This Poisson time inference
next time: vector

A note on conditional probability:

\[ P(A \mid BC) = \begin{cases} \frac{P(ABC)}{P(BC)} & \text{if } P(BC) > 0 \\ \text{undefined} & \text{if } P(BC) = 0 \end{cases} \]

if either \( P(B=0) \) or \( P(C=0) \), \( P(A \mid BC) = \text{undefined} \)

\[ P = (\theta, C) \stackrel{\text{maximize}}{\rightarrow} (\hat{\theta}, \hat{C}) \]

problem question(s)

\[ M = \{ P(\theta \mid B), P(D \mid \theta, B) \} \]

\[ \text{inference, prediction, statistical model} \]
\( M^* = \{ p(0|B), p(1|B B), (\rho|B), u(0, 0|B) \} \)

Inference, prediction, decision

**Theorem:** If can uniquely specify \( M^* \), then it is optimal inference & prediction.

**Theorem:** If \( M^* \), then it is optimal inference & decision.

It would be neat if \( C + B \) always uniquely specifies \( p(0|B) \) but it's actually somewhat for unique choices \( u(0, 0|B) \), of these 4 things to arise from \( C + B \).
ex. AMI case study:

\[ (I_i \mid \theta, B) \sim \text{Bernoulli}(\theta) \]
\[ (i = 1, \ldots, n) \]

This is uniquely determined by $C : \{ \text{defining } H_i's \text{ then} \}$. Before $D = ( y_1, \ldots, y_n )$ arrives, your uncertainty about the $y_i$ is exchangeable:

If \[ \theta \sim \text{Bernoulli}(\theta) \] but $p(\theta \mid B)$ was not unique in this case study.

In general, we have uncertainty about $(M, M^*)$: model uncertainty.
\[ u = x^2 + y^2 = 12 \]

\[ a = 1 \]

\[ b = 2 \]

\[ c = 3 \]

\[ \theta = \frac{\pi}{2} \]

\[ \text{joint dist.} \]

\[ (x_1, x_2, \ldots, x_n) \]

\[ (y_1, y_2, \ldots, y_n) \]
\[ \frac{d}{d\lambda} \mathbb{E} (z | \lambda) = \frac{5}{\lambda} - \frac{5}{\lambda^2} \]

\[ 2 = \hat{\lambda}_{\text{MLE}} \]

\[ \begin{bmatrix} \frac{d^2}{d\lambda^2} \mathbb{E} (z | \lambda) \end{bmatrix} \]

\[ x > 0 \quad \text{local max.} \]

\[ I(\hat{\lambda}_{\text{MLE}}) = \frac{\lambda}{\hat{\lambda}^2} \]

\[ \hat{\lambda}_{\text{MLE}} = \frac{5}{n} = o(n) \checkmark \]

\[ \text{Var}(\hat{\lambda}_{\text{MLE}}) = \frac{1}{I(\hat{\lambda})} \]

\[ \frac{7}{5} = o\left( \frac{1}{n} \right) \]
\[ SE_{\hat{\theta}}(\hat{\theta}_{MLE}) = \sqrt{\hat{\theta}_{MLE} - \hat{\theta}_{MLE}^2} \]

\[ = O\left(\frac{1}{\sqrt{n}}\right) \quad \text{for \ (n \ large)} \quad \text{CLT} \]

approx. 95% CI for \[ \hat{\theta}_{MLE} \pm 1.96 \ SE(\hat{\theta}_{MLE}) \]

\[ 100(1-\alpha)% \]

\[ \text{Bivariate Poisson} \]

\[ \mathcal{P}(r|\alpha) \sim \frac{\alpha^r}{r!} e^{-\alpha} \]

\[ \frac{\mathcal{P}(r|\alpha)}{\mathcal{P}(r'|\alpha)} = \frac{\alpha^r}{\alpha^{r'}} \frac{r!}{r!'!} e^{-\alpha} \]

\[ \alpha(\tilde{y}) = c \tilde{y} e^{-\tilde{y}} \quad \text{is \ sufficient} \]

\[ p(x) = c x^{\alpha-1} e^{-\beta x} \quad \text{for x } \]
\[ F(z) = c \left[ \exp \left( -\frac{z}{\theta} \right) \right] \left[ \exp \left( -\frac{\theta}{z} \right) \right] \]

post.

\[ -c \leq \frac{\theta}{z} \leq (d+s)-1 \quad \text{lik.} \]

\[ = \int (d+s, \beta+n) \]

\[ (X_i \mid \beta, \theta) \sim \Gamma(\alpha, \beta) \]

\[ (Y_i \mid \alpha, \beta, \theta) \sim \text{Poisson} (\theta) \]

\[ (i = 1, \ldots, n) \]

\[ + \prod_{i=1}^{n} \Gamma (\alpha_i, \beta_i, \theta) \sim \Gamma (d+s, \beta+n), \]

\[ s = \sum_{i=1}^{n} y_i \]

Principle of Stable Estimation (Edwards, Savage, Lindeman)
Any prior that is close to flat in the region in which the likelihood is appreciable will have low information content.