Recall that the probability of tosing a coin N times and observing x "heads" is determined from the binomail probability mass function

$$Pr(x|p) = \frac{N!}{x!(N-x)!}\, p^x (1-p)^{N-x}$$

where $p = Pr(\text{"heads"})$***. Say that we suspect a coin is not a "fair" coin, and so we wish to estimate this probability*** p ***by tossing it*** N ***times and observing the numner of heads*** x***.***

1. Determine the MLE estimate $\hat{p}(x)$***.***

2. Determine the bias of this esimate.

3. Determine the estimator/error variance of this estimate.

4. Compute the Fisher's Information and the CRLB. Use this to determine whether the MLE estimate is efficient.

MLE Estimate

Given the the binomial distribution of x heads occuring when flipping a unfair coin N times, the likelihood of function relating probability p to the number of heads observed is thus this same distribution.

$$L(x|p) = \frac{N!}{x!(N-x)!}\, p^x (1-p)^{N-x}$$

To find the maximum likelihood estimate (MLE), we must identify the extrema of this likelihood function and prove which one is a maximum. To do this, the natural logarithm is applied (Monotonically increase it; it does not affect the maximum), and then the derivative is taken and set equal to zero.

$$\ln(L(x|p)) = \ln\left(\frac{N!}{x!(N-x)!}\right) + x\ln(p) + (N-x)\ln(1-p)$$

$$\frac{\partial \ln(L(x|p))}{\partial p} = \frac{x}{p} - \frac{N-x}{1-p} = 0$$

Subtracting the second term to the right hand side of the equation and cross multiplying gives the only extrema and therefore the MLE.

$$x(1-p) = p(N-x) \quad \longrightarrow \quad x - xp = pN \quad \longrightarrow \quad \hat{p}_{MLE} = \frac{x}{N}$$

MLE Bias

Defining the estimator bias as $\text{bias}(\hat{p}_{MLE}) = \mathbb{E}[\hat{p}_{MLE}] - p$ and given that the expected value under the binomial distribution of x is $\mathbb{E}[x] = Np$, the bias of estimate $\hat{p}_{MLE}$ is found to be

$$\text{bias}(\hat{p}_{MLE}) = \mathbb{E}[\frac{x}{N}] - p = \frac{1}{N}(Np) - p = p - p = 0.$$

Estimator Variance

Similar to the calculation of the estimator bias, using the given variance for a binomial distribution as $\text{Var}(x) = Np(1-p)$, the variance of estimator $\hat{p}_{MLE}$ is therefore

$$\text{Var}(\hat{p}_{MLE}) = \frac{1}{N^2}Np(1-P) = \frac{p(1-p)}{N}$$

Estimator Efficiency

Using the log-likelihood definition of Fisher's Information found in lecture 39, and the previously computed log-likelihood function from part 1, Fisher's Information is expressed as

$$\mathbb{I}(p) = -\mathbb{E}[\frac{d^2}{dp^2}\ln(L(p|x))] = -\mathbb{E}[\frac{d^2}{dp^2}\left(\frac{x}{p} - \frac{N-x}{1-p}\right)] = \mathbb{E}[\frac{x}{p^2} + \frac{N-x}{(1-p)^2}].$$

Again, using the given definiation of $\mathbb{E}[x] = Np$, the expected value of $\mathbb{E}[N-x]$ is therefore $N - \mathbb{E}[x] = N(1-p)$. Building upon this intuition, Fisher's Information is computed as

$$\mathbb{I}(p) = \frac{\mathbb{E}[x]}{p^2} + \frac{\mathbb{E}[N-x]}{(1-p)^2} = \frac{N}{p} + \frac{N}{1-p} = \frac{N}{p(1-p)}.$$

Using this value, the Cramer-Rao Lower Bound (CRLB) theorm states that an efficient estimator is one with a variance that is the inverse of this informational value. In part 3, we found the estimator variance to be

$$\text{Var}(\hat{p}_{MLE}) = \frac{p(1-p)}{N}$$

which when inverted is directly equal to Fisher's informational value and therefore since $\hat{p}_{MLE}$ is also unbiased, it is an efficient estimator.

$$(\text{Var}(\hat{p}_{MLE}))^{-1} = \frac{N}{p(1-p)} = \mathbb{I}(p)$$