MLE and MAP

CS 3630

Perception

We’ve seen a lot of probability theory in the last few slide decks. How can we use these results to make inferences about the state of the world?

- Maximum Likelihood Estimation – simply use the likelihood
- MAP (Maximum A Posteriori) Estimation – Maximize the posterior given the sensor reading.

We’ll look now at each of these.

Maximum likelihood estimation

Given S = sensor reading, C = object category, what is the most likely object category given some sensor reading?

$$C^* = \arg\max_{\mathrm{C}} P(S|C)$$

in which the maximization is done w.r.t. the set

$$\mathrm{C} = \{\mathrm{Cardboard}, \mathrm{Paper}, \mathrm{Can}, \mathrm{Scrap\ Metal}, \mathrm{Bottle}\}$$

NOTE: MLE assumes all Categories C are equally likely (it does not account for the prior!)

Likelihood for continuous measurements

Recall that our weight sensor returns a continuous random variable from a Gaussian distribution:

$$f_{W|C}(w|C) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(w-\mu)^2}{2\sigma^2}}$$

Category (C)	$\boldsymbol{f_{W|C}(W|C)}$
Cardboard	$\boldsymbol{N(20, 10)}$
Paper	$\boldsymbol{N(5, 5)}$
Can	$\boldsymbol{N(15, 5)}$
Scrap metal	$\boldsymbol{N(150, 100)}$
Bottle	$\boldsymbol{N(300, 200)}$

$N(\mu, \sigma^2)$ denotes the Gaussian distribution with mean and variance μ *and* σ^2

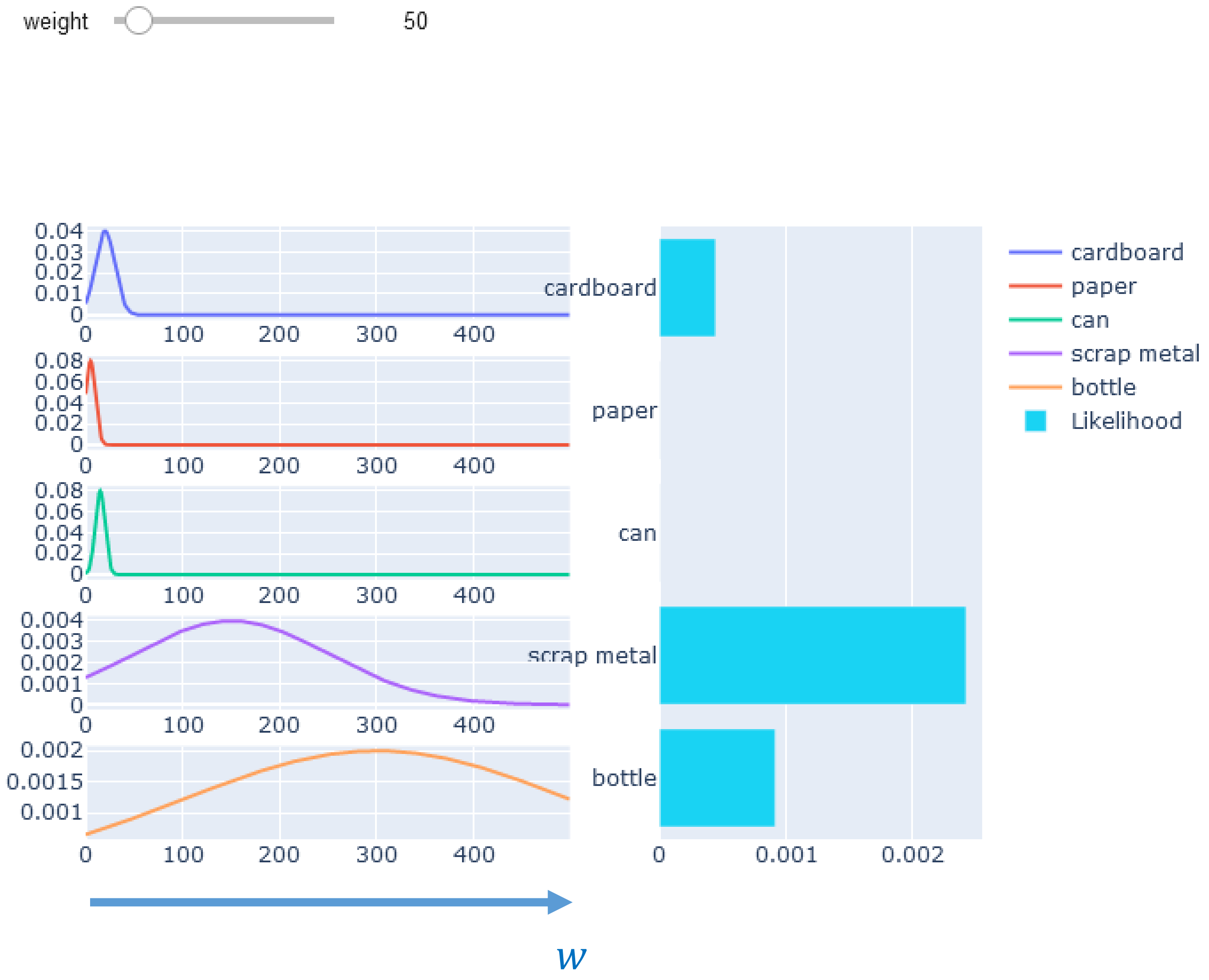
The likelihood function for category c is given by:

$$P(S|C) = \frac{1}{\sigma_c\sqrt{2\pi}} e^{-\frac{(w-\mu_c)^2}{2\sigma_c^2}}$$

For example,

$$P(w|ScrapMetal) = \frac{1}{10\sqrt{2\pi}} e^{-\frac{(w-150)^2}{200}}$$

Example



Example for $w = 50$.

- On the left are the five conditional probabilities for the categories
- On the right are the likelihood values for $w = 50$.

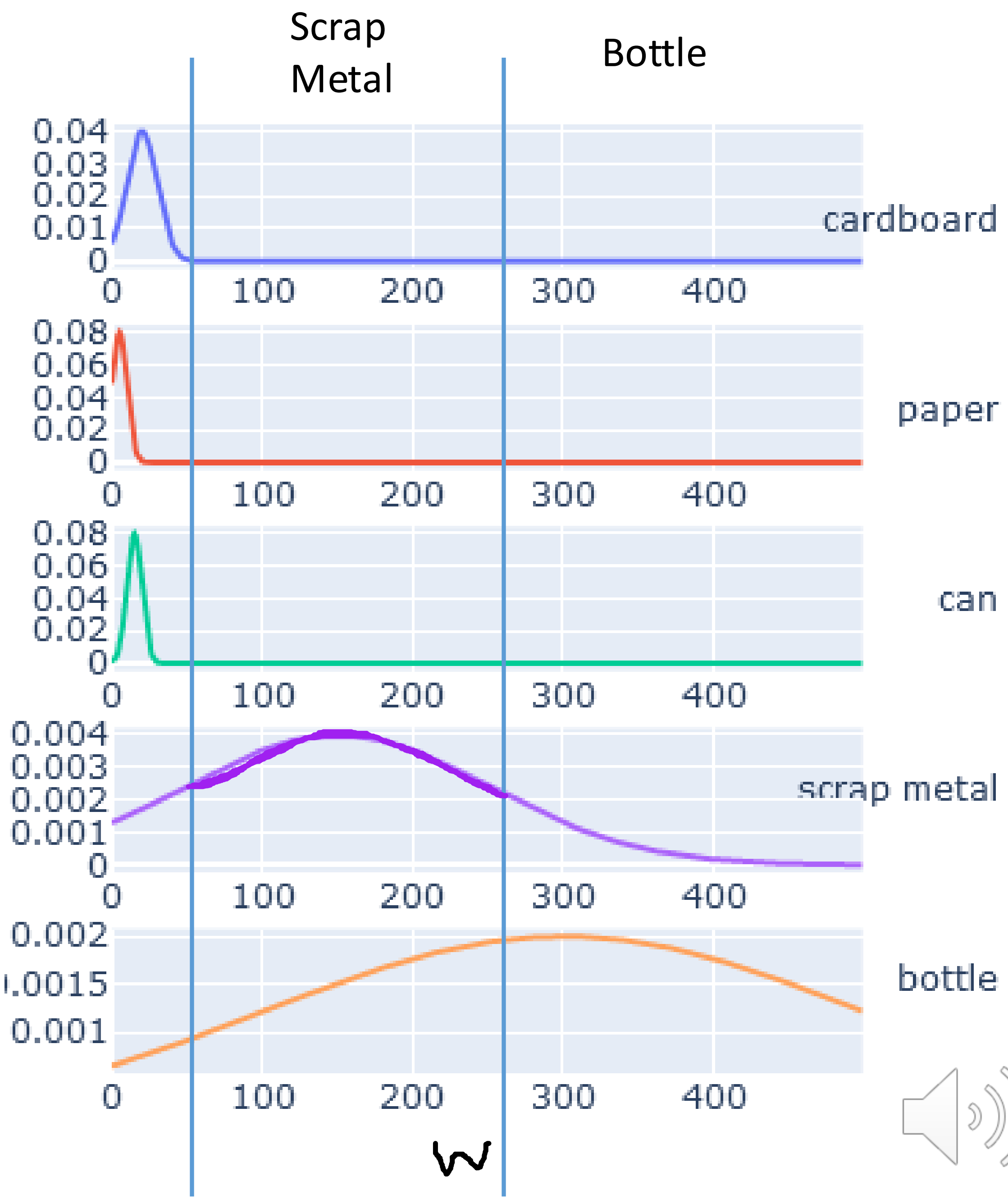
In this example, the maximum likelihood estimate is Scrap Metal

Example (cont)

- As the weight increases, the maximum likelihood category changes from Paper to Can to Cardboard to Scrap Metal to Bottle.
- For example, Scrap Metal wins out for a long interval between approx. 45g and 270g
- Bottle becomes the MLE above 270g.

The transition points are known as *decision boundaries.*

These represent the locations in measurement space where our ML estimator changes its estimate.



Maximum A Posteriori (MAP) Estimation

- The MAP estimate is the category that maximizes the posterior probability of the category, C, given the measurement, S, i.e., $C^* = \arg\max_{C} P(C|S)$
- Recall that Bayes gives the posterior as $P(C|S) = \frac{P(S|C)P(C)}{P(y)} = \eta P(S|C)P(C)$
- Since $\eta > 0$ is a constant,

$$\arg\max P(C|S) = \arg\max \cancel{\eta} P(S|C)P(C) = \underline{\arg\max P(S|C)P(C)}$$

- And therefore, the MAP estimate is given by

$$c^* = \arg\max_{c \in C} P(S|c)P(c)$$

Summary

Both Maximum Likelihood Estimation (MLE) and Maximum A Posteriori (MAP) Estimation solve

$$C^* = \arg\max_{\mathrm{C}} P(C|S)$$

The MAP estimate directly uses Bayes Rule, such that

$$c^* = \arg\max_{c \in \mathrm{C}} P(S|c)P(c)$$

MLE is the same as assuming that P(c) is constant, or that we don't have a prior. It is calculated by

$$c^* = \arg\max_{c \in \mathrm{C}} P(S|c)$$