Practice questions

1. Roll a 3-sided die twice. Let X be the larger number and Y be the smaller number you rolled. Find (a) the conditional PMF of X given Y and (b) the expected value of X given $Y = y$ for all values of y.

   Solution: From Homework 4 question 1, we have the joint PMF $p_{XY}(x, y)$:

$x\backslash y$	1	2	3
1	1/9	0	0
2	2/9	1/9	0
3	2/9	2/9	1/9

   and marginal PMF:

y	1	2	3
$p_Y(y)$	5/9	1/3	1/9

   $p_{X|Y}(x|y) = \frac{p_{XY}(x,y)}{p_Y(y)}$, we have the conditional PMF $p_{X|Y}(x|y)$:

$x\backslash y$	1	2	3
1	1/5	0	0
2	2/5	1/3	0
3	2/5	2/3	1

   $E[X|Y = 1] = \frac{11}{5}$, $E[X|Y = 2] = \frac{8}{3}$, $E[X|Y = 3] = 3$.

2. In 2017 there were 0.848 men for every woman in Hong Kong. Men and women had life expectancies of 81.7 years and 87.7 years, respectively. What was the life expectancy of a random person?

   Solution: Suppose a random person in 2017 lives L years. Let M be the event that person is a man. The life expectancies of men is the expected value of L, conditioned on that person is male, i.e. $\mathrm{E}[L \mid M]$. Assume every person in Hong Kong has equal probability to get picked, then $P(M)$ is the ratio of men to the entire population. By the law of total expectation, $\mathrm{E}[L] = \mathrm{E}[L \mid M]P(M) + \mathrm{E}[L \mid M^C]P(M^C) = 81.7 \cdot 0.848/1.848 + 87.7 \cdot 1/1.848 \approx 84.9$.

3. You toss a coin 100 times. Which of the following random variables are independent?

   (a) The number of consecutive heads `HH` and the number of consecutive tails `TT`.

   (b) The number of consecutive heads in the first 50 tosses and the number of consecutive tails in the last 50 tosses.

   (c) The random variables in part (b), conditioned on having exactly 50 heads in the 100 coin tosses.

   Solution: We denote the random variables by X and Y in each part.

(a) Not independent. $P(X = 99, Y = 99) = 0$ as there cannot be 99 consecutive heads and 99 consecutive tails. However, $P(X = 99) > 0$ and $P(Y = 99) > 0$ as each of these events may occur individually (and has probability 2^{-100}. Therefore $P(X = 99, Y = 99) \neq P(X = 99)\,P(Y = 99)$.

(b) Independent. It is not easy to calculate these numbers, but we can reason it out. The probability of the event $Y = y$ does not depend on what happens in the first 50 coin tosses, so all the conditional probabilities $P(Y = y|X = 0), P(Y = y|X = 1), \dots, P(Y = y|X = 49)$ have the same value p. By the total probability theorem,

$$\begin{aligned} P(Y = y) &= P(Y = y|X = 0)\,P(X = 0) + \cdots + P(Y = y|X = 49)\,P(X = 49) \\ &= p\,P(X = 0) + \cdots + p\,P(X = 49) \\ &= p, \end{aligned}$$

so $P(Y = y|X = x)$ and $P(Y = y)$ are always the same.

(c) Not independent. Let E be the event we are conditioning on. Conditioned on A, all $\binom{100}{50}$ balanced sequences of heads and tails are equally likely. In particular, $P(X = 49|A) = 1/\binom{100}{50}$, as $X = 49$ can occur in one possible way given A. For the same reason, $P(Y = 49|A) = 1/\binom{100}{50}$. But $P(X = 49, Y = 49|A)$ is also $1/\binom{100}{50}$. Therefore $P(X = 49, Y = 49|A) \neq P(X = 49|A)\,P(Y = 49|A)$ and so the two are not independent.

4. You go to the casino with \$3 to play roulette. (Roulette has 37 possible outcomes, out of which 18 are red, 18 are black, and one is green.) Calculate the expected value and standard deviation of your profit under the following two gambling strategies:

(a) You play for 3 rounds, where in every round you bet \$1 on red.

(b) You bet all your money on red. If you win, you bet everything on red again. If you win again, you bet everything on red one last time.

Solution:

(a) The profits X_1, X_2, X_3 in the three rounds are independent random variables with the following PMF:

x	-1	1
$p(x)$	$19/37$	$18/37$

Each of them has expected value $E[X_i] = -1/37$ and variance $\mathrm{Var}[X_i] = E[X_i^2] - E[X_i]^2 = 1 - 1/37^2 \approx 0.9993$. By linearity of expectation, the total profit $X = X_1 + X_2 + X_3$ has expected value $E[X] = E[X_1] + E[X_2] + E[X_3] = -3/37 \approx -0.0811$. By independence, $\mathrm{Var}[X] = \mathrm{Var}[X_1] + \mathrm{Var}[X_2] + \mathrm{Var}[X_3] \approx 2.9978$ so the standard deviation of X is $\sqrt{2.9978} \approx 1.7314$.

(b) Let Y be the profit after 3 games. You either earn \$21 by winning all 3 games, or lose \$3 otherwise. Therefore the PMF of Y is

y	-3	21
$p(y)$	$1 - (18/37)^3$	$(18/37)^3$

The expected value of Y is $E[Y] = -3(1 - (\frac{18}{37})^3) + 21(\frac{18}{37})^3 \approx -0.2367$. The expected value of Y^2 is $E[Y^2] = (-3)^2(1 - (\frac{18}{37})^3) + 21^2(\frac{18}{37})^3 \approx 58.7389$, so the variance is $\mathrm{Var}[Y] = E[Y^2] - E[Y]^2 \approx 58.6828$ and the standard deviation is about $\sqrt{58.6828} \approx 7.6605$.

In conclusion, the expected profit in strategy 1 is higher (the expected loss is lower), but strategy 1 is more risk-averse as the standard deviation in strategy 2 is a lot higher.

5. Consider 10 persons forming 5 couples who live together at a given time. Suppose that at some later time, the probability of each person being alive is p, independent of other persons. At that later time, let A be the number of persons that are alive and let S be the number of couples in which both partners are alive. Find $\mathrm{E}[S \mid A = a]$. *(Textbook problem 2.32)*

Solution: Let X_i be the random variable taking the value 1 or 0 depending on whether the first partner of the ith couple has survived or not. Let Y_i be the corresponding random variable for the second partner of the ith couple. Then, we have $S = X_1Y_1 + \cdots + X_5Y_5$ and by using the total expectation theorem, for any a,

$$\begin{aligned} \mathrm{E}[S \mid A = a] &= \mathrm{E}[X_1Y_1 \mid A = a] + \cdots + \mathrm{E}[X_5Y_5 \mid A = a] && (1) \\ &= 5\,\mathrm{E}[X_1Y_1 \mid A = a] && (2) \\ &= 5\,\mathrm{E}[Y_1 \mid X_1 = 1, A = a]\,\mathrm{P}(X_1 = 1 \mid A = a) && (3) \\ &= 5\,\mathrm{P}(Y_1 = 1 \mid X_1 = 1, A = a)\,\mathrm{P}(X_1 = 1 \mid A = a) && (4) \end{aligned}$$

Here, equation (1) is due to linearity of expectation. (Linearity of expectation works for conditional expectations as long as we don't change the conditional.) Equation (3) is due to total expectation theorem and the expectation $\mathrm{E}[X_1Y_1 \mid X_1 = 0, A = a] = 0$. In equation (4) we replace the expectation of an indicator (0-1) random variable with the probability that it takes value 1.

We can calculate $\mathrm{P}(Y_1 = 1 \mid X_1 = 1, A = a)$: This is the probability that my partner has survived, given that I have survived and a people have survived. As all 9 people have the same probability to be among the $a-1$ other survivors, the probability that my partner made it is $(a-1)/9$. We can similarly calculate $\mathrm{P}(X_1 = 1 \mid A = a)$ as $a/10$, as everyone including me is equally likely to be among the a survivors. Therefore $\mathrm{E}[S \mid A = a] = a(a-1)/2 \cdot 9 = a(a-1)/18$.