

审计总体的分层技术

胡桂华(博士)

(广西财经学院 南宁 530003)

【摘要】 本文分析了累计频数根技术和几何技术在审计总体分层中的具体应用过程,并将两者进行了比较,以期能促进我国审计总体分层技术的提高。

【关键词】 审计总体 累计频数根技术 几何技术

现代审计的一个显著特点是运用了概率抽样,其中分层抽样法应用得较为广泛。分层抽样能否提高审计效率、保证审计质量、降低审计风险、节约审计成本,在很大程度上取决于审计总体分层的结果。目前我国有一些审计人员往往凭经验和感觉确定审计总体分层的层数和层与层之间的边界值,这样做不够科学和准确。因此,本文拟介绍两种在国外应用十分普遍的分层技术,即累计频数根技术和几何技术,以促进我国审计总体分层技术的提高。

一、分层概述

设审计总体由N笔交易组成: X_1, X_2, \dots, X_N 为每笔交易的账面额; Y_1, Y_2, \dots, Y_N 为审计人员审定的每笔交易金额。我们的目标是估计审计总额 T_Y ,并与账面总额 T_X 进行比较。为实现目标,我们从审计总体中分层抽取用来估计 T_Y 容量为n的样本。

我们把审计总体分为L层,用 N_1, N_2, \dots, N_L 表示各层交易的总数目,则审计总体交易总数目 $N=N_1+N_2+\dots+N_L$ 。如果用 n_1, n_2, \dots, n_L 表示从各层抽取的样本单位数,那么样本总规模 $n=n_1+n_2+\dots+n_L$ 。

在分层抽样下, T_Y 的估计量 $\hat{T}_{Y, st}$ 为:

$$\hat{T}_{Y, st} = N \sum_{h=1}^L (N_h/N) (\bar{y}_h) \quad (1)$$

在这里, \bar{y}_h 是第h层的样本审计平均数, $\hat{T}_{Y, st}$ 的方差估计量为:

$$V(\hat{T}_{Y, st}) = N^2 \sum_{h=1}^L (N_h/N)^2 (1 - n_h/N_h) (S_h^2/n_h) \quad (2)$$

其中, S_h 是第h层的标准差,即:

$$S_h = \sqrt{\sum_{i=1}^{n_h} (Y_{hi} - \bar{Y}_h)^2 / (n_h - 1)} \quad (3)$$

二、分层技术

1. 累计频数根技术。该技术由Dalentius和Hodges于1957年共同提出。其具体应用步骤如下:①递升排列变量X,并把X分为J类;②确定每一类的频数 f_i ($i=1, 2, \dots, J$),计算每一类的频数平方根 $\sqrt{f_i}$;③汇总频数平方根并除以总层数L得到Q;④在变量X表中找到与Q, 2Q, ..., LQ近似相等的累计频数根;⑤把与累计频数根相对应的X取值区间的上限作为各层的上限值。

2. 几何技术。该技术由Gunning和Horgan于2004年共同提出。

其具体应用步骤如下:①递升排列研究变量X;②把交易额最小的项作为层间边界值 K_0, \dots, K_L 的第一项,交易额最大的项作为层间边界值的最后一项;③计算共同比率r, $r = \sqrt[L]{X_{\max}/X_{\min}}$;④把对应于共同比率的X值作为层间的边界值,即 $K_0=a, K_1=ar, \dots, K_L=ar^L$ 。

值得注意的是,这两种技术都要求审计总体服从右偏分布,而经济总体大多是服从这种分布的,这也是这两种技术被广泛采用的原因之一。

分层后,就可按式(1)和式(2)估计审计总额,并与账面总额进行比较,视差异大小发表不同的审计意见。

3. 分层技术的应用。我们以爱尔兰某公司债务人总体为例说明这两种技术如何应用。该总体共有3 369笔债务,每笔债务的数额在40~28 000之间。将这些债务额(X)分为30类,并由小到大排列。

(1)累计频数根技术。从表1可以看出,变量X服从右偏分布,即变量值小的频数大,变量值大的频数小。如果用统计直方图表示,则越往右频数值越小。

我们把审计总体分为3层。不难计算, $Q=133.67 \div 3 = 44.56$,则Q、2Q、3Q分别为44.56、89.11、133.68。表1中与它们近似相等的累计频数分别为52.49、88、133.67,对应的最大X值分别为972、3 768、28 000;相对应的3个层次分别为40~972, 972~3 768, 3 768~28 000。

不难看出,运用累计频数根技术得到的计算结果有比较大的不确定性,因为分层结果取决于初始类别个数及类别边界值的选择。

(2)几何技术。易知, $r = \sqrt[3]{28\,000 \div 40} = 8.88$, $K_0=a=40$, $K_1=ar=40 \times 8.88=355$, $K_2=ar^2=3\,152$, $K_3=ar^3=28\,000$;相应的三个层次分别为40~355, 355~3 152, 3 152~28 000。

可见,几何技术比累计频数根技术的计算过程简单,具有客观性。为全面了解、掌握和运用这两种技术,下面对它们的统计特性进行比较。

4. 两种技术的统计特性比较。

(1)各层变异系数与层数的关系。将审计总体分为3、4、5层。从表2、表3、表4可以看出,在层数一定的情况下,几何技术

表1

X 的值	f_i	$\sqrt{f_i}$	累计 $\sqrt{f_i}$
40~972	2 755	52.49	52.49
972~1 904	258	16.06	68.55
1 904~2 836	137	11.70	80.25
2 836~3 768	60	7.75	88.00
3 768~4 700	42	6.48	94.48
4 700~5 632	33	5.74	100.22
5 632~6 564	18	4.24	104.46
6 564~7 496	14	3.74	108.20
7 496~8 428	11	3.32	111.52
8 428~9 360	9	3	114.52
9 360~10 292	6	2.45	116.97
10 292~11 224	2	1.41	118.38
11 224~12 156	6	2.45	120.83
12 156~13 088	2	1.41	122.24
13 088~14 020	4	2	124.24
14 020~14 952	0	0	124.24
14 952~15 884	4	2	126.24
15 884~16 816	0	0	126.24
16 816~17 748	0	0	126.24
17 748~18 680	1	1	127.24
18 680~19 612	2	1.41	128.65
19 612~20 544	0	0	128.65
20 544~21 476	1	1	128.65
21 476~22 408	0	0	129.65
22 408~23 340	1	1	130.65
23 340~24 272	0	0	130.65
24 272~25 204	1	1	131.65
25 204~26 136	0	0	131.65
26 136~27 068	1	1	132.65
27 068~28 000	1	1	133.65

下审计总体各层之间变异系数的差异比累计频数根技术下审计总体各层之间变异系数的差异小许多。这意味着使用几何技术时,审计总体层之间的边界值是最佳的,分层结果是最理想的;在分层技术一定的情况下,层数越多,各层之间变异系数的差异越小。

(2)层数与方差的关系。本文通过从各层抽取容量为100的样本来对比这两个技术的方差,即计算相对效率(Eff):

$$\text{Eff} = V_{\text{几何技术}}(\hat{T}_{Y, \text{st}}) / V_{\text{累计频数根技术}}(\hat{T}_{Y, \text{st}}) \quad (4)$$

式(4)中: $\hat{T}_{Y, \text{st}} = \text{式}(1) / N$, $V_{\text{几何技术}}(\hat{T}_{Y, \text{st}})$ 或 $V_{\text{累计频数根技术}}(\hat{T}_{Y, \text{st}}) = \text{式}(2) / N^2$ 。

从表5可以看出:①在层数一定的情况下,相对效率接近于1,也就是说,使用这两种分层技术推断的审计总体参数的精度基本相同;②在分层技术一定的情况下,层

表2 审计总体分为3层情况下两种技术变异系数比较

技术	1	2	3
几何:层区间	40~354	354~3 152	3 152~28 000
变异系数	0.71	0.68	0.64
累计频数根:层区间	40~972	972~3 768	3 768~28 000
变异系数	0.70	0.42	0.76

表3 审计总体分为4层情况下两种技术变异系数比较

技术	1	2	3	4
几何:层区间	40~205	205~1 057	1 057~5 443	5 443~28 000
变异系数	0.45	0.44	0.48	0.50
累计频数根:层区间	40~972	972~1 904	1 904~5 632	5 632~28 000
变异系数	0.70	0.19	0.27	0.69

表4 审计总体分为5层情况下两种技术变异系数比较

技术	1	2	3	4	5
几何:层区间	40~147	147~549	549~2 037	2 037~7 552	7 552~28 000
变异系数	0.37	0.38	0.40	0.37	0.41
累计频数根:层区间	40~972	972~1 904	1 904~2 836	2 836~7 496	7 496~28 000
变异系数	0.52	0.30	0.20	0.25	0.57

表5 n=100时两种技术的相对效率

层次	技术	方差	效率
3	几何技术	2 659.74	1.02
	累计频数根技术	2 592.41	
4	几何技术	1 355.36	0.81
	累计频数根技术	1 664.86	
5	几何技术	917.17	1.06
	累计频数根技术	857.63	

数增加,方差下降。那么,是否在层数增加到一定程度后,方差还会有明显的减少呢?答案当然是否定的。Cochran(1960)经研究证明,在层数达到6层后,再增加层数,方差减少的幅度是十分有限的。这是不难理解的,因为在样本量一定的情况下,增加层数,就意味着每一层的样本量减少,在其他条件相同的情况下,这必然会增大抽样误差。

三、结论

几何技术和累计频数根技术是用于审计总体分层的两种主要技术,在实际审计工作中可以选择其一。从前面的比较分析中可以看到,几何技术下的分层结果比较客观,不受研究变量X分类的影响,由最大的和最小的变量值决定,而且计算简单;累计频数根技术的分层结果受研究变量X分类的影响,即不同的分类,分层结果就会不相同。因此,从总体来看,几何技术无疑优于累计频数根技术。

主要参考文献

1. Cochran, W.G. Comparison of Methods for Determining Stratum Boundaries. Bulletin of the International Statistical Institute, 1961-32-2
2. Hedlin, D.A. Procedure for Stratification by an Extended Ekman Rule. Journal of Official Statistics, 2000-16-15