POISSON LOG-LINEAR MODEL

observed number of events during t_i units of time ($Y_i \in \mathbb{N}_0$)

$$Y_i \mid V_i, t_i \sim \mathrm{Pois}(\lambda_i \cdot t_i) \quad \text{where } \lambda_i = \exp\{\alpha^T V_i\}$$

set of regressors of i-th individual

set of unknown coefficients (including the intercept)

$$\lambda_i = \frac{E[Y_i \mid V_i, t_i]}{t_i}$$ expected number of events observed during a unit of time

$$t_i \cdot \lambda_i = \exp\{\log t_i + \alpha^T V_i\}$$ expected number of events in t_i units of time

an offset regressor with fixed α-coefficient!

Likelihood: $L(\alpha) = \mathrm{const} \cdot \prod_{i=1}^{n} [t_i \lambda_i]^{Y_i} \exp\{-\lambda_i t_i\}$

Log-likelihood: $\ell(\alpha) = \mathrm{const} + \sum_{i=1}^{n} \left[Y_i (\log t_i + \alpha^T V_i) - \exp\{\log t_i + \alpha^T V_i\} \right]$

Estimating equation $\frac{\partial}{\partial \alpha}$: $0 = \sum_{i=1}^{n} \left[\left(Y_i - \exp\{\log t_i + \alpha^T V_i\} \right) V_i \right]$

offset

model formula

response

(R): glm(Y ~ model formula + offset(log(time)), family = poisson) — default canonical log link

intercept term $\alpha_0 = \log(E[Y_i \mid V_i = (1, 0, 0, \dots, 0)])$

$$e^{\alpha_j} = \frac{t_i \cdot \exp\{\alpha^T V_i + \alpha^T e_j\}}{t_i \cdot \exp\{\alpha^T V_i\}} = \frac{E[Y_i \mid V_i + e_j]}{E[Y_i \mid V_i]}$$

proportional increase in EY_i per unit difference in covariate V_j during the same time period

EXPONENTIAL REGRESSION with ARBITRARY RANDOM CENSORING

observed censored time $X_i = \min\{T_i, C_i\}$ — censoring time independent of T_i

independent survival times (identically distributed if not influenced by any regressor)

$$T_i \mid Z_i \sim \mathrm{Exp}(\lambda(Z_i, \beta)), \quad \text{where } \lambda(Z_i, \beta) = \lambda_0 \cdot \exp\{\beta^T Z\}$$

$\lambda_0 > 0$

set of regressors of i-th individual which does not include intercept!

$= \lambda_i$ hazard function — constant given by Z_i and β (constant in X_i)

baseline hazard ($\Leftrightarrow Z_i = 0$)

$\delta_i = \mathbb{1}(T_i \le C_i)$ ← failure indicator

Likelihood: $L(\beta) \underset{\text{Exp continuous}}{\overset{\text{lecture}}{=}} \mathrm{const} \cdot \prod_{i=1}^{n} [\lambda(Z_i, \beta)]^{\delta_i} \cdot \exp\left\{ - \int_0^{X_i} \underbrace{\lambda(Z_i, \beta)}_{\text{const in } y} \, dy \right\}$

$$= \mathrm{const} \cdot \prod_{i=1}^{n} \left[\exp\{\log \lambda_0 + \beta^T Z_i\} \right]^{\delta_i} \cdot \exp\{-X_i \lambda_0 \exp\{\beta^T Z_i\}\}$$

Log-likelihood: $\ell(\beta) = \mathrm{const} + \sum_{i=1}^{n} \delta_i [\log \lambda_0 + \beta^T Z_i] - X_i \exp\{\log \lambda_0 + \beta^T Z_i\}$

$\beta_0^*, Z_{0,i}^* = 1$

Estimating equation $\frac{\partial}{\partial \beta^*}$: $0 = \sum_{i=1}^{n} \left[\left(\delta_i - \exp\{\log X_i + \beta^{*T} Z_i^*\} \right) \cdot Z_i^* \right]$

where $\beta^* = (\log \lambda_0, \beta^T)^T$ and $Z_i^* = (1, Z_i^T)^T$

as offset

is of the same form as the one on the left! Therefore,

(R): glm(Delta ~ model formula + offset(log(X)), family = poisson) — default canonical log-link

intercept term $\beta_0^* = \log \lambda_0 = \log \lambda(0, \beta)$ … log of baseline hazard

$$e^{\beta_j} = \frac{\lambda_0 \cdot \exp\{\beta^T Z_i + \beta^T e_j\}}{\lambda_0 \cdot \exp\{\beta^T Z_i\}} = \frac{\lambda(Z_i + e_j, \beta)}{\lambda(Z_i, \beta)}$$

proportional increase in the hazard per unit difference in covariate Z_j