Jason Waseq
CSE 107
Lab 3

Instructions:

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 $pX,Y\ (x, y)$, the conditional PMF $pX|Y(x|y)$ and the conditional PMF $pY|X(y|x)$, for all pairs (x, y) satisfying $0 \leq y \leq x \leq n$. You will present your experimental values for these functions in three tables, whose size and contents depend on the parameters n, p and q. Your report will be submitted as a pdf file named lab3.pdf. It will present results corresponding to parameter values $\boldsymbol{n} = \mathbf{7}$, $\boldsymbol{p} = \mathbf{0.\ 6}$ and $\boldsymbol{q} = \mathbf{0.\ 7}$.

Each trial should simulate n weeks of play, as described above. Thus for each week, your simulation will flip a coin with probability p to decide if Joe plays that week. If so, increment the number of plays x, then flip another coin with probability q to decide if Joe wins that week. If so, increment the number of wins y. When the trial is over, increment the frequency of the pair (x, y). Continue the simulation for 100,000 trials. When all trials are complete, compute the relative frequencies constituting your estimates of the values for the joint PMF $pX,Y(x, y)$. By summing over appropriate rows or columns, you can determine the marginal PMFs $pX(x)$ and $pY(y)$. By dividing $pX,Y(x, y)$ by the appropriate marginals, you can determine the conditional PMFs $pX|Y(x|y)$ and $pY|X(y|x)$. All of your table entries should be rounded to 4 decimal places.

Code in Python:

```
import numpy as np

# Simulation parameters
n = 7  # Number of weeks
p = 0.6  # Probability of playing
q = 0.7  # Probability of winning if played
num_trials = 100000  # Number of simulations

# Storage for counts
joint_count = np.zeros((n + 1, n + 1))

# Run simulations
for _ in range(num_trials):
    x = 0  # Number of times played
    y = 0  # Number of wins
    for _ in range(n):
```

```
        if np.random.rand() < p:  # Joe plays this week
            x += 1
            if np.random.rand() < q:  # Joe wins this week
                y += 1
    joint_count[x, y] += 1

# Compute joint PMF
joint_pmf = joint_count / num_trials

# Compute marginal PMFs
px = np.sum(joint_pmf, axis=1)  # P(X=x)
py = np.sum(joint_pmf, axis=0)  # P(Y=y)

# Compute conditional PMFs
px_given_y = np.divide(joint_pmf, py, where=py > 0)  # P(X|Y)
py_given_x = np.divide(joint_pmf.T, px, where=px > 0).T  # P(Y|X)

# Function to format table output
def format_pmf_table(pmf):
    table = " y: " + "    ".join(f"{i:<8}" for i in range(n + 1)) + "\n"
    table += "x " + "-" * 82 + "\n"
    for i in range(n + 1):
        row = f"{i} | " + " ".join(f"{pmf[i, j]:.4f}" for j in range(n + 1) if j <= i)
        table += row + "\n"
    return table

# Print results
print("Joint PMF of X and Y")
print(format_pmf_table(joint_pmf))

print("Conditional PMF of X given Y")
print(format_pmf_table(px_given_y))

print("Conditional PMF of Y given X")
print(format_pmf_table(py_given_x))
```

Output:

```
Joint PMF of X and Y
y: 0      1      2      3      4      5      6      7
x --------------------------------------------------------------------------------
0 | 0.0015
1 | 0.0049 0.0123
2 | 0.0072 0.0327 0.0377
3 | 0.0050 0.0368 0.0867 0.0665
4 | 0.0025 0.0223 0.0771 0.1194 0.0689
5 | 0.0007 0.0074 0.0349 0.0782 0.0941 0.0447
6 | 0.0001 0.0012 0.0078 0.0235 0.0434 0.0395 0.0152
7 | 0.0000 0.0001 0.0008 0.0029 0.0058 0.0084 0.0075 0.0022

Conditional PMF of X given Y
y: 0      1      2      3      4      5      6      7
x --------------------------------------------------------------------------------
0 | 0.0696
1 | 0.2244 0.1091
2 | 0.3286 0.2904 0.1538
3 | 0.2262 0.3263 0.3540 0.2290
4 | 0.1129 0.1977 0.3146 0.4111 0.3249
5 | 0.0337 0.0653 0.1426 0.2691 0.4433 0.4826
6 | 0.0046 0.0106 0.0317 0.0809 0.2045 0.4266 0.6693
7 | 0.0000 0.0006 0.0033 0.0099 0.0273 0.0908 0.3307 1.0000

Conditional PMF of Y given X
y: 0      1      2      3      4      5      6      7
x --------------------------------------------------------------------------------
0 | 1.0000
1 | 0.2861 0.7139
2 | 0.0930 0.4216 0.4854
3 | 0.0255 0.1886 0.4447 0.3412
4 | 0.0085 0.0768 0.2655 0.4116 0.2376
5 | 0.0028 0.0283 0.1343 0.3007 0.3618 0.1720
6 | 0.0008 0.0091 0.0594 0.1799 0.3320 0.3023 0.1165
7 | 0.0000 0.0025 0.0289 0.1044 0.2095 0.3038 0.2720 0.0788
```

Conjecture:
In this experiment, Joe's lottery play follows a predictable pattern: the more he plays, the more he can win, but his wins are always limited by how often he plays.

- The joint PMF shows that Joe is most likely to play around np times, and his number of wins is always less than or equal to his plays.
- The conditional PMFs reveal that if we know how many times he played, we can estimate his wins, and vice versa.

Overall, the results confirm that Joe’s play and win patterns follow a structured probability distribution, where wins depend entirely on play frequency, and both follow binomial-like behavior.