Exercise: Independent Random Variables

(2/2 points)

In this exercise, we look at how to check if two random variables are independent in Python. Please make sure that you can follow the math for what's going on and be able to do this by hand as well.

Consider random variables W, I, X, and Y, where we have shown the joint probability tables $p_{W,I}$ and $p_{X,Y}$.

		I = 1	I = 0
W	sunny	1/2	0
	rainy	0	1/6
	snowy	0	1/3

		Y = 1	Y = 0
X	sunny	1/4	1/4
	rainy	1/12	1/12
	snowy	1/6	1/6

In Python:

```
prob_W_I = np.array([[1/2, 0], [0, 1/6], [0, 1/3]])
```

Note that here, we are not explicitly storing the labels, but we'll keep track of them in our heads. The labels for the rows (in order of row index): sunny, rainy, snowy. The labels for the columns (in order of column index): 1, 0.

We can get the marginal distributions p_W and p_I:

```
prob_W = prob_W_I.sum(axis=1)
prob_I = prob_W_I.sum(axis=0)
```

Then if W and I were actually independent, then just from their marginal distributions p_W and p_I, we would be able to compute the joint distribution with the formula:

$$\textbf{If } W \textbf{ and } I \textbf{ are independent:} \qquad p_{W,I}(w,i) = p_W(w)p_I(i) \qquad \textbf{for all } w, i.$$

Note that variables `prob_W` and `prob_I` at this point store the probability tables p_W and p_I as 1D NumPy arrays, for which NumPy does *not* store whether each of these should be represented as a row or as a column.

We could however ask NumPy to treat them as column vectors, and in particular, taking the outer product of `prob_W` and `prob_I` yields what the joint distribution would be if W and I were independent:

$$\begin{bmatrix} p_W(\text{sunny}) \\ p_W(\text{rainy}) \\ p_W(\text{snowy}) \end{bmatrix} \begin{bmatrix} p_I(1) & p_I(0) \end{bmatrix} = \begin{bmatrix} p_W(\text{sunny})p_I(1) & p_W(\text{sunny})p_I(0) \\ p_W(\text{rainy})p_I(1) & p_W(\text{rainy})p_I(0) \\ p_W(\text{snowy})p_I(1) & p_W(\text{snowy})p_I(0) \end{bmatrix}.$$

The left-hand side is an outer product, and the right-hand side is precisely the joint probability table that would result if W and I were independent.

To compute and print the right-hand side, we do:

```
print(np.outer(prob_W, prob_I))
```

- Are W and I independent (compare the joint probability table we would get if they were independent with their actual joint probability table)?

  - [ ] Yes
  - [x] No ✔

- Are X and Y independent?

  - [x] Yes ✔
  - [ ] No

Solution:

- Are W and I independent (compare the joint probability table we would get if they were independent with their actual joint probability table)?

Solution: The answer is **No**. When you run the code above, you should see that the joint probability distribution for W and I is *different* from the joint probability of W and I if they were independent. In fact, if they were independent, you'd end up with the joint probability table for X and Y.

- Are X and Y independent?

  Solution: You can repeat the code above for X and Y to see that indeed X and Y are independent.

You have used 1 of 5 submissions