Bayes

Bayesian Theorem:

$$P(A|B) = \frac{P(AB)}{P(B)} = \frac{P(B|A)P(A)}{P(B)}$$

Where,

P(A): the Prior Probability, i.e. the probability that fair A happens in general

P(B): the Marginal Probability, i.e. the general probability for the given restriction B, $P(B) = P(\Sigma A_i B) = \Sigma P(A_i B)$

P(B|A): the Likelihood Probability. i.e. if A happens, then the probability that B also happens

P(A|B): the Posterior Probability

Maximum Likelihood Estimation *v.s* Maximum A Posterior Estimation

MLE: without considering about the prior probability. It assumes that all prior fairs happen in a same probability

$$\theta_{MLE} = argmax_\theta logP(X \,\big|\, \theta)$$

$$= argmax_\theta log \prod_i P(x_i|\theta) = argmax_\theta \sum_i logP(x_i|\theta)$$

MAP: based on the Bayesian theorem, considering about the prior probability for the fair to be estimated

$$\begin{aligned}\theta_{MAP} &= argmax_\theta P(X|\theta)P(\theta)\\ &= argmax_\theta logP(X|\theta)P(\theta)\\ &= argmax_\theta log \prod_i P(x_i|\theta)P(\theta)\\ &= argmax_\theta \Sigma_i logP(x_i|\theta)P(\theta)\end{aligned}$$

When the prior fair (X)with all its possible outcomes (x_i) happen uniformly

with a constant probability, then $\theta_{MLE} = \theta_{MAP}$