

大数据在统计分析中的应用探究

高倩倩

对外经济贸易大学统计学院, 中国·北京 100029

【摘要】信息化技术应用规模的不断夸大,使得社会当中出现了越来越多的大数据,有效将其应用到统计分析当中,可有效提升统计分析工作的开展质量。基于此,本文通过对大数据的简单概述,进而介绍了大数据背景下的统计分析思想,之后以此为基础,阐述了大数据在统计分析中的重要性,探讨了大数据在统计分析中的具体应用,为更好地开展统计分析工作提供支持。

【关键词】大数据; 统计分析; 应用

Application of Big Data in Statistical Analysis

Gao Qianqian

School of Statistics, University of International Business and Economics Beijing 100029

[Abstract] With the continuous exaggeration of the application scale of information technology, there are more and more big data in the society. The effective application of big data into statistical analysis can effectively improve the quality of statistical analysis. Based on this, this paper introduces the idea of statistical analysis under the background of big data through a brief overview of big data, and then on this basis, expounds the importance of big data in statistical analysis, discusses the specific application of big data in statistical analysis, so as to provide support for better statistical analysis work.

[Keywords] Big data; Statistical analysis; application

引言

统计分析是现代社会当中重要的工作之一,通过统计分析,准确了解相关工作的开展情况,以此为决策的制定提供支持,为推动社会向着更加良好方向的发展提供支持。而想要开展高效的统计分析工作,则需要合理对大数据进行应用,通过大数据排除影响统计分析的各种因素,展示出统计分析的特点,增强统计分析结果准确性等,为相关决策的制定提供良好依据。

1 大数据概述

科学技术的迅猛发展,使得社会逐渐进入大数据时代。所谓的大数据,指的是一台传统计算机无法独立处理与分析的庞大数据源,有效对这些数据进行分析与应用,将会对各项工作的开展及社会发展提供巨大帮助。对于大数据来说,通常具备下述四个特征:(1)数量庞大。这是大数据最基本的特征,很早阶段,数据量可以用“PB”为单位进行表示,而到了当前,则是通过“ZB”为单位进行表示,与之前相比增长数数万倍,正是由于其数量较为庞大,使得传统计算机难以独立进行处理分析。(2)多样性。大数据当中,不仅涉及多个方面的信息,而且这些数据存在不同的类型,如图片、文字、音频等,具有不同的信息展示效果。(3)价值型。大数据的数量虽然庞大,但其中很多数据无任何应用价值,有价值数据的密度非常低,需要采用科学、合理的方式对大数据予以分析,才会使大数据在实际当中发挥出最大的作用^[1]。(4)速率快,不论是数据的传输,还是数据的处理与分析等,均具有较高的速度,以确保大数据的有效应用。

2 大数据背景下的统计分析思想

大数据背景下,若依然采用传统思维理念开展统计分析工作,则无法符合时代背景的需求,不利于现代社会的发展。所以,现代统计分析时,应形成全新的统计分析思想:(1)改

变“抽样统计”理念。通过样品的随机选取与分析,可以预估出一个事物基本情况,但预估出来的结果通常具有较大的偏差,可能与实际并不相符。而对于大数据来说,其本身就是样本,通过对该样本的分析,可得到更加准确的结果,有利于最终决策的制定。(2)无需过度关注数据精确性,而应注重数据纷繁复杂性。(3)无需刻意寻找事物间的因果关系,只要分析出事物内部各要素检的关联性即可。

3 大数据在统计分析中的重要性

3.1 排除统计分析影响因素

统计分析是现代社会发展中的重要工作之一,通过该项工作的开展,可使政府部门准确了解社会各个领域的发展情况,并以此为基础,制定出科学、合理的发展决策,为推动社会更好地发展提供支持,但需要注意的是,在传统统计分析工作当中,很容易受到各种因素的干扰,如人员因素、时间因素、空间因素的,导致统计分析结果并不是很准确,无法真实、准确展示出被调查事物的具体情况,不利于决策的制定。而应用大数据后,则可改变这一情况。大数据是人类生活与工作当中产生的,每次每刻都会产生各种各样的数据,不会受到时间、空间等方面等因素的限制,使得统计分析结果更加准确^[2]。

3.2 展示统计分析特点

以往很长一段时间内,相关人员开展统计分析工作时,主要是利用文字报告的方法进行传递,在信息传递之前,不仅需要投入大量人员与物资,而且还要花费较长的时间,导致统计分析效率较低,特别是遇到规模较为庞大的数据信息时,更是可能出现大量问题,降低统计分析结果的准确性。随着社会的不断发展,大数据规模得更加庞大,人工统计分析的错误率也在不断提升,逐渐被社会所淘汰。而应用大数据后,可利用其相匹配的大数据技术完成该项工作,可自动收集数据信息,快速对数据的分析,以此得到更加精确的结果,并以此为基础,通过文字、图片、

表格及其他各种形式体现出来, 直观性更强, 有利于人们对研究事物的了解。

3.3 增强统计分析结果准确性

统计分析工作的开展, 主要是通过对事物的分析, 找出其中的本质与联系, 以此为相关决策的制定提供支持。所以, 不论是哪一方面的统计分析, 均是以数据为基础完成的。在过去一段时间内, 统计分析工作质量差, 分析结果准确性不高, 导致这一情况的主要原因就是分析方法的不合理, 无法全面对整个事物加以分析。特别是进入大数据时代后, 由于数据量的迅猛提升, 加之数据类型更加丰富, 不仅大大增加了人员统计分析的工作量, 而且提升了有价值信息的隐秘性, 导致人员很难从庞大的数据群体中挖掘出那些真正有价值的信息, 从而降低统计分析结果的准确性^[3]。针对这一问题, 通过对数据收集、数据存取、数据处理、数据挖掘等相关技术的应用, 即可从根本上改变这一问题, 例如: 在大数据系统当中, 只需要录入相应的关键词, 即可在数秒时间内自动检索出与此相关的数据信息, 速度快, 且准确性高, 能够分析出更加精确的统计分析结果。

4 大数据在统计分析中的具体应用

4.1 国民经济统计分析

从上世纪 70 年代末期开始, 我国经济呈现出迅猛增长的态势, 以此推动了我国各个行业、领域的发展, 逐渐出现了越来越多的经济数据, 若仍然通过以往人工统计的方式予以分析, 将会导致人员工作量非常庞大, 降低工作效率, 增加问题发生率, 难以得到准确的国民经济分析结果, 不利于相关部门制定符合实际情况的发展决策。而应用大数据后, 则会有效改变这一情况。国民经济统计分析过程中, 可在互联网内, 自动提取与国民经济相关的数据, 并有计算机软件予以整理、分类与分析等, 以此分析出当前国民经济具体情况, 并以此为基础, 与上一年进行对比, 确定出本年度国民经济发展中存在的优势, 找出其中存在的不足与缺陷, 以此为相关部门决策的制定提供支持。

4.2 就业、失业率统计分析

现代居民生活水平评估时, 通常采用诸多评估指标, 就业率与失业率是其中较为重要的两个指标, 其中, 失业率呈现出高水平时, 表明很多居民无工作, 无法正常获得资金入手, 不仅影响居民生存, 而且还会干扰社会分工, 严重情况下, 甚至使经济发展瘫痪。而就业率呈现较高的水平时, 则说明社会较为稳定, 分工机制较为合理, 大部分居民均可正常获取收入, 有利于个人生活及社会乃至整个国家的正常发展^[4]。以往阶段, 对就业率与失业率统计分析时, 一般是以毕业生增长情况为基础, 结合市场对劳动力的需求来计算的, 一方面, 很容易泄露大量人员信息, 另一方面随着我国教育领域的快速发展, 每年毕业大学生数量明显增加, 依然采用传统方式予以分析, 则很难得到精确的结果。而这种情况下, 通过大数据技术, 则能够从多角度出发, 全面对社会居民失业与就业情况进行分析, 以计算出更加真实、准确的就业率与失业率, 以此为政府对人力市场的管理提供支持。

4.3 财务审计的统计分析

财务审计工作开展中, 也可对大数据进行应用, 具体包括: (1) 财务数据的收集与存储。由上述分析可知, 大数据存在诸多优势, 其中, 最主要的就是可以快速处理庞大的数据集, 并针对数据的类型, 将其放入到不同区域予以存储, 提升数据存储的质量。与此同时, 在审计过程中, 通过相关关键词的检索, 即可将有关数据提取出来, 并在压缩处理后, 快速传输给其他平

台, 提升了数据的安全性。(2) 提升数据分析速率。对财务数据分析时, 可利用系统内设置的算法与原理, 在较短的时间内, 分析出相应的结果, 有利于财务管理策略的制定。(3) 有利于问题溯源。财务审计时, 若发现问题, 可通过对大数据的溯源分析, 及时发现引发问题的主要原因, 以快速将问题解决。

4.4 医疗卫生的统计分析

医疗卫生是现代社会当中的重要组成部分, 可向居民提供相应的医疗卫生服务, 以确保人类身体保持健康。为了向人们提供高质量的医疗卫生服务, 医疗卫生领域也需要开展相关的统计分析工作。以往阶段, 医疗卫生统计分析时, 效率并不是很高, 即使在同一区域内, 由于缺少统一的信息平台, 难以将数据共享与传输, 影响整个医疗卫生事业的正常运营与发展, 一些医院资源明显不足, 而还有一些医院则存在资源浪费的情况, 出现严重的医疗资源浪费问题。这一情况下, 通过大数据技术的应用, 能够促进各医疗单位间的联系, 促进各方面的沟通, 以此为医疗资源的合理分配提供帮助。

4.5 人口普查统计分析

人口普查是现代民生工作之一, 通过人口普查, 掌握一定区域及国家总人口数量, 以此为人口政策的制定提供依据。对于我国来说, 当前共有 14 亿人口, 技术非常庞大, 若按照传统人工普查的方式, 速度较难, 需要投入大量的时间与资金, 而且在普查周期内, 有新的人口降生, 也有人口死亡, 在一定程度上干扰普查结果的准确性。而应用大数据技术后, 一方面, 可由人民自动在互联网内填写人口普查问卷, 另一方面, 各地方相关部门也可帮助人口普查对当地居民资料进行整理, 大大缩短了人口普查的时间, 且各性别比例、年龄占比、学历程度等角度出发, 全面对人口现状予以分析, 以更加准确的了解现代人口情况^[5]。

5 总结

综上所述, 大数据作为科学技术快速发展的产物, 具有诸多方面的优势, 有效将其应用到统计分析工作当中, 可大大提升统计分析结果的准确性, 以此为相关决策的制定提供支持。所以, 现代国民经济分析、就业及失业率计算、财务审计、医疗卫生事业发展、人口普查等工作开展过程中, 应加强对大数据的重视程度, 针对各领域工作的具体情况, 合理对大数据进行应用, 以得到更加真实、准确的统计分析结果。

参考文献:

- [1] 王秀芳. 大数据视角下统计分析在财务管理中的应用及创新[J]. 河北企业, 2022, 06(01): 125-127.
- [2] 朱孔春. 基于大数据背景下的统计分析在财务管理中的应用[J]. 营销界, 2021, 10(Z4): 135-137.
- [3] 李应硕. 对大数据统计分析方法在经济管理领域中的研究应用[J]. 商展经济, 2021, 06(20): 125-127.
- [4] 张旭. 大数据技术在金融统计分析中的应用初探[J]. 当代经济, 2021, 02(07): 26-29.
- [5] 陆瑜. 大数据时代下数据统计分析在高校教育管理中的应用探讨[J]. 无线互联科技, 2021, 18(05): 128-129.

作者简介:

高倩倩(1987-)女, 汉, 吉林白城人, 大学本科, 研究方向: 统计学、大数据科学与应用。对外经济贸易大学统计学院在职人员高级课程研修班学员。