Approximate confidence intervals in ML estimation

Model: $y=\{y_1,\dots,y_n\}$, obs, $\theta=\{\theta_1,\dots,\theta_p\}$, parameters

Likelihood $L(\theta;y)=p(y|\theta)$

Fisher information matrix: $\tilde{H}(\theta)=E_{y|\theta}\Big[-\underbrace{\nabla_\theta^2 \log[p(y|\theta)]}_{H(\theta)}\Big]$

Theorem: Under mild assumptions,

$$\hat{\theta}_{ML}-\theta \overset{d}{\approx} N(0,\tilde{H}(\theta)^{-1})$$

When $p=1$, $\theta=\theta_1$

$$\frac{\hat{\theta}_{ML}-\theta}{Var(\hat{\theta}_{ML})^{1/2}} \xrightarrow{d} N(0,1)$$

and $Var(\hat{\theta}_{ML})=\left[\tilde{H}(\theta)^{-1}\right]_{11}$

Ex) $y_1,\dots,y_n \overset{\text{indep}}{\sim} Po(\lambda)$

Let $\theta=\lambda$

$p(y|\theta)=e^{-\lambda n}\prod_{i=1}^{n}\lambda^{y_i}/(y_i!)$, $\log p(y|\theta)=-\lambda n+\sum_{i=1}^{n} y_i\log(\lambda)+\text{const}$

$\frac{\partial}{\partial\lambda}\log p(y|\theta)=-n+\frac{1}{\lambda}\sum_{i=1}^{n}y_i$, $\frac{\partial^2}{\partial\lambda^2}\log p(y|\theta)=-\frac{1}{\lambda^2}\sum_{i=1}^{n}y_i$

$\tilde{H}(\lambda)=\frac{n}{\lambda}$, $Var(\hat{\lambda}_{ML})\approx\frac{\lambda}{n}$

$$\hat{\lambda}_{ML}=\frac{1}{n}\sum_{i=1}^{n}y_i$$

$$\frac{\hat{\lambda}-\lambda}{Var(\hat{\lambda})^{1/2}}\approx\frac{\hat{\lambda}-\lambda}{\sqrt{\lambda/n}}\approx\frac{\hat{\lambda}-\lambda}{\sqrt{\hat{\lambda}/n}}\sim N(0,1)$$

$$P\left(z_{\alpha/2}<\frac{\hat{\lambda}-\lambda}{\sqrt{\hat{\lambda}/n}}<z_{1-\alpha/2}\right)\approx 1-\alpha$$

$$P\left(\hat{\lambda}-\sqrt{\hat{\lambda}/n}\,z_{1-\alpha/2}<\lambda<\hat{\lambda}-\sqrt{\hat{\lambda}/n}\,z_{\alpha/2}\right)=1-\alpha$$

$CI_\lambda=\left(\hat{\lambda}-\sqrt{\hat{\lambda}/n}\,z_{1-\alpha/2},\ \hat{\lambda}-\sqrt{\hat{\lambda}/n}\,z_{\alpha/2}\right)$ or $(0,\hat{\lambda}-\sqrt{\hat{\lambda}/n}\,z_{\alpha/2})$ when Left <0

$CI_{\lambda^2}=\left((\ \ -''-\ \ )^2,(\ \ -''-\ \ )^2\right)$ (since $\cdot^2$ is increasing and $\lambda>0$)

Alternative parametrisation:

$\theta=\sqrt{\lambda}$

$$\frac{\hat{\theta}_{ML}-\theta}{Var(\hat{\theta}_{ML})^{1/2}}\sim N(0,1) \text{ for large } n$$

$$p(y|\theta) = e^{-\theta^2 n} \prod_{i=1}^{n} \theta^{2y_i} / const \quad , \quad \log p(y|\theta) = -\theta^2 n + \sum_{i=1}^{n} 2y_i \log\theta + const$$

$$\frac{\partial}{\partial\theta} = -2\theta n + \frac{2}{\theta} \sum_{i=1}^{n} y_i \quad , \quad \frac{\partial}{\partial\theta^2} = -2n - \frac{2}{\theta^2} \sum_{i=1}^{n} y_i$$

$$\hat{\theta}_{ML} = \sqrt{\frac{1}{n} \sum_{i=1}^{n} y_i} \quad , \quad \hat{I}(\theta) = +4n$$

$$P\left(z_{\alpha/2} < \frac{\hat{\theta}_{ML} - \theta}{\sqrt{1/(4n)}} < z_{1-\alpha/2}\right) = P\left(\hat{\theta}_{ML} - \frac{1}{2\sqrt{n}} z_{1-\alpha/2} < \theta < \hat{\theta}_{ML} - \frac{1}{2\sqrt{n}} z_{\alpha/2}\right) = 1-\alpha$$

$$CI(\lambda) = \left(\max\left(0, \hat{\theta}_{ML} - \frac{1}{2\sqrt{n}} z_{1-\alpha/2}\right)^2 < \lambda < \left(\hat{\theta}_{ML} - \frac{1}{2\sqrt{n}} z_{\alpha/2}\right)^2\right)$$