3.10.7 → Poisson regression

→ Exposures

→ A poisson distribution models the frequency of an event w/in a specified scope. This scope is defined & measured in <u>units of exposure</u>.

→ If have non-constant exposures, then the model needs to take that into account

→ Poisson regression

→ Let $Y_1, \ldots, Y_n \overset{ii}{\sim} \text{Poisson}(\mu_i = a_i \lambda_i)$

where a_i is the exposure amount for the ith observation &
λ_i is the mean per exposure — — — — —

→ using a log link function, we get

$$\boxed{\begin{aligned} \ln(\mu_i) &= \ln(a_i \lambda_i) \\ &= \ln(a_i) + \ln(\lambda_i) \\ &= \ln(a_i) + x_i^T \beta \\ \Rightarrow \mu_i &= a_i \exp\{x_i^T \beta\} \end{aligned}}$$

→ This means the predictors do not directly explain the Poisson mean at the observation level. Instead, the predictors explain the the Poisson mean <u>per exposure</u>, which is then multiplied by an exposure amount to compute the poisson mean for an observation

→ The term $\ln(a_i)$ is called an <u>offset</u> since it adds to the usual linear component. One way to interpret the offset is as a predictor w/ a regression coefficient of 1.

→ Then, we obtain the following simplified expressions:

$$\hat{\mu}_i = a_i \exp\{x_i^T \beta\}$$

→ Log likelihood function → $\ell(\beta) = \sum_{i=1}^{n} [y_i \ln(\mu_i) - \mu_i - \ln(y_i!)]$

→ Score function → $u_j = \sum_{i=1}^{n} (y_i - \mu_i) x_{ij}$

→ Information matrix → $J = \sum_{i=1}^{n} \mu_i x_i x_i^T$

→ Deviance → $D = 2 \sum_{i=1}^{n} y_i \ln\left(\frac{y_i}{\hat{\mu}_i}\right) - (y_i - \hat{\mu}_i)\}$

$= 2 \sum_{i=1}^{n} y_i \ln\left(\frac{y_i}{\hat{\mu}_i}\right)$ ← simplified formula ?

→ Pearson residual → $e_i^P = \frac{y_i - \hat{\mu}_i}{\sqrt{\hat{\mu}_i}}$

→ Pearson chi-square statistic → $\sum_{i=1}^{n} (e_i^P)^2 = \sum_{i=1}^{n} \frac{(y_i - \hat{\mu}_i)^2}{\hat{\mu}_i}$

→ Deviance residual → $e_i^D = \pm \sqrt{2\left[y_i \ln\left(\frac{y_i}{\hat{\mu}_i}\right) - (y_i - \hat{\mu}_i)\right]}$

→ Interpretation of parameters

→ In general, the estimated mean changes by factor of $\exp(\hat{\beta}_j)$ per unit increase in x_j, assuming all other predictors are held constant

3.10.8 → Log-linear models

? (come back to it if time)