1994年5月 重庆大学学报 Vol. 17, №. 3
第17卷第3期 JOURNAL OF CHONGQING UNIVERSITY May. 1994

·研究简报· (22) 127-130

多重混合样本的平均受检次数及其猜测式*

The Proof of a Conjecture for the Formula of Expected Inspection of Multiple Pooled Sample Test

黄 薇
Huang Wei
（重庆大学系统工程及应用数学系，重庆，630044）

O212·2

A **摘 要** 温启愚首先提出二重混合检查策略的概念并对n重混合检查的平均受检次数作出了一个猜测，本文运用概率论的思想简洁地证明了这一结果；其次又独立地提出了n重混合平均受检次数公式的另一形式，并证明了两类公式的等价性。

关键词 检验；多重混合样本；平均受检次数

中国图书资料分类法分类号 O212.2；O211.1

ABSTRACT The double sample test was proposed first but no proof by Wen Qiyu. This test is more saving than the simple pooled sample test. The paper[1] also gave a conjecture for the formula of expected inspection number of n-fold pooled sample test. This paper first gives a simple proof of the conjecture by probabilistic methods and obtains another form of the conjecture independently and proves the two conjectures are equivalent.

KEYWORDS test; multiple pooled samples; expected inspection number

0 引 言

抽样检查对于判断整批是否合格可以节约人力物力，但在有的情况下，需逐个检查检出缺陷（或称阳性）个体。R. Dorfman[2]在第二次世界大战中提出了不必逐个检查的混合检查法（即一重混合检查策略），首先在军队中用来普查性病患者，对节约检查量收到了显著效果。Feller 在专著[3]中也提到了一重混合检查法，即：当受检样品个数较大，而每个样品的阳性率相同且较小，并只判定是否呈阳性时，将适当数量的样品混在一起检查，呈阳性时，才逐个检查，比一开始就逐个检查要节约检查量。这种方法后来被广泛地用于某些疾病的普查、大面积污染的调查、运动员服违禁药品的某些项目的检查、缺陷产品的检出等方面，节约了大量人力物力，提高了时效。温启愚[1]从增加节约量的角度出发首先提出了二重混合检查策

* 收文日期 1992-11-11

略、进一步更提出 n 重混合检查策略的概念。显然，样品的平均受检次数越小越能节约检查量，因此，在运用此策略时，首先必须求出平均受检次数。本文主要是解决有关 n 重混合检查的平均受检次数的理论问题。

文[1]给出了一重混合检查策略平均受检次数的公式。如设阳性率为 p，而 $q=1-p>0$，将 k 个样品混为一组检查，呈阳性时才逐个检查，此时每个样品的平均受检次数为：

$$N_1=\frac{1}{k}q^k+\frac{k+1}{k}(1-q^k)=1-\left(q^k-\frac{1}{k}\right)=1-\frac{1}{k}(kq^k-1)$$

例如，当 $p=0.1$ 时，取 $k=4$，则 $q^k-\frac{1}{k}=0.4$，即用一重混合策略平均能减少 40% 的工作量。

文[1]首先提出了二重混合检查策略，即：将 $2k$ 个样品合为一个大组检查，呈阳性时，再将大组对半分为 k 个一组检查，如呈阳性，才逐个检查。此时每个样品的平均受检次数为：

$$\begin{aligned}N_2&=\frac{1}{2k}q^{2k}+\frac{2k+3}{2k}(1-q^k)^2+\frac{k+3}{2k}\cdot 2(1-q^k)q^k\\&=1-\frac{1}{2k}(2q^{2k}+2kq^k-3)\\&=1-\left(q^k-\frac{1}{k}\right)-\frac{1}{k}(q^{2k}-0.5)\end{aligned}$$

例如，当 $p=0.01$，取 $k=2$，则 $\frac{1}{k}(q^{2k}-0.5)=0.23$，即此时用二重混合策略比用一重混合策略平均能减少 23% 的工作量。

温启愚在此基础上又进一步提出 n 重混合检查策略的概念[1]，即：将 $2^{n-1}k$ 个样品合为一组检查，如呈阳性，则对半分为 $2^{n-2}k$ 个样品一组检查，如再呈阳性，则再次对半分组，直至分为 k 个一组检查，呈阳性时才逐个检查。最后得到一个猜测，即 n 重混合策略的平均受检次数为：

$$N_n=1-\left(q^k-\frac{1}{k}\right)-\sum_{i=0}^{n-2}(q^{2^{i+1}k}-0.5)\qquad(n\geqslant 2)\tag{1}$$

1 公式(1)的证明

用数学归纳法证明。我们从节约量入手作出分析，二重策略比一重策略节约的量为：

$$N_1-N_2=\frac{1}{k}(q^{2k}-0.5)$$

故 $n=2$ 时公式(1)成立。

设 $n=m(m>1)$ 时结论成立。现考察 $n=m+1$ 时的情况。实际上，$m+1$ 重比 m 重节约的量是从 $m+1$ 重到 m 重的过程中产生的。换名话说，就是在这一过程中，由于 p 较小，可使样品获得比 N_m 更小的平均受检次数，即 $N_{m+1}<N_m$，从 $m+1$ 重到 m 重，具体分析如下：

1) $2^m k$ 个样品合组检查为阳性，需进入 m 重混合检查，则此时 N_{m+1} 比 N_m 多出 $\frac{1}{2^m k}(1-q^{2^m k})$。

2) $2^m k$ 个样品组检查为阴性，无需进入 m 重混合检查，则此时 N_{m+1} 比 N_m 少 $\left(\frac{1}{2^{m-1}k}-\frac{1}{2^m k}\right)q^{2^m k}=\frac{1}{2^m k}q^{2^m k}$.

由 1)，2) 知：

$$
\begin{aligned}
N_{m+1} &= N_m + \frac{1}{2^m k}(1-q^{2^m k}) - \frac{1}{2^m k}q^{2^m k} \\
&= N_m - \frac{1}{2^{m-1}k}(q^{2^m k}-0.5)
\end{aligned}
$$

故 $n=m+1$ 时，结论也成立。

因此，公式(1) 对 $n \geqslant 2$ 是成立的。证毕。

注：对不同的阳性率 p，可算出最能节约的合组重数(见[1])，以上的证明作为前提条件应有 $q^{2^n k}-0.5>0$，即使得 $n+1$ 重比 n 重仍能节约。

2　n 重混合的平均受检次数的另一计算式及其与上式的等价性

为了便于归纳，首先计算四重混合的平均受检次数。利用二项式展开，特别注意展开式的系数，可得：

$$
\begin{aligned}
N_4 = \frac{1}{8k}\{ & q^{8k} + 8(k+7)(1-q^k)q^{7k} + [4(2k+7)+8(2k+9)+16(2k+11)]\cdot q^{6k} \\
& \cdot (1-q^k)^2 + [8(3k+9)+16(3k+11)+32(3k+13)](1-q^k)^3 q^{5k} \\
& + [2(4k+9)+4(4k+11)+48(4k+13)+16(4k+15)](1-q^k)^4 q^{4k} \\
& + [24(5k+13)+32(5k+13)](1-q^k)^5 q^{3k} \\
& + [4(6k+13)+24(6k+15)](1-q^k)^6 q^{2k} \\
& + 8(7k+15)(1-q^k)^7 q^k + (8k+15)(1-q^k)^8 \}
\end{aligned}
$$

化简得：

$$
N_4 = 1 - \frac{1}{8k}(2q^{8k} + 4q^{4k} + 8q^{2k} + 8kq^k - 15).
$$

可见直接求 N_4 比较繁，重数再高些计算就更费时且易错，因而证明公式(1) 成立不仅在理论上有意义在实际运用中也是必要的。

注意到 N_1、N_2、N_3、N_4 的表达式，将其与重数联系起来：

$$N_1 = 1 - \frac{1}{2^{1-1}k}[2^{1-1}kq^{2^{1-1}k} - (2^1 - 1)]$$

$$N_2 = 1 - \frac{1}{2^{2-1}k}[2^1q^{2^{2-1}k} + 2^{2-1}kq^k - (2^2 - 1)]$$

$$N_3 = 1 - \frac{1}{2^{3-1}k}[2^1q^{2^{3-1}k} + k^{3-1}q^{2k} + 2^{3-1}kq^k - (2^3 - 1)]$$

$$N_4 = 1 - \frac{1}{2^{4-1}k}[2^1q^{2^{4-1}k} + 2^2q^{2^{4-2}k} + 2^{4-1}q^{2k} + 2^{4-1}kq^k - (2^4 - 1)]$$

可以猜测：当 $n \geqslant 2$ 时，有：

$$\begin{aligned} N_n &= 1 - \frac{1}{2^{n-1}k}[2q^{2^{n-1}k} + 2^2q^{2^{n-2}k} + \cdots + 2^{n-1}q^{2k} + 2^{n-1}kq^k - (2^n - 1)] \\ &= 1 - \frac{1}{2^{n-1}k}\left[\sum_{i=1}^{n-1} 2^i q^{2^{n-i}k} + 2^{n-1}kq^k - (2^n - 1)\right] \end{aligned} \tag{2}$$

下面，我们证明公式(1) 与公式(2) 等价。

事实上，将公式(1) 适当变形，由(1)：

$$\begin{aligned} N_n &= 1 - \left(q^k - \frac{1}{k}\right) - \frac{1}{k}(q^{2k} - 0.5) - \cdots - \frac{1}{2^{n-2}k}(q^{2^{n-1}k} - 0.5) \\ &= 1 - \frac{1}{2^{n-1}k}[2q^{2^{n-1}k} + 2^2q^{2^{n-2}k} + \cdots + 2^{n-1}q^{2k} + 2^{n-1}kq^k - (2^{n-1} + 2^{n-2} + \cdots + 1)] \\ &= 1 - \frac{1}{2^{n-1}k}\left[\sum_{i=1}^{n-1} 2^i q^{2^{n-i}k} + 2^{n-1}kq^k - (2^n - 1)\right] \end{aligned}$$

最后一个式子就是公式(2).

致谢：四川大学数学系温启愚副教授曾给予指导，重庆大学数学系杨万年教授给予悉心指点，作者在此表示衷心感谢！

参 考 文 献

1 温启愚等. 一、二重混合检查法——大批中剔出缺陷个体的方法. 四川大学学报，1992，29(1)

2 Dorfman R. The detection of defective members of large populations. Ann. Math. Stat.，1943

3 Feller W. An Introduction to Probability Theory and its Applications. New York：1957