

By looking at our residual plots, we can perform some model diagnosis as to how appropriate the final variables are for a Cox model.

From the Cox-Snell residual plot, we can see that the plotted line generally follows the 45° line, which indicates that the Cox model fits the data well.

Next, we can look at the Martingale residual plots for each covariate to evaluate if a possible transformation needs to be investigated. For tumor size, it doesn't appear like a transformation is necessary as the loess smooth curve is relatively linear. For the gleason index, the loess smooth curve is quite linear until 10, where afterwards the curve goes crazy. Given that it looks like a threshold value may be appropriate (say an index value above/below 10.5), this variable may have to be discretized. Lastly for treatment, the plot doesn't reveal anything as the variable is categorical and can't be transformed.

With the Schoenfeld residual plots for each variable, we can see how the proportional hazards assumption holds for each covariate. The plot for tumor size doesn't appear to have any issues as the residual points randomly fluctuate and gradually converge to 0. For the gleason index, the movement generally looks random but there is a residual that is quite far from 0. Overall though, there aren't any blatant issues with the plot. With the plot for treatment, the assumption looks like it holds since the two lines are more or less parallel. All three of these conclusions for the Schoenfeld residuals aren't very conclusive though, as there are only six data points included because most patients/observations were censored.