

产品质量 检验监管统计技术

■ 主 编 李华森 李宜轩 刘心同
■ 执行主编 周尊英



 中国标准出版社

产品质量 检验监管统计技术

主 编

李华森 李宜轩 刘心同

执行主编

周尊英

审 定

杨立明 郭合颜 冯立兵

中国标准出版社

北 京

图书在版编目(CIP)数据

产品质量检验监管统计技术/李华森,李宜轩,刘心同
主编. —北京:中国标准出版社,2008
ISBN 978-7-5066-5087-8

I. 产… II. ①李…②李…③刘… III. 产品质量-
质量检验-监督管理-统计方法 IV. F273.2

中国版本图书馆 CIP 数据核字(2008)第 199003 号

中国标准出版社出版发行
北京复兴门外三里河北街 16 号
邮政编码:100045

网址 www.spc.net.cn

电话:68523946 68517548

中国标准出版社秦皇岛印刷厂印刷

各地新华书店经销

*

开本 787×1092 1/16 印张 21 字数 522 千字

2008 年 12 月第一版 2008 年 12 月第一次印刷

*

定价 49.00 元

如有印装差错 由本社发行中心调换

版权专有 侵权必究

举报电话:(010)68533533

编写人员名单

名誉主编 葛志荣

主 编 李华森 李宜轩 刘心同

执行主编 周尊英

副 主 编 王 岩 袁晓鹰 李 磊

徐京隆 赵 军 刘海峰

编写成员(以姓氏笔画为序)

丁立平 孙 灿 冯立兵 金 伟

周素杰 杨立明 郭合颜 贺照云

胡顺峰 高峰长 徐 刚 葛兆兵

审 定 杨立明 郭合颜 冯立兵

前 言

产品质量事关国家形象、企业前途和人民健康安全。自从改革开放特别是加入 WTO 以来,我国产品质量的总体水平有了很大提高,许多产品性能已达到或接近国际同类产品的先进水平。产品质量整体水平的提高,较好地满足了我国经济、社会发展的需要和人民群众日益增长的消费需求,增强了我国产品、企业在国际上的竞争力。但是也应当看到,我国在产品质量方面还存在着不少问题,在产品质量检验监管方面也存在着薄弱环节。例如近几年来食品中加入苏丹红、多宝鱼中农药残留、奶粉中含有三聚氰胺等事件的屡屡发生,说明了问题的严重性。因此,加强对产品质量检验监管的力度,提高产品质量和检验监管效率,突破国外技术壁垒,维护国家经济安全,在不增加产品成本的情况下,实现提速、减负、增效、严密监管的目标,已成为当前急需解决的一项课题。

众所周知,合格的产品是生产出来的,不是检验出来的。生产优质产品的企业必须要有健全的质量管理体系作为保障,而统计技术则是质量管理体系的基础。产品生产的每一阶段都离不开统计技术理论和方法的支持。

产品质量的评价必须依靠检验的手段,通过检验得出的数据,对产品批质量做出正确的判断,进而提高产品质量,并防止不合格产品进入流通领域。为使合格评定准确可靠,就必须运用统计抽样、描述性统计、测量分析、试验误差理论、不确定度评定、假设检验、过程能力分析、SPC 图、回归分析等多种统计技术和方法。只有正确应用统计技术和方法,才能对产品质量做出正确的评价,进而指导生产,使产品质量得以改进和完善;才能对产品质量及生产过程进行有效的监管和控制,使其始终处于统计受控状态。

广大质量检验人员通过大量的实践和科学研究,创建了许多行之有效的检验监管模式,并应用于日常检验监管工作之中。产品质量检验监管模式体系即是产品质量检验监管形式的体现,因此,都属于质量管理的范畴。国际质量管理标准强调质量管理必须以统计技术为基础,质量管理必须以数据和事实为依据,尽可能地量化管理指标,因此,不管哪种质量检验监管模式,都应以统计技术为理论基础。

统计技术应用于产品质量检验、管理和控制在国际上已有近百年的历史,能有效地使产品质量稳步提高,有一整套切实可行的统计技术和方法可借鉴。应当承认我国在这一领域还比较落后,特别是与发达国家的差距更大,这也是我国产品质量相对落后的主要原因之一。为此,我们编写了《产品质量检验监管统计技术》一书,目的是使大家重视统计技术在生产过程、产品检验、质量监督管理方面的作用,加强统计技术在检验监管中的应用,进行科学化管理,促使我国产品质量的总体水平稳步提高,进一步扩大在国际市场上的占有率。

本书共分10章,第1章简要地介绍了统计技术必要的基础知识;第2章阐述了统计技术在质量管理体系中的应用,介绍了统计技术在六西格玛管理体系、食品安全管理体系和产品检验实验室质量管理体系中的应用;第3章讲述了产品质量检验与评价的统计方法和示例;第4、5、7、8章介绍了统计技术在产品质量检验监管各种模式中的应用,并着重阐述了在源头管理、过程检验和分类管理模式中的具体应用;第6章介绍了食品安全保证中统计技术的运用;第9、10章以大量的示例介绍了统计技术在产品质量检验结果准确度和检验数据统计处理方面的一些理论和方法。

本书作者参阅了大量的国内外最新标准和文献,应用了一些科研课题的研究成果和具体的检验监管实践,结合国际上推荐的统计技术理论和方法编著了该书。因此,书中的定义、概念标准规范,内容新颖,并在产品质量检验监管方面有较大的创新。该书以较多的实例和示例对深奥的统计技术理论加以说明,使该书做到深入浅出、通俗易懂。

本书对保证检验结果准确度,提高质量控制、监督管理水平和产品质量等方面将有很大的帮助、促进和推动作用;可供不同行业的产品生产企业、以及质量管理、检验等部门从事产品质量检验、质量管理、质量监督、管理体系运行和认证等人员参考阅读。

在本书的编写过程中,得到了国家质量监督检验检疫总局、山东出入境检验检疫局、日照出入境检验检疫局有关领导的大力帮助和支持,还得到了席华、周素君、李建、王辉、金立宁、孙艳辉等同志的大力协助,在此一并表示衷心感谢!本书由李华森、李宜轩、刘心同任主编,周尊英任执行主编。

由于作者水平有限,书中难免有不当之处,敬请读者批评指正。

编 者

2008年9月26日

目 录

第1章 统计技术基础知识

1.1 概述	1
1.2 产品质量检验数据的分类	1
1.2.1 计量值数据	2
1.2.2 计数值数据	2
1.3 统计方法的分类	2
1.3.1 统计型方法	3
1.3.2 情理型方法	3
1.4 随机变量及其分布	3
1.4.1 随机现象	3
1.4.2 随机变量	4
1.4.3 随机变量的概率分布	4
1.4.4 离散型随机变量的分布	8
1.5 正态分布	10
1.5.1 正态分布的定义与性质	10
1.5.2 标准正态分布	11
1.5.3 正态分布的概率计算	12
1.5.4 正态分布的特性值	14
1.6 抽样分布	17
1.6.1 χ^2 分布	17
1.6.2 t 分布	18
1.6.3 F 分布	19
1.7 大数定律和中心极限定理	20
1.7.1 大数定律	20
1.7.2 中心极限定理	21

第2章 管理体系与统计技术

2.1 概述	22
2.2 质量管理体系常用统计技术	26
2.2.1 描述性统计	26
2.2.2 试验设计	29
2.2.3 假设检验	30
2.2.4 测量分析	32

2.2.5 过程能力分析	33
2.2.6 回归分析	35
2.2.7 可靠性分析	36
2.2.8 抽样	38
2.2.9 模拟	39
2.2.10 SPC(统计过程控制)图	39
2.2.11 统计容差法	41
2.2.12 时间序列分析	42
2.3 六西格玛管理体系	43
2.3.1 六西格玛管理与 ISO 9000 标准	43
2.3.2 六西格玛管理的定义与 特点	45
2.3.3 六西格玛管理的实施	47
2.4 食品安全管理体系	50
2.5 产品检验实验室质量管理体系	51

第3章 产品质量检验与评价

3.1 概述	53
3.2 产品质量检验	54
3.2.1 质量检验	54
3.2.2 检验的分类	55
3.3 产品质量平均水平的评价	60
3.3.1 算术平均值	60
3.3.2 加权算术平均值	61
3.3.3 几何平均值	62
3.3.4 中位数	63
3.3.5 众数	63
3.3.6 产品质量指标均值的区间 估计	64
3.4 产品质量稳定性的评价	65
3.4.1 极差	65
3.4.2 标准差、方差	65
3.4.3 变异系数	66

3.4.4	趋势图	67
3.4.5	过程能力	69
3.4.6	SPC图(统计过程控制图)	69
3.4.7	产品质量变异的区间估计	71
3.4.8	SN(信噪)比	72
3.5	产品质量水平	74
3.5.1	产品质量缺陷的分级标准	74
3.5.2	产品质量缺陷指导书	75
3.5.3	产品质量水平的确定	76
3.6	产品的使用寿命和可靠性分析	77
3.6.1	产品的使用寿命	77
3.6.2	产品的可靠性分析	80
3.6.3	产品的维修性和可用性	82
3.7	产品安全性评价	82
3.7.1	工业产品的安全性评价	83
3.7.2	食品安全性评价	88
3.8	产品的适应性和经济性	89

第4章 源头管理

4.1	源头管理的重要性	90
4.2	对供方的要求	91
4.3	采购产品的验证	91
4.3.1	描述性统计	91
4.3.2	过程能力分析	94
4.3.3	假设检验	95
4.3.4	抽样检验	103

第5章 过程检验

5.1	概述	106
5.2	过程检验的内容	107
5.3	影响过程的因素、波动及统计受控	107
5.3.1	影响过程的因素与波动	107
5.3.2	过程的统计受控	108
5.4	统计技术在过程检验中的应用	109
5.4.1	原辅材料与采购	110
5.4.2	资源管理	110
5.4.3	开发和设计	110
5.4.4	设备	111
5.4.5	过程的监视和测量	113
5.4.6	产品的监视和测量	131
5.4.7	数据分析与改进	136
5.5	过程检验中“PDCA”模式的应用	142
5.5.1	PDCA模式的四个阶段八个步骤	142
5.5.2	PDCA模式(循环)的特点	143
5.6	出口企业过程检验模式的实施	145
5.6.1	实施范围	146
5.6.2	申请过程检验的出口企业应具备的基本条件	146
5.6.3	过程监督检验内容	146
5.6.4	过程监督检验的实施	146
5.6.5	不合格判定及处理	147

第6章 食品安全的保证

6.1	概述	148
6.2	食品安全、食品安全标准及技术性贸易壁垒	149
6.2.1	食品安全	149
6.2.2	食品安全标准	149
6.2.3	技术性贸易壁垒	149
6.3	食品安全(卫生)法规与国际食品安全管理组织	150
6.3.1	我国食品安全的相关法规	150
6.3.2	国外及国际食品安全管理组织及相关标准/法规	150
6.4	国内外食品安全监管体制简介	151
6.4.1	分散管理模式	151
6.4.2	集中管理模式	151
6.4.3	中国食品安全监管体制	151
6.5	食品质量安全管理与统计技术的运用	152
6.5.1	良好生产操作规范与卫生标准操作程序	152
6.5.2	HACCP食品安全管理体系	153
6.5.3	ISO 22000食品安全管理体系	159

6.6 风险分析	165
6.6.1 食品风险分析的有关概念	165
6.6.2 风险分析的过程	166
6.6.3 统计技术在风险分析中的应用	168
6.7 食品加工过程的质量控制	175
6.7.1 正交试验法	175
6.7.2 正交表及正交设计	176
6.7.3 正交试验方案设计的一般步骤	177
6.7.4 正交试验的试验结果分析	178
6.8 食品实验室检验工作质量的控制	180

第7章 分类管理

7.1 概述	183
7.2 企业分类	183
7.2.1 一类企业	183
7.2.2 二类企业	184
7.2.3 三类企业	184
7.2.4 免检企业及产品	184
7.2.5 按企业产品的过程能力分类	185
7.3 产品不合格品率的计算	185
7.3.1 利用正态分布计算(估计)不合格品率	185
7.3.2 利用过程能力指数 C_p 计算(估计)不合格品率	187
7.4 检验监督管理	190
7.4.1 批次检验合格率	190
7.4.2 批次抽检率	190
7.4.3 抽检率及抽检批次的确定	190
7.4.4 抽检批的检验	192
7.4.5 质量管理体系的审核	192
7.4.6 产品质量的稳定性	193
7.4.7 型式试验	193
7.4.8 检验机构的审核	194
7.5 分类管理的动态管理	206
7.6 分项管理	206
7.6.1 检测项目重要性的分类	207
7.6.2 检测项目过程能力指数 C_p 的分类	208
7.7 分项管理的检验分类	209
7.8 分项管理的实施	209
7.9 分项管理的动态管理	211
第8章 产品质量检验监管模式体系	
8.1 产品质量检验监管的模式与体系	212
8.1.1 产品质量检验与监管	212
8.1.2 检验监管模式	213
8.1.3 检验监管体系	214
8.1.4 检验监管模式与体系的关系	214
8.2 检验监管模式的创建、发展与现状	215
8.2.1 检验监管模式的创建	215
8.2.2 检验监管模式的发展与现状	215
8.3 检验监管模式必须以统计技术为基础	216
8.4 进出口商品检验监管模式体系	217
8.4.1 全数检验模式	218
8.4.2 抽样检验模式	219
8.4.3 型式试验模式	238
8.4.4 过程检验(监控)模式	240
8.4.5 登记备案模式	243
8.4.6 符合性验证模式	244
8.4.7 符合性评估模式	246
8.4.8 合格保证模式	248
8.4.9 免于检验模式	251
8.5 检验监管模式的动态调整与日常检验监管	252
8.5.1 检验监管模式的动态调整	252
8.5.2 日常检验与日常监管	253
第9章 产品质量检验结果的准确度	
9.1 测试误差	254
9.1.1 真值和约定真值	254

第1章 统计技术基础知识

1.1 概述

统计就是有目的地收集数据,整理数据,并使用相应的方法制图、列表与分析数据的过程。对于产品质量检验监管而言,必须要以数据为依据,才能做出对产品质量的评判。所以需要收集数据、整理数据以及使用相应的一些方法分析数据,因此,产品检验监管包含了大量的统计工作。

统计技术是以概率论为理论基础,研究随机现象中确定的数学规律的学科。是应用数学的一个分支。统计技术是一个大的概念,是指整个学科而言,指的是一门技术的总概括。

产品质量检验就是以样本来推断整批产品质量特性的过程,产品质量特性值的检验结果本身就是随机现象,因此,真值永远是未知的,所以只能以一定的概率来评估产品质量特性的真值。而产品质量监督管理则是根据这些检验结果在一定的概率下进行统计控制和管理。因此,产品检验监管也必须使用以概率论为理论基础的统计技术。

统计方法是指统计技术中的具体方法。如排列图是统计技术中的一种方法,显著性检验是统计技术中的一种方法,等等。原则上不应称排列图、显著性检验等为统计技术,但在实际应用和习惯上这两种概念往往相混淆,不那么严格区分。

统计工具是指简化的统计方法。统计工具的开发是日本质量管理专家在质量管理工作中的一大贡献。统计技术的理论基础是概率论,但概率论这一理论对于基层的工人甚至对于一般的技术人员来说也是难以理解的,因此也妨碍了统计技术的推广应用。为此,针对基层工人和初级技术人员的特点,20世纪60年代日本质量管理专家开发了:因果图、排列图、调查表、散布图、控制图和分层法,称之为质量管理常用的七种工具。随着质量管理的不断深化,20世纪70年代日本质量管理专家又开发出系统图、关联图、矩阵图、矢线图、KJ(亲和图)法、PDPC(过程决策程序图)法和矩阵数据解析法,称之为质量管理的新七种工具。这些方法简便、易行、易懂、直观、质量改进的效果明显,故亦称质量改进工具和技术,或称作描述性统计。

所谓的工具,指不讲统计方法的原理和设计,也不讲对统计结果的分析,只讲操作步骤。好比工人在生产过程中操作各种工具,如榔头、扳手、螺丝刀等,并不要求工人掌握这些工具的设计原理和制造工艺,只要能在适用条件下得心应手地使用这些工具就可以了。

因此,在一些国际标准如ISO 9000族标准中把这些统计工具作为产品质量改进的工具和技术,帮助质量改进项目和活动的成功。这些统计工具在产品生产的各个阶段,特别是在产品质量检验及监督管理中是应用最多的统计方法之一。

1.2 产品质量检验数据的分类

产品质量检验的数据是我们统计的对象,人们习惯上把由数字组成的数字数据称为数据。从广义上讲,数据还包括用语言和文字表示的非数字数据(产品质量感官检验中应用的较多)。

数字数据指由数字(0,1,2,3,4,5,6,7,8,9)和小数点组成的数据。

数字数据是对可定量描述的特性的表达。可以通过抽样、测量、测试、分析、计算、记录获得。任何数字数据又都可以形成(服从)一定的分布(统计规律)。

根据数据的性质,数据又可分为计量值数据和计数值数据。计数值数据又可分为计件值数据和计点值数据。

1.2.1 计量值数据

计量值数据是指可以连续取值,在有限的区间内可以无限取值的数据。如长度、面积、体积、重量、密度、电压、电流、强度、纯度、有毒有害元素含量等,产品质量检验中大部分质量特性的数值都属于计量值数据。计量值数据服从正态分布。

1.2.2 计数值数据

计数值数据是只能间断取值,在有限的区间内只能取有限数值的数据。计数值数据就是我们通常所说的数数,如产品质量感官检验中的大部分项目属于计数值数据,如布匹上的疵点数,铸件上的沙眼数,玻璃器皿上的气泡数,不合格产品的件数以及某些仪器所测试的项目如钢铁硬度计所测的洛氏硬度、布氏硬度等。所以计数值数据,是以正整数(自然数)的方式表现。

1.2.2.1 计件值数据

计件值数据是指产品(或其他物件)件数而得到的数值。计件值数据服从二项分布。

1.2.2.2 计点值数据

计点值数据是指产品的缺陷数而得到的数值。计点值数据服从泊松分布。

1.3 统计方法的分类

产品质量检验数据是统计方法应用中的统计对象,根据统计数据类别不同,统计方法可分为统计型方法(适用于数字数据的方法)和情理型方法(适用于非数字数据的方法)。表1-1总结了两类不同方法的特点。

表 1-1 两类统计方法的比较

方法类别 比较项目	统计型(数字数据)方法	情理型(非数字数据)方法
收集对象	收集大量定量描述的数字数据。收集方法: 取样、测试、计算、记录	收集大量定性描述的非数字数据。 收集方法:调查、研究
处理方式	对数字数据进行统计计算,取得反映客观规律的质量特性值	对非数字数据(语言资料)进行分类、 归纳、整理,得到有条理的思路
功 能	实施统计推断(预测)及统计控制	作为决策依据
方法类型	常见统计方法中大部分属于统计型方法,如 控制图、散布图、直方图、试验设计、方差分析、 回归分析等	常见统计方法中少数为情理型方法 如因果图、分层图、流程图、树图、水平 对比法、头脑风暴法等

1.3.1 统计型方法

1.3.1.1 收集的对象

统计型方法收集的是大量可以定量描述的数字数据。收集数据是通过取样、测量、测试、分析、计算和记录而实现的。

1.3.1.2 数据的处理

统计型方法应用一定的数学公式,对取得的数字数据进行统计计算,得到反映客观规律的质量特性值。

1.3.1.3 用途

数字数据分布的特征值反映了产品的质量波动状况,所以统计型方法常用于过程的统计控制和统计推断。

在常用的统计方法中统计型方法占有较大的比重。如描述性统计中的排列图、直方图、散布图、控制图、趋势图等以及显著性检验、方差分析、试验设计、过程能力分析、回归分析、测量分析、SPC图、时间序列分析等。在产品质量检验和监督管理中我们所用的统计方法大都是统计型方法。

1.3.2 情理型方法

1.3.2.1 收集的对象

情理型方法收集的是大量定性描述的非数字数据,非数字数据需通过广泛、深入的调查研究和利用感官观察获得。

1.3.2.2 数据的处理

非数字数据的处理又称为资料处理。即把复杂事物的多方面信息收集起来,利用其相互内在的联系,加以整理,通过分层、分类、归纳、总结,得到解决问题的思路。

1.3.2.3 用途

通过情理型方法的应用,所得到的有条理的思路,可作为决策过程的输入(决策依据)。情理型方法多应用于描述性统计方面,如因果图、流程图、分层图、关联图、水平对比法、头脑风暴法等。在产品质量检验监管中除描述性统计外,其他方法较少应用。

由于统计型方法大都需要仪器设备的测试、分析并涉及到大量的数字数据、公式、统计处理和计算等,比较复杂,所以就认为应用难度大。而情理型方法的应用过程比较简单,认为容易掌握。其实恰恰相反,前者如果进行产品检验的设备仪器、方法、操作、公式和计算都是正确的,检验数据则是准确的。对同一产品质量问题无论任何检验人员去检验,所得出的结论大都是相同的。而情理型方法的应用与检验人员的思维判别能力、知识面、素质等因素有很大的影响,同一产品质量问题用同样的方法而由不同的检验人员进行判断可能会得出绝然不同的结论。

1.4 随机变量及其分布

1.4.1 随机现象

当给定一组所需的条件后,所发生的结果是不确定的(随机的)现象称为随机现象。

所谓的现象是指事物的表现,任何现象都不会无缘故地发生。是在一定条件下所导致的结果。根据条件与结果的关系,现象可分为确定现象和随机现象两大类。

确定现象(必然现象)是当给定一组所需要的条件后,所发生结果是必然的(确定的)现象称为确定现象。如水在一个标准大气压、温度 0°C 时,必然结冰。这属于确定现象。

随机现象结果是不确定的,大部分产品的质量特性属于随机现象。例如在车床上加工一批圆轴,要求每一根圆轴的直径必须保证为 20.00mm 。显然,这样的加工任务任何人都无法完成,因为质量是有变异性的,它受若干不确定随机因素的影响,是一种随机现象。

产品检验结果虽然是一种随机现象,但并不是漫无边际的随机,而是在一定范围内符合一定数学规律的随机。

检验结果的随机性造成质量特性(试验数据)的波动,而表现为质量特性值(试验数据)的分布。

1.4.2 随机变量

量可分为常量和变量。常量即是取固定数值的量,如圆周率 $\pi=3.1416$,自然常数 $e=2.718$ 等。变量是数值可以变化的量,它是相对于常量而言的。

随机变量:取值随试验结果而定,且有一定的概率分布的变量。

随机变量是变量中的一种类型,随机变量的数值是由随机因素的作用而发生的。随机变量在相继取值的过程中,下一个数值的大小是不可预测的。

产品的质量是一种随机现象,反映质量的数据就是一种随机变量。反映质量的随机变量是数字数据,所以又有计量值数据、计数值数据(计件值数据和计点值数据)之分,各自服从一定的分布。

随机变量具有以下特点:

① 随机性:即在试验前只知道它的可能的取值范围,而不能预知它具体取什么值。

② 统计规律性:由于试验的各个结果的出现有一定概率,于是随机变量取值也有一定的概率。

③ 互斥性:随机变量所取的各个可能值是互斥的,取某个数值后,不能再取其他值。

最常见的随机变量有两大类,即离散型随机变量和连续型随机变量。如果一个随机变量只可能取数轴上有限个或可数个孤立的值,并且对应于这些有确定的概率,则称这个随机变量为离散型随机变量。离散型随机变量的取值是计数值数据。如果一个随机变量所有可能取的值是充满某一区间,甚至整个一段数轴,就称为连续型随机变量。连续型随机变量的取值是计量值数据。

1.4.3 随机变量的概率分布

1.4.3.1 概率分布的有关概念

1. 概率

概率即度量一随机事件发生可能性大小的实数,其值介于 0 与 1 之间。

注:随机事件即是在一定的条件下,可能发生也可能不发生的事件称为随机事件,简称事件。

例如,投钱币时,蒲丰(Buffon J. L.)投过 4040 次,得 2048 次正面,出现正面频率为 $2048/4040=0.5069$;皮尔逊(Pearson K.)投过 24000 次,得到 12012 次正面,出现的频率为

0.5005, 频率稳定到常数 0.5, 故出现的正面概率为 0.5。

事件越容易发生, 其出现的概率越大。事件越不易发生, 其出现的概率越小。必然事件出现的概率为 1, 不可能事件出现的概率为 0。

2. 随机变量的概率分布

随机变量的概率分布: 给出一个随机变量取任何给定值或取值于任何给定集合的概率的函数。

注: 随机变量在全部集合取值的概率为 1。概率分布简称分布。

结合产品质量检验我们可以以质量统计的观念来理解概率分布的概念:

1) 产品质量具有变异性(不均匀性、不一致性)

产品质量是一种随机现象, 由于影响产品质量的不确定因素(人、机、料、法、环)无时无刻都在变化着, 所以产品质量具有变异性。例如, 以上所述要求工人在车床上加工 100 根 $\phi 20\text{mm}$ 的圆轴, 要求每一根圆轴的直径必须保证恰好为 20.00mm , 恐怕任何技艺高超的工人也无法完成这样的加工任务。由于质量的变异性, 应当给出一个公差范围。如果下达的任务为车制 100 根 $\phi(20 \pm 0.10)\text{mm}$ 的圆轴, 则任何一名熟练的车工都可以完成。这是因为产品质量具有变异性。

2) 产品质量的变异具有规律性(分布)

产品质量虽然具有变异性, 但不是漫无边际的变异, 是在一定的范围内而且符合一定规律的变异。如上例所述, 工人在接到加工 $\phi(20 \pm 0.10)\text{mm}$ 圆轴的指令后, 必然要尽可能地加工出 $\phi(20 \pm 0.00)\text{mm}$ 的圆轴, 因为产品质量的变异性, 而不可能保证每一件产品都是 $\phi(20 \pm 0.00)\text{mm}$, 而是在 $19.90\text{mm} \sim 20.10\text{mm}$ 范围内分布, 但也不是平均分配, 而是在接近 20.00mm 处出现的数据最多, 偏离 20.00mm 越远数据出现的就越少。图 1-1 反映了以上的情况, 实际就是质量数据(随机变量)的分布。当产品质量特性属连续型随机变量即计量值数据时, 其分布大部分如图 1-1 所示。

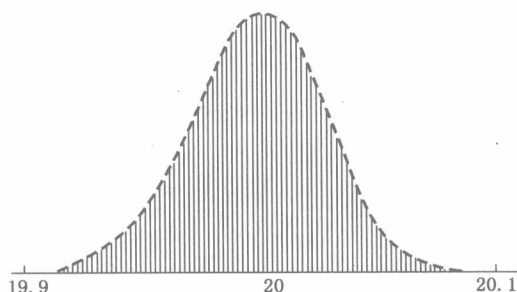


图 1-1 随机变量的分布

3. 分布函数

随机变量 X 小于或等于实数 x 的概率称为分布函数, 它是 x 的函数:

$$F(x) = P(X \leq x)$$

分布函数的意义就是: X 是一个随机变量, x 是任意实数, 那么“ X 取小于(或小于或等于) x 之值”时, 这个事件的概率。

4. 连续随机变量及[概率]密度函数

如果随机变量的分布函数 $F(x)$ 可表示为一非负函数 $f(x)$ 的积分:

$$F(x) = \int_{-\infty}^x f(t) dt$$

则称该随机变量为连续随机变量。 $f(x)$ 称为它的[概率]密度函数。

任何密度函数 $f(x)$ 必满足

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

注:连续随机变量的概率分布称为连续分布。

5. 独立

若两组随机变量有联合概率分布,其任何一组的条件分布都不随另一组的取值而变化,则称它们是独立的。否则称它们是相依的。

在产品质量检验中,广泛应用独立这一术语。

两次测试独立(两次重复测定)的意思是它们不相依,即它们不相互依赖,一次测试结果的取值与另一次测试结果取值无关。

独立的对立术语是相依,相依为不独立。在产品质量检验中如统计推断、回归分析、方差分析、正交试验、协同试验、能力验证、不确定度评定等统计技术应用时,都要求测试必须相互独立,否则以上的试验结果的有效性就难以令人信服。

6. 期望

对以概率 p_i 取值 x_i 的随机变量 X ,其期望的定义为:

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

式中 $\sum$ 是对 X 的所有可能的值 x_i 求和。

对密度函数为 $f(x)$ 的连续随机变量 X ,其期望的定义为:

$$E(X) = \int_{-\infty}^{\infty} x f(x) dx$$

随机变量的期望也称为它的概率分布的期望。同义词:均值(mean)、期望值或数学期望。

期望是随机变量的最重要的位置特征,它表示随机变量的平均大小和位置,期望越大,随机变量的平均愈大,位置愈靠数轴的右边。它与算术平均值在概念上有严格的区别,期望是对随机变量或分布或总体而言,而算术平均值是对样本而言。样本算术平均值对期望来说是一个无偏、一致、有效的估计值。

在产品质量检验中样本的算术平均值是对产品总体质量平均水平的评价。

7. 方差与标准差

方差 随机变量 X 的方差定义为:

$$V(X) = E[(X - E(X))^2]$$

它亦称为 X 的分布方差。

标准差 方差的正平方根 $\sqrt{V(X)}$ 。

标准差是重要的分散度量,它是方差的正平方根。标准差也称标准偏差。又称均方根,因为它是偏差 $X - E(X)$ 的平方平均再开方。

偏差 $X - E(X)$ 有正有负,但平方后恒正,再平均开方后取正值,它是各偏差的均方根值,代表了各个偏差而不计其正负,且能将绝对值大的偏差显露出来。

标准差愈大,随机变量愈分散,标准差愈小,随机变量愈集中。

在产品质量检验中,方差和标准差是度量产品质量变异性(稳定性、均匀性、一致性、分散性)的最佳指标。方差由于它具有可加性和传递性,因此,在样品的抽取中得到了广泛的应用。

8. 分位数

对随机变量 X , 满足条件 $P(X < x_p) \leq p$ 和 $P(X \leq x_p) \geq p$ 的实数 x_p 称 X 的或其分布的 p 分位数。

注: ① p 分位数可以不唯一。

② $x_{0.75}$ 与 $x_{0.25}$ 分别称为上、下四分位数。

分位数中比较重要的有以下两个:

中位数 $x_{0.5}$ 常用作集中性指标。

$x_{0.75}, x_{0.25}$ 在 $[x_{0.75}, x_{0.25}]$ 区间内有这组数据的一半, 而且 $x_{0.75} - x_{0.25}$ 是 50% 数据的极差, 所以也是一个离散性度量的指标。

为做假设检验, 常需临界值, 将它和统计量比较, 以做出假设为真或非真的判断, 临界值就是分位数。取定一个小概率 α (如 $\alpha = 0.05, 0.01$ 等), x_α 或 $x_{1-\alpha}$ 常用作单侧假设检验的临界值。 $x_{\alpha/2}$ 或 $x_{1-\alpha/2}$ 常用作双侧假设检验的临界值。在产品质量检验以及实验室检验结果的准确度评价中, 显著性检验使用较多的是双侧检验, 有时也用到单侧检验。

对不同随机变量, x_p 常列于统计技术表或附于专业书刊后, 这样由 p 可查出 x_p 值。本书的后面也附有一些我们日常常用的分位数表。

9. 统计技术应用中的两类错误

既然统计技术是以概率论为理论基础, 事件的发生都有一定的概率, 都有一定的随机性。特别是在产品质量检验中, 都是以样本来推断总体, 因此统计方法在应用过程中, 都会犯两类不同性质的错误。

第一类错误即是原假设为真而被拒绝。又称之为弃真的错误, 即以真当假。在统计方法应用中, 犯第一类错误的概率, 称为弃真概率, 记为 α 。

在产品检验过程中把合格产品误判为不合格品; 在产品验收中把应该接收的批误判为拒收批; 在过程监管控制中把正常的生产过程误判为异常; 在假设检验中把没有显著性差异误判为有显著性差异等等都属于犯第一类错误。犯第一类错误往往会给生产方造成损失, 所以又称为生产方风险。

第二类错误即是原假设不真但被接受。又称之为取伪的错误, 即以假当真。在统计方法应用中, 犯第二类错误的概率, 称为取伪概率, 记为 β 。

在产品检验过程中把不合格产品误判为合格品; 在产品验收中把应该拒收的批误判为接收批; 在过程监管控制中把异常的生产过程误判为正常; 在假设检验中把有显著性差异误判为没有显著性差异等等都属于犯第二类错误。犯第二类错误往往会给使用方造成损失, 所以又称为使用方风险。

任何统计方法的应用犯错误是不可避免的。当采取措施减少犯第一类错误的概率 α 的同时, 会导致犯第二类错误的概率 β 的增加, 反之亦然。无论是弃真概率 α 还是取伪概率 β , 都可以应用概率分布的理论进行计算, 得到两类错误风险的具体数值, 而且可以采取把两类错误的总和降到最低程度。

1.4.3.2 随机变量概率分布的类型

随机变量的概率分布分为离散型随机变量的概率分布和连续型随机变量的概率分布这两大类。

离散型随机变量的概率分布主要有二点分布(0-1 分布), 二项分布、负二项分布、泊松