Maximum Likelihood

Lab 2

Elad Zippory

Quant III

September, 11, 2014

Agenda

Jensen's inequality

- The invariance property: if $\hat{\theta}$ is MLE of θ_o, then $\tau(\hat{\theta})$ is MLE of $\tau(\theta_0)$

Jensen's inequality

- The invariance property: if $\hat{\theta}$ is MLE of θ_o, then $\tau(\hat{\theta})$ is MLE of $\tau(\theta_0)$
- And $E[\hat{\theta}] = \theta_0$.

Jensen's inequality

- The invariance property: if $\hat{\theta}$ is MLE of θ_o, then $\tau(\hat{\theta})$ is MLE of $\tau(\theta_0)$
- And $E[\hat{\theta}] = \theta_0$.
- But, unbiasedness is not transferable: $E[\hat{\theta}] = \theta_0 \nRightarrow E[\tau(\hat{\theta})] = \tau(\theta_0)$

Jensen's inequality

- The invariance property: if $\hat{\theta}$ is MLE of θ_o, then $\tau(\hat{\theta})$ is MLE of $\tau(\theta_0)$
- And $E[\hat{\theta}] = \theta_0$.
- But, unbiasedness is not transferable: $E[\hat{\theta}] = \theta_0 \nRightarrow E[\tau(\hat{\theta})] = \tau(\theta_0)$
- So, what is the direction of the bias?

Jensen's inequality

- The invariance property: if $\hat{\theta}$ is MLE of θ_o, then $\tau(\hat{\theta})$ is MLE of $\tau(\theta_0)$
- And $E[\hat{\theta}] = \theta_0$.
- But, unbiasedness is not transferable: $E[\hat{\theta}] = \theta_0 \nRightarrow E[\tau(\hat{\theta})] = \tau(\theta_0)$
- So, what is the direction of the bias?
- Jensen's Inequality says that if $\tau(\cdot)$ is convex, then $\tau(E[\hat{\theta}]) \leq E[\tau(\hat{\theta})]$; if concave, the inequality changes direction.

Expected Likelihood

- In asymptotia, $E[\hat{\theta}] = \theta_0$

Expected Likelihood

- In asymptotia, $E[\hat{\theta}] = \theta_0$
- Can we simulate this? open R code named `expected_likelihood`

Identification of Parameters

- When you did fixed effects, could you estimate the overall constant?

Identification of Parameters

- When you did fixed effects, could you estimate the overall constant?
- A trivial example: $Y \sim N(\mu_1 + \mu_2, \sigma)$

Identification of Parameters

- When you did fixed effects, could you estimate the overall constant?
- A trivial example: $Y \sim N(\mu_1 + \mu_2, \sigma)$
- What happens in MLE? open R code named `identification`

Analytical results and the Newton-Raphson algorithm

- The joint density is $f(y_1, ..., y_n) = \prod \frac{\lambda^{y_i} e^{-\lambda}}{y_i!}$

Analytical results and the Newton-Raphson algorithm

- The joint density is $f(y_1, ..., y_n) = \prod \frac{\lambda^{y_i} e^{-\lambda}}{y_i!}$
- The log likelihood is $log\lambda \sum y_i - n\lambda - \sum logy_i!$

Analytical results and the Newton-Raphson algorithm

- The joint density is $f(y_1, ..., y_n) = \prod \frac{\lambda^{y_i} e^{-\lambda}}{y_i!}$
- The log likelihood is $log\lambda \sum y_i - n\lambda - \sum logy_i!$
- The gradient is $\frac{\sum y_i}{\lambda} - n$

Analytical results and the Newton-Raphson algorithm

- The joint density is $f(y_1, ..., y_n) = \prod \frac{\lambda^{y_i} e^{-\lambda}}{y_i!}$
- The log likelihood is $log\lambda \sum y_i - n\lambda - \sum logy_i!$
- The gradient is $\frac{\sum y_i}{\lambda} - n$
- The Hessian is $\frac{-\sum y_i}{\lambda^2}$

Analytical results and the Newton-Raphson algorithm

- The joint density is $f(y_1, ..., y_n) = \prod \frac{\lambda^{y_i} e^{-\lambda}}{y_i!}$
- The log likelihood is $log\lambda \sum y_i - n\lambda - \sum logy_i!$
- The gradient is $\frac{\sum y_i}{\lambda} - n$
- The Hessian is $\frac{-\sum y_i}{\lambda^2}$
- The Newton-Raphson algorithm is $\lambda_{t+1} = \lambda_t - \frac{l'(\lambda_t)}{l''(\lambda_t)}$

Analytical results and the Newton-Raphson algorithm

- The joint density is $f(y_1, ..., y_n) = \prod \frac{\lambda^{y_i} e^{-\lambda}}{y_i!}$
- The log likelihood is $log \lambda \sum y_i - n\lambda - \sum log y_i!$
- The gradient is $\frac{\sum y_i}{\lambda} - n$
- The Hessian is $\frac{-\sum y_i}{\lambda^2}$
- The Newton-Raphson algorithm is $\lambda_{t+1} = \lambda_t - \frac{l'(\lambda_t)}{l''(\lambda_t)}$
- Let's do it! How would you start? (solution will be posted)