

大数据统计分析方法在企业经济管理中的应用

华文字¹ 陈鹏旭² 肖欣佑³ 张振华⁴ 徐卓君³

1 中央民族大学理学院, 北京海淀, 100081;

2 中央民族大学管理学院, 北京海淀, 100081;

3 中央民族大学经济学院, 北京海淀, 100081;

4 广州科技职业技术大学经济与管理学院, 广东广州, 510555;

摘要: 文章介绍了回归分析、推断性统计分析和聚类分析与分类分析等三种大数据统计分析方法, 分析三种方法在企业经济管理中的应用场景, 结合华为、阿里巴巴和沃尔玛的具体案例分析回归分析、推断性统计分析和聚类分析与分类分析在市场战略规划、财务风险管理和供应链优化中的应用。

关键词: 大数据统计分析; 企业; 经济管理

DOI: 10.69979/3029-2700.25.07.061

引言

近年来, 信息技术发展迅速, 大数据统计分析方法为企业提供了问题的解决方案, 利用高效的数据处理, 企业能够从海量信息中提取有价值的洞察, 推动市场战略、财务管理和供应链优化等各个领域的科学决策, 提升整体经营效能。然而, 当前许多企业在经济管理的过程中仍存在问题, 导致企业无法充分利用大数据的潜力。在此背景下, 文章分析大数据统计分析方法的分类及其在企业经济管理中的应用, 旨在为企业在复杂多变的市场环境中实现可持续发展提供重要保障。

1 大数据统计方法分类

1.1 回归分析

回归分析能够探索因变量与自变量之间的关系, 并借此进行预测, 被广泛应用于企业经济管理中^[1]。其核心在于利用最小二乘法等方法拟合数据, 得到回归方程, 预测未来的经济趋势或评估不同因素对企业绩效的影响程度。回归分析的基本目标是建立一个数学模型, 揭示一个或多个市场价格、广告投入、生产成本等自变量与销售额、利润率、市场份额等因变量之间的定量关系。

1.2 推断性统计分析

推断性统计分析是一种基于样本数据对总体特征进行估计的统计方法。核心目的是从样本数据中提取关键信息, 根据信息推断整体经济活动、市场行为或财务状况的规律性。在企业经济管理中, 推断性统计分析包

括假设检验、置信区间估计和模型推断等手段, 利用分析样本数据的不确定性, 为企业决策提供科学依据。该方法可以帮助企业识别潜在的风险, 评估政策或行动的效果, 有效支持战略规划、资源配置及风险管控等管理决策。

1.3 聚类分析与分类分析

聚类分析是一种无监督学习方法, 能够根据数据的内在相似性将数据对象分组, 揭示出在某些特征上相似的群体。分类分析是一种对历史数据进行学习并建立预测模型将新数据分配到已知类别中的监督学习方法, 常被企业用于细化市场分析并优化资源配置。聚类分析与分类分析能够对企业数据进行详细的归类, 帮助企业识别数据中的潜在趋势, 支持精细化管理决策。

2 大数据统计分析方法在企业经济管理中的应用场景

2.1 市场战略规划

企业可以利用回归分析评估价格弹性, 预测价格变动对需求量的具体影响, 从而优化定价策略。与此对广告费用、促销策略等市场营销活动与销售表现之间进行回归分析, 企业能够明确投资回报率 (ROI), 优化营销预算的配置。回归分析还可用于新产品推出前的市场需求预测, 对历史数据进行回归分析, 帮助企业估计新产品可能获得的市场份额, 为资源配置和生产计划提供参考。

2.2 财务风险管理

企业对财务数据进行假设检验和置信区间分析,识别财务报表中的异常波动,评估财务决策的风险,并预判未来财务状况的可能变化^[2]。例如,对历史财务数据进行分析,企业能够预测资本结构调整或税务政策变化对财务表现的潜在影响,为财务决策提供数据支撑。推断性统计分析还可以用来评估企业的财务稳健性,识别潜在的违约风险或资本流动问题,及时采取措施,防范财务危机的发生。

2.3 供应链优化

企业利用聚类分析可以根据供应商的地理位置、交货能力、产品质量等特征将供应商分类,制定针对性的采购策略,优化供应链管理。比如,企业可以对供应商进行区域性聚类,将地理位置相近的供应商归为一类,从而降低物流成本和供应链风险。分类分析可以用于需求预测和库存管理,企业可以根据历史销售数据、季节性波动、市场趋势等因素,建立预测模型,精确预测不同商品或原材料的需求量,从而优化库存水平,减少库存积压和缺货现象。

3 大数据统计分析方法在企业经济管理中的应用案例分析

3.1 案例 1: 华为利用回归分析进行市场战略规划

华为公司在市场战略规划过程中,应用了大数据统计分析方法中的回归分析,优化了其全球市场布局^[3]。在面对复杂多变的市场环境时,回归分析为华为提供了量化市场变量之间关系的能力,帮助华为公司精准把握不同因素对市场表现的影响,进而制定出最优的市场战略。

3.1.1 背景及数据收集

华为在全球市场的扩展过程中,扩展过程受价格策略、广告投入、竞争对手行为、地区经济增长、消费者需求变化等诸多因素影响。为了分析以上因素之间的关系,华为首先进行市场调研,获得了过去五年内在不同地区的销售数据、广告费用、价格策略、市场份额以及各类经济指标等数据。数据来源包括华为内部销售系统、行业报告以及第三方市场调研数据。

在此基础上,华为的市场分析团队决定利用回归分析来识别价格、广告投入、竞争强度、消费者偏好等因素对销售额和市场份额的影响。华为利用该过程揭示不

同市场因素的权重,提高营销投资回报率。

3.1.2 回归分析模型构建

为了进行回归分析,华为采用了多元线性回归模型。

假设销售额(因变量)受到以下因素(自变量)的影响:

广告投入(X_1): 市场推广的费用。

价格策略(X_2): 产品定价对需求量的影响。

竞争强度(X_3): 主要竞争对手的市场行为,如价格下调、产品创新等。

消费者需求(X_4): 目标市场的需求强度。

回归模型如式 1 所示:

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4 + \epsilon \quad (1)$$

其中, Y 代表销售额, β_0 为常数项, $\beta_1, \beta_2, \beta_3, \beta_4$ 为回归系数, ϵ 为误差项。

华为团队对收集到的数据进行回归分析,并得出回归系数,进而量化每个因素对销售额的影响。

3.1.3 数据分析与结果

通过回归分析,华为获得了回归系数见表 1。

表 1 华为公司的回归系数

自变量	回归系数 (β)	标准误差	t 值	P 值
广告投入(X_1)	0.45	0.12	3.75	0.000
价格策略(X_2)	-0.30	0.09	-3.33	0.001
竞争强度(X_3)	-0.25	0.10	-2.50	0.015
消费者需求(X_4)	0.60	0.14	4.29	0.000

数据来源: 2024 年华为公司企业财务年报

根据回归系数的显著性检验,回归系数分别为 0.45 和 0.60。结果显示,广告投入(X_1)和消费者需求(X_4)对销售额具有显著的正向影响,表明增加广告投入、满足消费者需求能够显著提高销售额;价格策略(X_2)和竞争强度(X_3)则显示出负向影响,回归系数分别为 -0.30 和 -0.25,表明价格上调、激烈的市场竞争会削弱销售业绩。

根据统计分析结果,回归模型的拟合优度(R^2)为 0.85,说明该模型能够解释 85%的销售额变异,表明回归模型具有较强的预测能力。

根据 F 检验的结果($P < 0.05$),整体回归模型具有显著性,能有效预测市场战略的实施效果。

3.2 案例 2: 阿里巴巴利用推断性统计分析进行财务风险管理

在复杂的市场环境背景下，阿里巴巴的财务风险管理面临汇率波动、资金链断裂、应收账款回收风险等诸多挑战。阿里巴巴公司利用推断性统计分析，及时识别和应对财务风险，有效提升了财务风险管理的精度，避免了财务危机。

3.2.1 推断性统计分析的应用背景

为了精确识别并预测潜在风险，阿里巴巴公司引入了假设检验、置信区间、方差分析等推断性统计分析方法。该方法帮助财务部门预测未来的财务波动，增强了公司在财务决策中的应对能力。

3.2.2 数据收集与推断性统计分析方法的选择

阿里巴巴首先收集了各季度的现金流量、营业收入、应收账款、货币市场利率、汇率变动数据以及不同行业的市场风险数据等大量关于其财务状况的历史数据，利用以下几种方法进行财务风险预测：

假设检验：用于验证销售收入、库存周期、应收账款等不同财务变量对财务风险的影响。

置信区间：为每个财务风险预测结果构建置信区间，以提供风险评估的可信度，降低由于单一预测结果带来的误差。

方差分析：分析不同地区、不同业务单元之间财务数据的差异，识别影响财务风险的关键因素。

3.2.3 数理分析与案例数据展示

假设其每个季度的应收账款数据呈现出明显的波动性（以 6 个季度阿里巴巴的季度应收帐款数据为例，详细数据见表 2），可能对现金流产生较大影响。对从 2023 年第一季度开始的 6 个季度应收账款数据进行假设检验和回归分析，评估不同因素对应收账款周期的影响。

表 2 6 个季度阿里巴巴的季度应收帐款数据

季度	销售收入 (X1) /万元	应收账款 (Y) /万元	季度利率 (X2) /%	收账周期 (X3) /天
Q1	11500	5110	4.00	85
Q2	12800	5520	4.20	88
Q3	13400	6030	4.10	90
Q4	14100	6510	4.30	92
Q5	15800	7050	4.50	93
Q6	16600	7525	4.70	95

数据来源：2023-2024 年阿里巴巴公司企业财务年报

结果表明，Q6 季度的销售额和收账周期均最高，可能存在财务风险，需要进一步分析。

3.2.4 假设检验

假设检验能够检验应收账款与销售收入之间的相关性，以及市场需求指数对应收账款的影响。

阿里巴巴公司使用 T 检验和方差分析检验分析应收账款波动的原因。例如，使用单样本 T 检验来判断应收账款是否高于某一基准值，从而评估风险。假设检验结果：

H0（零假设）：销售收入与应收账款之间没有显著关系。

H1（备择假设）：销售收入与应收账款之间有显著正相关。

计算 T 值和 P 值后，发现 P 值远小于 0.05（例如 0.001），因此拒绝零假设，表明销售收入与应收账款之间有显著正相关关系。

3.2.5 置信区间

阿里巴巴公司计算应收账款的置信区间，得到了应收账款的预测区间。例如，在 95% 置信度下，未来一个季度的应收账款预计将在 6200 万元~6800 万元之间波动。该区间的估算提供了更为精准的风险评估，有助于阿里巴巴提前做出调整措施。

3.2.6 方差分析

阿里巴巴利用方差分析检验 Q6 季度不同地区（东部、西部和南部）之间应收账款周期的差异。各地区的销售收入如图 1。

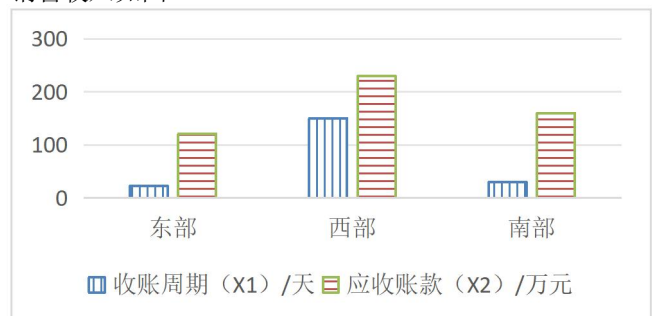


图 1 各地区之间应收账款周期的差异 (数据来源：2023-2024 年阿里巴巴公司企业财务年报)

方差分析结果显示，其他变量不变的情况下，西部地区的收账周期、应收款项较高，存在较大财务风险，且结果具有统计学意义 ($P < 0.05$)。分析结果帮助阿里巴巴公司进行区域性风险管理，针对性地调整其应收账款政策。

案例 3：沃尔玛通过聚类分析与分类分析进行供应链优化

利用大数据分析统计方法，沃尔玛公司对海量销售

数据、库存数据及顾客行为数据的深入分析，在供应链管理中取得了显著的成效。

沃尔玛公司利用聚类分析方法对其各个门店进行市场细分，实现了个性化的库存管理。聚类分析能够将门店按照销售特点、地理位置和顾客需求等维度分成若干组，帮助沃尔玛确定不同区域的供应链策略。

对过去 12 个月内门店销售数据进行分析，随机选取每组中具有聚类特征的 6 个有代表性门店分别编号为 A、B、C、D、E、F，分析其销售额和顾客偏好指数（见表 3），使得供应链更加精细化。

表 3 不同销售模式和供应链管理需求

门店编号	销售额（月均）/万元	顾客偏好指数（0-1）
A	513,250	0.8
B	607,330	0.7
C	450,410	0.6
D	721,000	0.9
E	550,100	0.85
F	481,100	0.65

数据来源：2024 年沃尔玛公司企业财务年报

结果表明，不同类型的门店在商品采购量、定价策略及物流安排等方面需采取差异化管理，以最大化库存周转效率。

分类分析是一种根据历史数据对商品类别或供应链环节进行分类，从而预测未来的需求波动的监督学习方法。沃尔玛公司利用分类分析方法进一步优化其供应链的预测精度。沃尔玛采用支持向量机（SVM）算法，将各个商品按照销售趋势分为 P001、P002 和 P003 三类，预测商品在不同促销期间的需求，提前安排采购。SVM 模型进行训练和预测的具体数据见表 4。

表 4 SVM 模型进行训练和预测

商品编号	历史销售额/万元	价格波动	促销活动	季节性因素	销售分类（高/低）
P001	1,209,000	高	是	夏季	高
P002	878,000	低	否	冬季	低
P003	1,053,000	中	是	秋季	高

数据来源：2024 年沃尔玛公司企业财务年报

沃尔玛利用分类分析为每个商品类别定义相应的库存水平、补货策略及物流配送计划。该优化措施提高了库存周转率，能够有效地进行了供应链优化。

4 结语

回归分析、推断性统计分析、聚类分析等大数据统计分析方法帮助企业深刻洞察市场动态、优化资源配置和预测财务风险，为决策提供科学依据，取得了较好的应用效果，逐渐被广泛应用于各类企业中。未来随着科学技术手段的不断进步，大数据统计分析将更加深入应用于经济管理各领域，成为企业提升运营效率的核心驱动力。

参考文献

[1] 赵璟璇. 大数据背景下企业风险管理和内部控制整合应用研究[D]. 内蒙古财经大学, 2022.

[2] 范雯雯. 数字化转型对企业经济管理模式的的影响与变革路径[J]. 中国经贸导刊, 2024, (18): 148-150.

[3] 傅建千. 大数据时代企业经济管理模式存在的问题及优化策略[J]. 中国商界, 2024, (10): 196-197.

作者简介：华文字（2004—），女，山东德州人，中央民族大学理学院本科在读，研究方向：应用统计。
邮箱：2955627767@qq.com