1 (March 14) Let $X \sim$ Binomial(n, p), where $n \geq 3$ and $p = 1/2$. Show that $\mathbb{P}(X = k) = \mathbb{P}(X = n - k)$ for all $k \in \{0, \ldots, n\}$.

Answer: By using the formula from the course book (see page 12), we get:

$$\begin{aligned}
\frac{\mathbb{P}(X = k)}{\mathbb{P}(X = n - k)} &= \frac{\binom{n}{k}(1/2)^k(1/2)^{n-k}}{\binom{n}{n-k}(1/2)^{n-k}(1/2)^k} \\
&= \frac{\frac{n!}{k!(n-k)!}(1/2)^n}{\frac{n!}{(n-k)!(n-n+k)!}(1/2)^n} \\
&= 1.
\end{aligned}$$

Comment: Have a look at the plots shown in the course book at the pages 13, 36, 43, 44, 47. Remark the symmetry of the graphs due to the property of the binomial distribution which is proven above.

2 (March 14) Let $X \sim$ Binomial(n, p), where $n \geq 3$ and $p \in (0, 1)$. For ease of writing, we introduce the notation $q = 1 - p$. Additionally, we define $P_n(k) = \mathbb{P}(X = k)$ for $k \in \{0, \ldots, n\}$. Prove the following:

a If $np + p$ is not an integer, then $P_n(k)$ reaches its maximum for $k_{\max} = \lfloor np+p \rfloor$, where $\lfloor \cdot \rfloor$ is the greatest integer less than or equal to the real number in the argument. For example, $\lfloor 2.0 \rfloor = 2$, $\lfloor 2.5 \rfloor = 2$, $\lfloor 2.999999 \rfloor = 2$.

b If $np + p$ is an integer, then $P_n(k)$ reaches its maximum for two consecutive values of k: $k_{\max,1} = np + p$ and $k_{\max,2} = np - q$.

Answer: The key idea is to investigate the behaviour of the ratio $P_n(k)/P_n(k - 1)$, where $k \in \{1, \ldots, n\}$. We have:

$$\begin{aligned}
\frac{P_n(k)}{P_n(k-1)} &= \frac{\binom{n}{k}p^k q^{n-k}}{\binom{n}{k-1}p^{k-1}q^{n-(k-1)}} \\
&= \frac{\frac{n!}{k!(n-k)!}}{\frac{n!}{(k-1)!(n-k+1)!}}\frac{p}{q} \\
&= \frac{n-k+1}{k}\frac{p}{q}.
\end{aligned}$$

Therefore, we get:

$$\begin{aligned}
\frac{P_n(k)}{P_n(k-1)} &> 1 \\
&\Updownarrow \\
\frac{n-k+1}{k}\frac{p}{q} &> 1 \qquad (1)\\
&\Updownarrow \\
\frac{n-k+1}{k} &> \frac{q}{p} \\
&\Updownarrow \\
pn - pk + p &> kq \\
&\Updownarrow \\
pn - k(p+q) + p &> 0 \\
&\Updownarrow \\
pn - k + p &> 0 \\
&\Updownarrow \\
k &< pn + p. \qquad (2)
\end{aligned}$$

a When $pn+p$ is not an integer, we have from (2) that $P_n(0) < P_n(1)$, $P_n(1) < P_n(2)$, $\ldots$, $P_n(\lfloor pn+p \rfloor - 1) < P_n(\lfloor pn+p \rfloor)$. From the calculations above, we also have that $P_n(\lfloor pn+p \rfloor) > P_n(\lfloor pn+p \rfloor + 1)$, $P_n(\lfloor pn+p \rfloor + 1) > P_n(\lfloor pn+p \rfloor + 2)$, $\cdots$, $P_n(n-1) > P_n(n)$. It follows that $P_n(k)$ reaches its maximum for $k_{\max} = \lfloor np+p \rfloor$.

b When $pn+p$ is an integer, we have from (2) that $P_n(0) < P_n(1)$, $P_n(1) < P_n(2)$, $\ldots$, $P_n(k_{\max,1}-2) < P_n(k_{\max,1}-1)$ and $P_n(k_{\max,1}) > P_n(k_{\max,1}+1)$, $P_n(k_{\max,1}+1) > P_n(k_{\max,1}+2)$, $\ldots$, $P_n(n-1) > P_n(n)$. All that remains is to compare $P_n(k_{\max,1}-1)$ and $P_n(k_{\max,1})$. Remark that $k_{\max,1} - 1 = pn+p-1 = pn-(1-p) = pn-q = k_{\max,2}$. Hence, we have:

$$\begin{aligned}
\frac{P_n(k_{\max,1})}{P_n(k_{\max,1}-1)} &= \frac{P_n(k_{\max,1})}{P_n(k_{\max,2})} \\
&= \frac{n-k_{\max,1}+1}{k_{\max,1}}\frac{p}{q} \qquad \text{[see (1)]} \\
&= \frac{n-np-p+1}{np+p}\frac{p}{q} \\
&= \frac{(n+1)(1-p)}{(n+1)p}\frac{p}{q} \\
&= 1.
\end{aligned}$$

It follows that $P_n(k_{\max,1}) = P_n(k_{\max,2})$ and this concludes the proof.

Comment: The most important conclusion of these calculations is that the value of k which maximizes $P_n(k)$ is always close to pn.

3 (March 19) According to the definition from the course book (see page 22), the events $A_1, A_2, \ldots, A_n$ $(n > 2)$ are mutually independent if $\mathbb{P}(A_1 \cap A_2 \cap \ldots \cap A_n) = \mathbb{P}(A_1)\mathbb{P}(A_2)\ldots\mathbb{P}(A_n)$ and the same multiplication rule holds for every subcollection of the events too. For example, if we want to verify if the events A_1, A_2, A_3 are mutually independent, then we should check that the following identities hold true:

$$\begin{aligned}
\mathbb{P}(A_1 \cap A_2 \cap A_3) &= \mathbb{P}(A_1)\mathbb{P}(A_2)\mathbb{P}(A_3),\\
\mathbb{P}(A_1 \cap A_2) &= \mathbb{P}(A_1)\mathbb{P}(A_2),\\
\mathbb{P}(A_1 \cap A_3) &= \mathbb{P}(A_1)\mathbb{P}(A_3),\\
\mathbb{P}(A_2 \cap A_3) &= \mathbb{P}(A_2)\mathbb{P}(A_3).
\end{aligned}$$

Therefore, in the case when $n = 3$, the number of equations which should be written down is four. Show that, in the general case, $2^n - (n+1)$ equations are needed to establish the mutual independence of n events.

Answer: We need to write down $\binom{n}{2}$ equations for checking the multiplication rule for pairs of events A_i, A_j, where $1 \le i < j \le n$. Similarly, for each k with property that $2 < k \le n$, we should write down $\binom{n}{k}$ equations for checking the multiplication rule for $A_{i_1}, A_{i_2}, \ldots, A_{i_k}$, where $1 \le i_1 < i_2 < \cdots < i_k \le n$. It follows that the total number of equations is

$$\begin{aligned}
\sum_{k=2}^{n} \binom{n}{k} &= \sum_{k=0}^{n} \binom{n}{k} - \sum_{k=0}^{1} \binom{n}{k}\\
&= 2^n - \binom{n}{0} - \binom{n}{1}\\
&= 2^n - 1 - n\\
&= 2^n - (n+1).
\end{aligned}$$

4 (March 21) Let $n > 2$. Suppose that $A_1, A_2, \ldots, A_n$ are events such that $\mathbb{P}(A_1 \cap A_2 \cap \cdots \cap A_{n-1}) > 0$. Then

$$\begin{aligned}
&\mathbb{P}(A_1 \cap A_2 \cap \cdots \cap A_n)\\
&\quad = \mathbb{P}(A_1)\mathbb{P}(A_2|A_1)\mathbb{P}(A_3|A_1 \cap A_2) \cdots \mathbb{P}(A_n|A_1 \cap A_2 \cdots \cap A_{n-1}).
\end{aligned}$$

Answer: It follows from the formula for conditional probabilities that the product of the probabilities on the right side of the equation above is equal to

$$\mathbb{P}(A_1)\frac{\mathbb{P}(A_1 \cap A_2)}{\mathbb{P}(A_1)}\frac{\mathbb{P}(A_1 \cap A_2 \cap A_3)}{\mathbb{P}(A_1 \cap A_2)}\cdots\frac{\mathbb{P}(A_1 \cap A_2 \cap \cdots \cap A_n)}{\mathbb{P}(A_1 \cap A_2 \cap \cdots \cap A_{n-1})}.$$

Since $\mathbb{P}(A_1 \cap A_2 \cap \cdots \cap A_{n-1}) > 0$, each of the denominators in this product are positive. All of the quantities involved in the product cancel each other except the final numerator $\mathbb{P}(A_1 \cap A_2 \cap \cdots \cap A_n)$ which is the left side of our equation.

Comment: The formula for the chains of events can be also found in the course book (see page 32), where is given for the particular case when $n = 3$.