**STAT 801: Mathematical Statistics**

**Optimality theory for point estimates**

Why bother doing the Newton Raphson steps?

Why not just use the method of moments estimates?

Answer: method of moments estimates not usually as close to right answer as MLEs.

**Rough principle**: A good estimate $\hat{\theta}$ of $\theta$ is usually close to $\theta_0$ if $\theta_0$ is the true value of $\theta$. Closer estimates, more often, are better estimates.

This principle must be quantified if we are to "prove" that the mle is a good estimate. In the Neyman Pearson spirit we measure average closeness.

**Definition**: The **Mean Squared Error** (MSE) of an estimator $\hat{\theta}$ is the **function**

$$MSE(\theta) = E_\theta[(\hat{\theta} - \theta)^2]$$

Standard identity:

$$MSE = \text{Var}_\theta(\hat{\theta}) + Bias^2_{\hat{\theta}}(\theta)$$

where the bias is defined as

$$Bias_{\hat{\theta}}(\theta) = E_\theta(\hat{\theta}) - \theta \,.$$

**Primitive example**: I take a coin from my pocket and toss it 6 times. I get $HTHTTT$. The MLE of the probability of heads is

$$\hat{p} = X/n$$

where $X$ is the number of heads. In this case I get $\hat{p} = \frac{1}{3}$.

Alternative estimate: $\tilde{p} = \frac{1}{2}$.

That is, $\tilde{p}$ ignores data; guess coin is fair.

The MSEs of these two estimators are

$$MSE_{\text{MLE}} = \frac{p(1-p)}{6}$$

and

$$MSE_{0.5} = (p - 0.5)^2$$

If $0.311 < p < 0.689$ then 2nd MSE is smaller than first.

For this reason I would recommend use of $\tilde{p}$ for sample sizes this small.

Same experiment with a thumbtack: tack can land point up (U) or tipped over (O).

If I get $UOUOOO$ how should I estimate $p$ the probability of $U$?

Mathematics is identical to above but is $\tilde{p}$ is better than $\hat{p}$?

Less reason to believe $0.311 \le p \le 0.689$ than with a coin.