

数学核心素养在初中阶段的主要表现之数据观念

鲍建生 章建跃

【摘要】统计与概率研究的是现实世界中的不确定现象,其研究对象、基础、方法等与确定性数学有较大区别.在明确统计的定义、课程标准在统计与概率领域内容与要求的基础上,对初中阶段“数据观念”的主要表现进行了细化分析,讨论了如何帮助学生逐步理解随机现象的特点,感悟统计的基本思想方法,感悟统计与数学的区别等,由此给出一些发展学生数据观念的教学策略,以帮助教师解决统计与概率教学中的困惑.

【关键词】数据观念;表现行为;教学策略

一、概述

统计进入中小学数学课程是国际中小学数学课程现代化的一个必然趋势,也反映了现代社会对数据意识及统计方法的普遍需求.

美国统计协会给出的统计定义是:统计是从数据中形成知识,并且测量、控制、表达不确定性的科学.该协会和美国数学教师协会(NCTM)合作研究了多个统计活动框架后,整合了其中的相似部分,并结合NCTM出版的数学课程标准,发布了《统计学评估和教育指导纲要:中小学课程框架》(以下简称“GAISE”),其中将中小学阶段的统计活动划分为四个阶段(如图1).

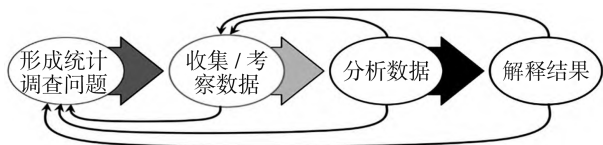


图1 美国统计协会界定的中小学统计活动

(1)形成统计调查问题:包括明确现实问题,提出能用数据回答的统计问题.

(2)收集/考察数据:设计有效收集适当数据的计划,实施数据收集的计划.

(3)分析数据:选择适当的统计图表或数值方法,使用选择的方法进行数据分析.

(4)解释结果:说明分析过程,结合原始问题与数据进行解释.

因此,数据是统计学的研究对象,其核心是数据分析,问题解决是统计活动的基本形式,而数据观念是统计思维的基础.史宁中指出,统计学的研究依赖于对数据的感悟,甚至是对一堆看似杂乱无章的数据的感悟.通过对数据的归纳整理、分析判断可以发现

其中隐藏的规律.因为可以用各种方法对数据进行归纳整理、分析判断,所以得到的结论也可能是不同的.而且,我们很难说哪一种方法是对的,哪一种方法是错的,我们只能说能够更客观地反映实际背景的方法要更好一些.可见,统计与数学在思维方式上有很大不同.

《义务教育数学课程标准(2022年版)》(以下简称《标准》)区分了小学与初中两个阶段在概率统计方面的教学要求,其中小学阶段侧重于数据意识,旨在加强学生对数据的意义的感悟.与小学阶段相比,初中阶段的概率统计课程增加了下列内容.

(1)增加了总体与样本的概念.学生将学习简单地获取数据的抽样方法,为以后的推断统计做准备.

(2)引入了更多统计量与统计图表.一方面,可以更全面地描述、分析数据;另一方面,有助于通过比较更好地理解不同统计量与统计图表的意义及适用场合.

(3)增加了统计活动的要求,使学习从调查问题与随机抽样开始,经历相对完整的用统计方法解决实际问题的过程,感悟统计的思维方式与实际意义.

(4)在概率方面,引入了定量刻画随机事件发生可能性大小的方法,使学生初步感悟概率与统计的联系.通过初中阶段概率统计课程的学习,希望帮助学生形成相对稳定的数据观念,感悟处理随机现象的思维方式.

从具体应用的角度来看,统计学包括三个部分:描述统计、推断统计和实验设计.其中,描述统计是指对所收集的大量数据资料进行整理、概括,寻找数据的分布特征,用以反映研究对象的内容和实质的统计方法;推断统计是指根据来自样本的数据推断总体的

性质,并标明可能发生的误差,以对随机现象做出估计、推断的统计方法,推断统计可根据已知材料,去估计、推测未知的可能性大小;实验设计是指研究者为揭示实际问题中自变量与因变量的关系,验证假设之前所制订的实验计划,内容包括研究步骤的制订、抽样、实验变量及实验条件的控制、对结果的统计处理方法等。在初中数学课程中,仍然以描述统计为主,但可以通过样本与总体的关系,初步感悟一点推断统计的意义,同时可以介绍一些简单的统计实验案例。

《标准》对初中阶段“数据观念”的说明如下:“数据观念主要是指对数据的意义和随机性有比较清晰的认识。知道数据蕴含着信息,需要根据问题的背景和所要研究的问题确定数据收集、整理和分析的方法;知道可以用定量的方法描述随机现象的变化趋势及随机事件发生的可能性大小。形成数据观念有助于理解和表达生活中随机现象发生的规律,感知大数据时代数据分析的重要性,养成重证据、讲道理的科学态度。”

下面以《标准》为依据,对初中阶段数据观念的具体表现指标及教学做进一步的解读。

二、数据观念的主要表现形式

在中小学课程体系中,初中阶段的课程介于小学与高中之间,具有承前启后的作用。

从“数据意识”到“数据观念”再到“数据分析”,是一个从定性描述逐步过渡到定量分析的过程,通过这个过程逐步帮助学生对数据及其意义加深认识、拓宽视野,形成稳定的统计素养。其中,小学阶段关注的是儿童通过活动对随机现象的感悟,本质上属于描述统计,高中阶段要求具备一定的运用模型解决随机现象问题的能力,初中阶段则介于两者之间。因此,需要关注小学阶段与初中阶段、初中阶段与高中阶段之间的衔接。

下面结合初中阶段的课程内容与教学要求,进一步细化初中阶段数据观念的主要表现。

(一)理解数据和抽样的意义

数据是统计的基本研究对象,抽样是用样本数据描述或推断总体的基本研究途径。理解数据和抽样意义的具体表现有如下几个方面。

(1)感悟数据的意义。知道数据与数量的区别,数据是统计的语言,统计是靠数据说话的;知道数据一般都具有实际的意义,是用来分析某种随机现象和解决问题的;知道数据具有多样性,不管是哪种数据,数据的来源与真实性是最重要的。

(2)体验数据收集的过程。知道不同的问题需要收集不同的数据,同样的问题也可以收集不同的数据;知道数据是抽样、测量与实验的结果,而不是推理

或猜想的结果,抽样、测量与实验的方法会影响到数据的质量,进而影响对数据的分析。

(3)感悟样本数据的意义和随机性,体会样本与总体的关系。知道数据的随机性是区分描述性统计与推断性统计的标志,知道现实世界中的许多随机现象是不能直接分析的,需要通过随机抽样,利用样本数据的统计特征来描述和推断总体的统计特征。例如,知道可以用样本平均数估计总体均值,但由于样本数据具有随机性,这种估计也具有随机性,样本数据的质量会影响估计的可靠性。

(4)体会抽样的必要性,能够通过实例了解简单的随机抽样过程。了解抽样过程涉及以下四个方面。

①抽样的目的:获得有关事件的可靠信息,这些信息对总体而言是不可及的,只能间接获得。

②抽样的方法:从总体中随机地产生一个子集,这个子集的信息与总体类似,并且可以直接获得。

③总体与样本的关系:在一定程度上,样本可以代表总体。

④统计推断:将从样本获得的信息推广到总体。

(5)感悟数据的变异性。知道导致数据变异的原因是很多的,如不同的抽样会得到不同的数据,相同的抽样也会因为测量误差而导致数据不同,但只要有足够的数据就可能从中发现规律,而且数据越多,所发现的规律就越可靠。因此,基于数据的推理是一种归纳规律的过程。

(6)进一步感悟统计活动的随机性。知道抽样是一个随机过程,用样本数据来描述和推断总体特征是一种估计过程,具有不确定性。因此,统计活动中,主要依靠的是归纳推理。

在一些国家的中小学统计课程中,特别重视对各种数据的认识,要求学生亲身经历数据收集的完整过程,其中包括现场调查、测量与简单的实验观察。

(二)初步掌握整理、描述与分析数据的方法,理解统计量的意义

虽然数据都蕴含着信息,但这些信息常常不是显而易见的,需要运用一定的统计工具去整理、描述与分析,并用一些可视化的方式把蕴含在数据中的信息挖掘和表达出来。具体表现包括如下几个方面。

(1)感悟数据整理的意义。会对数据进行分类,初步了解数据分类与分类标准的关系;能够依据数据的特点进行分组或者排序;对异常数据做出初步的解释或判断,知道数据的来源与真实性对解决问题的重要性。

(2)知道可以用图表的方式对收集到的数据进行简单的整理、分类、分组,以及频数与频率的统计等,通过图表对数据特征及其规律进行直观感悟,逐

步培养对数据的统计直觉。

(3)理解平均数、中位数、众数的意义与区别,能计算中位数、众数、加权平均数,知道它们是对数据集中趋势的描述。知道平均数可以反映数据的“整体水平”,但容易受极端数值的影响;中位数的一个优势就是不受极端值的影响,但忽略了整体性,不能反映数据和中位数的差异情况;平均数与中位数适合于定量数据,而对于一些定性数据、类别数据或定类数据,只能用众数作为集中趋势的统计量。

(4)体会刻画数据离散程度的意义,知道极差、离差平方和与方差都是描述数据离散程度的统计量。会计算一组简单数据的离差平方和、方差。经历对数据进行分类的活动,知道对数据按照组内离差平方和最小的原则进行分类的方法。

(5)感悟统计图表的意义,解释数据中蕴含的信息。能画条形统计图、折线统计图、扇形统计图与频数直方图,知道直方图与条形统计图的区别;能读懂报纸、电视、互联网等媒体中的简单统计图表;能够依据数据的特征合理选择统计图表,并用统计图表合理表示和分析数据,通过统计图表感受随机现象的变化趋势,解释所表达的意义。

(6)初步感悟分布的意义,知道分布是描述和预测随机现象规律的重要工具。知道平均数与方差、中位数与四分位数都可以不同程度地描述分布;频数直方图与箱线图(如图2所示的五数综合)也能够一定程度上表示分布。通过对分布概念的理解,可以使发现某种随机现象的模式及数据的整体特征。

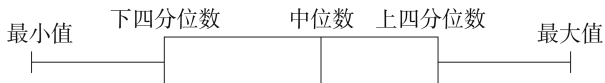


图2 用箱线图描述数据的分布

(7)初步了解大数据的特点。知道在数字化时代,每天都会产生海量数据,这些数据常常以多种方式被记录、收集与分析,如手机中的照片、行动轨迹、微信等;初步了解大数据的一些处理方法,如加权平均、分布计算等。

统计量在一定程度上反映了运用数学工具解决统计问题的特点,即研究与揭示数据的数量特征与规律。

(三)理解解决统计问题的特点、过程与意义

中小学统计课程的一个主要目的是帮助学生初步感悟真实背景下的解决统计问题的特点、过程与意义,在问题情境中去理解数据、统计量的意义和数据处理的基本方法。具体要求包括如下几个方面。

(1)形成问题。知道哪些问题可以用统计方法解决,能够在真实的情境中明确研究目标,提出能用数据回答的统计问题。

(2)收集数据。根据研究问题设计有效收集适当数据的计划,实施计划、收集数据;初步感悟统计问题解决与确定性数学问题解决的区别,即统计问题解决的基本思想方法是从数据出发,通过数据分析提出可供参考的结论,而确定性数学问题的解决,一般是从定义出发,得到确定的结论。

(3)分析数据。选择适当的统计图表或数值方法,使用选择的方法进行数据分析。根据史宁中的观点,利用数据分析不确定现象的目的是“在一堆看似杂乱无章的数据中提炼信息、寻找规律,抓住主要因素,对随机现象进行描述、预测,形成结论、决策、知识”。

(4)解释结果。通过明晰分析过程,结合原始问题与数据对统计分析的结果进行解释,并根据实际情境与需求做出预测或决策。知道在统计问题解决过程中,样本、数据等都具有变异性;同一个问题往往可以采取不同的统计量与统计方法;利用统计方法得到的结论也不会是百分之百的肯定。

真实的统计活动一般都比较复杂,也比较费时。因此,在中小学的实际教学中,一方面,需要在日常教学中适当简化统计过程,聚焦相关的统计概念与方法,便于学生理解与掌握统计的基础知识与基本技能;另一方面,还需要借助典型的统计案例,使学生相对完整地经历解决统计问题的过程,感悟统计中的一般思想方法,积累统计领域的活动经验。

(四)初步感悟随机事件的概率的意义

著名的法国数学家和天文学家拉普拉斯曾经说过,概率论实际上只是将常识归结为计算,它使我们能够用理性的头脑精确地评价凭某种直觉感受到的、往往又不能解释清楚的见解……引人注意的是,概率论这门起源于对机会游戏进行思考的科学,早就应该成为人类知识中最重要的组成部分……生活中那些最重要的问题绝大部分其实只是概率论的问题。

数学历来被认为是确定性的科学,这就意味着,从同样的条件出发就应当得到同样的结论,如果得到的结论不一致,就会认为其中至少有一个结论是错误的。但在日常生活中,人们却会遇到大量的不确定性事件,也就是说,事先无法确定这样的事件是否一定会发生、会发生到什么程度。例如,明天下雨的事件,期末考试得到90分以上的学生,彩票中奖的事件,等等,人们称这样的事件为随机事件。

小学阶段的概率学习主要是定性描述随机事件的可能性,初中阶段则需要对随机事件及其可能性的大小进行初步的定量刻画。具体表现包括如下几个方面。

(1)初步感知随机现象。知道在现实世界中随机

现象是普遍存在的,能够列举出各种随机现象的实际例子,通过具体实例感悟随机现象的客观性。

(2)初步感悟概率的意义.知道一个随机事件可能发生,也可能不发生;概率是刻画随机事件发生可能性大小的度量,反映了随机现象的数量规律;概率有助于对随机事件的预测,从而采取合理的行为或决策.例如,在知道本地今天的降水概率为85%时,出门就应做好防雨的准备。

(3)通过具体实例初步感悟概率的基本特征.知道概率大的随机事件也可能不发生,小概率事件也可能发生;影响随机事件概率的因素有很多,如高水平射击运动员中靶的概率就比一般人大,投掷一枚质地均匀的硬币,那么正面朝上和反面朝上的概率应该是一样的。

(4)能够描述简单随机事件的特征,即可能的结果有限,而且每一个可能结果的发生概率相同;能通过列表、画树状图等方法列出简单随机事件所有可能的结果,并在此基础上了解简单随机事件概率的计算方法。

(5)知道随机事件在进行大量的重复试验时,其发生的频率具有稳定性.因此,可以通过大量重复试验,用频率估计概率,体会数据的随机性以及概率与统计的关系。

与小学阶段的概率课程相比,初中阶段虽然给出了量化的概率,但仍然属于概率的初步认识.其中的许多概念与方法需要到高中阶段进一步地界定.例如,通过高中阶段的样本空间概念可以帮助学生更好地理解随机事件及简单随机现象。

三、通过真实、有意义的统计活动,使学生经历统计中问题解决的过程

虽然数与式、方程与不等式、函数、图形与几何等中小学数学课程内容大都在17世纪之前就已成形,但它们仍然是整个数学的基础,中小学数学教学的一个重要目标是帮助学生形成初步的知识体系,从中领悟数学的思想方法。

相比之下,中小学统计课程的主要目标不是构建逻辑体系,而是通过真实的统计活动,帮助学生初步感悟数据的意义及统计问题解决的一般过程,从中感悟统计的思想方法。

因此,设计真实的、有意义的统计活动是统计教学的基本方式。

(一)对数据的理解是形成数据观念的基础

虽然概念和符号对于统计学也是重要的,但统计学的研究建立在数据的基础上,通过数据进行推断.因此,统计学关心的是如何有效地获取数据,如何保证数据的有效性和可靠性,关心如何进行数据的整理

和分析,如何从数据中挖掘有用的信息;如何通过信息构建数据产生背景的知识,如何对随机现象进行预测.正因为如此,许多学者称统计学为数据分析的科学和艺术。

在教学中,应该通过各种统计活动,帮助学生了解数据的意义、多样性与变异性。

第一,数据是否具有统计意义取决于对应的统计问题,不同的问题往往需要不同的数据.例如,学校要定制校服,那就需要收集学生的身高与体重两类数据,因为一般服装的规格尺寸是依据这两类数据确定的;但如果要考察学生的肥胖情况,那么独立考察身高与体重的数据就意义不大,还需要在这两个数据的基础上利用身体质量指数(BMI)计算公式得出新的数据.因此,要帮助学生理解数据的统计意义,就必须了解想要解决的统计问题.好的统计问题对学生来说应该是真实的、有意义的,而且还应该是学生熟悉的、有兴趣的。

第二,应该关注数据的来源与质量,了解数据收集的基本途径.例如,要获得所在班级学生的身高数据,至少有三种途径:(1)现场调查,让每名学生自报一个身高数值,获得调查数据;(2)查找资料,利用统计资料,从学生的档案中获得身高的数据;(3)实际测量每名学生的身高获得数据.可以让学生比较不同来源数据的优缺点,并根据所考察的统计问题选择合适的途径.虽然判断数据的好坏没有统一标准,主要看是否符合研究问题的需求,但必须保证数据的真实性,并尽可能减少测量误差。

第三,初步感悟数据的随机性.通过观察、测量或试验得到的数据都具有随机性,这种随机性至少表现在以下两个方面.一是每个数据的获得具有随机性.例如,每名学生的身高不可能有一个准确值,不同时间测量的结果会不一样,不同测量工具也会产生各种误差.二是与选取的样本有关.如果我们想利用班级部分学生的身高来分析全班学生身高的情况,那么得到的数据就会因选法的不同而改变.尽管数据都具有随机性,但当数据足够多时就可以发现其中的规律.例如,在测量一名学生的身高时,尽管每次测量结果都有误差,但测量次数足够多时,结果的误差就会显示出一定的规律性(满足正态分布),利用其中的规律性就可以获得“更好”的数据,如把所有测量结果的平均数作为想要分析的数据。

第四,感悟数据整理与表示的意义.知道一组杂乱无章的数据很难看出其中的规律,数据越多就越需要整理.数据整理的途径包括分类、分组、排序与图表表示,除了探究数据规律的目的外,还要关注一些异常的数据,初步分析造成数据异常的可能原因,以保

证数据的真实性。在用统计图表表示数据时,要初步养成规范作图的习惯,知道统计图表的优点是直观与综合。好的统计图表让人一目了然,容易发现其中的规律。

此外,还要帮助学生初步感悟各种统计量的意义。知道引入统计量的目的是描述、分析和预测数据的集中趋势、离散程度等分布特征,不同的统计量有不同的功能,所适用的场合也不一样。例如,要比较两个班级学生的身高情况,可以利用两组数据的平均值,也可以利用两组数据各自所在的范围(极差);而要在两个班级中挑选最高的人进行比较,则只需考察两组数据的最大值。又如,要分析两个班级学生中各有多少名学生的身高超过1.4米,可以用百分数比较,用平均数就没有意义。

(二) 通过真实的统计活动,感悟统计的思想方法

在初中阶段,学生已经学习了描述性统计的基本概念,也初步了解了抽样的意义。因此,可以比较完整地实施具有现实意义的统计活动,在统计问题的解决过程中使学生逐步感悟统计的思想方法,形成一定的统计直觉。

史宁中指出,统计素养包括以下三个方面。首先,养成通过数据来分析问题的习惯。其实质是通过事实来分析问题。当遇到问题时,应当去调查研究,应当去收集数据,在此基础上进行的推断才可能客观地反映实际背景。其次,建立随机的概念。有些事情可能发生,有些事情可能不发生,这在日常生活中是大量存在的。即便如此,只要我们掌握的信息多了,也能够合理地推断实际背景。最后,学习如何去判断事情的主要因素。统计学能够在一堆看似杂乱无章的数据中提炼信息、寻找规律,这就需要抓主要因素。

一个相对完整的统计活动,一般经历以下几个环节。

(1) 在现实背景下发现和提出统计问题。

问题是研究的起点,针对不同的问题有不同的研究方法。如果一个问题中具有明确的数量关系或规律,一般情况下就可以用方程、不等式、函数这样的数学模型来解决;但如果面对的是不确定的随机现象,就需要用统计方法解决。在统计教学中,教师应创设情境,让学生自己提出一些真实的、感兴趣的、有意义的统计问题。例如,初中生每天应该保持多少户外运动的时间?我上学时可以乘公交,也可以骑自行车,哪种交通方式更合理?我们班学生的身高可以达到全国青少年身高的平均水平吗?在一个池塘边挂着“水深危险,请勿游戏”的牌子,那么“水深”是什么意思?我们要如何知道池塘有多深呢?等等。学生自己

发现、提出统计问题的过程,有助于更好地理解随机现象和数据分析的意义。

(2) 根据研究问题设计调查方案、收集数据。

许多研究都表明,让学生亲自收集数据,而不是仅仅分析现成的数据,可以加深对数据及其收集过程、数据背景、抽样及随机性的理解。

史宁中认为,在统计研究中,首先遇到的问题是获取“好”的数据。所谓“好”的数据,是指那些能更加客观地反映实际背景的数据,而获取好数据则要依赖“好”方法。通常,根据数据来源的不同,获取数据的方法可分为两大类:一是通过调查收集数据;二是通过实验制造数据。中小学统计教学中涉及的主要是前者,称为抽样调查(而后者通常被称为实验设计)。抽样调查又可分为两种:一种是收集已存在的数据,称之为抽样,如市场的物价、学生的身高、企业的产值等;另一种是需要通过询问才能获取的,称之为调查,如人们日常消费的主要项目、中小學生喜欢的歌手等。

收集数据的方法可因问题而异,但要遵循如下两个基本原则。

第一,采用能获取“好”的数据的方法。为了获取好的数据,我们需要尽可能多地利用已有的先验知识。例如,希望知道学生的身高,先验知识是“少年儿童的身高在不同年龄之间差别很大”。因此,最好是按年龄分段,在各年龄段按比例抽取样本,这就是“分层抽样”。可以看到,统计方法的直观想法是很明显的。如果对实际背景一无所知,那么一定要随机抽取样本,这便是“随机抽样”。例如,希望了解初中生喜欢的歌手,因差别可能不大,就可以采取“随机抽样”的方法。当然,也可以按性别不同采用“分层抽样”的方法。

第二,采用简单的方法。能够基于上述两个原则的方法就是一个“好”方法。我们不要小看第二个原则,一个好的方法往往能够节省很多调查经费。这就是为什么咨询公司非常欢迎统计学家的原因。

(3) 组织与表示所收集的数据,并利用统计模型分析数据,描述或推断数据的分布特征。

对数据的组织就像是写小说。文字本身不构成故事,但作家可以通过组织文字来完成一个故事。故事是否精彩与文字组织有一定的关系,文字组织得不好,可以让人看不懂故事。数据也一样,组织得不好就不容易理解数据蕴含的信息。

数据的组织包括分类、分组、排序、画统计图表等。一些研究者认为,在统计素养中,最重要的是“图表意识”。在初中阶段的统计活动中,涉及的统计图表主要包括条形统计图、折线统计图、扇形统计图、频

数分布直方图与箱线图。制作图表的主要目的包括:表示数据的特征与趋势,产生、确定或推翻某个假设,显示数据的结构,便于他人理解。“图表意识”不仅是指学生会正确地构建统计图表,更重要的是会根据问题的实际情境选择恰当的统计图表。因此,图表意识是数据观念的重要组成部分。

数据的分析主要包括:在数据集或统计图表内比较数据,在数据集或多个统计图表间比较数据,对给定的数据和统计图表进行推断等。初中阶段的数据分析一般是选择适当的统计量描述数据的分布特征,如数据的集中趋势、离散程度等,用以描述(推断)总体的分布特征,由此获得调查分析的结果。

(4) 获得统计结论,做出预测或决策,并结合原始问题与数据进行解释。

学生通过统计推断的方式得出结论,并结合情境对结论进行解释和说明,针对最初的问题做出推断或者决策。

通过经历相对完整的统计活动,可以使学生更好地理解统计的意义与思维特点,形成正确的数据观念。

四、帮助学生逐步理解随机现象的特点,感悟统计与数学的区别

统计作为一级学科,其基本研究对象及思想方法与确定性数学有较大差异。但由于中小学课程中将统计内容纳入数学课程,作为数学课程的一部分,而绝大多数数学教师缺乏应用性统计的学习与教学经验。因此,如何讲出“具有统计味道”的统计课是目前中小学统计教学的一个挑战。应对这一挑战的基础在于理解统计与数学的联系与区别。

爱因斯坦曾经说过,数学定律不能百分之百地用于现实世界。能用数学定律百分之百确实描述的,就不是现实生活。而统计中的数据通常都来自现实世界,具有变异性,所以统计结论不是绝对的。

下面,我们从几个方面初步讨论统计、概率与确定性数学的区别。

(一) 历史根源不同

数学的发生、发展是人类发展史的一个重要组成部分,古希腊时代的欧几里得《原本》中就已经构建了相对完整的演绎系统。相比之下,统计与概率的发展则要晚得多。

统计问题的产生源自对不确定性现象的估计,早期的研究者都是具有很强数学功底的其他学科领域的专家,其目的是运用具有确定性的数学工具去研究不确定性现象。在研究这些具有不确定性的现象时,数学方法可以给人一定的确定性。在统计学的早期发展过程中,数学起到了决定性的作用。今天,统计学的

方法也成了数学研究的重要工具。

概率问题的产生也源自现实世界的活动,但与统计不同的是,早期的研究者一般是数学家,其目的是把随机现象作为数学的研究对象,从而构建处理随机现象的确定性模型。对可能性大小进行量化的想法自古有之,但直到16—17世纪才初步形成。到16世纪时,意大利数学家吉罗拉莫·卡尔达诺(Girolamo Cardano, 1501—1576)讨论过赌博中的骰子问题,开始出现量化的思想;17世纪早期,伽利略也研究过骰子点数的计算。但真正把概率作为一种度量研究,并使概率论成为一个数学分支的应当归功于帕斯卡与费马。他们曾经围绕下面两个经典问题进行了一系列的书信讨论:(1) 骰子问题。一个玩家如果在投掷8次骰子中出现一个6点,那么他可以拿走桌上的赌金。现在假设他已经投掷了3次,都没有一次是6点,庄家想让他放弃第4次投掷的机会,应该给他多少赌金才合理?(2) 点数问题。两名水平相当的玩家进行赌博比赛,他们事先约定一个点数,先达到这个点数为胜,可以拿走全部赌金。在进行若干轮后,赌局因故中断,应该如何分配赌金才合理?帕斯卡与费马围绕上述问题的讨论,引发了数学界对随机现象的普遍关注,但直到柯尔莫哥洛夫构建了概率论的公理系统,才真正确立了概率论的数学基础。

(二) 立论基础不同

我国著名统计学家陈希孺院士在《数理统计学简史》的序中这样写道,统计学不仅仅是一种方法或技术,还含有世界观的成分——它是看待世界上万事万物的一种方法,我们常讲某事从统计观点看如何如何,指的就是这个意思。但统计思想也有一个发展过程。因此,统计思想(或观点)的养成,不单需要学习一些具体的知识,还要能够从发展的眼光,把这些知识连缀成一个有机的、清晰的图景,获得一种历史的厚重感。这里,陈院士强调的是“统计思想的养成”,是人的全面发展中的“历史厚重感”。

史宁中认为,数学建立在概念和符号基础上;为了合理地解释抽象出来的东西,或者为了使数学研究的基础脱离人的经验,人们构建了基于自然数公理体系在内的实数理论,最终归结到集合论公理体系(反过来说,概念符号和公理体系的基础还是人的经验)。统计学建立在数据的基础上:如何有效获取收集数据?如何保证数据有效和可靠?如何整理、分析数据?如何挖掘数据中的有用信息?如何预测随机现象(反过来说,有效数据的基础还是数据)?

由于研究的对象都具有不确定性,因此统计与概率之间具有天然的联系。统计研究对象的真实主体是数字的分布,这个分布可以通过参数来描述。将数字

的分布看作是概率理论的元素,这样就可以建立参数估计量的最优化标准,然后去解决用数据描述分布时遇到的数学问题。

可以看到,统计与数学之间在基本思想方法上既有所区别,又有联系。随着大数据时代的到来,统计结果也在一定程度上具备了确定性,并产生具有普遍意义的知识。数学与统计之间将有越来越多的联系。

(三)推理方法不同

在统计学发展早期,研究者很少向公众展示或解释统计分析的过程,他们所做的是论述自己的结论,并发表能证明结论真实性的数据。例如,格雷戈尔·约翰·孟德尔(Gregor Johann Mendel, 1822—1884)在其著名的豌豆实验中,从没有展示他全部豌豆培育实验的结果,他叙述了他的系列实验,然后写道:两组系列实验的前10个数据可以用来说明……在20世纪40年代,费歇尔检验了孟德尔用来说明结论的数据,发现这些数据过分完美,以至于失真,它们并没有表现出应该具有的随机程度。

在数学中,推理过程依据的是逻辑规则,在确定前提的基础上,可以运用逻辑的方法一步一步地得出结论,推理过程的每一步都必须清晰明了,确凿无疑。也就是说,数学结论是推理过程的一个自然结果,是不可分割的。

史宁中认为,数学证明基于公理和假设,公理体系最初虽然来源于人们的经验和直觉,但最终还是脱离了现实背景(反过来说,公理体系最终虽然脱离了现实背景,其来源还是经验直觉);数学推理的前提是基本原则,包括同一律、矛盾律和排中律;数学证明的过程依赖演绎推理,得到的结论是必然的(验证结论)。统计推理的基础是数据背景;统计推理的前提是历史经验;统计推理的过程依赖的是归纳推理,得到的结论是或然的(发现结论)。

统计与概率的研究对象虽然都是不确定现象,但两者在方法上是有区别的。在统计活动中,某个随机过程的概率模型是未知的,但是通过这个过程可以得到一系列观测值,我们希望依据这些观测值来推断出这个随机过程的分布特征,并对所得到的结论的可靠性进行判断;而在概率中,虽然研究对象也是不确定现象,但随机过程的概率模型通常是已知的,该模型的不确定性由相应的概率分布来描述;概率论要回答的问题是这个随机过程产生某个结果的可能性有多大,也即通过已知概率模型去精确计算各种结果的可能性。

(四)判断原则不同

史宁中认为,数学对结果的判断标准是对或错,数学来源于现实经验,但严格讲它是一门缜密的科学。统计学对结果的判断原则是好或坏,即使是同样

的数据,也允许人们根据自己对数据背景理解的不同提出不同的推断方法,给出不同的推断结果,所以统计学不仅是一门科学,也是一门艺术。

在统计问题解决中,研究者通常根据实际情境与需求自己确定一个标准,达成标准的方法不是唯一的,研究结论也不会百分之百符合标准,而是依据标准设置一个可以接受的程度(范围)。判断统计结论是否符合要求,除了考察其是否属于可接受范围外,还需要考虑简易性、变异性等多种与研究背景相关的因素。因此,统计结论不可能像传统数学题那样有“独一无二的正确答案”。

此外,在数学中使用抽象符号进行形式化表达的一个目的是尽可能不受生活经验的干扰,直觉是一种数学发现的基础,但不是数学论证的途径。但在统计中,一些研究者认为,对不确定性的理解都源自日常生活中的直觉。生活直觉可以通过真实的统计活动得到加强,逐步形成统计直觉,而统计直觉在统计学习与研究中都至关重要。因此,在统计与概率的教学中,应该从起步阶段就重视培养学生对不确定性的感悟。

参考文献:

- [1] American Statistical Association. Pre - K - 12 Guidelines for Assessment and Instruction in Statistics Education II (GAISE II): A Framework for Statistics and Data Science Education[R/OL]. [2022 - 10 - 08]. [https://www.amstat.org/education/guidelines-for-assessment-and-instruction-in-statistics-education-\(gaise\)-reports](https://www.amstat.org/education/guidelines-for-assessment-and-instruction-in-statistics-education-(gaise)-reports).
- [2] 史宁中,孔凡哲,秦德生,等. 中小学统计及其课程教学设计: 数学教育热点问题系列访谈之二[J]. 课程·教材·教法, 2005(6): 45 - 50.
- [3] 教育部基础教育课程教材专家工作委会. 《普通高中数学课程标准(2017年版2020年修订)》解读[M]. 北京: 高等教育出版社, 2020.
- [4] 史宁中. 数学基本思想18讲[M]. 北京: 北京师范大学出版社, 2016.
- [5] 陈希孺. 数理统计学简史[M]. 长沙: 湖南教育出版社, 2002.
- [6] 史宁中. 数形结合与数学模型: 高中数学教学中的核心问题[M]. 北京: 高等教育出版社, 2018.

【作者简介】 鲍建生, 华东师范大学数学科学学院; 章建跃, 人民教育出版社课程教材研究所。

【原文出处】 摘自《中国数学教育》: 初中版(沈阳), 2022. 11. 3 ~ 11, 21

原标题: 数学核心素养在初中阶段的主要表现之六: 数据观念
<https://www.rdfybk.com/>