

# Chapter 5: Evaluating Hypotheses

# What reasons do we have for evaluating hypotheses?

- to decide whether to use a particular hypothesis or not
- to use the evaluation as part of the learning algorithm

## **The Key Difficulties are**

- Bias in estimate
- Variance in estimate

# Estimating Hypothesis Accuracy

# Some Definitions

$X$  Space of possible instances

$D$  Unknown probability distribution that defines the probability of each instance in  $X$ .

$f(x)$  The target function

$h(x)$  The hypothesized function

# Interesting Questions

“

1. Given a hypothesis  $h$  and a data sample of size  $n$  drawn randomly according to the distribution  $D$ , what is the best estimate of the accuracy of  $h$  over future instances drawn from  $D$ ?
2. What is the probable error in this accuracy estimate?

”

# Sample Error and True Error

**sample error:**

$$\text{errors}(h) = \frac{1}{n} \sum_{x \in S} d(f(x), h(x))$$

where  $d(f(x), h(x))$  is 1 if  $f(x) \neq h(x)$

**true error:**

$$\text{errorD}(h) = \Pr_{x \in D} [f(x) \neq h(x)]$$

# Confidence Intervals for Discrete-Valued Hypotheses

$$\text{errors}(h) \pm Z_N \sqrt{(\text{errors}(h) (1 - \text{errors}(h)) / n)}$$

where  $Z_N$  is a constant that is chosen depending on the % confidence required. eg for  $N=95$   $Z_N = 1.96$

# Confidence Intervals

The previous equation is a good approximation when

$$n \text{ errors}(h)(1 - \text{errors}(h)) \geq 5$$

# Sampling Theory

# Sampling Theory

1. Error Estimation and Estimating Binomial Proportions
2. The Binomial Distribution
3. Mean and Variance
4. Estimators, Bias and Variance
5. Confidence Intervals
6. Two-Sided and One-Sided Bounds