分层抽样方法在保险索赔处理正确性评价中的应用

胡桂华(教授)

(广西财经学院　南宁　530003)

【摘要】本文研究了保险索赔处理正确性的评价问题,指出如果总体层的结构未知,可以使用达拉纽斯和霍基斯方法确定层间边界值,采用比例分配和最优分配的方法将样本总量在层间进行分配。

【关键词】保险索赔　分层抽样方法　索赔金额

近几年来,随着城乡居民保险意识的不断增强,购买保险的人越来越多。但随之而来的一个问题是,由于投保人掌握的保险方面的知识有限或保险公司故意隐瞒重大实情,以致保险合同存在对投保人不公平的条款的可能,而且投保人往往并未意识到这种可能的存在。显然,如果投保人不幸发生了保险事故,得到的索赔金额可能远远低于他们实际应该得到的部分。保险索赔就是投保人或其他权利人在保险事故发生后,保险系统根据保险合同的相关规定给予其经济补偿或支付保险金。那么,如何评价一个国家或地区的保险系统是否公平对待投保人的保险索赔要求,是否完全履行了保险合同规定的给付义务呢?

统计抽样方法是一种被广泛使用的评估工具,其中的分层抽样方法通常用来解决会计总体的参数估计问题。会计总体的一大特征是其频数分布向右偏,即金额较大的项目数少。在分层抽样方法下,保险索赔总体的所有索赔事项首先会被划分到各层中,然后从每一层中随机抽取样本,对抽取的样本(即索赔事项)进行审计,并根据审计结果来估计两个总体参数,即正确处理索赔事项的比例和正确处理索赔事项的金额。如果总体能够被有效地划分为若干同质的层,进行分层抽样便可以获得比较高的参数估计精度。然而,在总体层的结构未知的情况下(通常情况下都是这样的),必须解决两个极其重要的问题,即确定层的数目和层间边界值。样本总量分配应该充分考虑参数性质上的差异与层的内部索赔金额的变异程度,以使估计参数的抽样误差最小。

一、保险索赔总体及其估计参数

保险索赔总体由向保险系统投保的所有投保人的索赔事项构成,索赔金额可能是零美元,也可能是数百美元、数千美元甚至数百万美元。估计参数有两个:一个是保险系统正确处理的索赔事项占全部索赔事项的比例;另一个是保险系统应当支付给投保人的索赔金额(又称审计金额或正确索赔处理金额)。我们可以通过分层抽样对这两个参数进行估计。

二、分层抽样方案

设计总体的分层抽样方案时会面临许多选择。如果总体层的结构未知,审计人员首先必须选择形成层的分层变量,然后确定层的数目和层间边界值。样本总量确定后,要将其在各层间进行分配。比例分配、最优分配和奈曼分配是常用的样本分配方法,最优分配和奈曼分配能使总体参数估计量的方差达到最小。科克伦(1977)指出,随着层的数目和分层抽样经费的增加,参数估计的精度会提高。

马哈拉洛比斯(1952)和汉森(1953)提出了一种确定层间边界值的方法,就是将每一层的抽样权数与分层变量的平均值相乘。达拉纽斯和霍基斯(1959)提出,将分层变量频数分布平方根的累计数划分成相等的区间而得到近似的最佳边界值。艾克曼(1959)通过使每一层的抽样权数与层的范围的积相等而得到层间边界值。荷斯(1966)研究了有关保险索赔总体的四种样本分配法和四种边界值构造法。罗贝克等(1977)对会计总体抽样进行了经验研究,他们采用四个会计总体从而产生了几个具有不同误差率的研究总体,并且得出了参数估计者对总体可能实施的行为(在不同抽样方案下)。他们的研究结果表明,使用分层抽样方法对会计总体参数进行估计会得到比较满意的结果。

三、案例分析

本文引用1997年美国明尼苏达州保险索赔的例子进行分析,具体资料见表1:

表1　某季度索赔金额总体概括性统计数据和频数分布

索赔金额区间(美元)	索赔项目数
0	2 259 067
0.01～1 000	6 732 691
1 000.01～10 000	140 829
10 000.01～100 000	8 103
100 000以上	111

我们从表1中不难看出,频数分布向右偏,也就是金额较大的索赔事项少,而且每笔索赔金额差异很大,是典型的会计总体。

我们的目标是收集相关证据来评价保险系统对投保人索

赔事项的处理是否公正(以一个季度为例)。保险系统对投保人索赔的处理结果见表2。

表2　　索赔处理结果

项　目	金额(美元)
索赔总额	806 400 496
平均数	117.17
标准离差	1 209.73
偏度	115.04
最大值	672 796.59
最小值	0

在本文研究中,将误差定义为索赔金额和审计金额之差。如果审计金额大于索赔金额,表明保险系统少支付投保人的索赔款,这称为"支付不足金额",即为负误差;如果审计金额小于索赔金额,表明保险系统多支付投保人的索赔款,这称为"超额支付金额",即为正误差。从以往的情况来看,总体发生这两类误差的概率在2%~4%之间。

那么,索赔金额是否与误差发生的概率有显著关系呢?到目前为止,还没有收集到这方面的证据。本文研究提出两个假设:第一,总体各个部分的误差率在统计上是相同的,与索赔金额没有直接关系;第二,超额支付金额不会超过索赔金额,而支付不足金额则可能超过索赔金额。

1. *计算样本总量n。*样本总量来自三个类型层:第一类型层,索赔金额为零的层;第二类型层,索赔金额非零的层;第三类型层,索赔金额特别大但项目很少的层。

(1)计算样本量n_{ao}(n_a)。当索赔金额被分成两类(正确索赔金额和错误索赔金额)时,样本量使用公式(1)和(2)计算。

$$n_{ao}=(Z_{\alpha/2}/d)^2P(1-P) \tag{1}$$

其中:$Z_{\alpha/2}$为显著性水平α下正态分布的临界值,d为抽样极限误差(可允许误差),P为总体中正确索赔金额所占的比例。

$$n_a=n_{ao}/[1+(n_{ao}-1)/N] \tag{2}$$

在总体容量N不是特别大时,用公式(2)计算,否则用公式(1)计算。

(2)计算样本量n_b。当正确索赔金额用区间表示时,其样本量使用公式(3)进行计算。

$$n_b=(NZ_{\alpha/2}SD/d)^2 \tag{3}$$

其中,d为审计人员在审计过程中确定的抽样极限误差,SD为总体索赔金额的标准离差。

(3)计算样本量n_c。首先确定较大索赔金额的分界点,然后将所有超过这个分界点的索赔金额列入样本,即100%抽样。在通常情况下,审计人员依据工作量大小或分配到金额较大索赔层的资源量确定分界点。本文研究以100 000美元作为分界点。

(4)确定样本总量n。计算完三个样本量之后,首先将n_a按比例P_0(总体中索赔金额为0的索赔事项所占比例)分配到第一类型层,然后取$(1-P_0)n_a$与n_b中的较大值作为第二类型层的样本量,最后将n_c作为第三类型层的样本量。这样,样本总量按公式(4)计算。

$$n=P_0n_a+Max[n_b,(1-P_0)n_a]+n_c \tag{4}$$

为了提高总体的参数估计精度,通常需要更大的样本量,但这受资源的可获取程度的限制。那么一个切实可行的办法就是通过审计人员的判断来确定样本总量。尽管如此,样本总量至少应该为"n_a+n_b",以达到总体的参数估计精度。

2. *确定层的数目与层间边界值。*分层抽样的一个重要内容是对分层变量进行选择。本文研究以索赔金额作为分层变量,这是因为:一是数据容易得到;二是能够直接根据数据来计算总体参数估计量。

分层抽样理论表明,非正态总体(会计总体)层的数目超过20层时,再增加层的数目,估计精度不会再提高;正态总体层的数目以6为限。本文研究将保险索赔总体分为10层。

本文研究采用达拉纽斯和霍基斯方法确定层间边界值,具体分六步进行。第一步,确定索赔金额的区间数目L。L越大,层间边界值越小。第二步,确定索赔金额的每一个区间i(i=1,2,3,…,L),其宽度记为ω_i,每个区间的宽度不一定相等,具体根据频数分布确定。第三步,计算每一个区间i的索赔数目N_i。第四步,计算每一个区间i的频数函数$f(y)=\omega_i\times N_i\sqrt{f(y)}$以及每一个区间i的累计频数函数。第五步,计算$\sqrt{f(y)}$的累计数。第六步,将$\sqrt{f(y)}$的累计数除以计划的总层数H(本文研究中为10层),某一层的边界值就是层间边界值。

3. *样本分配。*通常采用比例分配法或最优分配法对样本进行分配。如果各层抽样比相等,就使用比例分配法。不难理解,比例分配法适用于对总体的正确索赔金额比例的估计,因为估计的比例在总体的各个部分被假定是近似相等的。

如果索赔金额在各层之间差异较大,就使用最优分配法,也就是说,估计总体的正确索赔金额时使用最优分配法较为恰当。最优分配法根据各层索赔金额差异的大小适当分配样本量,由于索赔金额较大的层的差异也大,所以应该分配比较大的样本量。对于索赔金额较大的层应当100%抽取样本。

本文研究将保险索赔总体分为10层,其中:第1层样本量为P_0n_a;第10层样本量为n_c;第2~9层的样本量使用公式(5)或公式(6)确定。如果$(1-P_0)n_a=n_b$,就使用公式(5),否则用公式(6)。

$$n_h=(N_hS_h/\sum_{h=1}^{H}N_hS_h)\times n_b \tag{5}$$

$$n_h=(N_hS_h/\sum_{h=1}^{H}N_hS_h)\times(1-P_0)n_a \tag{6}$$

样本分配结果见表3。

4. *总体的正确索赔金额比例及金额总和的估计。*

(1)总体的正确索赔金额比例的估计。估计量$\hat{P}$按公式(8)或公式(9)计算。具体分两步进行:第一步,使用公式(7)计算所有层的正确索赔金额的估计量$\hat{V}$;第二步,使用公式(8)或公式(9)计算$\hat{P}$。公式(9)是对公式(8)的修正。

$$\hat{V}=\sum_{h=0}^{H+1}\sum_{i=1}^{n_h}v_{hi} \tag{7}$$

其中:当h=0时,$n_0=n_a$;当h=H+1时,H=9,$n_{H+1}=n_c$。

我国审计理论研究各阶段的特征分析

李文贵

（浙江财经学院 杭州 310018）

【摘要】本文将改革开放以来我国的审计理论研究过程划分为四个阶段，并在此基础上对各个阶段的特征进行分析，希望能为我国未来的审计理论研究指明方向。

【关键词】审计理论研究 研究阶段 特征

一、我国审计理论研究的阶段划分

我们将改革开放以来我国的审计理论研究过程划分为四个阶段：①1980～1985年为借鉴和学习他国的审计理论阶段；②1985～1993年为传统的审计理论研究阶段；③1993～1999年为审计理论的系统化研究阶段；④1999年至今为构建审计理论研究的新模式阶段。这可以用下图表示：

1980年 1985年 1993年 1999年 至今

借鉴和学习他国的审计理论阶段 | 传统的审计理论研究阶段 | 审计理论的系统化研究阶段 | 构建审计理论研究的新模式阶段

我国审计理论研究的阶段划分图

二、各阶段的特征分析

由于理论发展受不同时期的环境特征影响，因此我们对改革开放以来我国审计理论研究阶段的特征进行分析时将着重考虑如下因素：①社会经济背景；②经济管理体制；③研究方法；④研究内容的主要指向。

1. 借鉴和学习他国的审计理论阶段（1980～1985年）的特征。这一阶段正值审计制度在我国的恢复时期，人们对现代审计各方面的认识才刚刚起步。这一阶段突出的特点是对其他国家实施的审计制度及审计经验、审计理论进行借鉴和学习。

1980年2月，《会计研究》对美国注册会计师协会将于当

表3 分层及样本分配结果

层次	层间边界值（美元）	各层分配的样本量
1	0	31
2	0～40	32
3	40～110	47
4	110～250	34
5	250～650	39
6	650～1 570	39
7	1 570～3 960	39
8	3 960～10 430	39
9	10 430～100 000	89
10	100 000以上	111

$$\hat{P}=(\hat{V}/n) \qquad (8)$$

$$\hat{P}=\hat{V}+(Z_{\alpha/2}^{2}/2)/(n+Z_{\alpha/2}^{2}) \qquad (9)$$

在置信概率$(1-\alpha)\times 100\%$下，$\hat{P}$的置信区间为：

$$\hat{P}\pm Z_{\alpha/2}\sqrt{\hat{P}(1-\hat{P})/n}$$

（2）总体的正确索赔金额总和的估计。我们可以使用差异方法进行估计。首先使用公式（10）计算X_h：

$$X_h=\sum_{i=1}^{N_h}X_{hi} \qquad (10)$$

然后使用公式（11）和（12）计算$\hat{Y}$和$S^2(\hat{Y})$：

$$\hat{Y}=\sum_{h=0}^{H+1}[X_h+N_h(\bar{y}_h-\bar{x}_h)] \qquad (11)$$

$$S^2(\hat{Y})=\sum_{h=0}^{H+1}\{N_h^2(1-n_h/N_h)[1-n_h(n_h-1)][\sum_{i=1}^{n_h}(x_{hi}-\bar{x}_h)^2+\sum_{i=1}^{n_h}(y_{hi}-\bar{y}_h)^2-2\sum_{i=1}^{n_h}(x_{hi}-\bar{x}_h y_{hi}+\bar{x}_h\bar{y}_h)]\} \qquad (12)$$

最后，在置信概率$(1-\alpha)\times 100\%$下，$\hat{Y}$的区间可根据公式（13）确定。

$$\hat{Y}\pm Z_{\alpha/2}\sqrt{S^2(\hat{Y})} \qquad (13)$$

5. 计算结果。使用公式（8）计算得出保险系统正确索赔金额的比例$\hat{P}=97\%$；利用公式（12）计算得出保险系统正确索赔金额$\hat{Y}=798\ 107\ 160$（美元）；正误差为11 525 104美元，负误差为3 231 768美元。

主要参考文献

1. Agresti A.，Coull B.A.. Approximate is Better than Exact for Interval Estimation of Binomial Proportions. The American Statistician，1998；2

2. Arens A.A.，Beck J. K.. Applications of Statistical Sampling to Auditing. Prentice Hall Inc.，1981

3. Dalenius T.，Hodges J.L.. Minimum Variance Stratific－ation. Journal of the American Statistical Association，1959；5

4. Ekman G.. An Approximation Useful in Univariate Stratification. The Analysis of Mathematical Statistics，1959；30