

广东统计论文选

第二辑

广东省统计学会编

广东统计论文选

(第二辑)

广东省统计学会编

一九八三年九月

前 言

为了交流统计科研成果，促进统计科学的发展，广东省统计学会不定期编印《广东统计论文选》，内部发行。

《广东统计论文选》第二辑包括二十四篇论文是广东省一九八三年八月举行的第二次统计科学讨论会上评选获奖的论文。目录次序是按照获奖等级排列的：第1—3篇为一等奖；第4—5篇为二等奖；第6—8篇为三等奖；第9—24篇为四等奖。

目 录

关于有关标识排队等距抽样方法的再探讨·····	龚鉴尧 (1)
论我国对外贸易发展的几个指标·····	王正宪 (13)
以资金利税率为中心建立多层次的工业经济 效益指标体系·····	叶荣基 (21)
试谈建安工程价格指数编制方法·····	张应文 袁略文 (30)
改革商品流转统计报表与采用多种调查方法 初探·····	梁煜灵 梁泽林 (42)
广泛应用经济预测方法加强外贸经济效益预测 ·····	朱慧强 (50)
试论统计体制改革·····	杜之道 (67)
智力开发程度综合指标的确定·····	任 平 (74)
计量经济模型预测方法·····	暴奉贤 (83)
对现行《工业统计学》的修改浅见·····	刘向集 (106)
承包工厂经济效益统计的探讨·····	陈重坚 (114)
浅谈工业企业经济指标如何计算经济效益的 方法·····	张已立 (119)
开展统计分析和预测,发挥统计的服务和监 督作用,提高企业经济效益·····	林炳填 (128)

建立统计数字质量考核制度问题的浅见·····	陈友莲 (139)
对外加工装配工业总产值计算方法应该具体化 ·····	陈凤基 (144)
关于如何综合评价工业经济效益的探讨·····	梁 响 (150)
运用多种方法, 搞准中选队产量是推断总体的 关键·····	郑观藻 (155)
主成分分析在社会经济学中的应用·····	滕成业 (161)
试编企业实物型投入产出表的初步探讨·····	陈伟明 (176)
论地区国民收入流出流入的计算·····	伍荣坤 (184)
抽样法在城乡私人建房综合统计工作上的应用 ·····	肖 荣 (196)
农业统计分析几个问题的探讨·····	黄盛伟 (210)
关于建立建筑安装企业主要经济效益指标的初 步探讨·····	莫建坤 (227)
对产区木材总产值计算的一些看法·····	赖其梁 (238)

关于有关标识排队等距抽样方法 的再探讨

蔡 鉴 尧

我的《关于有关标识排队等距抽样方法的探讨》论文发表后，在统计学界引起了很大的兴趣和反响。许多同志认为《探讨》一文“提出了新的见解”，用有关标识排队等距抽样方法抽选出来的样本，“完全符合优样本原则”；这种方法的形成与发展，是我国统计工作者“长期应用富有成果的划类选典方法”和“吸收新发展的概率抽样法所结出的丰盛果实”。在一些新近出版的统计教材和著作，如陕西财经学院、暨南大学、中央财政金融学院、四川财经学院等合编的高等财院试用教材《统计原理与经济统计》，天津财经学院编写的《社会经济统计学原理》，宋元村、黄玉喜编著的《数理统计学》等，都肯定了有关标识排队等距抽样方法；中国人民大学、北京经济学院、天津财经学院和国家统计局等主编的《社会经济统计学原理电视讲座学习材料》，还对这种方法作了详细论述。目前，这种方法在我国农产量调查、农民和职工家计调查、人口调查等各个领域，正在被越来越多地推广应用。但是，也有一些同志不同意这种方法，认为它“不符合随机原则”，“缺乏充分科学根据”；有同志甚至写信给我说，这种方法是否科学、站得住，关系着抽

样调查在我国的推广应用和农产量抽样调查等工作的前途，要求我进一步加以澄清。因此，我想确有进一步探讨的必要，以就正于关心抽样调查的同志们。

一、关于有关标识排队等距抽样方法的随机性问题

有的同志认为，有关标识排队等距抽样，在实践上可能是一种好办法，但在理论上似乎还缺乏充分的科学根据，主要是因为它“不符合随机原则”，只能算做是一种“主观抽样法”。

我认为，这种对有关标识排队等距抽样方法的认识是不正确的。我在《抽样法浅说》和《探讨》论文中都曾经说过：“任何抽样方法的首要之点，就是必须消除一切产生偏误的根源。要做到这一点，最简单的可靠方法是：或者抽取样本的方法完全随机，或者在提高正确性而又不致给调查结果带来偏误的情况下限制随机。”有关标识排队等距抽样方法，在我看来，就是这样一种符合随机原则而在一定程度上限制随机的方法。事实上，在实际工作中，采用完全随机办法是很少的，因为完全随机样本的抽样误差比较大，需要抽取的样本单位数比较多，因而所需的人力和费用也比较大，不经济。而类型抽样、等距抽样，以及任何多阶段抽样方法，它们都使随机性在不同程度上受到了限制，但却比完全随机抽样方法更为有效：即或是在抽取同样大小的样本时具有较小的抽样误差，或是在具有同样大小的抽样误差的情况下需要的样本单位数要少；同时，也更便于组织实施。这在

理论上和实践上都是已经被证明了的。不需再多作论证。

有些同志对抽样调查，片面地理解了随机性，认为随机性越大越好，因此把完全随机抽样看做是最理想的方法，其实是不正确的。因为现代统计科学和抽样技术的发展，越来越多地并不是采用简单随机抽样法，而是提倡尽可能地搜集辅助信息和利用已知信息，采用那些能提高正确性而又不致给调查结果带来偏误，限制随机而又不损害随机原则，同时更加便于组织和经济有效的方法。

有的同志说，无关标识排队等距抽样是限制随机的办法，但在抽取样本时，可以任意一点为起点，虽然限制了随机，却无损于随机原则，这种方法是可以赞成的；而有关标识排队等距抽样，从组中点抽取样本单位，一旦抽样数目确定后，整个样本也就确定了，只能抽出一套样本，其它单位再没有被抽中的机会，这样大的限制，是否还符合随机原则？

我们说，任何等距抽样方法，包括无关标识排队和有关标识排队在内，只要抽样数目和起点一经确定，就只能有一套样本。问题在于，无关标识排队，是以任意点为起点，在抽样起点确定以前，它有可能抽中距离内的任何一套样本。

比如说，总体单位数为 $N=100$ ，抽样数目为 $n=10$ ，抽样距离 $=\frac{N}{n}=\frac{100}{10}=10$ ，则有10套可能样本；而有关标识排队，我们通常使用的办法，是从组中点开始，只能抽取一套样本。

从这个意义上说，有关标识排队等距抽样的确比无关标识排队等距抽样的随机性小。但这只能说它的随机性更加受到限制，而不是“不符合随机原则”。

应当说明的一点是：有关标识排队等距抽样，由于总体单

位是按高低顺序排列的，通常对抽样调查单位的具体做法，都是在计算出抽样距离后，以距离的中间一个单位——组中点为起点。这样抽出来的样本，应当说是很理想的样本。而从其它起点抽取的样本，往往会出现偏高偏低的情况。这从直观上就可以看到它的好处，简便易行，比较接近我国习惯上所用的划类选点的做法，易于为调查人员和广大基层干部、群众所接受。因此，我们过去在谈到有关标识排队等距抽样时，一般都只谈到这种办法。

采用组中点为起点，不仅在实践上是有效的，而且在理论上也是可行的。潘孝瑞、许刘俊等同志的文章，对此都已作过充分的论证。

由于我们在五十年代就对有关标识排队等距抽样方法进行过研究和试点调查，六十年代初期采用这种方法进行了全国规模的农产量抽样调查，当时我曾经编写过一本《抽样调查方法介绍》的小册子，着重介绍这种方法，广为流传；加上长期闭关锁国，对国外抽样技术的新发展了解很少，也没有看到过有谈到这种方法的。因此，我和许多同志都认为：这种方法是我国独创的；国际统计学界似乎没有人论证过，可能对这种方法是持否定态度的。直到去年，北京农业大学刘宗鹤教授寄了一本W. G. 柯克岱写的《抽样方法》（英文本，1977年第三版）给我，我才注意到，在柯克岱的著作中，也谈到了这种方法。并且持肯定的态度。不过，他不是叫做“有关标识排队”，而是叫做“直线趋势总体”；并认为当总体是呈这种“直线趋势”排列时，采用中心位置样本比随机位置样本会更为正确，而且要比普通等距抽样（即无标识排队）优越。可见，我的关于有关标识排队等距抽样方

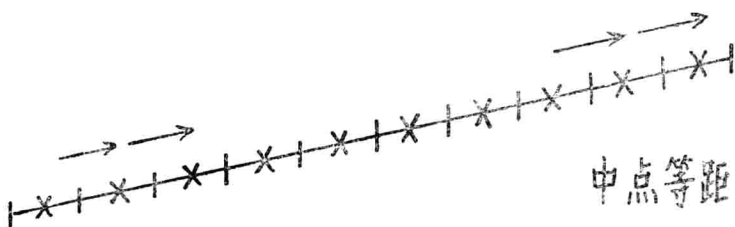
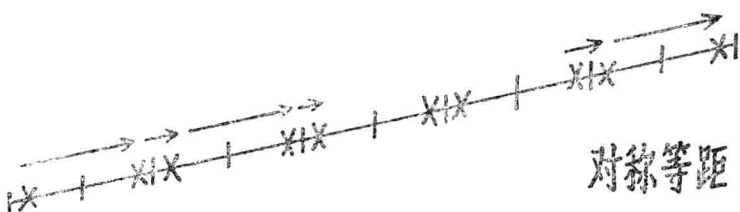
法的观点，在国际统计学界，也并非孤立的。所不同的是，柯克峦是把这种“直线趋势总体”作为可能遇到的一种特殊情况来对待，而我们则是自觉地把被研究的总体改造成为“直线趋势总体”。姑且不论这种方法是否我们创始的，但是，我深信就世界范围来说，对这种方法实践应用最多，最有经验，最有发言权的，仍然是我们国家。这是适合我国具体情况的一种行之有效的抽样方法。

二、改进或增强有关标识排队等距抽样方法随机性的办法

对有关标识排队等距抽样方法，能否不用组中点样本，而用任意起点随机样本，以改进或增强它的随机性呢？我认为，是可以的。在去年8月中国统计学会和广东省统计学会联合召开的“抽样调查在农业方面的应用”科学讨论会上，我和刘宗鹤同志等，都曾经提到过一些别的方法，譬如说，采用对称等距法，就可以和无关标识排队一样，以任意一点为起点抽取样本，并获得良好的无偏误推算数。例如，有100个综合商店，打算采用有关标识排队等距抽样方法，抽取其中10个商店来调查其营业状况，包括购、销、赢利等情况。根据有关材料，已知其投资额大小的排队次序，在求出抽样距离（ $R = \frac{n}{N} = \frac{10}{100} = 10$ ）后，即可采用对称等距方法，从第一个等距中的任意一点开始，对称地等距抽出样本单位。假定其起点为第2号商店，即第一个抽中单位，第二个抽中单位应为第二个距离中的倒数第二个单位，即第19号商店，第三

个抽中单位为22号商店，第四个单位为39号商店，以下抽中的商店为第42号，第59号，第62号，第79号，第82号，第99号。如以第7号为起点，则以下抽中的商店为第14号，第27号，第34号，……依此类推。

对称等距与中点等距的抽点方法示意图如下：



中点等距的每个样本单位的距离都是相等的；而对称等距实际是两两等距。

再一种办法是采用顺逆排列交替等距抽样法。仍用上例，已知 $N = 100$ ， $n = 10$ ， $R = 10$ ，则所有样本单位即可排列如下式：

划 分 等 距 数	分 样 本									
	1	2	3	4	5	6	7	8	9	10
1	1	2	3	4	5	6	7	8	9	10
2	20	19	18	17	16	15	14	13	12	11
3	21	22	23	24	25	26	27	28	29	30
4	40	39	38	37	36	35	34	33	32	31
5	41	42	43	44	45	46	47	48	49	50
6	60	59	58	57	56	55	54	53	52	51
7	61	62	63	64	65	66	67	68	69	70
8	80	79	78	77	76	75	74	73	72	71
9	81	82	83	84	85	86	87	88	89	90
10	100	99	98	97	96	95	94	93	92	91

从下表可以看出，用顺逆排列交替等距抽样法，以第一个距离中的任一点为起点，每样本单位的距离完全相等，抽出的10套分样本，实际上同前面用对称等距法所抽出的结果，是完全一致的。用这种方法，如果根据已知辅助信息，对每个单位不仅有次序上的了解，而且有了数量上的了解，例如，在前面的例子中，已经知道每个综合商店的投资额，则这种排队方法的好处，可以更明显地看出来。下面是根据各商店已知投资额按顺逆交替法排队的各套样本的情况：

按投资额 (万元) 顺逆交替排列的分样本

划 分 等 距 数	1	2	3	4	5	6	7	8	9	10
1	5.50	5.50	5.75	5.81	5.88	5.89	6.00	6.04	6.07	6.09
2	6.49	6.47	6.40	6.40	6.40	6.38	6.32	6.30	6.11	6.11
3	6.54	6.55	6.55	6.62	6.63	6.70	6.73	6.73	6.77	6.79
4	7.09	7.09	7.02	7.01	6.92	6.92	6.92	6.91	6.84	6.80
5	7.10	7.10	7.10	7.14	7.15	7.15	7.15	7.20	7.21	7.24
6	7.40	7.40	7.38	7.35	7.32	7.30	7.25	7.25	7.25	7.24
7	7.45	7.50	7.53	7.59	7.60	7.62	7.62	7.70	7.70	7.71
8	8.00	7.98	7.98	7.97	7.90	7.85	7.82	7.80	7.80	7.74
9	8.04	8.05	8.09	8.09	8.10	8.19	8.20	8.28	8.28	8.28
10	8.62	8.60	8.50	8.50	8.47	8.42	8.36	8.36	8.34	8.30
小 计	72.23	72.24	72.30	72.48	72.37	72.42	72.35	72.57	72.37	72.40
平 均	7.22	7.22	7.23	7.25	7.24	7.24	7.23	7.26	7.24	7.24

以上两种有关标识排队等距抽样的办法，和无关标识排队一样，都可以任意一点为起点，比中点等距更具有随机性，但抽样数目最好取偶数，否则就不能完全对称。

三、有关标识排队等距抽样和类型抽样的相同和不同之处

有关标识排队等距抽样和类型抽样有什么相同和不同的地方？为什么说这种方法是把等距抽样和类型抽样结合起来，兼有二者的优点，为什么不把它干脆归入类型抽样呢？

首先，让我们来看看有关标识排队等距抽样方法和类型抽样方法有什么相同之处。我认为，它们的相同之处，最基本的有以下三点：

（一）划分类型和有关标识排队，都需要事先搜集了解总体的有关资料，利用总体的已知信息和辅助信息。

（二）通过对总体的划类和按有关标识次序排队，都可以大大缩小总体的差异程度，增强样本对总体的代表性，提高调查结果的准确程度。

（三）类型抽样和有关标识排队等距抽样，都是属于限制随机性的抽样方法。

有关标识排队等距抽样方法和类型抽样方法的不同之处是：

（一）划分类型与抽样距离做法不一样。类型抽样通常是将总体单位按其属性特征分为若干类型，然后在各类型中抽取样本单位。例如，在进行工业企业抽样调查时，将工业企业划分为采掘、制造、加工、修理等几个类型；或是在进

行农业调查时，将一个地区划分为山区、丘陵平原等几个类型，然后在每个类型中分别抽选若干企业或生产队来进行调查。所谓属性特征，一般是指具有一定程度的质的差别的特征。因此，类型的划分，应有清楚的界限；同时，分类型的数目也不宜太多，否则将失去类型的特征，不便于在每类中抽样。而有关标识排队等距抽样，则是在对总体单位的次序有一个基本了解的情况下，进行排队抽样，对排队标识的了解，要求比划分类型更细；而距离的划分，是按抽样数目来确定的，因此每个距离之间，并没有清楚的质的界限。

（二）抽取样本单位的目的和方式不一样。类型抽样在划分好类型后，一般在每个类型中，再分别（按比例或不按比例）随机抽选样本单位。每个类型实际上就是一个小总体，抽选的样本单位，要能够用来代表各该类型，因而必须有一定的数量。而有关标识排队等距抽样，划分等距的数目就是抽取样本单位的数目，即从每个距离中只抽取一个样本单位，并不要求它对各该距离具有代表性，而是以全部样本单位作为一个整体来代表总体。当然，类型抽样也可以在各个类型中采用等距抽样方法，但这是在总体划分出各个类型以后。

（三）推算方法的不同。类型抽样原则上应分类抽样，分类代表，分类推算，即在分类的基础上推断总体，观察总体。而等距抽样则是以整个样本来代表总体，直接推算总体，观察总体。

（四）划类与分等距的作用不同。划类是本质的，有形的，在抽样、推算、计算抽样误差等每个环节，都有其实际的意义和作用。而有关标识排队等距抽样，等距的计算，只

是用以决定抽样起点和各个样本单位之间的距离，每个距离并不一定需要划分出来，它是虚的，无形的。只有在计算抽样误差时，这样的组距才有其实际意义。

因此，有关标识排队等距抽样与类型抽样是有所相同而又有所不同的，它在性质上，具有类型抽样的一些特点；而在组织形式上，则属于等距抽样。所以，我们说它是把等距抽样和类型抽样结合起来运用，兼有二者的优点。

能不能把等距抽样就看做是类型抽样呢？我认为，可以把它看做是一种特殊的类型抽样，而不是一般的类型抽样。因为，只有在这样的情况下：即把总体划分为若干相等的类型，每个类型只抽一个样本单位，则类型抽样就成了等距抽样。但是，在一般情况下，由于类型是根据总体的属性特征来划分的，这样，就不可能使每个类型都刚好相等，而且，按照类型抽样的定义，在一般情况下，每个类型中只抽一个样本单位来代表各该类型，是不符合类型抽样的原则的，如果类型的数目和抽样数目一样多，类型本身也就失去了意义。所以，我认为，有关标识排队等距抽样方法，作为等距抽样的一种形式，更能说明它本身的特点，如果把它归并入类型抽样，不仅将使我国公认并行之有效的一种好的抽样方法的特点消失，而且也可能给类型抽样方法带来混乱。既然如此，我们为什么要去干这种弄巧反拙的事情呢？

参考文献：

- 〔1〕龚鉴尧：《关于有关标识排队等距抽样方法的探讨》，见《统计研究》第5辑及《广东统计论文选》第一辑。
- 〔2〕潘孝瑞、许刘俊：《关于“有关标识排队等距抽样方

- 法”的统计分析》，《统计与预测》1982年第1期及《统计研究》第7辑。
- 〔3〕刘宗鹤：《我国农业现行多阶段等距抽样调查的又一个排队方法》，《统计与预测》1983年第2期。
- 〔4〕统计原理与经济统计编写组：《统计原理与经济统计》，中国财政经济出版社，1982年。
- 〔5〕天津财经学院教研室：《社会经济统计学原理》，1980年。
- 〔6〕宋元村、黄玉喜：《数理统计学》，湖南人民出版社，1982年。
- 〔7〕《社会经济统计学原理电视讲座学习材料》上册，国家统计局人事教育司印，1983年。
- 〔8〕龚鉴尧：《抽样法浅说》，中国财政经济出版社，1981年。
- 〔9〕W.G.Cochran: Sampling Techniques 3rd Ed. 1977, by John Wiley & Sons.