## Department of Mathematical Sciences Colloquium

## MING-YING LEUNG Department of Math. Sciences, UTEP

## Poisson Process Approximation for Palindrome Occurrences in Random DNA Sequences

In many bioinformatics problems, the mathematical model for DNA is a random sequence of letters sampled from the alphabet A, C, G, T. The occurrences of palindromes (symmetrical DNA words) can then be considered as a stochastic process. Using the Wasserstein distance, it can be shown that the palindrome process is approximated by a Poisson process under suitable conditions. Furthermore, a computable upper bound on the Wasserstein distance between the palindrome and the Poisson processes can serve as a guide to the accuracy of the approximation. The properties of the Poisson process provide a mathematical basis for formulating statistical criteria that help identify nonrandom patterns in DNA sequences which may be of biological importance. However, it should also be recognized that the utility of a simple Poisson process approximation is rather limited and extensions to compound Poisson process approximations need to be explored.

## Friday, February 10, 2006, at 3 pm in Bell Hall 143 The University of Texas at El Paso

Refreshments will be served in front of the colloquium room 15 minutes before the start of the colloquium.

For further information, please contact Dr. Pavel Šolín, Bell Hall 220. Phone: (915) 747-6770, email: solin@utep.edu.