件。“3 个中恰有 1 个不合格品”是在合格品 6 中取 3，在不合格品 4 中取 1，故根据乘法原理有 $C_6^2\cdot C_4^1=60$ 种可能试验结果。同样，“3 个中恰有 2 个是不合格品”有 $C_6^1\cdot C_4^2=36$ 种可能试验结果。“3 个中有 3 个是不合格品”有 $C_4^3=4$ 种可能试验结果，这些可能的试验结果共有 120 个，再也没有其它可能的试验结果了。它们是基本事件。它们的集合是样本空间。不管出现某种具体结果如何，我们总是在 10 个成品中抽取 3 个，因而样本空间的基本事件数可由成品 10 中取 3 决定，即有 $C_{10}^3=120$ 个基本事件。

例 1-1-5　在例 1-1-4 中若要回答“3 个中至多有 1 个不合格的概率”，则我们要关心“3 个中恰有 0 个不合格品”、“3 个中恰有一个不合格品”二类基本事件。例 1-1-4 和本例中每类基本事件显然是样本空间的子集，每类子集中包含样本空间中若干个同类基本事件。

2. 概率空间

用集合定义事件的图示。

在图 1-1-4 中将例 1-1-4，例 1-1-5 的样本空间用它的基本事件集 S 来表示，记为

图 1-1-4　样本空间

$$S=\{e_1,e_2,\cdots,e_{20}\}$$

将全部基本事件分为四类，每类分别表示：

$E_1=\{e_1,e_2,\cdots,e_{20}\}$——3 个中恰有 0 个不合格品

$E_2=\{e_{21},e_{22}\cdots,e_{80}\}$——3 个中恰有 1 个不合格品

$E_3=\{e_{81},e_{82},\cdots,e_{116}\}$——3 个中恰有 2 个不合格品

$E_4=\{e_{117}, e_{118}, e_{119}, e_{120}\}$——3 个中恰有 3 个不合格品

每类事件这个小集合 称为随机事件,因而从集合概念看,基本事件 e 相当于元素,基本事件集相当于全集,随机事件是全集的一个子集。

当随机事件仅含 1 个基本事件时,随机事件就是基本事件。

样本空间这个大集合,是 4 个随机事件这样的小集合(子集)的集合

$$S=\{E_1, E_2, E_3, E_4\}$$

样本空间包含着一切可能的试验结果,是必然事件,其概率为 1。

在例 1-1-5 中,若我们关心的是随机事件($E_1+E_2+E_3+E_4$)出现的概率,则有:

$$P_r(S)=P_r\sum_{i=1}^{4}E_i=\sum_{i=1}^{4}P_r(E_i)=$$

$$\frac{C_6^3+C_6^2C_4^1+C_6^1C_4^2+C_4^3}{C_{10}^3}=$$

$$\frac{20+60+36+4}{120}=1$$

式中,E_1 是不同的基本事件 $e_1, e_2, \cdots, e_{20}$的并(和)集,则 $P(E_1)=P(e_1)+P(e_2)+\cdots+P(e_{20})$。$E_2, E_3, E_4$ 具有同样的性质。

在例 1-1-5 中,我们关心的是随机事件(E_1+E_2)出现的概率

$$P_r(E_1+E_2)=P_r(E_1)+P_r(E_2)=$$

$$\frac{C_6^3+C_6^2C_6^1}{C_{10}^3}=\frac{20+60}{120}=\frac{2}{3}$$

因此从集合概念看,我们不仅关心基本事件,更要关心样本空间 S 中的某些子集,这些子集是可分割事件。

例 1-1-6 在表 0-2 中,列出了磨 50 个外圆时,直径尺寸落在七个区间中频数分布,磨削一次得某一直径 x, x 为基本事件,理论上 x 可以在〔$-\infty, +\infty$〕内取值,因而样本空间为数轴上无穷多个点的集合,若直径尺寸及公差修改为 $119.95^{+0.05}_{-0.01}$, x 落在 III、VI、V、VI、VII 及区间 II 的上界限。区间内该零件为合格,此时不经修配即可达到互换性要求,并能保证装配质量。在此,我们更为关心“外圆直径尺寸落在[119.4, 120] [$-\infty, +\infty$]中”这样一类事件,我们所要关心的样本空间的某些子集是区间 III、VI、V、VI、VII 及区间 II 的上界限。这些子集也是随机事件。当然区间 I、II 也是随机事件。

样本空间、随机事件及其概率称为概率空间。简言之,对样本空间所有随机事件都赋予概率时,称这个样本空间为概率空间。

例 1-1-4 中样本空间 $S=\{e_1, e_2, \cdots, e_{120}\}$,例 1-1-6 中样本空间 $S=$〔$-\infty, +\infty$〕都在赋予概率后构成概率空间。

由以上概率空间含义知,概率研究是以样本空间为基础的。

五、梵恩图(Venn Diagram)

一个样本空间及其中的事件可以用 Venn 图来表示,图中样本空间用一个矩形表示,事件 E 则在此矩形中用一个闭合区来表示,而在闭合区以外的矩形面积即相应的对立事件 $\bar{E}$,如图 1-1-5 所示。

随机事件用拉丁字母 A、B、C、D、E、…、U、V 来表示,并且用 S 表示必然事件,用 表示不

可能事件，用 E 表示对立事件。随机事件间的相互关系可以用一系列公式表示，其中基本的有

(1) A　B　事件 A 包含于事件 B，如图 1-1-6 所示。

所谓 A　B，就从 A 中任取基本事件 e，它一定是 B 的基本事件。

例如，在叙述随机变量时，如果 B 表示随机变量，A 表示连续随机变量，那么 A　B 表示连续随机变量包含于随机变量。此时 $P(B) \geq P(A)$ 成立。

图 1-1-5　对立事件

(2) A = B　两事件 A 与 B 中任一事件的发生必然导致另一事件的发生，即 A　B 且 B　A。

(3) $C = A \times B$　二个事件同时出现的事件，图 1-1-7。

(4) $C = A + B$　事件 A 与 B 中至少有一个出现的事件，图 1-1-8。

二个事件相交，表示二个事件同时发生。例如二人各掷一颗骰子，朝上一面出现二个事件：A 表示李君掷得少于 4 点，B 表示王君掷得不少于 3 点，则二人同时掷得 3 点用 $A \times B$ 表示。

图 1-1-6　包含事件　　图 1-1-7　事件的乘积　　图 1-1-8　事件的和

二个事件相加，意味着其中至少一个事件发生，例如无锡与北京之间客运方式：B 表示使用火车，A 表示使用飞机，那么无锡与北京之间客运方式可用 $(A + B)$ 表示。

一些事件可以由其它事件组合或推导出来，“相交”和“相加”是主要的两个基本方法。

(5) $C = A - B$　A 出现而 B 不出现的事件，如图 1-1-9 所示。

(6) $A \times B =$　　互斥事件，事件 A 与 B 不可能同时发生，如图 1-1-10 所示。

(7) $A_i A_j =$　$(1 \leq i \leq j \leq n)$　n 个事件互不相容，即 n 个事件 $A_1, A_2, \cdots, A_n$ 中任意二个事件不可能同时发生。

(8) $\sum_{i=1}^{n} A_i = S$，$A_i A_j =$　$(1 \leq i \leq j \leq n)$　完备群事件，如图 1-1-11 所示。

图 1-1-9　事件的差　　图 1-1-10　互斥事件　　图 1-1-11　完备群事件

n 个事件构成完备群是指这些事件的相加构成样本空间，且任意二个事件两两互不相容，即 n 个事件构成互不相容的完备群，这些事件称为完备群事件。