Ivan Miller - Problem Set 5 for Adversarial AI. Tuesday, Oct. 25, 2022

1. Consider taking an (evenly weighted) average of the following data points:
x_1 = 10, x_2 = 5, x_3 = 8, x_4 = 9, x_5 = 2, x_6 = 5, x_7 = 3, x_8 = 15, x_9 = 12, x_10 = 7
What's the lowest average you can get if you leave out two data points, but those two missing data points are not allowed to be consecutive (e.g. if you remove x_4, you're not allowed to also remove _x5 or x_3.

The lowest possible average could be calculated after removing the highest value (X_8 = 15) and the value which has an index that fits the following condition: |X_8 - X_n| > 1.
The second highest value at X_9 = 12 does not fit the above criteria, so we move on to the value that is the highest of remaining: X_1 = 10 and remove it. The resulting average after removing X_8 = 15 and X_1 = 10 is 6.375 or 1.225 lower than the original average.

X_num:	X_8	X_9	X_1	X_4	X_3	X_10	X_2	X_6	X_7	X_5
Value:	~~15~~	12	~~10~~	9	8	7	5	5	3	2

Original average: 7.600
Modified average: 6.375

2. Suppose you are given N data points (like in the first problem, but now N is general instead of being =10). Describe an algorithm that takes in the N data points and outputs the lowest average value you can get from them if you are allowed to remove 2 data points with the constraint that the two removed can't be consecutive.

1. Sort values in descending order
2. Find highest value, remember its index (idx1), remove the value itself
3. Find the next available highest value, check its index (idx2). If the index difference between that value and the one we removed in step #2 is higher than 1 (indices are not consecutive, | idx1 - idx2| != 1) then remove the value corresponding to the index. Otherwise, move to the next available highest value and remove it.

3. Suppose you test 10 null hypotheses on the same data set. Suppose that all of the null hypotheses are true, and the probability of each test wrongly outputting "Reject the null hypothesis" is 5%. Assume the outcomes of the tests are independent events. What's the probability that you wrongly reject at least one null hypothesis?

We need to calculate the number of desired outcomes (errors) using the formula for binomial

$$P_x = \binom{n}{x} p^x q^{n-x}$$

probability:
Number of events n = 10
Number of desired outcomes X = 1
P$error$ = 0.05
Probability of one or more error:

$$\frac{10!}{1!*9!} * 0.05 * (1 - 0.05)^9 = 10 * 0.05 * 0.95^9 = 0.05 * 0.63 \sim 0.32$$