EXERCISE 1. Show that the randomised response mechanism, as originally defined, is not differentially private with respect to the addition/deletion neighbourhood.

The randomized response mechanism is a technique in surveys to allow respondents to answer sensitive questions while maintaining confidentiality. A randomized response mechanism that is differentially private typically ensures that the probability of a specific output does not change much when an individual's data is added or removed. This is formalized by the concept of ϵ-differential privacy, which requires that for all outputs o and for any two neighboring datasets D and D' (datasets that differ by one element):

$$\frac{Pr[\text{Mechanism}(D) = o]}{Pr[\text{Mechanism}(D') = o]} \leq e^{\epsilon}$$

The original randomized response mechanism does not satisfy this definition. Let's consider a simplified mechanism where a respondent flips a coin in private and, if it comes up heads, they tell the truth; if tails, they say "yes" regardless of the truth. This mechanism is not ϵ-differentially private with respect to the addition or deletion of an individual's data because the probabilities of the outputs can change significantly with the presence or absence of a single individual's truthful response.

For instance, if the true answer for an individual added to the dataset would be "yes", the probability of the mechanism outputting "yes" can increase substantially compared to the dataset without that individual's data. The ratio of these probabilities is not bounded by e^{ϵ} for any ϵ, because the probability could increase from 0.5 (if the true answer was "no" for everyone else) to 1 (if the coin flip is tails or the added individual's true answer is "yes"), which is an unbounded increase.

Hence, the original randomized response mechanism as described does not satisfy the condition for ϵ-differential privacy with respect to the addition or deletion of a single individual's data.

EXERCISE 2. Define a variant of the binary randomised response mechanism that satisfies differential privacy with respect to the insertion and deletion neighbourhood. More precisely, it should hold that for any dataset pair $\boldsymbol{x} = (x_1, \ldots, x_n)$, and $\boldsymbol{x}' = (x_1, \ldots, x_n)$, $\boldsymbol{x} = (x_1, \ldots, x_n, x_{n+1})$, the mechanism satisfies

$$\pi(A|\boldsymbol{x}) \leq e^{\epsilon}\pi(A|\boldsymbol{x}'), \qquad \forall A \subset \mathcal{A}.$$

To define a variant of the binary randomized response mechanism that satisfies ϵ-differential privacy for the insertion and deletion neighborhood, one common approach is to adjust the probabilities of the randomized response such that the presence or absence of any single data point does not significantly change the probability distribution of the output.

Let's denote the original binary response as x_i, which can be 0 or 1. The randomized mechanism π will output a response based on the following:

1. With probability p, output the true response x_i.
2. With probability $1 - p$, output a random bit, which is 0 or 1 with equal probability of 0.5.

To satisfy ϵ-differential privacy, we want:

$$\frac{\pi(A|x)}{\pi(A|x')} \leq e^{\epsilon}, \quad \forall A \subseteq \{0, 1\}$$

This can be achieved by setting p according to the privacy budget ϵ. Specifically, we can choose:

$$p = \frac{e^{\epsilon}}{1 + e^{\epsilon}}$$

And the probability of outputting the opposite of the true value would be:

$$1 - p = \frac{1}{1 + e^{\epsilon}}$$

With this setting, no matter what the true value is, the output probabilities will differ by at most a multiplicative factor of e^{ϵ}, satisfying the differential privacy condition.

The proof is short and involves checking that the ratio of probabilities under any outcome A for datasets x and x' that differ in one element does not exceed e^{ϵ}, which is guaranteed by the chosen probabilities.

EXERCISE 3. The mean estimator $\hat{\theta}$ for the randomised response mechanism with flipping probability p, where the true data generation process is Bernoulli with parameter θ, is defined as

$$\hat{\theta}(a) = \frac{\bar{a} - p}{1 - 2p},$$

where $\bar{a} = \frac{1}{T}\sum_{t=1}^{T} a_t$ is the mean of the observed answers. Prove that this estimator is unbiased, i.e. that

$$\mathbb{E}[\hat{\theta}] = \theta,$$

where the expectation is taken over the unobserved data randomness and the randomness of the mechanism.

To prove that the estimator $\hat{\theta}(a)$ is unbiased, we need to show that the expected value of $\hat{\theta}$ is equal to the true parameter θ. The estimator is defined as:

$$\hat{\theta}(a) = \frac{\bar{a} - p}{1 - 2p},$$

where $\bar{a}$ is the mean of the observed answers a_t from the randomized response mechanism, and p is the probability of flipping the response.

Let Y be the observed random variable from the randomized response mechanism. Then Y is a mixture of the true response X (which is Bernoulli distributed with parameter θ) and the noise added by the random flipping process. Specifically:

- With probability p, the response is flipped (i.e., $Y = 1 - X$).
- With probability $1 - p$, the response is truthful (i.e., $Y = X$).

Therefore, the expected value of Y is:

$$\begin{aligned} E[Y] &= p(1 - \theta) + (1 - p)\theta \\ &= p + (1 - 2p)\theta. \end{aligned}$$

The observed mean $\bar{a}$ is the average of T observed values of Y, so $E[\bar{a}] = E[Y]$. Plugging the expected value of Y into the estimator gives us:

$$\begin{aligned} E\left[\frac{\bar{a} - p}{1 - 2p}\right] &= \frac{E[\bar{a}] - p}{1 - 2p} \\ &= \frac{p + (1 - 2p)\theta - p}{1 - 2p} \\ &= \frac{(1 - 2p)\theta}{1 - 2p} \\ &= \theta. \end{aligned}$$

Since the expected value of the estimator $\hat{\theta}$ is equal to the true parameter θ, the estimator is unbiased.

EXERCISE 4. Prove that the exponential mechanism is uniform when $\epsilon \to 0$ and deterministically returns the maximising action when $\epsilon \to \infty$.

Given a utility function $u(d, r)$ which measures the utility of outcome r for the dataset d, and a privacy parameter ϵ, the probability that the exponential mechanism $M(d, u, \epsilon)$ outputs a particular result r is:

$$Pr[M(d, u, \epsilon) = r] \propto \exp\left(\frac{\epsilon u(d, r)}{2\Delta u}\right),$$

where Δu is the sensitivity of the utility function, the maximum change in u that can be caused by changing a single element in the dataset.

Now, let's prove the two statements:

1. **As $\epsilon \to 0$, the exponential mechanism becomes uniform.**
   As $\epsilon \to 0$, the exponent $\frac{\epsilon u(d,r)}{2\Delta u}$ approaches 0 for all outcomes r, since $u(d,r)$ and Δu are fixed with respect to ϵ. Therefore, $\exp\left(\frac{\epsilon u(d,r)}{2\Delta u}\right)$ approaches 1 for all r, making the probability distribution uniform since all outcomes are equally likely when the exponent is 0.
2. **As $\epsilon \to \infty$, the exponential mechanism deterministically returns the maximizing action.**
   As $\epsilon \to \infty$, the term $\exp\left(\frac{\epsilon u(d,r)}{2\Delta u}\right)$ grows without bound for the outcome r with the highest utility $u(d,r)$, much faster than for any other outcomes. This is because, for any r' with $u(d,r') < u(d,r)$, the ratio of their probabilities becomes:

$$\frac{\exp\left(\frac{\epsilon u(d,r)}{2\Delta u}\right)}{\exp\left(\frac{\epsilon u(d,r')}{2\Delta u}\right)} = \exp\left(\frac{\epsilon(u(d,r) - u(d,r'))}{2\Delta u}\right),$$

   which approaches infinity as $\epsilon \to \infty$. Hence, the probability that the exponential mechanism selects the outcome r with the highest utility approaches 1, meaning that the mechanism deterministically selects the action that maximizes the utility function.

EXERCISE 5. Prove that the exponential mechanism, when we used to calculate a noisy version of the function $q(x)$ with $q : \mathcal{X} \to \mathbb{R}$ and utility $U(a, q, x) = -|q(x) - a|^p$, results in the Laplace mechanism for $p = 1$ and the Gaussian mechanism for $p = 2$.

1. **Laplace Mechanism for $p = 1$:**

   When $p = 1$, the utility function becomes $U(a, q, x) = -|q(x) - a|$. The exponential mechanism selects an outcome a with probability proportional to $\exp\left(\frac{\epsilon U(a,q,x)}{2\Delta q}\right)$, where Δq is the sensitivity of the query function q.

   The probability of selecting outcome a is:

$$Pr[M(d, q, \epsilon) = a] \propto \exp\left(-\frac{\epsilon|q(x) - a|}{2\Delta q}\right)$$

   This is the probability density function of the Laplace distribution, centered at $q(x)$ with scale parameter $b = \frac{2\Delta q}{\epsilon}$. Thus, the exponential mechanism with $p = 1$ and the given utility function is equivalent to adding Laplace noise to $q(x)$.

1. **Gaussian Mechanism for $p = 2$:**

   When $p = 2$, the utility function becomes $U(a, q, x) = -(q(x) - a)^2$. Similarly, the probability of selecting outcome a is:

$$Pr[M(d, q, \epsilon) = a] \propto \exp\left(-\frac{\epsilon(q(x) - a)^2}{4\Delta q^2}\right)$$

   This is the probability density function of the Gaussian distribution, centered at $q(x)$ with variance proportional to $\frac{4\Delta q^2}{\epsilon^2}$. Therefore, the exponential mechanism with $p = 2$ and the given utility function is equivalent to adding Gaussian noise to $q(x)$.

   In both cases, the selected outcome a is a noisy version of $q(x)$, with the noise distribution determined by the parameter p in the utility function. The exponential mechanism with these specific utility functions thus results in mechanisms that are equivalent to the Laplace mechanism for $p = 1$ and the Gaussian mechanism for $p = 2$, which are commonly used differentially private mechanisms to add noise to the function outputs.

EXERCISE 6. Consider the following relaxation of differential privacy to KL divergences. A mechanism is ϵ-KL-private if

$$\sum_a \ln \frac{\pi(a|x)}{\pi(a|x')} \pi(a|x) \leq \epsilon$$

Prove that two-folded composition of such a mechanism is 2ϵ-KL-private.

Let M_1 and M_2 be two mechanisms that are ϵ-KL-private. The composition of these mechanisms, applied to a dataset x, gives a joint distribution $M_1(\cdot|x) \times M_2(\cdot|x)$.

We need to compute the KL divergence of the composition, which is:

$$\begin{aligned}
&KL\left(M_1(\cdot|x) \times M_2(\cdot|x) || M_1(\cdot|x') \times M_2(\cdot|x')\right) \\
&= \sum_{a_1} \sum_{a_2} \pi(a_1, a_2|x) \ln\left(\frac{\pi(a_1, a_2|x)}{\pi(a_1, a_2|x')}\right) \\
&= \sum_{a_1} \sum_{a_2} \pi(a_1|x)\pi(a_2|x) \ln\left(\frac{\pi(a_1|x)\pi(a_2|x)}{\pi(a_1|x')\pi(a_2|x')}\right) \\
&= \sum_{a_1} \sum_{a_2} \pi(a_1|x)\pi(a_2|x) \left[\ln\left(\frac{\pi(a_1|x)}{\pi(a_1|x')}\right) + \ln\left(\frac{\pi(a_2|x)}{\pi(a_2|x')}\right)\right] \\
&= \sum_{a_1} \pi(a_1|x) \ln\left(\frac{\pi(a_1|x)}{\pi(a_1|x')}\right) \sum_{a_2} \pi(a_2|x) + \sum_{a_2} \pi(a_2|x) \ln\left(\frac{\pi(a_2|x)}{\pi(a_2|x')}\right) \sum_{a_1} \pi(a_1|x) \\
&= 2 \sum_{a} \pi(a|x) \ln\left(\frac{\pi(a|x)}{\pi(a|x')}\right) \\
&\leq 2\epsilon
\end{aligned}$$

Since both M_1 and M_2 are ϵ-KL-private, their individual divergences are each less than or equal to ϵ, and hence the total divergence for the composition is less than or equal to 2ϵ, proving the composition is 2ϵ-KL-private.

EXERCISE 8. Let us have some data generated from a null treatment policy π_0 of the form (a_t, y_t). There is a simple model that explains the data of the form

$$y_t \mid a_t = a, \theta \sim \mathcal{N}(a + \theta, 1),$$

where the actions are distribution according to $\pi(a_t)$.

- Assume that $\pi_0 \in [0, 1]$ is given and it is $a_t \mid \pi = \pi_0 \sim \textit{Bernoulli}(\pi_0)$. First, estimate θ. Then, calculate the distribution of $y_t \mid \pi_0, \theta$ for any other policy and plot the resulting mean and variance as π changes. You can do this first in a maximum-likelihood manner. Advanced: estimate the posterior distribution of θ for a normal prior on θ.
- Now assume that π_0 is not given. This means that you must also estimate π_0 itself before estimating the effect of any other policy π on the data.
- In this exercise, can you learn about other actions when you are not taking them? Why?

1. **Estimate θ when π_0 is known:**
   If π_0 is known, and actions follow $a_t|\pi = \pi_0 \sim \text{Bernoulli}(\pi_0)$, you can estimate θ using maximum likelihood estimation (MLE) by maximizing the likelihood of the observed data under the normal distribution given the Bernoulli-distributed actions.
2. **Calculate the distribution of $y_t|\pi_0, \theta$ for any policy π:**
   Once θ is estimated, you can compute the distribution of y_t for any other policy π. Since y_t is normally distributed with a mean of $a_t + \theta$, where a_t is either 0 or 1 depending on the action taken, the mean of y_t would be θ when $a_t = 0$ and $1 + \theta$ when $a_t = 1$. The variance would remain 1, as specified by the normal distribution.
3. **Estimate the posterior distribution of θ for a normal prior on θ:**
   This involves using Bayesian methods to estimate the posterior distribution of θ given a prior distribution for θ that is normal. You would update the prior based on the likelihood of the observed data to get the posterior distribution.
4. **Estimate π_0 when it is not given:**
   If π_0 is not given, you must estimate it from the data, likely by finding the proportion of actions equal to 1 in your data set. With this estimate of π_0, you can then proceed as before to estimate θ.
5. **Learning about other actions when not taking them:**
   The question of whether you can learn about other actions when you are not taking them is about inferring the effects of actions not present in your data. If your policy π only samples a subset of possible actions, you have no direct information about the outcomes of the unobserved actions. However, if there is some structure or assumption you can make about the relationship between actions and outcomes (like a linear model), you might extrapolate to some extent. The certainty of such extrapolation would generally be lower than for actions directly observed.

EXERCISE 9. Let us have some data generated from a null treatment policy π_0 of the form (a_t, y_t). Let us now consider a slightly model where $\theta \in \mathbb{R}^2$.

$$y_t \mid a_t = a, \theta \sim \mathcal{N}(\theta_a, 1),$$

where the actions are distribution according to $\pi(a_t)$.

- Assume that $\pi_0 \in [0, 1]$ is given and it is $a_t \mid \pi = \pi_0 \sim \textit{Bernoulli}(\pi_0)$. First, estimate θ. Then, calculate the distribution of $y_t \mid \pi_0, \theta$ for any other policy and plot the resulting mean and variance as π changes. You can do this first in a maximum-likelihood manner. Advanced: estimate the posterior distribution of θ for a normal prior on θ.
- Now assume that π_0 is not given. This means that you must also estimate π_0 itself before estimating the effect of any other policy π on the data.
- In this exercise, can you learn about other actions when you are not taking them? Why?

1. **Estimate θ when π_0 is known:**
   First, estimate the parameters θ using maximum likelihood estimation. Since θ is now a vector, this involves finding estimates for both components of θ (say θ_0 and θ_1), corresponding to the two possible actions a_t being 0 or 1.
2. **Calculate the distribution of $y_t|\pi_0, \theta$ for any policy π:**
   Once you have estimates for θ, calculate the distribution of y_t for any other policy π by considering the Bernoulli distribution of actions. This will give you a mixture of two normal distributions, where the means are θ_0 and θ_1, and the weights are $1 - \pi$ and π respectively.
3. **Plot the resulting mean and variance as π changes:**
   The mean of y_t will be a weighted sum of θ_0 and θ_1, and the variance will be 1 due to the normal distribution assumption, irrespective of π.
4. **Estimate the posterior distribution of θ with a normal prior:**
   If you're given a normal prior on θ, use Bayesian methods to update this prior with the data to obtain a posterior distribution. This can be done analytically if the prior is conjugate to the normal likelihood or through numerical methods like MCMC.
5. **Estimate π_0 when it is not given:**
   If π_0 is not given, estimate it from the data, likely by the proportion of actions equal to 1 in your dataset.
6. **Learning about other actions:**
   Regarding learning about other actions when you are not taking them, it refers to the extrapolation problem. If you have a model for how actions lead to outcomes, you can infer the potential outcomes of unobserved actions. However, the certainty of these inferences is typically much lower than for observed actions due to model and estimation uncertainties.

EXERCISE 10. Given your estimates, find the optimal policy for each one of those cases. Measure the quality of this policy on

- The actual data you have already (e.g. using importance sampling)
- On new simulations (using the testing framework).

Advanced: The optimal policy when θ is known is to always take the same action. Does that still hold when θ is not known and you are estimating it all the time from new data?

1. **Finding the Optimal Policy:**
   The optimal policy π^* would be the one that maximizes the expected outcome based on your model. If θ is known and represents the effect of taking action 1, then the optimal policy is to always take the action that maximizes the expected value of y_t. If θ_0 and θ_1 are the effects of actions 0 and 1 respectively, and $\theta_1 > \theta_0$, the optimal policy is to always take action 1, and vice versa.
2. **Measuring the Quality of the Policy on Actual Data:**
   You can use importance sampling to estimate the expected outcome of following the optimal policy on the actual data you have. This involves reweighting the outcomes in the actual data by the likelihood ratio of the action being taken under the optimal policy versus the behavior policy that generated the data.
3. **Measuring the Quality on New Simulations:**
   You would generate new data using the estimated model and then apply the optimal policy to this simulated data. The quality of the policy can be measured by the average outcome y_t when actions are taken according to π^*.
4. **Optimal Policy when θ is Unknown:**
   If θ is not known and you are estimating it from new data, the statement that the optimal policy is to always take the same action may not hold. The optimal policy in this case would depend on the current estimates of θ, which may change as new data is observed. If the estimates are uncertain, the optimal policy might involve exploration, i.e., sometimes taking different actions to learn more about their effects.

EXERCISE 11 (Advanced). Let us have some data generated from a null treatment policy π_0 of the form (x_t, a_t, y_t), with $a_t, x_t \in \{0, 1\}$.

$$y_t \mid a_t = a, x_t = x, \theta \sim \mathcal{N}(\theta_{a,x}, 1),$$

where the actions are distribution according to $\pi_0(a_t \mid x_t)$.

- Assume that π_0 is given and it is $a_t \mid x_t = x, \pi = \pi_0 \sim \mathit{Bernoulli}(w_x)$. First, estimate θ. Repeat your analysis.
- Now assume that π_0 is not given. Again, repeat your analysis.
- Is there now globally better action a_t? Should it depend on x_t, like in the observed policy? Can you estimate the optimal policy?

1. **When π_0 is given:**
   - **Estimate θ:** You'd estimate the parameters θ using maximum likelihood estimation or Bayesian methods, considering that you have the conditional distributions $\pi_0(a_t|x_t)$.
   - **Repeat analysis:** Use the estimated θ to analyze the outcomes under different policies by computing the expected value of y_t given x_t, according to the probabilities defined by π_0.
2. **When π_0 is not given:**
   - You'd first need to estimate π_0 itself, likely using the observed frequencies of actions for each covariate x_t in your dataset.
   - After estimating π_0, you would then re-estimate θ taking into account the estimated policy, and repeat your analysis of outcomes under this estimated policy.
3. **Determining if there is a globally better action a_t:**
   - Whether there's a globally better action would depend on how θ varies with x_t and a_t. If θ has a consistent trend (e.g., $\theta_{1,x} > \theta_{0,x}$ for all x), then the action associated with the higher θ would be globally better.
   - However, if θ varies with x_t, then the better action might also vary with x_t. In this case, the optimal policy would not be a simple constant action but a function of x_t, potentially mirroring the observed policy.
4. **Estimating the Optimal Policy:**
   - The optimal policy π^* would be the one that chooses the action that maximizes the expected outcome given the current estimate of θ and the covariate x_t.
   - To estimate π^*, you would use your estimates of θ to calculate the expected outcome for each possible action given each possible value of x_t and choose the action with the higher expected outcome.