4.3. GLMs for Counts

4.3.1. Poisson Loglinear Model

\( Y \sim \text{Poisson}(\mu) \), \( \mu > 0 \)

Use the log link.

why: (i) \( \log(\mu) \) can take any real value

like \( \alpha + \beta x \)

(ii) \( \log(\mu) \) is the natural parameter, thus

log link is the canonical link

\[
\log \mu = \alpha + \beta x
\]

\[
\mu = \exp(\alpha + \beta x)
\]

\[
\frac{\mu(x+1)}{\mu(x)} = \frac{e^\beta e^{\beta(x+1)}}{e^\beta e^{\beta x}} = e^\beta
\]

or \( \mu(x+1) = e^\beta \mu(x) \)

4.3.2. Horseshoe Crab Mating data

width of female crab vs. # of satellites

\[
\log \hat{\mu} = -3.305 + .164 x
\]

\[
\hat{\mu} = e^{\log \hat{\mu}} = e^{-3.305 + .164 x} = e^{-.33 + .164 x}
\]

1 cm increase in width yields an 18% increase in the estimated mean

Compare w/ the Poisson GLM w/ identity link

\[
\hat{\mu} = \hat{\alpha} + \beta x = -11.53 + .55 x
\]

4.3.3. Overdispersion

Source: mixture \( Y \mid X \sim \text{Poisson}(\mu(x)) \)

But if \( \mu(x) \) is random, then overdispersion arises

Consequence: invalid results
4.3.4. Negative binomial GLMs

4.3.6. Poisson GLM of independence in 2xJ contingency tables

\[ M_{ij} = \mu \alpha_i \beta_j \]

\[ \alpha_i, \beta_j > 0, \quad \sum \alpha_i = \sum \beta_j = 1 \]

\[ E_M = \lambda + \alpha_i * + \beta_j * \]

Poisson model for independence