高等学校试用教材

压缩机可靠性

西安交通大学 金光熹 杨绍侃 编

机械工业出版社

前　言

本教材是根据1984年5月高等工业学校流体动力机械教材分编审委员会第二次委员（扩大）会议制定的新教学大纲以及审定的编写大纲编写的。

机器失效的原因，主要是构件的断裂和磨损，但在国内外现有的机器可靠性著作中，有的专门阐述机械零件结构强度计算，有的专门阐述机械零件的磨损计算。为了使读者对压缩机可靠性有较为全面的认识，在新的编写大纲中将两方面的内容同时列入。此外，可靠性技术是以概率统计为理论依据的，考虑到尚未学过概率论和数理统计学读者的需要，我们还简要地介绍了概率统计的基本知识。

本教材共分八章，系统地介绍了可靠性的基本理论、压缩机的可靠性设计与试验、压缩机故障分析与诊断。针对我国的实际情况，本书还专章阐述了提高压缩机可靠性的途径。

本书由西安交通大学金光熹、杨绍侃编写。其中杨绍侃编写第二章中的第六节、第七章；其余部分由金光熹编写。

本书主审王迪生教授对原稿提出了许多宝贵意见，特此致谢。本书由林强、李红旗协助整理图稿，在此一并致谢。

本书是压缩机专业必选课程“压缩机可靠性”的教材，也可供制冷设备与低温技术、内燃机和化工机械等专业师生以及从事压缩机研究、设计和制造的工程技术人员参考。

由于编者水平有限，书中难免存在一些缺点或错误，请读者予以批评指正。

编者

1988年9月

目　录

第一章　压缩机可靠性的基本理论

第一节　可靠性的基本名词、定义和特征量

一、可靠性

可靠性（reliability）是指产品在规定条件下和规定时间内完成规定功能的能力。

在这一定义中，“产品”是指作为单独研究和分别试验对象的任何元件、设备或系统。研究和试验对象中能独立完成给定功能的最简单的组成部分称为元件；产品的单元称为设备；各元件或设备的组合称为系统。元件和系统是一个相对性概念，随研究对象的变化而变化。例如，当研究动力站或化工流程的可靠性时，动力站或化工流程是系统，其中的压缩机视作元件。当以压缩机作为研究对象时，压缩机是系统，而压缩机中的零部件就是元件。

产品的“规定功能”是指产品应具有的技术指标的总和。压缩机的技术指标主要有体积流量（排气量）、体积比能（比功率）、噪声以及这些功能参数的保持性。只要有一项达不到，就是没有完成压缩机规定的功能。

“规定时间”是指要求产品处于完成规定功能状态的时间。随着研究任务的不同，“时间”这个概念还可以用周期、次数和距离等其它物理量表示。

“规定条件”是指产品在使用中的环境条件（如温度、压力、湿度、振动和噪声等）和工作条件（如工质、工况、维护和操作者技术水平等）。

产品通常分为可修复产品和不可修复产品两类。压缩机中大部分零件如曲轴、连杆、气缸和活塞等，在失效后一般均能修复，属于可修复产品。压缩机就整体而言，也属于可修复产品。活塞环、阀片、弹簧和密封填料等零件，在失效后将不能或不值得去修复，属于不可修复产品。

由于产品性质不同，对可靠性的理解也就不同。对于不可修复产品，主要考虑产品在规定时间内发生失效的难易程度，这就是狭义可靠性。对于可修复产品，还需考虑产品出现故障后修复的难易程度。这种在规定条件下使用的产品，在规定时间内、按规定的程序和方法进行维修时，保持或恢复到规定功能的能力，称为维修度（maintainability）。因此，对于可修复产品，应把不发生故障的狭义可靠性和排除故障的维修性结合起来，称为有效度（availability），这就是广义可靠性（图1-1）。

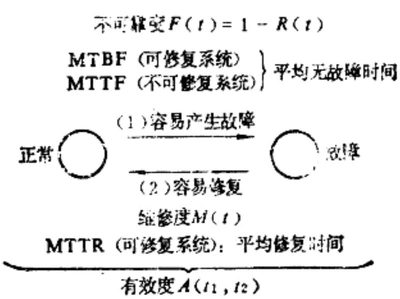

图1-1　狭义可靠性和广义可靠性

可靠度$R(t)$是用来表示产品可靠性高低的数量指标，它是指产品在规定条件下和规定时间内，完成规定功能的概率。

由可靠度的定义可知，可靠度是对一定时间而言的。同一产品，如果规定的时间不同，其可靠度

也就不同。产品在某一时刻 t 的可靠度等于该时刻仍然完好的产品数与最初投入工作的产品总数之比。

设有 N 个产品，从开始工作到 t 时刻的失效数为 $n(t)$，则当 N 足够大时，产品在 t 时刻的可靠度为

$$R(t)=\frac{N-n(t)}{N} \tag{1-1}$$

随着时间的增长，产品的可靠度越来越低。它是介于 1 与 0 之间的数值，即 $0\leqslant R(t)\leqslant 1$。

把抽象的可靠性用数学形式的概率表示，这就是可靠性技术发展的出发点。有了统一明确的尺度后，产品可靠性的评估、预测、分配、保证和管理等才有了基础。

二、失效

失效（failure）是指产品丧失规定的功能。对可修复的产品通常也称故障。为了便于分析，可以根据产品情况和实际需要适当分类（表1-1）。

表1-1 失效的分类

分类特征	分类	定义
失效原因	误用失效	不按规定条件使用产品而引起的失效
	本质失效	由于产品本身固有的弱点而引起的失效
失效阶段	早期失效	由于设计制造上的缺陷而发生的失效
	偶然失效	由于偶然因素发生的失效
	耗损失效	由于老化、磨损、疲劳等原因引起的失效
失效程度	完全失效	产品性能超过规定值，以致完全丧失规定功能的失效
	部分失效	产品性能超过规定值，但没有完全丧失规定功能的失效
预测可能性	突然失效	通过事前测试或监控不能预测到的失效
	渐变失效	通过事前测试或监控可以预测到的失效
失效的时间特性	间歇失效	产品失效后不经修复，而在限定时间内能自行恢复功能的失效
	突变失效	突然而完全的失效
	退化失效	渐变而部分的失效
失效后果	致命失效	可能导致人或物重大损失的失效
	严重失效	可能导致复杂产品完成规定功能能力降低的产品组成单元的失效
	轻度失效	不致引起复杂产品完成规定功能能力降低的产品组成单元的失效
失效的关联性	关联失效	在解释试验结果或计算可靠性特征量的数值时，必须计入的失效
	非关联失效	在解释试验结果或计算可靠性特征量的数值时，不应计入的失效
失效的独立性	独立失效	不是由于另一个产品失效而引起的失效
	从属失效	由于另一个产品失效而引起的失效

失效的特征量主要有以下几种：

1．累积失效概率（不可靠度）$F(t)$

累积失效概率（cumulative failure probability）是指产品在规定条件下和规定时间内失效的概率。$F(t)$的定义和$R(t)$的定义是相互对应的，属于相互独立的对立事件，两者不可能同时出现，也不可能同时都不出现，两者必居其一。因此，它们的概率之和恒等于1，即

$$R(t)+F(t)=1 \tag{1-2}$$

如果N个产品，从开始工作到t时刻的失效数为$n(t)$，则当N足够大时，产品在该时刻的累积失效率为

$$F(t)=\frac{n(t)}{N} \tag{1-3}$$

例1-1 有100台压缩机，在开始工作的1000 h内，累积失效数为10台，而工作到2000 h内，累积失效数为30台，求1000 h内和2000 h内的累积失效率和可靠度。

解 由题意知$N=100$，$n(1000)=10$，$n(2000)=30$

故

$$F(1000)=\frac{10}{100}=10\%$$

$$R(1000)=1-F(1000)=\frac{100-10}{100}=90\%$$

$$F(2000)=\frac{30}{100}=30\%$$

$$R(2000)=1-F(2000)=\frac{100-30}{100}=70\%$$

2．失效率$\lambda(t)$

失效率（failure rate）是指工作到某时刻尚未失效的产品，在该时刻后单位时间内发生失效的概率。

设有N个产品从$t=0$开始，工作到t时刻失效数为$n(t)$，又工作到$t+\Delta t$时刻，失效数为$n(t+\Delta t)$，则失效率为

$$\lambda(t)=\frac{n(t+\Delta t)-n(t)}{[N-n(t)]\Delta t}=\frac{\Delta n(t)}{[N-n(t)]\Delta t} \tag{1-4}$$

失效率是标志产品可靠性常用特征量之一。失效率越低，可靠性就越高。失效率用时间表示，通常用1000 h的百分数（$=10^{-5}/\mathrm{h}$）作为单位。对于高可靠性的产品，失效率用Fit（failure unit）表示，

$$1Fit=10^{-9}/\mathrm{h}=10^{-6}/\mathrm{kh}$$

例1-2 根据例1-1数据，求失效率。

解 $\Delta t=2000\ \mathrm{h}-1000\ \mathrm{h}=1000\ \mathrm{h}$，$\Delta n(1000)=30$台

故

$$\lambda(2000)=\frac{30}{(100-10)\times 1000\ \mathrm{h}}=33.3\times 10^{-5}/\mathrm{h}$$

复杂产品的典型失效率曲线见图1-2，这个图形实际上是由人的生命现象借用来的。若以横坐标为年龄，纵坐标为死亡率，则在区域Ⅰ（幼儿期），幼儿由于缺乏对外界的抵抗力，因此死亡率较高。到了区域Ⅱ（青壮年期），青壮年的生命力最旺盛，死亡率低而稳定，其死亡原因往往是由偶然事故等外界因素引起。进入老年期（区域Ⅲ），由于人的主要器官发生了本质上的老化现象，死亡率急剧增加。将这三个时期绘制成连续曲线，其形状似浴盆，故又称浴盆曲线。

对于产品，区域Ⅰ称为早期失效期，它是由于产品本身存在的设计错误，原材料有缺陷，制造工艺不佳等原因造成的失效，它出现在产品使用的初期。新产品在研究和试验阶段出现的失效，通常都是早期失效。这种失效应尽早发现，尽早解决，以便使工作稳定。进行合理的筛选，可以把早期失效产品淘汰掉，使出厂产品的失效率达到或接近偶然失效期的水平。

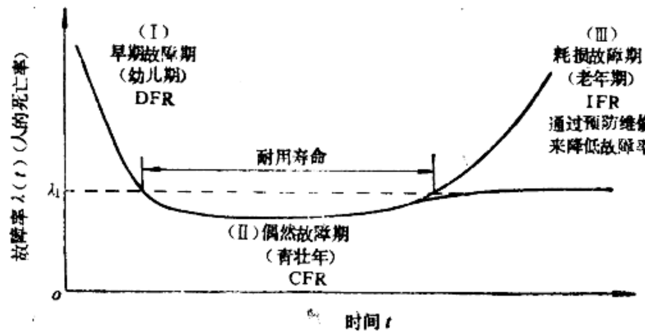

图1-2　典型失效率曲线

区域Ⅱ称为偶然失效期，在此期间产生的失效，是由于零部件中某些无法排除的缺陷在偶然因素作用下引起的，其失效率较低且接近于常数，这是产品最佳的工作阶段。

在耗损失效期（第Ⅲ区域），产品由于老化、磨损和疲劳等原因引起失效，其失效率随时间而上升。对于可以修复的产品，应在进入耗损期前采取维修措施，进行预防维修，换掉即将失效的元件，使图1-2中不断上升的失效率降低。

若规定产品的失效率为λ_1，则由图1-2的失效率曲线求得相应的使用寿命（图中称耐用寿命）。

三、可靠度函数

在研究可靠性技术时，应注意两个显著特点：时间性和概率性。可靠度$R(t)$和不可靠度$F(t)$就是两个与时间密切相关的概率分布函数。就概率分布而言，它们表示在该时间内完好产品或失效产品占全部工作产品的百分比。例如，经过筛选的压缩机开始工作时，$t=0$，所有压缩机都是完好的。因此，失效数$n(0)=0$，$R(0)=1$，$F(0)=0$。随着时间的不断增加，失效概率也不断增加，而完好的概率却相应减小。所有产品，不管寿命有多长，在使用过程中总是要失效的。因此，$n(\infty)=N$，$R(\infty)=0$，$F(\infty)=1$。在任何时刻，$R(t)+F(t)=1$。如图1-3a所示，可靠度$R(t)$为取值范围〔1，0〕的递减函数；不可靠度$F(t)$为取值范围〔0，1〕的递增函数。在$t=t_1$时，若$R(t_1)=95\%$，则$F(t_1)=5\%$。

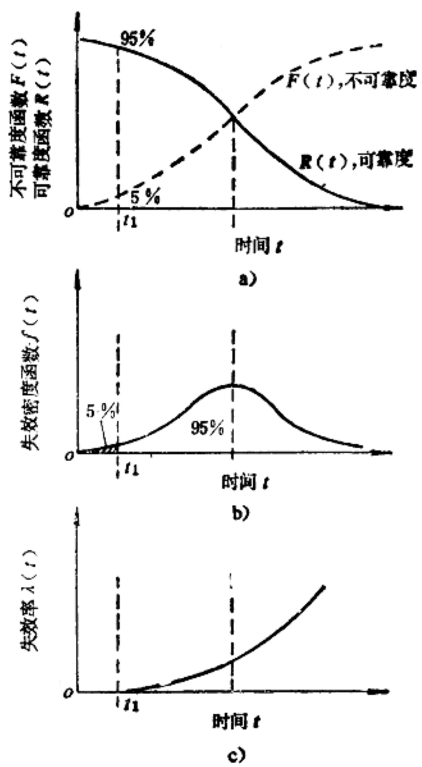

图1-3　$R(t)$、$f(t)$和$\lambda(t)$

为了描述产品在$0\sim+\infty$的整个时间内失效概率的分布情况，可将累积失效概率（不可靠度）

$F(t)$对时间微分，即

$$f(t)=\frac{\mathrm{d}F(t)}{\mathrm{d}t}=-\frac{\mathrm{d}R(t)}{\mathrm{d}t} \tag{1-5}$$

$f(t)$称为失效密度函数，它是指产品失效发生在t时刻的单位时间内的概率，表示瞬时的失效百分比。图1-3b中$f(t)$为一正态分布函数，曲线下面积就是失效的总概率，它应等于1。由式（1-5）及图1-3b可知，当$t=t_1$时，相应的$F(t)$和$R(t)$分别为

$$F(t)=\int_0^{t_1}f(t)\mathrm{d}t \tag{1-6}$$

$$R(t)=\int_{t_1}^{\infty}f(t)\mathrm{d}t \tag{1-7}$$

因此，图中阴影面积即表示$t=t_1$时的不可靠度$F(t)$，而无阴影线的面积即为可靠度$R(t)$。

在实际数据统计中，$(t, \Delta t)$时间间隔内的失效密度函数可按下式求得：

$$f(t)=\frac{n(t+\Delta t)-n(t)}{N\Delta t}=\frac{\Delta n(t)}{N\Delta t} \tag{1-8}$$

可靠度函数$R(t)$、$F(t)$、$f(t)$和$\lambda(t)$之间存在着如下关系：

$$\lambda(t)=\frac{n(t+\Delta t)-n(t)}{N\Delta t}\cdot\frac{N}{N-n(t)}=\frac{f(t)}{R(t)}=\frac{f(t)}{1-F(t)} \tag{1-9}$$

由此可见，失效率通过失效密度函数对可靠度或不可靠度的比值来表达。在失效率已知的情况下，可靠度也可以用$\lambda(t)$来表达。由式（1-5）和式（1-9）可得

$$\lambda(t)=\frac{1}{R(t)}\left[-\frac{\mathrm{d}R(t)}{\mathrm{d}t}\right]=-\frac{\mathrm{d}\ln R(t)}{\mathrm{d}t} \tag{1-10}$$

两边积分，便得

$$\int_0^t\lambda(t)\mathrm{d}t=-\ln R(t)\Big|_0^t=-\ln R(t)$$

$$\therefore\quad R(t)=e^{-\int_0^t\lambda(t)\mathrm{d}t} \tag{1-11}$$

式（1-11）为可靠度的一般表达式，它是从$\lambda(t)$的时间积分为指数的指数型函数。当已知失效率$\lambda(t)$时，就可以根据上式求得可靠度$R(t)$。

典型失效率曲线（图1-2）由三部分组成：递减型、恒定型和递增型，与此相应的$R(t)$和$f(t)$曲线见图1-4。产品在正常运行条件下，失效率属于恒定型，此时$\lambda(t)=\lambda=$常数。故

$$R(t)=e^{-\lambda\int_0^t\mathrm{d}t}=e^{-\lambda t} \tag{1-12}$$

四、寿命

平均寿命（mean life）是最常用的寿命特征量。在概念上，平均寿命对不可修复产品和可修复产品的含义不同。不可修复产品失效后，一般进行更新而不再修复，因此其寿命是指产品失效前的工作时间，平均寿命是指产品失效前的平均工作时间，通常记作MTTF（mean time to failure）。对于可修复产品，寿命是指两次相邻故障间的工作时间，而不是指每个产品工作到报废的时间。因此，对这一类产品，平均寿命就是平均无故障工作时间，通常记作MTBF(mean time between failures)。但是不管对哪一类产品，它们在理论上上的意义都是类似的，在数学上的表达式也是一致的。因此在以后的讨论中均统称为平均寿

命。

如果通过试验得到N个产品的全部寿命数据为t_1，t_2，…，t_N，则其平均寿命为

$$\bar{t}=\frac{1}{N}\sum_{i=1}^{N}t_i \qquad (1-13)$$

如果试验的产品数N较大，计算起来不太方便，可将观察值分组，然后按组取其中值来计算平均寿命。设N个观察值分成k组，以每组的中值t_i作为该组中每个观察值的近似值，则总工作时间就可用各组的t_i与相应频数（即组内观察值的数量）n_i的乘积和来近似表达，其平均寿命为

$$\bar{t}=\frac{1}{N}\sum_{i=1}^{k}t_i n_i \qquad (1-14)$$

由于n_i/N为第i组的频率(频数与总数之比）f_i，故上式可改写成

$$\bar{t}=\sum_{i=1}^{k}t_i f_i \qquad (1-15)$$

当分组数很大，分组很细时，$\Delta t\to 0$，上式的求和过程就变为求积分的过程，故

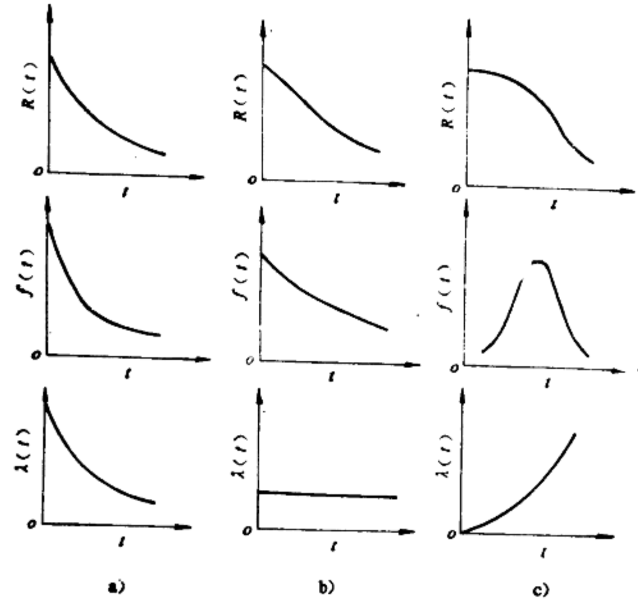

图1-4 失效率的三种基本类型及相应分布

$$\bar{t}=\int_0^\infty t f(t)\mathrm{d}t \qquad (1-16)$$

如前所述，产品在正常运行阶段出现的失效属于恒定型失效，服从指数分布，其可靠度$R(t)=e^{-\lambda t}$，失效率$\lambda(t)=\lambda$，由式（1-9）得

$$f(t)=\lambda(t)R(t)=\lambda e^{-\lambda t}$$

因此，失效规律服从指数分布的产品的平均寿命θ是失效率的倒数，即

$$\theta=\int_0^\infty \lambda t e^{-\lambda t}\mathrm{d}t=-te^{-\lambda t}\Big|_0^\infty+\int_0^\infty e^{-\lambda t}\mathrm{d}t=-\frac{1}{\lambda}e^{-\lambda t}\Big|_0^\infty=\frac{1}{\lambda} \qquad (1-17)$$

当$t=\theta=1/\lambda$时，$R(t)=e^{-1}=0.37$。因此，对于这类产品，能够工作到平均寿命的仅占37%左右，约有1－37%＝63%的产品在达到平均寿命前将发生失效（或故障)。

五、维修度和有效度

产品的可靠度随时间增加而减少，故障率则随产品的老化而增加，对于可修复产品，可以通过及时维修来防止其老化，降低递增的故障率。

维修的特征量是维修度$M(t)$，其定义已在第一部分中阐明，简言之，它是表征维修难易程度的概率。当维修度服从指数分布时

$$M(t)=1-e^{-\mu t} \qquad (1-18)$$

其中参数μ与可靠度函数中的故障率（失效率）相应，称为修复率（repair rate)，它是指

修理时间已达到某个时刻但尚未修复的产品，在该时刻后的单位时间内完成修理的概率。修复率的倒数称为平均修复时间MTTR(mean time to repair)

$$\mathrm{MTTR}=\frac{1}{u} \tag{1-19}$$

它与MTTF和MTBF相对应。

$M(t)$与$R(t)$的对比见表1-2。

表1-2 $M(t)$与$R(t)$的对比

比较项目	可靠度	维修度
累积分布函数	可靠度函数$R(t)$	$1-M(t)$
	不可靠度函数$F(t)$，失效状态的概率，失效难易程度	维修度函数$M(t)$，由于维修而恢复正常状态的概率，维修难易程度
密度函数	$f(t)=\frac{dF(t)}{dt}$	$m(t)=\frac{dM(t)}{dt}$
(单位时间的)率	故障率 $\lambda(t)=\frac{f(t)}{R(t)}$	修复率 $u(t)=\frac{m(t)}{1-M(t)}$
指数分布时的累积分布与平均时间	$F(t)=1-e^{-\lambda t}$ $\mathrm{MTBF(MTTF)}=\frac{1}{\lambda}$	$M(t)=1-e^{-ut}$ $\mathrm{MTTR}=\frac{1}{u}$

有效度$A(t)$是综合可靠度和维修度的广义可靠性尺度。对于可修复产品，虽然发生故障，但可以在规定时间内修复，重新正常工作，故有效度与单纯可靠度相比，增加了产品正常工作的概率，更合乎实际情况。实践证明，通过提高维修度来达到规定的有效度，要比通过提高可靠度来达到同样的目的更为经济。对于不可修复产品，有效度不受维修度影响，此时$R(t)=A(t)$。

有效度也是时间的函数，不同情况使用不同尺度：

（1）瞬时有效度$A(t)$　瞬时有效度（instantaneous availabilty）为产品在某时刻具有或维持其规定功能的概率。

（2）平均有效度$A(t_1,t_2)$　平均有效度（mean availability）指在某个规定时间区域内有效度的平均值。

在t_1，t_2时间区间内

$$A(t_1,\ t_2)=\int_{t_1}^{t_2}A(t)\mathrm{d}t/(t_1-t_2) \tag{1-20}$$

（3）极限有效度$A(\infty)$　极限有效度（li mi ing availability）又称稳态有效度或时间有效度，它是指当时间趋于无限长时，瞬时有效度的极限值。在可靠性问题中研究的主要是产品长时间使用中的问题，故极限有效度应用较多。通常在使用时将$A(\infty)$简写为A：

$$A=\frac{U}{U+D} \tag{1-21}$$

式中　U——产品因故障而不能工作的时间；

D——产品能工作的时间。

式（1-21）也可写成

$$A=\frac{\mathrm{MTBF}}{\mathrm{MTBF}+\mathrm{MTTR}} \tag{1-22}$$

若可靠度、维修度均属指数分布，$R(t)=e^{-\lambda t}$，$M(t)=1-e^{-ut}$，则式(1-21)还可改写为

$$A=\frac{u}{\lambda+u} \tag{1-23}$$

第二节　概率统计的基本知识

上述可靠性的各种特征量都是用概率表示的，可靠性设计和可靠性试验所用参数和公式，都是根据多次试验和测定所得数据，通过概率统计推导而得到的。因此，概率论和数理统计是可靠性理论的基础。为此，本节对可靠性技术中常用的概率统计知识作一简单介绍。

一、概率的基本概念

当我们抛掷硬币时，可能出现正面向上，也可能出现反面向上。当我们测试一台压缩机的性能时，其结果可能合格，也可能不合格。这种在一定条件下可能发生，也可能不发生，而在大量重复进行中却具有某种规律性的事件，称为随机事件（简称事件）。

上例中，硬币正面向上或反面向上，压缩机合格或不合格，都是孤立的事件，称为离散事件。我们在测量一个零件的尺寸时，在一定范围内的任何尺寸都可能出现，这种事件称为连续事件。压缩机的特性和可靠性的分布都是连续的，只是采用离散事件的分布可以简化计算，而其计算结果在一定条件下又适用于连续分布，因此往往从考虑离散事件着手。

概率（probability）是用一个具体数来表示某一随机事件出现可能性的大小。若在一次试验中，共有N个等可能而又互斥的事件，其中有M个能使事件A出现，则事件A出现的概率为

$$P(A)=\frac{M}{N} \tag{1-24}$$

硬币抛掷一次，出现正面向上和反面向上是两个等可能事件，但出现了正面就不会出现反面，反之亦然，故它们又是两个互斥事件。若事件A表示正面向上和反面向上同时出现，则该事件不可能发生，故$N=2$，$M=0$，$P(A)=0/2=0$。若事件A只表示正面向上，或只表示反面向上，则该事件属于可能发生，也可能不发生的随机事件，故$N=2$，$M=1$，$P(A)=1/2$。当事件A表示出现正面向上或反面向上时，则该事件必然发生，故$N=2$，$M=2$，$P(A)=2/2=1$。因此可知，任何事件A出现的概率总是介于0与1之间的一个数值，即

$$0\leqslant P(A)\leqslant 1 \tag{1-25}$$

在实际试验中，设某一试验重复进行n次，事件A在这n次试验中发生k次，则其频率为k/n。若试验次数越来越多时，事件A出现的频率始终在某一数字p附近作微小的摆动，则事件A的概率p，即

$$P(A)=p \tag{1-26}$$

这就是概率的统计定义。从统计定义出发，当试验次数很多时（即n很大时），事件A的概率p可用频率k/n来代替。

二、概率的基本运算

我们用字母A、B表示事件；用符号“$A+B$”表示事件A和B至少出现一件的事件，

称为事件和；用符号"AB"表示事件A和B同时出现的事件，称为事件积。如果事件AB为不可能事件，则A与B互斥，称A与B互斥事件；如果事件$A+B$为必然事件，而且事件A与B互斥，则A与B互逆（或互补），即A为B的逆事件，B为A的逆事件。

1。加法定理

有A、B两个事件，其中至少有一个事件要发生（即事件和）的概率为

$$P(A+B)=P(A)+P(B)-P(AB) \tag{1-27}$$

若A、B事件互斥，则

$$P(AB)=0$$

所以

$$P(A+B)=P(A)+P(B) \tag{1-28}$$

对于互斥事件A和B，我们有

$$P(A)+P(B)=1 \tag{1-29}$$

称为概率的互补定理。

例1-3 设某厂在一定条件下生产的一批压缩机，二级品出现的概率为3%，三级品出现的概率为1%，其余均为一级品。求出现二级品或三级品的概率。

解 设出现二级品的事件为A，出现三级品的事件为B，且A、B事件互斥，则

$$P(A+B)=P(A)+P(B)=\frac{3}{100}+\frac{1}{100}=\frac{1}{25}$$

故出现二级品或三级品的概率为1/25。

2。乘法定理

A和B两个事件同时发生（即事件积）的概率为

$$P(AB)=P(A)P\left(\frac{B}{A}\right)=P(B)P\left(\frac{A}{B}\right) \tag{1-30}$$

其中，$P\left(\frac{B}{A}\right)$是在事件$A$出现的条件下，事件$B$出现的概率，称为$B$对于$A$的条件概率。同理，$P\left(\frac{B}{A}\right)$为$A$对$B$的条件概率。若事件$A$和$B$的出现互不影响（独立事件），则$P\left(\frac{B}{A}\right)=P(B)$，$P\left(\frac{A}{B}\right)=P(A)$，代入式（1-30）得

$$P(AB)=P(A)P(B) \tag{1-31}$$

例1-4 从一批压缩机中任取一台进行测试，其性能指标合格的概率为95%，求连续抽取三台压缩机测试结果均合格的概率。

解 设第一次抽取压缩机合格为事件A，第二次为事件B，第三次为事件C，由于三次连续抽取，互不影响，属于独立事件，故

$$P(ABC)=P(A)P(B)P(C)=0.95^3$$
$$=0.86$$

故连续抽取三台压缩机均为合格品的概率为86%。

例1-5 某零件从毛坯到成品需经六道自动加工工序，若每一道工序的废品率均为0.6%，求成品的废品率。

解 此题应先求成品的合格率，再求其废品率。这是因为只有六道工序都合格，最后成

品才合格，故成品的合格率$P(A)$为事件积的概率。

设六道工序的合格率分别为$P(A_1)$，$P(A_2)$，…，$P(A_6)$，且互相独立，则成品合格率为

$$P(A)=P(A_1A_2\cdots A_6)=P(A_1)P(A_2)\cdots P(A_6)$$

按题意，六道工序的废品率$P(B_1)=P(B_2)=\cdots P(B_6)=0.6\%$，而废品率与合格率为相互对立的互补事件，故

$$P(A_1)=1-P(B_1)=1-0.6\%=99.4\%$$
$$P(A_1)=P(A_2)=\cdots P(A_6)=99.4\%$$

故成品的废品率$P(B)$为

$$\begin{aligned}P(B)&=1-P(A)=1-P(A_1)P(A_2)\cdots P(A_6)\\&=1-P(A)^6=1-(0.994)^6\\&=3.5\%\end{aligned}$$

3. 排列与组合

在计算较复杂事件的概率时，经常要遇到排列与组合问题。例如，求在编号为1，2，…，10的十个球中，一次取二个，其中一个为2号球，一个为5号球的概率。设想第一次取到2号球（不再放回）的概率为1/10，第二次取到5号球的概率为1/9，故此事件的概率为1/10×1/9=1/90。也可以设想第一次取到5号球，第二次取到2号球，则此事件的概率也为1/90。若要求取得2号球和5号球有一定顺序时，就涉及排列问题。

设从n个不同事件中取出r个排成一列，每次取出后不再放回，则排列数为

$$P_r^n=n(n-1)(n-2)\cdots\Big[n-(r-1)\Big]=\frac{n!}{(n-r)!} \tag{1-32}$$

由此可知，上例中10球一次取2个的排列数有

$$P_2^{10}=\frac{10!}{(10-2)!}=\frac{10\times9\times8\times\cdots\times2\times1}{8\times7\times6\times\cdots\times2\times1}=90$$

因此，取得某一排列的概率为1/90，与前面计算一致。

当一次取十个球，而十个球的编号按任何一个指定的顺序出现时，其概率为

$$P_{10}^{10}=\frac{10!}{0!}=10\times9\times8\times\cdots\times2\times1=3628800$$

若先取得2号球或先取得5号球不作区别，只要取出这两个球就行，则此事件的概率为$(1/90)\times2=1/45$，相当于将两个成功的排列的概率相加。此类问题可以按组合方法计算。

不考虑顺序时，从n个不同事件中取出r个为一组，其组合数为

$$C_r^n=\frac{P_r^n}{r!}=\frac{n!}{(n-r)!r!} \tag{1-33}$$

上例中，从十个球中一次取两个球的组合数为

$$C_2^{10}=\frac{10!}{(10-2)!2!}=\frac{10\times9\times8\times\cdots\times2\times1}{(8\times7\times6\times\cdots\times2\times1)\times2\times1}=45$$

因此，一次取出2号球和5号球成为一组的概率为1/45。

对于复杂事件的概率，可将事件成功的组合数除以可能的组合总数求得。

例1-6　一个箱子中有八个白球和六个黑球，随机取四个球，求正好二个白球和二个黑球的概率。

解 八个白球中取两个的组合数为

$$C_2^8=\frac{8!}{(8-2)!\,2!}=28$$

六个黑球中取两个的组合数为

$$C_2^6=\frac{6!}{(6-2)!\,2!}=15$$

从这个箱子中同时取出二个白球和二个黑球的组合数为

$$C_2^8\times C_2^6=28\times 15=420$$

从这个箱子中取出四个球的组合总数为

$$C_4^{14}=\frac{14!}{(14-4)\,4!}=1001$$

因此，从这个箱子中正好取出二个白球和二个黑球的概率为

$$\frac{C_2^8\times C_2^6}{C_4^{14}}=\frac{420}{1001}=0.42=42\%$$

三、随机变量与概率分布

可靠性技术研究的对象是随机事件，对随机事件进行定量分析时就需引入随机变量的概念。

随机变量是指随偶然因素变化并按一定概率取值的变量。如欲测试一批压缩机的排气量，每台压缩机的测试结果必然是随各种偶然因素而变化的一个变量，这些变量在一定条件下都是以一定概率出现的。因此，所测试每台压缩机的排气量属于随机变量。通过测试，我们可以通过以下所述直方图及分布曲图线，知道该随机变量的变化范围以及每一取值出现的频率或概率。

随机变量有离散型和连续型两种。前者为研究离散事件的变量，后者为研究连续事件的变量。

随机变量的概率分布是对随机变量最完善的描述。确定离散随机变量的概率（即频率）分布，就是要计算出随机变量每一可能取值的概率。确定连续随机变量的概率分布，就是要求出随机变量在某一区间内取值的概率分布规律，但不论是何种随机变量的概率分布，都是对某随机事件进行研究（试验或观察）得到的。

为了便于研究，在数理统计中把研究对象的全体称为母体，把组成母体的每个基本单元称为个体。例如，在评估1000台压缩机的平均寿命时，1000台压缩机的全体就是母体，其中的每一台压缩机则为个体。在工业生产和科学研究等领域里，对母体的每一个个体进行试验观察，往往需要消耗大量时间和人力物力，而且有些试验还具有破坏性，因此只能在母体中任意抽取一些个体进行试验观察，由观察所得数据资料对母体进行估计或推断，这种研究随机事件的方法称为抽样。

我们把从母体中抽取的个体的集合，称为母体的子样或样本，所抽取的个体数目 n 称为子样容量。例如，从1000台压缩机中任意抽取20台进行试验研究，则这20台压缩机的全体构成一个子样，其子样容量为 $n=20$。抽样本身是随机性的，因此抽样观察的结果也是随机性的，每个观察值均为随机变量。简单的随机抽样应满足两个条件：（1）在抽取子样的每个个体时，母体中每一个体被抽取的概率应相等；（2）在抽取子样的每个个体时，母体中的个体成分不变。因此，采取简单随机抽样的方法抽取容量为 n 的子样，实质上就是做 n 次独立的重

复试验。

在压缩机可靠性研究中，为了保证随机抽样的第一个条件，可将母体中的每一个个体编号，然后用抽签的方法来决定用几号个体进行试验。为了保证第二个条件，必须将试验后的个体重新放回母体，但若母体中个体的总数N远远大于子样容量n，例如当$N/n\geqslant 10$时，从母体中去掉n个个体，对母体成分的改变影响极小，可以忽略不计。在此条件下，逐个抽取n个个体进行试验，与一次抽取n个个体进行试验，其结果是没有区别的。因此在实践上，我们总是从一批压缩机中随机抽取n台同时进行试验。

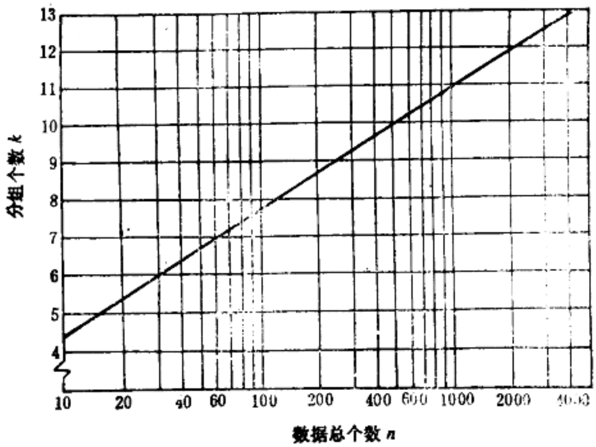

图1-5　确定分组数

将子样中总数为n的每一试验观察值记录下来，依大小次序排列好，并按一定间隔分组。分组数k可按图1-5确定，其关系式为

$$k=1+3.3\lg n \tag{1-34}$$

因此，当观察值的最小值为a，最大值为b时，组距$\Delta x_i=(b-a)/k$。求出每个组距Δx_i内出现的频数（即次数）n_i或频率与组距之比$f_i/\Delta x_i$后，便可绘出图1-6所示的频数直方图或频率直方图。频率直方图更为清晰方便。图中每一小矩形面积，等于随机变量在相应组距内出现的频率。直方图是子样的频率分布图，也是离散随机变量的频率分布图。

显然，子样容量n越大，组距Δx_i越小，直方图就越细密，当每个组距趋于零时，直方图的上缘将以光滑的曲线为极限，如图1-6中所示曲线。这种曲线就是母体的概率密度曲线，它表示母体的概率分布或连续随机变量的概率分布。概率密度曲线通常用概率密度函数$f(x)$来表示，简称为密度函数。

设有一个连续随机变量ξ，它的取值范围为$(-\infty<\xi<+\infty)$，其概率密度函数为$f(x)$，则ξ在任一区间(a,b)内的概率（图1-7）为

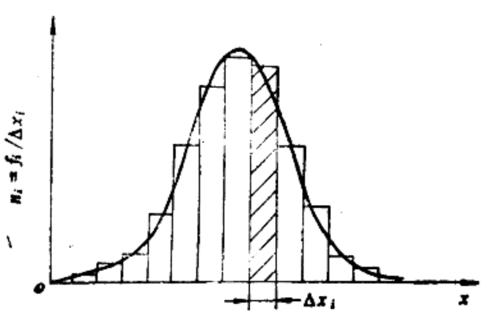

图1-6　频数、频率直方图

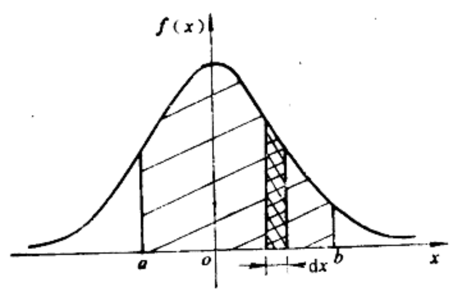

图1-7　分布曲线图

$$P(a<\xi<b)=\int_a^b f(x)\mathrm{d}x \tag{1-35}$$

分布曲线与横坐标所围成面积为全部概率之和，它应等于1，即

$$\int_{-\infty}^{\infty} f(x)\mathrm{d}x=1 \tag{1-36}$$

对于不同的随机变量，可以有不同的概率分布，也就有不同的密度函数$f(x)$。当已知某一随机变量的$f(x)$后，就可以按式（1-35）计算概率了。因此，对于连续随机变量，确定其概率分布也就是要确定其概率密度函数$f(x)$。

例如，测量中经常遇到的读数凑整误差和计算中的舍入误差，均属取值范围为(a,b)的连续随机变量，在$a\leqslant\xi\leqslant b$范围内，ξ取任一值时均有相同的概率，即

$$f(x_1)\mathrm{d}x=f(x_2)\mathrm{d}x=\cdots$$

$$\therefore\quad f(x_1)=f(x_2)=\cdots=C$$

而在(a,b)区间外，其取值概率为0。这种概率分布称为均匀分布。由式（1-35）得

$$P(a\leqslant\xi\leqslant b)=\int_a^b f(x)\mathrm{d}x=\int_a^b C\mathrm{d}x=C(b-a)=1$$

$$\therefore\quad \left.\begin{aligned} &f(x)=\frac{1}{b-a}, && a\leqslant x\leqslant b\\ &f(x)=0, && \text{其他}\end{aligned}\right\} \tag{1-37}$$

式（1-37）就是服从均匀分布的随机变量的概率密度函数，当$a=-0.5$，$b=+0.5$时，$f(x)=1/(b-a)=1$。

四、随机变量的特征量

当已知随机变量的概率分布时，其可能取值以及用什么概率取这些值就确定了。但在解决某些实际问题时，往往只需知道随机变量的某些特征量就可以了。在可靠性问题中，最常用的特征量是均值和标准离差。

1．均值

均值又称平均值或数学期望，它是表示概率分布集中趋势的特征量。

对于离散随机变量，均值是把一组数值相加后再以数值的个数相除而得到的算术平均值。若有n个数值的离散随机变量x_1，x_2，…，x_n，则其均值为

$$\bar{x}=\frac{x_1+x_2+\cdots+x_n}{n}=\frac{1}{n}\sum_{i=1}^{n}x_i \tag{1-38}$$

当数值个数较多时，为减少计算工作量，可将这些数值组成k组，用频数n_i或频率f_i计算，此时x_1，x_2，…，x_k代表各组中值：

$$\bar{x}=\frac{x_1n_1+x_2n_2+\cdots+x_kn_k}{n}=\frac{1}{n}\sum_{i=1}^{k}x_in_i \tag{1-39}$$

或

$$\bar{x}=\sum_{i=1}^{k}x_if_i \tag{1-40}$$

对于连续随机变量，均值可视为分布面积重心的x坐标：

$$\bar{x}=\frac{\int_{-\infty}^{\infty} x f(x)\mathrm{d}x}{\int_{-\infty}^{\infty} f(x)\mathrm{d}x}$$

因为

$$\int_{-\infty}^{\infty} f(x)\mathrm{d}x=1$$

所以

$$\bar{x}=\int_{-\infty}^{\infty} x f(x)\mathrm{d}x \tag{1-41}$$

以上为子样均值的求法。母体的均值μ与子样的均值$\bar{x}$的求法相同，只是计算μ时n代表母体的个体总数，而$\bar{x}$中的n则代表子样中的个体数。统计理论证明，μ是母体均值的精确值，$\bar{x}$是根据子样得到的母体均值的估计值，子样容量越大，$\bar{x}$值就越接近μ值。一般说来，子样均值是母体均值的最佳估计。

2. 标准离差

标准离差又称均方根偏差，简称标准差，它表示概率分布分散性的特征量。分散性的大小是指各个观测值离开均值的距离远近。

母体标准偏差的定义为

$$\sigma=\sqrt{\frac{\sum_{i=1}^{n}(x_i-\bar{x})^2}{n}} \tag{1-42}$$

式中，$x_i(i=1,2,\cdots,n)$表示一组观测值，n是观测值的个数（即母体的个体总数）。

对于连续随机变量，母体标准离差在数值上等于概率分布曲线下这部分图形的回转半径：

$$\sigma=\sqrt{\frac{\int_{-\infty}^{\infty}(x-\bar{x})^2 f(x)\mathrm{d}x}{\int_{-\infty}^{\infty} f(x)\mathrm{d}x}}=\sqrt{\int_{-\infty}^{\infty}(x-\bar{x})^2 f(x)\mathrm{d}x} \tag{1-43}$$

如前所述，子样均值是其母体均值的最佳估计，但标准离差却并非如此。为了使子样的标准离差作为母体的标准离差的最佳估计，子样的标准离差定义为

$$s=\sqrt{\frac{\sum_{i=1}^{n}(x_i-\bar{x})^2}{n-1}} \tag{1-44}$$

当子样容量较大时，$\sqrt{n-1}$便接近于$\sqrt{n}$。

在可靠性技术中，我们总是抽取一定容量的子样求其均值及标准离差，以此来估计母体的均值及标准偏差。子样容量的大小，关系到估计的正确程度，即关系到子样特征量的可靠程度。我们把这种估计正确性的概率称为置信度或置信水平。

另一种表示概率分布分散性的特征量是方差，它是标准离差的平方，即s^2或σ^2。

此外，表示随机变量特征量的还有中位数和极差等。将观察值x_1，x_2，$\cdots$，x_n按大小次序排列，当n为奇数时，取正中间的一个数；当n为偶数时，取正中间两个数的平均数，这个数称为中位数。观察值中最大值和最小值之差为极差。

五、置信度与自由度

1. 置信度

如前所述，用有限容量的子样来推断母体时，总存在一定误差。这种误差的大小表示由