PROBABILITY BY ALAN F. KARR

Solutions

Leo Wang

Contents

1 Probability

Exercise 1.1. Prove *de Morgan's law*:

$$(A \cup B)^c = A^c \cap B^c$$

and

$$(A \cap B)^c = A^c \cup B^c$$

both directly and using indicator functions.

Solution. Let's prove $(A \cup B)^c = A^c \cap B^c$, then $(A \cap B)^c = A^c \cup B^c$ can be proved in a similar way.

Directly
($\subseteq$) For an element $x \in (A \cup B)^c$, then $x \notin A \cup B = \{y | y \in A \vee y \in B\}$. In this case, $x \notin A$ and $x \notin B$, that is $x \in A^c$ and $x \in B^c$. Hence, $x \in A^c \cap B^c$. Given that this holds for all $x \in (A \cup B)^c$, we have $(A \cup B)^c \subseteq A^c \cap B^c$.
($\supseteq$) For an element $x \in A^c \cap B^c$, it means $x \in A^c$ and $x \in B^c$. In this case, $x \notin A$ and $x \notin B$, hence $x \notin A \cup B$ by definition. Therefore, $x \in (A \cup B)^c$. Given that this holds for all $x \in A^c \cap B^c$, we have $(A \cup B)^c \supseteq A^c \cap B^c$.

Indicator functions

$$\begin{aligned}
\mathbb{1}_{(A\cup B)^c} &= 1 - \mathbb{1}_{A\cup B} \\
&= 1 - \max\{\mathbb{1}_A, \mathbb{1}_B\} \\
&= \min\{1 - \mathbb{1}_A, 1 - \mathbb{1}_B\} \\
&= \min\{\mathbb{1}_{A^c}, \mathbb{1}_{B^c}\} \\
&= \mathbb{1}_{A^c \cap B^c}
\end{aligned}$$

□

Exercise 1.2. Prove that for events A and B, $A\Delta B = A^c \Delta B^c$.

Solution. By definition

$$\begin{aligned}
A\Delta B &= A \setminus B + B \setminus A \\
&= A \cap B^c + B \cap A^c \\
&= A^c \cap (B^c)^c + B^c \cap (A^c)^c \\
&= A^c \setminus B^c + B^c \setminus A^c \\
&= A^c \Delta B^c
\end{aligned}$$

□

Exercise 1.3. Let B and C be events (A_n) and let $A_n = B$ if n is odd and $A_n = C$ if n is even. Calculate $\limsup_n A_n$ and $\liminf_n A_n$.

Solution. Intuitively while n goes to ∞, $\limsup_n A_n$ is the set where A_n will *"never leave forever"*, and $\liminf_n A_n$ is the set where A_n will *"eventually stay forever"*. Since A_n is alternating between B and C, $\limsup_n A_n = B \cup C$ and $\liminf_n A_n = B \cap C$.

Now, let's prove it. By definition

$$\begin{aligned}
\limsup_n A_n &= \bigcap_{k=1}^{\infty} \bigcup_{n=k}^{\infty} A_n \\
&= \bigcap_{k=1}^{\infty} \Big(\bigcup_{n=2k-1}^{\infty} A_n \cup \bigcup_{n=2k}^{\infty} A_n \Big) \\
&= \bigcap_{k=1}^{\infty} (B \cup C) \\
&= B \cup C
\end{aligned}$$

And

$$\begin{aligned}
\liminf_n A_n &= \bigcup_{k=1}^{\infty} \bigcap_{n=k}^{\infty} A_n \\
&= \bigcup_{k=1}^{\infty} \Big(\bigcap_{n=2k-1}^{\infty} A_n \cap \bigcap_{n=2k}^{\infty} A_n \Big) \\
&= \bigcup_{k=1}^{\infty} (B \cap C) \\
&= B \cap C
\end{aligned}$$

□

Exercise 1.4. Prove part b) of proposition 1.9: Let $A_1, A_2, \ldots$ be subsets of Ω. If $A_1 \supseteq A_2 \supseteq \cdots$, then $A_n \to A = \cap_{n=1}^{\infty} A_n$. (This is written as $A_n \downarrow A$)

Solution. Let $A = \cap_{n=1}^{\infty} A_n$. To prove $A_n \to A$, we can show 1) $\liminf_n A_n = A$, and 2) $\limsup_n A_n = A$.

1) Given $A_1 \supseteq A_2 \supseteq \cdots$, for each k, $\cap_{n=k}^{\infty} A_n = A$. Hence by definition

$$\liminf_n A_n = \bigcup_{k=1}^{\infty} \bigcap_{n=k}^{\infty} A_n = \bigcup_{k=1}^{\infty} A = A.$$

2) On the other hand, for each k, $\cup_{n=k}^{\infty} A_n = A_k$, by definition

$$\limsup_n A_n = \bigcap_{k=1}^{\infty} \bigcup_{n=k}^{\infty} A_n = \bigcap_{k=1}^{\infty} A_k = A.$$

□

Exercise 1.5. Prove that
a) if $t_n \downarrow t$, then $(-\infty, t_n] \downarrow (-\infty, t]$, but that
b) if $t_n \uparrow t$ and $t_n < t$ for each n, then $(-\infty, t_n] \uparrow (-\infty, t)$.

Solution. Let's prove a), then b) can be proved in a similar way. By Proposition 1.9,

$$(-\infty, t_n] \downarrow \bigcap_{n=1}^{\infty} (-\infty, t_n].$$

It is left to show that $\cap_{n=1}^{\infty}(-\infty, t_n] = (-\infty, t]$.

Let $x \in \cap_{n=1}^{\infty}(-\infty, t_n]$, then $x \leq t_n$ for all n. As $n \to \infty$, $t_n \to t$ and $x \leq t$, so $x \in (-\infty, t]$. On the other hand, if $x \in (-\infty, t]$, then $x \leq t \leq t_n$ for all n. Therefore, we have $\cap_{n=1}^{\infty}(-\infty, t_n] = (-\infty, t]$. □

Exercise 1.6. Show that the disjointification procedure used to prove Proposition 1.24 actually yields $\cup_{n=1}^{k} A_n = \Sigma_{n=1}^{k} B_n$ for every k.

Solution. Let $k = 1$, then $A_1 = B_1$. Assuming $\cup_{n=1}^{k} A_n = \Sigma_{n=1}^{k} B_n$ holds for some $k > 1$, then

$$\begin{aligned}
\cup_{n=1}^{k+1} A_n &= A_{k+1} \cup (\cup_{n=1}^{k} A_n) \\
&= (A_{k+1} \setminus \cup_{n=1}^{k} A_n) + \cup_{n=1}^{k} A_n \\
&= B_{k+1} + \Sigma_{n=1}^{k} B_n \\
&= \Sigma_{n=1}^{k+1} B_n
\end{aligned}$$

Therefore, the statement is proved by induction. □

Exercise 1.7. Prove (1.1) through (1.4):
(1.1) $\mathbb{1}_{A \cup B} = \max\{\mathbb{1}_A, \mathbb{1}_B\}$
(1.2) $\mathbb{1}_{A \cap B} = \min\{\mathbb{1}_A, \mathbb{1}_B\} = \mathbb{1}_A \mathbb{1}_B$
(1.3) $\mathbb{1}_{A^c} = 1 - \mathbb{1}_A$
(1.4) $\mathbb{1}_{A \Delta B} = |\mathbb{1}_A - \mathbb{1}_B|$

Solution. Note that both sides of the above equations are either or 0 or 1. If one side equals to 1, then it must be greater and equal to another side. Therefore, we just need to focus on the case when one side equals to 0. Let $\omega \in \Omega$.

(1.1) ($\geq$) If $\mathbb{1}_{A \cup B}(\omega) = 0$, by definition $\omega \notin (A \cup B)$. This means that $\omega \notin A$ and $\omega \notin B$, that is $\mathbb{1}_A(\omega) = \mathbb{1}_B(\omega) = 0$.
($\leq$) If $\max\{\mathbb{1}_A, \mathbb{1}_B\}(\omega) = 0$, it must be true that $\mathbb{1}_A(\omega) = \mathbb{1}_B(\omega) = 0$. By definition, $\omega \notin A$ and $\omega \notin B$. Hence $\omega \notin A \cup B$ and $\mathbb{1}_{A \cup B}(\omega) = 0$.

(1.2) ($\geq$) If $\mathbb{1}_{A \cap B}(\omega) = 0$, then $\omega \in A$ and $\omega \in B$ cannot happen simultaneously and $\min\{\mathbb{1}_A, \mathbb{1}_B\}(\omega) = 0$.
($\leq$) If $\min\{\mathbb{1}_A, \mathbb{1}_B\}(\omega) = 0$, then it is either $\mathbb{1}_A(\omega) = 0$ or $\mathbb{1}_B(\omega) = 0$. This means that either $\omega \notin A$ or $\omega \notin B$. In either of these cases, $\mathbb{1}_{A \cap B}(\omega) = 0$.

(1.3) ($\geq$) If $\mathbb{1}_{A^c}(\omega) = 0$, then $\omega \notin A^c$ but $\omega \in A$, and $1 - \mathbb{1}_A(\omega) = 0$.
($\leq$) If $1 - \mathbb{1}_A(\omega) = 0$, then $\mathbb{1}_A(\omega) = 1$ and $\omega \in A$. Thus $\omega \notin A^c$ and $\mathbb{1}_{A^c}(\omega) = 0$.

(1.4) ($\geq$) If $\mathbb{1}_{A\Delta B}(\omega) = 0$, then it is either $\omega \in (A \cup B)^c$ or $\omega \in (A \cap B)$. The former case gives $\mathbb{1}_A(\omega) = \mathbb{1}_B(\omega) = 0$ and the latter case gives $\mathbb{1}_A(\omega) = \mathbb{1}_B(\omega) = 0$. Both cases yield $|\mathbb{1}_A - \mathbb{1}_B|(\omega) = 0$.
($\leq$) If $|\mathbb{1}_A - \mathbb{1}_B|(\omega) = 0$, similar to the previous case, it is either $\omega \in A$ and $\omega \in B$ occur simultaneously, or $\omega \notin A$ and $\omega \notin B$ occur simultaneously. In either case, $\mathbb{1}_{A\Delta B}(\omega) = 0$. $\square$

Exercise 1.8. Let $A_1, A_2, \ldots$ be events. Prove that $A_n \to A$ if and only if $\mathbb{1}_{A_n} \to \mathbb{1}_A$ as functions on Ω (that is, $\mathbb{1}_{A_n}(\omega) \to \mathbb{1}_A(\omega)$ for every ω).

Solution. Using (1.1) and (1.2) from Exercise 1.7, we have

$$\begin{aligned}
\mathbb{1}_{\liminf_n A_n} &= \mathbb{1}_{\cup_{k=1}\cap_{n=k}^{\infty} A_n} \\
&= \sup_{k\geq 1} \mathbb{1}_{\cap_{n=k}^{\infty} A_n} \quad \text{by (1.1)} \\
&= \sup_{k\geq 1} \inf_{n\geq k} \mathbb{1}_{A_n} \quad \text{by (1.2)} \\
&= \liminf_n \mathbb{1}_{A_n}
\end{aligned}$$

Similarly, we can prove $\mathbb{1}_{\limsup_n A_n} = \limsup_n \mathbb{1}_{A_n}$. Using these two results, the statement can be easily proved. $\square$

Exercise 1.9. Given subsets A and B of Ω, identify all sets in $\sigma(A, B)$.

Solution. Let $A_1 = (A \cup B)^c, A_2 = A \backslash B, A_3 = A \cap B, A_4 = B\ A$ be finite partitions of Ω. Then

$$\begin{aligned}
\sigma(A, B) = \{&\emptyset, A_1, A_2, A_3, A_4, \\
&A_1 \cup A_2, A_1 \cup A_3, A_1 \cup A_4, A_2 \cup A_3, A_2 \cup A_4, A_3 \cup A_4, \\
&A_1 \cup A_2 \cup A_3, A_1 \cup A_2 \cup A_4, A_2 \cup A_3 \cup A_4, \Omega\}
\end{aligned}$$

$\square$

Exercise 1.10. Prove that $\{x\}$ is a Borel set for every $x \in \mathbb{R}$.

Solution. By definition, a Borel set is an element of $\mathcal{B}(\mathbb{R})$, so the statement can be proved by representing $\{x\}$ by countable intersections and complement of intervals $(a, b]$, where $a < b$ in $\mathbb{R} \cup \{-\infty, +\infty\}$:

$$\{x\} = (-\infty, x] \backslash \cap_{n=1}^{\infty} (-\infty, x - \frac{1}{n}].$$

$\square$

Exercise 1.11. Let A, B and C be disjoint events with $P(A) = .6$, $P(B) = .3$ and $P(C) = .1$. Calculate the probabilities of all events in the σ-algebra generated by $\{A, B, C\}$.

Solution. Given A, B, C are disjoint events, $\sigma(\{A, B, C\}) = \{\emptyset, A, B, C, A \cup B, A \cup C, B \cup C, A \cup B \cup C\}$. Then

$$\begin{aligned}
P(\emptyset) &= 0 \\
P(A) &= .6 \\
P(B) &= .3 \\
P(C) &= .1 \\
P(A \cup B) &= P(A) + P(B) = .6 + .3 = .9 \\
P(A \cup C) &= P(A) + P(C) = .6 + .1 = .7 \\
P(B \cup C) &= P(B) + P(C) = .3 + .1 = .4 \\
P(A \cup B \cup C) &= P(A) + P(B) + P(C) = .9 + .6 + .3 = 1
\end{aligned}$$

□

Exercise 1.12. Verify that (1.8) defines a probability.

Solution. To verify ε_ω (the point mass at $\omega \in \Omega$) is a probability, we need to verify the three conditions in the Definition 1.19.
a) As ε_ω has values in $\{0, 1\}$, $\varepsilon_\omega(A) \geq 0$ for all $A \in \mathcal{F}$.
b) As $\omega \in \Omega$, $\varepsilon_\omega(\Omega) = 1$.
c) Let $A_1, A_2, \ldots$ be disjoint sets in $\mathcal{F}$. It is either $\omega \in \Sigma_{n=1}^{\infty} A_n$ or $\omega \notin \cup_{n=1}^{\infty} A_n$.
In the former case, $\varepsilon_\omega(\Sigma_{n=1}^{\infty} A_n) = 1$, and ω must exist in only A_k for some $k \geq 1$, so $\Sigma_{n=1}^{\infty} P(A_n) = P(A_k) + \Sigma_{n \neq k} P(A_n) = 1 + 0 = 1$. □

Exercise 1.13. Confirm that (1.9) defines a probability.

Solution. a) Since p_i are defined to be positive, $\Sigma_{i \in I} p_i \geq 0$ for all $I \subseteq \{1, \ldots, n\}$.
b) By definition of p_i, $\Sigma_{i=1}^{n} p_i = 1$.
c) The countable distjoint sets can only be finite in this problem. By definition of p_i, c) can be easily verified. □

Exercise 1.14. Prove that for all events A and B,

$$\begin{aligned}
P(A \Delta B) &= P(A \cup B) - P(A \cap B) \\
&= P(A) + P(B) - 2P(A \cap B).
\end{aligned}$$

Solution. For the first equality, since $A \cup B = A \Delta B + A \cup B$, the sum of two disjoint sets, we have $P(A \cup B) = P(A \Delta B) + P(A \cup B)$. Simply rearranging gives the first equality.
For the second equality, given $A = A \backslash B + A \cap B$ and $B = B \backslash A + A \cap B$, we can get $P(A \backslash B) = P(A) - P(A \cap B)$ and $P(B \backslash A) = P(B) - P(A \cap B)$. Since $A \Delta B = A \backslash B + B \backslash A$, we have $P(A \Delta B) = P(A \backslash B) + P(B \backslash A) = P(A) + P(B) - 2P(A \cap B)$. □

Exercise 1.15. Let Ω be a finite set and let $\mathcal{F} = \mathcal{P}(\otimes)$ be the σ-algebra of all subsets of Ω. For $A \subseteq \Omega$, let $|A|$ be the number of elements in A. Prove that the formula $P(A) = |A|/|\Omega|$ defines a probability on $(\Omega, \mathcal{F})$ (the uniform distribution on Ω).

Solution. Assume Ω is not empty.
a) Since $|A| \geq 0$ and $|\Omega| \geq 1$, $P(A) \geq 0$ for all $A \in \mathcal{F}$.
b) $P(\Omega) = |\Omega|/|\Omega| = 1$.
c) Let $A_1, A_2, \ldots, A_n$ be a finitne number of disjoint sets in $\mathcal{F}$. The number of elements in $\Sigma_{k=1}^{n} A_k$ must be equal to the total number of elements in A_k for all $k \in \{1, \ldots, n\}$, i.e. $|\Sigma_{k=1}^{n} A_k| = \Sigma_{k=1}^{n} |A_k|$. Thus, c) can be easily verified. □

Exercise 1.16. Prove that (1.15) and (1.16) are equivalent.

Solution. (1.15) $\Rightarrow$ (1.16):

$$\limsup_{n\to\infty} P(A_n) = 1 - \liminf_{n\to\infty} P(A_n^c) \leq 1 - P(\liminf_n A_n^c) = P(\limsup_n A_n)$$

(1.16) $\Rightarrow$ (1.15):

$$P(\liminf_n A_n) = 1 - P(\limsup_n A_n^c) \leq 1 - \limsup_{n\to\infty} P(A_n^c) = \liminf_{n\to\infty} P(A_n)$$

□

Exercise 1.17. Prove that there does not exist a uniform distribution on the set $\mathbb{N} = \{0, 1, \ldots\}$.

Solution. Assume there exists a uniform distribution on $\mathbb{N}$, which has the probability $P(A) = k|A|$ for some constant $k \geq 0$.
If $k = 0$, $P(\mathbb{N}) = 0$ contradicts the definition of probability.
If $k > 0$, $P(\mathbb{N}) = k|\mathbb{N}| = +\infty > 1$ also contradicts the definition of probability.
Therefore, by contradiction, there does not exist a uniform distribution on $\mathbb{N}$. □

Exercise 1.18. Suppose that P_1 and P_2 are probabilities on $(\Omega, \mathcal{F})$ and that $0 \leq \alpha \leq 1$. Prove that the set function

$$P(A) = \alpha P_1(A) + (1-\alpha) P_2(A)$$

is also a probability.

Solution. a) As $\alpha \geq 0$ and $1 - \alpha \geq 0$, $P(A) \geq 0$ for all $A \in \mathcal{F}$.
b) $P(\Omega) = \alpha P_1(\Omega) + (1-\alpha) P_2(\Omega) = \alpha + (1-\alpha) = 1$.
c) Let $A_1, A_2, \ldots$ are disjoint sets in $\mathcal{F}$,

$$\begin{aligned}
P(\Sigma_{n=1}^{\infty} A_n) &= \alpha P_1(\Sigma_{n=1}^{\infty} A_n) + (1-\alpha) P_2(\Sigma_{n=1}^{\infty} A_n) \\
&= \alpha \Sigma_{n=1}^{\infty} P_1(A_n) + (1-\alpha) \Sigma_{n=1}^{\infty} P_2(A_n) \\
&= \Sigma_{n=1}^{\infty} (\alpha P_1(A_n) + (1-\alpha) P_2(A_n)) \\
&= \Sigma_{n=1}^{\infty} P(A_n)
\end{aligned}$$

□

Exercise 1.19. Prove that if $P(A_i) = 1$ for each i, then

$$P(\cap_{i=1}^{\infty} A_i) = 1.$$

Solution. By Boole's inequility, $P(\cup_{i=1}^{\infty} A_i^c) \leq \Sigma_{i=1}^{\infty} P(A_i^c) = 0$. Since $P \geq 0$ by difinition, $P(\cup_{i=1}^{\infty} A_i^c) = 0$. Thus

$$P(\cap_{i=1}^{\infty} A_i) = 1 - P(\cup_{i=1}^{\infty} A_i^c) = 1.$$

□

Exercise 1.20. Show that for events A and B,

$$P(A \cap B) \geq P(A) + P(B) - 1.$$

Solution. By Boole's inequality,

$$P(A \cap B) = 1 - P(A^c \cup B^c) \geq 1 - P(A^c) - P(B^c) = P(A) + P(B) - 1.$$

□

Exercise 1.21. Let P be a probability on $\mathbb{R}$. Prove that for every $\varepsilon > 0$ there is a compact set K such that $P(K) > 1 - \varepsilon$.

Solution. Consider compact intervals $K_n = [-n, n]$, $n \in \mathbb{N}$. As $n \to \infty$, $K_n \uparrow \mathbb{R}$ and $P(K_n) \uparrow 1$. For an arbitrary $\varepsilon > 0$, there must exist a sufficiently large n such that $P(K_n) > 1 - \varepsilon$. □

Exercise 1.22. Prove that a distribution function on $\mathbb{R}$ has at most countably many points of discontinuity.

Solution. [TODO] □

Exercise 1.23. Prove that if $P(A) > 0$, then the set function

$$P_A(B) = P(B|A)$$

is a probability on $(\Omega, \mathcal{F})$ satisfying $P_A(A^c) = 0$.

Solution. By Bayes' theorem,

$$P_A(A^c) = \frac{P(A^c \cap A)}{P(A)} = \frac{P(\emptyset)}{P(A)} = 0.$$

□

Exercise 1.24. Let P be the uniform distribution on a fininte set Ω and let A be a subset Ω. Prove that $P(\cdot|A)$ is the uniform distribution on A.

Solution. For an arbitrary set $B \subseteq A$, by Bayes' theorem,

$$P(B|A) = \frac{P(B \cap A)}{P(A)} = \frac{P(B)}{P(A)} = \frac{|B|}{|A|},$$

which coincides the probability of a uniform distribution on A. □

Exercise 1.25. Let $A_1, \dots, A_n$ be events, and for $J \subseteq \{1, \dots, n\}$, let $B_J = \cap_{j \in J} A_j$. For $k \geq 1$, let $S_k = \Sigma_{|J|=k} P(B_J)$, where the sum is over all subsets J of $\{1, \dots, n\}$ with $|J| = k$.
a) Prove the *inclusion-exclusion principle*:

$$P(\bigcup_{i=1}^{n} A_i) = \Sigma_{k=1}^{n} (-1)^{k-1} S_k.$$

b) Suppose that $P(B_J)$ depends only on $|J|$, i.e., there are numbers $q_0, \dots, q_n$ such that $P(B_J) = q_k$ whenever $|J| = k$. (This is true, for example, in Exercise 1.15.) Prove that $S_k = \binom{n}{k} q_k$.

Solution. a) Prove the inclusion-exclusion principle by induction is such a pain. Here we show a simpler proof by borrowing Definition 4.1 of expectation.
Using indicator functions, we have

$$\mathbb{1}_{\cup_{i=1}^{n} A_i} = 1 - \mathbb{1}_{\cap_{i=1}^{n} A_i^c} = 1 - \prod_{i=1}^{n} (1 - \mathbb{1}_{A_i}).$$

Expanding this equation, we can get

$$\mathbb{1}_{\cup_{i=1}^{n} A_i} = \Sigma_{k=1}^{n} (-1)^{k-1} \Sigma_{|J|=k} \prod_{j \in J} \mathbb{1}_{A_j}.$$

Taking expectation from both sides tives the inclusion-exclusion principle.
b) As $\binom{n}{k}$ represents the number of combinations of choosing k from n:

$$S_k = \Sigma_{|J|=k} P(B_J) = \Sigma_{|J|=k} q_k = \binom{n}{k} q_k.$$

□

Exercise 1.26. This is known as the *coupon collector's problem.* There are t different types of coupons available and the collector is seeking to collect one of each (for example, in order to win some premium). Show that if n coupons have been collected, then the probability p_n of having at least one of each type is

$$p_n = \Sigma_{k=1}^{t} (-1)^{k-1} \binom{t}{k} \left(1 - \frac{k}{t}\right)^n.$$

Solution. Let A_i be the event that the i-th type coupon is missing when n coupons have been collected, and $\cup_{i=1}^{t} A_i$ is the event that at least one type is missing. Then the event of having at least one of each type is the complement of $\cup_{i=1}^{t} A_i$. By applying the inclusion-exclusion principle (from exercise 1.25),

$$P(\cup_{i=1}^{n} A_i) = \Sigma_{k=1}^{n} (-1)^{k-1} S_k$$

where $S_k = \Sigma_{|J|=k} P(B_J)$ and $B_J = \cap_{j \in J} A_j$. In this problem, B_J is the event that all the types in J are missing once n coupons have been collected, and

$$P(B_J) = \left(1 - \frac{k}{t}\right)^n.$$

Since $P(B_J)$ depends only on $|J| = k$, applying exercise 1.25 b), $S_k = \binom{n}{k} \left(1 - \frac{k}{t}\right)^n$ and

$$P(\cup_{i=1}^n A_i) = \Sigma_{k=1}^n (-1)^{k-1} \binom{n}{k} \left(1 - \frac{k}{t}\right)^n.$$

Therefore, the probability p_n of having at least one of each type is

$$p_n = 1 - P(\cup_{i=1}^n A_i) = 1 - \Sigma_{k=1}^n (-1)^{k-1} \binom{n}{k} \left(1 - \frac{k}{t}\right)^n.$$

(Note: It seems like there is an erratum in this exercise.) □

Exercise 1.27. Consider an urn that initially contains r red balls and b black balls. At each trial one ball is drawn. It is replaced and $c \geq 0$ balls of the same color added to the urn. Let A_j be the event that the j th ball drawn is black. Show that $P(A_j) = b/(b+r)$ for every j.

Solution. [TODO] □

2 Random Variables

Exercise 2.1. Let $\Omega = \{1, 2, 3, 4, 5, 6\}$, and let$\mathcal{F} = \sigma(\{1, 2, 3, 4\}, \{3, 4, 5, 6\})$.
a) List all sets in $\mathcal{F}$.
b) Is the function

$$X(\omega) = \begin{cases} 2 & \omega = 1, 2, 3, 4 \\ 7 & \omega = 5, 6 \end{cases}$$

a random variable over $(\Omega, \mathcal{F})$?
c) Given an example of a function on Ω that is *not* a random variable over $(\Omega, \mathcal{F})$.
d) Show that there exists a probability P on $(\Omega, \mathcal{F})$ such that $P(A)$ is zero or one for all $A \in \mathcal{F}$, yet P is not a point mass.

Solution. a)

$$\mathcal{F} = \{\emptyset, \{1, 2\}, \{3, 4\}, \{5, 6\}, \{1, 2, 3, 4\}, \{1, 2, 5, 6\}, \{3, 4, 5, 6\}, \Omega\}.$$

b)

$$\{X \leq t\} = \begin{cases} \emptyset & t < 2 \\ \{1, 2, 3, 4\} & 2 \leq t < 7 \\ \Omega & t \geq 7 \end{cases}$$

Given $\{X \leq t\}$ for every $t \in \mathbb{R}$, by Proposition 2.12, X is a random variable over $(\Omega, \mathcal{F})$.

c) Consider a function Y such that

$$Y(\omega) = \begin{cases} 2 & \omega = 1 \\ 7 & \omega = 2, 3, 4, 5, 6 \end{cases}$$

Since $\{X \leq 2\} = \{1\} \notin \mathcal{F}$, Y is not a random variable over $(\Omega, \mathcal{F})$.

d) Consider a probability meausre P such that $P(A) = \mathbb{1}(A \supseteq \{1,2\})$. It can be easily verified that P is a probability measure, yet P is not a point mass.

$\square$

Exercise 2.2. Prove that if X and Y are random variables, then $\{X \le Y\}$, $\{X < Y\}$ and $\{X = Y\}$ are events.

Solution. To prove $\{X \le Y\}$ is an event, we can show that $\{X \le Y\} \in \mathcal{F}$. By Proposition 2.13, $X - Y$ is a random variable. By Proposition 2.12, $\{X \le Y\} = \{X - Y \le 0\} \in \mathcal{F}$. Similarly, $\{X < Y\} \in \mathcal{F}$. Hence $\{X = Y\} = \{X \le Y\} \backslash \{X < Y\} \in \mathcal{F}$. $\square$

Exercise 2.3. Let X and Y be random variables and let A be an event. Prove that the function

$$Z(\omega) = \begin{cases} X(\omega) & \text{if } \omega \in A \\ Y(\omega) & \text{if } \omega \in A^c \end{cases}$$

is a random variable.

Solution. Recall that indicator functions of an event are random variables (Example 2.4). As $Z = \mathbb{1}_A X + (1 - \mathbb{1}_A)Y$, by Proposition 2.13, Z is a random variable. $\square$

Exercise 2.4. Let $\mathcal{G} = \{A_1, \ldots, A_n\}$ be a finite partition of Ω, and let $\mathcal{F} = \sigma(\mathcal{G})$.
a) Prove that a function $X : \Omega \to \mathbb{R}$ is a random variable if and only if X is constant over each partition set A_i.
b) Use part a) to show that provided $\mathcal{F} \ne \mathcal{P}(\Omega)$, there exist functions Y on Ω such that $|Y|$ is a random variable but Y is not.

Solution. [TODO] $\square$

Exercise 2.5. Let X^+ and X^- be the positive and negative parts of the function $X : \Omega \to \mathbb{R}$. Prove that $X = X^+ - X^-$ and $|X| = X^+ + X^-$.

Solution. For $\omega \in \Omega$ such that $X(\omega) \ge 0$, $X^+(\omega) = X$ and $X^-(\omega) = 0$. Then, $X = X - 0 = X^+ - X^-$ and $|X| = X + 0 = X^+ + X^-$. For $\omega \in \Omega$ such that $X(\omega) < 0$, $X^+(\omega) = 0$ and $X^-(\omega) = -X$. Then, $X = 0 - (-X) = X^+ - X^-$ and $|X| = 0 + (-X) = X^+ + X^-$. As the equations hold for all $\omega \in \Omega$, the proof is complete. $\square$

Exercise 2.6. Show that if X is discrete with values in the countable set C, then for every $B \in \mathcal{B}(\mathbb{R})$, $P\{X \in B\} = \Sigma_{a \in C \cap B} P\{X = a\}$.

Solution. We can write $B = (C + C^c) \cap B = C \cap B + C^c \cap B$. Since X is discrete with values only in C, $P\{X \in C^c \cap B\} = 0$ and $P\{X \in B\} = P\{X \in C \cap B\}$. As C is countable, then $C \cap B = \cup_{a \in C \cap B}\{a\}$ is also countable. Hence

$$P\{X \in B\} = P\{X \in \cup_{a \in C \cap B}\{a\}\} = \Sigma_{a \in C \cap B} P\{X = a\}$$

$\square$

Exercise 2.7. Consider a random permutation of the intergers $\{1, \ldots, n\}$, with all $n!$ permutations equally likely. For each i, let X_i be the integer in the ith position, and let $A_i = \{X_i = i\}$. (Physically, this means that there is a *match* in the ith position.)

a) Use the inclusion-exclusion principle (Exercise 1.25) to show that the probability of at least one match is

$$P(\cup_{i=1}^n A_i) = \Sigma_{k=1}^n (-1)^{k-1}/k!.$$

b) Let p_n be the probability of no match. Show that $\lim_n p_n = 1/e$.

Solution. a) Recall the inclusion-exclusion principle:

$$P(\cup_{i=1^n} A_i) = \Sigma_{k=1}^n (-1)^{k-1} S_k$$

where $S_k = \Sigma_{|J|=k} P(\cap_{j \in J} A_j)$ and $J \subseteq \{1, \dots, n\}$. As $P(\cap_{j \in J} A_j)$ is the probability of there is a match in the jth position for every $j \in J$, we can derive it by counting. Consider the k positions in J, the first position has n intergers to choose, and the second position then has $n-1$ integers to choose, etc. Hence

$$P(\cap_{j \in J} A_j) = \frac{1}{n(n-1)\cdots(n-k+1)} = \frac{(n-k)!}{n!}.$$

For $|J| = k$, there are $\binom{n}{k}$ numbers of distinct J. Thus

$$S_k = \Sigma_{|J|=k} P(\cap_{j \in J} A_j) = \binom{n}{k} \frac{(n-k)!}{n!} = \frac{n!}{k!(n-k)!} \cdot \frac{(n-k)!}{n!} = \frac{1}{k!}.$$

b)

$$\begin{aligned} p_n = 1 - P(\cup_{i=1}^n A_i) &= 1 - \Sigma_{k=1}^n \frac{(-1)^{k-1}}{k!} \\ &= 1 + \Sigma_{k=1}^n \frac{(-1)^k}{k!} \\ &= \Sigma_{k=0}^n \frac{(-1)^k}{k!} \to \frac{1}{e} \quad \text{as } n \to \infty. \end{aligned}$$

The limit is justified by the Taylor expansion of e^{-1}. □

Exercise 2.8. Let X have distribution $P(\lambda)$. Show that the function $i \mapsto P\{X = i\}$ is first increasing and then descreasing, with its maximum value at $\lfloor \lambda \rfloor$, the largest integer less than or equal to λ.

Solution. Let us first assume that λ is not an integer. For $k \geq 1$, let $P\{X = k\} - P\{X = k-1\} > 0$, that is

$$P\{X = k\} - P\{X = k-1\} = \frac{e^{-\lambda}\lambda^k}{k!} - \frac{e^{-\lambda}\lambda^{k-1}}{(k-1)!} = \frac{e^{-\lambda}\lambda^{k-1}}{(k-1)!}\left(\frac{\lambda}{k} - 1\right) > 0.$$

As $\frac{e^{-\lambda}\lambda^{k-1}}{(k-1)!} > 0$, this gives $\frac{\lambda}{k} - 1 > 0$ and $k < \lambda$, which shows that the function $i \mapsto P\{X = i\}$ is increasing for $k < \lambda$. Next, we let $P\{X = k\} - P\{X = k+1\} > 0$ for $k \geq 0$. Similarly, we get $k > \lambda - 1$, which shows that $i \mapsto P\{X = i\}$ is descreasing for $k > \lambda - 1$. The maximum value is at a k such that $\lambda - 1 < k < \lambda$. As k is an integer, it must be $\lfloor \lambda \rfloor$ (as we assumed that λ is not an integer).
If λ is an integer, by a similar argument, we can show that $i \mapsto P\{X = i\}$ is non-decreasing for $k \leq \lambda$ and non-increasing for $k \geq \lambda - 1$. Then the maximum value is at k such that

$\lambda - 1 \le k \le \lambda$. Since both k and λ are integers, the same maximum value is at both $k = \lambda$ and $k = \lambda - 1$. $\square$

Exercise 2.9. Let X has density

$$f_X(x) = \frac{1}{\pi(1+x^2)}, \qquad x \in \mathbb{R},$$

known as a *Cauchy density*. Show that

$$F_X(t) = 1/2 + (1/\pi)\arctan t.$$

Solution.

$$F_X(t) = \int_{-\infty}^{t} f_X(x)dx = \int_{-\infty}^{t} \frac{1}{\pi(1+x^2)}dx = (1/\pi)\arctan t + C$$

where C is a constant. To determine C, using the symmetry of $f_X(x)$, we have $F_X(0) = 1/2$, which gives $C = 1/2$. $\square$

Exercise 2.10. Show that the analogue of (2.8) for g continuous and strictly decreasing is

$$P\{Y \le t\} = 1 - F_X(h(t)-).$$

Solution.

$$P\{Y \le t\} = P\{g(x) \le t\} = P\{X \ge h(t)\} = 1 - P\{X < h(t)\} = 1 - F_X(h(t)-).$$

The last equality is justified as F_X is right-continuous. $\square$

Exercise 2.11. a) Let X have distribution function F. For each k, calculate the distribution function of $|X|^k$.
b) Let X be absolutely continuous. Compute the density of $|X|$.

Solution. a) If $k > 0$,

$$\begin{aligned}
P(|X|^k \le t) &= P(|X| \le t^{1/k}) \\
&= P(-t^{1/k} \le X \le t^{1/k}) \\
&= P(X \le t^{1/k}) - P(X < -t^{1/k}) \\
&= F(t^{1/k}) - F((-t^{1/k})-).
\end{aligned}$$

If $k < 0$,

$$\begin{aligned}
P(|X|^k \le t) &= P(|X| \ge t^{1/k}) \\
&= P(\{X \ge t^{1/k}\} \cup \{X \le -t^{1/k}\}) \\
&= P(X \ge t^{1/k}) + P(X \le -t^{1/k}) \quad \text{as disjoint} \\
&= 1 - P(X < t^{1/k}) + P(\le -t^{1/k}) \\
&= 1 - F(t^{1/k}-) + F(-t^{1/k}).
\end{aligned}$$

If $k = 0$, $P(|X|^0 \geq t) = P(1 \leq t)$. Since $\{1 \leq t\} = \emptyset, \forall t < 1$ and $\{1 \leq t\} = \Omega, \forall t \geq 1$,

$$P(1 \leq t) = \begin{cases} 0 & t < 1 \\ 1 & t \geq 1 \end{cases},$$

which is a step function.

b)

$$\begin{aligned} F_{|X|}(t) = P(|X| \leq t) &= P(-t \leq X \leq t) \\ &= P(X \leq t) - P(X < -t) \\ &= F_X(t) - F_X((-t)-) \\ &= F_X(t) - F_X(-t). \end{aligned}$$

The last equality is justified as X is absolutely continuous. Differentiating the above equation with respect to t on both sides gives the density

$$f_{|X|}(t) = f_X(t) + f_X(-t).$$

□

Exercise 2.12. Let X have distribution function F. Calculate the distribution functions of X^+ and X^-.

Solution. For X^+, if $t < 0$, then $\{X^+ \leq t\} = \emptyset$ and $P(X^+ \leq t) = 0$. If $t \geq 0$, then $\{X^+ \leq t\} = \{0 < X \leq t\} + \{X \leq 0\}$. Hence

$$\begin{aligned} P(X^+ \leq t) &= P(0 < X \leq t) + P(X \leq 0) \\ &= P(X \leq t) - P(X \leq 0) + P(X \leq 0) \\ &= P(X \leq t) \\ &= F(t). \end{aligned}$$

Combining both cases, $F_{X^+}(t) = P(X^+ \leq t) = F(t)\mathbb{1}\{t \geq 0\}$.

For X^-, $P(X^- \leq t) = 1 - P(X^- > t)$. If $t \geq 0$, then $\{X^- > t\} = \emptyset$ $P(X^t > t) = 0$. If $t < 0$, then $\{X^- > t\} = \{t < X \leq 0\} + \{X > 0\}$. Hence

$$\begin{aligned} P(X^- > t) &= P(t < X \leq 0) + P(X > 0) \\ &= P(X \leq 0) - P(X \leq t) + 1 - P(X \leq 0) \\ &= 1 - P(X \leq t) \\ &= 1 - F(t). \end{aligned}$$

Combining both cases, $F_{X^-}(t) = P(X^- \leq t) = \mathbb{1}\{t \geq 0\} + (1 - \mathbb{1}\{t \geq 0\})F(t)$. □

Exercise 2.13. Assume that X is positive and absolutely continuous with density f and distribution F satisfying $F(s) < 1$ for all $s < \infty$, and let $H(s) = -\log(1 - F(s))$.

a) Prove that H is differentiable with derivative $h = f/(1-F)$, which is termed the *hazard function* of T (or F).

b) Prove that for each t,

$$h(t) = \lim_{h \downarrow 0} \frac{1}{h} P\{X \leq t+h | X > t\}.$$

c) Prove that $\int_0^\infty h(t)dt = \infty$.

d) Prove that

$$P\{X > t+s | X > t\} = \exp\left[- \int_t^{t+s} h(u)du \right]$$

for each t and s.

e) Prove that $X \stackrel{d}{=} E(\lambda)$ if and only if $h \equiv \lambda$.

Solution. a) Since $F(s) < 1$ for all $s < \infty$, we have $1 - F(s) > 0$. As $\log(x)$ is continuous and differentiable for all $x > 0$, it is clear that H is differentiable. Applying chain rule, it can be easily showed that $h = f/(1-F)$.

b)

$$\begin{aligned}
\lim_{h \downarrow 0} \frac{1}{h} P(X \leq t+h | X > t) &= \lim_{h \downarrow 0} \frac{1}{h} \frac{P(t < x \leq t+h)}{P(x > t)} \\
&= \lim_{h \downarrow 0} \frac{1}{h} \frac{P(x \leq t+h) - P(x \leq t)}{1 - P(X \leq t)} \\
&= \lim_{h \downarrow 0} \frac{1}{h} \frac{F(t+h) - F(t)}{1 - F(t)} \\
&= \frac{1}{1-F(t)} \lim_{h \downarrow 0} \frac{F(t+h) - f(t)}{h} \\
&= \frac{f(t)}{1-F(t)} \\
&= h(t).
\end{aligned}$$

c) As X is positive, $F(0) = 0$. Then,

$$\begin{aligned}
\int_0^\infty h(t)dt &= \lim_{n \to \infty} H(n) - H(0) \\
&= \lim_{n \to \infty} -\log(1 - F(n)) + \log(1 - F(0)) \\
&= \lim_{n \to \infty} -\log(1 - F(n)).
\end{aligned}$$

Given $F(n) \to 1$ as $n \to \infty$, it is clear that $\int_0^\infty h(t)dt = \infty$.

d)

$$
\begin{aligned}
P(X > t+s | X > t) &= \frac{P(X > t+s)}{P(X > t)} \\
&= \frac{1 - P(X \le t+s)}{1 - P(X \le t)} \\
&= \frac{1 - F(t+s)}{1 - F(t)} \\
&= \frac{\exp\{-H(t+s)\}}{\exp\{-H(t)\}} \\
&= \exp\{-(H(t+s) - H(t))\} \\
&= \exp\left[- \int_t^{t+s} h(u) du \right].
\end{aligned}
$$

e) For sufficiency, if $X \stackrel{d}{=} E(\lambda)$, we have $f(x) = \lambda e^{=\lambda x}, x > 0$ and $F(x) = 1 - e^{-\lambda x}, x > 0$. Thus,

$$
h(x) = \frac{f(x)}{1 - F(x)} = \frac{\lambda e^{-\lambda x}}{e^{-\lambda x}} = \lambda, \quad \forall x > 0.
$$

For necessity, if $h \equiv \lambda$, we have $\lambda = f/(1 - F)$. This is equivalent to the following ODE (ordinary differential equation):

$$
\frac{\partial F}{\partial x} + \lambda F = \lambda.
$$

Taking derivatives of $e^{\lambda x} F$ gives

$$
\frac{\partial (e^{\lambda x} F(x))}{\partial x} = e^{\lambda x} \frac{\partial F(x)}{\partial x} + \lambda e^{\lambda x} F(x) = \lambda e^{\lambda x}.
$$

The last equality is obtained by using the above ODE. Therefore, we have

$$
e^{\lambda x} F(x) = e^{\lambda x} + C \quad \Rightarrow \quad F(x) = 1 + C e^{-\lambda x},
$$

where C is a constant. Since X is positive and $F(0) = 0$, we have $C = -1$ and $F(x) = 1 - e^{-\lambda x}$ that is the distribution function of $E(\lambda)$. $\square$