

月度调查的另一种平衡的两层次样本轮换模式*

侯志强¹, 李勇²

(1. 北方工业大学 统计系, 北京 100041; 中国人民大学 统计学院, 北京 100872;
2. 重庆工商大学 数学与统计学院, 重庆 400067)

[摘要] 对于一级单元规模较大的两阶段抽样月度调查给出了一种两层次样本轮换模式, 该模式中一级单元的轮换模式为 8-4-8(16), 二级单元的轮换模式为 2-10-2(4)。

[关键词] 月度调查; 样本轮换; 两层次轮换; 轮换模式

[中图分类号] C8 [文献标识码] A [文章编号] 1672-0598(2007)06-0042-03

一、引言

连续调查又称为经常性调查, 是每隔一定时间(比如一个月或一个季度)进行一次的抽样调查。由于其需要对被调查对象连续进行访问, 容易使被调查对象产生抵触情绪, 从而对调查敷衍了事或拒绝回答, 所以被调查对象不宜接受过多次数的调查。那么, 被调查对象究竟应该接受多少次调查为宜呢? 尽管对这个问题没有统一的答案, 但可以肯定的是每个被调查对象只接受一次调查就退出样本是不太合适的, 因为这样既会增加调查的费用, 又会给两个时期变化量的估计带来困难。因此, 一个合理的选择是每个被调查对象接受一定次数的调查后由新的未接受过调查的被调查对象替代, 这即是样本轮换。对于采用单阶段抽样的连续调查, 只需要对一级抽样单元实施样本轮换就可以了, 但是对于实施多阶段抽样的连续调查, 就需要对每级抽样单元均实施样本轮换。现在的问题是各级抽样单元应该怎样轮换才能保证任何级别单元的任何一次轮换不改变各级单元在不同时间间隔时的样本拼配率呢? 这即是多阶段抽样连续调查所遇到的平衡多层次样本轮换问题。

对于两阶段抽样月度调查中的平衡两层次样本轮换问题, 日本的劳动力调查^[1]或中国台湾的人

力资源调查^[2]给出了一种解决方案, 即一级单元的轮换模式为 4-8-4(8), 二级单元的轮换模式为 2-10-2(4)。该平衡两层次样本轮换模式适用于一级单元规模较小的情况。如在日本的劳动力调查中, 一级单元约包含 50 个住户。当一级单元规模较大时, 世界上并无其他可以借鉴的实践案例。本文试图构造一种新的平衡两层次样本轮换模式。

二、平衡两层次样本轮换模式的构造

对于采用两阶段抽样设计的月度调查(以下本文均以采用两阶段抽样设计的月度调查为例), 在一级单元规模较大时, 可以采用如下一种两层次样本轮换模式, 即各级单元的轮换每月进行一次, 且一级单元的轮换模式为 8-4-8(16), 一级单元内二级单元的轮换模式为 2-10-2(4)。在该两层次样本轮换模式下, 任何级别单元的任何一次轮换均不改变各级单元在不同时间间隔时的样本拼配率, 因而该两层次样本轮换模式是一种平衡的两层次样本轮换模式, 其样本轮换图如图 1 所示。

三、样本轮换模式的解释

(一) 图形说明

图 1 中所展示的调查时间为第 t 年一月到第 $t+1$ 年九月, 所涉及子样本有三个, 分别为子样

* [收稿日期] 2007-09-23

[基金项目] 本文得到国家社会科学基金项目“劳动力调查方法的国际比较研究”(07CTJ006)的资助。

[作者简介] 侯志强(1970-), 男, 山西曲沃人, 北方工业大学统计系, 讲师, 中国人民大学统计学院博士生, 研究方向: 抽样技术。

李勇(1970-), 男, 重庆南川人, 重庆工商大学数学与统计学院, 讲师, 研究方向: 统计理论与应用。

本 m 、子样本 $m+1$ 和子样本 $m+2$ 。每次调查的样本来自于两个子样本或三个子样本,如第 t 年一月的样本来自于子样本 m 和子样本 $m+1$,第 $t+1$ 年二月的样本则来自于子样本 m 、子样本 $m+1$ 和子样本 $m+2$ 。

图 1 中每个大写英文字母表示一个一级轮换组,每个阿拉伯数字表示一个二级子样本,每个一级轮换组由若干个一级单元组成,每个二级子样本由若干个二级单元组成。每个子样本包含 16 个一级轮换组,分别用大写英文字母“A”到“P”表示,每个一级轮换组包含四个二级子样本,分别用阿拉伯数字“1”到“4”表示。

年份	月份	子样本 m	子样本 $m+1$	子样本 $m+2$
第 t 年	一月	ABCDEFGH MNOP	ABCD	
		44332211 4433	2211	
	二月	BCDEFGHI NOP	ABCDE	
		44332211 443	32211	
	三月	CDEFGHIJ OP	ABCDEF	
		44332211 44	332211	
	四月	DEFGHIJK P	ABCDEF	
		44332211 4	4332211	
	五月	EFGHIJKL	ABCDEF	
		44332211	44332211	
	六月	FGHIJKLM	BCDEFGH	
		44332211	44332211	
七月	GHIJKLMN	CDEFGHIJ		
	44332211	44332211		
八月	HIJKLMNO	DEFGHIJK		
	44332211	44332211		
九月	IJKLMN	EFGHIJKL		
	44332211	44332211		
十月	JKLMNOP A	FGHIJKLM		
	44332211	44332211		
十一月	KLMNOP AB	GHIJKLMN		
	44332211	44332211		
十二月	LMNOP ABC	HIJKLMNO		
	44332211	44332211		
第 $t+1$ 年	一月	MNOP ABCD	IJKLMN	
		4433 2211	44332211	
	二月	NOP ABCDE	JKLMN	A
		443 32211	44332211	1
	三月	OP ABCDEF	KLMN	AB
		44 332211	44332211	11
	四月	P ABCDEF	LMNOP	ABC
		4 4332211	44332211	211
	五月	ABCDEFGH MNOP	ABCD	
44332211 4433		2211		
六月	BCDEFGHI NOP	ABCDE		
	44332211 443	32211		
七月	CDEFGHIJ OP	ABCDEF		
	44332211 44	332211		
八月	DEFGHIJK P	ABCDEF		
	44332211 4	4332211		
九月	EFGHIJKL	ABCDEF		
	44332211	44332211		

图1 两层次样本轮换图

一级轮换组总是隶属于某个特定的子样本,二级子样本总是隶属于某个特定的一级轮换组。如第 t 年一月样本的第一列,一级轮换组“A”隶属于子样本 m ,二级子样本“4”隶属于一级轮换组“A”。

每次调查的样本用 2 行 16 列共 32 个字符表示,其中,第一行的 16 个字符表示参与本次调查的 16 个一级轮换组,第二行的 16 个字符表示参与本次调查的 16 个二级子样本。

(二) 轮换模式的解释

1、一级轮换组的轮换

每个一级轮换组从首次进入样本到永久退出样本历时一年零八个月,共接受 16 次调查。每个一级轮换组首次进入样本后连续接受 8 次调查,接着暂时退出样本 4 个月,然后再次进入样本连续接受 8 次调查,之后永久退出样本。以子样本 $m+1$

内的一级轮换组“D”为例,其首次进入样本的时间为第 t 年一月,从第 t 年一月到第 t 年八月连续接受 8 次调查,接着暂时退出样本四个月,然后于第 $t+1$ 年一月再次进入样本,并从第 $t+1$ 年一月到第 $t+1$ 年八月连续接受 8 次调查,之后永久退出样本。

每个月有 1 个一级轮换组首次进入样本接受第 1 次调查,有 1 个一级轮换组接受第 2 次调查,……,有 1 个一级轮换组接受第 8 次调查并暂时退出样本,有 1 个一级轮换组返回样本接受第 9 次调查,有 1 个一级轮换组接受第 10 次调查,……,有 1 个一级轮换组接受第 16 次调查并永久退出样本。如第 $t+1$ 年九月,子样本 $m+2$ 内的一级轮换组“H”首次进入样本接受第 1 次调查,“G”接受第 2 次调查,……,“A”接受第 8 次调查并暂时退出样本,子样本 $m+1$ 内的一级轮换组“L”返回样本接受第 9 次调查,“K”接受第 10 次调查,……,“E”接受第 16 次调查并永久退出样本。

2、二级子样本的轮换

每个二级子样本从首次进入样本到永久退出样本历时一年零两个月,共接受 4 次调查。每个二级子样本首次进入样本后连续接受两次调查,然后暂时退出样本十个月,最后再次进入样本连续接受两次调查并永久退出样本。以子样本 $m+1$ 内一级轮换组“D”内的二级子样本“1”为例,第 t 年一月它首次进入样本,从第 t 年一月到第 t 年二月连续接受两次调查,然后暂时退出样本十个月,第 $t+1$ 年一月返回样本,从第 $t+1$ 年一月到第 $t+1$ 年二月连续接受两次调查,之后永久退出样本。

每个月有 4 个二级子样本接受第 1 次调查,有 4 个二级子样本接受第 2 次调查,有 4 个二级子样本接受第 3 次调查,有 4 个二级子样本接受第 4 次调查。如第 $t+1$ 年九月,子样本 $m+2$ 内“H”内的“1”、“F”内的“2”、“D”内的“3”和“B”内的“4”接受第 1 次调查,“G”内的“1”、“E”内的“2”、“C”内的“3”和“A”内的“4”接受第 2 次调查,子样本 $m+1$ 内“L”内的“1”、“J”内的“2”、“H”内的“3”和“F”内的“4”接受第 3 次调查,“K”内的“1”、“I”内的“2”、“G”内的“3”和“E”内的“4”接受第 4 次调查,

3、二级子样本轮换与一级轮换组轮换之间的关系

每个二级子样本的轮换均在某个特定的一级轮换组内进行。以子样本 $m+1$ 内一级轮换组“D”内的二级子样本轮换为例。当第 t 年一月到第 t

年八月“D”连续接受八次调查的时候,其内的二级子样本“1”接受一月到二月的调查,“2”接受三月到四月的调查,“3”接受五月到六月的调查,“4”接受七月到八月的调查。第 t 年九月到第 t 年十二月,“D”暂时退出样本,其内的二级子样本也停止接受调查。当第 t+1 年一月到第 t+1 年八月“D”再次接受八次调查的时候,其内的二级子样本“1”再次接受一月到二月的调查,“2”再次接受三月到四月的调查,“3”再次接受五月到六月的调查,“4”再次接受七月到八月的调查。当“D”在第 t+1 年八月接受最后一次调查并永久退出样本的时候,其内的二级子样本也随之永久退出样本。

表 1: 两种模式下各级单元的样本拼配率

时间间隔 (月份)	拼 配 率			
	模式一		模式二	
	一级单元	二级单元	一级单元	二级单元
1	6/8	4/8	14/16	8/16
2	4/8	0	12/16	0
3	2/8	0	10/16	0
4	0	0	8/16	0
5	0	0	7/16	0
6	0	0	6/16	0
7	0	0	5/16	0
8	0	0	4/16	0
9	1/8	0	5/16	0
10	2/8	0	6/16	0
11	3/8	2/8	7/16	4/16
12	4/8	4/8	8/16	8/16
13	3/8	2/8	7/16	4/16
14	2/8	0	6/16	0
15	1/8	0	5/16	0
16	0	0	4/16	0
17	0	0	3/16	0
18	0	0	2/16	0
19	0	0	1/16	0
>=20	0	0	0	0

四、各级单元的样本拼配率

为了叙述方便,将日本劳动力调查中所使用的

平衡两层次样本轮换模式简称为模式一,将本文中构造的平衡两层样本轮换模式简称为模式二。模式一与模式二中各级单元在不同时间间隔时的样本拼配率如表 1 所示。

从表 1 中可以看出,不论模式一还是模式二,各级单元的样本拼配率总是随着间隔时间的扩大而先下降,后上升,然后再下降。对于任何一种模式,一级单元的样本拼配率总是大于等于二级单元的样本拼配率。模式二中一级单元的样本拼配率总是大于等于模式一中一级单元的样本拼配率。两种模式中二级单元在不同时间间隔时的样本拼配情况完全一样,其中相邻两个月份和相邻两年的相同月份时二级单元的样本拼配率均为 50%。

五、结束语

日本劳动力调查的平衡两层次样本轮换模式中一级单元的轮换速度比较快,这在一级单元规模较小时是必需的。当一级单元规模较大时,仍然以这么快的轮换速度轮换一级单元将造成调查费用的大量上升。本文提出的平衡两层次样本轮换模式在不影响二级单元样本拼配率的前提下明显放慢了一级单元的轮换速度,因而适用于一级单元规模较大时的两阶段抽样月度调查。

[参考文献]

- [1] Statistics Bureau of Japan. Sampling Method, Estimation Method, and Sampling Errors of Estimates [EB/OL]. <http://www.stat.go.jp/english/data/roudou/pdf/3.pdf>, 2007-03-01.
- [2] 黄奕嘉. 人力资源(运用)样本轮替方式说明 [EB/OL]. <http://tqlda.sinica.edu.tw/modules/tinyd0/index.php?id=4>, 2007-08-02.

(责任编辑:朱德东)

A balanced two-layer Yotation pattern for montly survey

HOU Zhi-qiang¹, LI Yong²

(1. Department of Statistics, North China University of Technology, Beijing 100041, China;

School of Statistics, Renmin University of China, Beijing 100872, China;

2. Mathematics and Statistics College, Chongqing Technology and Business University, Chongqing 400067, China)

Abstract: A two-layer rotation pattern is presented for the monthly surveys when the primary sampling units are relatively large, in which the rotation pattern for the primary sampling units is 8-4-8(16), and for the secondary sampling unit is 2-10-2(4).

Keywords: monthly survey; sample rotation; two-layer rotation; rotation pattern