

By Special Request: Firth's Penalized Likelihood!

Epid 814

Firth's Penalized Likelihood

- Way of handling separability, small sample sizes, and parameter estimate bias
- Penalizes the likelihood function based on the Fisher information

Separation & Monotone Likelihood

- When outcome values (e.g. response vs. non-response) can be perfectly separated by a single risk factor or by a non-trivial linear combination of risk factors
 - E.g. if all cases have purple hair and all non-cases do not
- This results in a monotone likelihood—leads to infinite or near infinite parameter estimates
- With smaller sample sizes, this can happen even for relatively low effect sizes

Firth's Penalized Likelihood

- For regression parameters $\theta = (\theta_1, \dots, \theta_n)$, the penalized likelihood is:

$$\mathcal{L}(\theta)^* = \mathcal{L}(\theta) \cdot \sqrt{|I(\theta)|}$$

- $\sqrt{|I(\theta)|}$ is the Jeffrey's prior
- Firth's method is equivalent to maximum *a posteriori* with a Jeffrey's prior
- Helps to handle the bias from small sample sizes; prevents parameter estimates from going to infinity