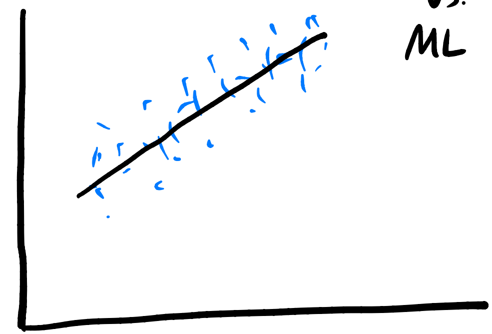
Lecture Note 6: Maximum Likelihood

Random var X with pdf $f(x;\theta)$

$\{X_i\}_{i=1}^{N}$ iid, each i has realization x_i

LS vs. ML



Likelihood:

$$\max_{\tilde{\theta}} L = \max_{\tilde{\theta}} \prod_{i=1}^{N} f(x_i;\tilde{\theta})$$

$$\max_{\tilde{\theta}} \ln L = \max_{\tilde{\theta}} \sum_{i=1}^{N} \ln\left[f(x_i;\tilde{\theta})\right]$$

Solution $\hat{\theta}^{MLE}$ satisfy $\frac{\partial \ln L}{\partial \tilde{\theta}} = 0$.

Properties:

1. Consistency: $\hat{\theta}^{MLE} \xrightarrow{p} \theta$
2. Asymptotic normality (CLT): $\hat{\theta}^{MLE} \xrightarrow{d} N(\theta, \Sigma)$
3. Asymptotic efficiency

Bernoulli $X_i = \begin{cases} 1 & \text{w.p. } p \\ 0 & \text{w.p. } 1-p \end{cases}$

Sample with 3 iid observations: $(1,1,0)$

$$L = \Pr[X_1=1] \cdot \Pr[X_2=1] \cdot \Pr[X_3=0]$$

$$= p \cdot p \cdot (1-p)$$

$$= p^2(1-p)$$

MLE: $\max_{\tilde{p}} L = \max_{\tilde{p}} \tilde{p}^2(1-\tilde{p})$

ln MLE: $\max_{\tilde{p}} \ln L = \max_{\tilde{p}} 2\ln(\tilde{p}) + \ln(1-\tilde{p})$

FOC: $\frac{d \ln L}{d\hat{p}} = \frac{2}{\hat{p}} - \frac{1}{1-\hat{p}} = 0 \Rightarrow \hat{p} = \frac{2}{3}$

Sample of size N (iid), S successes, $N-S$ failures

$$L = p^{S}(1-p)^{N-S}$$

$$\ln L = S\ln(p) + (N-S)\ln(1-p)$$

$$\hat{p} = \frac{S}{N}$$

Approaches to MLE

① Analytic optimization

② ~~(Undirected) grid search~~

③ (Directed) numerical optimization