

Project: Evaluation of survival estimates

Supervisor: Ørnulf Borgan

In medical statistics, one is often interested in studying the lifetimes T_i (or times to onset of a specific disease) for a group of individuals, and how the lifetimes depend on vectors of explanatory variables \mathbf{x}_i . In particular, one would like estimate the conditional survival function $S(t|\mathbf{x}_i) = P(T_i > t|\mathbf{x}_i)$. One complication, which makes survival analysis different from other parts of statistics, is that the lifetimes T_i are not observed for all individuals. For some individuals one only knows that $T_i > C_i$ for some censoring time C_i .

There exist a number of methods for estimation of $S(t|\mathbf{x}_i)$ in the presence of censoring; two important methods are Cox's and Aalen's regression models. Assume that one has two or more estimates of the conditional survival function $S(t|\mathbf{x}_i)$. One would then like to evaluate these estimates in order to determine which of them is the best.

There exist a number of criteria for evaluating survival estimates. A popular method is the Brier score, and that will be the focus of this project. When there are no censored observations, the Brier score at time t is given by (for n observations)

$$BS(t) = \frac{1}{n} \sum \{I(T_i > t) - \hat{S}(t|\mathbf{x}_i)\}^2.$$

In the presence of censoring, the Brier score has to be modified, and there exist two ways of doing this. The main aim of the project is to study these two modifications of the Brier score and make recommendations of which one should be used.

The work on the project will probably be done according to the following plan:

- Read about basic concepts, models and methods in survival analysis, and summarize this in writing. (No previous knowledge of survival analysis is assumed.)
- Try out survival regression models (at least Cox's) and estimation of conditional survival functions on real data, and document this in writing.
- Try out the two modifications of the Brier score for censored data (including the necessary programming) and document this in writing.
- Study the two modifications of the Brier score by simulations, and (perhaps) by theoretical considerations, and summarize this in writing.
- Write the conclusions of the project.