Proposition 9.1 Assume \( X \) is a r.v. on a countable sample space \( S \) (i.e., \( X \) is a function from \( S \) to \( \mathbb{R} \)).

Then \( E(X) = \sum_{w \in S} X(w) P(\{w\}) \).

Note: The sum is a countable sum, hence a series, because \( S \) is countable. Assume \( S \) is finite the first time you read this. \( P(\{w\}) \) is the probability of the particular outcome \( w \).

Proof: Let \( S_i = \{ X = x_i \} \) where \( x_1, \ldots, x_n \) are the distinct possible values of \( X \). (If \( X \) can take on countably many values, the proof is readily adapted by considering an \( \omega \) sequence \( x_1, x_2, \ldots \)).

\[ S_i = \{ w : X(w) = x_i \} = \bigcup_{w \in S_i} (\text{a countable disjoint union}) \]

(ii) \( P(S_i) = \sum_{w \in S_i} P(\{w\}) \) by countable additivity.

By definition:
\[ E(X) = \sum_{i=1}^{n} x_i P(S_i) \]

\[ = \sum_{i=1}^{n} x_i \sum_{w \in S_i} P(\{w\}) \quad \text{by (i)} \]

\[ = \sum_{i=1}^{n} \sum_{w \in S_i} X(w) P(\{w\}) \quad \text{by (i)} \]

The line holds since \( S_1, S_2, \ldots \) partition \( S \) into \( n \) disjoint sets and so summing \( w \) over each \( S_i \) is just summing \( w \) over \( S \).