

统计学原理

参考资料

陈仁恩 黄良文 编

中央广播电视大学出版社

统计学原理参考资料

陈仁恩 黄良文 编

中央广播电视大学出版社

统计学原理参考资料

陈仁恩 黄良文 编

*

中央广播电视大学出版社出版

新华书店北京发行所发行

三二〇七工厂印刷

*

开本 787×1092 1/32 10.125 印张 224千字

1986年7月第1版 1986年10月第1次印刷

印数: 1—148,500

书号: 4300.162 定价: 1.35元

目 录

统计学的性质、内容和方法	黄良文 (1)
统计方法的性质	(美) 保罗G·侯埃 雷蒙J·杰森 (23)
社会经济统计学和数理统计学的关系	黄良文 (33)
谈谈统计指标的概念和特点	陈仁恩 (42)
典型调查在统计工作中的作用	刘心铨 (50)
抽样调查的随机原则	黄良文 (57)
统计分组的几个理论问题	(苏) 尼·舍列梅特 (63)
分组的种类和组数组距问题	陈仁恩编译 (72)
论统计相对数的分类	陈仁恩 (78)
论调和平均数	周雄鹏 (84)
浅谈算术平均数与调和平均数的关系	陈炳林 (93)
中位数和众数的应用	颜金锐 (99)
计算经济指标的平均发展速度和平均增 长速度的新公式	(苏) И·巴斯哈维尔 (119)
经济指数的两种编制方法	翁礼馨 陈仁恩 (125)
指数法发展的三个阶段	(苏) 格·柯瓦列夫斯基 (139)
相关分析和它在统计研究中的 运用(一)	(苏) И·范尼茨基 (153)
相关分析和它在统计研究中的 运用(二)	(苏) И·范尼茨基 (161)
抽样方法及其对抽样结果的影响	黄良文 (184)

关于抽样平均误差·····	黄良文	(195)
抽样中的概率度和置信度·····	黄良文 王美今	(208)
总体方差的估计·····	黄良文	(216)
类型抽样的效果分析·····	黄良文	(228)
抽样法百题判断·····	(苏)A·M·叶琳娜	(246)
统计预测中的指数平滑法·····	华伯泉	(265)
略论三点法在统计预测中的应用·····	颜金锐	(284)
中华人民共和国统计法·····		(296)
国务院关于加强统计工作的决定·····		(303)
中国的统计体制与统计机构·····	孙义仁 杨本全	(308)
中国统计体制的改革问题·····	黄良文	(313)

统计学的性质、内容和方法

黄良文

统计对象的性质

统计科学的性质决定于统计研究对象的性质。作为社会经济统计研究对象的社会经济数量方面具有哪些性质，它为什么要用专门的统计方法来研究呢？

首先要考虑社会经济现象的性质问题。我们知道，人是社会经济现象的主体，但社会经济现象的性质却不可能从每一个人的自然特征中寻找答案。一切社会经济现象都表现为一定人与人或人与物之间的关系，实质是人们相互之间的关系，其中最基本的是生产关系。所以社会经济数量的性质也只能从这些占有关系、交换关系、分配关系等等去探讨研究。所有这些关系都具体地存在于现实社会生活当中，不可能脱离周围的环境和条件而孤立地存在。因而对于社会经济现象数量关系的研究也就不能把它从现实生活中抽取出来，放在实验室里进行物理测试或化学反应，也很难用生物模拟或数学演绎的方法进行抽象的分析。而只能深入实际，向社会作调查，占有丰富的材料，从现实存在的复杂联系中探讨其内在的关系。正如马克思所说的：“研究必须充分地占有材料，分析它的各种发展形式，探寻这些形式的内在联系。”^① 这就是社会经济现象的社会性，它导致统计研究的第一个特点，在于它是一门社会调查研究的科学，它是关于对社会、政治、经济、文化各项数量和情况进行调查、整理和分析的理论和方法。

正基于这一点，我们不能同意多数西方统计学所持的观点，他们把统计看为单纯数据处理和数量的抽象研究。因而统计学是从数据整理的次数分配开始，而把资料的搜集作为已知条件，或推给社会学来研究，其结果便是抽掉了社会经济统计学的基础。

其次，要考虑社会经济数量的性质。社会经济现象的规模和数量对比关系是在一系列复杂因素的作用下形成的。在这些复杂的因素中，有些是主要的因素，它起着决定的、普遍的作用；有些是次要的因素，它起着局部的、偶然的作用。但不论是主要因素或次要因素，对各单位实际发生作用的程度和方向，以及各个因素相结合的方式都不可能完全相同。这就使得社会经济现象的数量具有这些性质，即对于每一单位的个体来说带有一定的随机性质，而对于一定多数的总体来说，又具有相对稳定的共同趋势。这就是社会经济现象的随机性，它导致统计研究另一特点在于它对同类大量社会经济现象的观察和综合分析，排除次要的、偶然的因素影响，以研究主要的，共同起作用的因素所呈现的规律性。统计和登记核算既有联系又有区别，统计以原始的登记核算为基础，但如果只停留在这一阶段，而没有汇总综合和分析过程，也就不能算做统计。

社会经济现象的随机性问题长期来被视为禁区，不许“妄加评论”，至今仍存在很大的分歧意见。这个统计理论的重要课题，需要作进一步的探讨。

社会自然现象可以分为两种类型。在一定的条件下，某种事件必然出现，而且其数量表现也为一定的，这种事件称为必然事件。必然事件的对立面称为不可能事件，必然事件和不可能事件属于肯定型的现象。在一定条件下，某种事件

可能出现，也可能不出现，其出现的数量可能大一些，也可能小一些，这种事件称为随机事件，随机事件属于非肯定型的现象。对社会自然界中肯定型现象的数量研究用数学分析的方法，它在人类的认识中已经有很长的历史，这种研究方法已经达到相当精密的程度。对社会自然界中的非肯定型现象，长期以来被认为是属于无法控制的偶然现象。随着科学的进步，认识能力的提高，人们发现对于个别现象可能出现也可能不出现的随机事件而在大量的总体中却具有某种规律性。在表面上是偶然性在起作用的地方，这种偶然性始终是受内部的隐蔽着的规律支配着，因而社会自然界中的随机事件也是可以为人们所掌握，研究这种非肯定型现象数量变化的就是专门的统计方法。由于自然现象和社会现象有不同特点，所以所用的统计方法也就不同，例如自然现象用试验统计的方法，而社会现象则用调查统计的方法。但由于它们都具有随机的性质，所以统计的方法也有共同的一面。

非肯定型现象是现象的一种客观属性，人们对事物的认识水平和技术水平的提高，只能改变随机事件的条件，而不能改变随机事件的属性，更不能消灭随机事件。那么，社会经济活动中为什么存在着非肯定型现象呢？有一种事件它的出现决定于多种多样的原因，决定着每一个体事物出现的原因对于总体的作用又都都比较小，而且这些原因的产生受许多偶然的因素影响，因而它对各个个体所起的作用不论在程度上，甚至在方向上都可能不同，这就导致事件的随机性。例如火车站每天发售客票，其数量多寡决定于乘客的多少，而旅客乘车当然都有各种原因，但这些原因对于发售客票总量来说作用又都相对地小，旅客乘车原因如此复杂，因而一种原因对旅客或起作用或不起作用，或起这样作用或起那样作

用。所以客票发售不发售是随机事件，而售票的张数则是随机变数，像这一类非肯定型的现象在社会经济生活中是大量存在的。在统计工作中，对社会经济现象进行观察、测量、登记，除了可能存在系统性的误差外(如瞒产、虚报等等)，由于主观的疏忽和客观条件的限制，引起各种偶然性的误差，也增加了统计资料的随机性。当然我们也要看到，社会经济现象不但受随机因素影响，而且受必然的因素影响，这两种因素总是交错结合在一起的。例如上述发售火车票的例子，(在年关时节乘客特别拥挤，由于铁路事故而停止售票等等这都不能说是随机的因素)。在统计分析上相应有两种处理方法，其一是根据某一必然因素的标志对现象进行分类，而形成一定条件下的随机现象。其二是对现象结果的因变量根据形成原因的性质，用一定的数学方程，区分为解析变量影响和随机变量影响，列出随机项(亦称干扰项)，并且对随机项作种种假设(例如随机误差总和等于零，随机误差有一定的标准差等等)。虽然由于数学模型的完善程度不同，有可能将必然因素挤入随机项，但随机项的假设则是根据统计现象的随机性质而来的。

构成统计总体的先决条件是单位的标志值存在着差异，这种差异必然是随机性的。如果不存在这种差异，也就不存在统计方法。统计的基本方法是大量观察法、分组法和综合指标法，这些方法都是和统计现象的随机性联系在一起的。大量观察法的目的在于取得了解全面情况的资料，而在于根据随机现象大数稳定的原则，只有观察足够多的单位，才可能通过偶然性发现必然性，认识现象规律性的表现形式。统计分组法表面上似乎和现象的随机性关系不大，其实不然。分组法就其数量关系来说，正是运用了上面所提的条件、

随机性的原理，来探讨现象之间的内在联系。例如按施肥量分组的粮食亩产量，就是假定同一施肥量/各组内决定粮食亩产的其他非肥料因素都是随机的，它们的误差总和为零，这样才能显示施肥量和粮食产量之间的关系，否则分组法就成为没有意义的计算。综合指标法和统计现象的随机性更有密切的联系。例如计算平均指标的前提乃在于总体各单位存在着数量上的随机差异，因而一个单位的标志值并不能代表总体的一般水平。也正由于这种差异的随机性，才有可能通过平均化的计算来消除偶然的偏差，显示一般的典型水平。而平均指标能消除偶然偏差，则由于总量指标已把个别的差异舍象掉了。大量观察法、分组法、综合指标法是统计普遍适用的方法，也说明统计现象的随机性也是普遍存在的。

出于种种原因，人们对于社会经济现象随机性的认识还有很大差距。其中之一在于对随机性的理解上。有些同志没有认识到随机性的非肯定型现象是现象的一种客观属性，而把它看为人们主观的认识程度，这实是一种误解。例如他们把非肯定型理解为不肯定性，进而断定其中一种是由于人们对事物真实情况的无知造成的。人类从无知到有知，从片面认识到全面认识，也就是从不肯定到肯定过程，这是所有的科学研究都要经历的过程。这种表明人类主观认识程度的不肯定性，和统计对象的属性又有什么联系呢？另一种认识，把随机偶然性看成一种神秘不可知，无确定原因的偶然性。事实上任何一种现象的发生都是有原因，而且有确定的原因，随机性的现象也同样是有原因的。人们可以有两种不同的认识途径：对于肯定型的现象，由于因果联系比较单纯、固定，可以用解剖分析的方法，探讨各种原因的具体联系，并且掌握和控制各种原因所形成的条件。对于非肯定型现象，由于

因果联系错综复杂；而且每个原因对于总体的作用都相对地小，所以我们可以用综合统计的方法，从总规模、水平上认识各种原因所引起的共同趋势，认识事物变化的规律性。统计成为研究因果联系的一种重要方法，我们没有理由因此认为随机现象是没有原因的。

怎样认识人类社会有意识有目的的行动对于现象随机性的影响呢？社会经济现象和自然现象不同，它不是盲目的、无意识的各种动力彼此互相作用的结果，而是人们有意识有目的行动的结果。这个不同点曾被人们作为否认社会经济现象具有随机性的有力论据。然而，人们的主观愿望如何是一回事，而主观愿望能实现到什么程度又是另一回事。由于所处的环境不同，根据的条件各异，相同的愿望对各人的结果可能完全不同，更何况各人的愿望本身就是千差万别，乃至相互矛盾冲突。因此人类社会历史的进程同样要依靠偶然性来开辟自己的道路。正如恩格斯所指出的：“即使在 这一领域内，尽管各个人都有自觉期望的目的，在表面上，总的说来好象也是偶然性在支配着”，“这样，无数的个别愿望和个别行动的冲突，在历史领域内造成了一种同没有意识的自然界中占统治地位的状况完全相似的状况。”^②

魁特莱第一个用统计的方法，论证了社会生活中随机偶然现象贯彻着必然的规律性。他在1836年出版的《社会物理学》中，用大量的事实说明社会生活中表面上偶然的事物，在它的定期的复现和一定的平均数中，呈现出其内在的必然性。马克思高度评价魁特莱所做的贡献，认为“他过去有很大的功绩。”^③此后一百年来，人们对于社会经济随机现象的认识和控制能力都有了迅速的提高。怎能无视这些历史事实，否定社会经济现象的随机性呢？

有人说，那是资本主义社会的情况，在那里社会经济规律作为自发强制的力量发生作用，所以社会经济现象的随机性是可以理解的。而在社会主义社会情况就不同了。在社会主义条件下，人们可以组织起来，自觉地利用客观社会经济规律，实行计划经济制度，保证国民经济有计划按比例发展。这里必然性占据着统治地位，社会经济现象的随机性再没有容身之处了。他们为社会经济现象设置种种限制，并断定马克思、恩格斯的话只适用于资本主义社会。只需稍为分析一下就会发现，在社会主义制度下，国民经济计划也是一种人们预期的目的，而计划的组织制度则是人们掌握控制社会经济随机现象的一种方法。由于实现了社会主义公有制，使预期的主体从个体扩大到集体，这有可能大大地提高人们认识和掌握社会经济随机现象的能力，执行计划就是这种认识和掌握的过程，而计划的执行结果则表明认识和掌握的程度。不能认为计划制订了，一切都会按计划办事，社会就没有矛盾了，或矛盾都是事先规定，可以预见的。事实上，一种计划可以有各种各样的结果，这不但和计划制订的完善程度有关，而且和执行计划的条件有关，从整个社会看来，这些条件所发生的影响都是带有随机性的。社会经济现象随机性使认识和掌握这类现象产生许多困难，以至计划的制订和统计的核算常有失误，这只能从提高计划统计工作水平来解决，而不是去否认社会经济现象随机性本身。在社会主义社会还存在着大量的非计划现象，或由于主观条件的限制，我们的计划还不能不是粗线条的、大部类的，在计划的安排上也必须有更大的灵活性和自主性。例如居民收入、家庭消费、市场经济乃至文化科学、思想意识等领域，这些现象的随机性就更不用说了。

有些同志并不否认社会经济现象的随机性，也不否认社会主义条件下随机现象的存在，但他们认为在社会主义统计中随机性只存在于抽样调查现象，至于全面统计则无随机性可言。因此得出结论是，社会经济生活中决定性的现象是主要的、大量的，而随机性现象只是次要的、少数的。他们举例说，工厂年初实有职工500人，本年退休离职20人，招工补员50人，年末实有职工530人，这里 $500 - 20 + 50 = 530$ 只有必然性没有随机性！这是一种很有代表性的看法。实际上现象的随机性不可能从运算的方法中表现出来，数学运算的严格准确也并不能否定每个数量的随机性。年初职工为什么恰好是500人，本年退休离职为什么是20人，招工补员为什么又恰好是50人，这都是带有随机性质的，而这三个数字中的任何一项随机性又决定着年终人数的随机性，这才需要用统计的方法来研究职工人数的变动情况。进一步说，所谓全面资料也都只是相对而言，如利用一个地区的资料、一个年度的资料进行推断和预测，又会产生代表性的随机性误差。再就抽样调查的随机性来说，那是指抽样平均数或成数，更一般地说即抽样统计量的分布具有随机性，抽样误差的分布具有随机性，至于抽样平均数、成数的计算毫无例外是准确精密的。现象的随机性和计算的严密性并没有矛盾。现象的随机性是客观存在的，并不因为运用全面调查方法，或抽样调查方法而改变其性质。把统计对象划分为两部分，其中全面调查的对象无随机性，抽样调查的对象有随机性，是缺乏根据的。

有些同志出于好心，顾虑承认社会经济统计现象的随机性，势必数理统计学代替社会经济统计学，概率论代替历史唯物论和政治经济学作为统计学的理论基础，社会经济统计

学就行将消亡。所以认为社会经济统计现象的随机性是一种异端邪说大逆不道。他们扬言社会主义统计应该立足于必然性，而不应该立足于偶然性，就如同建设社会主义不能侥幸取胜，不存在什么概率论或大数法则起作用的问题。这里，我们需要进一步研究，大数法则在社会主义统计学中能起什么作用，不能起什么作用；概率论在社会主义统计学中所起的作用和哲学政治经济学在社会主义统计学中所起的作用有什么联系和区别。已如上面指出的，社会经济统计对象不但有随机性一面，而且还有社会性一面，属于随机性的研究大数法则提供了理论依据和方法基础这是不可否认的，属于社会性的研究则大数法则就无能为力这也是明显的。在社会经济统计中，大数法则不是一开始就能发挥作用，更不是自始至终都发挥作用。我们知道，构成统计总体的前提是现象的同质性，而判断现象是否同质，只能由社会政治经济理论来回答，大数法则无法解决这个问题。统计的科学性首先体现在设置切合实际的统计指标体系和统计分组体系。统计指标是一定社会经济范畴的数量表现。如何根据社会经济问题提出一定的统计任务，又归结一定的统计指标和指标体系，並由此规定统计指标的涵义、口径、计算方法、计量标准等等都必须在正确的社会经济理论的指导下才能解决。统计分组必须反映现象的类型特征，说明事物变化的数量界限，分组标志如何选择、组数组距怎样确定，这些也不是大数法则所能解决的问题。再就统计的数量分析来说，大数法则的作用在于通过大量数据的综合汇总，抵消各单位变量的偶然误差，而显现出现象所表现的共同趋势，反映现象发展规律性在具体时间、地点、条件下的表现。就是说大数法则的作用是帮助我们通过偶然性达到发现必然性，认识规律性的表现形式。

然而它不能说明现象的本质，因为现象规律性的内容决定于现象本身事物的矛盾性质，不可能用概率论来解释。社会经济现象的本质只能由历史唯物论和政治经济学来说明。魁特莱运用概率论和大数法则发现了社会现象的必然性，做了很大贡献，但他对社会现象必然性的解释却是失败的。他把社会规律和自然规律等同起来，把历史的社会规律性看为永恒不变，甚至归结到神定的秩序。所以马克思在肯定他的功绩，同时还指出他从来没有能对社会现象的必然性作出解释，三十年来他实际上並沒有进步。我们论证社会经济现象的随机性在于说明社会经济统计可以充分利用数理统计的成就，为更有效地认识社会服务，但数理统计並不能代替社会经济统计，概率论也不可能代替历史唯物论和政治经济学作为社会经济统计的理论基础。

概率论为统计方法提供理论基础和唯物辩证法对统计方法的指导並不是互斥的两码事。概率论是关于随机变量的理论和方法，它在数量领域中具体地体现了偶然与必然、个别与总体、特殊与一般的辩证关系，统计正是运用这些科学的理论和方法来研究统计规律性，这怎能说接受概率论的理论方法基础便排斥了唯物辩证法对统计方法的指导呢？必须提出，过去长期以来，将概率论和哲学政治经济学对统计学的指导机械地对立起来，使社会经济统计学和数理统计分了家，断了缘，其结果便是划地为牢，作茧自缚，这个精神枷锁应该彻底解脱。

总而言之，由于社会经济统计对象具有社会性和随机性，一方面它要向社会作调查，所以它必须依据马列主义哲学和政治经济学的理论指导，另一方面又是对大量社会现象的数量分析，所以它又必须充分运用数理统计所提供的方法。两

者既不相互矛盾，也不能相互代替。

统计理论的内容

在统计学原理与专业统计的关系上，过去流行的看法是，统计学原理讲方法，专业统计讲指标，两者是分工和补充的关系。这样，统计有没有理论？理论的内容是什么？就成了问题。

我认为，统计学原理应该研究指导一切统计工作的最基本理论，是各专业统计工作经验的最一般概括。统计学原理与专业统计的关系应该是基本理论与具体应用的关系。

那么什么是统计的基本理论呢？统计基本理论不能只是统计的方法论，统计本体也应该有它的理论。什么又是统计的本体呢？统计研究社会经济现象的数量方面，具体说就是通过统计所特有的指标和指标体系，以及相应的分组和分组体系来反映这些现象的规模、水平、比例、速度，以阐明社会经济发展的规律性。所以统计指标和统计分组构成统计本体的两要素。（统计指标和分组的理论应该是统计本体的基本理论。然而我们在这方面的研究是很不够的。

统计指标是对现实生活中大量反复出现的具体社会经济现象的某种共同数量特征加以概括而形成的概念，所以它是一定社会经济范畴的反映。统计指标虽然表现为一定的数量，但并不是任何数量都可以构成指标，当然也不是任何范畴都能够用数量来表现的。作为统计指标首先必须符合政治经济理论对于范畴所作的一般概括，而且进一步加以具体化，使它有明确的界限和统一的方法。另一方面，统计指标又是对总体各单位的具体数量表现进行综合汇总而归纳形成能说明

总体的综合数量特征。所以统计指标既是对一定社会经济范畴的具体化过程，又是对具体事实的抽象化过程。这也显示社会经济统计既不同于数理统计，也不同于会计核算和业务核算的地方。

一个统计指标反映着社会经济现象的某一特征，说明矛盾的一个侧面。要反映客观现象的各个方面，事物发展的全过程，就必须运用一系列相互联系的指标体系。有科学根据的国民经济指标体系，应该是指标相互衔接，反映国民经济活动的一切主要现象和过程，能表明社会主义再生产的客观规律性。统计指标的科学界定，以及指标体系的建立是统计理论必须深入研究的问题。

从统计指标的作用来看，基于统计任务的不同，可以有各类不同的指标。例如检查计划指标、考核指标、观察指标等等。检查计划指标主要是为了加强国民经济的计划管理，对计划执行情况进行检查监督，这类指标是社会主义计划经济制度的一个组成部分，是社会主义统计所特有的指标，在国民经济指标体系中占着主导地位。考核指标是国家对基层企业主要的生产经营活动规定必须实现的最低目标，它带有指令性的要求，也是评价企业和个人经济利益的直接依据。观察指标主要是为分析国民经济活动情况而设置的指标。例如有些社会现象虽然未能加以计划控制，但却是反映社会的基本情况，如人口的出生死亡变动、消费的构成变化、气象资源变化等等，有些则是分析计划执行情况所必须的，如出勤率、废品率、劳动装备率等等。国民经济中这三类指标基本上是统一的，但也存在着种种矛盾。有些指标从检查计划的角度来看，要求严格按计划的口径和计算方法，而从反映总规模水平来看，又要求按现象的性质进行统计。有些指