Probability and Statistics

Solution sheet 1

Solution 1.1

(a) $\Omega = \{1,2,3,4,5,6\}^2$. The first coordinate will represent the green dice and the second one the red dice.

(b)
- $W_1 := \{(x,y) \in \Omega : x \leq 2, y \leq 2\} = \{1,2\}^2$.
- $W_2 := \{(x,y) \in \Omega : x = y\} = \{(1,1),(2,2),(3,3),(4,4),(5,5),(6,6)\}$.
- $W_3 := \{(x,y) \in \Omega : x = 3y\} = \{(3,1),(6,2)\}$.
- $W_4 := \{(x,y) \in \Omega : x+1 = y\} = \{(1,2),(2,3),(3,4),(4,5),(5,6)\}$.
- $W_5 := \{(x,y) \in \Omega : x \geq y\}$.

(c) We can define the equivalence relation $\sim$ as

$$(x,y) \sim (z,w) \Leftrightarrow \{x,y\} = \{z,w\},$$

then $\tilde{\Omega} = \Omega/\sim = \{\{x,y\} : x,y \in \Omega\}$. The W_i that can still be defined on $\tilde{\Omega}$ are the sets W such that if $(x,y) \in W$ then $(y,x) \in W$. This is the case for W_1, and W_2 only.

Solution 1.2

(a)
1. $A = \bigcap_{j=1}^n A_j^c$.
2. $B = \bigcup_{j=1}^n B_j$.
3. $C = \displaystyle\bigcup_{\substack{C,D \subseteq \{1,\dots,n\} \\ |C|=|D|}} \left(\bigcap_{j\in C} A_j \cap \bigcap_{j \in D^c} B_j^c \right)$.

(b)
1. $\left(\displaystyle\bigcup_{j=1}^n (A_j)^c \right)^c = \displaystyle\bigcap_{j=1}^n A_j$: All the numbers drawn are larger than $2k$.
2. $\displaystyle\bigcup_{j=1}^{N-2} (A_j \cap A_{j+1} \cap B_{j+2})$: There exists one point in time, where in two consecutive draws we got a number larger than $2k$ and in the following draw we got an even number.
3. $\bigcup_n^N \bigcap_{j=n}^N A_j \cap B_j$: There is a point in time after which, we only draw numbers that are even *and* bigger than $2k$.

Solution 1.3

(a) It's clear that

$$\begin{aligned}
\mathbb{1}_{\bigcup_{j=1}^n A_j} &= 1 - \mathbb{1}_{\bigcap_{j=1}^n A_j^c} \\
&= 1 - \prod_{j=1}^n \mathbb{1}_{A_j^c} \\
&= 1 - \prod_{j=1}^n \left(1 - \mathbb{1}_{A_j}\right),
\end{aligned}$$

Then we can just use the product formula to obtain

$$\prod_{j=1}^{n}\left(1-\mathbb{1}_{A_j}\right)=1+\sum_{k=1}^{n}\sum_{1\leq i_1<\ldots<i_k\leq n}\prod_{j=1}^{k}-\mathbb{1}_{A_{i_j}}.$$

Taking expectations on both sides

$$\begin{aligned}\mathbb{P}\left[\bigcup_{j=1}^{n}A_j\right]&=\mathbb{E}\left[\mathbb{1}_{\bigcup_{j=1}^{n}A_j}\right]\\&=-\mathbb{E}\left[\sum_{k=1}^{n}\sum_{1\leq i_1<\ldots<i_k\leq n}\prod_{j=1}^{k}-\mathbb{1}_{A_{i_j}}\right]\\&=\sum_{k=1}^{n}(-1)^{k+1}\sum_{1\leq i_1<\ldots<i_k\leq n}\mathbb{P}\left[A_{i_1}\cap\ldots\cap A_{i_k}\right].\end{aligned}$$

(b) We want to use induction. For $n=1$ both formulas are true. For the inductive step suppose that the formulas are true for an integer n and we will prove that the formulas hold for $n+1$. We first compute

$$\begin{aligned}\mathbb{P}\left[\bigcup_{j=1}^{n+1}A_j\right]&=\mathbb{P}\left[\bigcup_{j=1}^{n}A_j\cup A_{n+1}\right]\\&=\mathbb{P}\left[\bigcup_{j=1}^{n}A_j\right]+\mathbb{P}\left[A_{n+1}\right]-\mathbb{P}\left[\bigcup_{j=1}^{n}A_j\cap A_{n+1}\right]\\&\leq\mathbb{P}\left[\bigcup_{j=1}^{n}A_j\right]+\mathbb{P}\left[A_{n+1}\right]-\mathbb{P}\left[A_n\cap A_{n+1}\right]\\&\leq\sum_{j=1}^{n}\mathbb{P}[A_j]-\sum_{j=1}^{n-1}\mathbb{P}\left[A_j\cap A_{j+1}\right]+\mathbb{P}\left[A_{n+1}\right]-\mathbb{P}\left[A_n\cap A_{n+1}\right]\\&=\sum_{j=1}^{n+1}\mathbb{P}[A_j]-\sum_{j=1}^{n}\mathbb{P}\left[A_j\cap A_{j+1}\right],\end{aligned}$$

where we used the additivity property of a probability measure for the first inequality and the induction hypothesis for the second. For the second formula we have

$$\begin{aligned}
\mathbb{P}\left[\bigcup_{j=1}^{n+1} A_j\right] &= \mathbb{P}\left[\bigcup_{j=1}^{n} A_j \cup A_{n+1}\right] \\
&= \mathbb{P}\left[\bigcup_{j=1}^{n} A_j\right] + \mathbb{P}\left[A_{n+1}\right] - \mathbb{P}\left[\bigcup_{j=1}^{n} (A_j \cap A_{n+1})\right] \\
&\geq \mathbb{P}\left[\bigcup_{j=1}^{n} A_j\right] + \mathbb{P}\left[A_{n+1}\right] - \sum_{j=1}^{n} P(A_j \cap A_{n+1}) \\
&\geq \sum_{j=1}^{n} \mathbb{P}[A_j] - \sum_{i,j=1,i<j}^{n} \mathbb{P}\left[A_j \cap A_i\right] + \mathbb{P}\left[A_{n+1}\right] - \sum_{j=1}^{n} \mathbb{P}(A_j \cap A_{n+1}) \\
&= \sum_{j=1}^{n+1} \mathbb{P}[A_j] - \sum_{i,j=1,i<j}^{n+1} \mathbb{P}\left[A_j \cap A_i\right],
\end{aligned}$$

where we used again the additivity property of a probability measure for the first inequality and the induction hypothesis for the second.

Solution 1.4 We prove the statement by induction. For $n = 2$, we use the definition of conditional probabilities: if $\mathbb{P}(B_1) > 0$, then

$$\mathbb{P}(B_2|B_1) = \mathbb{P}(B_2 \cap B_1)/\mathbb{P}(B_1),$$

therefore $\mathbb{P}(B_2 \cap B_1) = \mathbb{P}(B_1)\mathbb{P}(B_2|B_1)$.

Assume the statement true for $n - 1$ and apply the previous step to $B_2 = A_n$ and $B_1 = A_1 \cap A_2 \cap \cdots \cap A_{n-1}$, we get

$$\mathbb{P}(A_1 \cap \cdots \cap A_n) = \mathbb{P}(A_1 \cap \cdots \cap A_{n-1})\mathbb{P}(A_n|A_1 \cap \cdots \cap A_{n-1}).$$

Since $A_1 \cap \cdots \cap A_{n-2}$ contains $A_1 \cap \cdots \cap A_{n-1}$, the former event has positive probability, we can inductively deduce the multiplication rule as announced.