抽样框总体容量变动下的复合抽样设计

张维群[1], 杨澜泳[1,2], 杨善学[1*]
(1. 西安财经大学统计学院, 陕西 西安 710100; 2. 西安培华学院, 陕西 西安 710100)

摘要: 考虑到“三新”企业在短时间内新生和消亡的企业数量较大, 采用固定的抽样框对“三新” 企业进行连续抽样存在代表性较差问题, 为解决此问题, 提出了基于总体容量变动下的复合抽样设计方法. 采用复合抽样设计方法有效解决了总体容量变动下抽样样本的代表性问题, 给出了总体总值估计量并验证了其无偏性. 结果表明: 使用复合抽样设计方法对总体的出生率消亡率估计, 能够实现对总体容量变动下抽样框进行更新; 复合抽样设计方法比较适合总体总量、单元或单位特征发生变化的调查总体; 在总体总量发生变动时, 复合随机抽样方法优于传统抽样方法.

关键词: 复合抽样设计; 总体容量变动; 参数估计; 传统抽样方法
中图分类号: C811 **文献标识码**: A **文章编号**: 1000−5781(2021)04−0566−11
doi: 10.13383/j.cnki.jse.2021.04.010

Composite sampling design with changing population capacities of the sampling frame

Zhang Weiqun[1], Yang Lanyong[1,2], Yang Shanxue[1*]
(1. School of Statistics Xi'an University of Finance and Economics, Xi'an 710100, China;
2. Xi'an Peihua University, Xi'an 710100, China)

Abstract: Due to the large number of newborns and extinctions in a short period of time in the “three-new” enterprises, a fixed sampling frame method would result in the problem of poor representativeness. In order to solve the problem, this article proposes a composite sampling design method with changing population capacities. The method solves the representativeness problem of sampling when the population capacity changes. Further, the population value estimation is given and its unbias property is proved. Results show that: using the composite sampling design method to estimate the population birth rate and death rate can update the sampling frame when the population capacity changes. The composite sampling design method is more suitable for investigations with changing population capacities and changing unit or unit characteristics. The compound random sampling method is superior to the traditional sampling method with changing population capacities.

Key words: composite sampling design; changing population capacities; parameter estimation; traditional sampling method

收稿日期: 2020−07−07; 修订日期:2021−03−22.
基金项目: 国家社会科学基金资助项目(20ATJ002).
*通信作者.

1 引　言

2018年党中央、国务院出台了关于加快发展“三新”的有关要求, 国家统计局印发了《新产业新业态新商业模式统计分类(2018)》, “三新”企业是指具有出现新产业、新业态和新商业模式特征的企业, 在我国经济高质量发展中扮演了重要角色, 越来越受到政府的重视, 因此“ 三新”企业数据的统计已经成为政府统计研究的重要工作. 目前, 中国“ 三新”企业具有跨界多元化、规模小、数量多、变化快等特点, 而变化快的特点给长期统计监测带来困难[1]. 例如, 通过淘数据网站获取的淘宝女装行业店铺数据, 在2019年7月至2019年8月, 店铺数分别为 80 628户和 83 168户, 净增加 2 540户, 环比增长3.15%; 2019年9月的店铺数为 85 141, 净增加 1 973 户, 环比增长2.37%, 从这三个月数据可以明显看出当前女装行业店铺数量变化快且增长幅度大. 若对九月份女装行业现状进行分析, 采用七月份抽样框进行抽样, 将会导致4 513户店铺无法入样, 得到的结论会存在较大偏差. 运用动态思想研究我国经济形势, 并选择最优模型对其分析, 能够得出最精确的结论[2,3]. 在实践中, 使用固定抽样框对“ 三新”企业进行抽样时, 下期可能抽取到部分企业已“消亡”, 而部分“ 新生”企业则无法被抽取到, 从而出现样本采集困难和样本代表性等问题, 难以及时准确地反映“三新”企业实际情况. 因此, 对总体容量变动下抽样框抽样方法的研究, 能够有效地解决总体容量变化的抽样框问题, 对当前“三新”企业的统计实践具有理论指导意义.

本文将已有文献对抽样框的研究分为以下两个方面: 第一是关于多重抽样框的研究. 为解决传统抽样框无法完全覆盖总体信息的问题, Hartley[4] 提出了多重抽样框的方法, 主要是通过权重系数的确定从而进一步得出总体的估计量, 即H估计量. 随后, Lund[5]和Fullller等[6]对H估计量进行改进, 得到L估计量与FB估计量, 进一步完善了多重抽样框理论. 然而以上三个估计量最显著的问题在于其权重系数的确定过度依赖研究变量间的协方差. Skinner等[7]利用极大似然思想修正了之前的估计量, 提出了PML估计量, 有效地避免了研究变量不同而导致估计量不同的问题, 使其最优权重系数仅与不同部分的总体容量有关. 以上学者将双重抽样框看做三个部分进行总体总值估计量的构造. 期间, Bankier[8]、Kalton等[9]、Skinner[10]将双重抽样框视为单一抽样框进行统计量的构造, 构造了SF 估计量, 将权重系数按照不同部分区域进行计算, 可以有效避免了重叠区域样本重复的困难. 贺建风[11]利用简单估计方法将现有的研究成果进行统一, 提出了HT估计量, 并应用于“三农” 问题的调查研究中. 贺建风[12]和张勇[13]利用多重抽样框方法分别对我国服务业和农产量进行抽样设计. Metcalf 等[14]对心血管疾病危险因素的调查中使用多重抽样框调查方法, 并对全体毛利人进行了合理的估计. 以上学者的整体思路基于传统抽样框无法完整覆盖总体时, 构建由多个单一抽样框组合成一个能够完整覆盖总体的多重抽样框, 通过更新其中某些抽样框, 以达到对多重抽样框更新的目的. 第二是关于当总体发生动态变化时完善总体辅助信息的研究. 张维群等[15]从由周期性普查和经常性调查组成的中国统计数据入手, 讨论了综合动态数据库的建立, 利用周期性普查数据对经常性调查数据修正的时序模型以及周期性普查指标的数据插补方法, 说明建立综合动态数据库是一个随着时间增加而不断完善的过程. 基于总体容量动态变化特征, 朱钰等[16]在大数据背景下对动态抽样框的建立进行了研究, 利用“重复等分”确定时间节点函数及流速函数, 进而通过积分定理确定抽样总体, 构建动态抽样框的具体步骤.

以上文献未涉及以下三个方面: 在对抽样框进行更新的过程中未剔除消亡样本信息、抽样框总体容量发生动态变化的具体情形以及总体容量变动下抽样框估计量的构造. 因此本文在已有文献的基础上需要解决以下三个问题: 1) 如何设计一个总体容量变动情形下抽样框模型, 能够对抽样框总体信息进行补充同时剔除已经消亡的样本信息. 2) 如何根据抽样框的动态变化特点对总体容量变动下抽样框变动过程进行细分. 3) 如何构造一个符合总体容量变动下抽样框复合抽样设计步骤的估计量.

基于“三新”企业总体容量变动下抽样框抽样设计, 本文从“三新”企业规模等具有变动快的特点入手, 提出了总体容量变动下抽样框的抽样方法, 与以往的研究相比内容存在以下三点不同: 首先, 传统抽样框仅能

够对总体信息进行补充, 并未将已经消亡的样本信息进行删除, 总体容量变动下抽样框的提出在覆盖总体信息的同时删除消亡样本信息、补充出生样本信息, 减小了使用传统抽样框进行抽样设计时产生的偏差. 其次, 赋予总体容量变动下抽样框准确的定义, 能够在实际应用中对符合该定义的调查总体进行合理地抽样设计. 最后, 确定总体容量变动下抽样框的估计量, 使得总体容量变动下抽样框抽样方法的应用得到完善. 研究结果表明: 1) 使用复合随机抽样对总体容量变动下抽样框进行研究, 能够通过出生率和消亡率将抽样框进行更新. 2) 当总体容量、单元或单位特征发生变化, 对调查总体进行抽样设计时可以使用该抽样框方法. 3) 在总体容量发生变动时, 使用抽样框复合随机抽样优于传统抽样方法.

2 总体容量变动下的抽样框

为了研究抽样框容量变动下的抽样方法, 需要引入总体容量变动下抽样框的相关理论, 其中包括总体容量变动下抽样框的定义, 抽样框总体容量变动下的模型结构及总体容量变动下抽样框与传统抽样框的区别.

在赋予总体容量变动下抽样框定义之前首先整理本文所使用的符号及注释, 见表1.

表 1 符号及注释

Table 1 Symbols and comments

符号	释义	符号	释义
P	总体	$S_j(j=t,t+1)$	第j期抽样框
$N_{\rm ret}$	保留样本总体容量	$N_j(j=t,t+1)$	第j期总体容量
$N_{\rm dea}$	消亡样本总体容量	$Y_j(j=t,t+1)$	第j期总体总值
$N_{\rm bir}$	出生样本总体容量	$y_{j,i}(j=t,t+1)$	第j期样本单元
$Y_{\rm ret}$	保留样本总体总值	$n_j(j=t,t+1)$	第j期简单随机抽样抽取的样本数
$Y_{\rm dea}$	消亡样本总体总值	$n_{j,{\rm ret}}(j=t,t+1)$	第j期简单随机抽样保留样本数
$Y_{\rm bir}$	出生样本总体总值	$n_{t,{\rm dea}}$	第t期简单随机抽样消亡样本数
$\bar{Y}_{\rm ret}$	保留样本总体均值	$n_{t+1,{\rm bir}}$	第$t+1$期简单随机抽样出生样本数
$\bar{Y}_{\rm bir}$	出生样本总体均值	$\bar{y}_{j,{\rm ret}}(j=t,t+1)$	第j期简单随机抽样保留样本均值
q_t	总体消亡率	$y_{j,{\rm ret}}(j=t,t+1)$	第j期简单随机抽样保留样本总值
p_t	总体出生率	$y_{t+1,{\rm bir}}$	第$t+1$期简单随机抽样出生样本总值
r_t	总体容量变化率	$\bar{y}_{t+1,{\rm bir}}$	第$t+1$期简单随机抽样保留样本均值

定义 1 在相邻两个时期内总体内的总体容量、单元或者单位特征等单位结构发生变化, 则称该总体为动态总体.

定义 2 为了按一定概率抽取调查单位而编制的、由全部符合规定特征的有限总体单位所组成的集合称作抽样框[17].

定义 3 存在总体P在第t期构成抽样框S_t, 在第$t+1$期总体P的抽样框为S_{t+1}, S_t包含N_t个总体单元, 其中样本为$\{y_{t,i}\}, i=1,2,...,N_t$; S_{t+1}包含N_{t+1} 个总体单元, 其中样本为$\{y_{t+1,i}\}, i=1,2,...,N_{t+1}$. 若对于任意时期有$S_t \neq S_{t+1}$ 时, 则称总体P的抽样框S_t为动态总体抽样框, 也称为总体容量变动下抽样框.

总体容量变动下的抽样框是前后两期总体单元及总体容量发生了变动, 图1为总体容量变动下抽样框的模型结构, "○" 表示在t期抽到的样本, "△" 表示在$t+1$期抽到的样本, 抽样框S_t, 空白区抽到的"○"即为无法在抽样框S_{t+1}中匹配到的消亡样本, 抽样框S_{t+1} 空白区抽到的"△" 即为无法在抽样框S_t中匹配到的出生样本, 而阴影区域的样本则为保留样本. 显然, 第$t+1$ 期总体中包含了t 期在$t+1$期中保留单元、第$t+1$ 中新

增加的单元两个部分. 如果需要依据上一期辅助信息对样本容量调整时, 则样本单位就包括了t期在$t+1$期中保留单元、第$t+1$中新增加的单元和因样本容量调整补充的单位.

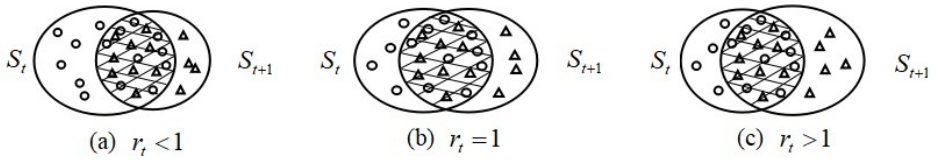

图 1　总体容量变动下抽样框的模型结构图

Fig. 1　The model structure diagram of the sampling frame under the population capacity change

图1中三种模型图均为出生率、消亡率同时存在的总体容量变动下抽样框, 其都包含了重叠区域, 故能够使用复合随机抽样方法使得通过前后两期抽样得到的样本相关联. 由图(a) 知, 当$r_t<1$ 时, 抽样框S_t随着时间变化总体容量减少; 由图(b)知, 当$r_t=1$时, 抽样框S_t 随着时间变化总体容量不变, 但抽样框的部分样本单元发生了变化; 由图(c) 知, 当$r_t>1$时, 抽样框S_t随着时间变化总体容量增多. 本文所研究的总体容量变动下的动态抽样框可分为三种情况, 通过总体容量变动率[1]表示, 即$r_t<1$、$r_t=1$、$r_t>1$, 分别表示随时间变化总体容量变小、存在出生样本和消亡样本但总体容量不变以及总体容量变大三种情形. 归纳整理得到在总体容量变动下的动态抽样框包含以下三种情形:

情形1　当$N_t>N_{t+1}$时, 认为$S_t\neq S_{t+1}$, 则S_t为一总体容量变动下抽样框.

情形2　当$\exists\left(y_{t,i}\notin S_{t+1}\right)\wedge\left(y_{t,i}\in S_t\right)$ 且$\exists\left(y_{t+1,i}\in S_{t+1}\right)\wedge\left(y_{t+1,i}\notin S_t\right)$, $N_t=N_{t+1}$时, 认为$S_t\neq S_{t+1}$, 则S_t为一总体容量变动下抽样框[2].

情形3　当$N_t<N_{t+1}$时, 认为$S_t\neq S_{t+1}$, 则S_t为一总体容量变动下抽样框.

由于传统抽样框被认为在较长时间是稳定不变的, 通过对总体容量变动下抽样框定义的描述以及其所包含不同的情形, 可以得到总体容量变动下抽样框不同于传统的固定抽样框, 第一, 总体容量变动下抽样框由前后两期抽样框设计得到, 并且前后两期抽样框包含的总体测量指标具有一致性; 第二, 总体容量变动下抽样框的总体特征数据来源是已知的基期总体信息以及两次抽样得到的样本测度指标数据; 第三, 总体容量变动下抽样框所包含的样本信息会随时间的变化而发生改变.

3　抽样框总体容量变动下复合抽样设计

在抽样框无法及时更新的问题上, 诸多调查采用多重抽样框进行抽样设计, 利用普查数据建立名录框, 再根据调查对象所在的区域构建地域框, 将两个或两个以上的抽样框组合得到一个能够覆盖总体信息的多重抽样框. 万舒晨[18]在对小微企业的研究中, 建立了多个名录框与区域框组合使得在小微企业调查中抽样总体与目标总体一致. 然而金勇进[19] 在对规模以下工业调查的若干问题研究中, 提到了由于市场竞争加快, 企业新生和消亡现象发生频繁, 使得目录企业抽样框变动较大, 通过实地调研发现, 对规模以下工业的调查使用多重抽样框(针对目录企业)存在较大困难. 说明了在某些调查研究中多重抽样框已经不再适用, 本文基于总体容量变动下抽样框提出了新的解决思路, 包含以下两个步骤:

1) 针对总体容量变动下抽样框的特点, 采用了两次随机抽样方法对总体容量变动下抽样框进行抽样. 使用两次随机抽样方法能够获取出生样本、消亡样本及保留样本, 能够使得在两个时期抽样框内所抽取的样本间存在关联性. 复合随机抽样方法包括两点: 首先在第t期抽样框进行第一次简单随机抽样, 其次在

[1] 总体容量变动率: 第$t+1$期总体容量与第t期总体容量的比值, 通过其数值可以反映随时间变化总体容量的变动情况.

[2] 情形2存在某些样本不包含于抽样框S_{t+1}但包含于抽样框S_t且存在某些样本不包含于抽样框S_t但包含于抽样框S_{t+1}, 两者样本数相同使得总体容量无变动, 认为抽样框S_t 与抽样框S_{t+1}不同, 则S_t为一总体容量变动下抽样框.

第$t+1$ 期抽样框进行第二次简单随机抽样, 两次随机抽样均为独立进行. 将两次简单随机抽样得到的样本在第$t+1$ 期进行调查, 便可得到出生样本、消亡样本及保留样本. 以此类推, 对于以后各期的抽样设计中均可以通过上期辅助信息进行抽样设计.

2) 对所抽取的样本采用多水平调查以减少调查所需的成本. 多水平调查指在调查单位接受调查时, 除了搜集当期的信息, 也搜集过去若干期的信息[20]. 在第$t+1$ 期对第一次简单随机抽样得到的样本进行二水平调查, 能够在$t+1$第期被调查到的样本为保留样本, 未被调查到的为消亡样本, 将保留样本第t期以及第$t+1$ 期的观测值登记记录. 同样, 在第$t+1$期对第二次简单随机抽样得到的样本进行二水平调查, 能够获取到第t期观测值的样本为保留样本, 未能获取到的为出生样本, 将保留样本第t 期、第$t+1$ 期的观测值以及出生样本第$t+1$期的观测值登记记录.使用该方法能够计算得到出生率和消亡率进而有效获取到抽样框在前后两期的变化情况.

传统的一次抽样方法与复合随机抽样方法的区别: 首先一次抽样方法在通过多水平调查后只能获取到出生率或消亡率其中一个数据, 复合随机抽样方法可以估算出生率与消亡率, 进而得到总体容量变动率使得第$t+1$期总体容量估计值更为准确; 其次一次抽样方法在进行总体总值估算时未对样本类型进行区分, 而复合随机抽样方法将第$t+1$ 期总体分为保留样本及出生样本, 分别通过两部分的样本信息加权计算得到总体总值估计量, 因此复合随机抽样方法更适用于总体容量变动下抽样框的抽样设计.

在对于“三新”企业的实际应用中, 需要注意以下两点: 1) 两个时间点的选择. 抽样设计前依据以往数据得到企业数量的历史变动情况, 选择合适的时间点, 以此作为基期总体. 2) 分析当下经济变化形势. 通过历史数据分析得到时间间隔后, 仍需分析两个时期外部经济环境是否存在差异, 例如政策的出台会影响企业的发展[21], 甚至关乎企业的新生与消亡, 相反的经济环境下需考虑重新调整时间间隔以达到对“三新”企业采用总体容量变动下抽样框复合随机抽样方法更为准确. 通过该抽样设计步骤计算消亡率、出生率进而得到两个时期内企业数量的变化情况, 并且得到“三新”企业营业收入总额估计值.

4 估计量的确定

本节基于总体容量变动下抽样框运用复合随机抽样方法, 对出生率、消亡率及报告期的总体容量进行估计, 从而实现对总体总值特征进行复合抽样估计, 讨论了估计量的无偏性, 并在其方差最小时求解得到最优权重系数.

4.1 总体容量的估计量

使用复合随机抽样方法可以得到消亡样本数与出生样本数, 通过与两次抽样得到的样本数的比值计算出消亡率与出生率的估计值, 即

$$\hat{q}_t=\frac{n_{t,\mathrm{dea}}}{n_t},\ \ \hat{p}_t=\frac{n_{t+1,\mathrm{bir}}}{n_{t+1}}.$$

由消亡率和出生率可以得到第$t+1$期总体容量的估计

$$\begin{aligned}\hat{N}_{t+1}&=(1-\hat{q}_t)\,N_t+\hat{p}_t\hat{N}_{t+1}\\&=\left(\frac{1-\hat{q}_t}{1-\hat{p}_t}\right)N_t=\hat{r}_tN_t.\end{aligned}\tag{1}$$

4.2 总体总值的估计量

Hartley[4]在对双重抽样框的研究中提出了H估计量, H估计量采用分离抽样框估计方法, 按照每个总体单位归属抽样框的特征进行域虚拟分离, 利用各域内的样本信息对域层面的子总体信息进行推断, 然后将所有子总体信息组合起来对目标总体变量进行估计. 将双重抽样框看作三个部分, 即$\hat{Y}_a$为抽样框A中与抽

样框B 不重叠的总体总值, $\hat{Y}_b$为抽样框B中与抽样框A不重叠的总体总值, 以及两个抽样框的重叠部分, 其中$\hat{Y}_{ab}^{A}$为两个抽样框重叠部分来自抽样框A的总体总值, $\hat{Y}_{ab}^{B}$为两个抽样框重叠部分来自抽样框B的总体总值, 即

$$\hat{Y}_H(\theta)=\hat{Y}_a+\hat{Y}_b+\theta\hat{Y}_{ab}^{A}+(1-\theta)\hat{Y}_{ab}^{B}, \tag{2}$$

其中 θ为H估计量的权重系数, $\hat{Y}_H(\theta)$为抽样框B的总体总值估计量.

随后Fullller和Burmeister[6]在H估计量的基础上提出了FB估计量, N_A为抽样框A的总体容量, N_B为抽样框B的总体容量, $\hat{N}_{ab,srs}$ 为两个抽样框重叠部分的总体容量估计量, $\hat{\mu}_{a,srs}^{A}$为抽样框A中非重叠部分的均值的估计量, $\hat{\mu}_{b,srs}^{B}$为抽样框B中非重叠部分的均值的估计量, $\hat{\mu}_{ab,srs}$为两个抽样框重叠部分均值的估计量, n_A为在抽样框A中抽取的样本量, n_B为在抽样框B中抽取的样本量, n_{ab}^{A} 为在两个抽样框重叠部分来自抽样框A抽取的样本量, n_{ab}^{B}为在两个抽样框重叠部分来自抽样框B抽取的样本量, 即

$$\hat{Y}_{srs}=\left(N_A-\hat{N}_{ab,srs}\right)\hat{\mu}_{a,srs}^{A}+\left(N_B-\hat{N}_{ab,srs}\right)\hat{\mu}_{b,srs}^{B}+\hat{N}_{ab,srs}\hat{\mu}_{ab,srs}, \tag{3}$$

其中 $\hat{Y}_{srs}$为抽样框B的总体总值估计量, 并给出了$\hat{N}_{ab,srs}$的计算公式

$$\left(n_A+n_B\right)x^2-\left(n_AN_B+n_BN_A+n_{ab}^{A}N_A+n_{ab}^{B}N_B\right)x+\left(n_{ab}^{A}+n_{ab}^{B}\right)N_AN_B=0, \tag{4}$$

其中 x的最小值就为$\hat{N}_{ab,srs}$的值.

根据H估计量以及FB估计量的构造思路, 在得到第$t+1$期总体容量的估计值后, 采用分离抽样框估计方法将抽样框S_{t+1}分为出生样本部分及保留样本部分并分别测量两个部分的变量信息, 利用两部分样本的复合估计总体目标变量值, 则第$t+1$期总体总值的估计量为

$$\begin{aligned}\hat{Y}_{t+1}&=w_{t+1}\hat{N}_{t+1}\hat{\bar{Y}}_{\mathrm{ret}}+\left(1-w_{t+1}\right)\hat{N}_{t+1}\hat{\bar{Y}}_{\mathrm{bir}}\\&=\left[w_{t+1}\frac{1}{n_{t,\mathrm{ret}}+n_{t+1,\mathrm{ret}}}\left(y_{t,\mathrm{ret}}+y_{t+1,\mathrm{ret}}\right)+\left(1-w_{t+1}\right)\bar{y}_{t+1,\mathrm{bir}}\right]\hat{r}_tN_t,\end{aligned} \tag{5}$$

式(5)对于总体总值的估计量构造是基于第$t+1$期的样本框进行的.

由于二次抽样都采用独立的放回简单随机抽样, 对式(5)进行无偏性检验, 容易得到

$$\mathrm{E}\left(\hat{Y}_{t+1}\right)=Y_{t+1}. \tag{6}$$

由式(6)可知$\hat{Y}_{t+1}$是Y_{t+1}的无偏估计量, 因此所构造的总体总值估计量是有效的. 式(6)的详细推导过程见附录A.

4.3　最优权重系数的确定

由于保留样本与出生样本间是相互独立的, 第一次简单随机抽样的保留样本与第二次简单随机抽样的保留样本间是相互独立的, 依据简单随机放回抽样中样本均值方差的计算公式[23]$\mathrm{Var}\left(\bar{\mathrm{y}}\right)=\frac{\sigma^2}{\mathrm{n}}$, 因此$\hat{Y}_{t+1}$的方差为

$$\begin{aligned}\mathrm{Var}\left(\hat{Y}_{t+1}\right)=&N_{t+1}^2\left[\left(\frac{1}{n_{t,\mathrm{ret}}+n_{t+1,\mathrm{ret}}}\right)^2w_{t+1}^2\left(n_{t,\mathrm{ret}}\sigma_{\mathrm{ret}}^2+n_{t+1,\mathrm{ret}}\sigma_{\mathrm{ret}}^2\right)+\left(1-w_{t+1}\right)^2\frac{\sigma_{\mathrm{bir}}^2}{n_{t+1,\mathrm{bir}}}\right]+\\&N_t^2\left[\left(\left(\frac{1}{n_{t,\mathrm{ret}}+n_{t+1,\mathrm{ret}}}\right)^2w_{t+1}^2\left(n_{t,\mathrm{ret}}\sigma_{\mathrm{ret}}^2+n_{t+1,\mathrm{ret}}\sigma_{\mathrm{ret}}^2\right)+\left(1-w_{t+1}\right)^2\frac{\sigma_{\mathrm{bir}}^2}{n_{t+1,\mathrm{bir}}}\right)+\bar{Y}_{t+1}^2\right]\mathrm{Var}\left(\hat{r}_t\right),\end{aligned} \tag{7}$$

其中 σ_{ret}^2为保留样本总体方差, σ_{bir}^2为出生样本总体方差, $\bar{Y}_{t+1}$为第$t+1$期总体均值, $\mathrm{Var}\left(\hat{r}_t\right)$为总体容量变动率估计量的方差. 式(7)的详细推导过程见附录B.

由于权重系数w_{t+1}是一个未知参数, 复合估计中需要确定w_{t+1}最优的值, 因此, 可以在保证总体总值估计的有效性的前提下确定复合随机抽样的权重. 即计算$\mathrm{Var}\left(\hat{r}_t\right)$ 最小值时的权重w_{t+1}. 令$\dfrac{\partial\mathrm{Var}\left(\hat{\mathrm{Y}}_{\mathrm{t+1}}\right)}{\partial w_{t+1}}=0$,

可得最优权重系数w_{t+1}的值为

$$w_{t+1}=\frac{\frac{\sum_{i=1}^{n_{t+1,\mathrm{bir}}}\left(y_{i,t+1,\mathrm{bir}}-\bar{y}_{t+1,\mathrm{bir}}\right)}{n_{t+1,\mathrm{bir}}\left(n_{t+1,\mathrm{bir}}-1\right)}}{\left(\frac{1}{n_{t,\mathrm{ret}}+n_{t+1,\mathrm{ret}}}\right)^2\left(\frac{n_{t,\mathrm{ret}}\sum_{i=1}^{n_{t,\mathrm{ret}}}\left(y_{i,t,\mathrm{ret}}-\bar{y}_{t,\mathrm{ret}}\right)}{n_{t,\mathrm{ret}}-1}+\frac{n_{t+1,\mathrm{ret}}\sum_{i=1}^{n_{t+1,\mathrm{ret}}}\left(y_{i,t+1,\mathrm{ret}}-\bar{y}_{t+1,\mathrm{ret}}\right)}{n_{t+1,\mathrm{ret}}-1}\right)+\frac{\sum_{i=1}^{n_{t+1,\mathrm{bir}}}\left(y_{i,t+1,\mathrm{bir}}-\bar{y}_{t+1,\mathrm{bir}}\right)}{n_{t+1,\mathrm{bir}}\left(n_{t+1,\mathrm{bir}}-1\right)}},\tag{8}$$

式(8)的详细推导过程见附录C.

5 数值分析

为讨论总体容量变动下抽样框的复合随机抽样方法的优良性, 本小节生成模拟数据对该方法优良性进行分析, 由于未能获取到“三新”企业详细数据, 本次实验依据“三新”企业的变化特点进行数据模拟, 使其能够用于“三新”企业发展规律等研究.

随机模拟分为以下步骤, 首先设置总体容量变化率, 在实验中添加了$r_t<1$, $r_t=1$, $r_t>1$三种情况. 随后开始生成第t期总体, 给定消亡率值, 系统随机删除t 期总体的部分样本单元即为消亡样本单元, 删除后得到保留样本单元, 并对保留样本单元进行20%左右的随机增幅. 同时给定出生率值, 系统随机生成新的样本单元即为出生样本单元, 结合出生样本单元与增幅后的保留样本单元得到第$t+1$期总体. 最后对生成的两期总体进行抽样, 在t期总体进行第一次简单随机抽样过程, 将所抽取样本与第$t+1$ 期总体进行匹配, 得到保留样本单元及其在第期的观测值; 在第$t+1$期总体进行第二次简单随机抽样过程, 将所抽取样本与第t 期总体进行匹配, 得到出生样本单元, 依据两次抽样得到的样本信息进行分析. 由于第$t+1$ 期营业额是总体待估计目标变量, 在样本容量确定中, 往往根据第t期辅助变量的方差大小确定, 因此, 第$t+1$ 期样本容量仍然依据第t 期辅助变量的方差大小确定; 以此类推, 在第$t+2$期的抽样中可以依据第$t+1$期抽样样本对总体目标变量方差的估计值进行样本容量的调整.

随机模拟过程中既定参数值如表2.

表 2 生成模拟数据的参数设置

Table 2 Parameter settings for generating simulation data

$r_t<1$		$r_t=1$		$r_t>1$	
$r_t=0.82$	$N_t=1\,000$	$r_t=1$	$N_t=1\,000$	$r_t=1.143$	$N_t=1\,000$
$N_{t+1}=820$	$Y_t=51\,713$	$N_{t+1}=1\,000$	$Y_t=59\,826$	$N_{t+1}=1\,143$	$Y_t=51\,516$
$Y_{t+1}=48\,632$	$q_t=29\%$	$Y_{t+1}=69\,032$	$q_t=20\%$	$Y_{t+1}=66\,385$	$q_t=20\%$
$p_t=13.4\%$	$n_t=200$	$p_t=20\%$	$n_t=200$	$p_t=30\%$	$n_t=200$

本次随机模拟实验使用Python软件进行了随机模拟. 对总体容量变动下抽样框使用复合随机抽样估计方法以及将其看作静态抽样框使用传统的抽样方法进行实验分析, 包含多重抽样框估计总体总值、第一次简单随机抽样保留样本估计总体总值和第二次简单随机抽样样本估计总体总值三种传统的抽样方法, 其中多重抽样框的实验模拟将第t期的总体当作基期名录框, 将第$t+1$期总体当作现期地域框进行数值分析. 可以得到100轮实验不同的总体总值, 并计算各个方法的偏差, 实验结果如图2.

通过图2四种不同方法得到总体总值估计的结果比较可知, 图(a)为$r_t<1$时的实验结果, 总体容量变动下抽样框复合随机抽样方法计算得到的平均偏差为0.047, 第一次简单随机抽样保留样本计算得到的平均偏差为0.138, 第二次简单随机抽样样本计算得到的平均偏差为0.38, 多重抽样框方法计算得到的平均偏差为0.351; 图(b)为$r_t=1$ 时的实验结果, 总体容量变动下抽样框复合随机抽样方法计算得到的平均偏差为0.034, 第一次简单随机抽样保留样本计算得到的平均偏差为0.162, 第二次简单随机抽样样本计算得到的

平均偏差为0.203, 多重抽样框方法计算得到的平均偏差为0.221; 图(c)为$r_t > 1$时的实验结果, 总体容量变动下抽样框复合随机抽样方法计算得到的平均偏差为0.051, 第一次简单随机抽样保留样本计算得到的平均偏差为0.263, 第二次简单随机抽样样本计算得到的平均偏差为0.136, 多重抽样框方法计算得到的平均偏差为0.161. 实验结果表明在总体容量发生变化时抽样框的复合随机估计方法明显优于简单随机抽样和多重抽样框的抽样方法, 说明总体容量变动下抽样框复合随机抽样方法对于抽样框变动频繁的情况下具有较好的应用价值.

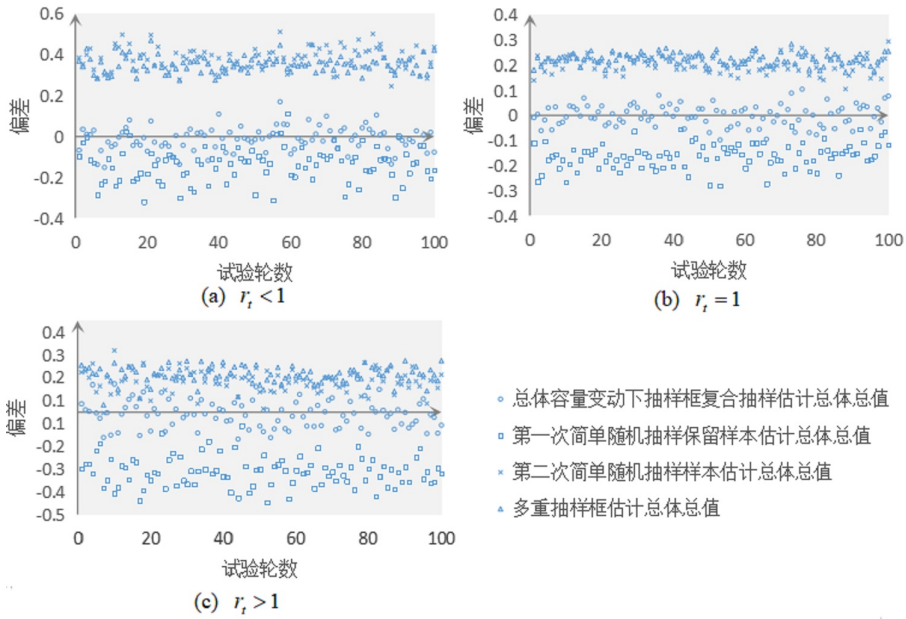

图 2　总体容量变动下抽样框复合抽样估计与其他抽样估计的优良性比较

Fig. 2　Comparison of the superiority of the composite sampling estimation of the sampling frame with other sampling estimations under the population capacity

6　结束语

本文从“三新”企业变化频繁的现实出发, 提出了基于总体容量变动下抽样框的复合随机抽样方法, 该方法的提出解决了传统抽样框在处理总体容量发生变动时的缺陷. 本文主要研究了动态总体及总体容量变动下抽样框的理论内涵, 给出了总体容量变动下抽样框的出生率、消亡率和第$t+1$期总体总值的估计量, 并讨论了该估计量的优良性及复合估计权重的确定方法. 实验结果显示, 抽样框复合随机抽样估计在总体容量变动时抽样估计精度更高. 本文仅在总体容量变动下抽样框抽样方法进行讨论, 对于抽样框更复杂变动情形的理论研究仍需进一步讨论.

参考文献:

[1] 赵顺招. 做好“三新”统计. 中国统计, 2016, 64(8): 55–57.
Zhao S Z. Do a good job of “Three-New” statistics. China Statistics, 2016, 64(8): 55–57. (in Chinese)

[2] 韩志刚, 王洪桥. 动态系统预报的多模型多算法综合模式. 系统工程学报, 1991, 7(2): 26–33.
Hang Z G, Wang H Q. Multi-models and multi-algorithms synthetic pattern of dynamic system prediction. Journal of Systems Engineering, 1991, 7(2): 26–33. (in Chinese)

[3] 王意冈, 王浣尘. 动态模式经济控制论模型的模糊形式研究. 系统工程学报, 1994, 10(1): 84–90.
Wang Y G, Wang H C. On fuzzy form of dynamic pattern economic cybernetics model. Journal of Systems Engineering, 1994, 10(1): 84–90. (in Chinese)

[4] Hartley H O. Multiple frame surveys. Proceedings of the Social Statistical Section, American Statistical Association, 1962: 203–206.

[5] Lund R E. Estimators in multiple frame surveys. Proceedings of the Social Statistics Sections, American Statistical Association, 1968: 282–288.

[6] Fulller W A, Burmeister L F. Estimators for samples selected from two overlapping frames. Proceedings of the Social Statistics Section, American Statistical Association, 1972: 245–249.

[7] Skinner C J, Rao J N K. Estimation in dual frame surveys with complex designs. Journal of the American Statistical Association, 1996, 91(433): 349–356.

[8] Bankier M D. Estimators based on several stratified samples with applications to multiple frame surveys. Journal of the American Statistical Association, 1986, 81(396): 1074–1079.

[9] Kalton G, Anderson D W. Sampling rare populations. Journal of the Royal Statistical Society, 1986, 39(149): 65–82.

[10] Skinner C J. On the efficiency of raking ratio estimation for multiple frame surveys. Journal of the American Statistical Association, 1991, 86(415): 779–784.

[11] 贺建风. 多重抽样框方法及其应用研究. 广州: 暨南大学, 2010.
He J F. The Study About Method and Application of Multiple Sampling Frame. Guangzhou: Jinan University, 2010. (in Chinese)

[12] 贺建风. 基于多重抽样框的服务业抽样调查体系改革研究. 暨南学报(哲学社会科学版), 2015, 37(3): 25–33.
He J F. Research on the reform of sampling survey system of service industry based on multiple sampling frames. Jinan Journal, Philosophy & Social Science Edition, 2015, 37(3): 25–33. (in Chinese)

[13] 张 勇. 中国农产量抽样调查若干问题研究. 北京: 中国统计出版社, 2007.
Zhang Y. Research on Several Issues of Sampling Survey of Agricultural Production in China. Beijing: China Statistics Press, 2007. (in Chinese)

[14] Metcalf P, Scott A. Using multiple frames in health surveys. Statistics in Medicine, 2009, 28(10): 1512–1523.

[15] 张维群, 王佐仁. 论综合动态数据库的建立与应用. 统计与信息论坛, 2008, 23(1): 72–75.
Zhang W Q, Wang Z R. Construction and application of comprehensive dynamic data base of periodic census & regular survey data. Statistics & Information Forum, 2008, 23(1): 72–75. (in Chinese)

[16] 朱 钰, 王 恬. 网络数据环境下动态抽样框的构建及其应用. 统计与决策, 2019, 35(2): 5–8.
Zhu Y, Wang T. Construction and application of dynamic sampling frame in network data environment. Statistics & Decision, 2019, 35(2): 5–8. (in Chinese)

[17] 《规模以下工业抽样框问题研究》课题组. 我国规模以下工业抽样框的问题与解决思路. 统计研究, 2003, 20(7): 41–45.
Research Group on “Sampling Frames for Sub-scale Industries”. Problems in design of sampling frame of under scale industry firms in China and solutions. Statistical Research, 2003, 20(7): 41–45. (in Chinese)

[18] 万舒晨, 金勇进. 大数据背景下小微企业多重抽样框应用研究. 数量经济技术经济研究, 2018, 35(9): 126–140.
Wan S C, Jin Y J. Multiple frame surveys of the small and micro enterprises in the context of big data. The Journal of Quantitative & Technical Economics, 2018, 35(9): 126–140. (in Chinese)

[19] 金勇进, 姜天英. 规模以下工业调查的若干问题研究. 统计研究, 2019, 36(3): 42–50.
Jin Y J, Jiang T Y. A study on some issues in industrial surveys under a designated size. Statistical Research, 2019, 36(3): 42–50. (in Chinese)

[20] 陈光慧, 刘建平. 我国基本单位多水平连续性抽样调查体系改革研究: 以规模以下工业调查为例. 统计研究, 2014, 31(7): 50–57.
Chen G H, Liu J P. The reform research about China’s basic unit multilevel successive sampling survey system: A case study on industrial survey under the designated size. Statistical Research, 2014, 31(7): 50–57. (in Chinese)

[21] 张宇翔, 谭德庆. 碳政策对厂商技术稳定性及租售策略的影响. 系统工程学报, 2019, 34(5): 632–643.
Zhang Y X, Tan D Q. Impact of carbon policy on technical stability and rental strategy of manufacturers. Journal of Systems Engineering, 2019, 34(5): 632–643. (in Chinese)

[22] 贺建风. 基于多重抽样框的校准估计方法研究. 统计研究, 2018, 35(4): 104–116.
He J F. The Research on the calibration estimation method based on multiple sampling frames. Statistical Research, 2018, 35(4): 104–116. (in Chinese)

[23] 谢邦昌. 抽样调查的理论及其应用方法. 北京: 中国统计出版社, 1998.
Xie B C. Theory and Application of Sampling Survey. Beijing: China Statistics Press, 1998. (in Chinese)

[24] 贺建风. 基于双重抽样框的抽样估计方法研究. 统计研究, 2011, 28(12): 89–96.
He J F. The sampling estimation methods study based on sampling frame. Statistical Research, 2011, 28(12): 89–96. (in Chinese)

作者简介:

张维群 (1969—), 男, 陕西旬邑人, 博士, 教授, 硕士生导师, 研究方向: 抽样理论, 数据挖掘, Email: xazhangwq@163.com;

杨澜泳 (1995—), 女, 陕西铜川人, 硕士生, 研究方向: 抽样理论, Email: yangly05@163.com;

杨善学 (1981—), 男, 山东济宁人, 博士生, 讲师, 研究方向: 供应链管理, Email: shxyang@xaufe.edu.cn.

附录

附录A. 公式(6)的详细推导过程

$$\begin{aligned}
\mathrm{E}\left(\hat{Y}_{t+1}\right) &= \left(w_{t+1}\frac{1}{n_{t,\mathrm{ret}}+n_{t+1,\mathrm{ret}}}\left(n_{t,\mathrm{ret}}\mathrm{E}\left(\bar{y}_{t,\mathrm{ret}}\right)+n_{t+1,\mathrm{ret}}\mathrm{E}\left(\bar{y}_{t+1,\mathrm{ret}}\right)\right)+\left(1-w_{t+1}\right)\mathrm{E}\left(\bar{y}_{t+1,\mathrm{bir}}\right)\right)\mathrm{E}\left(\hat{r}_t\right)N_t\\
&= \left(w_{t+1}\frac{1}{n_{t,\mathrm{ret}}+n_{t+1,\mathrm{ret}}}\left(n_{t,\mathrm{ret}}\bar{Y}_{\mathrm{ret}}+n_{t+1,\mathrm{ret}}\bar{Y}_{\mathrm{ret}}\right)+\left(1-w_{t+1}\right)\mathrm{E}\left(\bar{y}_{t+1,\mathrm{bir}}\right)\right)r_tN_t\\
&= \left(w_{t+1}\bar{Y}_{\mathrm{ret}}+\left(1-w_{t+1}\right)\bar{Y}_{\mathrm{bir}}\right)r_tN_t\\
&= \left(w_{t+1}\bar{Y}_{\mathrm{ret}}+\left(1-w_{t+1}\right)\bar{Y}_{\mathrm{bir}}\right)N_{t+1}\\
&= w_{t+1}Y_{t+1}+\left(1-w_{t+1}\right)Y_{t+1}\\
&= Y_{t+1}.
\end{aligned}$$

附录B. 公式(7)的详细推导过程

$$\begin{aligned}
\mathrm{Var}\left(\hat{Y}_{t+1}\right) &= \mathrm{Var}\left[\left(\frac{1}{n_{t,\mathrm{ret}}+n_{t+1,\mathrm{ret}}}w_{t+1}\left(y_{t,\mathrm{ret}}+y_{t+1,\mathrm{ret}}\right)+\left(1-w_{t+1}\right)\bar{y}_{t+1,\mathrm{bir}}\right)\hat{r}_tN_t\right]\\
&= N_t^2\mathrm{E}\left[\left(\frac{1}{n_{t,\mathrm{ret}}+n_{t+1,\mathrm{ret}}}w_{t+1}\left(y_{t,\mathrm{ret}}+y_{t+1,\mathrm{ret}}\right)+\left(1-w_{t+1}\right)\bar{y}_{t+1,\mathrm{bir}}\right)^2\right]\mathrm{E}\left(\hat{r}_t^2\right)-\\
&\quad N_t^2\left[\mathrm{E}\left(\frac{1}{n_{t,\mathrm{ret}}+n_{t+1,\mathrm{ret}}}w_{t+1}\left(y_{t,\mathrm{ret}}+y_{t+1,\mathrm{ret}}\right)+\left(1-w_{t+1}\right)\bar{y}_{t+1,\mathrm{bir}}\right)\right]^2\mathrm{E}\left(\hat{r}_t^2\right)\\
&= \left[\left(\left(\frac{1}{n_{t,\mathrm{ret}}+n_{t+1,\mathrm{ret}}}\right)^2w_{t+1}^2\left(n_{t,\mathrm{ret}}\sigma_{\mathrm{ret}}^2+n_{t+1,\mathrm{ret}}\sigma_{\mathrm{ret}}^2\right)+\left(1-w_{t+1}\right)^2\frac{\sigma_{\mathrm{bir}}^2}{n_{t+1,\mathrm{bir}}}\right)+\right.\\
&\quad \left.\bar{Y}_{t+1}^2\right]\left[\mathrm{Var}\left(\hat{r}_t\right)+r_t^2\right]N_t^2-Y_{t+1}^2\\
&= N_{t+1}^2\left[\left(\frac{1}{n_{t,\mathrm{ret}}+n_{t+1,\mathrm{ret}}}\right)^2w_{t+1}^2\left(n_{t,\mathrm{ret}}\sigma_{\mathrm{ret}}^2+n_{t+1,\mathrm{ret}}\sigma_{\mathrm{ret}}^2\right)+\left(1-w_{t+1}\right)^2\frac{\sigma_{\mathrm{bir}}^2}{n_{t+1,\mathrm{bir}}}\right]+\\
&\quad N_t^2\left[\left(\left(\frac{1}{n_{t,\mathrm{ret}}+n_{t+1,\mathrm{ret}}}\right)^2w_{t+1}^2\left(n_{t,\mathrm{ret}}\sigma_{\mathrm{ret}}^2+n_{t+1,\mathrm{ret}}\sigma_{\mathrm{ret}}^2\right)+\left(1-w_{t+1}\right)^2\frac{\sigma_{\mathrm{bir}}^2}{n_{t+1,\mathrm{bir}}}\right)\bar{Y}_{t+1}^2\right]\mathrm{Var}\left(\hat{r}_t\right).
\end{aligned}$$

附录C. 公式(8)的详细计算过程

$$\begin{aligned}
\frac{\partial\mathrm{Var}\left(\hat{Y}_{t+1}\right)}{\partial w_{t+1}} &= N_{t+1}^2\left[\left(\frac{1}{n_{t,\mathrm{ret}}+n_{t+1,\mathrm{ret}}}\right)^2 2w_{t+1}\left(n_{t,\mathrm{ret}}\sigma_{\mathrm{ret}}^2+n_{t+1,\mathrm{ret}}\sigma_{\mathrm{ret}}^2\right)-2\left(1-w_{t+1}\right)\frac{\sigma_{\mathrm{bir}}^2}{n_{t+1,\mathrm{bir}}}\right]+\\
&\quad N_t^2\left[\left(\frac{1}{n_{t,\mathrm{ret}}+n_{t+1,\mathrm{ret}}}\right)^2 2w_{t+1}\left(n_{t,\mathrm{ret}}\sigma_{\mathrm{ret}}^2+n_{t+1,\mathrm{ret}}\sigma_{\mathrm{ret}}^2\right)-2\left(1-w_{t+1}\right)\frac{\sigma_{\mathrm{bir}}^2}{n_{t+1,\mathrm{bir}}}\right]\mathrm{Var}\left(\hat{r}_t\right)=0,
\end{aligned}$$

上式化简可得

$$\begin{aligned}
&N_{t+1}^2\left(\frac{1}{n_{t,\mathrm{ret}}+n_{t+1,\mathrm{ret}}}\right)^2 2w_{t+1}\left(n_{t,\mathrm{ret}}\sigma_{\mathrm{ret}}^2+n_{t+1,\mathrm{ret}}\sigma_{\mathrm{ret}}^2\right)+2w_{t+1}N_{t+1}^2\frac{\sigma_{\mathrm{bir}}^2}{n_{t+1,\mathrm{bir}}}+\\
&N_t^2\mathrm{Var}\left(\hat{r}_t\right)\left(\frac{1}{n_{t,\mathrm{ret}}+n_{t+1,\mathrm{ret}}}\right)^2+2w_{t+1}\left(n_{t,\mathrm{ret}}\sigma_{\mathrm{ret}}^2+n_{t+1,\mathrm{ret}}\sigma_{\mathrm{ret}}^2\right)+2w_{t+1}N_t^2\mathrm{Var}\left(\hat{r}_t\right)\frac{\sigma_{\mathrm{bir}}^2}{n_{t+1,\mathrm{bir}}}\\
&=2N_{t+1}^2\frac{\sigma_{\mathrm{bir}}^2}{n_{t+1,\mathrm{bir}}}+2N_t^2\mathrm{Var}\left(\hat{r}_t\right)\frac{\sigma_{\mathrm{bir}}^2}{n_{t+1,\mathrm{bir}}},
\end{aligned}$$

上式可以解得w_{t+1}的值为

$$w_{t+1}=\frac{\frac{\sigma_{\mathrm{bir}}^2}{n_{t+1,\mathrm{bir}}}\left(N_{t+1}^2+N_t^2\mathrm{Var}\left(\hat{r}_t\right)\right)}{\left[\left(\frac{1}{n_{t,\mathrm{ret}}+n_{t+1,\mathrm{ret}}}\right)^2\left(n_{t,\mathrm{ret}}\sigma_{\mathrm{ret}}^2+n_{t+1,\mathrm{ret}}\sigma_{\mathrm{ret}}^2\right)\right]\left(N_{t+1}^2+N_{t+1}^2\mathrm{Var}\left(\hat{r}_t\right)+\frac{\sigma_{\mathrm{bir}}^2}{n_{t+1,\mathrm{bir}}}\left(N_{t+1}^2+N_t^2\mathrm{Var}\left(\hat{r}_t\right)\right)\right)}$$
$$=\frac{\frac{\sigma^2}{n_{t+1,\mathrm{bir}}}}{\left(\frac{1}{n_{t,\mathrm{ret}}+n_{t+1,\mathrm{ret}}}\right)^2\left(n_{t,\mathrm{ret}}\sigma_{\mathrm{ret}}^2+n_{t+1,\mathrm{ret}}\sigma_{\mathrm{ret}}^2\right)+\frac{\sigma_{\mathrm{bir}}^2}{n_{t+1,\mathrm{bir}}}},$$

若用样本进行权重估计, 则有

$$w_{t+1}=\frac{\frac{\sum_{i=1}^{n_{t+1,\mathrm{bir}}}\left(y_{i,t+1,\mathrm{bir}}-\bar{y}_{t+1,\mathrm{bir}}\right)}{n_{t+1,\mathrm{bir}}\left(n_{t+1,\mathrm{bir}}-1\right)}}{\left(\frac{1}{n_{t,\mathrm{ret}}+n_{t+1,\mathrm{ret}}}\right)^2\left(\frac{n_{t,\mathrm{ret}}\sum_{i=1}^{n_{t,\mathrm{ret}}}\left(y_{i,t,\mathrm{ret}}-\bar{y}_{t,\mathrm{ret}}\right)}{n_{t,\mathrm{ret}}-1}+\frac{n_{t+1,\mathrm{ret}}\sum_{i=1}^{n_{t+1,\mathrm{ret}}}\left(y_{i,t+1,\mathrm{ret}}-\bar{y}_{t+1,\mathrm{ret}}\right)}{n_{t+1,\mathrm{ret}}-1}\right)+\frac{\sum_{i=1}^{n_{t+1,\mathrm{bir}}}\left(y_{i,t+1,\mathrm{bir}}-\bar{y}_{t+1,\mathrm{bir}}\right)}{n_{t+1,\mathrm{bir}}\left(n_{t+1,\mathrm{bir}}-1\right)}}.$$

****

2022大数据时代交通与物流国际会议

暨第十届国际决策科学高峰论坛(TL&DS 2022)

2022年7月1日~3日

黑龙江·哈尔滨

2022大数据时代交通与物流国际会议(2022 International Conference on Intelligent Transportation and Logistics with Big Data) 暨第九届国际决策科学高峰论坛(The Tenth International Forum on Decision Sciences) 将于2022年7月1日~3日在黑龙江哈尔滨召开. 该会议由中国优选法统筹法与经济数学研究会船海经济管理分会主办, 哈尔滨工程大学承办, 采用线上线下结合方式(海外报告嘉宾线上, 国内参会者线下).

本次会议将围绕“交通、物流、决策”等研究领域进行广泛交流, 旨在为从事交通运输、供应链物流、决策科学领域研究的科研人员和从业人员提供一个学习交流的平台, 探讨交通运输、供应链物流、决策科学前沿理论、创新应用及其在大数据时代面临的机遇与挑战. 本次会议将邀请美国、加拿大、意大利、新加坡、中国香港等多个国家和地区的知名专家学者参会, 并贡献精彩的主题报告. 结合会议主题, 将组织一系列的SCI、SSCI期刊的Special issue.

经过6年的不断努力, 中国优选法统筹法与经济数学研究会船海经济管理分会现已发展成为全国交通物流、船舶工业、海洋经济、港航工程等领域相关科技工作者的群众性专业学术组织和学术交流平台. 由其主办的国际会议“大数据时代的智慧交通与物流国际会议(International Conference on Intelligent Transportation and Logistics with Big Data)”2018年入选中国科协首届重要学术会议指南, 其后连续三年持续入选中国科协重点学术会议指南. 目前, 该会议已经成长为国内“交通与物流”领域有影响力的国际会议.