Survival Data Analysis (BIOS 7210)
Breheny

<div align="center">

Assignment 10

Due: Thursday, November 19

</div>

1. In class, we derived the relationship

$$\lambda_{ij} = 1 - (1 - \lambda_{0j})^{w_i}.$$

   Show that in the continuous case, this relationship still yields the usual Cox proportional hazards assumption in the limit. In other words, show that

   $$\lim_{\epsilon \to 0} \frac{1 - [1 - \{\Lambda_0(t + \epsilon) - \Lambda_0(t)\}]^{w_i}}{\epsilon} = \lambda_0(t)w_i.$$

2. (a) Under independent censoring, the ith subject contributes a $\lambda_i(t_i)^{d_i} S_i(t_i)$ term towards the likelihood, where λ_i and S_i are the hazard and survival functions for the ith subject. Show that, under the proportional hazards assumption, the ith contribution towards the log-likelihood is given by

   $$\ell_i(\eta_i) = d_i\eta_i - e^{\eta_i}\Lambda_0(t_i),$$

   up to a constant not depending on η_i.

   (b) Show that

   $$\max_{\eta_i} \ell_i(\eta_i) = -d_i \log \Lambda_0(t_i) - d_i.$$

   (c) In class, we used $\tilde{\ell}_i$ to denote the term from (b) and expressed the deviance residual in terms of $2(\tilde{\ell}_i - \ell_i)$. Show that

   $$\max_{\eta_i} \ell_i(\eta_i) - \ell_i(\hat{\eta}_i) = -\hat{m}_i - d_i \log(d_i - \hat{m}_i);$$

   in other words, the deviance residual can be written in terms of the martingale residual, $\hat{m}_i$, and the failure indicator d_i.

3. The course website contains a data set (Byar1980) from a randomized controlled trial in which men with stage 3 or stage 4 prostate cancer were assigned to either a therapy involving estrogen (at various doses) or a placebo. The outcome is all-cause mortality (time until death from any cause). The purpose of this problem is to assess the functional form of a few key covariates. For each covariate, (i) briefly describe why you chose the functional form that you did, (ii) present the estimate of its effect, along with confidence interval(s), and (iii) in one sentence, describe the nature of the covariate's effect on survival. For (ii), if the effect is linear, a hazard ratio will suffice. If the effect is nonlinear, a plot will be necessary.

   (a) Serum hemoglobin (hg)
   (b) Tumor size (sz)
   (c) Age (Age)
   (d) Serum prostatic acid phosphatase (ap)