

## 第二次作业（Logistic 回归）

请用 Logistic 回归模型估计违约概率，数据集为德国信贷数据集“germancredit.csv”。数据集包含了 1000 个贷款申请的属性和结果，由汉堡大学统计研究所提供。关于变量的描述请见文件“germancreditDescription.doc”。

要求：请在 1000 个观测值中随机选取 900 个作为训练集，建立合适的 logistic 回归模型，进行解读。并用剩下的 100 个作为测试集，展示你的预测效果。

### 报告的具体要求：

- （1） 报告框架要清晰，格式符合数据分析报告要求、撰写流畅。  
研究背景、研究问题以及研究意义，研究方法，探索过程，以及结论有清晰阐述（不是追求多）。
- （2） 数据的描述性分析完整、合理，呈现方式符合规范，并且有清楚的解释。
- （3） 建模分析部分，需要给出必要的统计分析过程以及方法。所选模型的合理性要说明。对模型的结果需要进行合理的解读。