Cox PH Model

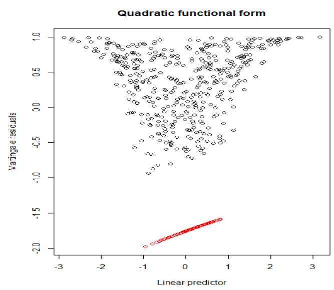
Martingale residuals

Martingale residuals

- Martingale residuals are defined as
$$r_{M,i} = \delta_i - r_{CS,i} = \delta_i - \hat{H}(t_i \mid Z_i)$$
  - δ_i counts number of events for the ith subject in $[0, t_i]$
  - $\hat{H}(t_i \mid Z_i)$ is the expected number of events for the ith subject in $[0, t_i]$ under the estimated model
  - $r_{M,i}$ is the "excess" number of events
- Martingale residuals represent the discrepancy between the observed data and fitted model.
- It takes value in $(-\infty, 1]$.
- Interpretation of martingale residuals
  - Positive values mean the patient died sooner than expected (according to the model) e.g. $r_{M,i} = 1 \Rightarrow$ "die too soon"
  - Negative values mean that the patient lived longer than expected. E.g. a large negative value of $r_{M,i} \Rightarrow$ "live too long"

Martingale residuals

- Martingale residual can be used to access the functional form of covariates.
- Suppose we generate data from an exponential regression with a quadratic form of x
$$h(t) = \exp(-x^2/2).$$
- But instead fit the data with a Cox model with
$$h(t) = \exp(-\beta x).$$
- Martingale residual plot reveals true function form.



Martingale residuals

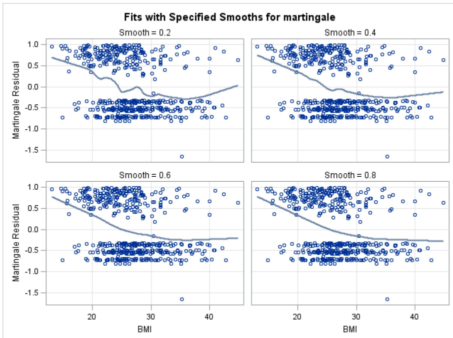
- To find approximate functional form of a continuous variable, say, Z_1,
  - Fit a Cox model without Z_1, and compute the martingale residuals.
  - Draw the scatter plot of $\{r_{M,i}\}$ versus $\{Z_{1i}\}$ and overlay it with a smoothed curve (LOWESS)
  - The smoothed curve suggest function forms for Z_1.

SAS Example

- Assess functional form of a continuous covariate: bmi

```
proc phreg data = whas500;
class gender;
model lenfol*fstat(0) = ;
output out=residuals resmart=martingale;
run;
proc loess data = residuals
plots=ResidualsBySmooth(smooth);
model martingale = bmi / smooth=0.2 0.4 0.6 0.8;
run;
```

The plots=ResidualsBySmooth option on the proc loess statement allows us to examine residual plots for each smooth (with loess smooth themselves)



SAS Example

- Plot indicates a quadratic functional form for BMI.
- One can repeat this procedure for all continuous covariates.
- It was found that linear form for other variables seem fine.
- We will fit a quadratic function of BMI.

```
proc phreg data = whas500;
class gender;
model lenfol*fstat(0) = gender bmi|bmi age hr;
output out=residuals resmart=martingale;
run;
proc loess data = residuals
plots=ResidualsBySmooth(smooth); model
martingale = age / smooth=0.2 0.4 0.6 0.8;
run;
```