# American Statistical Association

## Kansas-Western Missouri Chapter 2006 Fall Meeting

Speaker

**Dr. Morgan C. Wang**
**Department of Statistics and Actuarial Science**
**University of Central Florida**
**Orlando, FL 32816-2370**

# Maximum Likelihood Regression Trees

# (See abstract next page)

Monday, October 16th, 2006
at
University of Missouri Kansas City
University Center
5000 Holmes
Room 106

6.00 p.m. – 6.30 p.m. Social Time
Dinner starts at 6.30 p.m.

Dinner:     Triple Entree Dinner Buffet: Tossed seasonal salad served with ranch and Italian salad, pecan chicken with honey mustard on the side, London broil with mushroom sauce, vegetable stuffed Portobello mushrooms, rice pilaf, vegetable medley, Green bean armondine, fresh wheat and white, dinner rolls served with butter, cheesecake, brewed regular and decaffeinated coffee, iced tea, and iced water

Cost:     $30.00 ($20.00 for students)
Contact information: Ananda Jayawardhana (ananda@pittstate.edu)

# Maximum Likelihood Regression Trees

**Morgan C. Wang**
**Department of Statistics and Actuarial Science**
**University of Central Florida**
**Orlando, FL 32816-2370**

**Abstract** Least squares regression trees have been widely studied and well implemented in many automated statistical methods. One limitation of using least squares optimization criterion to construct trees, however, is the data type for response variable must be continuous. For discrete response, other optimization criteria such as misclassification rate (e.g., CHAID like trees or CART like trees) are used. This limitation can be overcome with the proposed maximum likelihood approach that can deal with various types of response variables. In addition, many standard likelihood related methods such as the AIC model selection criterion and the likelihood ratio test can be naturally incorporated into the tree analysis. Simulation study shows that the proposed maximum likelihood approach tends to select more accurate tree size than traditional least squares approaches. The analysis of a baseball data set is given as an illustration example. We also discuss how to use the proposed tree method to "boost" ordinary regression models.

*Key Words:* *Regression Trees, AIC, $AIC_C$, Maximum Likelihood, and CART.*