国家标准《橡胶与橡胶制品 统计学在物理试验上的应用指南》征求意见稿编制说明

一、任务来源

根据《国家标准化管理委员会关于下达2021年第一批推荐性国家标准计划及相关标准外文版计划的通知》（国标委发〔2021〕12号文），由全国橡标委归口组织制定国家标准《橡胶与橡胶制品 统计学在物理试验上的应用指南》，计划项目编号20211093-T-606。

计划下达时，为等同采用ISO19003：2006《橡胶与橡胶制品 统计学在物理试验上的应用指南》，项目周期为18个月。在标准转化过程中发现，该项目所采用的国际标准ISO 19003:2006,标准内容体量较大，并在消化其标准内容的过程中发现了大量的问题，如公式表达、示例计算上的一些错误，这些都需要论证并请教统计学专业人士。但由于疫情的原因，无法进行走访调研、线下会议研讨；同时也影响了试验验证的开展。于2022-05-30提交项目延期申请，申请延期6个月。2022-06-06国标委批准项目延期。整个项目周期30个月。

也由于ISO19003：2006存在的问题较多，故将等同采用调整为修改采用。

二、制定本标准的意义

统计学在橡胶试验中的应用越来越多，越来越广。实践中不仅用于样品的检验，而且用于产品工艺控制和质量控制；还用于新研发的产品的技术指标的确定。

统计学方法在试验过程的各个阶段，从实验的设计到结果的解释，都具有重要的作用。因此，介入试验的人需要对统计学原理和所需使用的统计学技术的知识有一个基本的了解。有许多教科书和标准讲述统计学方法，但是，有一本易于检索最常用的方法和公式、且还能考虑其对各种橡胶试验方法特定应用的指南，是很方便的。本文件的制定是对统计学通用标准和橡胶试验方法标准两者的补充。

目前，在橡胶与橡胶制品领域ISO/TC45已发布了4项与统计学相关的标准：

ISO 19003:2006（2019）《橡胶与橡胶制品 统计学在物理试验上的应用指南》

ISO 19983:2022 《橡胶 试验方法精密度的确定》

ISO/TR 12134:2010 《橡胶 试验方法不确定度的估算 非函数参数》

ISO/TR 9272:2005（2018废止）《橡胶制品 试验方法标准精密度的测定》

其中，ISO/TR 9272:2005于2018年因发布ISO 19983:2017而废止。

本文件所转化的国际标准ISO19003:2006《橡胶与橡胶制品 统计学在物理试验上的应用指南》，是在本领域内适用面最广的一个文件。它给出了在各种橡胶物理试验中如何利用统计学判定试验结果在给定的置信水平是否符合产品规范要求的方法和指南。

本文件并非意在与现行涉及统计学技术的国际标准相矛盾或将其替代，而是要提供补充，并给出这些技术应用于具体橡胶试验情况的示例。

本文件编写的方法是，针对每个主题，分别介绍原理、方法和在橡胶试验上的应用。在原理下，概述基本概念。方法考虑能够使用的统计学技术；给出基本程序和公式。如果适用；对于不常用的方法或更先进的处理，则引用其他出版物。“在橡胶试验中的应用”指这些方法如何、何处可以应用，并给出针对橡胶性能和试验的示例。

本文件所涉及的内容包括：术语和定义、符号、试验结果的限制、结果的分布和集中趋势的度量、置信限和显著性差异、排序法、拒绝异常值的标准、方差分析、回归分析、不确定度的测量、抽样、试样数量、结果表示、精密度陈述、实验设计和统计学质量控制等。

本文件符合《中国制造2025》第三章"战略任务和重点"第四条“加强质量品牌建设”中提出的“加快提升产品质量。组织攻克一批长期困扰产品质量提升的关键共性质量技术，加强可靠性设计、试验与验证技术开发应用”，属于基础通用项目。

三、主要工作过程

1、组成起草工作组

接到任务用，由全国橡标委秘书处组织协调组成标准起草工作组，主要起草单位为：

主要起草人：

2、翻译国际标准、把握专业上的准确性

在标准起草过程中，随着对标准的理解，不断地对标准的术语、用词进行校对修改，力求专业上的准确性。

3、计算验证标准中各计算公式、计算结果的正确性，对错误的进行修正。

4、走访、请教统计学专业人士

咨询统计学专业人士

5、查阅统计学相关标准、书箱

* 1. 贾俊平编著，统计学基础（第四版），中国人民大学出版社，2016.
  2. 贾俊平等编著，统计学（第7版），中国人民大学出版社，2018.
  3. 陈珍珍主编，统计学（第六版），厦门大学出版社，2018
  4. 孙淑萍编，误差与数据处理，辽宁大学化学系，1989.
  5. GB/T 3358.1-2009 统计学词汇及符号 第1部分：一般统计术语与用于概率的术语
  6. GB/T 3358.2-2009 统计学词汇及符号 第2部分：应用统计
  7. GB/T 3358.3-2009 统计学词汇及符号 第3部分：实验设计
  8. GB/T 3359-2009 数据的统计处理和解释 统计容忍区间的确定 (ISO 16269-6:2005,IDT)
  9. GB/T 3361-1982 数据的统计处理和解释 在成对观测值情形下两个均值的比较（ISO 3301:1975，NEQ）
  10. GB/T 4087-2009 数据的统计处理和解释 二项分布可靠度单侧置信下限
  11. GB/T 4088-2008 数据的统计处理和解释 二项分布参数的估计与检验
  12. GB/T 4089-2008 数据的统计处理和解释 泊松分布参数的估计和检验
  13. GB/T 4882-2001 数据的统计处理和解释 正态性检验(ISO 5479:1997,IDT)
  14. GB/T 4883-2008 数据的统计处理和解释 正态样本离群值的判断和处理
  15. GB/T 4889-2008 数据的统计处理和解释 正态分布均值和方差的估计与检验(ISO 2854:1976,MOD)
  16. GB/T 6380-2019 数据的统计处理和解释 Ⅰ型极值分布样本离群值的判断和处理
  17. GB/T 8055-2009 数据的统计处理和解释 Г分布（皮尔逊Ⅲ型分布）的参数估计
  18. GB/T 8056-2008 数据的统计处理和解释 指数分布样本离群值的判断和处理
  19. GB/T 10092-2009 数据的统计处理和解释 测试结果的多重比较
  20. GB/T 14838-2009《橡胶与橡胶制品 试验方法标准精密度的确定》（Idt ISO/TR 9272:2005）
  21. GB/T 17560-1998 数据的统计处理和解释 中位数的估计（ISO 8595:1989, EQV）
  22. ISO 2602《试验结果的统计学解释——均值的估计——置信区间》，
  23. I SO 2854《数据的统计学解释——与均值和方差相关的估计和检验技术》
  24. ISO 19983：2020《橡胶 试验方法精密度的确定》
  25. Alan G. Veith，A New Approach to Evaluating Inter-laboratory Testing Precision，Polymer Testing，1993，12，113-184

6、征求意见稿

于2013年元月形成征求意见稿，公开征求意见。

——发全国橡标委全体委员

——相关专业人员

——国家标准委网站、全国橡标委网站。

征求意见期：二个月。

7、送审稿

8、标准审查形成报批稿

四、标准编写原则

本文件采用ISO19003:2006，主要是介绍统计学在橡胶物理试验上的应用，对如何利用统计学的方法分析试验方法、试验结果的准确性、正确性给予指导。涉及大量的统计学运算和计算结果的使用。所以本文件的编制原则是：

——遵循GB/T 1.1—2020《标准化工作导则 第1部分：标准化文件的结构和起草规则》的规则，使标准在结构、语言表述和编排格式上符合GB/T 1.1的要求；

——遵循GB/T 1.2-2020《标准化工作导则 第2部分：以ISO/IEC标准化文件为基础的标准化文件起草规则》采用规定；

——努力保证词汇符合统计学术语或用语；

——验算ISO19003中给出的统计学实例，确保公式正确，计算正确

——进行适当的试验，并按ISO19003示例的方法进行验算。

五、主要技术内容及对ISO19003相关内容的修正

（一）本文件的结构

本文件一共有18章、9个资料性附录和一个参考文献，界定了27个术语（第3章）、41个符号（第4章）。本文件的主要结构目录如下：

前言

1 范围

2 规范性引用文件

3 术语和定义

4 符号

5 试验结果的限制

6 结果的分布和集中趋势的度量

7 置信限和显著性差异

8 排序法

9 拒绝离群值的标准

10 方差分析（ANOVA）

11 回归分析

12 不确定度的测量

13 抽样

14 试样数量

15 结果表示

16 精密度陈述

17 实验设计

18 统计学质量控制

附录A（资料性） 本文件中引用的分布函数的数学形式

附录B 资料性） 平均值的其他形式

附录C 资料性） 在双指数和威布尔分布中集中趋势度量的相互关系

附录D（资料性） 计算标准差的公式6

附录E（资料性） 构建威布尔概率纸

附录F（资料性） 计算学生t值的公式

附录G（资料性） 方差分析

附录H（资料性）计算回归系数的方程

附录I（资料性） Intercal介入法

参考文献

（二）第5章 试验结果的限制

这一章介绍了试验结果的变异性，准确度、正确度和精密度，相关性和显著性的基本概念。

a) 所有测量都可能受变异影响。有必要知道变异的来源，并对其大小做出可靠的估计。变异可能来源于混炼程序、配合剂中的任何变化，不同的实验室、不同的仪器，操作者本自身的波动等。根据这种信息，应该可以判断结果的可靠性，从而判断其不确定度和显著性。

b) 准确度是试验结果与可接受的参照值之间的一致性（见3.14），而正确度是大量试验结果平均值与真值或可接受的参照值之间的一致性（见3.15）。另一方面，精密度是试验结果之间的一致性(见3.16)，与可能存在的任何参照值无关。

c) 并非所有的试验都能很好的表征产品的性能，有些试验比另外一些试验更具相关性。“显著性”一词有时用来表示相关性，并适用于实际试验或所测量的性质，但在本文件中，显著性用于统计意义上，例如，一种材料的强度显著高于另一种材料。

（三）第6章 结果的分布和集中趋势的度量

试验结果的分布，是指将试验结果按大小绕其平均值左右进行排列，这些试验结果所呈现出的一种排列状态。通常这种结果的分布可以用特定的数学定律表示，如正态分布。

绘制成曲线，则可从形状得到关于趋向中心值的趋势、结果离散的程度以及偏离中心超过一定程度的比例的有用度量。

本章主要描述介绍了分布函数、集中趋势的度量、离散的度量，不直接服从正态分布试验结果转换为正态分布，以及偏离正态的检验。

1、 分布函数

本文件描述了三种分布，正态分布、双指数分布和威布尔分布。

正态分布（也称高斯分布）函数，其应用最广泛，它可以用两个参数完整的表征，即均值*μ*（见3.4）和标准差*σ*（见3.9），其分布图形（密度分布）为钟形。均值*μ*决定了中心值的位置，标准差*σ*决定了图形中黄线的陡峭程度。当*σ*大时，曲线趋于平缓；当*σ*小时，，曲线趋于陡峭。

有必要说明的是，在6.2.1.1正态分布中z和Z的区别。在标准中并未给予说明，但根据相关书箱中，大写Z似应为经标准化后的变量，即将均值转换为0，将标准差转换为1后的变量。应称之为标准化变量Z。

设*X*～*N*(*μ,σ*2)，则：

在第4章中，Z定义为“假设检验中的Z分数”

本章对双指数分布和威布尔分布仅做了简单介绍。

2、 集中趋势的度量

即使在最仔细的实验条件下在相同材料上进行重复测量也出现结果分散。但当数据按数字顺序排列时典型值趋于出现在数据的中段，这叫做集中趋势。集中趋势是指一组数据向某一中心值靠拢的程度，它反映了一组数据中心点的位置所在。

集中趋势的度量包括：平均值、中位数、众数。

1) 平均数

平均数，也称均值，英文对应的是mean，是最常见的集中趋势度量。平均值包括算术平均值、几何平均值等。算术平均值是常见的一种，因此常常省略“算术”二字。在本文件中，除另有说明外，平均值即为算术平均值。

总体的均值用*μ*表示，取自总体的样本的均值用‾*x*表示，对总体平均值*μ*的估计平均值用*μ*ˆ表示。

2)中位数

一组数据排序的后处于中间位置上的数据值，称为中位数。如果数据的个数为偶数时，则取两个中间值的平均值。

中位数的好处是使用方便，当测量结果为3或5个时，不需计算就可以判断而取值。对于预期适用双指数分布函数的拉伸强度或断裂伸长率，其试验结果的取值，现在已应用于GB/T 528中。

3)众数

众数是一组数据中出现频率最多的数值。一般情况下，只有数据量较大时众数才有意义。从分布的角度看，众数是一组数据分布的最高点所对应的数值。如果有两个或多个最高峰点，也可以有两个或多个众数。

虽然众数在实践中很少使用，但是为了完整，本文件中还是给出了其介绍。

4）三种集中趋势度量的关系

均值、中位数和众数是描述数据集中趋势的三个主要统计量。平均数是最易被多数人所理解和接受的，实际中应用的也多，但其主要缺点是易受极端值的影响，对于严重偏态分布的数据平均值代表性差。中位数和众数提供的信息不像平均数那样多，但它们也有优点，如不受极端值影响，具有统计上的稳定性。

当数据为对称分布时，平均数、中位数和众数是相等的，即：

平均值=中位数=众数

3、离散的度量（6.2.3）

数据的分散程度是数据分布的另外一种重要特征，它反映的是各变量值远离其中心值的程度，因此也称为离中趋势。数据的离散程度越大，集中趋势的度量值对该组数据的代表性就越差，离散程度越小，其代表性就越好。

描述数据离散程度度量值有标准差、方差、极差、四分位数、离差、变异系数等。本文件只介绍了标准差、极差和变异系数。在介绍标准差的同时，还介绍了标准误差。

1）标准差（6.2.3.2.1）

标准差是指一组数据中的每一个数据与组平均数之差的平方的均值的平方根。也就是用每个数据与组均值相减再平方，然后求和；这个和再除以数据的个数，所得商称为方差，方差再开平方所得值称为“标准差”。概括地说，各变量与其平均数离差平方的平均数称为方差；方差的平方根称为标准差。

与方差不同，标准差是具有量纲的，量纲与变量值的计量单位相同，因此，标准差实际意义要比方差更清楚，也更多地被使用。

对总体标准差*s*的估计，按下式计算：

而实际统计检验时通常使用的是对取自总体的样本的标准差*s*′估计:

ISO 19003中公式（5）和公式（6）恰恰弄反了。已予以更正。即，基于样品的标准差分母中用*n*-1；基于总体的，分母中则用*n*，而不是*n*-1。

2) 标准误差（6.2.3.2.3）

标准差除以观测次数的平方根所得商称为标准误差（standard error）：

标准误差是对各观测值在预期值周围分散程度的一种度量。

3）极差（6.2.3.3）

极差，英文词对应的是range(范围)，表达的是样本观测值中最大值与最小值之差，也称全距。是描述数据离散程度的最简单的度量，用于不太精准的估计。可以用极差来估计标准差，即将极差乘以一个因子来估计标准差。

4）变异系数（6.2.3.4）

一组数据的标准差与其相应的平均数之比，称为变异系数，也称为离散系数。

标准差反映的是数据差异水平的绝对值。一方面，标准差的大小受原始数据的影响较大，绝对值大的，标准差就大，绝对值小的，标准差就小；另一方面，标准差与原始数据的计量单位相同，采用不同计量单位计量的数据，其标准差的值也就不同。因此，对于不同组别的数据，如果原始数据的绝对值相差较大或计量单位不同时，就不能用标准差直接比较其离散程度，这时需要计算变异系数。

*C*v = (*s* / ‾*x* ×100)

4、转换为正态分布（6.2.4）

由橡胶试验获得的观测结果可能不会直接符合正态分布函数，这样就需要将其通过某种运算将其转换为符合正态分布的形式。这需要掌握一些描述数据的分布函数的相关知识。

本文件在这里简单地给出了几种常用的转换方法，如对数、开方或求倒数。

5、偏离正态的检验（6.2.5）

本条介绍了用对数-正态纸、威布尔(Weibull)分布纸以及双指数分布作图检验分布形式的简单快速的方法。

这里有一句话“对数-正态纸、威布尔(Weibull)纸也方便得到”，但确实不好买，所以将其删掉了。好在，这些方法都可以在excel是完成。

其中，6.2.5.2 双指数分布中：

c) 中位数强度，对应于横坐标-0.3665的值；

d) 均值，对应于横坐标-0.5772的值。

-0.5772是欧拉常数，附录C。

6、应用示例（6.3）

本条主要的基于6.2所介绍的方法给出应用的示例。

1）拉伸试验（6.3.2）

本示例的目的是用表5的数据计算均值、标准差、均值的标准误差，以及中位数。

计算步骤如下：

a) 按6.2.2.2计算均值：‾*x* =（∑*x*）/*n*

b) 按6.2.3计算标准差和均值的标准误差：

和 ，其中，*n*=12

c)用a)和b)得到的值按C.1计算（即，双指数分布）“计算出的中位数”：

即：中位数=均值+0.164 28*σ*

标准表6中，混炼胶A计算出的中位数为25.9，不是原文中的26.6。

中位数=均值+0.164 28*σ* =25.8+0.16428×0.46=25.8755688=25.9

通过中位数分布图（图5）综合判断，应为文献中数据录入错误。据此，更正标准表6中相应的数据（见表1）。

表1 标准中表6 拉伸强度测量的计算值

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 混炼胶 | 平均值 | 标准差 | 平均值的标准误差 | 观察到的  中位数 | 计算出的  中位数 |
| A | 25.8 | 0.46 | 0.13 | 25.8 | 25.9 |
| B | 26.4 | 1.24 | 0.36 | 26.4 | 26.6 |
| C | 17.6 | 1.99 | 0.57 | 18.3 | 17.9 |

2）疲劳（6.3.3）

按ISO 6943《硫化橡胶 拉伸疲劳的测定》用10个试样进行试验，得出一组疲劳寿命的数据（表7）。

该示例证明了疲劳寿命服从威布尔分布（图7）。当然这里涉及了第11章回归分析的内容。

3）转换为正态分布（6.3.4）

将标准表8中的电阻率数据按升序排列并计算概率纸上横坐标Pm，形成表2：

表2 标准表8数据升序排序

|  |  |  |  |
| --- | --- | --- | --- |
| 序号 | 电阻率  Ω·cm | log（电阻率） | Pm=100\*m/(n+1) |
| 1 | 3.54×108 | 8.55 | 16.67 |
| 2 | 2.75×109 | 9.44 | 33.33 |
| 3 | 1.20×1010 | 10.08 | 50.00 |
| 4 | 2.68×1010 | 10.43 | 66.67 |
| 5 | 2.81×1011 | 11.45 | 83.33 |

以Pm为横坐标、电阻率为纵坐标作图，得正态分布图。如果数据线呈直线，则为正态分布，否则为非正态分布。很明显，图1不是一个直线。当用log(电阻率)对Pm作图时，则形成一近似的直线（图2）。与标准中6.2.5.1.3所述相符，经转换后服从正态分布。

图1 电阻率在概率纸上对Pm作图

图2 对数电阻率对Pm作图

（四）第7章 置信限和显著性差异

1、置信限和置信区间（7.2.1）

a) 7.2.1 介绍了置信限和置信区间的计算，以及t-分布表的使用。其中，

对于标准中公式（14）*C*I = 2(*tα s*)/ 中符号*C*I的说明，应为置信区间长度，而不是置信水平。即应为：

*C*I =置信区间长度。

如果是置信水平，则*C*I =1-α而不是*C*I = 2(*tα s*)/ 。而置信区间的表达为：

[‾*x* – (*tα s*)/，‾*x* + (*tα s*)/]

也可写成：‾*x* ± (*tα s*)/

b) 在置信限的计算中都用到了*t*分布。值得指出的是，本标准中表11与*t*分布表在第一列上是不一致主的，*t*分布的第一列是自由度*df*，而表11是观察数*n*。本标准中自由度df与观察数n的关系为：

*df*=*n*-1

3) 7.2.1.2中讨论标准差时，讨论了**标准差***s*的置信限，并给出了**标准差**的置信上限和置信下限的计算公式，即标准中的式（16）和（17）：

和

但对符号的说明有误：

——是*s*的上置信水平，应为置信上限；

——是*s*的下置信水平，应为置信下限。

应特别指出的是，标准中表13的第1列给出的应为自由度*df*，用*n*表示会产生误解，所以根据GB/T 4086.2-1983《统计分布数值表 χ2分布》予以了更正。

2、显著性差异（7.2.2）

显著性，又称统计显著性（Statistical significance）， 是指[零假设](https://baike.baidu.com/item/%E9%9B%B6%E5%81%87%E8%AE%BE/8078898)为真的情况下拒绝零假设所要承担的风险水平，又叫概率水平，或者显著水平。 [1]

显著性的含义是指两个群体的态度之间的任何差异是由于系统因素而不是偶然因素的影响。我们假定控制了可能影响两个群体之间差异的所有其他因素，因此，余下的解释就是我们所推断的因素，而这个因素不能够100%保证，所以有一定的概率值，叫显著性水平（Significant level）。

总的来说，它表示群体之间得以相互区别的能力。在统计[假设检验](https://baike.baidu.com/item/%E5%81%87%E8%AE%BE%E6%A3%80%E9%AA%8C)中，公认的[小概率事件](https://baike.baidu.com/item/%E5%B0%8F%E6%A6%82%E7%8E%87%E4%BA%8B%E4%BB%B6/7339473)的[概率值](https://baike.baidu.com/item/%E6%A6%82%E7%8E%87%E5%80%BC/12726566)被称为统计假设检验的[显著性水平](https://baike.baidu.com/item/%E6%98%BE%E8%91%97%E6%80%A7%E6%B0%B4%E5%B9%B3/1383148),对同一量，进行多次计量，然后算出平均值。

1）利用两组数据的均值判断显著性差异（7.2.2.2）

这里给出了判断两组数据是否具有显著性差异的一种方法，即用均值进行比较。在给定水平下比较两组数据平均值之差是否小于t检验值与加权标准误差的乘积*t*α*S*。如果是，则没有显著性差异；否则，有显著性差异。

公式（18）有误，应为|‾*x*1 -‾*x*2|＜*t*α*S*

对于公式（18），|‾*x*1 -‾*x*2|＜*t*α*S* 中符号*t*α的说明，“*t*α是双侧学生*t*检验值(表11)，自由度（*n*1+ *n*2 - 2），对于置信水平95%的是97.5，对于置信水平99%的是99.5”。对照7.2.1.2.3以及表11，明显应为“……对于置信水平95%的*α*是0.975，对于置信水平99%的*α*是0.995”。

2）利用两组数据的标准差判断显著性差异（7.2.2.3）

该方法是利用两组数据的标准差*s*1和*s*2的比值，通过查F检验表，对比判断。即，如果计算值大于表中临界值，那么这两个标准差在选定的置信水平下是有差异的：

*F*计算=*s*1/*s*2 >*F*临界，式中*s*1>*s*2。

3、应用示例

1）置信限和指标限值（7.3.2）

标准中7.3.2给出了利用置信限判断三个混炼胶的压缩应力松弛性能是否符合产品标准指标要求的示例。但在此处存在着两处错误：

错误1：根据ISO 3384和GB/T 1685-2008,橡胶混炼胶的压缩应力松弛是施加以初始应力*F*0，经过一段时间*t*后，应力下降百分率来计算的，即：

压缩应力松弛百分率*R*(*t*)的数值是随着时间而增大的。所以，对该性能的要求通常为不大于某个值（如，不大于20%），也就是规定上限，而不规定下限。

因此，ISO 19003中7.3.2提出计算“下限”是不正确的，应计算上限。

错误2：计算错误，计算公式（22）应为计算上限的公式，表17也应为上限的计算结果。

据此，对计算的示例和表17进行了修改。具体修改如下：

“每个混炼胶在95%显著性水平的置信下限用式（22）计算：

*c*L = *x* - (*t*0.95*s*)/ （22）

因此，对于*s*为0.36的混炼胶1的值为：

*c*L = 22.5 - (2.92 × 0.36/)

*c*L = 21.9”

修改为：

“每个混炼胶在95%显著性水平的置信上限用式（22）计算：

*C*u = *x* + (*t*0.95*s*)/ （22）

因此，对于*s*为0.36的混炼胶1的值为：

*C*u = 22.5 + (2.92 × 0.36/)

*C*u = 23.1”

并将表17进行重新计算，表题修改为“应力松弛百分率的置信上限”。

ISO19003表17 应力松弛百分率的置信下限

|  |  |  |  |
| --- | --- | --- | --- |
| 置信水平 | 混炼胶 | | |
| % | 1 | 2 | 3 |
| 90 | 22.1 | 19.8 | 15.8 |
| 95 | 21.9 | 20.5 | 16.5 |
| 99 | 21.1 | 23.0 | 19.2 |

本标准表17 应力松弛百分率的置信上限

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 置信水平 | *t*α | 混炼胶 | | |
| % |  | 1 | 2 | 3 |
| 90 | *t*0.90=1.8856 | 22.89 | 19.77 | 15.87 |
| 95 | *t*0.95=2.9200 | 23.11 | 20.42 | 16.55 |
| 99 | *t*0.99=6.9646 | 23.95 | 23.00 | 19.21 |

其中*t*α列没有列入标准正文表17中，在这里给出是为了给出计算时用到的*t*值。

从修改后的表17中，可以看出与ISO 19003中7.3.2对计算出的结果的解释是相符的。即混炼胶1都不符合要求，混炼胶3都符合要求，混炼胶2只有90%的符合要求。

2）结果的比较（7.3.3）

本条给出了对7.2.2.2的应用，即95%时，混炼胶1的|‾*x*1 -‾*x*2|＞*t*α*S，*说明其“平均值的差超过了*t*0.95*S*的乘积，所以超过95%可以确定两个实验室没有产生统计学上等效的数据” 。

同时分析了混炼胶2的数据，但没有给出99%的计算结果，只指出“混炼胶2，平均值的差接近99%的显著性水平。”其计算结果如下：

查t分布表，双侧检验，99%时查*t*0.005=3.3554，*t*0.005*S*=3.523。与表18中平均值差|‾*x*1 -‾*x*2|=3.3接近， |‾*x*1 -‾*x*2|＜*t*α*S*，

（五）第8章 排序法

排序法，当观察不能精确地量化，还必须做出主观的评价时所采用的一种检验方法。

这里给出了一个“秩”的概念，英文中用的是rank。秩，是次序，表达的是一种排序，所以有秩序一词。

排序法是将数据按规定的标准、按大小进行排列。多用于客观的对事物进行的评定，如气味、斑痕水印的深浅等不能量化的结果的评估。

本文件给出的示例是硫化胶耐臭氧试验结果龟裂长度的评价（8.3）。

本章还给出了一个简便的更加粗糙的方法——外计数法，快速简便。

（六）第9章 拒绝离群值的标准

离群值(outlier)，也称逸出值，是指在数据中有一个或几个数值与其他数值相比差异较大。有时会考虑将这样的结果作为离群值排除掉，否则会扭曲其他结果所代表的真实数据。但这个动作，是不希望发生的，要慎重处理。如果处理不当，会导致结果的失真，影响最终的判断。

本章给出了判断离群值，也就是可以被拒绝的结果的条件。同时明确了，只有超过99%置信水平的结果才被标记为离群值（outlier）。介于95%和99%置信水平之间的异常结果称之为歧离值(straggler)。

本章介绍了两种对离群数据进行检验的方法，即狄克逊检验（Dixon′s test）和柯克伦方差检验（Cochran’s test）.

1、狄克逊检验（9.2.2）

狄克逊检验（Dixon′s test）可用于对若干组试验结果的组平均值进行检验。其方法是使用计算出的狄克逊商（Q）与狄克逊检验临界值（表22）进行比较。

2、柯克逊检验（9.2.3）

柯克伦检验适用于利用观测集的方差进行检验，宜在狄克逊均值检验之前进行。这里用柯克伦商*C*与（表23）柯克伦检验临界值比较。

3、应用示例（9.3）

1）狄克逊检验用于单个结果异常（9.3.2）

用一组8个压缩永久变形试验结果进行了狄克逊检验，说明了尽管看起来有一个结果不正常、很离谱，但也不能仅凭直观感觉就将其轻易的剔除。

2）柯克伦方差检验（9.3.3）

对不利于七个实验室的试验结果进行了检验。首先是，发现有一个实验室的数据不正常，即出现了低均值、高标准差。鉴于此，进行了柯克伦方差检验。结果证明，检验统计结果柯克伦比C=0.764,大于表23查到的临界值0.664。所以可以拒绝实验室5的试验结果。

这里需要说的是，在柯克伦比的计算中，ISO 19003中计算出最后一位数字与验证计算的不同，见下列计算，括号中的是验算的结果。

*s*max2 = 0.8522 （0.8502）

= 0.722 5 （0.723 3）

∑*s*2 = 0.1732 + 0.0582 + ... + 0.1532

= 0.946 2（0.946 7）

Cochran比 *C*=0.722 5/0.946 2（0.723 3/0.946 7）

= 0,764

这可能是与计算过程中小数点后数字保留位数或录入错误有关。好在最后结果不受影响，结果一致。故在标准正文中予以更正。

3）Dixon检验应用于某一组的平均值异常（9.3.4）

这是对来自于6个实验室，每个实验室三个平行结果利用平均值进行狄克逊检验。这组试验结果（见标准中表27）似乎有一个很高的平均值19.8；标准差看起来也是起伏变化的，但没有一个结果是突出的异常大。

所以对平均值进行狄克逊检验。在进行平均值检验前，应先进柯克伦方差检验，以确定这组数据中没有标准差大到足以拒绝该数据。计算结果如下：

标准中表27 体积膨胀试验2

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| 实验室 |  | 结果 |  | **平均值** | **标准差** | *s*2 |
| 1 | 2 | 3 |
| 1 | 13.5 | 13.8 | 13.8 | 13.7 | 0.173 | 0.0299 |
| 2 | 10.8 | 13.0 | 12.6 | 12.1 | 1.173 | 1.3759 |
| 3 | 12.9 | 13.0 | 12.7 | 12.9 | 0.153 | 0.0234 |
| 4 | 10.9 | 11.2 | 14.2 | 12.1 | 1.825 | 3.3306 |
| 5 | 14.2 | 14.2 | 14.4 | 14.3 | 0.115 | 0.0132 |
| 6 | 19.7 | 20.8 | 18.9 | 19.8 | 0.954 | 0.9101 |
| *s*2 |  |  |  |  |  | 5.6832 |
| =3.3306/5.6832 =0.5860 | | | | | | |

柯克伦比C=0.5860,小于99%和95%置信水平柯克伦临界值0.722和0.616，柯克伦检验通过。

然后，再进行狄克逊平均值检验，结果如标准中所述，狄克逊商Q=0.714,大于表查临界值95%置信水平的0.628，小于99%置信水平的是0.740。此为歧离值，不是离群值，不宜拒绝。

（七）第10章 方差分析（ANOVA）

方差分析（analysis of variance, ANOVA），用于两个及两个以上样本均值差别的显著性检验，即多个总体均值是否相等的统计方法。

本章不介绍方差分析的理论，只注重于应用。

有两个术语需要弄清楚，因素和水平。我们常说，影响混炼胶某一性的因素有哪些（如炭黑、操作油、硫化温度和时间等）。这个因素，就是方差分析中的因素，也叫因子。这些炭黑、操作油等（因素）的用量的变化对性能也是有影响的；那么，这些不同的用量（如10份、20份、30份等）就称之为水平。

本章主要是给出了进行方差分析过程中涉及的计算公式，并用这些计算公式分析了炭黑和操作油用里的变化对混炼胶耐磨性能的影响。但是，计算中出现了些错误。现将这一示例重新计算如下：

标准表29 磨耗体积

单位：mm3

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| a)原始结果 | | | | | |
| 结果 | 油水平（含量） | 炭黑水平（含量） | | | |
| 60 | 80 | 100 | 120 |
| 1 | 0 | 273 | 256 | 202 | 188 |
| 2 |  | 233 | 262 | 215 | 195 |
| 3 |  | 273 | 242 | 261 | 177 |
| 1 | 5 | 288 | 257 | 244 | 242 |
| 2 |  | 260 | 271 | 229 | 203 |
| 3 |  | 313 | 311 | 245 | 201 |
| 1 | 10 | 269 | 247 | 249 | 217 |
| 2 |  | 317 | 253 | 220 | 215 |
| 3 |  | 245 | 262 | 232 | 203 |
| 1 | 20 | 231 | 270 | 222 | 230 |
| 2 |  | 298 | 307 | 227 | 214 |
| 3 |  | 287 | 278 | 203 | 242 |
| b) 调整后的结果 | | | | | |
| 结果 | 油水平（含量） | 炭黑水平（含量） | | | |
| 60 | 80 | 100 | 120 |
| 1 | 0 | 2.73 | 2.56 | 2.02 | 1.88 |
| 2 |  | 2.33 | 2.62 | 2.15 | 1.95 |
| 3 |  | 2.73 | 2.42 | 2.61 | 1.77 |
| 1 | 5 | 2.88 | 2.57 | 2.44 | 2.42 |
| 2 |  | 2.60 | 2.71 | 2.29 | 2.03 |
| 3 |  | 3.13 | 3.11 | 2.45 | 2.01 |
| 1 | 10 | 2.69 | 2.47 | 2.49 | 2.17 |
| 2 |  | 3.17 | 2.53 | 2.20 | 2.15 |
| 3 |  | 2.45 | 2.62 | 2.32 | 2.03 |
| 1 | 20 | 2.31 | 2.70 | 2.22 | 2.30 |
| 2 |  | 2.98 | 3.07 | 2.27 | 2.14 |
| 3 |  | 2.87 | 2.78 | 2.03 | 2.42 |
| 注：在b)中的调整后的结果是原始数据除以100后的结果，这样做是为了便利用小数字制表。 | | | | | |

将表29调整后的将每个炭黑水平中的每个油水平三次测量结果加和，结果列入新表中，得表30。

标准表30 和之表

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 油水平（含量）  (因素 B) | 炭黑水平（含量） (因素 A) | | | | A之和 |
| 60 | 80 | 100 | 120 | (B*X*j) |
| 0 | 7.79 | 7.60 | 6.78 | 5.60 | 27.77 |
| 5 | 8.61 | 8.39 | 7.18 | 6.46 | 30.64 |
| 10 | 8.31 | 7.62 | 7.01 | 6.35 | 29.29 |
| 20 | 8.16 | 8.55 | 6.52 | 6.86 | 30.09 |
| ﻿B之和 (A*X*i) | 32.87 | 32.16 | 27.49 | 25.27 |  |

按本章给出的和附录G给出的相关参数的公式处理本章示例中的所有数据。计算结果与ISO19003中的数据对比列入下表，为方便对比仅将与ISO19003不同的结果列出。

表x ISO19003第10章示例计算结果与验算结果对比

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 参数 | 公式 | ISO19003中的值 | 计算值 | 修约值 |
| 因素T |  | 17.79 | 117.79 |  |
| A值的个数 | *a* | 4 |  |  |
| B值的个数 | *b* | 4 |  |  |
| 重复的次数 | *r* | 3 |  |  |
| 校正因子 |  | 289.07 | 289.0517521 | 289.05 |
| 所有观测值的平方和 |  | 879.43 | 879.4219 | 879.42 |
| 因素A平方和 |  | 3.352 | 3.362872917 | 3.363 |
| 因素B平方和 |  | 0.388 | 0.389472917 | 0.389 |
| 因素A和B共平方和 |  | 0.337 | 0.336535417 | 0.337 |
| 残差平方和 |  | 1.483 | 1.49666666666661 | 1.497 |
| 总平方和 |  | 5.559 | 5.58554791666671 | 5.586 |
| 因素A自由度 | =*a –* 1 | 3 |  |  |
| 因素B自由度 | =*b -* 1 | 3 |  |  |
| 因素A和B共自由度 |  | 9 |  |  |
| 残差自由度 |  | 33 | 32 |  |
| 总自由度 |  | 48 | 47 |  |
| 因素A方差 | *Ma*= | 1.117 | 1.120957639 | 1.121 |
| 因素B方差 | *Mb*= | 0.129 | 0.129824306 | 0.130 |
| *Mab*= |  | 0.037 | 0.037392824 | 0.037 |
| *Mr*= |  | 0.045 | 0.046770833 | 0.047 |
| *Mt*= |  | 0.116 | 0.118841445 | 0.119 |

据此，对标准中示例的计算结果予以校正，其中**表31校正为：**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 源 | 平方和 | 更正为 | 自由度 | 方差估计 | 更正为 |
| 因素 A | 3.352 | 3.363 | 3 | 1.117 | 1.121 |
| 因素 B | 0.388 | 0.389 | 3 | 0.129 | 0.130 |
| 相互作用 | 0.337 | 0.337 | 9 | 0.037 | 0.037 |
| 残差 | 1.483 | 1.497 | 32 | 0.045 | 0.047 |
| 合计 | 5.560 | 5.586 | 47 | 0.116 | 0.119 |

**10.3.5和10.3.6各参数的计算值见下表，标准中按修约值更正。**

|  |  |  |  |
| --- | --- | --- | --- |
| 参数 | ISO19003的值 | 计算值 | 修约值 |
| *Mab*/*Mr*= | 0.83 | 0.799490225 | 0.80 |
| S*r* = Sab + Sr | 1.829 | 1.83320208333333 | 1.833 |
| *DFr* =*DFab* + *DFr* | 42 | 41 |  |
| *Mr* = S*r*/ DF*r* | 0.043 | 0.044712246 | 0.045 |
| *Ma*/*Mr* | 25.98 | 25.07048383 | 25.07 |
| *Mb* /*Mr* | 3.00 | 2.903551428 | 2.90 |

**尽管数据计算出现了一些错误（失误），但是，方法没有错误，总的结论没有错误。**

注：利用Excel进行方差分析的步骤可参见贾俊平《统计学》（第七版）一书

（八）第11章 回归分析

回归分析主要的任务是在考察变量之间的数量依存关系的基础上，通过一定的数学表达式将这种依存关系描述出来，进而确定一个或几个变量（自变量）对另外一个特定变量（因变量）的影响程度。回归分析主要解决以下几个方面的问题：

——从一组样本数据出发，确定变量之间的数学关系式；

——对这些关系式的可信程度进行各种统计检验，并从影响某一特定变量的诸多变量中找出哪些变量的影响是显著的，哪些是不显著的；

——利用所求得的关系式，根据一个或几个变量的取值来估计或预测另一个特定变量的取值，并给出这种估计或预测的可靠程度。

本标准主要介绍了用最小二乘法建立一元回归方程的方法，包括一元一次回归方程、一元二次回归方程和一元三次回归方程。其中主要介绍的是建立一元一次回归方程的方法——线性最小二乘法（11.2.2）。建立一元二次回归方程的二次最小二乘法（11.2.3）和建立一元三次回归方程的立方最小二乘法（11.2.4）则以附录的形式给出。

建议计算一元二次函数、一元多次函数的回归方程中的系数利用软件进行，手工计算量太大，易出错。

1、　线性最小二乘法（11.2.2）

此为最简单的线性回归方程：

*y* = *a* + *bx*

在11.2.2中ISO 19003给出了计算线性回归方程中系数a和b的因子。并通过计算检验统计量F，使其与F分布表中的值进行比较来判断两个变量x和y之间的关系是否显著。

其中公式（52）有误，应为：。因为：

计算统计量*F*r的公式为：，其中：

SSR（回归平方和）= =D

SSE（残差平方和）= SST – SSR = Cyy – *b*2C11

SST（总平方和）=SSR + SSE = Cyy

,令,则：

2、二次最小二乘（11.2.3）

这里假设回归线形式为一元二次方程：

*y*=*a*+b*x*+c*x*2

这种分析所需的因子计算在附录H中给出。

3、立方最小二乘（11.2.4）

这里假设回归方程为一元三次方程：

该分析所需的因子计算在标准附录H中给出。

4、应用示例（11.3）

1）温度对压缩永久变形的影响（11.3.2）

ISO19003给出了按照ISO 815进行测试的、在不同温度下老化7d后测定压缩永久变形数据，标准中表32。这是一组橡胶热老试验，一般服从化学反应动力学定律——阿伦尼乌斯Arrhenius定律。所以可以用Arrhenius公式建立数学模型，以预测其性能变化。

Arrhenius公式的形式为：

式中，*α*和*β*是常数；*T*是温度，单位为开尔文，K(绝对温度)。

这个函数是不能直接利用最小二乘法确定*α*和*β*的,但通过取自然对数它很容易转变成可用的公式：

式中，*θ*是温度，℃。

将该公式转换为线性方程的形式：

式中，

，cs为三个平行数据的均值；

。

由此，将标准中表32数据转化为如下表，

标准中表33　转换的压缩永久变形变量

|  |  |  |
| --- | --- | --- |
| *x，*10–3 | | *y* |
| ISO 19003 | 验算excel计算出的 | 平均值 |
| 2.92 | 2.915452 | 3.21 |
| 2.79 | 2.793296 | 3.42 |
| 2.68 | 2.680965 | 3.60 |
| 2.51 | 2.512563 | 3.84 |
| 2.36 | 2.364066 | 4.05 |

用11.2.2中给出的各因子计算公式，计算出的结果如下。

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| *C*11 | 1.93×10-7 | *a* | 7.66 |  | -20.17 |
| *C*yy | 0.446 | *b* | -1520.71 | *Fr* | -2.9350 |
| *C*y1 | -2.93×10-4 |  |  |  |  |
|  |  |  |  |  |  |
| SSR=D= | 0.445 |  |  |  |  |
| SSE= *C*yy –*b*2C11 | 0.001 | *F*r | 1389.849191 |  |  |
| SST= *C*yy | 0.446 |  |  |  |  |

其中，及由此D值计算出的*F*r=-2.9350，与ISO 19003计算的结果不符；而用D=计算的结果D=0.445是正确的，由此得*F*r=1389.85。

回归系数*a*=7.66，*b*=-1520.71

因为：*a* = ln(α)，*b* = *β*，所以：

α= EXP(*a*)= EXP(7.66)=2121.76

*β*=-1520.71

这样得到的回归方程为：

据此对ISO19003中的数据进行了修正。

下列数据是利用excel<数据><数据分析><回归>计算的结果，与用ISO19003给出计算因子计算所得结果一致。

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 方差分析 |  |  |  |  |  |
|  | df | SS | MS | F | Significance F |
| 回归分析 | 1 | 0.445288 | 0.445288 | 1389.849 | 4.25E-05 |
| 残差 | 3 | 0.000961 | 0.00032 |  |  |
| 总计 | 4 | 0.44625 |  |  |  |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | Coefficients | 标准误差 | t Stat | P-value | Lower 95% | Upper 95% |
| Intercept | 7.661148 | 0.108525 | 70.59339 | 6.26E-06 | 7.315773 | 8.006523 |
| X Variable 1 | -1520.71 | 40.79096 | -37.2807 | 4.25E-05 | -1650.53 | -1390.9 |

查F分布表，95%置信度（显著性水平α=0.05），分子自由度*df*1=1,分母自由度*df*2=5-2=3, 临界值*F*α=10.13。由于*F*r>*F*α，表明所得到回归方程是显著的。

在利用excel计算所得结果中，有一项Significance F，它是用于检验的P值。将significance F的值与给定的显著性水平α的值进行比较，如果Significance F<α，则表明因变量y与自变量x之间有显著的线性关系；否则，就没有证据表明因变量y与自变量x之间有显著的线性关系。本示例，Significance F=4.25×10-05<α=0.05。

利用回归方程验算，其结果与ISO19003的计算结果一致（见下表）。

表34　压缩永久变形值的比较

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 温度 | 观测值 | 计算值 | 验算结果 | |
| ℃ | % | % | cs | 修约 |
| 70 | 24.7 | 25.2 | 25.18993 | 25.2 |
| 85 | 30.7 | 30.4 | 30.33214 | 30.3 |
| 100 | 36.7 | 36.0 | 35.98243 | 36.0 |
| 125 | 46.6 | 46.5 | 46.48463 | 46.5 |
| 150 | 57.6 | 58.3 | 58.26154 | 58.3 |

ISO19003中的中间计算过程所得一些计算结果与我们验算的结果经常出现的不一致，或略有差异，除录入错误外，主要是修约问题，也是ISO19003中指出的那样，其一些计算精度不够，过早地进行了修改，造成的计算误差的传递。

2）老化对拉伸强度的影响

一种橡胶混炼胶在70℃下热老化1个月。每隔一周，将五个哑铃样样从烘箱中取出，冷却一晚，在23℃下进行测试，结果在标准中表35给出。

这是一个一元二次方程，标准中利用附录H给出的各因子，得到由下述方程（60）定义形式的回归线:

TS = 11.6 + 1.16*t* − 0.0511*t*2

但是，验算时得到的结果与之不符：

*a*=12.79561427，*b*=-1.16628×10-04，*c*=6.29769×10-08

利用excel进行计算所得结果则与之相符：

*a*=11.59742857 =11.6 （Intercept，截距）；

*b*= 1.158734694 =1.16 （X Variable 1，变量x1）;

*c*= -0.051107872=-0.0511 ( X Variable 2,变量x2)。

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 方差分析 |  |  |  |  |  |
|  | df | SS | MS | F | Significance F |
| 回归分析 | 2 | 124.1286 | 62.06431 | 602.2654 | 0.001657645 |
| 残差 | 2 | 0.206103 | 0.103051 |  |  |
| 总计 | 4 | 124.3347 |  |  |  |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | Coefficients | 标准误差 | t Stat | P-value | 下限 95.0% | 上限 95.0% |
| Intercept | 11.59742857 | 0.302116 | 38.38733 | 0.000678 | 10.29752802 | 12.89733 |
| X Variable 1 | 1.158734694 | 0.051126 | 22.66441 | 0.001941 | 0.938758406 | 1.378711 |
| X Variable 2 | -0.051107872 | 0.001751 | -29.1891 | 0.001172 | -0.058641483 | -0.04357 |

利用excel解一元一次、二次或三次方程，方法如下：

本示例中，在excel中建立一个数据表格如下：

|  |  |  |
| --- | --- | --- |
| 中位数 | 老化时间/d |  |
| y | x | x2 |
| 11.6 | 0 | 0 |
| 17.3 | 7 | 49 |
| 17.5 | 14 | 196 |
| 13.7 | 21 | 441 |
| 3.87 | 28 | 784 |

然后，

a) 在数据集添加**"临时"**列,使用Excel的数据→数据分析→回归。.

b) 如果做一个简单的线性回归，只需要2列,X和Y。

c) 如果要做二次或三次回归分析，关键在于定义输入X范围:如果要做二次方，则需要X\_1,X\_2和Y，其中X\_1是x变量，X\_2是x 2 ； 同样，如果做一个立方体，就需要X\_1，X\_2，X\_3和Y，其中X\_1是x变量，X\_2是x 2，X\_3是x 3。注意输入X范围是从A1到Bnn，跨越2列；或从A1到Cnn,跨越3列。

3）回缩试验的温度（11.3.4）

ISO 19003中，根据ISO 2921进行的回缩温度试验中，将三个试样置于热交换槽中，使温度从- 70℃上升到室温，这期间每隔2min测量一次这三个试样的回缩百分率，以估算硫化橡胶恢复10% (TR10)、50% (TR50)和70% (TR70)时的温度。标准中提供了16个测试温度下测定的三个试样的平均值，即一共有16×3=48个数据（见标准中表36）。

ISO19003中“测得了每组三个试样的总计44个数据对（A total of 44 data pairs for each of the three test pieces was produced）”，有误，修改为“测得了每组三个试样的总计48个数据”表36中给出了平均值的简表

根据标准表36的数据，用excel和**SPSS**（Statistical Product Service Solutions，“统计产品与服务解决方案”软件）对其进行了回归计算，都得出如下回归方程：

ISO19003中式（61）：

有误，并根据上述计算结果予以了更正。

ISO 19003中对于此回归方程的计算利用的附录H中的计算因子。如前述一元一次、一元二次回归方程计算时所发现，其中计算因子有误，又无从查找其来源，无法对这些计算因子进行勘误。

建议使用统计学软件建立二次、三次回归函数，如excel、spss等。

以下是用excel计算的过程

回缩率-温度回归方程计算

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | 回缩 | 温度 |  | |
| % | ℃ |
| *R* | *θ* | *θ2* | *θ3* |
| 1 | 0 | -68.3 | 4664.89 | -318611.987 |
| 2 | 0 | -62.2 | 3868.84 | -240641.848 |
| 3 | 0.7 | -56.6 | 3203.56 | -181321.496 |
| 4 | 2.7 | -50.6 | 2560.36 | -129554.216 |
| 5 | 7 | -44.8 | 2007.04 | -89915.392 |
| 6 | 11 | -38.7 | 1497.69 | -57960.603 |
| 7 | 17 | -32.9 | 1082.41 | -35611.289 |
| 8 | 23.3 | -26.7 | 712.89 | -19034.163 |
| 9 | 29 | -20.6 | 424.36 | -8741.816 |
| 10 | 36 | -14.9 | 222.01 | -3307.949 |
| 11 | 41.3 | -8.4 | 70.56 | -592.704 |
| 12 | 55.7 | -0.8 | 0.64 | -0.512 |
| 13 | 66 | 3.8 | 14.44 | 54.872 |
| 14 | 75.7 | 9.2 | 84.64 | 778.688 |
| 15 | 85.3 | 15.3 | 234.09 | 3581.577 |
| 16 | 90.7 | 22.2 | 492.84 | 10941.048 |

|  |  |
| --- | --- |
| 回归统计 | |
| Multiple R | 0.998164271 |
| R Square | 0.996331912 |
| Adjusted R Square | 0.99541489 |
| 标准误差 | 2.157237165 |
| 观测值 | 16 |

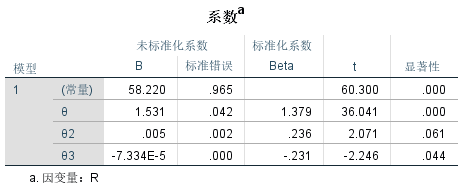
|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 方差分析 |  |  |  |  |  |
|  | df | SS | MS | F | Significance F |
| 回归分析 | 3 | 15168.45 | 5056.151 | 1086.486 | 7.13E-15 |
| 残差 | 12 | 55.84407 | 4.653672 |  |  |
| 总计 | 15 | 15224.3 |  |  |  |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | Coefficients | 标准误差 | t Stat | P-value | Lower 95% | Upper 95% |
| Intercept | 58.21975348 | 0.965497 | 60.3003 | 2.86E-16 | 56.11612 | 60.32339 |
| X Variable 1 | 1.530955979 | 0.042479 | 36.04062 | 1.33E-13 | 1.438403 | 1.623509 |
| X Variable 2 | 0.004991273 | 0.00241 | 2.071073 | 0.060573 | -0.00026 | 0.010242 |
| X Variable 3 | -7.33413E-05 | 3.27E-05 | -2.2456 | 0.044347 | -0.00014 | -2.2E-06 |

其中：a=58.2(Intercept), b=1.53(X Variable 1), c=0.004 991(X Variable 2),

d=-7.33413E-05=-0.000 0733 4(X Variable 3)。

Spss系数估计表如下：



另，F值为1086.486，自由度为3×(48-3-1)=3×44（为什么与方差分析的结果不一致？）。

并据此，将11.3.4.3中的F值和自由度进行了更正。

用 对标准中表37进行计算验证，结果如下：

表37计算对比

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 回缩值 | 温度 | | 计算验证 | |
|  | ℃ | | ℃ | |
|  | 三次方 | S形 | 三次方 | S形 |
| 10 | -39 | -37 | -39.59 | -38.34 |
| 50 | -5 | -6 | -5.46 | -5.46 |
| 70 | +8 | +7 | 7.55 | 7.03 |

数据与ISO19033数据不一致的主要原因是小数位数不同导致的，属于计算误差。S形由累计正态分布函数计算得出，分别偏-1.28、0、0.52，一般不取S形计算结果，不测算也无妨。

（九）第12章 不确定度的测量

不确定度的评定就是要得到观测结果的离散性。用它来表征测量结果的离散性，即测量结果有多大概率会现在某个范围之内。其使用是在出具检验报告时，将不确定度（*U*）与检测结果一起报告，其形式如下：

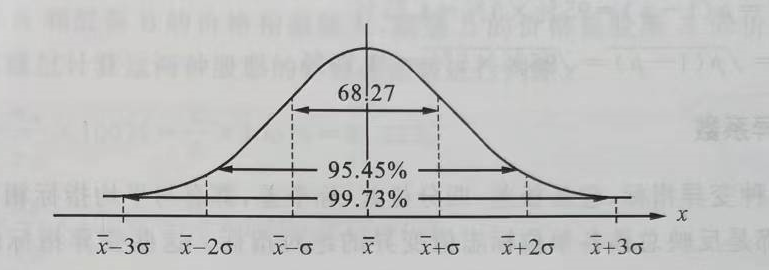
其中，是经过修正的检测结果。

报告中还应给出“本不确定度的估计置信概率不低于95%”这样的关于置信度的陈述。否则，该报告的价值不大。

不确定度包括两个部分，随机不确定度(*U*r)和系统不确定度（*U*s）。

随机不确定度(*U*r)通常被认为是正态分布的，所有可能测量值的68.3%会落在限值±σ之间，其中σ被定义为标准偏差。95%位于±1 96σ之间, 99.7%位于±3σ的范围内。

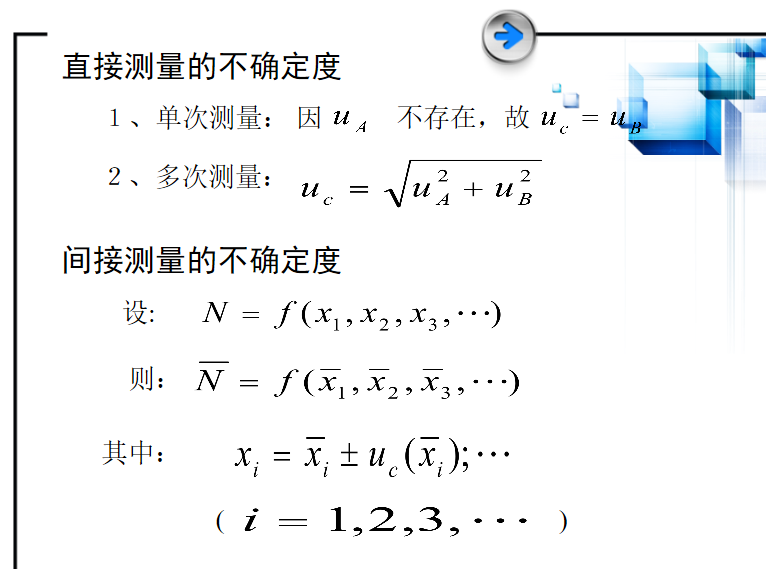
当考虑整个总体的正态分布时，对于95%的置信水平，不确定度（±*U*）,可以视为等于±1,96σ。事实上，这也能处理成±2σ。



均值的随机不确定度用标准中公式62） 计算。而σ/ ，当小样本时，即为标准误差，通过查*t*分布表就可计算；当样本量足够大时，在95%的置信水平下，*t*=1.96。

系统不确定度*U*s是非统计学处理所得，其来源为较多，如使用了不正确的校正、温度影响和校准的不确定度等。还有一个主要来源是试验或测量仪器附带的校准证书。

总不确定度*U*是随机不确定度和系统不确定度的合并， 。



（十）第13章 抽样

抽样，是一个很大的命名。其他国际标准已很好地讨论了这个题目。有两种抽样方法，计数抽样和计量抽样。本标准只给出了一个大纲，有关详细信息，感兴趣的用户请参考计数抽样标准ISO 2859的各个部分和计量抽样标准ISO 3951的各个部分。

橡胶工业由于本身的特点，多采用计数抽样。计数抽样的详细规定请参见ISO 2859 Sampling procedures for inspection by attributes的所有部分，与之对应的国家标准为GB/T 2828《计数抽样检验程序》的所有部分。

本章简单地介绍了与抽样有关的几个概念：检验水平、计数抽样计划、随机抽样等。

还介绍了简单介绍了ISO 2859-1:1999的使用。

关于AQL的说明

ISO 19003中将其称为acceptable quality level（接收质量水平）与ISO 2859-1;1989相同，但ISO 2859-1已修订为1999版，Sampling procedures for inspection by attributes — Part 1: Sampling schemes indexed by acceptance quality limit (AQL) for lot-by-lot inspection重新定义为acceptance quality limit（接收质量限）

ISO 2859-1:1989 Sampling procedures for inspection by attributes — Part 1: Sampling plans indexed by acceptable quality level (AQL) for lot-by-lot inspection中对于AQL的定义为：acceptable quality level (AQL)：When a continuous series of lots is considered, the quality level which for the purposes of sampling inspection is the limit of a satisfactory process average.

当考虑一个连续批时，抽样检验的质量水平，即令人满意的过程平均值的极限。

GB/T 2828.1-1987AQL的定义，在抽样检验中，认为可以接收的连续提交检验批的过程平均上限值。

GB/T 2828.1-2013/ISO 2859-1:1999中对AQL的定义为：

acceptance quality limit

AQL

quality level that is the worst tolerable process average when a continuing series of lots is submitted for acceptance

当一个连续系列批被提交抽样时，可容忍的最差过程平均质量水平。

两者的含义是相同的。

鉴于此，本文件中，将AQL“接收质量水平”替换为“接收质量限”与国家标准GB/T 2828.1-2013和现行的国际标准ISO 2859-1:1999保持一致。

（十一）第14章 试样数量

在一些需要通过对大量的试样进行试验来改进统计精密度的时候，需要估计需要多少试样既能满足精度的要求，又能降低成本。

本章给出了简单的估算方法。更准确的方法在第17章中给出。

（十二）第15章 结果表示

1、试验报告的基本内容（15.2.1）

标准中15.2.1给出了试验报告应包括的内容，当试验方法标准中指明报告中要引用的统计数据时，建议在报告中加入以下内容：

1. 均值‾*x*(见6.2.2.2)；
2. 总体的估计标准差s(见6.2.3.2)；
3. 变异系数，Cv(见6.2.3.4)；
4. 每个试验结果；
5. 不确定性的估计，如果有。

试验报告中最好将每个试验结果都报告出来，并给出处理方法、所作的假设等，以便对将要作出的结果进行独立的核查

在遇到大数据集时，可以用某种形式的图表（如直方图）来代替。

2、数值修约

数值的修约很重要，应给予足够的重视。15.2.2简单地介绍了数据修约的一些常识。15.3介绍了一些参数测量值报告的有效数字的一般规则和注意事项，并给出了修约示例。

报告参数的值时的一般规则:

1. 线性尺寸、体积、力和应力不宜超过三位有效数字;
2. 应变和能量不宜超过两位有效数字;
3. 质量可以报告至四位有效数字，尽管通常三位就足够了。

（十三）第16章 精密度陈述

精密度陈述，主要是作为试验方法标准中的一章的内容。本章主要是介绍了一些基本要求和注意事项，精密度试验及其数据的处理应按ISO19983进行。

用名义相同的材料在名义相同的条件下进行的试验通常不产生相同的结果。重复试验之间的这种差异被冠以一个通用的名字——精密度。

不同操作员和（或）不同设备之间的差异通常比在短时间内使用给定设备的给定操作员的差异大。因此，精密度包括两个方面的度量——重复性和再现性。

与所有统计因素一样，重复性和再现性是基于接受的置信水平的估计。一般都是取95%的置信水平。

开展精密度试验时，实验室的选择应避免具有专门技能的实验室或被认为是标准实验室的实验室占多数。

所得数据以类似于方差分析(第10章)的表格形式排列，所得的统计量如下:

1. 实验室间方差的估计，*s*L2；
2. 实验室内方差的估计，*s*w2；
3. 试验方案中所有实验室的实验室间方差的平均值，sm2；
4. 重复性方差的估计，*s*r2；
5. 再现性方差*s*R2的估计

1. 重复性*r*，

1. 再现性*R*，

（十四）第17章 实验设计

1、内容概述

本章主要了三种实验设计，描述性设计、比较性设计和响应性设计

1）、描述性设计

对一个全部试样从标准材料中随机抽取的统计学样本进行试验以确定该材料某一特性的基本统计量。例如，能够报告的该标准材料拉伸强度的平均值和标准差。

2）、比较性设计

比较性设计包括与特定标准进行比较、两材料与独立样本的比较和用成对样本比较两种材料。

比较实验中，实验者只关心平均值是否不同，则使用双侧检验；如果想证明一种新材料的均值比标准的更大，则使用单侧检验。

双侧检验的假设为：

*H*0，当*μ*1=*μ*2

*H*a，当*μ*1≠*μ*2

单侧检验的假设为：

*H*0，当*μ*1=*μ*2

*H*a，当*μ*1＞*μ*2

3）、响应设计

响应设计一章包括因素实验、复合设计和实验室间试验

a) 因素实验

研究几个变量的影响的实验。

b）复合设计

复合设计亦称为增强设计。系在研究的最后阶段增加某因素的额外试验点，以便估计响应曲线。

c）实验室间试验

实验室间试验旨在评估一组实验室中每个实验室内试验结果的重复性和实验室之间试验结果的再现性。也是试验方法标准中重要一章“精密度”的内容。

ISO 19003中提到精密度的确定在ISO 5725和ISO/TR 9297中有详细的描述。ISO/TR 9297已被ISO 19983所代替，所以本标准中替换为ISO 19983。

2、问题与修改

1） 17.2.2和17.2.3 给出了描述性实验、比较性实验和响应实验中试样量N的计算的示例。在这些计算示例中都在给出的试样数量N的计算公式前增加一个常数项0.5，

0.5+(*Zα*σ/cI)2 ，17.2.2 e)

, 17.2.3.1 e)

, 17.2.3.2 e)

, 17.2.3.3 e)

增加这常数项0.5的目的是确保公式中第二项计算出的结果取整修约时，都向上进位。这样做，在此处是方便了，但因没有给出说明，则造成了与17.3的实例中的不统一，使读者难于理解。

在17.3中的实例计算中的公式分别为：

17.3.1.1 e)， (*Z*ασ/*δ*)2 = 5.52；

17.3.1.2 e) ，(*Z*ασ/*δ*)2 = 4.7；

17.3.2.1 e) ，[(*Z*α + *Zβ* )*σ*/*δ* ]2 = 22.5；

17.3.2.2 e)，2(*Z*α + *Zβ* )2*σ*2/*δ*2 = 17.08；

17.3.2.3.1 e) ，(*Z*α + *Zβ* )2（2*σ*2）/*δ*2 = 13.88；

17.3.2.3.2 e)， (*Z*α + *Zβ* )2（2*σ*2）/*δ*2 = 86.59。

用17.2.2的数据对比两种方法计算的结果：

a）用0.5+(*Zα*σ/cI)2计算并按四舍五入修约

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| *Zα* |  | *C*l | n | 简单修约N | *t*α |
| 1.96 | 1.414214 | 1 | 8.1832 | 8 | 2.365 |
|  |  |  | 11.68645 | 12 | 2.201 |
|  |  |  | 10.1888 | 10 | 2.262 |
|  |  |  | 10.73329 | 11 | 2.228 |
|  |  |  | 10.42797 | 10 | 2.262 |
|  |  |  | 10.73329 | 11 |  |
| 取N=11。 | | | | | |

b）用(*Zα*σ/cI)2计算并向上圆整修约

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| *Zα* |  | *C*l | n | 向上修约N | *t*α |
| 1.96 | 1.414214 | 1 | 7.6832 | 8 | 2.365 |
|  |  |  | 10.63527 | 12 | 2.201 |
|  |  |  | 9.927968 | 10 | 2.262 |
|  |  |  | 10.23329 | 11 | 2.228 |
|  |  |  | 9.927968 | 10 | 2.262 |
|  |  |  | 10.23329 | 11 |  |
| 取N=11。 | | | | | |

c）计算比较17.2.3中的其他示例，计算结果相同。

d）结论，对比两种计算的结果，都取相同的结果，所以为减少误解，将计算示例中的常数项0.5取消。这也与统计学中的样本圆整法则保持一致，即：将小数点后面的数值进位成整数（见贾俊平《统计学（第7版）》，中国人民出版社，P149）。

2) 17.3.2.1示例中e)项的计算结果

[(*Z*α + *Zβ* )*σ*/*δ* ]2 = 22.5 应为22.8。

3）17.3.2.2示例中e）项的计算结果

2(*Z*α + *Zβ* )2*σ*2/*δ*2 = 17.08应为17.03

4）17.2.3.3中b)选择*α*、*β*、*δ*并估计(*σ*diff)2

注：如果潜在总体方差是*σ*2，并且假设两个总体的方差相同，则两个值(每个总体一个)之间的差的方差为2*σ*2。

注的意思是(*σ*diff)2 =2*σ*2，即两个总体的方差是这两个总体方差的和。

因此，e)项中的 应为：

，去掉常数项，则为：

由此，

17.3.2.3.1例1 两实验室撕裂强度的比较中b)项 σdiff = σ√2 的计算才有意义。在e)项的计算中，其公式 (*Z*α + *Zβ* )2（2*σ*2）/*δ*2 = 13.88，应表达为

(*Z*α + *Zβ* )2/*δ*2 = 13.88

直接用ISO19003的原式计算也可行相同的数值，但是就失去了在注中给出定义的意义。

17.3.2.3.2 例2 同上修改，e)项的计算式改为：

(*Z*α + *Zβ* )2/*δ*2 = 86.57

（十五）第18章 统计学在质量控制上应用

1、内容概述

本章主要介绍了质量控制上常用的控制图，计数控制图和计量控制图。

控制图，一般是以某一特征性能均值对时间作图的，通常不以连线的方式，而是以散点的方式绘制，如图14。

还有一种控制图，按均值和目标值之间的累积差对时间作图以累积和对时间作图，如图17。

2、错误与更正

本章18.2.3.2.2中的计算公式有误，根据计算上下限的公式

——1) 对于内控限，±1.96 *s*/*n*；应为±1.96 *s/*；1.96是双侧概率α/2=1/40=0.025的z分数（即置信度95%）；

——2) 对于外控限，±3.09 *s*/*n*。应为±3.09 *s/*；3.09是双侧概率α/2=1/1000=0.001的z分数（即，置信度为99.9%）

由此计算的结果与ISO 19003的最终结果一致：

内控限：59.5+1.96×1.2/=60.6 IRHD 和59.5-1.96×1.2/= 58.4 IRHD;

外控限：59.5+1.96×3.09/=61.158 IRHD和59.5-1.96×3.09/=57.842 IRHD

（十六）附录

1、附录A(资料性)

附录A为资料性的，给出了本文件中引用的分布函数的数学形式：

——正态密度分布函数

——双指数密度分布函数

——威布尔（Weibull）密度分布函数。

2、附录B(资料性)

附录B为资料性的附录，介绍了平均值的其他形式:

——几何均值;

——平方均值的平方根。

3、附录C（资料性）

附录C为资料性的，给出了双指数和威布尔分布中集中趋势度量的相互关系

4、附录D ISO19003附录D “计算标准差的公式”

ISO19003附录D为资料性的附录。介绍了另外一个计算标准差的公式，但又不推荐使用。同时介绍了有效数字的大小、使用的计算机软件的精度问题可能引用的误差。

其给出的计算示例，但起草工作组无法重复其关键的计算结果。

对表D.1用6.2.3.2中给出的公式（5）和（6）以及附录中的公式（D.1）进行了计算，结果与原表中的第三列的结果相差很大。

|  |  |  |  |
| --- | --- | --- | --- |
| 值 | 小 | 中 | 大 |
| 1 | 109.585 8 | 10 09.586 | 10 009.59 |
| 2 | 106.336 9 | 10 06.337 | 10 006.34 |
| 3 | 107.006 0 | 10 07.006 | 10 007.01 |
| 4 | 106.607 0 | 10 06.607 | 10 006.61 |
| 5 | 109.854 0 | 10 09.854 | 10 009.85 |
| 6 | 103.337 5 | 10 03.337 | 10 003.34 |
| 7 | 105.088 8 | 10 05.089 | 10 005.09 |
| 8 | 107.605 3 | 10 07.605 | 10 007.61 |
| 9 | 103.619 6 | 10 03.620 | 10 003.62 |
| 10 | 104.260 0 | 10 04.260 | 10 004.26 |
| 11 | 108.163 3 | 10 08.163 | 10 008.16 |
| 12 | 103.354 7 | 10 03.355 | 10 003.35 |
| 13 | 103.952 0 | 10 03.952 | 10 003.95 |
| 14 | 103.021 2 | 10 03.021 | 10 003.02 |
| 15 | 105.104 8 | 10 05.105 | 10 005.10 |
| 均值 | 105.793 1 | 10 05.793 | 10 005.79 |
| SD1 | 2.210 593 | 2.210 586 | 2.210 653 |
| 公式（6）验算 | 2.210 578 | 2.210 562 | 2.211 282 |
| 公式（5）验算 | 2.288 165 | 2.288 149 | 2.288 895 |
| SD2 | 2.210 874 | 2.245 459 | 4.532 608 |
| 验算 | 2.210 578 | 2.210 562 | 2.211 282 |

验算是用公式（5）、（6），以及excel中的函数STED.P、STED.S进行的。

验算的结果虽然都有差异，但主要是在第三例，特别是SD2的结果。

无法考证（重复）ISO19003中的计算，因其声称是“用一个具有单精度变量的BASIC程序计算的”，目前已无法找到BASIC程序软件，且其也声称，其是单精度的，所以会产生误差，且很大。

单精度（single-precision）和双精度(double-precision)型数值为浮点数值，它表示的是带小数的实数。通常单精度型能精确到七位，而双精度能精确到15位。而这两者的误差往往是由计算机的硬件决定的，即便是同一个软件、都是单精度的情况下，在不同的计算机上，也会有产生误差。

大家都知道身份证号是18位，如果是纯数字的身份证号，当输入excel表格时，在数字格式下其最后三位数会自动变为0。即输入的是123456789012345678，结果excel会自动将其变为123456789012345000。

要解决精度引起的误差，不是本文件的范围。

在实际应用中，单精度的有效数字已经足够了，excel等常用的具有很强计算功能的软件有很多，使用时注意修约问题就可以了，多说反而会引起不必要的误解。故删掉附录D这一资料性附录。正文中相应的提及也一并删除。

5、附录E

附录E为资料性的，给出了构建威布尔概率纸的方法。

6、附录F

删除ISO19003的附录F，因与7.2.1.1.8内容重复

附录F是资料性的附录，给出了简约的计算学生*t*值的公式（F.1）。该公式计算所得与*t*-值的误差不超过0.5%。

其实没有必要给出这种计算方法，一是本文件给出了*t*-值表，二是从统计学的书上很易查到，最方便的是可以用常见软件计算。我国有关统计学的教材很多都推荐使用软件计算，如excel。用此公式计算还不如用excel计算来的方便。所以删除ISO19003的附录F.

以下是用(F.1)计算的结果与查t-值表所得值的对比

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | *n* =5 | | *n* =10 | |
| *t* | 计算 | excel/t-值表 | 计算 | excel/t-值表 |
| *t*0.95 | 2.1310 | 2.1318 | 1.8325 | 1.8331 |
| *t*0.975 | 2.7737 | 2.7764 | 2.2604 | 2.2622 |
| *t*0.99 | 3.7373 | 3.7469 | 2.8167 | 2.8214 |
| *t*0.995 | 4.5819 | 4.6041 | 3.2405 | 3.2498 |

7、附录G 方差分析

本附录为资料性附录，介绍了单因素方差分析的方法和多因素分析的方法，给出的计算公式。

对于方差分析，统计学的教材中都有详细介绍的。如贾俊平的《统计学（第7版）》，其中还给出了用Excel进行方差分析的操作步骤。

8、附录H(资料性)

本附录给出了用于计算一元多次回归方程系数的公式。

正如（七）中计算验证时所述，利用附录H并不能给出正确的结果，利用excel进行计算所得结果与之一致。同时，也很难查到ISO 19003所列各回归方程系数公式的出处。所以将其删除，换之以用excel计算多元函数回归的方法。

9、附录I(资料性)

本附录给出了一种叫做“intercal”的方法，用以校正实验室间的检验试验经常产生令人惊讶的大的重现性R值的问题。

针对相对于整个试验的总平均值，一个特定试验持续地倾向于产生低或高的平均值的特定的特定实验室，进行干预、校正。

Intercal是这种方法的名子，字典中找不到，经查文献(参考文献[13])，Intercal是INTERlaboratory CALibration的缩写，即“实验室间校准”的意思。

INTERCAL的概念是一种使用独立获得的校准曲线来校正实验室间有偏差差异的程序，它大大降低了实验室间的再现性R值;前提是:(1)所有实验室都具有良好的实验室内精度或较小的r;(2)检测水平在一段时间内保持稳定。

但是，本附录中的一个公式（I.1）：

校正值 = 观测值 -（截距+斜率-观测值）？

在逻辑上存疑，故拟删除本附录I。

如果想了解此方法，可以参阅参考文献[13]-[15]。正文中对应的16.2.3中也予以提示。