Fall 2024	**Survival Analysis**	**BST 222**
David M. Rocke		**October 22, 2024**

Homework Assignment 4
Due November 7, 2024

The addicts data set is from a study by Caplehorn et al. ("Methadone Dosage and Retention of Patients in Maintenance Treatment," Med. J. Aust., 1991). These data comprise the times in days spent by heroin addicts from entry to departure from one of two methadone clinics. There are two further covariates, namely, prison record and methadone dose, believed to affect the survival times.

The data set and R input code are on the website. The variables are as follows: **id:** Subject ID; **clinic:** Clinic (1 or 2); **status:** Survival status (0 = censored, 1 = departed from clinic); **time:** Survival time in days; **prison:** Prison record (0 = none, 1 = any); **methadone:** Methadone dose (mg/day).

1. Test proportionality of hazards with `cox.zph` on the model with clinic, prison, and methadone. Make the three plots vs. time of the Schoenfeld residuals for clinic, prison, and methadone. Interpret the results

2. Plot the cumulative hazard of the Cox-Snell residuals and interpret the apparent goodness of fit.

3. Plot the martingale residuals (omitting methadone) vs. methodone. The original paper using these data treated dose as a categorical variable with categories < 60, 60–79, and 80+. Make a categorized methadone variable using `cut`. Try fitting the model using categorized dose as a factor and compare to dose in statistical significance of the variable using the LR test (maybe with `drop1`) . Compare the two full models by AIC. Does it make an important difference to replace the quantitative variable by a 3-category factor? Note that you cannot compare the two models directly by a LR test since they are not nested.

4. Plot the martingale residuals vs. the linear predictor, the deviance residuals vs. the linear predictor, and the three dfbeta values by observation order. Identify possibly interesting observations and try to interpret the results. Since the clinic hazards are clearly not proportional, this is not the final answer, but it will do for the moment.