DIFFERENTIAL PRIVACY - 1

1. Which of the following methods is the best method to efficiently protect the data to preserve the privacy of the users?
    a. Anonymization
    b. Cryptographical Solution
    c. Statistical solution
    d. Data Compression
    e. Data Duplication

2. Between a randomized response (with epsilon>0) and a fair coin toss response, which algorithm would you use to preserve privacy but have a better utility?
    a. Randomized response because the chance of falsehood is 50%
    b. Randomized response because the chance of truth is greater than 50%
    c. Randomized response because the chance of falsehood is greater than 50%
    d. Coin toss response because the chance of falsehood is 50%
    e. Coin toss response because the chance of truth is greater than 50%
    f. Coin toss response because the chance of falsehood is greater than 50%

3. Consider the equation in the context of privacy guarantees (The notations used are the same as used during the lecture).

    $$P(RR(x') = b) * e^{-\varepsilon} \leq P(RR(x) = b) \leq P(RR(x') = b) * e^{\varepsilon}$$

    To maximize the privacy gains, which of the following values should be changed and how?
    a. ε should be maximum for privacy, ε should be minimum for utility
    b. ε should be minimum for privacy, ε should be minimum for utility
    c. ε should be maximum for privacy, ε should be maximum for utility
    d. ε should be minimum for privacy, ε should be maximum for utility
    e. ε is unrelated

4. Consider the below values:
   X = {x1,x2,.....xN} is the truth of an experiment
   Y = {y1, y2,.....yN} is the revealed values instead of the truth
   To identify the average of truth, Y as an estimator cannot be used for the process by which it was obtained. You derive new values Z where Z = {z1,z2,.....zN} from Y which are better estimators of X. How do you arrive at the values Z?
   a. Removing the bias from Y introduced through the random process
   b. Adding the bias to Y removed through the random process
   c. Removing the variance from Y introduced through the random process
   d. Adding the variance to Y removed through the random process

5. If ε is fixed, given a privacy guarantee, to improve the utility, which of the following values can be modified?
   a. Increase the number of experiments
   b. Increase the amount of randomness
   c. Increase the amount of bias introduced in the random process
   d. Increase the amount of variance introduced in the random process

6. Identify the equation for the ε-differential mechanism (The notations used are the same as used during the lecture): :
   a. $\frac{P(M(x)\epsilon S)}{P(M(x')\epsilon S)} \leq e^{\varepsilon}$
   b. $\frac{P(M(x')\epsilon S)}{P(M(x')\epsilon S)} \leq e^{\varepsilon}$
   c. $\frac{P(M(x)\epsilon S)}{P(M(x)\epsilon S)} \leq e^{\varepsilon}$
   d. $\frac{P(M(x)\epsilon S)}{P(M(x')\epsilon S)} \geq e^{\varepsilon}$
   e. $\frac{P(M(x')\epsilon S)}{P(M(x')\epsilon S)} \geq e^{\varepsilon}$
   f. $\frac{P(M(x)\epsilon S)}{P(M(x)\epsilon S)} \geq e^{\varepsilon}$

7. Identify the correct scenario in the case of differential privacy
    a. Trust the curator; Trust the world
    b. Do not trust the curator; Trust the world
    c. Trust the curator; Do not trust the world
    d. Do not trust the curator; Do not trust the world

8. Identify all the values representing sensitivity in a laplacian mechanism where the function under consideration is an average of n binary values {0,1} (The notations used are the same as used during the lecture).
    a. $\frac{1}{n}$
    b. $\frac{1}{n} | x_n' - x_n |$
    c. ε
    d. $\frac{\varepsilon}{n}$
    e. $\frac{-1}{n}$
    f. $\frac{\Delta}{\varepsilon}$

9. Identify the distribution from which the noise is derived in a laplacian mechanism. The representation is of the form Laplacian(a,b) where a is the mean and b is the spread parameter. (The notations used are the same as used during the lecture)
    a. laplacian(1, $\frac{\Delta}{\varepsilon}$)
    b. laplacian($\frac{\Delta}{\varepsilon}$,0)
    c. laplacian($\frac{\Delta}{\varepsilon}$,1)
    d. laplacian(0, $\frac{\Delta}{\varepsilon}$)
    e. laplacian(1,1)
    f. laplacian(0,0)

10. Higher privacy guarantees can be achieved in which of the following scenarios? Identify all the possible scenarios.
    a. Epsilon should be high
    b. Inverse Sensitivity should be high
    c. Variance should be high
    d. Noise should be high
    e. Utility should be high

11. Identify the deviation of the value from the truth in the scenario of a laplacian mechanism. (The notations used are the same as used during the lecture).
    a. $O(\frac{1}{\varepsilon n})$
    b. $O(\frac{n'}{\varepsilon n})$
    c. $O(\frac{\varepsilon}{n})$
    d. $O(\frac{e}{\varepsilon n})$
    e. $O(\varepsilon n)$

12. In the scenario of a privacy-utility trade-off, for fixed privacy, the number of samples required for a particular utility varies between the Laplacian mechanism and Randomized response is different by what factor?
    a. Constant factor
    b. Linear factor
    c. Exponential factor
    d. Logarithmic factor
    e. Quadratic factor