Useful definitions and properties

If $v_i = 1$ then x_i is an observed failure time; if $v_i = 0$ then x_i is a censoring time.

$H(t)$ is the *integrated hazard*. If we define a new random variable U by $U := H(T)$ then U has an exponential(1) distribution. (See 'A useful result' below.)

$\hat{H}(t)$ is the estimated integrated hazard (after fitting a model).

The *Cox-Snell* residual y_i is defined by

$$y_i := \hat{H}(x_i).$$

The y_i will look like a (censored) sample from an exponential(1) distribution if the model fits well.

The *modified* Cox-Snell residual y'_i is defined by

$$y'_i := (1 - v_i) + \hat{H}(x_i).$$

We are effectively adding 1 to all censored values (note that we would have expected a censored observation from exponential(1) to have a survival time shortened by 1). The expectation of y'_i is therefore 1, irrespective of censoring.

The *Martingale* residual y''_i is defined by

$$y''_i := 1 - y'_i = v_i - \hat{H}(x_i).$$

The Martingale residual has expectation zero and can be thought of as 'observed' (v_i) minus 'expected' ($\hat{H}(x_i)$).

A useful result

To show $U := H_T(T)$ has an exponential(1) distribution (I have subscripted the H to emphasize it is the integrated hazard belonging to T):

Let F_U and F_T be the survivor functions for U and T respectively. Then, by definition:

$$F_U(u) = \mathcal{P}\{U > u\} = \mathcal{P}\{H_T(T) > u\}.$$

H_T is increasing and and has an inverse H_T^{-1}, so we can write:

$$F_U(u) = \mathcal{P}\{T > H_T^{-1}(u)\} = F_T(H_T^{-1}(u))$$

where the second equality is by definition. We now use $F_T = \exp(-H_T)$ to obtain

$$F_U(u) = \exp\{-H_T(H_T^{-1}(u))\} = \exp(-u)$$

and $\exp(-u)$ is the survivor function for an exponential(1) distribution.