Practice questions

1. Let X and Y be independent random variables with PMFs $P(X = 1) = P(X = 2) = P(X = 3) = 1/3$ and $P(Y = 3) = P(Y = 4) = 1/2$. Let $M = X + Y$ and $N = Y - X$.

   (a) What is the PMF of M given N?

   Solution: We first calculate the joint PMF $p_{MN}(m, n)$: of M and N:

$m\backslash n$	0	1	2	3
4	0	0	$\frac{1}{6}$	0
5	0	$\frac{1}{6}$	0	$\frac{1}{6}$
6	$\frac{1}{6}$	0	$\frac{1}{6}$	0
7	0	$\frac{1}{6}$	0	0

   Then we calculate the marginal PMF $p_N(n)$ of N:

n:	0	1	2	3
$p_N(n)$:	$\frac{1}{6}$	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{6}$

   The conditional PMF $p_{M|N}(M = m|N = n)$ is then $\frac{p_{MN}(m,n)}{p_N(n)}$:

$m\backslash n$	0	1	2	3
4	0	0	$\frac{1}{2}$	0
5	0	$\frac{1}{2}$	0	1
6	1	0	$\frac{1}{2}$	0
7	0	$\frac{1}{2}$	0	0

   (b) Are M and N independent? Justify your answer.

   Solution: No. For example, $p_{MN}(4, 0) = 0$, $p_M(4) = 1/6$, and $p_N(0) = 1/6$, so $p_{MN}(4, 0) \neq p_M(4)p_N(0)$.

   (c) What is the expectation of M given $N < 2$?

   Solution: We first calculate the joint PMF $p_{MN|N<2}(m, n)$ of M and N given $N < 2$. This is obtained from the joint PMF of M and N by discarding the columns $n = 2$ and $n = 3$ and rescaling so that the probabilities add up to one:

$m\backslash n$	0	1
4	0	0
5	0	$\frac{1}{3}$
6	$\frac{1}{3}$	0
7	0	$\frac{1}{3}$

   We then obtain the conditional PMF $p_{M|N<2}(m)$ by adding up the rows of this table:

m:	5	6	7
$p_{M|N<2}(m)$:	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{3}$

   The conditional expectation is then

$$E[M|N < 2] = 5 \cdot \frac{1}{3} + 6 \cdot \frac{1}{3} + 7 \cdot \frac{1}{3} = 6.$$

2. A fair coin is tossed 9 times. Which of the following random variables are independent?

   (a) The number of consecutive heads (`HH`) in the first 5 tosses and the number of consecutive tails (`TT`) in the last 5 tosses.

Solution: Not independent. $\mathrm{P}(X = 4, Y = 4) = 0$ while $\mathrm{P}(X = 4) \cdot \mathrm{P}(Y = 4) = 2^{-5} \cdot 2^{-5} = 2^{-10}$.

(b) The number of changes in the first 5 tosses and the number of changes in the last 5 tosses. A *change* is a head followed by a tail or a tail followed by a head (`TH` or `HT`).

Solution: They are independent. Let F and L be the random variables of interest and E be the event "the fifth toss is a head." By the total probability theorem,

$$\mathrm{P}(F = f, L = l) = \mathrm{P}(F = f, L = l|E)\,\mathrm{P}(E) + \mathrm{P}(F = f, L = l|E^c)\,\mathrm{P}(E^c).$$

The events $F = f$ and $L = l$ are conditionally independent given E, so $\mathrm{P}(F = f, L = l|E) = \mathrm{P}(F = f|E)\,\mathrm{P}(L = l|E)$. The events $F = f$ and $L = l$ are individually independent of E because the fifth toss being a head does not affect the number of changes in the first and last five tosses. Therefore $\mathrm{P}(F = f|E)\,\mathrm{P}(L = l|E) = \mathrm{P}(F = f)\,\mathrm{P}(L = l)$. By the same reasoning $\mathrm{P}(F = f, L = l|E^c) = \mathrm{P}(F = f)\,\mathrm{P}(L = l)$. Therefore

$$\mathrm{P}(F = f, L = l) = \mathrm{P}(F = f)\,\mathrm{P}(L = l)\,\mathrm{P}(E) + \mathrm{P}(F = f)\,\mathrm{P}(L = l)\,\mathrm{P}(E^c) = \mathrm{P}(F = f)\,\mathrm{P}(L = l).$$

(c) (**Optional**) The number of tails followed by a head (`TH`) in the first 5 tosses and the number of heads followed by a tail (`HT`) in the last 5 tosses.

Solution: Not independent. The event $X = 0$ can happen if the first five flips consists of zero or more heads followed by tails. By the equally likely outcomes formula $\mathrm{P}(X = 0) = 6/2^5$. By symmetry $\mathrm{P}(Y = 0) = 6/2^5$. On the other hand, by the product rule the event "$X = 0$ and $Y = 0$" has $5 \cdot 5$ possible outcomes if there is least one tail in the first half and one more outcome (all heads) otherwise. Therefore $\mathrm{P}(X = 0, Y = 0) = (5^2 + 1)/2^9$, which is different from $\mathrm{P}(X = 0)\,\mathrm{P}(Y = 0) = 6^2/2^{10}$.

3. 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.

Solution: 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 $\mathrm{E}[X_i] = -1/37$ and variance $\mathrm{Var}[X_i] = \mathrm{E}[X_i^2] - \mathrm{E}[X_i]^2 = 1 - 1/37^2 \approx 0.9993$. By linearity of expectation, the total profit $X = X_1 + X_2 + X_3$ has expected value $\mathrm{E}[X] = \mathrm{E}[X_1] + \mathrm{E}[X_2] + \mathrm{E}[X_3] = -3/37 \approx -0.0811$. 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) 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: 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 $\mathrm{E}[Y] = -3(1-(\frac{18}{37})^3) + 21(\frac{18}{37})^3 \approx -0.2367$. 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 $\mathrm{Var}[Y] = \mathrm{E}[Y^2] - \mathrm{E}[Y]^2 \approx 58.6828$ 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.

4. Find the expectation of the following random variables without calculating their PMF:

   (a) The first time X at which both the patterns `TH` and `HT` have appeared in a sequence of fair coin flips.

   Solution: We can write $X = 1 + T_1 + T_2$, where T_1 is the time after the first flip it takes for the first pattern to appear, and T_2 is the additional time we have to wait for the second pattern. Then both T_1 and T_2 are Geometric(1/2) random variables, so $\mathrm{E}[T_1] = \mathrm{E}[T_2] = 2$. By linearity of expectation, $\mathrm{E}[T] = 1 + 2 + 2 = 5$.

   (b) The first time Y at which all three face values have appeared in a sequence of rolls of a fair 3-sided die.

   Solution: We write $Y = 1 + R_2 + R_3$, where R_2 is the time after the first roll that the second value appears and R_3 is the remaining time it takes for the last value to appear. Then R_2 is a Geometric(2/3) random variable, while R_3 is a Geometric(1/3) random variable. By linearity of expectation,

$$\mathrm{E}[Y] = 1 + \mathrm{E}[R_2] + \mathrm{E}[R_3] = 1 + 3 + \frac{3}{2} = \frac{11}{2}.$$

Additional ESTR 2018 questions

5. Let T be the number of times a 6-sided die is rolled until a 6 appears. What is the expected value of T conditioned on all rolls producing even numbers? *(due to E. Mossel)*

   Solution: 3/2.