ALGORITHMS TO CALCULATE SAMPLE SIZES FOR INSPECTION SAMPLING PLANS

核视察中取样数量的计算方法

乔盛忠　薛维明　译

原子能出版社

核视察中取样数量的计算方法

乔盛忠
薛维明 译

原 子 能 出 版 社

图字:01-2002-5087 号

图书在版编目(CIP)数据

核视察中取样数量的计算方法/乔盛忠,薛维明编著.—北京:原子能出版社,2002.9

ISBN 7-5022-2665-6

Ⅰ.核… Ⅱ.①乔…②薛… Ⅲ.核工程-工程材料-采样-数量-计算方法 Ⅳ.TL34

中国版本图书馆 CIP 数据核字(2002)第 052271 号

内 容 简 介

本书介绍了国际原子能机构核材料视察工作中定量及定性分析随机取样数量的精确及近似两种计算方法,这对我国开展核材料视察工作及核燃料循环中随机取样分析和抽样复验均具有指导作用和参考价值。

原子能出版社出版 发行

责任编辑:张梅

社址:北京市海淀区阜成路 43 号 邮政编码:100037

北京朝阳科普印刷厂印刷 新华书店经销

开本:787×1092mm 1/16 印张 9.25 字数 230 千字

2002 年 9 月北京第 1 版 2002 年 9 月北京第 1 次印刷

印数:1—300

定价:18.50 元

目 录

1. 前　言

本报告用于确定在一个给定的核物料层中随机抽取视察样品的数量，它们是属性检验中申报核实数据的基础。众所周知，如果探测目标量的探测概率 $1-\beta$ 对于每层物料都能达到，那么，不管目标量在物料层中的缺损如何分配，对于所有的物料层，这个探测概率都能很好地达到，并且假定所检查的每个缺损物项，都具有较高的探测概率。

在历来的样品数量的表述中，最精确的测量方法（偏倚检验）具有一个计算值，利用该计算值可获得不同统计量的足够小的方差 $\hat{D}$，以致允许分配在所有物料层中的目标量，如同很小的错误数据（偏差）均可被探测。样品数量问题的方差 $\hat{D}$ 不是新算法的一部分，因为经验证明探测一个属性检验的小缺损所需要的偏倚检验的样品数量一般大于 $\hat{D}$ 检验所需要的样品数量。即使某些个别物料层不是这种情况，本报告所描述的采用本算法所找到的所有物料层样品数量的 $\hat{D}$ 方差，将很少大于所希望需要探测目标量的概率。实际上，能给出可探测的目标量。

上面提出的建议是，将不对物料层中的 $\hat{D}_k$ 进行方差检验。因此，在物料层 k 中具有 $\hat{D}_k$ 的目标量的探测概率未被列入视察计划。对于所有物料层，$\hat{D}$ 很重要，正如本报告所指出的，可根据有关属性，通过逐层选择视察样品的数量，把 $\hat{D}$ 的方差自动控制在某个适当值。在研究一个很大的 $\hat{D}$ 时，对于单个物料层，把更仔细地检查 $\hat{D}_k$ 作为一个诊断工具是很重要的，但不能作为一种探测设备。

本报告的结构如下，下一节列出了表示方法，输入数据和输出数据的要求相同。在第 3 节中，提出了计算属性检验视察样品数量的一般方法，并描述了历来机构所采用的与运行者所采用方法的不同之处，以及新近的一些建议算法。第 4 节提供了用于这些算法的关键方程，这些方程没有给出描导，为方便感兴趣的读者，在附录中给出了有关推导。所应用的两套方程分别对应上限和下限两种情况的每一物项的材料量。在第 5 节中讨论了一种初始样品设置的反复迭代，以产生最后的样品数量。

在迄今的讨论中，有一个前提，那就是视察员在计算样品数量时，能使用个人计算机，实际上，本算法要求采用个人计算机。如果没有个人计算机，则可使用第 6 节中介绍的近似的算法。这些近似的样品数量能有效地用作本算法中的初始样品数量，以减少所需要达到最终样品数量设置的迭代数量。

本报告还包括一些附录，这些附录列出了演算结果，并讨论了一个非定量 1 级检验（总体缺损检验）如何影响样品数量测量的问题，附录中也给出了许多实例。

2. 符号和术语

β＝设计非探测概率；

Q＝某些特殊分额缺损数的非探测概率；

M＝目标量（一般 1 倍的显著量）；

x = 每一物项的平均(^{235}U)量,与 M 具有同样的单位;

注:x 的使用与更通常的 $\bar{x}$ 不同,x 是一个简化符号。

N = 在物料层中的物项数。

假设有三种测量方法或检验内容,运行者与视察员之间的相对标准偏差对于方法 i 为 δ_i,$i=1\sim3$。为便于理解,如果 $\delta_i=0.05$,也就是说标准偏差为物项量(元素或同位素)真实未知值的 0.05 倍(或 5%)。在一个给定的测量方法中,产生一个定量或半定量的结果,见附录 4。

一个附加参数为 A,$A=t_a$,定义为属性检验的显著性水平(如果 $A=4$,临界值为 4 倍的标准偏差;如果 $A=3$,临界值为 3 倍的标准偏差等。见 4.4.1 节)。

样品数量的计算,其输入参数有 β, M, x, N, δ_i 和 A。主要感兴趣的输出对于方法 i 是样品的数量,用 n_i 表示,$i=1,2,3$。在后面的报告中还描述了一个表征取样计划的附加输出,以及在制定表征取样计划所要求的补充说明。

3. 确定样品数量的一般方法

3.1 物料层中的总样品数量

物料层中总样品数量由下式给出

$$n = N(1-\beta^{x/M}) \tag{3.1}$$

问题在于获得样品数量 n_i,即在测量方法中配置的 n。

3.2 针对测量方法,样品数量 n 的配置

在仅有一种测量方法的情况下,很简单,样品数量 n_1 就是 n。在存在两种测量方法的情况下,对于一种和二种测量方法的样品数量分别为 n_1 和 n_2。

$$n_1 = n - \eta_2 \tag{3.2}$$

$$n_2 = \eta_2 \tag{3.3}$$

本报告提供了计算 η_2 的程序。

如果存在三种测量方法,那么样品数量 n_1,n_2 和 n_3 分别为:

$$n_2 = \eta_2 - n_3 \tag{3.4}$$

$$n_1 = n - n_2 - n_3 \tag{3.5}$$

对于所有的情况,本报告提供了一个计算 n_3 的程序

$$\sum_i n_i = n \tag{3.6}$$

3.2.1 历史上的配置方法

历史上寻找 η_2,把这个程序定义为部分缺损,这是一种量的缺损,或许不能被 1 级检验(一种测量方法)所探测。如果任意设置一个问题的概率为 0.025,那么这个固定的概率值导致一个定义为实际的部分缺损。

用符号 $\gamma_2 x$ 表示一个 2 级部分缺损的数量,也就是说一个缺损不能被具有低概率的 1 级

检验所探测，那么 0.025 的指定概率导出了 γ_2 的定义。在属性检验中，方程假定了 4 倍标准偏差($A=4$)的剔除限制。

$$\gamma_2 = \frac{2\delta_1}{1 \pm 2\delta_1} \tag{3.7}$$

上式中的减号用于上限，加号用于下限。上限描述的是存在的材料少于申报材料的情况；下限描述的是存在的材料多于申报材料的情况。这个熟知的样品数量的公式对于 η_2 类似于式(3.1)，所不同的是包括因于 γ_2 作为 x 的一个系数。

$$\eta_2 = N(1 - \beta^{\gamma_2 x/M}) \tag{3.8}$$

实际上，$M/\gamma_2 x$ 是部分缺损的一个数，正如在式(3.1)中的 M/x 是总体缺损中的一个数。两种描述都假定全部缺损是大小相同的，即对于部分缺损等于 $\gamma_2 x$，对于总体缺损等于 x。

方程(3.8)的应用，要求 δ_2 应小到 $\gamma_2 x$ 缺损的探测具有高的探测概率。究竟 δ_2 要小到什么程度，若干年前的文献[3,4]已给出量化的结果。遗憾的是 δ_2 的要求不总是能满足，并在这种情况下，为了满足$(1-\beta)$的探测概率，式(3.8)给出的样品数量太少。但是，通常当缺损的数量为 $\gamma_2 x$ 时，找到符合要求的$(1-\beta)$值的样品数量是可能的。文中给出了一个二级缺损的数量 $\gamma_2 x$，令 p_2 为概率，缺损将被适当地分类，如果采用二种测量方法，那么表示在附录 1 中的公式是

$$\eta_2 = \frac{\ln\beta}{\ln\left[1 - \frac{m(\gamma_2)p_2}{N}\right]} \tag{3.9}$$

$m(\gamma_2)$被称为二级缺损数，显然可由下面的方程给出

$$m(\gamma_2) = M/\gamma_2 x \tag{3.10}$$

上面的开发加以扩展，可给出 n_3 的计算公式。用 γ_3 和 p_3 分别代替式(3.9)中的 γ_2 和 p_2，借助这种方法可以任意定义部分缺损。γ_3 的公式具有与式(3.7)同样的形式，用 δ_2 代替 δ_1 给出 γ_2 和 γ_3。概率 p_2 和 p_3 按如下公式进行计算，参数 p_2 是采用二种测量方法对缺损数量 $\gamma_2 x$ 进行探测的概率。这样，p_2 的值用如下的积分进行计算。

$$p_2 = \frac{1}{\sqrt{2\pi}}\int_{u_2}^{\infty} e^{-x^2/2}dx = 1 - \phi(u_2) \tag{3.11}$$

在这里

$$u_2 = \frac{A\delta_2 - \gamma_2}{\delta_2(1 \pm \gamma_2)} \tag{3.12}$$

减号用于上限，加号用于下限。参数 p_3 可类似地进行定义，用角标 3 代替式(3.11)和式(3.12)中的 2。

总之，方程(3.1)～(3.12)组成了一种确定视察样品数量的方法；但是这种方法未被普遍使用。工作单已制备以允许样品数量计算的使用，但它们不是机构官方程序文件的一部分[5]。方程(3.7)的简单使用和它相应的 3 级缺损与具有 γ_2(或 γ_3)(3.1)方程的扩展相结合，引入系数 x，如果 p_2(或 p_3)接近于 1，那么它能给出上述公式等效的样品数量。但是，如果 δ_2 接近于 δ_1，那么 p_2 显著不同于 1，并且二级检验采用式(3.8)计算的样品数量太小，导

致一个小于$(1-\beta)$的探测概率。

3.2.2 历史方法的评估

正如概要中对一些方法的评价，γ_2 和 γ_3 可任意定义。基于一个很小概率的任意定义，缺损数量 $\gamma_2 x$ 应由一种测量方法检验所探测，缺损数量 $\gamma_3 x$ 应由二种测量方法检验所探测。属于复述，基于起发性的原因，这个小的概率被历史性地设置为 0.025，并将在下节要讨论的建议计算方法移到了本节。

事实上，在方程中任意指定 0.025 的概率是偶然的，它导致在许多情况下合理的样品数量。样品数量的配置能导致非探测概率接近于规定的 β 值，在某种意义上说是合理的。选择 0.025 值的实际建议具有重要的见识，它不伤害某些人为的需要、样品数量配置的计算和提供一种基于本算法的诊断措施。

一般说来，这是特别重要的，因为 0.025 的值对于二种测量方法和三种测量方法(部分和偏倚缺损检验)会趋于产生过多的样品，以至消耗了多于测量所需要的有价值的资源。而且，无论什么时候 δ_i 和 $\delta_{i+p}(i=1,2)$，成为接近的值，这样式(3.11)中的 $p_2(p_3)$显著地不同于 1。因此，采用一般的样品数量公式(3.9)与 0.025 的概率值一道，将引起严重的增大，取样数量；反之采用方程(3.8)会导致严重的缩小取样数量。在下面的三节中介绍的计算方法将校正这些缺陷。

4. 初始样品数量的计算

4.1 最佳化程序的概要

在一定的意义上，在此所描述的计算方法是量佳的，它找到了一套最少的初始样品数量。这些最小的样品数量保证在三种缺损数量(在一个物料层中的两种情况，具有两种测量方法)，具有$(1-\beta)$的探测概率。三种缺损数量是 x(总体缺损)，$\gamma_2 x$ 和 $\gamma_3 x$，在这里特别选择 γ_2 和 γ_3，以使 η_2 和 n_3 分别减至最小。

事实上，在这三种缺损数量情况下，探测概率都达到了$(1-\beta)$，但不保证在别的情况下也能很好地达到。正如所知，尽管 η_2 和 n_3 将很可能不小于由最佳化程序所计算的值，但它们其中的一个或两个有必要具有较大的值，以确保在所有的缺损数量达到$(1-\beta)$的探测概率。为了评价这个，非探测概率 Q 采用初始的计算值 n_1，n_2 和 n_3(对应于三种测量方法)，对于大量的缺损数量进行了计算。找到了 Q 的最大值，并且如果 Q 小于 β，样品数量的初始设置成为最后的设置。如果 Q_{max}的最大值超过 β，那么不是 η_2 就是 n_3 适当地增加，保持 n 不变，以便在此问题中的缺损数量使 Q 小于 β。样品数量的修改可引起 Q 趋过 β，此后迭代过程继续进行直到 Q_{max}小于 β，如果通过合理地视察尝试，假设的非探测概率 β 是可能达到的。

正如引言中所指出的，已经开发了近似的计算方法，加上一个袖珍计算器能进行简单的应用，这个开发是 6 节的主题。作为近似方法的另一应用，可减少样品数量最后设置的重复迭代。由本计算方法及 6 节中的近似公式所确定的初始样品数量的讨论将在 4.4 节中介绍，最后的计算方法结合了这个特点。然而，为了开发本近似方法，4.4 节的讨论，即初始样品数量

的计算方法是需要的。

4.2 强制限制

对于某些物料层,注意到了小的缺损数量要达到$(1-\beta)$的非探测概率是不可能的。当这个物料层的物项数是很大时,这种现象是可能发生的,因此目标量 M 可以在大量物项中分配,并且为了探测这些很少的缺损,即使用最精确的测量方法 δ_i 也太大。当然这种情况对于本计算方法不是特殊的,它是一个可以辨别的潜在问题。在这里所讨论的计算方法固定在一个物料层中的总样品数量为 n(正如描述在3.2.1节中的计算方法)。处理在某种意义上不能达到$(1-\beta)$的问题将在5节中介绍。

作为一个二级强制,当两个相继的 δ 值,即 δ_1 和 δ_2 是很接近的值,从统计的观点看,它与所采用的测量方法有关。具体地说,如果 $\delta_i/\delta_{i+1}<1.25$,测量方法 $i+1$ 要从样品数量计算方法中除去。当然,在应用中,方法 $i+1$ 可被用于 i 的位置,或者某些测量可用各自的方法(比较好的方法)来完成。

这种计算方法有两个特点值得注意:第一,因为 M/x 实际上是一级缺损(总体缺损),必然它是一个整数,并且在完成式(3.1)的计算之前,它应当被四舍五人到下一个整数。$M/\gamma_i x$ 也是如此,但它在此情况下四舍五人并不重要,因为通常 $M/\gamma_i x$ 是一个大数。第二,为了在物料层中累积目标量 M,缺损数量不能小于 M/Nx,因为产品 $N\gamma_i x$ 必须等于M。这意味着对于 γ_i 的限制,在用一个方程如式(3.7)计算 γ_i 时,有时会被忽略,但在这里提供的只是计算方法的一部分。

4.3 误差模型:随机误差或系统误差

在样品数量计算方法的开发中,对于 δ_i,它是做了两种极端假设条件下的数学模型。一方面 δ_i 被假设为随机误差;另一方面 δ_i 被假设为系统误差。实际上,真值存在于两个极值之间的一个范围内,在那里问题的实际处理变得很复杂。

在两种假设下,计算了大量的例子。正如所期望的,对于二种或三种测量方法的样品数量,系统误差模式比随机误差模式更大。但也有例外,样品数量的差别不大。作为这种现象的一个解释,是由于在计算感兴趣范围的非探测概率的重要参数时,在取样物项中存在一个或更多给定数量的缺损。第二个重要因素是缺损的适当分级,随机误差和系统误差模型之间的差别仅在于影响分级问题。

大量的测量数据表明,随机误差模型是更合理的。如果误差是纯系统的,那么同一物项的重复测量应当给出相同的结果;但结果与经验是相冲突的。按照上述的讨论,本报告做了基于一种随机误差模型样品数量配置计算方法的决定,在本报告中没有进一步地进行系统误差模型计算方法的叙述。

4.4 初始样品数量的确定

4.4.1 属性检验剔除限制的选择

为定义在属性检验中的剔除限制,在第2节中的符号和术语引入了参数 A。具体地说,一个4倍标准偏差的剔除限制,$A=4$;一个3倍标准偏差的限制,$A=3$ 等等。对于 γ_2,也注

意到熟悉的方程(3.7),假设 $A=4$ 作为通常的实践。

经进一步的考虑,决定选择 $A=3$ 作为视察计划的基础,这个决定主要基于以下三点考虑:

(a)不同 NDA 方法灵敏度和稳定性的改进,提高了这些测量方法的置信度。原先所选择的便于保持很小误报警概率的历史值 $A=4$,看来似乎是不正确的;

(b)研制在事件中精密的跟踪程序[6],找到样品中的部分缺损,以减少保持极低误报警率所需的样品;

(c)$A=3$ 和 $A=4$ 一对一地应用,导致二种测量方法或三种测量方法(部分或偏依缺损检验)样品数量明显地减少[(a)和(b)是选择 $A=3$ 的强制理由;(c)是这个决定所希望的结果]。

此后,在本报告的主体中,假设 $A=3$。但在附录 1 的推导中,参数 A 允许普遍性的应用。

4.4.2 初始样品数量的计算

本节中结果的推导包括在附录 1 中。首先找到了 η_2 和 η_3 最小的初始值。对于式(3.9)中的 η_2 是近似的,当产品 $m(\gamma_2)p_2$ 达到最大时,η_2 是最小值。为了简化,不管脚标,并用式(3.10)的等效形式代替 $m(\gamma_2)$,解方程 γ 是

$$\frac{e^{-M^2/2}}{\sqrt{2\pi}}\,\frac{(1\pm 3\delta)}{\delta(1\pm\gamma)^2}=\frac{1}{\gamma\sqrt{2\pi}}\int_M^{\infty}e^{-x^2/2}\mathrm{d}x \tag{4.1}$$

注意这个解不依赖于 M/x 的比值。在式(4.1)中减号应用于上限,加号应用于下限。γ,u 和 δ 的脚标不是 2 就是 3,取决于是解式(4.1)中的 γ_2 还是 γ_3。

虽然求解方程(4.1)似乎是很困难的,但用逐次逼近法(试凑法)容易被求解。下表给出了 γ_i 作为 δ_i 的函数,$i=2$ 或 3。

表 4.1 最小 mp,期望可探测缺损数的 γ 值

δ	上限 γ 值	下限 γ 值
0.15	0.564	0.605
0.14	0.531	0.568
0.13	0.497	0.530
0.12	0.463	0.491
0.11	0.428	0.452
0.10	0.392	0.412
0.09	0.356	0.371
0.08	0.318	0.330
0.07	0.280	0.289
0.06	0.242	0.249
0.05	0.203	0.208
0.04	0.163	0.166
0.03	0.123	0.125
0.02	0.082	0.083
0.01	0.041	0.041
0.005	0.020	0.020

表 4.1 给出了 γ 和 δ 之间的关系,也可由下列方程严密地逼近。

对于上限：

$$\gamma = -0.002 + 4.259\delta - 3.24\delta^2 \tag{4.2}$$

对于下限：

$$\gamma = -0.002 + 4.284\delta - 1.508\delta^2 \tag{4.3}$$

方程(4.2)和(4.3)的 γ 和 δ，不是具有脚标 2 就是 3。值得注意的是当 η_2 的初始值最小时，方程(3.7)指出，γ_2 的值取决于 δ_2，而不依赖于 δ_1。进而 γ_3 取决于 δ_3，而不依赖于 δ_2；但是，$\gamma_2(\gamma_3)$的最终值不依赖于 $\delta_1(\delta_2)$值的结论使人误解。由式(4.2)或(4.3)给出的 γ_i 值使最后样品数量 η_2 或 n_3 成为最小的值。对于这个最小的样品数量，要求的非探测概率 β 对于所有的部分缺损数量可以不被达到，并且样品数量可以需要一个向上的调整。β 是否达到具有最小的样品数量，取决于所有 δ 值的一部分。实际上 γ_2 取决于 δ_1，γ_3 取决于 δ_2；但是这个依赖不是任意地，其目的在于选择它的相关情况。

为了小结在讨论中的 η_2 已被达到，其计算步骤如下：

(1) 由(4.2)或(4.3)计算 γ_2；

(2) 由(3.10)计算 $m(\gamma_2)$；

(3) 由(3.11)和(3.12)计算 p_2；

(4) 由(3.9)计算 η_2；

(5) 由(3.1)和(3.2)计算 n_1。

实际上，对于三种测量方法，计算程序是相同的。η_2 的计算按照上面所给出的(1)到(4)步。n_3 的计算重复(1)到(4)步，脚标由 2 变到 3。那么样品数量 n_1 和 n_2 的计算用方程(3.1)，(3.4)和(3.5)。

5. 用迭代法选择最后样品数量

已找到了样品数量的初始值，下一步是确定在全部缺损数量下，所要求的非探测概率的最小值是否足够大。为此，计算作为部分缺损数量函数的非探测概率是需要的。

5.1 非探测概率的计算

5.1.1 用于概率计算的部分缺损数量

根据非探测概率计算缺损数，并用 m 表示。假设单一物项的缺损为 m，那么全部物项的缺损数量是 M/m，m 的最小值是

$$m = M/x \tag{5.1}$$

即是总体缺损数。m 的最大值是 N，即是所有的物项都被缺损，接近量 M/N。对于 m 的中间值 M/x 和 N，已完成非探测概率的计算。计算方法给在附录 7 中，80 个 m 中间值的形式是

$$m = (1/m_0 - ir)^{-1}; i = 0,1,2,\cdots,79 \tag{5.2}$$

式中

$$r = \frac{(1/m_0 - 1/N)}{80} \tag{5.3}$$

$$m_0 = \frac{M(1 \pm 3\delta_1)}{6\delta_1 x} \tag{5.4}$$

在式(5.4)中,加号用于上限,减号用于下限。参数 m_0 表示根据非探测概率所计算的第二个最小 m 值,方程(5.4)意味着缺损 M/m_0 具有高的探测概率。所以,对于 m_0 缺损的非探测概率远小于 M/x 缺损的非探测概率。因此,Q 在缺损M/x 和 m_0 之间不能超过 β,在计算中本范围内存在少数非探测概率的点。

5.1.2 每种方法的非探测概率

对于一个给定的缺损数 m,如果存在至少 1 个缺损用至少一种测量方法探测,那么定义了产生数据失真的探测概率。对于每一个 m 值,计算非探测概率比计算探测概率更简单,该计算是

$$\begin{aligned} Q &= \beta_1\beta_2 \qquad \text{两种测量方法} \\ Q &= \beta_1\beta_2\beta_3 \qquad \text{三种测量方法} \end{aligned} \tag{5.5}$$

β_1 = 检验方法 1 无缺损的探测概率

β_i = 检验方法 i 无缺损的有条件的探测概率,对于所给定的检验方法 $j<i$ 无探测,$i=2,3$。

在 β_i 的计算中,采用了一种近似,Q 的精确计算能被完成;但对于精确的计算结果是很繁琐的,而近似公式经鉴定是可接受的。附录 2 给出了问题的精确计算公式,并提供了一些与结果近似的数字进行比较。

对于给定的 m,可用式(5.6)或式(5.7)计算 β。

如果 $n_1 \leqslant mq_1$,

$$\beta_1 = \left[1 - \frac{mq_1}{N - 0.5(n_1 - 1)}\right]^{n_1} \tag{5.6}$$

如果 $n_1 > mq_1$,

$$\beta_1 = \left[1 - \frac{n_1}{N - 0.5(mq_1 - 1)}\right]^{mq_1} \tag{5.7}$$

[见附录 1,方程(A.9)和(A.10)]。

在这些方程中,如果测量已完成,那么 q_1 是缺损为 M/m 的检验方法 1 的探测概率。它由如下公式给出

$$q_1 = 1 - \phi(\nu_1) \tag{5.8}$$

$$\nu_1 = \frac{3\delta_1 x - M/m}{\delta_1(x \pm M/m)} \tag{5.9}$$

式中,减号用于上限,加号用于下限。

下面考虑 β_2 的计算,采用最简单的形式,用 β_1 一样的方法计算 β_2,分别用 n_2 和 q_2 代替 n_1 和 q_1。在某些应用中,它是适当的,它代表第二个从整个物项数 N 中抽取 n_2 个独立的样品,而不是从剩余物项 $N-n_1$ 中抽样。如果 n_1 小于 N,那么取样无论是用 $N-n_1$ 代替 N 或者不代替,它仅存在很小的差别;但在某些应用中,特别是当 x 大于M 时,n_1 相对于 N 是很大的。这样,通常需要进行更多的校正,因此,在用式(5.6)和式(5.7)计算 β_2 时,应当用 $N-n_1$ 代替 N。然而,这样将减少 n_1(即取样无替换),我们必须研究需从 m 中减去某个量。记住 β_2 是有条件的概率,它是在第二样品数量 n_2 中没有缺损时被探测,给出在初始样品数

量 n_1 中没有缺损被探测。如果在初始样品中未探测到缺损，也就是说未发生缺损，那末所给出的被测量的一个适当分类缺损的概率仅为 1。对于 m 的某些值，这个概率不是 1，但很小。因此，在样品数量 n_1 中估算缺损数的某些方法，给出没有缺损被探测是可能的。

令缺损数为 w_1，表示在附录 3 中的 w_1 是以下二次方程的解

$$q_1 w_1^2 + (N - n_1 q_1 - m q_1) w_1 - m n_1 (1 - q_1) = 0 \tag{5.10}$$

应用已知的二次公式，很容易解出此方程中的 w_1，其解 w_1 是最接近的一个整数。然后计算 m_1，在剩余样品数 $N - n_1$ 中估算缺损的数目，即是

$$m_1 = m - w_1 \tag{5.11}$$

于是，用式(5.6)或式(5.7)计算 β_2，用 m_1 代替 m，用 N_1 代替 N。

$$N_1 = N - n_1 \tag{5.12}$$

β, q 和 n 的脚标用 2 代替 1。

当 n_3 物项从剩余物项$(N - n_1 - n_2)$中随机抽取时，可用同样的推导进行剩余缺损数的估算。w_2 是式(5.10)所表示的二次方程的解，w_1, q_1 和 n_1 的脚标是 2 而不是 1，并用 m_1 代替 m，N_1 代替 N，那么，

$$N_2 = N_1 - n_2 \tag{5.13}$$

$$m_2 = m_1 - w_2 \tag{5.14}$$

采用式(5.6)和式(5.7)对 β_3 进行计算，用 m_2 代替 m，N_2 代替 N，β、q 和 n 的脚标用 3 代替 2。

5.1.3 总的非探测概率

已找到了 β_i，$i = 1, 2$ 和 3(在事件中存在三种测量方法)，那么，当存在 m 个缺损时，总的非探测概率由方程(5.5)给出。

5.2 迭代法

已找到的非探测概率 Q，由方程(5.2)～(5.4)表示 m 的所有的值。那么，可找到 Q 的最大值，并用 Q_{max}表示。假设初始的 Q_{max}在 $m = M/x$ 和 $m = N$ 之间，如果 $Q_{max} < \beta$，就选择初始样品数量作为视察样品数量。如果 Q_{max}超过 β，样品数量的分配也会改变，但总的样品数量保持由式(3.1)给出的固定值 n。

首先，对于两种测量方法，样品数量做了如下的调整。对于这些情况，Q 是$\beta_1 \beta_2$ 的乘积。保持 n 固定，Q 值仅能随 n_2 的增加和 n_1 的减少而减少。于是，β_2 将减少，而 β_1 将增加，但乘积 $\beta_1 \beta_2$ 将减少。

表明 n_2 的量将被增加的关键方程将从式(5.6)导出，它出现适当的下标改变。

$$\beta_2 = \left[1 - \frac{m_1 q_2}{N_1 - 0.5(n_2 - 1)}\right]^{n_2} \tag{5.15}$$

目前不管 β_1 的变化，减少 β_2 的值到 β_2'，如

$$\beta_1 \beta'_2 = \beta \tag{5.16}$$

或者

$$\beta'_2 = \beta / \beta_1 \tag{5.17}$$

现在 β_2 是

$$\beta_2 = Q_{max}/\beta_1 \tag{5.18}$$

结合式(5.17)和式(5.18),β_2' 的解是

$$\beta_2' = \beta_2\beta/Q_{max} \tag{5.19}$$

令所需要达到 β_2' 值的样品数量是(n_2+c),它近似地遵循式(5.15)

$$\beta_2' = \beta_2^{1+c/n_2} \tag{5.20}$$

由式(5.19),c 的解是

$$c = \frac{n_2(\ln\beta - \ln Q_{max})}{\ln\beta_2} \tag{5.21}$$

于是,调整后的样品数量是:

$$\begin{aligned} &n_1 - c \quad (\text{一种测量方法}) \\ &n_2 + c \quad (\text{二种测量方法}) \end{aligned} \tag{5.22}$$

这些成为新的样品数量和非探测概率像以前一样采用这些样品数量进行了计算,迭代过程被重复到 $Q_{max} \leqslant \beta$。

对于三种测量方法,采用一种类似的计算;但不是 n_3 就是 n_2 将被增加,这决取于在 Q_{max} 之内的 β_2 和 β_3 的相对数量。如果在 Q_{max}之内,$\beta_3 > \beta_2$,c 用式(5.21)进行计算,并且标记为 c_2。如果 $\beta_3 \leqslant \beta_2$,$c$ 用式(5.21)计算,n 和 β 的下标从 2 变成 3,其结果为 c_3。下一个迭代的样品数量如下:

对于 $\beta_3 > \beta_2$,新样品数量是

$$\begin{gathered} n_1 - c_2 \\ n_2 + c_2 \\ n_3 \end{gathered} \tag{5.23}$$

对于 $\beta_3 \leqslant \beta_2$,新样品数量是

$$\begin{gathered} n_1 \\ n_2 - c_3 \\ n_3 + c_3 \end{gathered} \tag{5.24}$$

采用新的样品数量对非探测概率进行了重新计算,如果 Q_{max}在第一次迭代之后,仍然超过 β,则随第一次迭代之后的第二次迭代继续完成,迭代过程重复进行直到 $Q_{max} \leqslant \beta$,而 n_2 进一步地减小致使 $Q_{max} > \beta$ 点时迭代过程停止。

正如先前所注意到的,在某些情况下,因为测量误差的限制,要达到要求的 β 是不可能的。在大的 m 值时,这种情况将会发生,即当存在许多小的缺损(偏倚缺损)和当用最精确的方法(第二或第三种测量方法)探测如此小的偏倚缺损,标准偏差太大时,要达到要求的 β 概率是不可能的。在此种形势下,也隐含着一个大物料层的条件(即此物料层具有若干个目标量),因为对于小物料层,小的偏倚缺损将不会累积到目标量。

首先,确定 m 的上限,如果 m 大于这个上限,则$(1-\beta)$的概率不能达到。特别是,用 m_{00} 表示的该上限由缺损数量 M/m_{00}所确定,这个缺损数量用最精确的测量方法进行了探测,其探测概率为 0.16(即一倍的标准偏差)。对于小的概率,即使很明显的缺损包括在样品中,也

很可能鉴别不出来。表示在附录 2 中的 m_{00} 的方程是

$$m_{00} = \frac{M(1 \pm \delta_i)}{2\delta_i x} \tag{5.25}$$

公式里的减号用于上限,加号用于下限。当存在两种测量方法时,i 为 2;当有三种测量方法时,i 为 3。

已找到了 m_{00},所计算的 Q 最大缺损数是 N_0,

$$N_0 = \min(m_{00}, N) \tag{5.26}$$

如果 $N_0 = m_{00}$,打印出信息 β 是做不到的。在这个事件中,关于 Q_{max} 存在三种可能性:

(a) Q_{max} 不是发生在 $m = M/x$,就是发生在 $m = N_0$,即没有内部峰值;

(b) 具有一个内部峰值,即 Q_{max} 发生在 $m = M/x$ 和 $m = N_0$ 之间,但 $Q_{max} \leqslant \beta$;

(c) 具有一个内部峰值,并且 $Q_{max} > \beta$。

进一步的程序步骤取决于(a),(b)或(c)中的哪一种情况。对于(a)和(b)是两种测量方法的情况,方法可确定样品的数量 n_2。在这里,Q 等于由方程(4.2)或(4.3)给出的 β,γ_2 的初始值。如果有三种测量方法,可采用一个类似的程序确定 n_3,Q 在 γ_3 的初始值处等于 β。

对于情况(c),迭代过程描述在 5.2 节中,即样品数量 n_2 和 n_3 被调整到 $Q_{max} \leqslant \beta$。

如果 $N_0 = N$,至少用实际的样品数量不能获得 β 的概率仍是可能的。方法采用限制搜索 Q_{max} 处理这种情况,令 m 的最大值为 N(即 $m = N$),并用 L 表示,下一个最大值用 $L-1$ 表示等等,在找到 Q_{max} 之前的所有

$$Q_{i-1} < Q_i; i = L, L-1, L-2, \cdots$$

Q 值都被取消。这种方法确保在可控范围内,Q_{max} 将不会超过 β;但对于大量小的缺损的可能性,将不会达到所希望的 β 概率。

6. 近似样品数量的计算公式

计算方法对于第 4 和第 5 节的应用,实际上要求用户具备一台具有存取能力的个人计算机(PC)。借助于 PC 机,计算方法要求很少的输人数据和执行时间为数秒的程序。

然而如果用户没有存取 PC 机,则算法中的方程不适合于日常的应用。虽然单个计算不复杂,但是对于一个完全的解所要求的计算数量是很大的。为了解决这个问题、不排除使用一种近似的计算方法。

本节的目的在于提供一种计算方法的近似形式,这些近似的形式可容易地用一个袖珍计算器进行计算。

在近似计算的情况下,样品数量的计算公式将被开发,基于 PC 机计算样品数量的方法力图能被很好地简化。这种简化的计算方法,采用 4.4.2 节中的样品数量近似计算方法代替初始样品数量的计算,这种改进减少了要求达到最后样品数量迭代的次数、完整的计算方法的修改,在 6.2 节中进行了讨论。在此,首先讨论近似计算方法的开发。

6.1 近似计算方法的开发

开发近似计算方法的第一步是采用 4 和 5 节中的算法在 PC 机上对于大量的例子进行样

品数量的计算，这个练习鉴别了初始样品数量也在最后样品数量的范围内。它也指出在具体而变化的条件下，样品数量的配置如何不同于初始样品的配置。

近似计算方法的一般形式如下，对于方程(3.9)和(3.10)，无论 η_2 或 n_3 一般样品数量方程，不管下标，可写成

$$n = \frac{\ln\beta}{\ln(1 - M/\gamma' Nx)} \tag{6.1}$$

这里

$$\gamma' = \gamma/p \tag{6.2}$$

目标是开发 γ' 的近似计算方法，则公式(6.1)对于已知的 β, M, N, x 和 γ' 可以采用一个袖珍计算器进行计算。

方程(4.2)和(4.3)分别用于对应 γ 上限和下限，这些方程用于 γ_2 和 γ_3 的计算。它有待找到 p(即 p_2 或 p_3)，以便采用式(6.2)导出 γ' 的表达式。对于给定的 γ 和 δ，并令 $A=3$，采用方程(3.11)和(3.12)计算 p，一个计算例子将有助理解本近似计算的开发。

假定 $\delta=0.10$　考虑上限情况，然后由式(4.2)得

$\gamma=0.392$；

由式(3.12)得：$u=-1.513$；

由式(3.11)得：$p=0.935$；

由式(6.2)得：$\gamma'=0.419$。

表 1　γ' 作为 δ 的函数

δ	上限				下限			
	γ	p	γ'	PRED	γ	p	γ'	PRED
0.15	0.564	0.959	0.588	0.587	0.605	0.739	0.819	0.821
0.14	0.531	0.955	0.556	0.556	0.568	0.748	0.759	0.760
0.13	0.497	0.950	0.523	0.523	0.530	0.758	0.669	0.700
0.12	0.463	0.945	0.490	0.489	0.491	0.767	0.640	0.641
0.11	0.428	0.941	0.455	0.455	0.452	0.776	0.582	0.582
0.10	0.392	0.935	0.419	0.419	0.412	0.785	0.525	0.525
0.09	0.356	0.931	0.382	0.382	0.371	0.794	0.467	0.468
0.08	0.318	0.924	0.344	0.344	0.330	0.802	0.411	0.413
0.07	0.280	0.918	0.305	0.305	0.289	0.810	0.357	0.358
0.06	0.242	0.913	0.265	0.264	0.249	0.821	0.303	0.304
0.05	0.203	0.908	0.224	0.223	0.208	0.832	0.250	0.251
0.04	0.163	0.900	0.181	0.181	0.166	0.839	0.198	0.199
0.03	0.123	0.894	0.138	0.137	0.125	0.851	0.147	0.148
0.02	0.082	0.885	0.093	0.093	0.083	0.855	0.097	0.098
0.01	0.041	0.875	0.047	0.047	0.041	0.855	0.048	0.048
0.005	0.020	0.846	0.024	0.024	0.020	0.836	0.024	0.024

表 6.1 给出了进一步的计算结果，表栏 PRED 列出的是 γ' 为 δ 函数经验方程的结果。方程的上限和下限分别是

上限：
$$\gamma' = 4.737\delta - 5.49\delta^2 \tag{6.3}$$

下限：
$$\gamma' = 4.802\delta + 4.47\delta^2 \tag{6.4}$$

当初始样品数量与最后样品数量相同时，找到了 γ' 作为 δ 的函数的表达式，下一步是鉴别发生这种现象的范围，以及导出在鉴别范围之外的应用方程。为了达到这个目标，计算了大量的例子，还给出了一些推导的结果和另外的设置以确定它们的应用能力。

对于一套给定的最后样品数量，如果已知 n,β,M,N 和 x，采用方程(6.1)可反算 γ'。n 的值必须是整数，因为真实的 β 将不等于实际的输入 β。近似计算方法开发的目的，n 取整数值是用于达到实际的 β。

忽略详细的推导，开发了下列经验公式

对于上限：

$$\gamma'_2 = 4.737(\Delta_2) - 5.49(\Delta_2^2) \tag{6.5}$$

这里

$$\Delta_2 = \max(\delta_2, 0.0075 - 0.0531\delta_1 + 2.369\delta_1^2) \tag{6.6}$$

$$\gamma'_3 = 4.737(\Delta_3) - 5.49(\Delta_3^2) \tag{6.7}$$

这里

$$\Delta_3 = \max(\delta_3, 0.331\delta_2) \tag{6.8}$$

对于下限：

$$\gamma'_2 = 4.802(\Delta_2) + 4.47(\Delta_2^2) \tag{6.9}$$

这里

$$\Delta_2 = \max(\delta_2, 0.162\delta_1) \tag{6.10}$$

$$\gamma'_3 = 4.802(\Delta_3) + 4.47(\Delta_3^2) \tag{6.11}$$

这里

$$\Delta_3 = \max(\delta_3, 0.208\delta_2) \tag{6.12}$$

式(6.5)，式(6.9)和式(6.11)中的系数被看作与式(6.2)和式(6.4)中的系数相同。这四个方程的鉴别范围，初始样品数量是在最后样品数量之内。

方程(6.5)～(6.12)在表 6.2(上限)和 6.3(下限)中的表格的形式表示。这些表给出了需要计算近似样品数量的所有信息。

为举例说明，考虑如下的输入设置：

$$\beta = 0.05 \qquad \delta_1 = 0.05$$

$$M = 25 \qquad \delta_2 = 0.03$$

$$x = 0.328(\text{上限情况})$$

$$N = 800$$

从表 6.2 于表 1 可知，$\Delta_2 = 0.0108$，因为 $\delta_2 > 0.0108$，用于表 2 可找到 $\gamma'_2 = 0.137$。那么

$$n = 800(1 - 0.05^{1/77}) = 30.53 = 31 \quad \text{由式(3.1)}$$

$$\eta_2 = \frac{\ln 0.05}{\ln[1 - 25/(0.137)(800)(0.0328)]} = 2.52 = 3$$

$$n_1 = 28 \quad \text{由式(3.2)}$$

$$n_2 = 3 \quad \text{由式(3.3)}$$

第二个例子包括三种测量方法

$$\beta = 0.10 \qquad \delta_1 = 0.12$$