Frequentist v.s. Bayesian

Usually, we face problems that given samples to find out the the hypothesis conditions (parameters). In statistics, that is to maximize:

$$P(Hypotheses|Samples)$$

In general, there're 2 major parties for this type of question: Frequentist and Bayesian.

Frequentist

Frequentist inference based on likelihood, and does not require a prior hypothesis on probability, which is actually the MLE routine, Maximum Likelihood Estimation.

$$L(\theta|x_i) = \prod_{\{i=1\}}^{n} P(x_i|\theta)$$

To maximize the likelihood is to maximize its 'log' transformation:

$$\log\big(L(\theta|x_i)\big) = \log\left(\prod_i P(x_i|\theta)\right) = \sum_i \log\big(P(x_i|\theta)\big)$$

take partials with respect to θ and set it equal to 0:

$$\frac{\partial log(L(\theta|x_i))}{\partial\theta} = \frac{\partial log\prod_i P(x_i|\theta)}{\partial\theta} = \frac{\partial\sum_i logP(x_i|\theta)}{\partial\theta} = 0$$

simplify the function and get the MLE estimator for θ.

Therefore, for MLE estimator:

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

Bayesian

Bayesian inference is based on Bayes' Theorem, where it considers about probabilities for both samples and hypotheses, thus requires the probability for prior hypothesis. It follows a MAP (Maximum A Posteriori Estimation) routine.

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