Exercise 1. Realize that
\[ E(X) = E(X1(X > 0)) \]
and apply Cauchy-Schwartz inequality:
\[ E(X)^2 = E(X1(X > 0))^2 \leq E(X^2)E((1(X > 0))^2) = E(X^2)P(X > 0). \]

Exercise 2. Start from \( g \) as indicator function (which is easy), and extend to \( g \) as simple functions. Now assume \( g \geq 0 \). Use a sequence of simple functions \( g_n \) to approach \( g \) pointwisely from below, and use MCT on both sides. Then for general \( g \), consider decomposition \( g = g^+ - g^- \) where both functions of the right hand side are non-negative. From \( \int |g(x)|\mu(dx) < \infty \) we may have both \( \int g^+(x)\mu(dx) \) and \( \int g^-(x)\mu(dx) \) finite. So we can do the subtraction and finish the work.

Exercise 3. Consider a 4 element space \( \Omega = \{a, b, c, d\} \) with each element weights with probability \( 1/4 \). Now consider \( A_1 = \{\{a, b\}, \{b, c\}\} \) and \( A_2 = \{\{a, c\}, \{b, d\}\} \). It is easy to check the independence but \( \sigma(A_1) \) contains \( \{a\} \) which is not independent of \( \{b, d\} \in A_2 \).

Exercise 4. First we show \( Y_n \) are distributed as described. First extend \([0, 1]\) to \([0, 2^n]\) and see that for \( \omega * 2^n \in [2k, 2k+1) \) interval, \( Y(\omega) = 1 \). Clearly these \( \omega \) takes up half of the intervals (since intervals like \([2k, 2k+1) \) takes up half portion in \([0, 2^n]\). Therefore \( P(Y_n = 1) = 0.5 \).

Then we show the independence. This is equivalent to show that for a finite collection of integers \( n_1, \ldots, n_k \) and binaries \( i_1, \ldots, i_k \in \{0, 1\} \), we have \( P(Y_{n_1} = i_1, \ldots, Y_{n_k} = i_k) = \prod_j P(Y_{n_j} = i_j) \). Without loss of generality we assume \( n_j \) are monotone increasing, and this can be done by induction.

WLOG, set all \( i_j = 1 \). For \( Y_{n_1} \), find out the intervals such that \( Y_{n_1} = 1 \) and realize that inside each interval, there are half portion that makes \( Y_{n_2} = 1 \). Induct this to the general case and the proof is complete.

Exercise 5. (i) WLOG assume \( E(Y) = 0 \). Assume \( P(Y \neq 0) > 0 \). Then there exists some \( n \) and set \( A \) such that \( Y > 1/n \) on \( A \), with \( P(A) > 0 \). Then
\[ P(X \in A, Y > 1/n) = P(X \in A)P(Y > 1/n) \]
by independence. However \( \{X \in A\} \) implies \( Y > 1/n \) therefore the LHS equals \( P(X \in A) \). Since \( P(A) \neq 0 \), we have \( P(Y > 1/n) = 1 \) which contradicts with \( E(Y) = 0 \).

(ii) Take \( Y = 1(X \geq 1/2) \). The critical point is that by observing \( Z \) only, one piece of \( X \) is missing and that is whether \( X \) is at the left or the right side of \( 1/2 \).