

**STANFORD UNIVERSITY**  
**DEPARTMENT OF STATISTICS**  
**DEPARTMENTAL SEMINAR**

4:15 p.m., Tuesday, March 14, 2006  
Sequoia Hall Room 200  
(Cookies at 3:45 in 1st Floor Lounge)

*Lutz Duembgen*  
University of Berne

**P-Values for Classification**

Abstract: Let  $(X, Y)$  be a complete observation consisting of a feature vector  $X$  and a class label  $Y \in \{1, 2, \dots, K\}$ . Classification means to predict the unobserved class label  $Y$  from the observed feature vector  $X$ . The joint distribution of  $(X, Y)$  is typically unknown and estimated from certain training data. We propose to replace classifiers (i.e. point predictors) or the (estimated) posterior distribution  $L(Y|X)$  with a tuple of p-values for the  $K$  potential class memberships. After a brief discussion of the potential benefits of this approach, we extend the classical theory of optimal classifiers to optimal p-values. Thereafter we discuss the impact of estimating the joint distribution from training data and describe a general method based on permutation tests. Some theoretical results and numerical examples show the method's potential.