

导 言

“社会经济统计是社会认识的最有力的武器之一。”

(列宁语)

社会经济现象包括质量和数量两个方面。研究社会经济现象不仅要研究它的质量方面，而且要分析它的数量方面。没有数量也就没有质量。

西方应用经济统计就是研究西方国家分析社会经济现象数量方面的统计理论和应用的统计方法。根据西方应用经济统计的研究对象和应用范围，又可以分为描述统计和推断统计。描述统计是通过调查，收集大量的数字资料，进行加工、计算并且通过规定和编制经济统计指标体系和计算经济统计指标数值来反映社会经济现象的数量表现、数量关系和数量变化的。它反映了社会经济现象的水平、规模、结构、发展速度、普遍程度以及比例关系。推断统计则是以概率论为基础，借助抽样方法，以局部推断整体，从而对不确定的事物作出决策。推断统计不仅以样本估计总体，而且对估计值进行假设检验以确定其有效性，从而有助于决策的选择和决断。

自本世纪20年代以来，特别是第二次世界大战以后，随着计算方法和计算技术的发展，西方经济统计学界开始把重点从描述统计转移到推断统计。但是，在微观经济分析和宏观经济分析中，描述统计和推断统计各有其不可替代的作用。

统计的语言是数字的语言。但是，应用经济统计绝不是为数量而研究、计算数量的。它是要通过对社会经济现象或对样本的数量分析来研究社会经济现象的本质，从而认识和推断微观或者宏观经济现象的质的特点及其规律性，并为正确的决策提供科学的依据。

西方应用经济统计的理论和方法不仅适用于西方国家，而且具有一般意义。它不仅可以用于认识和分析西方国家经济体制中的社会经济现象，而且对我国经济的分析也有借鉴意义。

目 录

导 言	1
第一章 统计调查和统计资料的整理	1
一、统计调查	1
1. 统计调查的基本方式	1
2. 统计调查的主要方法	5
二、统计资料的整理	7
1. 统计分组	7
2. 统计表	9
3. 统计图	11
第二章 统计资料的综合分析	18
一、绝对数	18
1. 时期绝对数和时点绝对数	18
2. 绝对数的计量单位	19
二、相对数	20
1. 结构相对数	20
2. 比较相对数	21
3. 动态相对数	22
4. 强度相对数	22
5. 计划完成相对数	23
三、每增长百分之一的绝对值	23

四、平均数.....	25
1. 算术平均数.....	25
2. 中位数.....	27
3. 众数.....	29
4. 几何平均数.....	31
5. 倒数平均数.....	32
五、离差.....	33
1. 全距.....	34
2. 平均差.....	34
3. 方差.....	36
4. 标准差.....	37
5. 离散系数.....	39
 第三章 指 数.....	43
一、个体指数.....	44
二、综合指数.....	49
1. 简单指数法.....	49
2. 加权指数法.....	51
3. 加权指数公式的变形.....	54
三、平均数指数.....	57
四、指数的基期、权数时期和年率.....	61
1. 指数的基期.....	61
2. 指数的权数和权数时期.....	62
3. 指数的年率.....	63
 第四章 动态数列分析法.....	66
一、发展水平.....	66

二、增长量	67
三、发展速度	68
四、增长速度	69
五、平均发展水平	72
1. 绝对数动态数列的平均发展水平	72
2. 相对数动态数列的平均发展水平	76
3. 平均数动态数列的平均发展水平	79
六、平均增长量	80
七、平均发展速度	81
1. 水平法	82
2. 累计法	83
八、平均增长速度	84

第五章 动态数列的分解 87

一、长期趋势及其测定	87
1. 移动平均法	88
2. 数学曲线配合法	91
二、季节变动及其测定	95
1. 按月（或按季）平均法	95
2. 移动平均法——相除法	99
3. 移动平均法——相减法	103
三、循环变动及其测定	107
1. 剩余法	107
2. 指标分析法	109
四、偶然因素	114

第六章 基本概率和概率分布 116

一、基本概率	116
1. 概率和概率树	116
2. 结合事件和条件事件的概率	120
3. 贝叶斯定理	126
4. 决策树	127
二、概率分布	131
1. 离散随机变量及其概率分布	131
2. 离散随机变量的平均值和方差	134
3. 二项式随机变量的概率及其分布	135
4. 连续随机变量及其概率分布	137
5. 正态分布	140
三、两个随机变量的概率及其分布	143
1. 联合分布和边缘分布	143
2. 两个随机变量的相互独立	146
3. 协方差	147
第七章 平均数和成数推断	156
一、样本和总体	156
1. 样本平均值的期望值和方差	156
2. 样本平均值的概率分布	160
3. 成数的概率分布	163
二、点估计	166
三、置信区间估计	169
1. 单个平均值的置信区间估计	170
2. 小样本和 t 值	171
3. 两个独立样本的平均值的置信区间	172
4. 成对样本平均值的置信区间	175

5. 成数的置信区间	178
6. 一侧的置信区间	179
四、假设检验	181
1. 用置信区间进行检验	182
2. 用一侧的P值进行检验	184
3. 用临界值进行检验	187
五、方差分析	189
1. 单因子方差分析	189
2. 两因子方差分析	193
第八章 回归分析和相关分析	202
一、简单回归分析	203
1. 简单回归方程和回归线	203
2. 简单回归模型的推断	208
二、复回归	217
1. 复回归方程	217
2. 复回归方程中 β 的置信区间估计与假设 检验	220
三、回归分析的外延	224
1. 虚变量 (0—1)	224
2. 简单的非线性回归	227
四、相关分析	231
1. 简单相关	231
2. 回归中的相关	238
附表 1: 随机数字表	246
附表 2: 平均增长速度查对表 (选页)	248
附表 3: 二项式概率P (S)	250

附表 4: 标准正态分布, 右侧累积概率	253
附表 5: t 表	254
附表 6: F 表	255

附录 8 — 1	259
附录 8 — 2	260
附录 8 — 3	260

第一章 统计调查和统计资料的整理

要如实地描述客观经济现象就需要用科学的方式和方法对客观经济现象的数量方面进行统计。一般地说，经济统计工作分为三个阶段，即统计调查，统计资料的整理和统计资料的综合分析。

一、统计调查

统计调查是整个统计工作中非常重要的阶段。它是收集社会经济现象数量表现原始资料的工作阶段，是进行统计资料整理和分析的前提和基础，也是保证整个统计工作科学性和真实性的基本环节。统计调查的目的就是取得客观经济现象的科学数据。

1. 统计调查的基本方式

统计调查的方式是指对调查对象总体组织统计调查的方式。组织统计调查的方式一般分为全面调查和非全面调查。全面调查是对总体中每个总体单位都进行的调查。非全面调查只对总体中的部分总体单位进行的调查。非全面调查的结果常常用来推断总体。

① 普查。普查是西方国家对总体进行全面调查的唯一方式。普查是一种专门组织的全面调查。它一般是相隔若干年定期进行的。其相隔年限，不同国家根据不同的普查内容

而不同，如10年、5年、2年或1年。西方国家组织的普查种类很多，有人口普查、住房普查、制造业普查、矿业普查、农业普查、商业普查等，总称之为国情普查。在一些国家常常结合人口普查同时进行有关住房、工业、农业、商业等方面的调查。

西方国家普查的内容既包括一定时点下的状态性现象，（如人口、资产、库存等），也包括属于过程性的现象（如产量，销售额等）。由于普查是对总体的全部总体单位的全面调查，因此通过普查获得的数据比较全面、准确，称之为基准统计数字或者标志性统计数字。

最早的普查是法国1669年举行的生产普查。而以法律形式明确规定普查的周期，并设立专门机构负责举行定期普查的现代普查制度则始于美国1790年的人口普查。1790年3月美国国会通过普查法案，规定1790年进行第一次人口普查，以后每10年定期进行一次。

② 抽样调查。在两次普查的间隔期间以及对不适宜或者不必要进行全面调查的总体，如具有破坏性的产品质量检查等，就需要采用非全面调查方式，即抽样调查。抽样调查还常常用于对普查资料进行检验和修订。

非全面调查只对总体中的部分总体单位进行调查，因此能节省时间、人力、物力和财力。随着抽样理论和抽样技术的发展，西方国家日益广泛地应用抽样调查。通过抽样调查获得的数据称之为现时性统计数字。

抽样调查主要有以下几种类型：

第一，随机抽样调查。随机抽样调查是西方国家应用最广的非全面的调查方式。它是根据抽样理论，按随机原则抽选总体中的部分总体单位作为样本来进行调查的一种方式。

在随机抽样调查中，总体中的每个总体单位都有同等被抽取的机会，不受任何主观意图的影响。这就是说，构成样本的总体单位完全是凭偶然性抽取的。样本的随机性使样本对总体具有充分的代表性，从而保证了从样本推断总体的科学性。

随机抽样调查的组织方式有：a. 纯随机抽样或简单随机抽样。这是完全凭偶然性抽取样本。具体方法可以是抽签或者在总体单位较多时，把总体单位随意编排上数字，然后运用随机数字表（附表1）抽取样本。纯随机抽样还分为重复抽样和不重复抽样。重复抽样是指总体单位被抽取后将放回总体，从而使总体中的总体单位总数在整个抽样过程中保持不变，而每一个总体单位都有被重复抽取的可能。不重复抽样是指总体单位一旦被抽取后就不再放回总体。总体中的总体单位总数随抽取的次数而相应减少，而每一个总体单位只有一次被抽取的机会。b. 机械随机抽样或系统随机抽样。机械随机抽样要求把总体单位按一定标准顺序排列，如按字母顺序，或按现象出现的时间先后顺序排列。然后按一定间隔抽取若干总体单位组成样本。c. 分类抽样。分类抽样是先把总体按一定标准分为若干类别，然后按类别分别随机抽取一定单位作为样本。例如，西德的家庭收支调查就是先把全部家庭按人口和收入分为五个类别，然后在每一个类别中抽取若干家庭作为调查样本。d. 多级随机抽样。这是把总体按一定标准分为若干级，然后逐级抽取，直到抽出总体的基本单位组成样本。例如，美国普查局在进行现时人口调查时就是采用多级随机抽样调查方式。他们首先把美国全国划分为2000个地区作为基本抽样单位，每个基本抽样单位又分为100个枚举地区，每个枚举地区再分为100个最终抽样单位，每个

最终抽样单位大约包括 3—4 个家庭，约 10 人。在抽取样本时，首先在 2000 个基本抽样单位中随机抽取 500 个单位，然后在抽中的 500 个基本抽样单位中各随机抽取 5 个枚举地区，最后再在每个抽中的枚举地区中各抽取 5 个最终抽样单位，通过这样逐级抽取的 12500 个最终抽样单位，约 37500 到 50000 户家庭，约 12500 人，就组成了美国现时人口调查的样本。

e. 整群随机抽样。它不是以一个基本总体单位为抽样单位，而是以若干总体单位组成抽样单位，然后用纯随机或机械随机抽样方式整群抽取。如美国现时人口调查中，每一步都是整群抽样，一直到最终抽样单位。

f. 序列抽样。序列抽样的特点是样本规模不确定。如果小样本已明确得出结论，调查就可以到此结束。但是，如果小样本的结论不明确或者模棱两可，那末就要按序列扩大样本规模，直到得出确切结论。各种随机抽样调查方式可以单独应用也可以结合应用以提高抽样调查的效果。

第二，代表抽样调查。代表抽样调查是西方国家经常应用的非随机抽样调查方式。它是按照一定标准有意识地在总体中选择若干合乎标准的代表单位作为样本来进行调查的一种方式。代表单位的标准因不同统计内容而异。例如，在美国，统计批发价格指数时就以交易额的大小作为抽取代表商品的标准，共选择了大约 2800 种交易额大的商品作为代表单位组成样本进行调查。再如，日本在推算工业生产指数时，则是以增加值的大小作为选择代表产品的标准，共选择了约 460 种增加值大的工业产品作为统计工业生产指数的代表产品组成样本。组成样本的代表单位的标准和数量在不同国家、不同时期不尽相同。

代表抽样调查的结果通常能反映所要统计的基本情况，

因此，具有较高的代表性。西方国家常常以此来推断总体或代表总体。但是，代表抽样调查的结果不能计算误差。

第三，广泛抽样调查。广泛抽样调查也是一种非随机抽样调查方式。在广泛抽样调查中，选作样本的单位具有任意性，对调查结果也不严格要求，因此是代表性不高的非全面调查方式。但是，这种方式简便易行，因此，在实际工作中仍常常采用。例如，美国劳工部劳工统计局就采用这种方式来调查就业、工时和收入等。这种调查以全部大企业为调查对象并常常根据没有全部交回的调查表来计算结果。

第四，典型调查。典型调查也是一种非全面调查，是对典型现象所作的专门调查。典型调查有它的独特作用。它有助于区别个别与一般，从而能具体深入地反映所研究的现象。例如，对美国或世界最大的几百家公司的资产额，销售额等的调查资料就能够反映当代资本主义经济中资本集中的程度。

典型调查的结果既不能用以推断总体，更不能代表总体，并且在一般的统计资料中也往往没有反映。

2. 统计调查的主要方法

统计调查的方法是指收集原始资料的具体方法。西方国家进行统计调查的方法主要有三种，即访问法，登记法和通讯法。

① 访问法。访问法就是由统计机构雇佣调查员按照调查内容访问被调查者，并由调查员或被调查者当场填写调查表取得调查资料的方法，例如人口普查。采用访问法让调查员亲自主动访问被调查者可以较好地解决调查中的各类问题，消除疑虑，解释和说明调查项目和要求等，从而促使调

查表填答全面、准确。因此，用访问法取得的原始资料一般填写比较完整，数据比较准确，质量比较高。但是，访问法需要耗费较多的时间、人力、物力和财力。

② 登记法。登记法就是由统计机构发出公告，规定应当登记的事项，由当事人或与当事人有关的人，于事件发生后主动到统计机构登记，填写登记表（调查表）。例如，人口的出生和死亡，婚姻状况登记等。登记法也可以是被调查者按统计机构规定的调查内容自己登记发生的情况，并定期提供给统计机构，例如，家庭收支统计。采用登记法调查时，统计机构处于被动的地位，因此往往会出现不登记或漏登记的情况。因此，通过登记法取得的数据会有偏差。

③ 通讯法。通讯法也称报告法，是由统计机构采用通讯的方法把调查表寄给被调查者填写，再寄回统计机构的方法。例如，美国劳工统计局的就业调查就是采用通讯法进行的。采用通讯法寄出的调查表往往不能全部寄回，而且寄回来的调查表也常常由于各种原因填写得不完整。因此，其准确程度比较低。

除了通过上述专门组织的统计调查方式和方法可以取得原始统计资料之外，西方各国政府业务活动的统计数字也是各国统计资料的重要来源。各国政府的行政管理业务常常和经济活动有密切关系。这些政府业务活动的记录数字综合整理而成的资料是西方各国统计资料的重要组成部分。例如，西方各国的财政统计资料和金融统计资料几乎完全是政府财政、金融机构行政管理业务活动资料整理而成的。

二、统计资料的整理

统计资料的整理是统计工作的第二阶段。统计资料的整理就是把调查到的原始资料进行汇总，分组，列表或制图。

统计调查得来的资料是反映个别调查单位事实情况的材料，而统计资料的整理则是把大量个体的原始资料经过科学的分组和计算，综合加工成描述总体数量特征的统计资料。

1. 统计分组

统计分组是统计资料整理的一个必要步骤。统计分组就是根据原始资料中各总体单位的标志进行分组。统计分组分为品质标志分组和数量标志分组。

① 品质标志分组。品质标志分组就是根据总体单位的自然特征或社会特征进行分组。例如，人口的男、女分组，商品按耐用品和非耐用品分组，以及食品、药品等分组即是按其自然特征进行的分组；而人口的文化水平，人们的职业状况以及按批发或零售来区分商品则是按其社会特征进行的分组。

品质标志分组的原则是“互相排斥”，即一种经济现象只能属于某一个组，而不能同时属于两个以上的组。因此，按品质标志分组后，各组数据之和与总量相等。

② 数量标志分组。数量标志分组就是根据总体单位的数量特征进行分组。例如，人口的年龄分组，企业按销售额或者资产额进行分组，家庭按收入多少进行分组等。

数量标志分组分为单项数列和组距数列。单项数列是用于数量标志不多的经济现象的分组的。例如，某个国家规定18

岁到21岁的男青年应根据需要服兵役。现在要统计符合服兵役男青年的总人数及其年龄结构。这时可以用单项数列，按18岁、19岁、20岁、21岁分成四个组，然后分别计算各年龄组的男青年人数以及四个组的总计以描述该国的兵源及其年龄结构。

对数量标志较多而且变动范围较大的经济现象则需要用组距数列进行分组。例如，全部人口的年龄分组在美国分为5岁以下，5—15岁，16—19岁，20—24岁，25—44岁，45—64岁，和65岁以上共七个组距数列组。再如，1973年，美国家庭按收入分为2500美元以下，2500—4999美元，5000—7499美元，7500—9999美元，10000—12499美元，12500—14999美元，15000—24999美元，以及25000美元以上，共八个组。

在组距数列中，每一组中的较小数值为该组的下限，较大数值为上限，如上述人口组距数列分组中，第2组的下限是5岁，上限为15岁。上限和下限之差为该组组距，如上述组的组距为 $15 - 5 = 10$ 。上限和下限之和再除以2，即为该组的组中值，如上述组的组中值为 $(15 + 5) / 2 = 10$ 。

编制理想的组距数列应努力遵循以下原则：

(1) 避免分组过多或者过少，一般以不少于5组，不多于15组为宜；

(2) 一般情况下应尽力作到各组组距相等，以便于比较分析；

(3) 应尽量避免出现开口组，如上述美国家庭收入分组中的第一组2500美元以下和最后一组25000美元以上。因为，开口组忽略了最小数值或最大数值，从而影响对比分析的准确程度。

对总体单位或者样本单位很多，数值变动范围又较大的经济现象可以通过一定的计算方法来求得较为合理的组数和组距，即：

$$\text{组数} = 1 + 3.322 \log N$$

其中，N为总体单位或样本单位的数量，对于样本也就是样本规模；log为自然对数。

$$\text{组距} = (\text{最大数值} - \text{最小数值}) / \text{组数}$$

例如，把10000个工人的工资编制成组距数列，其中最低工资为100，最高工资为800。根据前述计算方法可得：

$$\text{组数} = 1 + 3.322 \log 10000 = 14.288 \approx 14$$

$$\text{组距} = (800 - 100) / 14 = 50$$

这样就可以编制成各组组距均为50的，从100—150，到750—800的共14个组的组距数列。

品质标志分组和数量标志分组可以分别用于统计资料的汇总和整理，也可以结合使用。例如，在人口的各年龄分组中，同时统计该年龄组的男、女人口。

2. 统计表

通过分组、计算、汇总，原始统计资料就整理成了统计表。统计表是比较、分析、研究经济现象数量特征的基础。只反映某一时点或某一时期的经济现象的统计表叫作次数分配表或次数分配数列。而按时间先后顺序把次数分配数列排列起来的统计表则是时间分配数列，或称动态数列。西方各国统计资料中绝大多数统计表是动态数列。经济动态数列描述了经济现象随时间推移而发生的变化。

统计表主要由以下几个部分组成：

① 总标题。总标题表明统计资料反映的主体对象，或