

# SPLINE SMOOTH ESTIMATES OF SURVIVAL

Jerome Klotz

Ohio State University

## 1. Introduction

Let  $X$  be survival time with continuous distribution  $F_X(x)$  and density  $f_X(x)$ . Similarly, let  $Y$  be time to censoring, independent of  $X$ , with continuous distribution  $F_Y(y)$  and density  $f_Y(y)$ . We observe time on trial,  $T$ , and death or censoring indicator,  $D$ , where

$$T = \min(X, Y)$$
$$D = \begin{cases} 1 & \text{if } X \leq Y \quad (\text{death}) \\ 0 & \text{if } X > Y \quad (\text{censoring}) \end{cases} .$$

Using a sample  $\{T_i, D_i\}; i=1, 2, \dots, n\}$  we wish to find a smooth estimate of the survival distribution  $1 - F_X(x) = P[X > x]$ .

Define the hazard function by

$$h_X(x) = f_X(x) / (1 - F_X(x))$$

and the integrated hazard function by

$$H_X(x) = \int_0^x h_X(u) du = - \int_0^x d \ln (1 - F_X(u)) \quad ,$$

which is related to survival by  $1 - F_X(x) = e^{-H_X(x)}$ . Defining the indicator function  $I[A]$  (1 or 0 according as the event  $A$  holds or not), the sample cumula-