ADVANCED BAYESIAN MODELING

Rate Models

Consider responses y that are counts of some kind of event.

Typically, there is a potential explanatory variable t representing the amount of exposure to such events (e.g. time, population, or spatial extent)

Examples:

- $y =$ times a player scores, $t =$ total playing time of player
- $y =$ cases of rare disease in country, $t =$ population of country
- $y =$ number of bird nests on island, $t =$ area of island

Instead of modeling the mean of y, often want to model the *rate* per unit of t:

$$\mathrm{E}(y) \,/\, t$$

Examples:

- Player scoring rate per minute of playing time
- Cases of disease per million people
- Number of bird nests per square kilometer

In this case, we use a **rate model** ...

Poisson Rate Model

$$y_i \mid \beta, t_i, X_i \quad \sim \quad \text{indep. Poisson}(t_i\, r_i)$$

$$\log r_i \;=\; X_i\beta \qquad\qquad r_i \;=\; e^{X_i\beta}$$

Notes:

- r_i is rate of events per unit of t
- t_i should *not* be part of X_i
- Need all $t_i > 0$

In the rate model:

$$\begin{aligned} \log \mathrm{E}(y_i) &= \log(t_i\, r_i) &&= \log t_i + X_i\beta \\ &&&= \log t_i + \beta_1 x_{i1} + \cdots + \beta_k x_{ik} \end{aligned}$$

Notice: $\log t$ is like an explanatory variable with a fixed (known) coefficient of 1.

Such a fixed term is sometimes called an **offset**.

An offset is *not* an intercept – the model probably still needs an intercept (either explicit or implicit).