prior = $P(H)$. posterior = $P(H/D)$.

there are 10 hypothesis. what is the probability of h being the underlying hypothesis?

prior = $P(H=h)$ or $P(h)$.

this is computed without looking at the data

what is the probability of the Data D to occur given that the underlying hypothesis is h? dn

$$= \prod_{d=d_1} P(D=d \mid H=h)$$ ( $\Pi$ -product)

or

$P(D/h)$

this is likelihood

$P(h/D) \propto$ likelihood x prior

as per bayesian theorem

$$p(h/D) = \frac{p(D/h) * \overbrace{p(h)}^{\text{prior}}}{p(D)}$$

$p(h/D)$ → posterior

$p(D/h)$ → likelihood

$p(D)$ → same for all h. so normalization.

$p(D)$ can also be written as

$$p(D) = \int_{\text{Domain of } h} p(D/h) * p(h) \; dh$$

Given the Data-D, given that we chose hypothesis = h,
What is $P(X=x)$?

$$P(X=x \mid D) = P(X=x \mid H=h)$$

here h could be chosen based on
(1) max prior
(2) max likelihood
(3) max posterior

What if we use all the hypothesis weighted by posterior?

$$P(X=x \mid D) = \sum_{h=1}^{n} (X=x \mid h)\, P(h \mid D)$$

or $h \in H$

Bayesian Averaging ↑
saves in cases where posterior prob of multiple hypothesis are close by.