

Undergraduate Research Symposium May 17, 2013 Mary Gates Hall

Online Proceedings

SESSION 1N

MCNAIR SESSION - EXPLORING THE NATURAL WORLD: FROM NUMBERS TO NANOPARTICLES AND BATS TO BACTERIA

*Session Moderator: Todd Sperry, Office of Minority Affairs
& Diversity*

287 MGH

1:15 PM to 2:45 PM

* Note: Titles in order of presentation.

Statistical Designs of Experiment: Construction of Generalized Minimum Aberration Designs of Size 32

*Laura White, Senior, Mathematics, Arkansas State
University*

McNair Scholar

*Mentor: Hong Zhou, Mathematics and Statistics, Arkansas
State University*

Statistical designs are a way to layout experiments before conducting your experiment. A well-chosen experimental design allows one to get more information out of the experiment. Regular fractional factorial designs are types of experimental designs that are used for studying effects of two or more constraints simultaneously, but leave large gaps in run size. Run sizes are the amount of experiments an experimenter has to conduct. Non-regular fractional factorial designs can be constructed for every run size that is a multiple of four, which allows run size flexibility and economy. My research focuses on construction of optimal designs of size 32 runs using graphic processing unit (GPU) technology. I have been working on creating a design table of non-regular fractional factorial of size 32 runs. Creating design tables make it possible for engineers and scientist to plan experiments for any combination of runs and number of variables to be studied.

POSTER SESSION 2

MGH 241, Easel 149

12:45 PM to 2:15 PM

Uncertainty in Ranking the Hottest Years of U.S. Surface Temperatures

Tae Yen (Amy) Kim, Senior, Statistics, Economics

Mentor: Peter Guttorp, Statistics

People have often heard statements like "2010 is the hottest year on record in the world." However, this ranking in order of average global temperature is uncertain, because the global temperature is not a direct measurement but a statistical combination of data, which by itself has uncertainty. We have developed new methods to rank each year taking into account the standard error of a real mean temperature. The methods give us distributions of the rank of each year's temperature, and allows us to compute the probability that a given year is the hottest year on record. We can arrange the years in order of these probabilities. The results are consistent with regular ranking of continental US annual mean temperature, but allows us to express the uncertainty of the ranks. We can also estimate, for example, the probability that 12 of the last 15 years have been among the hottest ever, as President Obama said in his State of the Union address in February.

POSTER SESSION 3

Commons East, Easel 57

2:30 PM to 4:00 PM

Reliability in Mechanical Analysis of Twitter Pictures

Gilad Amitai, Senior, Statistics

Mentor: Tyler McCormick

Twitter is an exciting social network to discover trending topics, opinions, and behaviors. While Twitter has already been mined for this sort of information, these analyses have failed to collect data that is integral to the social sciences. The research group I am working with is exploring collecting demographic information from Twitter in order to apply information from Twitter to research in the social sciences. The demographic data was collected using the Amazon Mechanical Turk service, where participants were paid a small fee to determine characteristics including race, age, and sex from pictures posted on Twitter. As a member of the group, I examined the reliability of these results and developed different methods to determine the demographic information associated with a set of tweets. These methods were tested and compared to national survey data to determine precision. I hope that my results will show that the Amazon Mechanical

Turk service is an accurate and efficient method for collecting demographic data on Twitter, thus leading to a systematic way to use Twitter in the social sciences.