

商务数据分析实验任务书一

一、实验主要目的

练习使用统计软件工具对商务数据进行回归统计分析。

二、实验主要任务概述

1. 练习和理解课程中三个回归模型案例 (多元线性回归, 带交叉项的多因素方差分析, 逻辑回归) 的统计分析方法.
2. 阅读上述三个案例的分析报告并学习报告写法.
3. 使用自选统计软件工具 (R, SAS, SPSS, Matlab 等) 进行回归建模分析, 并撰写分析报告, 具体要求见案例描述.

注: 如使用 Rstudio 导入含中文数据的 csv 文件时报错或出现乱码, 可改用如下代码导入数据文件(例如<E2.csv>)

```
install.packages('readr'); library('readr')  
a<-read_csv('E2.csv')
```

如果仍然报错, 使用记事本文件打开 E2.csv, 点击<另存为>, 编码选择<UTF-8>, 保存覆盖原文件.

三、回归模型数据分析案例

案例 1: 上市公司财务报表分析与预测

1. 研究目的

通过对上市公司公开财务报表信息的统计分析, 预测该企业来年的盈利状况。

2. 数据介绍

随机抽取某年度 500 家上市企业的相关数据。对所有样本, 解释变量来自当年, 因变量来自下一年。因变量是该企业的净资产收益率, 解释变量有: 资产周转率, 利润率=主营业务利润/主营业务收入, 债务资本比率, 成长速度=主营业务收入增长率, 市倍率, 收

入质量=应收账款/主营业务收入, 存货率=存货/资产总计, 对数变换后的资产总计, 以及当年净资产收益率。该数据存放在数据文件<E1.csv>中。

3. 作业要求

- **问题理解:** 请参阅相关金融、会计教材, 对本案例所涉及的所有指标的常见定义和会计含义进行简介。
- **做完整的多元线性回归分析:** 包括描述性统计, 参数估计及假设检验。
- 将分析结果汇总成如回归模型应用案例-分析报告 1 所示的简短研究报告。

案例 2:非学位职业培训机构的学员数据分析

1. 研究目的

通过对某非学位职业培训机构的 178 名学员进行数据分析, 了解什么样的学员有可能获得更好的学习效果(成绩)。以此为依据, 来指导未来的招生计划。

2. 数据介绍

随机抽取的某年度 178 名已经结业学员。对所有样本, 因变量是学员整个培训过程中所通过的各个考试的平均成绩, 用于解释该因变量的自变量有: 性别、出生日期、企业性质、最高学历, 以及最高学历的毕业时间。该数据存放在数据文件<E2.csv>中。

3. 作业要求

- **问题理解:** 请上网搜索国内主要的非学位培训机构, 登录其网站, 了解相关信息。收集相关媒体报道, 为本案例提供良好的背景支持。
- **做完整的方差分析,** 包括参数估计和假设检验。
- 将分析结果汇总成如回归模型应用案例-分析报告 2 所示的简短研究报告。

案例 3: 移动通信客户流失规律分析

1. 研究目的

通过对某移动通信公司客户的流失数据分析, 了解客户流失规律, 建立流失预警系统, 为客户关系管理服务。

2. 数据介绍

某年度随机抽取的 1000 个移动通信客户。因变量是他们来年的流失行为（0=未流失，1=流失）。为了能够预测客户的未来行为，我们采集了下面这些来自当年的指标：客户等级（区分 VIP 客户等级）：1，2，3，4；主叫次数（%）：7 日内日均主叫次数/90 日内日均主叫次数；被叫次数（%）：7 日内日均被叫次数/90 日内日均被叫次数；通话时长（%）：7 日内日均通话时长/90 日内日均通话时长；费用（%）：7 日内日均通话费用/90 日内日均通话费用。该数据存放在数据文件<E3.csv>中。

3. 作业要求

- **问题理解：**请参阅相关媒体报道，理解客户流失对移动通信类企业的重大意义。
- **做完整的逻辑回归分析，**包括描述分析、参数估计、假设检验，以及预测评估。
- 将分析结果汇总成为如回归模型应用案例-分析报告 3 所示的简短研究报告。