CSE 107
Lab Assignment 3

In this project you will simulate the process described in hw4 problem 2 and hw5 problem 1, quoted below.

> Joe Lucky plays the lottery on any given week with probability p, independently of whether he played on any other week. Each time he plays, he has a probability q of winning, again independently of everything else. During a fixed time period of n weeks, let X be the number of weeks that he played the lottery and Y be the number of weeks that he won.

Your goal is to determine, through means of simulation, the joint PMF $p_{X,Y}(x,y)$, the conditional PMF $p_{X|Y}(x|y)$ and the conditional PMF $p_{Y|X}(y|x)$, for all pairs (x,y) satisfying $0 \le y \le x \le n$. You will present your experimental values for these functions in three tables. The size and contents of these tables will of course depend on the parameters n, p and q. Your report, submitted as a pdf file, will compute probabilities corresponding to parameter values $\boldsymbol{n = 8, p = 0.6}$ **and** $\boldsymbol{q = 0.7}$. The following example demonstrates the required output format, and corresponds to parameters $n = 6$, $p = 0.5$ and $q = 0.5$.

Example $n = 6$, $p = 0.5$ and $q = 0.5$.

```
Joint PMF of X and Y
    y: 0     1     2     3     4     5     6
 x ---------------------------------------------------
 0 | .0158
 1 | .0467 .0469
 2 | .0601 .1157 .0588
 3 | .0390 .1168 .1167 .0391
 4 | .0149 .0592 .0893 .0576 .0140
 5 | .0030 .0151 .0285 .0290 .0148 .0030
 6 | .0003 .0017 .0037 .0046 .0038 .0015 .0002


Conditional PMF of X given Y
    y: 0     1     2     3     4     5     6
 x ---------------------------------------------------
 0 | .0877
 1 | .2599 .1321
 2 | .3344 .3255 .1981
 3 | .2168 .3285 .3928 .2999
 4 | .0828 .1667 .3008 .4420 .4297
 5 | .0167 .0425 .0960 .2225 .4542 .6667
 6 | .0018 .0047 .0124 .0356 .1161 .3333 1.0000


Conditional PMF of Y given X
    y: 0     1     2     3     4     5     6
 x ---------------------------------------------------
 0 | 1.0000
 1 | .4989 .5011
 2 | .2563 .4930 .2508
 3 | .1251 .3748 .3745 .1255
 4 | .0633 .2519 .3799 .2452 .0597
 5 | .0321 .1614 .3051 .3104 .1586 .0323
 6 | .0202 .1061 .2318 .2931 .2394 .0954 .0139
```

You will perform 100,000 trials of the above process to produce the relative frequencies reported in your tables. Thus, for each trial, and for each of n weeks, will flip a coin with probability p to decide if Joe plays that week. If so, increment the number of plays x and flip another coin with probability q to decide if Joe wins that week. If so, increment the number of plays y. When the trial is over, increment the frequency of the pair (x, y). When all 100,000 trials are complete, compute the relative frequencies constituting your estimates of the values for the joint PMF $p_{X,Y}(x, y)$. By summing on appropriate rows or columns, you can determine the marginal PMFs $p_X(x)$ and $p_Y(y)$. Upon dividing by the appropriate marginals, you can determine the conditional PMFs $p_{X|Y}(x|y)$ and $p_{Y|X}(y|x)$. All of your table entries should be rounded to 4 decimal places.

In principle this project is no more difficult than prior assignments, but there is a lot more to compute in this case. Appropriate data structures will be needed to store the frequencies and relative frequencies to do your calculations. If you want advice on how to make these choices, feel free to bring it up in office hours or discussion sections. Since there is so much more to do here, be sure to start early.