1. We have to verify (A.1)–(A.3). As $\Omega^c = \emptyset$ is at most countable, (A.1) is satisfied. Now if $A \in \mathcal{F}$ then either $A$ or $A^c$ is countable. This is clearly equivalent to either $A^c$ or $(A^c)^c \equiv A$ being countable, so that $A^c \in \mathcal{F}$, (A.2) holds.

Finally, let $A_1, A_2, \ldots \in \mathcal{F}$. First suppose that all the $A_j$’s are countable. Then $\bigcup_{j \geq 1} A_j$ is also countable (do you remember how we showed that $\mathbb{N}^2$ is?), and so the union is in $\mathcal{F}$. If not all of the $A_j$’s are countable then at least one of them (say, $A_m$) is uncountable, and since $A_m \in \mathcal{F}$, $A_m^c$ must be at most countable. But then $(\bigcup_{j \geq 1} A_j)^c = \bigcap_{j \geq 1} A_j^c \subset A_m^c$, meaning that the set on the LHS is also at most countable, i.e. $\bigcup_{j \geq 1} A_j \in \mathcal{F}$, thus proving (A.3).

So $\mathcal{F}$ is a $\sigma$-algebra.

2. (i) As $B \subset A \cup B$, one has $P(B) \leq P(A \cup B) \leq P(A) + P(B)$ from monotonicity (the 1st inequality) and subadditivity (the 2nd inequality) of probability. The RHS here $= P(B)$ as $P(A) = 0$, so we are done.

(ii) Look: $P(A^c \cap B) = 1 - P(A^c \cap B)^c = 1 - P(A \cup B^c) = 1 - P(B^c)$ from part (i) since $P(A^c) = 0$. Finally, $1 - P(B^c) = P(B)$. QED.

3. Only one of $A$ and $B$ occurs iff the symmetric difference $A \Delta B = (A \cap B^c) \cup (B \cap A^c)$ occurs. As the events on the RHS are disjoint and $P(A) = P(A \cap B^c) + P(A \cap B)$ (the events appearing on the RHS forming a partition of $A$), one has

$$P(A \Delta B) = P(A \cap B^c) + P(B \cap A^c) = P(A) + P(B) - 2P(A \cap B)$$

(which is the same as $P(A \cup B) - P(A \cap B)$, of course).

4. Using de Morgan’s law (for the 2nd equality),

$$P \left( \bigcap_{n \geq 1} A_n \right) = 1 - P \left( \left( \bigcap_{n \geq 1} A_n \right)^c \right) = 1 - P \left( \bigcup_{n \geq 1} A_n^c \right) = 1,$$

in view of Boole’s inequality since all $P(A_n^c) = 1 - P(A_n) = 0$.

5. Can prove that claim using mathematical induction.

Initial step: take $n = 2$ (we could start with $n = 1$, but in that case the set of all $i, j$ s.t. $1 \leq i < j \leq n$ would be empty, which might confuse some people; so it’s kind of nicer to start with $n = 2$). Then the desired inequality becomes $P(A_1 \cup A_2) \geq P(A_1) + P(A_2) - P(A_1 \cap A_2)$. Hey, it does hold — with equality instead of $\geq$!
Induction step: \( n \mapsto n + 1 \) (we suppose that the inequality holds for that value of \( n \) and then show that that implies that it holds for \( n + 1 \) as well). Look:

\[
\mathbb{P} \left( \bigcup_{k=1}^{n+1} A_k \right) = \mathbb{P} \left( \bigcup_{k=1}^{n} A_k \right) + \mathbb{P} (A_{n+1}) - \mathbb{P} (B \cap A_{n+1}) \\
\geq \sum_{k=1}^{n} \mathbb{P} (A_k) - \sum_{1 \leq i < j \leq n} \mathbb{P} (A_i \cap A_j) + \mathbb{P} (A_{n+1}) - \mathbb{P} (B \cap A_{n+1}),
\]

where the last inequality follows from the induction assumption. It remains to note that

\[
\mathbb{P} (B \cap A_{n+1}) = \mathbb{P} \left( \bigcup_{k=1}^{n} (A_k \cap A_{n+1}) \right) \leq \sum_{k=1}^{n} \mathbb{P} (A_k \cap A_{n+1})
\]

by subadditivity, and that

\[
\sum_{1 \leq i < j \leq n} \mathbb{P} (A_i \cap A_j) + \sum_{k=1}^{n} \mathbb{P} (A_k \cap A_{n+1}) = \sum_{1 \leq i < j \leq n+1} \mathbb{P} (A_i \cap A_j),
\]

right? Right. QED.

6. Just verify probability axioms (P.1)–(P.3):

(P.1) is obvious: \( \forall A \in \mathcal{F}, \mathbb{P}(A) = pP_1(A) + (1-p)P_2(A) \geq 0 \) as both \( P_i(A) \geq 0 \), \( i = 1, 2 \).

(P.2) is obvious as well: \( \mathbb{P}(\Omega) = pP_1(\Omega) + (1-p)P_2(\Omega) = p + (1-p) = 1 \).

(P.3) For disjoint \( A_j, i \geq 1 \),

\[
P \left( \bigcup_{j \geq 1} A_j \right) = pP_1 \left( \bigcup_{j \geq 1} A_j \right) + (1-p)P_2 \left( \bigcup_{j \geq 1} A_j \right) \\
= p \sum_{j \geq 1} P_1(A_j) + (1-p) \sum_{j \geq 1} P_2(A_j) \quad \text{[from (P.3) for } p_j, j = 1, 2] \\
= \sum_{j \geq 1} (pP_1(A_j) + (1-p)P_2(A_j)) = \sum_{j \geq 1} P(A_j), \quad \text{bingo},
\]

where \( \sum \) holds since each of the series \( \sum_{j \geq 1} P_i(A_j), i = 1, 2 \), converges\(^1\).

Now clearly (P.1) and (P.2) hold for \( P_1P_2 \) as well, but (P.3) fails in the general case. For instance, if \( P_i = \varepsilon_i, i = 1, 2 \), and \( A_1 = \{1\}, A_2 = \{2\} \), then \( A_1 \cup A_2 = \{1, 2\} \)

\[
1 = P_1(\{1, 2\})P_2(\{1, 2\}) = (P_1P_2)(A_1 \cup A_2) \\
\neq (P_1P_2)(A_1) + (P_1P_2)(A_2) = P_1(\{1\})P_2(\{1\}) + P_2(\{2\})P_2(\{2\}) = 0.
\]

\(^1\)Can you give a counterexample to that in the case of general series, i.e., show that, in the general case, one cannot claim that \( \sum_{j \geq 1} (x_j + y_j) = \sum_{j \geq 1} x_j + \sum_{j \geq 1} y_j \)?
7. (a) First plot the density \( f \): DIY (\( f \) is a continuous piece-wise linear “triangular” function, vanishing outside \([0, 2]\), growing linearly from 0 to 1 on \([0, 1]\) and then decaying to 0 on \([1, 2]\)). The DF of \( P_a \) is given by the integral

\[
F_a(t) = \int_{-\infty}^{t} f(x)dx = \begin{cases} 
0, & t \leq 0, \\
\int_{0}^{t} x \, dx = \frac{t^2}{2}, & t \in (0, 1], \\
F_a(1) + \int_{1}^{t} (2 - x) \, dx = 1 - \frac{1}{2}(t - 2)^2, & t \in (1, 2], \\
1, & t \geq 2.
\end{cases}
\]

Plot this (quadratic spline) function yourself, please!

(b) \( F(t) = P((\infty, t]) = \frac{1}{2} P_d((\infty, t]) + \frac{1}{2} P_a((\infty, t]) = \frac{1}{2} F_d(t) + \frac{1}{2} F_a(t) \), where

\[
F_d(t) = \begin{cases} 
0, & t < 0, \\
0.04, & t \in [0, 1), \\
0.04 + 0.32 = 0.36, & t \in [1, 2), \\
0.36 + 0.64 = 1, & t \geq 2.
\end{cases}
\]

So,

\[
F(t) = \begin{cases} 
0, & t < 0, \\
0.02 + \frac{t^2}{4}, & t \in [0, 1), \\
0.18 + \frac{1}{2}(1 - \frac{1}{2}(t - 2)^2), & t \in [1, 2), \\
1, & t \geq 2.
\end{cases}
\]

To plot \( F \), take the plot of \( F_a \), “squash” it by taking the plot of \( \frac{1}{2} F_a \), and then “insert” jumps of sizes \( \frac{1}{2} \times 0.04 = 0.02, \frac{1}{2} \times 0.32 = 0.16 \) and \( \frac{1}{2} \times 0.64 = 0.32 \) at the points 0, 1 and 2, respectively. DIY!

(c) \( P(A) = \frac{1}{2} P_d(A) + \frac{1}{2} P_d(A) = \frac{1}{2} \times 0.04 + \frac{1}{2} \times 0.5 = 0.27 \),

\[
P(B) = \frac{1}{2} P_d(B) + \frac{1}{2} P_d(B) = \frac{1}{2} P_d(\{1\}) + \frac{1}{2} [F_a(1.5) - F_a(0.5)]
\]

\[
= 0.16 + \frac{1}{2} [(1 - \frac{1}{8}) - \frac{1}{8}] = 0.16 + \frac{3}{8} = 0.535,
\]

\[
P(C) = [F(1) - F(0.5)] + [F(10) - F(1.5)]
\]

\[
= [0.43 - (0.02 + \frac{1}{2} \times \frac{1}{8})] + [1 - F(1.5)]
\]

\[
= 0.41 - \frac{1}{16} + [1 - (0.18 + \frac{1}{2} \times (1 - \frac{1}{8}))] = 0.41 + 0.5 - 0.18 = 0.73.
\]