1. **Discrete Probability Spaces.**
Let Ω be a countable set.

(a) Show that its power set 2^Ω is a σ-algebra.

Solution.

i. The power set of Ω is the set of all subsets of Ω. The entirety of Ω is a subset of Ω and thus $\Omega \in 2^\Omega$. (Note: that the empty set is also a subset of Ω and thus $\Phi \in 2^\Omega$as well.)

ii. Take an event $A \in 2^\Omega$. Then $A \subseteq \Omega$. The complement of this event $A^C \in \Omega \setminus A \subseteq \Omega$ by definition. Thus, we have that for every $A \in 2^\Omega$ that $A^c \in 2^\Omega$ as well.

iii. Let $A_1, A_2, ...$ be a countable number of sets such $\forall A_i$, we have $A_i \in 2^\Omega$ then $\bigcup_i A_i \subseteq \Omega$ and thus $\bigcup_i A_i \in 2^\Omega$ by definition of 2^Ω.

$\blacksquare$

(b) Let $p : \Omega \to [0, 1]$ be a probability mass function. That is,

$$\sum_{\omega \in \Omega} p(w) = 1.$$

Define a function

$$\mathbb{P}_p : 2^\Omega \to [0, 1]$$

by

$$\mathbb{P}_p(A) \triangleq \sum_{w \in A} p(w)$$

for $A \in 2^\Omega$. Show that $(\Omega, 2^\Omega, \mathbb{P}_p)$ is a probability space.

Solution.
2^Ω is a valid sigma algebra as shown in part (a). It must now be shown that $\mathbb{P}_p$ is a valid probability law.

i. $\mathbb{P}(\Omega) = \sum_{\omega \in \Omega} p(\omega) = 1$

ii. $0 \le \mathbb{P}_p(A) \le 1$ for $A \subseteq \Omega$ is clear by definition.

iii. Let $\{A_i\}_{i=1}^n$ be a sequence of n mutually exclusive events. Then

$$\mathbb{P}\left(\bigcup_{i=1}^n A_i\right) = \sum_{w \in \bigcup_{i=1}^n A_i} p(\omega) = \sum_{i=1}^n \mathbb{P}_p(A_i)$$

$\blacksquare$

2. **σ-Algebra.**
Let Ω be an arbitrary set. All σ-algebras below are collections of subsets of Ω.

(a) Consider a a sequence of σ-algebras $\{\mathcal{F}_n\}_{n=1}^{\infty}$ such that $\mathcal{F}_n \subset \mathcal{F}_{n+1}$ for all $n \in \mathbb{N}$. Is

$$\mathcal{G} := \bigcup_{n=1}^{\infty} \mathcal{F}_n$$

a σ-algebra?

Solution. We conclude that $\mathcal{G}$ is **not** a σ-algebra.
Let $\Omega = \mathbb{R}$ and $J_n := \{1, ..., n\}$ for $n \in \mathbb{N}$.
Consider the sequence of σ-algebras $\{\mathcal{F}_n\}_{n=1}^{\infty}$

$$\mathcal{F}_n = \{\phi, \mathbb{R}\} \cup 2^{J_n} \cup \{\mathbb{R} \setminus s : s \in 2^{J_n}\}.$$

For all $n \in \mathbb{N}$, we have n in some sigma algebra – concretely, $\{n\} \in \mathcal{F}_n$, which implies that $\{n\} \in \bigcup_{n=1}^{\infty} \mathcal{F}_n$. Assume towards a contradiction that $\bigcup_{n=1}^{\infty} \mathcal{F}_n$ is a sigma algebra, then $\bigcup_{m=1}^{\infty}\{m\} \in \bigcup_{n=1}^{\infty} \mathcal{F}_n$. In other words, $\mathbb{N} \in \bigcup_{n=1}^{\infty} \mathcal{F}_n$. However, $\nexists i$ s.t. $\mathbb{N} \in \mathcal{F}_i$ which implies that $\mathbb{N}$ cannot be in the infinite union, which is a contradiction. So $\bigcup_{n=1}^{\infty} \mathcal{F}_n$ is not a sigma algebra. ∎

(b) Let $\mathcal{F}_{i \in I}$ be an arbitrary (possibly uncountable) collection of σ-algebras. Show that

$$\mathcal{H} := \bigcap_{i \in I} \mathcal{F}_i$$

is a σ-algebra.

Solution.
We must show three things to show $\mathcal{H}$ is a sigma algebra:

i. $\Omega \in \mathcal{H}$.
Since each $\mathcal{F}_i$ is a sigma algebra, $\Omega \in \mathcal{F}_i$ for all i giving that Ω is in the intersection.

ii. If $A \in \mathcal{H}$ then $A^c \in \mathcal{H}$.
Let $S \in \mathcal{H}$. This implies that S is in all the sigma algebras, and since all of them are sigma algebras, then they must contain S^c giving that S^c is also in the intersection.

iii. If $\{A_i\} \in \mathcal{H}$ then $\cup_i A_i \in \mathcal{H}$.
Let $\{A_i\} \in \mathcal{H}$, be some countable sequence of events. We need to show that $\cup_i A_i$ is contained in every $\mathcal{F}_i$. Let $\mathcal{F}_i$ be some sigma algebra in the collection. We know that the elements of $\{A_i\}$ are in $\mathcal{F}_i$, and since $\mathcal{F}_i$ is a sigma algebra, $\cup_i A_i \in \mathcal{F}_i$ giving that $\cup_i A_i$ must be in $\mathcal{H}$.

∎

(c) Let $\mathcal{F}_1$ and $\mathcal{F}_2$ be arbitrary σ-algebras. Define their Cartesian product

$$\mathcal{F}_1 \times \mathcal{F}_2 := \{A_1 \times A_2 \mid A_1 \in \mathcal{F}_1, A_2 \in \mathcal{F}_2\}$$

where

$$A_1 \times A_2 := \{(a_1, a_2) \mid a_1 \in A_1, a_2 \in A_2\}.$$

Either prove that $\mathcal{F}_1 \times \mathcal{F}_2$ is a σ-algebra or provide a counterexample.

Solution.
Let the sigma algebras $\mathcal{F}_1, \mathcal{F}_2$ be over the sample space $\mathbb{R}$. Consider the example where $\mathcal{F}_1$ and $\mathcal{F}_2$ consist of the Borel algebra of the half-infinite planes presented in lecture. Then for some $a_1, a_2 \in \mathbb{R}$, we have

$$A = ((-\infty, a_1], (-\infty, a_2)) \in \mathcal{F}_1 \times \mathcal{F}_2$$

Note that $\{A\} \in \mathcal{F}_1 \times \mathcal{F}_2$ by construction. Assume towards a contradiction that $\mathcal{F}_1 \times \mathcal{F}_2$ is a sigma algebra, then we must have $\{A^c\} \in \mathcal{F}_1 \times \mathcal{F}_2$. However, $A^c = ([a_1, \infty] \times (-\infty, a_2] \cup [a_1, \infty] \times [a_2, \infty) \cup (-\infty, a_1] \times [a_2, \infty))$. But, there are no intervals $I_1 \in \mathcal{F}_1$ and $I_2 \in \mathcal{F}_2$ s.t. $I_1 \times I_2 = A^c$ since A^c cannot be written as a single two dimensional interval. ■

3. **Properties of Probability Measures.**
Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space. Prove the following properties of $\mathbb{P}$.

(a) **Law of Complement Probability:**

$$\mathbb{P}(A) = 1 - \mathbb{P}(A^c), \qquad (\forall A \in \mathcal{F})$$

where $A^c = \Omega \setminus A$ is the complement of A.

Solution.
We note that by definition A and A^c are mutually exclusive events which partition Ω. Thus

$$\Pr(\Omega) = \mathbb{P}(A \cup A^c) = \mathbb{P}(A) + \mathbb{P}(A^c) = 1$$
$$\implies \mathbb{P}(A) = 1 - \mathbb{P}(A^c)$$

■

(b) **Monotonicity:** If $A, B \in \mathcal{F}$ with $A \subseteq B$, then $\mathbb{P}(A) \leq \mathbb{P}(B)$

Solution.
We can write event B as $B = A \cup (A^c \cap B)$ and note that A and $(A^c) \cap B$ are mutually exclusive events. Thus

$$\mathbb{P}(B) = \mathbb{P}(A \cup (A^c \cap B)) = \mathbb{P}(A) + \mathbb{P}(A^c \cap B) \geq Pr(A)$$

■

(c) **Union Bound:** *Note: There was a slight notation change from the original statement.* For any $\{A_i\}_{i=1}^{\infty} \subseteq \mathcal{F}$, we have

$$\mathbb{P}\left(\bigcup_{i=1}^{\infty} A_i\right) \leq \sum_{i=1}^{\infty} \mathbb{P}(A_i)$$

Solution.
Define

$$B_i := A_i \setminus \bigcup_{k=1}^{i-1} A_k$$

Then

$$\bigcup_i B_i = \bigcup_i A_i$$

and $B_i \subseteq A_i$ for all i. Noting that all of the B_i are disjoint, it follows that

$$\mathbb{P}\left(\bigcup_{i=1}^{\infty} A_i\right) = \mathbb{P}\left(\bigcup_{i=1}^{\infty} B_i\right) = \sum_{i=1}^{\infty} \mathbb{P}(B_i) \leq \sum_{i=1}^{\infty} \mathbb{P}(A_i)$$

by monotonocity. ■

(d) **Continuity of Probability:**

i. Let $\{A_n\}_{n=1}^{\infty} \subseteq \mathcal{F}$ be a sequence of events increasing to $A \in \mathcal{F}$. That is, $A_1 \subseteq A_2 \subseteq A_3 \subseteq ...$ and $\bigcup_{n=1}^{\infty} A_n = A$. Prove that $\lim_{n\to\infty} \mathbb{P}(A_n) = \mathbb{P}(A)$. Deduce that for any $\{A'_n\}_{n=1}^{\infty} \subseteq \mathcal{F}$, we have

$$\lim_{m\to\infty} \mathbb{P}\left(\bigcup_{n=1}^{m} A'_n\right) = \mathbb{P}\left(\bigcup_{n=1}^{\infty} A'_n\right)$$

Solution.
For more clarification, see appendix (a)
We want to show $\lim_{n\to\infty} \mathbb{P}(A_n) = \mathbb{P}(A)$. Note that $\mathbb{P}$ is a monotonic function for increasing sets and bounded from above and below. As a result, from basic real analysis we get that for any increasing sequence $B_1 \subseteq B_2 \subseteq ...$ we have $\lim_n \mathbb{P}(B_n)$ to converge to some value. So now we know that LHS converges to some value, call it p, and we have RHS to be some value, call it $p* \in [0, 1]$. We claim that $p = p*$. Assume towards a contradiction that $p \neq p*$. Let $\Delta = |p * -p|$. The first case is that $p* > p$. In this case, $\exists N$ s.t. $\forall n \geq N$ we have $|p - \mathbb{P}(A_n)| < \frac{\Delta}{2}$. However, that implies that $\exists S \subseteq A$ with

$\mathbb{P}(S) > \frac{\Delta}{2}$ s.t. $\forall s \in S, \nexists A_i$ s.t. $s \in A_i$. Which is a contradiction since that implies $s \notin A$.
The second case, assume $p > p*$. Let $\Delta = p - p*$. Then, $\exists N$ s.t. $\forall n \geq N$ we have $|p - \mathbb{P}(A_n)| < \frac{\Delta}{2}$. So $\forall n \geq N$, we have $\mathbb{P}(A_n \setminus A) > \frac{\Delta}{2}$ which is a contradiction since if an element is in some A_n then it must be in A and $A_n \setminus A = \phi$ which should have probability zero.
As a result, we get $p* = p$ as desired.
To see $\lim_{m\to\infty} \mathbb{P}\left(\bigcup_{n=1}^{m} A'_n\right) = \mathbb{P}\left(\bigcup_{n=1}^{\infty} A'_n\right)$, denote $\bigcup_{n=1}^{m} A'_n$ as C_n and $\bigcup_{n=1}^{\infty} A'_n$ to be C. Note that $C_1 \subseteq ... \supseteq C_n \subseteq ...$, and that $C = \bigcup_{n=1}^{\infty} C_n$. Using the above, we see that $\lim_{m\to\infty} \mathbb{P}(C_m) = \mathbb{P}(C)$ as desired. ■

ii. Let $\{B_n\}_{n=1}^{\infty} \subseteq \mathcal{F}$ be a sequence of events decreasing to $B \in \mathcal{F}$. That is, $B_1 \supseteq B_2 \supseteq B_3 \supseteq ...$ and $\bigcap_{n=1}^{\infty} B_n = B$. Prove that $\lim_{n\to\infty} \mathbb{P}(B_n) = \mathbb{P}(B)$. Deduce that for any $\{B'_n\}_{n=1}^{\infty} \subseteq \mathcal{F}$, we have

$$\lim_{m\to\infty} \mathbb{P}\left(\bigcap_{n=1}^{m} B'_n\right) = \mathbb{P}\left(\bigcap_{n=1}^{\infty} B'_n\right)$$

Solution.
For more clarification, see appendix (b)
We want to show $\lim_{n\to\infty} \mathbb{P}(B_n) = \mathbb{P}(B)$. Note that $\mathbb{P}$ is a monotonic decreasing function for decreasing sets and bounded from above and below. As a result, from basic real analysis we get that for any decreasing sequence $B_1 \subseteq B_2 \subseteq ...$ we have $\lim_n \mathbb{P}(B_n)$ to converge to some value. So now we know that LHS converges to some value, call it p, and we have RHS to be some value, call it $p* \in [0, 1]$. We claim that $p = p*$. Assume towards a contradiction that $p \neq p*$. Let $\Delta = |p * -p|$. The first case is that $p* > p$. In this case, $\exists N$ s.t. $\forall n \geq N$ we have $|p - \mathbb{P}(B_n)| < \frac{\Delta}{2}$. However, that implies that $\exists S \subseteq B$ with $\mathbb{P}(S) > \frac{\Delta}{2}$ s.t. $\forall s \in S, \nexists B_i$ s.t. $s \in B_i$. However, this is a contradiction since B is the intersection and each element in B must be in all B_i.
The second case, assume $p > p*$. Let $\Delta = p - p*$. Then, $\exists N$ s.t. $\forall n \geq N$ we have $|p - \mathbb{P}(B_n)| < \frac{\Delta}{2}$. Note that $\mathbb{P}(B_n)$ is monotonically decreasing function, so $\forall n \in N$, we have $\mathbb{P}(B_n \setminus B) > \frac{\Delta}{2}$. This implies that $\exists S \in B_i \forall i \in \mathbb{N}$ s.t. $Pr(S) > \frac{\Delta}{2}$ and $\forall x \in S, x \notin B$. However, this is a contradiction since $\forall x \in S$, we have $x \in A_i \forall i \in \mathbb{N}$ which implies that $x \in B$.
To see $\lim_{m\to\infty} \mathbb{P}\left(\bigcap_{n=1}^{m} B'_n\right) = \mathbb{P}\left(\bigcap_{n=1}^{\infty} B'_n\right)$, denote $\bigcap_{n=1}^{m} B'_n$ as C_n and $\bigcap_{n=1}^{\infty} A'_n$ to be C. Note that $C_1 \supseteq ... \supseteq C_n \supseteq ...$, and that $C = \bigcap_{n=1}^{\infty} C_n$. Using the above, we see that $\lim_{m\to\infty} \mathbb{P}(C_m) = \mathbb{P}(C)$ as desired. ■

(e) **Law of Total Probability:** Let $\{A_n\}_{n=1}^{\infty} \subseteq \mathcal{F}$ be a partition of Ω and $B \in \mathcal{F}$.

Then

$$\mathbb{P}(B) = \sum_{n=1}^{\infty} \mathbb{P}(A_n)\mathbb{P}(B \mid A_n).$$

Is the argument valid when $\mathbb{P}(A_{n'}) = 0$ for some $n' \in \mathbb{N}$?

Solution.
We first begin by noting that

$$B = B \cap \Omega$$

and next note that since $\{A_n\}_{n=1}^{\infty}$ is a partition of Ω that B can be expressed as

$$B = (A_1 \cap B) \cup (A_2 \cap B) \cup ... \cup (A_n \cap B) = \bigcup_{n=1}^{\infty}(A_n \cap B)$$

from which we use the fact that the probability of the union of mutually exclusive events is the same as the sum of the probability of each event to obtain

$$\mathbb{P}(B) = \mathbb{P}\left(\bigcup_{n=1}^{\infty}(A_n \cap B)\right) = \sum_{n=1}^{\infty} \mathbb{P}(A_n \cap B).$$

By Bayes' Rule this becomes

$$\mathbb{P}(B) = \sum_{n=1}^{\infty} \mathbb{P}(A_n)\mathbb{P}(B \mid A_n).$$

We note that the argument is still valid if $\mathbb{P}(A_{n'}) = 0$ for some $n' \in \mathbb{N}$ since that intersection need not be considered in the final result.

■

4. **Measurability of Indicators.**
Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space. For $A \in \mathcal{F}$, define the function $\mathbb{1}_A : \Omega \to \mathbb{R}$ by

$$\mathbb{1}_A(w) \triangleq \begin{cases} 1, & \omega \in A \\ 0, & \omega \notin A \end{cases}$$

Prove that $\mathbb{1}_A$ is a random variable over $(\Omega, \mathcal{F}, \mathbb{P})$.

Solution.
To show that $\mathbb{1}_A$ is a random variable with respect to $(\Omega, \mathcal{F}, \mathbb{P})$, we need

$$\mathbb{1}_A^{-1} : \mathbb{R} \to \mathcal{F}$$

such that

$$\mathbb{1}_A^{-1}(B) \in \mathcal{F} \qquad (\forall B \in \mathcal{B}(\mathbb{R}))$$

We define

$$\mathbb{1}_A^{-1}(B) = \begin{cases} A & 0 \notin B, 1 \in B \\ \Omega \setminus A & 0 \notin B, 1 \notin B \\ \Omega & 0 \in B, 1 \in B \\ \phi & \text{otherwise} \end{cases}$$

Note that $\{\phi, A, \Omega \setminus A, \Omega\}$ must be in $\mathcal{F}$ since it's a sigma algebra. Thus, $\mathbb{1}_A$ is a random variable. ■

5. **Induced Probability Measure vs. Probability Law.**
Let $X : \Omega \to \mathbb{R}^d$ be a random variable on $(\Omega, \mathcal{F}, \mathbb{P})$ with probability law $\mathbb{P}_X$. That is,

$$P_X(B) := P(X^{-1}(B) \qquad (\text{for } B \in \mathcal{B}(\mathbb{R}^d))$$

is a probability measure on $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$.

(a) Assume X is a discrete random variable ($supp(\mathbb{P}_X)$ is countable and let

$$p_X : supp(\mathbb{P}_X) \to [0, 1]$$

be its PMF, defined by

$$p_X(x) := \mathbb{P}_X(\{x\}) \qquad (x \in supp(\mathbb{P}_X))$$

Prove that the probability measure that is induced by the pmf p_X, $\mathbb{P}_{p_X}$, coincides with the law of $X, \mathbb{P}_X$. In other words, show

$$\mathbb{P}_{p_X}(B) = \mathbb{P}_X(B) \qquad (\forall B \in \mathcal{B}(\mathbb{R}^d))$$

Solution.
First note that only subsets of the form

$$A := B \cap S \qquad (B \in \mathcal{B}(\mathbb{R}^d))$$

where $S = supp(\mathbb{P}_X)$ are considered since this will remove all the zero probability events in the space induced by X. In other words,

$$\begin{aligned} \mathbb{P}_X(B) &= \mathbb{P}_X(B \cap (S \cup S^c)) \\ &= \mathbb{P}_X(B \cap S \cup B \cap S^c) \\ &= \mathbb{P}_X(B \cap S) + \mathbb{P}_X(B \cap S^c) \\ &= \mathbb{P}_X(B \cap S) \end{aligned}$$

This narrowing is important because it allows us to have only countable sets. Now we note that there are two perspectives on probability measures: the probability measure induced by the pmf, and the pushforward measure of $\mathbb{P}$ through a random variable.

i. Given a pmf $p : \Omega \to [0,1]$ for $(\Omega, \mathcal{F})$ such that $\sum_{\omega\in\Omega} p(\omega) = 1$ we can define the probability measure with respect to the pmf, p, by $\mathbb{P}_p(A) \triangleq \sum_{\omega\in A} p(\omega)$ for $\omega \in A$. We now consider the pmf induced by a random variable X, p_X defined by

$$p_X(x) \triangleq \mathbb{P}_X(\{x\}) = \mathbb{P}(X^{-1}(\{x\})) = \mathbb{P}(\{\omega \in \Omega \mid X(\omega) = x\})$$

So the probability of event A occurring using the probability measure with respect to the pmf induced by the random variable X is

$$\mathbb{P}_{p_X}(A) = \sum_{a\in A} p_X(a)$$

ii. The other perspective is in the context of being in a probability space, $(\Omega, \mathcal{F}, \mathbb{P})$, and there being a random variable $X : \Omega \to \mathbb{R}^d$. An event A can be written as the union of the singleton events in A. That is, if A is comprised of events $\{a_i\}_{i=1}^{\infty}$, we can say $A = \cup_{i=1}^{n} a_i$ (n could be infinity). Thus

$$\begin{aligned}\mathbb{P}_X(A) = \mathbb{P}_{p_X}\left(\bigcup_{i=1}^{n} a_i\right) &= \sum_{i=1}^{\infty} \mathbb{P}_X(\{a_i\}) \\ = \sum_{i=1}^{\infty} \mathbb{P}_X(\{a_i\}) &= \sum_{i=1}^{\infty} p_X(a_i) = \sum_{a\in A} p_x(a)\end{aligned}$$

Thus, the two perspectives are equivalent. The probability measure induced by the pmf of a random variable, X, is the same as the probability measure induced by the pushforward measure of $\mathbb{P}$ through X. ■

(b) Assume X is continuous with probability density function $f_X : \mathbb{R}^d \to \mathbb{R}_{\geq 0}$. Any PDF f (a nonnegative integrable function with $\int_{\mathbb{R}^d} f(x)dx = 1$) induces a measure $\mathbb{P}_f$ on $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$ by defining

$$\mathbb{P}_f(B) := \int_B f(x)dx \qquad (\forall B \in \mathcal{B}(\mathbb{R}^d))$$

Prove that $\mathbb{P}_{f_X}$ coincides with $\mathbb{P}_X$ on the generating set of the Borel σ-algebra. I.e

$$\mathbb{P}_{f_X}((-\infty, a_1] \times ... \times (-\infty, a_d]) = \mathbb{P}_X((-\infty, a_1] \times ... \times (-\infty, a_d])$$

for all $a_1, ..., a_d \in \mathbb{R}$.

Solution.
Assume $d = 1$.
Let $a \in \mathbb{R}$. We want to show that

$$\mathbb{P}_{f_x}((-\infty, a]) = \mathbb{P}_X((-\infty, a]).$$

We have

$$\mathbb{P}_{f_x}((-\infty, a]) = \int_{(-\infty,a]} f_x(x)dx.$$

From here we use the Fundamental Theorem of Calculus from real analysis to obtain

$$\int_{(-\infty,a]} f_x(x)dx = F_x(a) - \lim_{t \to -\infty} F_x(t) = F_x(a)$$

since f_x is the derivative of CDF F_X. We now realize that this is $\mathbb{P}_X((-\infty, a])$ by the definition of the CDF. ■

6. **Generation of Samples with Arbitrary Distribution.**
Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and $X : \Omega \to \mathbb{R}$ be a random variable on it. Denote the probability law of X by $\mathbb{P}_X$, and let

$$F_X : \mathbb{R} \to [0, 1]$$

defined by

$$F_X(x) = \mathbb{P}((-\infty, x]) \qquad ((\forall x \in \mathbb{R}))$$

be the cumulative distribution function.

(a) Show that $Y := F_X(X)$ is a random variable.

Solution.
For Y to be a random variable, we show that the composition of measurable functions is measurable, making the composition $F_X \circ X$ a valid random variable. To see this, let X be a random variable mapping from $(\Omega_x, \mathcal{F}_x)$ to $(\Omega_y, \mathcal{F}_y)$, and let Y be a random variable mapping from $(\Omega_z, \mathcal{F}_z)$. To have $Z = Y \circ X$ be a random variable, we need $Z^{-1}(B) \in \mathcal{F}_x, \forall B \in \mathcal{B}(\mathcal{F}_z)$. However, we know $Y^{-1}(B) \in \mathcal{F}_y, \forall B \in \mathcal{B}(\mathcal{F}_z)$ and $X^{-1}(B) \in \mathcal{F}_x, \forall B \in \mathcal{B}(\mathcal{F}_y)$. So we have $Z^{-1}(B) = X^{-1}(Y^{-1}(B)) \in \mathcal{F}_x \forall B \in \mathcal{B}(\mathcal{F}_z)$, since $Y^{-1}(B) \in \mathcal{F}_y$, and $X^{-1}(B) \in \mathcal{F}_x$. We have F_X to be a measurable function since it is monotonic and bounded by construction and monotonic functions are measurable since they have at most countable discontinuities (and thus the measure of the discontinuities is 0), so the composition $F_X \circ X$ is a random variable, with the preimage being $X^{-1} \circ F_X^{-1}$. ■

(b) Prove that Y is uniformly distributed on $[0, 1]$.It suffices to show that F_Y, the CDF of Y, is given by $F_Y(y) = y$ for all $y \in [0, 1]$ and then explain how the CDF argument extends to the entire probability law $\mathbb{P}_Y$.

Solution.
We compute the CDF of Y to be

$$\begin{aligned} F_Y(y) &= \mathbb{P}_Y((-\infty, y]) \\ &= \mathbb{P}_X(Y^{-1}((-\infty, y])) \\ &= \mathbb{P}_X(F_X^{-1}(-\infty, y])) \\ &= F_X(F_X^{-1}(y)) = y \end{aligned}$$

where $y \in [0, 1]$ for invertibility of CDF. ■

(c) Let $U \sim U[0, 1]$ be a uniformly distributed random variable. Propose a measurable function $T : [0, 1] \to \mathbb{R}$ such that $T(U)$ has distribution $\mathbb{P}_X$ (or equivalently, CDF F_X). Prove your answer, but the measurability of T need not be established.

Solution.
From parts a and b, we know $U = F_X(X)$. In part a we showed F_X is a measurable function and thus it's preimage is measurable. So by applying the preimage on U, we get $F_X{-}1(U) = F_X^{-1}(F_X(X)) = X$. By construction, we have $\mathbb{P}_X$ as the probability measure defined by X, and since $F_X^{-1}(U) = X$, then the probability measure defined by $F_X^{-1}(U)$ is also $\mathbb{P}_X$ ■

7. **Transition Kernels.**
Let $\Omega = \{0, 1\}$ and consider the following construction of the transition kernel $\kappa(\cdot \mid \cdot)$ from $(\Omega, 2^\Omega)$ to itself. For $\omega \in \{0, 1\}$, let $Ber(\alpha_\omega)$ be a Bernoulli distribution with parameter $\alpha \in [0, 1]$. The PMF of this distribution is

$$p_\omega(1) = 1 - p_\omega(0) = \alpha_\omega$$

Set $\kappa(\cdot \mid 0) = Ber(\alpha_0)$ and $\kappa(\cdot \mid 1) = Ber(1-\alpha_1)$. This construction of $\kappa(\cdot \mid \cdot)$ is called the *binary channel with flip parameters* $(\alpha_0, \alpha_1) \in [0, 1]^2$. Prove the following claims:

(a) $\kappa(\cdot \mid \cdot)$ is a transition kernel (verify the conditions in the definition).

Solution.
Have measurable spaces $(\Omega, 2^\Omega)$ and $(\Omega, 2^\Omega)$. Need to show that

$$\begin{aligned} &(i). && \kappa(\cdot \mid x) \in \mathcal{P}(\Omega) && (\forall x \in \Omega) \\ &(ii). && \kappa(B \mid \cdot) : \Omega \to \mathbb{R} \text{ is a random variable} && (\forall B \in 2^\Omega) \end{aligned}$$

i. $\kappa(\cdot \mid x)$ takes a $w \in \Omega$ and maps it to a Bernoulli distribution with parameter α_w. Since the Bernoulli distribution is a probability measure it follows that $\kappa(\cdot \mid x)$ is as well.

ii. We need $\kappa(B \mid \cdot)^{-1} : \mathbb{R} \to \Omega$ such that $\kappa(B \mid \cdot)^{-1}(B) \in 2^{\Omega}, \forall B \in \mathcal{B}(\mathbb{R})$. However, since the sigma field is the power set of omega, then any element in the image of $\kappa(B \mid \cdot)^{-1}$ is in the sigma algebra.

■

(b) Let $P_X = Ber(0.5)$ and define

$$P_Y := \mathbb{E}\left[\kappa((B \mid X)\right] \qquad (B \in 2^{\Omega})$$

Show that whenever $\alpha_0 = \alpha_1$ we have $P_Y = Ber(0.5)$.

Solution.

$$\begin{aligned}
P_Y := \mathbb{E}[\kappa(\cdot \mid X)] &= \sum_{x \in supp(P_X)} \kappa(\cdot \mid x) P_X(x) \\
&= \kappa(\cdot \mid 0)\frac{1}{2} + \kappa(\cdot \mid 1)\frac{1}{2} \\
&= \frac{1}{2}\left(Ber(\alpha_0) + Ber(1-\alpha_1)\right) \\
&= \frac{1}{2}\left(Ber(\alpha_1) + Ber(1-\alpha_1)\right)
\end{aligned}$$

From here we now analyze the output of the probability measure.

$$\begin{aligned}
P_Y(\{0\}) &= \frac{1}{2}\left(Ber(\alpha_1)(\{0\}) + Ber(1-\alpha_1)(\{0\})\right) \\
&= \frac{1}{2}(\alpha_1 + 1 - \alpha_1) = \frac{1}{2} \\
P_Y(\{1\}) &= \frac{1}{2}\left(Ber(\alpha_1)(\{1\}) + Ber(1-\alpha_1)(\{1\})\right) \\
&= \frac{1}{2}(\alpha_1 + 1 - \alpha_1) = \frac{1}{2} \\
P_Y(\{0,1\}) &= \frac{1}{2}\left(Ber(\alpha_1)(\{0,1\}) + Ber(1-\alpha_1)(\{0,1\})\right) \\
&= \frac{1}{2}(1+1) = 1 \\
P_Y(\{\}) &= \frac{1}{2}\left(Ber(\alpha_1)(\{\}) + Ber(1-\alpha_1)(\{\})\right) \\
&= \frac{1}{2}(0+0) = 0
\end{aligned}$$

and conclude that $P_Y = Ber(0.5)$. ■

(c) Give a counterexample to the symmetry of P_Y when $\alpha_0 \neq \alpha_1$.

Solution.
Let $\alpha_0 = 0$ and $\alpha_1 = 1$. Then,

$$P_Y = \frac{1}{2}\big(Ber(1) + Ber(1)\big)$$

giving that

$$P_Y(\{1\}) = \frac{1}{2}(1+1) = 1$$
$$P_Y(\{0\}) = \frac{1}{2}(0+0) = 0$$

Thus $P_Y \neq Ber(0.5)$. ∎

8. **Appendix.**

(a)

We first note that the sequence $\{\mathbb{P}(A_i)\}_{i=1}^{\infty}$ is monotonically increasing. That is,

$$\mathbb{P}(A_i) \leq \mathbb{P}(A_{i+1}) \qquad (\forall i)$$

by the monotonicity property proved above. Next we note that $\{\mathbb{P}(A_i)\}_{i=1}^{\infty}$ is bounded above since $\mathbb{P}(A_i) \leq 1\ \forall i$.

Letting E be the range of $\{\mathbb{P}(A_i)\}_{i=1}^{\infty}$, define $s = \sup E$. Then

$$\mathbb{P}(A_i) \leq s. \qquad (\forall i)$$

For every $\epsilon > 0$, there exists $N \in \mathbb{N}$ such that

$$s - \epsilon < \mathbb{P}(A_N) < s$$

for otherwise $s - \epsilon$ would be an upper bound of E. Since $\{\mathbb{P}(A_i)\}_{i=1}^{\infty}$ increases, $n \geq N$ implies that

$$s - \epsilon < \mathbb{P}(A_n) \leq s$$

giving that $\{\mathbb{P}(A_i)\}_{i=1}^{\infty}$ converges to s.

It must now be shown that $s = \mathbb{P}(A)$. Assume towards a contradiction that $s \neq \mathbb{P}(A)$. There are two cases: $(i).\mathbb{P}(A) < s$ and $(ii).\mathbb{P}(A) > s$.

If $\mathbb{P}(A) < s$ then that implies $A \subset A_{n'}$ for some $n' \in \mathbb{N}$. But this would mean that $\mathbb{P}(A) < \mathbb{P}(A_{n'})$ which contradicts the assumption that $\{A_n\}$ increases to A and thus $\mathbb{P}(A) \not< s$.

If $\mathbb{P}(A) > s$ then that means $A_n \subset A$ for all n again contradicting that $\{A_n\}$ increases to A and thus $\mathbb{P}(A) \not> s$.

Thus it must be the case that $\sup E = \mathbb{P}(A)$ and it finally follows that

$$\lim_{n\to\infty} \mathbb{P}(A_n) = \mathbb{P}(A).$$

(b)

We first note that the sequence $\{\mathbb{P}(B_i)\}_{i=1}^{\infty}$ is monotonically decreasing. That is,

$$\mathbb{P}(B_i) \geq \mathbb{P}(B_{i+1}) \qquad (\forall i)$$

by the monotonicity property proved above. Next we note that $\{\mathbb{P}(B_i)\}_{i=1}^{\infty}$ is bounded below since $\mathbb{P}(B_i) \geq 0\ \forall i$.

Letting E be the range of $\{\mathbb{P}(B_i)\}_{i=1}^{\infty}$, define $s = \inf E$. Then

$$\mathbb{P}(B_i) \geq s. \qquad (\forall i)$$

For every $\epsilon > 0$, there exists $N \in \mathbb{N}$ such that

$$s + \epsilon > \mathbb{P}(B_N) > s$$

for otherwise $s + \epsilon$ would be a lower bound of E. Since $\{\mathbb{P}(B_i)\}_{i=1}^{\infty}$ decreases, $n \geq N$ implies that

$$s + \epsilon > \mathbb{P}(B_n) \geq s$$

giving that $\{\mathbb{P}(B_i)\}_{i=1}^{\infty}$ converges to s.
It must now be shown that $s = \mathbb{P}(B)$. Assume towards a contradiction that $s \neq \mathbb{P}(B)$. There are two cases: $(i).\mathbb{P}(B) > s$ and $(ii).\mathbb{P}(B) < s$.
If $\mathbb{P}(B) > s$ then that implies $B_{n'} \supset B$ for some $n' \in \mathbb{N}$. But this would mean that $\mathbb{P}(B) > \mathbb{P}(B_{n'})$ which contradicts the assumption that $\{B_n\}$ decreases to B and thus $\mathbb{P}(B) \not> s$.
If $\mathbb{P}(B) < s$ then $\{B_n\} \supset B$ for all n and we again contradict that B_n decreases to B.
Thus it must be the case that $\inf E = \mathbb{P}(B)$ and it finally follows that

$$\lim_{n \to \infty} \mathbb{P}(B_n) = \mathbb{P}(B).$$