

Survival Data Analysis (BIOS 7210)
Breheny

Assignment 7

Due: Thursday, October 22

1. Suppose that the hazard function $\lambda(y)$ for a variable Y satisfies

$$\log \lambda(y) = \alpha + \beta y.$$

Show that $T = e^Y$ follows a Weibull distribution. State the parameters of that Weibull distribution (λ and γ) in terms of α and β .

2. Suppose that we fit the Weibull regression model stated at the bottom of slide 18 on the 10-13 notes to obtain estimates $\hat{\alpha}$, $\hat{\beta}^*$, and $\hat{\sigma}$. Based on these estimates, we are interested in predicting the survival distribution of the time-to-event T for an individual with covariates \mathbf{x}_i . State what distribution T follows, along with the parameters of that distribution, in terms of $\hat{\alpha}$, $\hat{\beta}^*$, and $\hat{\sigma}$.
3. As we saw in last week's homework, the score test for the hypothesis $H_0 : \beta_j = 0$ is based on

$$u_j \sqrt{(\mathbf{I}^{-1})_{jj}} \sim N(0, 1),$$

where the score \mathbf{u} and information \mathbf{I} are both evaluated at $\beta_j = 0$ and where β_{-j} maximizes $\ell(\beta_{-j} | \beta_j = 0)$.

- (a) Suppose that the information matrix \mathbf{I} is partitioned according to

$$\mathbf{I} = \begin{bmatrix} I_{11} & \mathbf{I}_{12} \\ \mathbf{I}_{21} & \mathbf{I}_{22} \end{bmatrix},$$

where \mathbf{I} is a $p + 1 \times p + 1$ matrix, I_{11} is a scalar, and \mathbf{I}_{22} is a $p \times p$ matrix. Then

$$(\mathbf{I}^{-1})_{11} = (I_{11} - \mathbf{I}_{12} \mathbf{I}_{22}^{-1} \mathbf{I}_{21})^{-1}.$$

Suppose we fit a model with \mathbf{X} as a design matrix, thereby obtaining $\hat{\beta}$, \mathbf{W} , and $\mathbb{V}(\hat{\beta}) = (\mathbf{X}^T \mathbf{W} \mathbf{X})^{-1}$. We are considering adding a new variable, β^* . In terms of \mathbf{X} , \mathbf{W} , and $\mathbb{V}(\hat{\beta})$ from the original fit, give expressions for $(\mathbf{I}^{-1})^*$ and the score statistic for testing $H_0 : \beta^* = 0$. These expressions should not contain any matrix inverses.

- (b) Fit an exponential regression model to the `pbcc` data, with `stage` and `bili` as explanatory variables. Carry out score tests for the significance of `trt`, `hepato`, and `ascites`. Note that carrying out these tests can be performed without actually fitting a model with any of the terms present.