Bump Hunting for Survival Outcomes

Vladimir Minin, Xueli Liu, Steve Horvath
Correspondence: shorvath@mednet.ucla.edu

Here we propose a method on how to use bump hunting (Friedman and Fisher, 1999) for censored (failure time) data (Liu et al. 2004).

The idea is simple: turn the censored outcome into an uncensored 'surrogate' outcome. A similar idea was suggested by Therneau et al (1990): use martingale residuals corresponding to an intercept only Cox regression model in regresssion trees models. First, we fit an intercept only Cox regression model to the outcome.
Next we consider the following four surrogate outcomes in bump huntinng

Surrogate1 =- (predicted number of events binned into 4 quartiles)
Surrogate2 = martingale residual
Surrogate3 = deviance residual
Surrogate4 = discretized martingale residual

To be more specific, surrogate1 is minus the predicted number of events binned into 4 quartiles (use the predict.coxph(...,type="expected") function in R;

surrogate 2 is the martingale residual $M_i = \delta_i - \hat{\Lambda}_0(t_i)$,

surrogate3 is the deviance residual (Therneau et al., 1990 LeBlanc and Crowley, 1992)

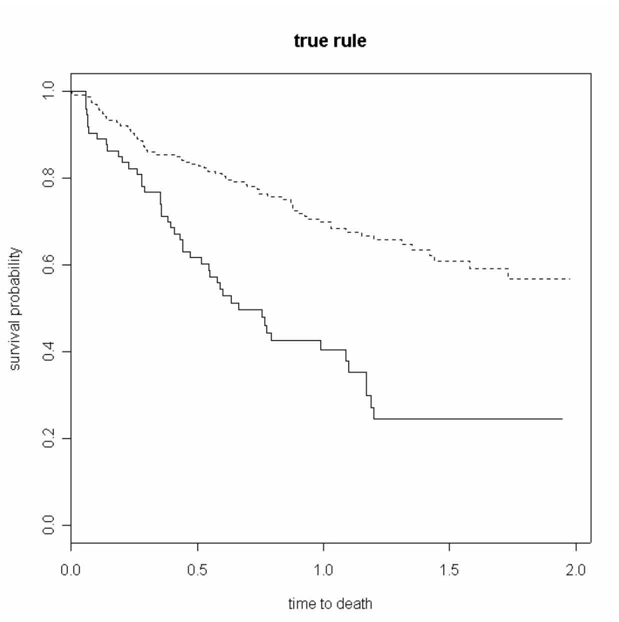
$$d_i = sign(M_i)\sqrt{2\left[\delta_i \log\left(\frac{\delta_i}{\Lambda_0(t_i)}\right) - M_i\right]}.$$

where δ_i is the censoring indicator, t_i is the possibly censored survival time, $\hat{\Lambda}_0$ is the estimated baseline cumulative hazard function;
surrogate4 is the martingale residual binned into its 4 quartiles.

SIMULATIONS

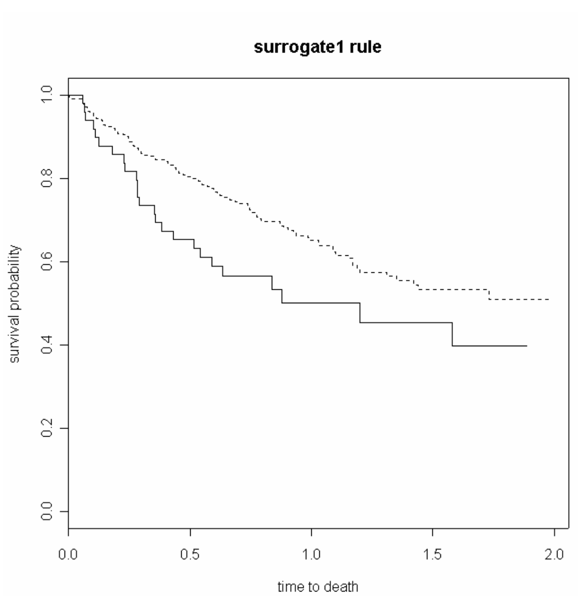
We performed a series of simulations to test the performance of bump hunting and to choose an appropriate outcome variable for survival analysis.
To test which of these surrogate outcomes performs best when used in bump hunting, we simulated survival data as follows. We generated 10 covariates X1, …, X10 ~ Beta(1,5). The first 3 covariates contain the signal: observations had increased risk when X1 > 0.1 and X2 > 0.1 and X3 > 0.1. Failure times were generated using the exponential distribution with random uniform censoring. Below are the rules, produced by bump hunting for each of the surrogates. We visualize the results of each rule with Kaplan-Meier curves and list the corresponding logrank test p-values.

True rule: X1 > 0.1 and X2 > 0.1 and X3 > 0.1

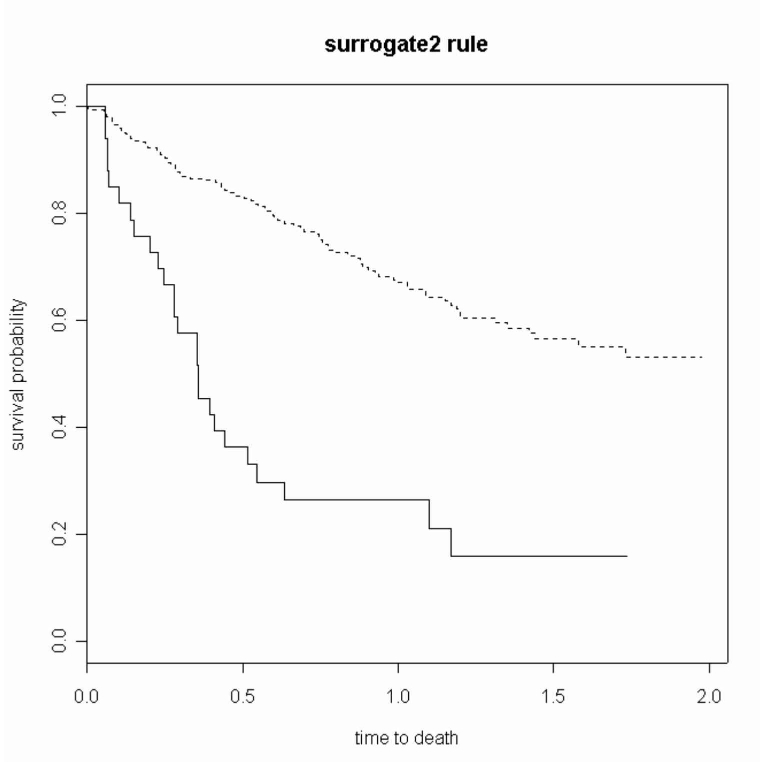


logrank test p-value = 6.32e-08

Surrogate1 leads to the following rule: X3 > 0.02506 and X2 > 0.1632 and X2 < 0.4002 and X7 < 0.2636 and x5 > 0.0609 and the corresponding Kaplan Meier plot

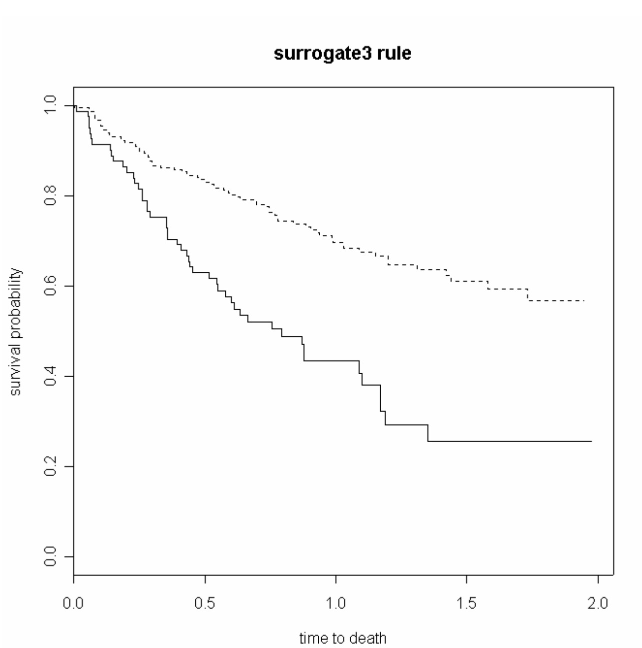


logrank test p-value = 0.0319
This is not particularly close to the ground truth.
Surrogate2 (martingale residual) leads to the follwing rule: X3 > 0.1177 and X1 > 0.1377 and X8 > 0.05616 and X8 < 0.3450 and X2 > 0.07522



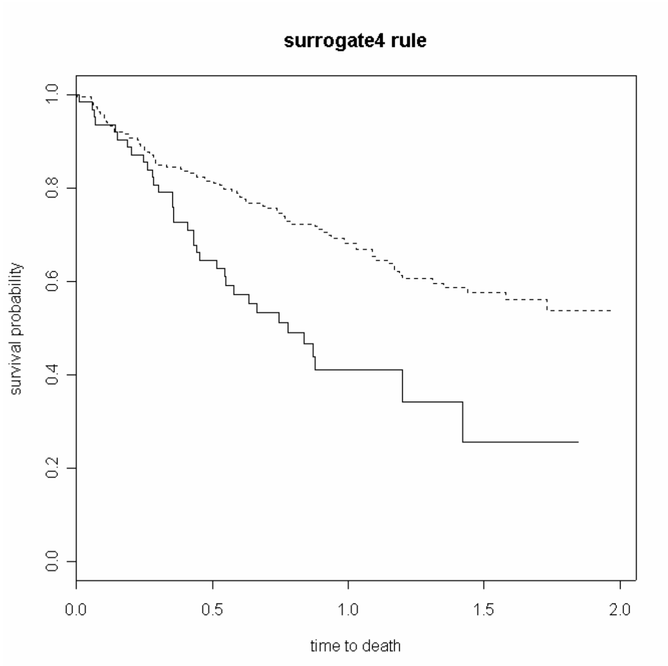
logrank test p-value = 5.11e-11
The rule contains too many “noise” covariates and reflects overfitting.

Surrogate3 rule (deviance residual): X3 > 0.07832 and X1 > 0.1111 and X6 < 0.2751



logrank test p-value = 4.61e-07
This rule contains 2 of the 3 signal predictors and sensible cut-offs. We find that it works best in this study.

Surrogate4 rule: X1 > 0.1286 and X3 > 0.02203 and X9 < 0.1814 and X6 < 0.2437



logrank test p-value = 0.000108
This result is not as good as surrogate 3 since it contains more noise covariates.

Discussion:
In this and similar examples, we find that surrogate 3 (the deviance residual) leads to the best rule induction results for bump hunting. The results should be compared on the basis of the recovered rule and the corresponding p-values. But the p-values should not be used for inference since severe overfitting took place. Instead, we use them here as descriptive measures of survival curve separation. The overly impressive p-values that results from surrogate 2 reflects overfitting as can be seen from the corresponding rule: it contains several noise covariates. We have also found that using the deviance residuals in regression trees (rpart function in R) leads to better results than the martingale residuals, which were suggested by Therneau et al (1990). There are dozens of survival tree methods and we do not claim that using deviance residuals is optimal. However, we have found in multiple unpublished simulation studies that it works well.

References:

- Liu X, Minin V, Huang Y, Seligson DB, Horvath S (2004) Statistical Methods for Analyzing Tissue Microarray Data. Submitted to a Special Issue of the Journal of Biopharmaceutical Statistics.
- Friedman, J. and Fisher, N. (1999). Bump hunting in high dimensional data. Statistical Computing, 9, 123-143.
- LeBlanc, M. and Crowley, J. Relative risk regression trees for censored survival data. Biometrics, 1992, 48(2), 411--425.
- Therneau, T., Grambsch, P. and Fleming, T. Martingale based residual for survival models. Biometrika, 1990, 77, 147--160.