Generative Models for discrete data:

feature vector : $\underline{X}$ → classify using Bayes rule

generative classifier: of form:

$$p(y=c \mid \underline{x}, \theta) \propto p(\underline{x} \mid y=c, \theta)\, p(y=c \mid \theta)$$

Posterior predictive distribution:

$p(\hat{x} \mid D)$ $\hat{x} \in C$ {its prob.} given $\hat{x} \in \{1, 2, \ldots, 100\}$ D;

Likelihood: $p(D \mid h) = \left[\frac{1}{size(h)}\right]^N$

Prior: low prior for unnatural concept $p(h)$

Posterior: likelihood times prior, normalized

$$p(h \mid D) = \frac{p(D \mid h)\, p(h)}{\sum_{h' \in H} p(D, h')} = \frac{p(h)\, \mathbb{I}(D \in h) / |h|^N}{\sum_{h' \in H} p(h')\, \mathbb{I}(D \in h') / |h'|^N}$$

$\mathbb{I}(D \in h)$ is.

>

MAP estimate:-

$$p(h \mid D) \to \delta_{\hat{h} MAP}(h)$$

~~$\hat{h}$ MAP~~ $\hat{h}$ MAP $= \underset{h}{\operatorname{argmax}}\, p(h \mid D)$ = posterior mode

δ = Dirac measure

$$\delta_x(A) = \begin{cases} 1 & \text{if } x \in A \\ 0 & \text{if } x \notin A \end{cases}$$

MLE (Maximum likelihood estimate)

$$\hat{h}_{mle} \triangleq \underset{h}{\operatorname{argmax}}\, p(D \mid h) = \underset{h}{\operatorname{argmax}} \log p(D \mid h)$$

Next ✱ The Beta Binomial Model