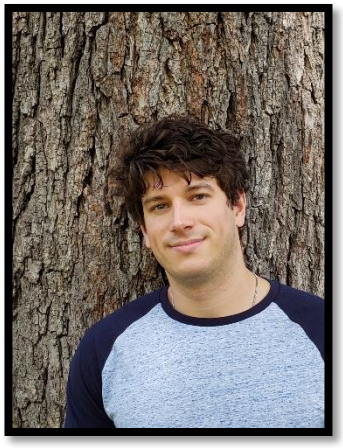
# STATISTICS SEMINAR

UW-Department of Statistics www.stat.wisc.edu



## TITLE:

**Differentially Private Inference for Binomial Data**


Abstract: Differential Privacy (DP) is a rigorous framework which quantifies the disclosure risk of statistical procedures computed on sensitive data. DP methods/mechanisms require the introduction of additional randomness beyond sampling in order to limit the disclosure risk. However, their implementations often introduce excessive noise, reducing the utility and validity of statistical results, especially in small samples.

In this paper, we first derive uniformly most powerful (UMP) tests for simple and one-sided hypotheses for a population proportion within the framework of DP, optimizing finite sample performance. By studying the privacy constraints, we prove a `Neyman-Pearson Lemma' for binomial data under DP, resulting in a DP-UMP test. Our tests can also be stated as a post-processing of a DP summary statistic, whose distribution we coin ``Truncated-Uniform-Laplace'' (Tulap), a generalization of the Staircase and discrete Laplace distributions.

We show that by post-processing the Tulap statistic, we are able to obtain exact p-values corresponding to the DP-UMP, uniformly most accurate (UMA) one-sided confidence intervals, uniformly most powerful unbiased (UMPU) two-sided tests, and uniformly most accurate unbiased (UMAU) two-sided confidence intervals. As each of these quantities are a post-processing of the same summary statistic, there is no increased privacy risk by including these additional results, allowing for a complete statistical analysis at a fixed privacy cost.


## Speaker:


**Jordan Awan**
PhD Student
Pennsylvania State University


## Time & Place:

Friday, February 14, 2020 **4pm**,
Room 133 SMI
Cookies & Coffee @ **3:30**, Rm 1210 MSC

THE UNIVERSITY of WISCONSIN MADISON