SNS COLLEGE OF TECHNOLOGY

**Coimbatore-35**

**An Autonomous Institution**

Accredited by NBA – AICTE and Accredited by NAAC – UGC with 'A+' Grade

Approved by AICTE, New Delhi & Affiliated to Anna University, Chennai

**DEPARTMENT OF AI&ML**

# FOUNDATIONS OF ARTIFICIAL INTELLIGENCE

II YEAR - III SEM

UNIT 5 – Statistical methods

# Statistical methods

# Statistical Learning Methods

- Statistical Learning based on the Learning of uncertainty in real environments.
- The methods probability and decision theory are used to handle uncertainty by the Agents
- First the agent must learn its probabilistic theories of the world from **experience**.
- A Bayesian view of learning is extremely powerful, providing general solutions to the problems of noise, overfitting, and optimal prediction.

# Statistical Learning Methods...

- Statistical Learning is about inferences
- The idea is generated from the **Data and Hypothesis** and these are called as key terms of statistical learning.
- Data (Samples and Population) are **Evidence**

# Surprise Candy

- Let us consider a very simple example.
- Our favorite Surprise candy comes in two flavors:
- Cherry (yum) and 
- Lime (ugh). 
- The candy manufacturer has a peculiar sense of humor and wraps each piece of candy in the same opaque wrapper, regardless of flavor.

- The candy is sold in very large bags, of which there are known to be five kinds—again, cannot identify from the outside:

1. h1: 100% cherry

2. h2: 75% cherry + 25% lime

3. h3: 50% cherry + 50% lime

4. h4: 25% cherry + 75% lime

5. h5: 100% lime

- Suppose there are five kinds of bags of candies:
- 10% are h1: 100% cherry candies

- 20% are h2: 75% cherry candies + 25% lime candies

- 40% are h3: 50% cherry candies + 50% lime candies

- 20% are h4: 25% cherry candies + 75% lime candies

- 10% are h5: 100% lime candies

- Then we observe candies drawn from some bag:
- What kind of bag is it? What flavor will the next candy be?

# Surprise Candy...

- Given a new bag of candy, the random variable H (for *hypothesis*) denotes the type of the bag, with possible values h1 through h5.
- H is not directly observable, of course.
- As the pieces of candy are opened and inspected,
- data are revealed—$D_1$, $D_2$, ...., $D_N$,
- where each $D_i$ is a random variable with possible values cherry and lime.
- The basic task faced by the agent is to **predict the flavor of the next piece of candy.**
- The agent need to learn a theory of its world,

- **Bayesian learning** simply calculates the probability of each hypothesis, given the data, and makes predictions on that basis.
- Let **D** represent all the data, with observed value **d**; then the probability of each hypothesis is obtained by Bayes' rule:
- $P(h_i/\mathbf{d}) = \alpha P(\mathbf{d}/h_i)P(h_i)$ :

- suppose we want to make a prediction about an unknown quantity X

$$\mathbf{P}(X|\mathbf{d}) = \sum_i \mathbf{P}(X|\mathbf{d}, h_i)\mathbf{P}(h_i|\mathbf{d}) = \sum_i \mathbf{P}(X|h_i)P(h_i|\mathbf{d})$$

- where each hypothesis determines a probability distribution over X.
- This equation shows that predictions are weighted averages over the predictions of the individual hypotheses.
- The key quantities in the Bayesian approach are the **prior hypothesis**, $P(h_i)$, and the **likelihood** of the data under each hypothesis, $P(\mathbf{d}/h_i)$.

- Our candy example, we will assume for the time being that the prior distribution over
- h1, h2, h3, h4, h5 is given by
- <0.1 0.2 0.4 0.2 0.1>

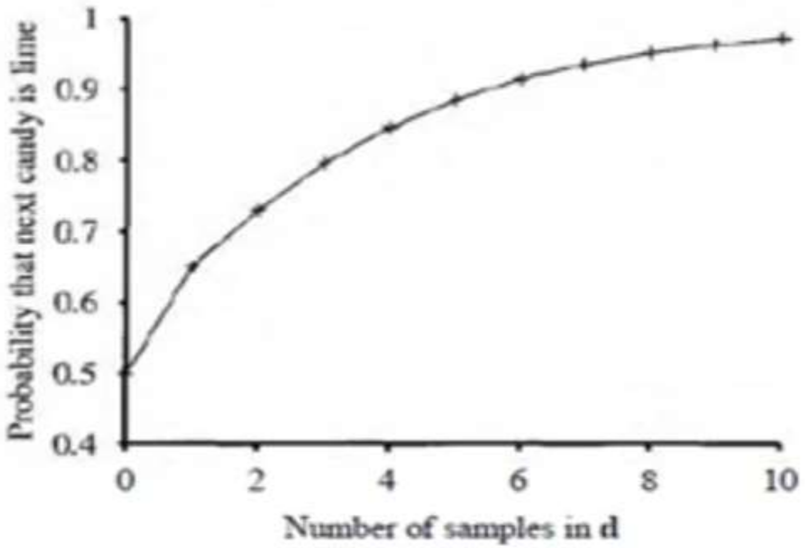
- as advertised by the manufacturer.
- The likelihood of the data ($P(d/h_i)$) is calculated under the assumption that the observations are **i.i.d.**—that is, **independently and identically distributed**—so that

$$P(\mathbf{d}|h_i) = \prod_j P(d_j|h_i)$$

- The predicted probability that the next candy is lime, based on Equation with respect to h5

$$\mathbf{P}(X|\mathbf{d}) = \sum_i \mathbf{P}(X|\mathbf{d}, h_i)\mathbf{P}(h_i|\mathbf{d}) = \sum_i \mathbf{P}(X|h_i)P(h_i|\mathbf{d})$$

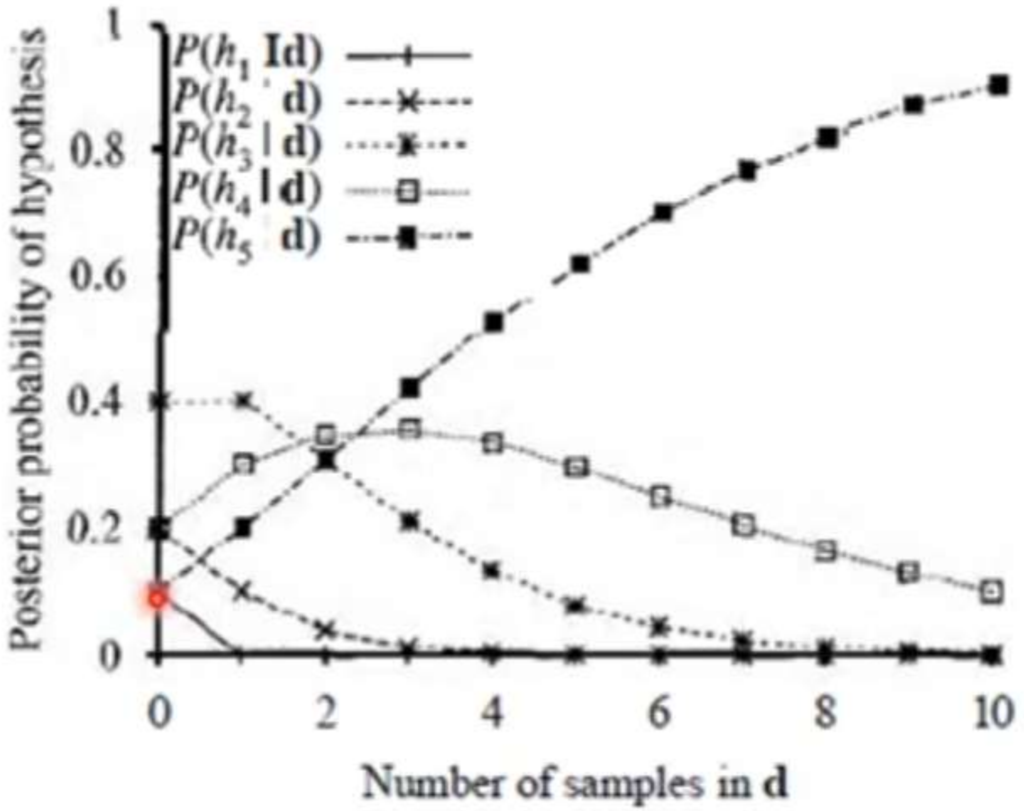
- As we would expect, it increases monotonically toward 1.



(b)

- the posterior probabilities of the five hypotheses change as the sequence of *10* lime candies is observed.

- H1
- H2
- H3
- H4
- h5



# THANK YOU