E4.1. There are 274 students on the list of the module MATH1715. Denote by \( X \) the number of students who have their birthday on the 15th of November. Compare the probabilities \( P(X = k) \) for a few values \( k = 0, 1, 2, \ldots \) with those given by the Poisson approximation.

**Solution.** We assume that a student’s birthday falls on 15 November with probability \( \frac{1}{365} \), independently of other students. That is, we have a sequence of \( n = 274 \) Bernoulli trials with probability of success \( p = \frac{1}{365} = 0.0027397 \ldots \) Then the random variable \( X \) (the number of students, out of 274, having birthday on 15 November) has the binomial distribution with parameters \( n = 274, p = \frac{1}{365} \), so that

\[
P(X = k) = \binom{274}{k} \left( \frac{1}{365} \right)^k \left( \frac{364}{365} \right)^{274-k}, \quad k = 0, 1, \ldots, 274.
\]

Set \( \lambda = np = \frac{274}{365} = 0.750685 \), then the Poisson approximation yields

\[
P(X = k) \approx \frac{\lambda^k}{k!} e^{-\lambda}, \quad k = 0, 1, 2, \ldots
\]

The numerical results for a few small values of \( k \) are gathered in the following table:

<table>
<thead>
<tr>
<th>( k )</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>( \geq 6 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Binomial</td>
<td>0.47156</td>
<td>0.35496</td>
<td>0.13311</td>
<td>0.03316</td>
<td>0.00617</td>
<td>0.00092</td>
<td>0.00013</td>
</tr>
<tr>
<td>Poisson</td>
<td>0.47204</td>
<td>0.35436</td>
<td>0.13300</td>
<td>0.03328</td>
<td>0.00625</td>
<td>0.00094</td>
<td>0.00013</td>
</tr>
</tbody>
</table>

Here \( P(X \geq 6) \) may be evaluated using the formula

\[
P(X \geq 6) = 1 - P(X \leq 5) = 1 - \sum_{k=0}^{5} P(X = k).
\]

Note that, as known from theory, the error in the Poisson approximation is estimated by

\[
\frac{\lambda^2}{n} = np^2 = \frac{274}{365^2} \leq 0.0021,
\]

but the actual error is even better than that.

E4.2. Let a random variable \( X \) have the geometric distribution with parameter \( p \), that is,

\[
P(X = k) = p(1 - p)^{k-1}, \quad k = 1, 2, \ldots
\]

Find the probability that \( X \) is greater than \( n \), where \( n = 0, 1, 2, \ldots \)

**Solution.** Using the formula for the sum of geometric series, we obtain

\[
P(X > n) = \sum_{k=n+1}^{\infty} p(1 - p)^{k-1} = p \left\{ (1 - p)^n + (1 - p)^{n+1} + \cdots \right\}
\]

\[
= p(1-p)^n \left\{ 1 + (1 - p) + (1-p)^2 + \cdots \right\}
\]

\[
= p(1-p)^n \cdot \frac{1}{1-(1-p)} = p(1-p)^n \cdot \frac{1}{p} = (1-p)^n.
\]
So the answer is given by

\[ P(X > n) = (1 - p)^n. \]

Note that for \( n = 0 \) this gives an obvious result \( P(X > 0) = (1 - p)^0 = 1 \).

**E4.3.** Suppose that the random variables \( X \) and \( Y \) are independent, and let \( X \) and \( Y \) have Poisson distributions with parameters \( \lambda \) and \( \mu \), respectively.

(a) Show that the sum \( Z = X + Y \) has the Poisson distribution with parameter \( \lambda + \mu \).

(b) Find the conditional probability \( P(X = k \mid Z = n) \) for \( k = 0, 1, \ldots, n \).

**Solution.** (a) We need to show that

\[ P(Z = n) = \frac{(\lambda + \mu)^n}{n!} e^{-(\lambda + \mu)}, \quad n = 0, 1, 2, \ldots \quad (1) \]

Note that in order for the sum \( Z = X + Y \) to take the value \( n \), the random variable \( X \) may take any value \( k = 0, 1, \ldots, n \) (we use that both \( X \) and \( Y \) are non-negative). Therefore,

\[ P(Z = n) = P(X + Y = n) = \sum_{k=0}^{n} P(X = k, X + Y = n) = \sum_{k=0}^{n} P(X = k, Y = n-k). \]

Due to independence of \( X \) and \( Y \) we can rewrite the right-hand side in the form

\[
\begin{align*}
\sum_{k=0}^{n} P(X = k) \cdot P(Y = n-k) &= \sum_{k=0}^{n} \frac{\lambda^k}{k!} e^{-\lambda} \cdot \frac{\mu^{n-k}}{(n-k)!} e^{-\mu} \\
&= e^{-(\lambda + \mu)} \sum_{k=0}^{n} \frac{1}{k!(n-k)!} \lambda^k \mu^{n-k} \\
&= \frac{1}{n!} e^{-(\lambda + \mu)} \sum_{k=0}^{n} \frac{n!}{k!(n-k)!} \lambda^k \mu^{n-k} \\
&= \frac{1}{n!} e^{-(\lambda + \mu)} (\lambda + \mu)^n,
\end{align*}
\]

where at the last step we used the binomial theorem. Hence,

\[ P(Z = n) = \frac{(\lambda + \mu)^n}{n!} e^{-(\lambda + \mu)}, \]

and formula (1) is proved.

(b) By definition of conditional probabilities and using the result of part (a) we have

\[
\begin{align*}
P(X = k \mid Z = n) &= \frac{P(X = k, X + Y = n)}{P(Z = n)} = \frac{P(X = k, Y = n-k)}{P(Z = n)} \\
&= \frac{P(X = k) \cdot P(Y = n-k)}{P(Z = n)} = \frac{\lambda^k e^{-\lambda}}{k!} \cdot \frac{\mu^{n-k} e^{-\mu}}{(n-k)!} \cdot \frac{n!}{(\lambda + \mu)^n e^{-(\lambda + \mu)}} \\
&= \frac{n!}{k!(n-k)!} \cdot \frac{\lambda^k}{(\lambda + \mu)^k} \cdot \frac{\mu^{n-k}}{(\lambda + \mu)^{n-k}} = \left( \frac{n}{k} \right) \left( \frac{\lambda}{\lambda + \mu} \right)^k \left( \frac{\mu}{\lambda + \mu} \right)^{n-k}.
\end{align*}
\]

So the conditional probabilities \( P(X = k \mid Z = n) \) are given by the binomial distribution with parameters \( n \) and \( p = \lambda/\lambda + \mu \).
E4.4. At each flip, a certain coin shows heads with probability \( p \) and tails with probability \( 1 - p \). Suppose that you flip the coin a random number \( N \) of times, where \( N \) has the Poisson distribution with parameter \( \lambda \) and is independent of the outcomes of the flips. Let \( X \) be the number of resulting heads and \( Y \) the number of resulting tails. (If \( N = 0 \) then we don’t flip the coin and so in this case \( X = Y = 0 \).) Find the distributions of \( X \) and \( Y \) and show that \( X \) and \( Y \) are independent.

Remark. This is quite surprising, because if the number of trials is non-random, \( N = n \), then the random variables \( X \) and \( Y \) are not independent because their sum is fixed, \( X + Y = n \), and hence there is a relation between them. So an additional “randomization” of the experiment may lead to additional independence.

Solution. Let us find the probability \( P(X = i) \) \((i = 0, 1, 2, \ldots)\). The idea is to condition on the value of \( N \), because if the number of trials is fixed then the number of heads is a binomial random variable. Noting that the number of trials \( N \) cannot be less than the number of heads observed, and applying the formula of total probability we obtain

\[
P(X = i) = \sum_{n=i}^{\infty} P(X = i \mid N = n) \cdot P(N = n)
\]

\[
= \sum_{n=i}^{\infty} \binom{n}{i} p^i (1 - p)^{n-i} \cdot \frac{\lambda^n}{n!} e^{-\lambda}
\]

\[
= \sum_{n=i}^{\infty} \frac{n!}{(n-i)!i!} p^i (1 - p)^{n-i} \cdot \frac{\lambda^n}{n!} e^{-\lambda}
\]

\[
= \frac{p^i}{i!} e^{-\lambda} \sum_{n=i}^{\infty} \frac{1}{(n-i)!} (1 - p)^{n-i} \cdot \lambda^n.
\]

Introduce another summation index \( k = n - i \) and rewrite expression (12) as

\[
\frac{p^i}{i!} e^{-\lambda} \sum_{k=0}^{\infty} \frac{1}{k!} (1 - p)^k \lambda^{i+k} = \frac{\lambda^i p^i}{i!} e^{-\lambda} \sum_{k=0}^{\infty} \frac{[\lambda(1-p)]^k}{k!}
\]

\[
= \frac{(\lambda p)^i}{i!} e^{-\lambda} \cdot e^{\lambda(1-p)} = \frac{(\lambda p)^i}{i!} e^{-\lambda + \lambda(1-p)} = \frac{(\lambda p)^i}{i!} e^{-\lambda p}.
\]

Thus \( X \) has the Poisson distribution with parameter \( \lambda p \). Similarly, one can show that \( Y \), the number of tails, has the Poisson distribution with parameter \( \lambda(1-p) \).

To check that \( X \) and \( Y \) are independent, consider the joint probability

\[
P(X = i, Y = j) = P(X = i, N = i + j)
\]

\[
= P(X = i \mid N = i + j) \cdot P(N = i + j)
\]

\[
= \binom{i + j}{i} p^i (1 - p)^j \cdot \frac{\lambda^{i+j}}{(i+j)!} e^{-\lambda}
\]

\[
= \frac{(i+j)!}{i!j!} p^i (1 - p)^j \cdot \frac{\lambda^{i+j}}{(i+j)!} e^{-\lambda}
\]

\[
= \frac{(\lambda p)^i}{i!} e^{-\lambda p} \cdot \frac{[\lambda(1-p)]^j}{j!} e^{-\lambda(1-p)}
\]

\[
= P(X = i) \cdot P(Y = j),
\]

and the definition of independence is satisfied.

E4.5. Two fair dice are rolled once. Let \( X \) be the smallest value obtained on the dice and \( Y \) be the sum of the two values, so that \( X = \min(X_1, X_2) \), \( Y = X_1 + X_2 \), where \( X_1, X_2 \) are the readings on the dice.

(a) Find the joint probability mass function of \( X \) and \( Y \).
(b) Find the marginal distributions of $X$ and $Y$.

(c) What is the expected value of $Y - X$?

**Solution.** (a) Note that if the random variable $X = \min(X_1, X_2)$ assumes value $i$ ($1 \leq i \leq 6$), then the random variable $Y = X_1 + X_2$ may take values $j \in \{i + i, i + (i + 1), \ldots, i + 6\}$. If $j = 2i$ then

$$P(X = i, Y = 2i) = P(X_1 = i, X_2 = i) = \frac{1}{36},$$

whereas for $j \geq 2i + 1$ we have

$$P(X = i, Y = j) = P(X_1 = i, X_2 = j - i) + P(X_1 = j - i, X_2 = i) = \frac{1}{36} + \frac{1}{36} = \frac{1}{18}.$$

Hence, the joint probability mass function $p_{XY}(i, j) = P(X = i, Y = j)$ is given by the following table:

\[
\begin{array}{cccccccccccc}
X & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & 11 & 12 \\
\hline
1 & \frac{1}{36} & \frac{1}{18} & \frac{1}{18} & \frac{1}{18} & \frac{1}{18} & 0 & 0 & 0 & 0 & 0 & 0 \\
2 & 0 & 0 & \frac{1}{36} & \frac{1}{18} & \frac{1}{18} & \frac{1}{18} & \frac{1}{18} & 0 & 0 & 0 & 0 \\
3 & 0 & 0 & 0 & 0 & \frac{1}{36} & \frac{1}{18} & \frac{1}{18} & \frac{1}{18} & 0 & 0 & 0 \\
4 & 0 & 0 & 0 & 0 & 0 & \frac{1}{36} & \frac{1}{18} & \frac{1}{18} & 0 & 0 & 0 \\
5 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{36} & \frac{1}{18} & 0 & 0 \\
6 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{36} & 0 & 0
\end{array}
\]

(b) The marginal distributions of $X$ and $Y$ are obtained by summing up columns and rows of the above table, that is,

$$p_X(i) = \sum_{j=1}^{6} p_{XY}(i, j),$$

$$p_Y(j) = \sum_{i=2}^{12} p_{XY}(i, j).$$

Therefore, we find

\[
\begin{array}{cccccccc}
i & 1 & 2 & 3 & 4 & 5 & 6 \\
p_X(i) & \frac{11}{36} & \frac{1}{4} & \frac{7}{36} & \frac{5}{36} & \frac{1}{12} & \frac{1}{36} \\
\hline
j & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & 11 & 12 \\
p_Y(j) & \frac{1}{36} & \frac{1}{18} & \frac{1}{12} & \frac{1}{9} & \frac{5}{36} & \frac{1}{6} & \frac{5}{36} & \frac{1}{9} & \frac{1}{12} & \frac{1}{18} & \frac{1}{36}
\end{array}
\]

(c) We use the formula

$$E(Y - X) = E(Y) - E(X).$$

Each of the expectations on the right-hand side can be computed using the marginal distributions:

$$E(X) = 1 \cdot \frac{11}{36} + 2 \cdot \frac{1}{4} + 3 \cdot \frac{7}{36} + 4 \cdot \frac{5}{36} + 5 \cdot \frac{1}{12} + 6 \cdot \frac{1}{36} = \frac{91}{36} = 2.52777 \ldots.$$
Finally, comparing equations (3) and (4), we obtain the equation

\[ \sigma_1^2 + \sigma_2^2 = \sigma_1^2 + 2\sigma_2^2 + \sigma_3^2 + 2 \text{Cov}(Y_1, Y_2), \]

whence

\[ \text{Cov}(Y_1, Y_2) = -\sigma_2^2. \]

Finally,

\[ \text{Cor}(Y_1, Y_2) = \frac{\text{Cov}(Y_1, Y_2)}{\sqrt{\text{Var}(Y_1) \text{Var}(Y_2)}} = -\frac{\sigma_2^2}{\sqrt{(\sigma_1^2 + \sigma_2^2)(\sigma_2^2 + \sigma_3^2)}}. \]

Note that the expectation \( E(Y) \) could be computed easier using that \( Y = X_1 + X_2 \), where \( X_1 \) and \( X_2 \) are the values obtained on the dice:

\[ E(Y) = E(X_1 + X_2) = E(X_1) + E(X_2) = 3.5 + 3.5 = 7. \]

Hence,

\[ E(Y - X) = E(Y) - E(X) = 7 - \frac{91}{36} = \frac{161}{36} = 4.47222 \ldots \]

**Remark.** Note that

\[ Y - X = X_1 + X_2 - \min(X_1, X_2) = \max(X_1, X_2). \]

It is therefore not surprising that while

\[ E(X) = E(\min(X_1, X_2)) = 2.52778 < 3.5, \]

we have

\[ E(Y - X) = E(\max(X_1, X_2)) = 4.47222 > 3.5, \]

where \( 3.5 = E(X_1) = E(X_2) \).

**E4.6.** Let \( X_1, X_2, \) and \( X_3 \) be independent random variables having finite non-zero variances \( \sigma_1^2, \sigma_2^2 \) and \( \sigma_3^2 \), respectively. Find the correlation between \( Y_1 = X_1 - X_2 \) and \( Y_2 = X_2 + X_3 \).

**Solution.** Since \( X_i, X_j \, (i \neq j) \) are independent, we have

\[ \text{Var}(Y_1) = \text{Var}(X_1 - X_2) = \text{Var}(X_1) + \text{Var}(X_2) = \sigma_1^2 + \sigma_2^2, \]

\[ \text{Var}(Y_2) = \text{Var}(X_2 + X_3) = \text{Var}(X_2) + \text{Var}(X_3) = \sigma_2^2 + \sigma_3^2. \]

We also note that \( Y_1 + Y_2 = (X_1 - X_2) + (X_2 + X_3) = X_1 + X_3 \) and hence

\[ \text{Var}(Y_1 + Y_2) = \text{Var}(X_1 + X_3) = \sigma_1^2 + \sigma_3^2. \]

(3)

On the other hand, a general formula for the variance of the sum yields

\[ \text{Var}(Y_1 + Y_2) = \text{Var}(Y_1) + \text{Var}(Y_2) + 2 \text{Cov}(Y_1, Y_2) = \sigma_1^2 + \sigma_2^2 + \sigma_2^2 + \sigma_3^2 + 2 \text{Cov}(Y_1, Y_2) \]

\[ = \sigma_1^2 + 2\sigma_2^2 + \sigma_3^2 + 2 \text{Cov}(Y_1, Y_2). \]

(4)

Comparing equations (3) and (4), we obtain the equation

\[ \sigma_1^2 + \sigma_3^2 = \sigma_1^2 + 2\sigma_2^2 + \sigma_3^2 + 2 \text{Cov}(Y_1, Y_2), \]

whence

\[ \text{Cov}(Y_1, Y_2) = -\sigma_2^2. \]

Finally,

\[ \text{Cor}(Y_1, Y_2) = \frac{\text{Cov}(Y_1, Y_2)}{\sqrt{\text{Var}(Y_1) \text{Var}(Y_2)}} = -\frac{\sigma_2^2}{\sqrt{(\sigma_1^2 + \sigma_2^2)(\sigma_2^2 + \sigma_3^2)}}. \]