

生存分析小论文

注意事项:

1. **提交期限: 2020 年 7 月 22 日 11: 59pm** 合并成一个 PDF, 通过 email 提交; **不能逾期**, 否则不计分;
2. 可以讨论, 但**不得抄袭**, 否则等同中山大学考试作弊处理;
3. 评价要点:
 - (1) 达到学校要求: 不能纯谈体会, 要包含思考过程、内容总结、体会, 字数不少于 3000 字 (或: 页数在 5 页-20 页范围内);
 - (2) 数据分析过程、R 代码要完整清晰;
 - (3) 至少包含 5 个生存分析知识点, 例如可从下面选取:
 - a. KM-plot;
 - b. Log-rank test;
 - c. Confounding;
 - d. Interaction;
 - e. Standard Cox model;
 - f. Log-log plot;
 - g. Observed-expected plot;
 - h. Goodness of fit test;
 - i. Schoenfeld residual plot;
 - j. Stratified analysis (stratified log rank test, stratified Cox model, etc.);
 - k. Extended Cox model;
 - l. AFT models;
 - m. Adjusted survival curves;
 - n. Prediction of survival probabilities;
 - o. Counting process;
 - p. Likelihood construction;
 - q. Censoring and truncation.
 - (4) 可涉及本课程未教的相关知识点, 例如: **competing risk** (该数据集适用), **random survival forest**, 等等;
 - (5) 最后评分会类比本科毕业论文, 按照一定比例将所有同学的小论文分成以下档次: 优、良、中等、及格、不及格, 再进行百分制转换。除非写的太差/逾期不交, 否则你的小论文不会落入最后两档次。

数据(Framingham Heart Study):

1. handouts_fhs.csv;
2. 数据说明:
Framingham Longitudinal Data Documentation_v2.pdf
3. 参考文献与网站:
<https://www.nature.com/articles/s41569-019-0202-5>
<https://framinghamheartstudy.org/>

要求:

1. 对 Framingham heart study 数据集进行描述;
2. 对你感兴趣的生存分析相关的科学问题进行描述, 并转化成统计问题;
3. 运用本课程涉及的方法解决你描述的统计问题, 并解释结果;
4. 总结成一个报告, 字数不少于 3000 字 (或: 页数在 5 页-20 页范围内), 在报告最后附上所有 R code。