P8108 Group 2 Survival Analysis Project

Yiming Zhao (yz3955) Wenshan Qu (wq2160) Tucker Morgan (tlm2152)
Junzhe Shao (js5959) Benjamin Goebel (bpg2118)

2022-10-19

```
library(survival)
library(tidyverse)
library(tidymodels)
```

Train Test Split

```
set.seed(2022)

rotterdam_split <- initial_split(rotterdam, prop = 0.8, strata = death)
rotterdam_training <- training(rotterdam_split)
rotterdam_test <- testing(rotterdam_split)
```

Perform 10-fold Cross-Validation

The output contains 1 row for each fold/repeat. So, 10 folds * 5 repeats = 50 rows. The split_analysis column is a list column containing a data frame for each row with 9 folds combined, and the split_assessment column is a list column containing a data frame for each row with 1 fold.

```
set.seed(2022)

rotterdam_folds <- vfold_cv(rotterdam_training, v = 10, repeats = 5,
                            strata = death)

rotterdam_folds <- rotterdam_folds %>%
  mutate(split_analysis = map(splits, analysis),
         split_assessment = map(splits, assessment))
```

Introduction

Methods

Exploratory Data Analysis

Cross-Validation

Cox/Cox with elastic net

Random survival forest

Conformalized survival analysis

Supplemental analyses

Results

Exploratory Data Analysis

Cross Validation

Cox/Cox with elastic net

Random survival forest

Conformalized survival analysis

Supplemental analyses

Discussion

How our results compare with past research

Conclusion