

管理统计学与不确定性现象和数据分析之间存在着密切的关系。管理统计学为应对不确定性现象提供了有效的策略与方法,同时也为数据分析提供了理论基础与方法支持。在管理决策过程中,充分用好管理统计学和数据分析,将有助于提升决策的科学性。

我们强调管理统计学的重要作用,不仅是要通过数据分析为决策赋能,更是社会责任的体现。学习统计学,要秉持严谨求实的科学态度,践行党的二十大报告中的诚信精神,确保数据真实可靠,为国家发展和社会进步提供坚实支撑。随着数字化时代的到来,数据及数据分析的价值将越来越凸显,将成为推动社会进步和创新的重要力量。

1.1 统计学的性质

统计学是研究如何收集、整理、分析和解释涉及社会、经济、管理问题的数据,并对研究对象进行统计推断的一门科学。

人们往往将统计学误解为数据资料的收集以及对这些数据作一些简单的均值、百分比等运算,或用图和表等形式加以表示。其实这些仅是统计学工作的一小部分。统计学还包括假设检验、回归分析、时间序列分析、方差分析等许多工作。

在自然界里,在生产、管理实践中,人们观察到的现象可以大体归结为两种类型:一种是确定性现象;另一种是不确定现象,也称随机现象。

随机现象是指在相同条件下重复进行试验,每次结果未必相同的现象;或是知道它过去的状况,但未来的发展事前不能完全确定的现象。例如,某种新产品投放市场后消费者对其欢迎程度、对某个房地产项目投资后的获利情况、顾客在某服务系统接受服务的等待时间等,都是不确定现象,即随机现象。

实际上人们发现,所谓不确定,只是对一次或少数几次观察或实践,当在相同条件下进行大量观察时,随机现象都呈现某种规律,这被称为随机现象的统计规律性。例如,通过大量的统计发现,某些服务系统对顾客的服务时间是服从指数分布规律的,人的身高是服从正态分布规律的,等等。从而人们便可以根据所发现的规律,对所研究的问题进行科学决策。统计学就是研究随机现象的统计规律性,所以统计学对于科学决策有着重要的意义。

一般来说,统计学的研究对象具有以下特点:

- (1) 随机性。随机性有两个解释,一个是事件发生的结果不确定;另一个是不同的个体有差异。
- (2) 群体性。研究对象应包含不止一个个体,换句话说,单个个体不需要统计分析。
- (3) 数量性。事件的发生情况应由数量表示,以便于进行统计分析。这在大部分情

况下是可以做到的。如果事件的发生是以定性的方式表述的,则应该将其转换为数量表述。例如,某产品的销售可能有好、中、差三种情况,可以将三种情况分别记为 1、2、3;再例如,对某种突发事件的发生记为 1,不发生记为 0;等等。

1.2 统计学在管理中的应用

统计学的应用领域非常广,几乎所有的研究领域都要用到统计方法。例如,经济学、管理学、气象学、医学、遗传学、地质学、历史学、语言学、教育学、金融学等。当统计学应用于管理学领域时,就被称作管理统计学。本书主要将从管理统计学的角度进行论述。

古语说得好:“知己知彼,百战不殆。”企业为了生存、发展,除了要了解自身特点、内在优劣势之外,还必须时刻掌握市场、竞争对手、经销商、顾客等多方面的情况,收集那些能反映企业目前状况的相关数据,并且最重要的是将这些数据转化为对企业有用的信息,为企业科学决策提供有效的依据。统计学在企业管理中所扮演的角色就是这样一个将原始数据转化为有价值信息的过程。可见,企业要真正做到知己知彼,百战不殆,离不开科学有效的管理统计。

随着大数据的发展,数据驱动管理决策的重要性愈发凸显,统计学在管理中的应用日益广泛深入,主要包括以下几个方面。

(1) 库存确定。公司经理会根据对原材料的需求和库存状态的分析,确定原材料的进货量(这里需要用到统计方法,分析需求量变化的分布规律)。

(2) 生产控制。生产经理根据对样本产品的质量检验情况,确定是否对生产过程进行调整(这里需要用到统计学中的抽样检验方法,选取样本产品)。

(3) 市场分析。市场部经理根据对某种新产品在样本消费者中试销情况的调查结果,确定该产品可能的销售量(这里需要用到统计学中的统计推断方法,进行销售量预测)。

(4) 客户管理。市场部经理通过调查,对公司客户的类别进行细分,根据不同类别客户的特点,选择不同的服务策略(这里需要用到统计学中的聚类分析方法)。

(5) 风险投资。投资部经理通过对各种项目投资回报率及其风险的分析,并根据企业所处的情况,确定是否对某项目进行风险投资(这里需要用到统计学中的参数估计等方法)。

(6) 审计查账。检查一个大公司账目的审计员,会运用统计方法抽样检查公司的账目,然后根据抽样检查的结果,确定公司的账目是否有问题(这里需要用到统计学中的抽样检验方法)。

(7) 考核标准。例如,公司总部希望了解设立在全国多个城市中的分公司的经营业绩是否与所在城市的 GDP 有关。如果有关,公司总部准备以分公司所在城市的 GDP 作为考核标准,对各个分公司进行绩效考核(这里需要用到统计学中的回归分析)。

(8) 品牌服务。通过对社交信息数据、客户交互数据等进行大数据计算,帮助企业进行品牌信息的水平化设计和碎片化扩散,并通过大数据信息交叉验证技术、分析数据内容之间的关联度等面向社会化用户开展精细化服务。

(9) 商业智能。通过对数据进行抽取、转换和装载(即 ETL 过程),合并到数据仓库里,利用合适的数据挖掘技术、查询和分析模型、联机处理(OLAP)工具等进行分析 and 处理,将数据转化为知识,为管理决策服务。

1.3 统计学方法的分类

统计学的方法很多,一般可以分为两类:描述统计学方法和推断统计学方法。

1.3.1 描述统计学

描述统计学方法主要包括利用获得的数据,绘制统计图(直方图、条形图、饼图、雷达图等),并计算一些数字特征值(均值、方差、中位数、调和均值、偏态系数等)。人们根据这些统计图可以比较直观地对研究对象的规律有一个大体的粗略认识,而通过数字特征值则可以从数据的集中趋势、分散程度、偏态情况等方面对研究对象的存在和演变规律有一个基本的了解。描述统计学方法相对比较简单。

1.3.2 推断统计学

推断统计学方法主要包括利用获得的样本数据,进行区间估计、假设检验、回归分析、方差分析、时间序列分析等。人们根据这些样本数据的分析结果,对研究对象的总体进行统计推断,包括推断研究对象的分布规律、推断不同因素间的相关性、确定多个因素间的统计关系、判断多个总体间的差异性等。推断统计学包括的内容比较多,相对复杂一些,也是本书的重点内容。

1.4 统计学的基本概念

1.4.1 总体

统计学将构成研究对象全部元素的集合称为总体,而将组成总体的每个元素称为个体。

例如,在研究某种产品的寿命时,该批产品的全体就组成了总体,其中的每个产品就是个体。在调查某公司员工对工作的满意程度时,该公司中的全体员工组成了总体,公司中的每个员工是个体。在统计学的分析中,我们只是关心每个产品的寿命指标,或每个员工的满意程度指标,而不是产品本身,或员工本身。所以我们把这一批产品的寿命指标集合,或这个公司全体员工的满意度指标集合作为总体。即总体是由一些实数构成的集合。我们不能把总体看成一群人或一些物品的集合。一般情况下,我们将总体和数量指标所有可能的取值组成的集合等同起来。

总体分为有限总体和无限总体。有限总体是指总体的范围能够明确确定,而且元素的个数有限。例如,一批待检验的产品就是有限总体,被调查的公司员工也是有限总体。无限总体是指总体所包含的元素是无限的。例如,在科学试验中,每一个试验数据可以看

做总体的一个元素,而试验可以无限地进行下去,因此由这种试验数据构成的总体就是一个无限总体。在统计学的研究中,对于有限总体,为保证每次抽样试验是独立的,需要根据具体情况对抽样试验进行相应的设计,这些将在后面的抽样方法中介绍。

1.4.2 样本

为了对总体的分布规律进行各种研究,需要对总体进行抽样观察。通过多次抽样观察可以得到总体指标 X 的一组数值 (x_1, x_2, \dots, x_n) , 其中每个 x_i 是一次抽样观察的结果。 (x_1, x_2, \dots, x_n) 称为容量为 n 的一个样本,也称样本观察值。由于样本是用来对总体的分布特征进行推断的,因而从总体中抽样进行观察时必须是随机的,即每个个体具有同等的可能性被抽到。因为只有这样才能经过多次观察比较全面地了解总体。

例 1.1 现准备调查某批产品的质量,假设质量分为合格、不合格,分别用数字 1,0 表示。用 X 记录这数字, X 就是要研究的指标。

该批产品共有 N 件,使用放回抽取方法,从这些产品中任意抽取 $n=30$ 件进行检查。即每抽一件产品,经检查,记录下质量情况 X 后,放回去,然后再抽下一件。从抽得的 30 件产品检查得到的质量指标 X 的值 $(x_1, x_2, \dots, x_{30})$ 构成一个样本,样本的容量 $n=30$ 。

1.4.3 总体参数和统计量

总体参数表示总体分布的某些特征,如分布位置、分布离散程度等。一般来说,估计总体参数的统计量是样本的函数。

常用的统计量有:

- (1) 表示位置的统计量——样本均值;
- (2) 表示离散程度的统计量——样本标准差、样本方差。

1.5 统计学的发展

统计学的发展经历了如下几个阶段:古典统计学、近代统计学、现代统计学。

1.5.1 古典统计学时代

这个时代大致是从 17 世纪中叶至 19 世纪初,其代表学派是“政治算术派”和“国势学派”。

“政治算术派”创始人是英国的威廉·配第,他于 1690 年出版了《政治算术》一书,用数字来描述英国的经济情况;另一代表人物是英国的约翰·格朗特,其代表作是《关于死亡的自然和政治观察》。

“国势学派”又称“记述学派”或“国情学派”,其创始人是德国的海尔曼·康令,这一学派最早提出了“统计学”的名词。

1.5.2 近代统计学时代

这个时代大致是从 18 世纪末到 19 世纪末。著名的大数法则、最小二乘法、相关与回

归分析、指数分析法、时间序列分析法以及正态分布等理论都是这个时期建立和发展起来的。代表学派主要有数理统计学派和社会统计学派。

数理统计学派产生于19世纪中叶,创始人是比利时学者阿道夫·凯特勒。凯特勒对统计学的主要贡献是将自然科学的研究方法引入社会现象的研究中,认为统计学就是数理统计学。他创立了大数法则,发展了大量观察法并为数理统计学的发展奠定了基础。

社会统计学派产生于19世纪末期,首创者是德国人克尼斯,主要代表人物是梅尔、恩格尔等人,他们认为统计学的研究对象是社会现象,研究方法是大量观察法,提出统计学是一门实质性的社会科学。因此人们试图通过社会调查,收集、整理、分析资料以揭示社会现象和问题,并提出解决问题的具体办法。如恩格尔对比利时工人家庭进行了持续的调查,通过不断地观察总结,终于在其《比利时工人家庭的生活费》一文中,提出了著名的“恩格尔法则”,即家庭收入越多,则饮食消费支出占家庭收入的百分比越少;家庭收入越少,则饮食消费支出占家庭收入的百分比越大,并引申出“恩格尔系数”,以此作为衡量生活水平标准。这一指标至今仍有很大实用价值。

1.5.3 现代统计学时代

1900年,英国统计学家卡尔·皮尔生提出了 χ^2 检验法。1908年,“学生”(William Seely Gosset,戈塞特的笔名)发表 t 分布的论文,创立了小样本代替大样本理论,费雪又对小样本理论进一步研究,发展为实验设计理论,标志着现代统计学的开端。1930年,尼曼与小皮尔生共同对假设检验理论进行了系统的研究,创立了“尼曼-皮尔生”理论,同时尼曼又创立了区间估计理论。美国统计学家瓦尔德把统计学中的估计和假设理论予以归纳,创立了“决策理论”。这些研究和发现大大充实了现代统计学的内容。

近几十年来,计算机技术以及信息论等现代科学技术,尤其是大数据应用的迅猛发展,给统计学带来了巨大的影响,新的领域层出不穷,例如多元统计分析、现代时间序列分析、贝叶斯统计、非参数统计等。各种统计软件如SPSS、SAS、MATLAB、R语言在趋于完善的同时,也加大了统计分析的功能,如数据挖掘功能。

大数据时代,对统计学的发展既是机遇又是挑战,机遇在于大数据的分析主要建立在对数据进行处理、分析的基础上,随着大数据技术的不断涌现和发展,数据获取、数据存储和数据“可视化”等技术手段日益丰富,推动传统统计学更好发展;而挑战在于,许多传统统计学的统计方法应用到大数据上,出现许多不适应现象,如巨大计算量和存储量往往使其难以承受。因此,需要我们进一步对统计学进行发展与创新。

本章小结

管理统计学是研究如何收集、整理、分析和解释涉及社会、经济、管理问题的数据,并对研究对象进行统计推断的一门科学。近年来统计学在许多管理问题如生产控制管理、客户关系管理、市场分析、投资风险管理等领域中得到了广泛的应用。

统计学由描述统计学和推断统计学组成,描述统计学内容将在本书的第3章、第4章

中讨论,后续章节主要围绕推断统计学展开。样本和总体、总体参数和统计量是统计学中最基本的概念,要学好本书的内容,必须充分理解这些基本概念。

关键术语

统计学(statistics)

总体(population)

描述统计学(descriptive statistics)

样本(sample)

推断统计学(inferential statistics)

总体参数(population parameter)

管理统计学(management statistics)

统计量(statistic)

即练即测



习题

1. 简述统计学的基本概念。
2. 怎样理解描述统计学和推断统计学在探索数量规律性方面的作用?
3. 简述统计学与企业管理的关系。
4. 统计学的发展经历了哪些阶段? 说明每个阶段的特点。
5. 大数据时代,对管理统计学的发展有哪些机遇和挑战?