Wald's Lemma

In class, we proved the following:

Proposition 1. *Suppose $X_1, X_2, \ldots$ are i.i.d., and let $S_n = X_1 + \cdots + X_n$. If T is a $\{0, 1, 2, \ldots\}$-valued random variable independent of $\{X_i : i \geq 1\}$, then*

$$E[S_T] = E[T]E[X_1],$$

as long as the expectations exist.

There is a more general version that is applicable to hitting times of Markov chains. Instead of T independent of $\{X_i\}$, it requires that $E[X_n 1_{\{T \geq n\}}] = E[X_n]E[1_{\{T \geq n\}}]$ for every n. This is the case if, for instance, X_n is independent of the event $\{T \geq n\}$.

Proposition 2. *Suppose $X_1, X_2, \ldots$ are random variables with $E[X_n] = \mu$ for all n, and let $S_n = X_1 + \cdots + X_n$. Suppose that T is a $\{0, 1, 2, \ldots\}$-valued random variable with finite mean and such that*

$$E[X_n 1_{\{T \geq n\}}] = E[X_n]P(T \geq n) \quad \text{for all} \quad n.$$

Then,

$$E[S_T] = \mu E[T].$$

Proof. Write

$$S_T = \sum_{n=1}^{T} X_n = \sum_{n=1}^{\infty} X_n 1_{\{T \geq n\}}.$$

Then,

$$\begin{aligned} E[S_T] &= E\left[\sum_{n=1}^{\infty} X_n 1_{\{T \geq n\}}\right] \\ &= \sum_{n=1}^{\infty} E[X_n 1_{\{T \geq n\}}] \\ &= \sum_{n=1}^{\infty} E[X_n]P(T \geq n) = \mu \sum_{n=1}^{\infty} P(T \geq n) = \mu \sum_{k=0}^{\infty} P(T > k) = \mu E[T]. \end{aligned}$$

□

Since a simple random walk is the sum of i.i.d. random variables, the last proposition can be applied to hitting times, even though they are not independent of the walk. This is given in the following corollary.

Corollary 3. *Suppose $X_1, X_2, \ldots$ are i.i.d. random variables, and let $S_n = X_1 + \cdots + X_n$. For $A \subset \mathbb{R}$, let*

$$\tau_A = \inf\{n \geq 1 : S_n \in A\},$$

and suppose $P(\tau_A < \infty) = 1$. Then,

$$E[S_{\tau_A}] = E[\tau_A]E[X_1],$$

as long as the expectations exist.

Proof. $\{\tau_A \geq n\}$ is the event that the first time the sum is in A happens after time $n-1$. That is,

$$\{\tau_A \geq n\} = \{S_1 \notin A, S_2 \notin A, \ldots, S_{n-1} \notin A\}.$$

Since X_n is independent of $S_1, S_2, \ldots, S_{n-1}$,

$$\begin{aligned} E[X_n 1_{\{\tau_A \geq n\}}] &= E[X_n 1_{\{S_1 \notin A, \ldots, S_{n-1} \notin A\}}] \\ &= E[X_n] E[1_{\{S_1 \notin A, \ldots, S_{n-1} \notin A\}}] \\ &= E[X_n] P(\tau_A \geq n). \end{aligned}$$

Then Proposition 2 says

$$E[S_{\tau_A}] = E[\tau_A] E[X_1].$$

□