MS&E 233: HW #2

Spring 2012-13, Prof. Ashish Goel, Mukund Sundararajan

Due in-class Tuesday, May 28

1. Suppose that a database has a single column that represents the income of the individual; recall that rows correspond to individual. Suppose that income is bounded by K and consider the query of finding the minimum income in the table. What is the sensitivity of this query? To recall the definition of sensitivity, see the top of page 4 of this paper linked of the second slide of the second lecture on privacy. What noise would you add to the query to get ϵ-Differential privacy? In what sense is this noise addition conservative for most databases?

2. Suppose a mobile app records the following information about its US users each time the user opens the app: mobile OS type, mobile OS version, and zip code of the phone's billing location. Assume there are 3 OS types, 10 OS version for each OS, 41,811 zip codes, and they are all independent and equally likely. What is the expected number of uniquely identifiable users using this information given that the app has 3 million users? (Note, suppose that each user has one phone, and does not change or update its OS or their billing zip code during the observation time.)

3. Consider a prediction market similar to what described in class, but assume that for a user i with prediction $p_i > c_Y$, the number of "Yes" shares the user holds is proportional to $p_i - c_Y$. The symmetric condition holds for c_N. Characterize the market clearing price c_Y in terms of the distribution of p_i.

4. Research Parimutuel markets on the web. Then consider a parimutuel market where there are only two outcomes, and every user makes a bid of \$1 on exactly one event. Assume that betting is open for an hour during which no user gets any new information, and users are allowed to switch their money from one option to the other as often as they like. Assume that the system reaches equilibrium (i.e. no user would

like to switch her bid at the end of the hour). Relate the amount of money bid on the "Yes" outcome to the distribution of p_i.

5. Outline one societal problem that you could solve using a social networks or networked marketing approach, and describe the proposed solution. (< 200 words).

6. **Widescope**: Follow instructions of the first phase of Widescope project sent to you via email.