

 质量管理体系中**统计技术**应用指导与培训教材

·主编/王毓芳 肖诗唐

统计抽样检验 的策划与实施



中国经济出版社
CHINA ECONOMIC PUBLISHING HOUSE

质量管理体系中统计技术应用培训与指导教材

统计抽样检验的策划与实施

主编 王毓芳 肖诗唐

主审 王宗凯 郝 凤

中国经济出版社

北 京

图书在版编目 (CIP) 数据

统计抽样检验的策划与实施/王毓芳主编. —北京: 中国经济出版社, 2005.4

ISBN 7-5017-6883-8

I. 统… II. 王… III. 抽样调查 IV. C811

中国版本图书馆 CIP 数据核字 (2005) 第 022002 号

出版发行: 中国经济出版社 (100037·北京市西城区百万庄北街 3 号)

网 址: WWW.economyph.com

责任编辑: 李晓岚 (电话: 010-68353496, E-mail: lxlan@netease.com)

责任印制: 张江虹

封面设计: 谭雄军

经 销: 各地新华书店

承 印: 人民文学印刷厂

开 本: 787×1092 毫米 1/16 印张: 25.5 字数: 350 千字

版 次: 2005 年 4 月第 1 版 印次: 2005 年 4 月第 1 次印刷

书 号: ISBN 7-5017-6883-8/F·5507 定价: 50.00 元

版权所有 盗版必究 举报电话: 68359418 68319282

服务热线: 68344225 68369586 68346406 68309176

随着 2000 年版 ISO 9000 族标准的发布, 人们将更加重视企业质量管理体系建立和运行的有效性及效率。1994 年版标准中的某些不足之处, 必将通过 2000 年版标准的宣传贯彻得到完善和改进。同时, 人们对统计技术的应用在质量管理体系建立、运行及业绩改进中的作用, 将会有更加深刻的理解和认识。

应当认识到, 质量管理起源于质量检验, 质量检验作为质量管理的重要组成部分, 也必然随着质量管理的发展而发展。在高科技提出高质量要求的现代化企业生产中, 过程控制的作用固然很重要, 但没有现代化质量检验的保证, 长期、稳定地向顾客提供优质产品和服务的要求也很难实现。

在大量生产过程中以及特殊产品的检验项目(如破坏性检验)的检验过程中, 抽样检验是非常必要的。但只有在统计抽样理论指导下的统计抽样检验才具有实际意义。目前尚有一些企业依然采用落后的、不科学的、不合理的非统计抽样检验(如百分比抽样检验), 无法实现现代化质量检验的基本要求。

质量检验系统是企业质量管理体系的重要组成部分。对质量检验工作以及统计抽样检验, 若欲取得实效, 必须经过精心的策划和认真的实施。因此, 对质量检验和统计抽样检验的基本理论、策划的要求和实施的方法, 必须对相关的管理人员、工程技术人员、质量专业技术人员, 甚至于具体执行检验操作的人员, 均应进行有计划的、系统的分层培训教育。质量检验岗位的检验员应经过应知、应会培训, 考核合格持证上岗。近几年国家人事部和国家质检总局在对质量工程师的考核及继续教育中, 已经将质量检验、统计抽样检验作为重要的培训内容。

本教材的编写为企业及质量监督系统组办质量检验人员、质量检验专业技术人员、统计抽样检验国家标准宣贯等培训, 提供实用性很强的教材。同时, 由于内容比较详细, 提供了参考的培训日程要求和配备了大量实用性练习题, 也将成为企业管理人员、工程技术人员、质量检验人员自学或组织培训的教材和工作参考书。

统计抽样检验的理论自 1924 年由美国贝尔实验室的罗米格博士和道吉博士提出后，多年来在实践中不断完善和发展，而且还在发展之中。就我国已发布的统计抽样检验国家标准也已多次换版，有些国家标准还正在修订过程之中。虽然本教材力求以最新的内容提供给读者，但总有不尽人意之处。特别是由于编审人员水平有限，对教材中的不足之处以及错误，真诚期望各方面专家予以指教。

编 审 者

2005 年元月

编审人员

主编：王毓芳 肖诗唐

主审：王宗凯 郝 凤

参加编审人员（以姓名笔画为序）

王 颖	王 蓉	王 刚	王宗凯	王淑玲
王淑敏	王毓芳	王毓泉	李志秀	刘宗森
刘隶放	肖诗唐	张立新	张铁锁	张润芝
张 霞	陈清香	尚文兰	郝 凤	梁淑田

目 录

前言	(1)
----------	-------

第一章 统计抽样检验的预备知识

1. 随机事件及概率	(1)
2. 质量的统计观念及数据的分类	(3)
3. 随机变量的分布	(4)
4. 排列与组合的计算	(42)
5. 统计模拟试验 (蒙特卡罗试验) 简介	(70)

第二章 统计抽样检验的概念

1. 质量管理发展过程中的质量检验	(74)
2. 统计抽样检验理论概述	(86)
3. 统计抽样检验常用的名词术语	(100)
4. 统计抽样检验的类型	(104)

第三章 统计抽样检验的策划与实施

1. 企业质量检验系统的建立	(112)
2. 统计抽样检验的策划与实施	(149)

第四章 统计抽样检验国家标准

1. GB/T 2828.1—2003 《计数抽样检验程序第一部分： 按接收质量限 (AQL) 检索的逐批检验抽样计划》	(171)
2. GB/T 2829—2002 《周期检验计数抽样程序及抽样表 (适用于对过程稳定性的检验)》	(206)

3. 跳批计数抽样检验国家标准简介	(230)
4. 孤立批计数抽样检验国家标准简介	(239)
5. 计量抽样检验国家标准的应用	(254)

第五章 统计抽样检验与统计过程控制

1. 质量管理体系中统计技术应用的两大支柱 ——统计抽样检验与统计过程控制	(313)
2. 过程能力的概念	(315)
3. 统计过程控制的基本原理及控制工具的应用	(324)

第六章 教学日程、练习及答案

1. 质量检验员培训	(366)
2. 质量检验专业技术人员培训	(374)
3. 统计抽样检验国家标准宣传、贯彻培训	(394)

参考文献	(397)
后记	(398)



统计抽样检验的预备知识

一、随机事件及概率

1. 随机事件的概念

(1) 现象

现象是事物的表现，任何现象的发生都是有原因（条件）的，根据现象发生的因果关系不同，现象可分为“确定现象”和“随机现象”。

① 确定现象

确定现象又称为必然现象，是指当给定一组条件（原因）后，所发生的结果是确定的（必然）的。例如水在常温常压下处于液态，当给定摄氏零度和一个标准大气压的条件时，水必然会结冰（转化为固态）。所以，水的结冰属于确定现象。

② 随机现象

随机现象又称为不确定现象，是指当给定一组条件（原因）后，所发生的结果是不确定的。质量就是一种随机现象，如同一名工人、操作同一台设备、采用同一种原材料、应用完全相同的工艺方法（即生产条件完全相同）的情况下，所加工产品的质量特性值不是惟一的，这种现象称为产品质量的变异性。又如加工一批产品，由于生产过程稳定，批不合格品率是确定的，但究竟哪一件产品合格，哪一件产品不合格，是不能确定的。

随机现象中的变量称为随机变量，如产品质量特性值。

显然，随机现象具有两个特点：

- i. 随机现象的结果至少有两个；
- ii. 随机现象的结果是哪一个出现，事先不可知。

(2) 随机事件

随机事件是指随机现象某些结果（样本点）的集合。如交验批批量 $N = 50$ ，假设已知批不合格品率为 6%，则抽样检验过程中检出不合格品的件数可能为 0、1、2、3。则检出不合格品件数是随机事件，若记为事件 A，则可表示为： $A = \{0, 1, 2, 3\}$ 。其中包含有四种结果。

2. 随机事件的概率

(1) 概率的概念

概率是指事件在给定条件下发生的几率(可能性)的度量。

随机事件某种结果的发生是不确定的,带有偶然性,但这种偶然性是可度量的。在统计抽样检验过程中,人们必然关心样本中检出几个不合格品件数这样的随机事件发生的可能性,即概率。

(2) 概率的定义

根据概率定义的方法不同有古典定义和统计定义。

① 概率的古典定义

概率的古典定义确定,应给定以下三个条件:

- i. 所涉及的随机现象只有有限个样本点,假设总计有 N 个样本点;
- ii. 每个样本点出现的可能性是完全相同的;
- iii. 被考察的事件 A 发生时含有 n 个样本点。

则事件 A 的发生概率为:

$$P_{(A)} = \frac{n}{N}$$

② 概率的统计定义

概率的统计定义确定,应满足以下几个条件:

- i. 与事件 A 有关的随机现象是可以大量重复试验的;
- ii. 若在 n 次重复试验(给定条件完全相同情况下的多次试验),事件 A 的发生次数为 k ,则称事件 A 发生的频率为 $f_{(A)}$

$$f_{(A)} = \frac{k}{n}$$

- iii. 频率 $f_{(A)}$ 会随着重复试验的次数 n 的增加而趋于稳定。理论上认为只有 n 趋近于无穷大时的频率才是事件 A 的发生概率 $P_{(A)}$ 。

$$P_{(A)} = \lim_{n \rightarrow \infty} f_{(A)}$$

实际在实践工作中不可能进行无穷多次重复试验,在数理统计技术应用中认为,当 n 足够大时频率 $f_{(A)}$ 趋近于稳定时的值就是事件 A 的概率 $P_{(A)}$ 。

二、质量的统计观念及数据的分类

1. 质量的统计观念

从统计的观点看待质量，有以下两个观念：

- (1) 质量具有变异性（不均匀性）；
- (2) 质量变异具有规律性（质量数据的分布）。

2. 质量变异规律性的反映

质量变异的规律性反映在质量数据的分布上，数据的性质不同时其服从的分布也不同。

要做好一件事情，必须按客观规律办事，任何违背客观规律的行为注定要失败的。分布理论是质量变异的客观规律，要做好质量工作，包括质量检验工作就必须学习和掌握分布理论。

3. 数据的分类及分布

按数据的性质不同，对量化的数字数据可分为计量值数据和计数值数据。

(1) 计量值数据

计量值数据指在一定区间内可以连续取值，可以取无穷多个数值的数据。大多数质量特性值都属于计量值数据，如长度、面积、体积、重量、强度、电压、电流等。

计量值数据服从正态分布。

(2) 计数值数据

计数值数据指在一定区间内只能间断取值，只能取有限个数值的数据。计数值数据可以理解为“数个数”的数据。因此，计数值数据为自然数。根据计数对象不同，计数值数据又可以分为计件值数据和计点值数据。

① 计件值数据

计件值数据是对成件的单位产品计数的结果，如人数、产品件数、合格品数、不合格品数等。

计件值数据服从二项分布。

② 计点值数据

计点值数据是对缺陷的计数结果，如钢板上的划痕、布匹上的疵点、造纸或印刷过程

的断头等。

计点值数据服从泊松分布。

三、随机变量的分布

1. 随机变量

量可分为常量和变量。

常量是取固定数值的量，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。

随机变量是变量中的一种类型，随机变量的数值变化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能预测的。

质量是一种随机现象，反映质量的数据就是一种随机变量。因此，对所研究的质量特性值（数字数据）常称为随机变量 X ，随机变量 Y 等。

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。

2. 分布的概念

（1）质量的统计观念（回顾）

① 产品质量具有变异性（不一致性）

质量是一种随机现象，由于影响产品质量的因素（人、机、料、法、环）无时无刻不在变化着，所以产品质量具有变异性。例如，要工人在车床上加工 100 根 $\phi 20\text{mm}$ 的圆轴，要求每一根圆轴的直径必须保持恰好为 20mm，恐怕任何技艺高超的工人也无法完成这样的加工任务。必然提出，应当给一个范围（公差）。如果下达的任务为车制 100 根 $\phi 20 \pm 0.1\text{mm}$ 的圆轴，则任何一名熟练的车工都可以完成。

② 产品质量的变异具有规律性（分布）

产品质量的变异不是漫无边际的变异，是在一定范围内而且符合一定规律的变异。

如上面给出的事例，一名好的工人，必然想到，下达的加工指令所要求的是 20mm，

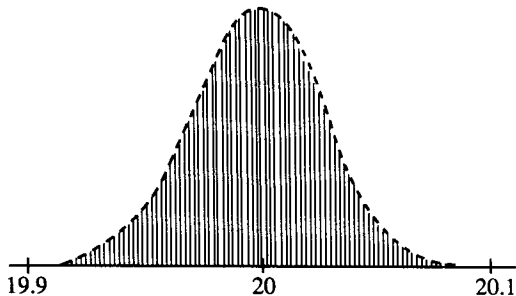


图 1-1 随机变量的分布

因为产品质量的变异性而不可能保证每一件产品都是 20mm，迫不得已才给出的公差范围。因此，对所测轴径得到的 100 个质量数据并不是在 19.9mm 至 20.1mm 范围内平均分配，而是在 20mm 处出现的数据最多，偏离 20mm 越远数据出现得就越少。图 1-1 反映的以上情况，实际就是质量数据（随机变量）的分布。

(2) 频数直方图

受图 1-1 反映现象的启示, 找出频数与质量特性值之间的关系就可以了解质量特性变异的规律性 (分布)。

例如某食用植物油生产厂，对灌装溢出量的统计，得到数据表（表 1-1）、频数分布表（表 1-2）和频数直方图（图 1-2）。显然，所做频数直方图的图形形状，基本上符合一种分布（正态分布）。如果我们取得无穷多数据，分组数为无穷多组时，所做的频数直方图就是质量数据（随机变量）的分布图形。

表 1-1

溢 出 量 (克)									
47	32	44	35	20	19	37	31	25	34
19	11	16	11	44	29	29	42	59	38
28	12	45	36	25	40	35	11	38	33
15	30	12	33	26	35	44	32	11	38
27	37	26	20	35	23	45	26	37	32
28	44	40	31	18	31	25	24	32	22
22	37	19	47	14	37	32	12	38	30
25	40	24	50	19	7	31	23	18	32
38	0	41	40	37	35	12	29	48	20
31	20	35	24	47	12	27	38	40	31
52	42	52	24	25	20	31	15	3	28
29	47	41	32	22	28	27	22	32	54
42	34	15	29	21					

$$N = 125$$

表 1-2

频数分布表

组号	分组界限	组中值	频数记录	f_i	u_i	$f_i u_i$	$f_i u_i^2$
1	-0.5~4.5	2	┐	2	-6	-12	72
2	4.5~9.5	7	—	1	-5	-5	25
3	9.5~14.5	12	正正	10	-4	-40	160
4	14.5~19.5	17	正正	10	-3	-30	90
5	19.5~24.5	22	正正┐	16	-2	-32	64
6	24.5~29.5	27	正正正正	20	-1	-20	20
7	29.5~34.5	32	正正正正—	21	0	0	0
8	34.5~39.5	37	正正正正┐	19	1	19	19
9	39.5~44.5	42	正正正┐	14	2	28	56
10	44.5~49.5	47	正┐	7	3	21	63
11	49.5~54.5	52	┐┐	4	4	16	64
12	54.5~59.5	57	—	1	5	5	25
$h = 5, x_0 = 32$				$N = 125$		$\sum f_i u_i$ = -55	$\sum f_i u_i^2$ = 658

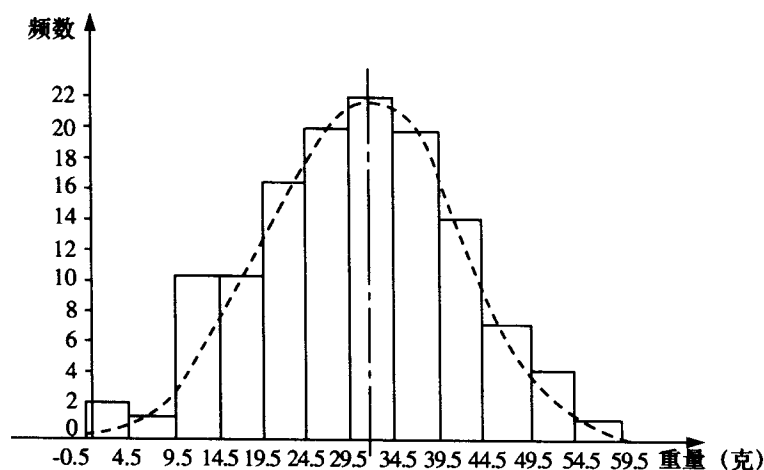


图 1-2 溢出量直方图

3. 计量值数据的正态分布

(1) 正态分布的图形——正态分布曲线

一般属于计量值数据的质量特性值服

从如图 1-3 所示的正态分布。

从图 1-3 所示的正态分布曲线可见：

① 正态分布曲线如同扣放的一口钟，所以又称为钟形曲线。

② 正态分布曲线在 $x = \mu$ 处有对称轴，且有最大值（最大频数）。

③ 正态分布曲线以 x 轴为渐近线，频数 $f(x)$ 永远为正值。

④ 正态分布曲线的拐点（凸曲线与凹曲线的交点）到对称轴的距离为 σ 。

⑤ 正态分布曲线向 $\pm \infty$ 无限延伸。

(2) 正态分布的密度函数

在平面直角坐标系中的曲线是某一函数的图像。正态分布曲线是正态分布密度函数的图像，二者是一一对应的。

正态分布的密度函数表达了频数 $[f(x)]$ 与随机变量 (x) 之间的关系。

$$f(x) = \frac{1}{\sigma \sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

式中： $f(x)$ 为频数；

x 为随机变量（质量特性值）；

π 为圆周率；

e 为自然常数；

μ 称为分布中心；

σ 称为标准差。

在正态分布密度函数中， π 和 e 为常量，不影响频数 $f(x)$ 与随机变量 x 的关系。而 μ 和 σ 为变量，所以会影响频数与随机变量之间的关系。这种影响反映在正态分布曲线的形状及其在平面直角坐标系中的位置。

图 1-4 反映了分布中心 μ 的影响，显然分布中心 μ 表征了质量特性值分布中心的位置。

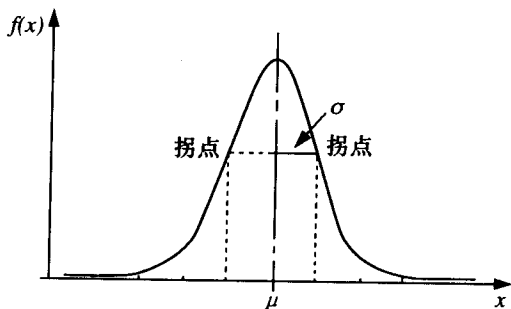


图 1-3 正态分布曲线

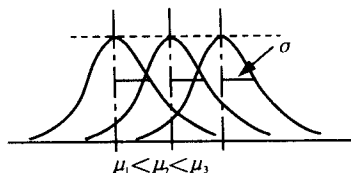


图 1-4 μ 值不同、 σ 值相同的正态分布曲线

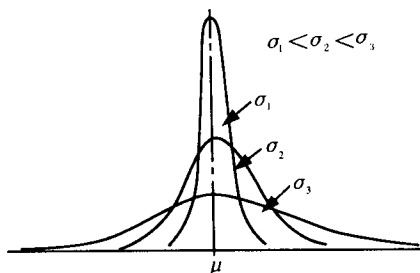


图 1-5 μ 值相同、 σ 值不同的正态分布曲线

图 1-5 反映了标准差 σ 的影响，显然标准差 σ 表征了质量特性值的离散程度。

分布中心 μ 和标准差 σ 称为正态分布的特征值（分布参数）。在研究质量特性时，其特性值分布的特征值 μ 和 σ 反映了质量波动的状况。因此分布特征值是质量保证、质量控制和质量改进工作的研究对象。

（3）正态分布的概率分布

概率即可能性或几率，关于概率的概念将在第二章中讨论。

从理论上讲，正态分布曲线是向 $\pm \infty$ 无限延伸的。但在实际工作中质量特性值的取值总是一个有限的范围。在质量工作中，要求工程技术人员及有关的管理干部必须了解质量变异的幅度及发生这种幅度的质量变异的概率。因此，研究正态分布的概率分布就非常有必要。如在 $[x_1, x_2]$ 区间内质量数据占全部质量数据的百分数，实际就是在 $[x_1, x_2]$ 区间内的正态分布的概率。

正态分布的密度函数 $f(x) > 0$ ，永远取正值，所以有

$$\int_{-\infty}^{\infty} \frac{1}{\sigma \sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx = 1,$$

在 $[x_1, x_2]$ 区间的正态分布概率为

$$\int_{x_1}^{x_2} \frac{1}{\sigma \sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx,$$

一般对正态分布要进行三种计算：

① 概率计算

概率计算是在已知随机变量区间 $[x_1, x_2]$ 的情况下，通过积分计算求概率值。

② 密度函数值计算

密度函数值计算是求与已知随机变量值(横坐标值)相对应的密度函数值(纵坐标值)。

③ 分位数计算

分位数计算是在已知概率值的情况下求与之对应的随机变量值(单侧)或区间(双侧)的计算。

对任何企业而言,所能见到的质量特性值的正态分布将会有千千万万,甚至无穷多个。面对无穷多个正态分布是无法计算的,即使对某一种质量特性值的一个正态分布而言,上述计算也是很复杂的。

为了研究方便,需要对正态分布进行标准变换,把千千万万个正态分布转换为一个正态分布——标准正态分布。

① 正态分布的标准变换

若随机变量 X 服从正态分布,其分布中心为 μ ,标准差为 σ ,可记为: $X \sim N(\mu, \sigma)$ 。

对随机变量 X 的每一个数值 x_i 做如下的变换

$$u_i = \frac{x_i - \mu}{\sigma}$$

则随机变量 U 服从标准正态分布,可记为 $U \sim N(0, 1)$ 。

标准正态分布的分布中心为“0”,标准差为“1”。

标准正态分布的分布曲线图形如图 1-6 所示。

标准正态分布的密度函数为

$$\varphi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$$

对标准正态分布的概率计算,可以通过计算机将计算结果作成数学用表——正态分布表具体计算时通过查正态分布表简化计算过程,对计算结果再利用数学变换进行还原。

对标准正态分布的概率计算表示为

$$\Phi[x_1, x_2] = \int_{x_1}^{x_2} \varphi(x) dx = \int_{x_1}^{x_2} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx,$$

符号 Φ 表示查正态分布表。

正态分布表见表 1-4。

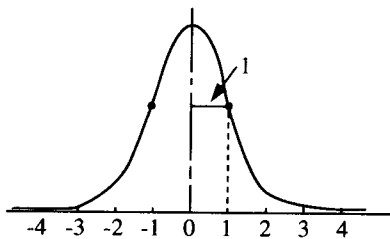


图 1-6 标准正态分布曲线