

17.3: Stopping Times

Basic Theory

As in the Introduction, we start with a stochastic process $\mathbf{X} = \{X_t : t \in T\}$ on an underlying probability space $(\Omega, \mathcal{F}, \mathbb{P})$, having state space \mathbb{R} , and where the index set T (representing time) is either \mathbb{N} (discrete time) or $[0, \infty)$ (continuous time). Next, we have a filtration $\mathfrak{F} = \{\mathcal{F}_t : t \in T\}$, and we assume that \mathbf{X} is adapted to \mathfrak{F} . So \mathfrak{F} is an increasing family of sub σ -algebras of \mathcal{F} and X_t is measurable with respect to \mathcal{F}_t for $t \in T$. We think of \mathcal{F}_t as the collection of events up to time $t \in T$. We assume that $\mathbb{E}(|X_t|) < \infty$, so that the mean of X_t exists as a real number, for each $t \in T$. Finally, in continuous time where $T = [0, \infty)$, we make the standard assumption that \mathbf{X} is right continuous and has left limits, and that the filtration \mathfrak{F} is right continuous and complete.

Our general goal in this section is to see if some of the important martingale properties are preserved if the deterministic time $t \in T$ is replaced by a (random) stopping time. Recall that a random time τ with values in $T \cup \{\infty\}$ is a *stopping time* relative to \mathfrak{F} if $\{\tau \leq t\} \in \mathcal{F}_t$ for $t \in T$. So a stopping time is a random time that does not require that we see into the future. That is, we can tell if $\tau \leq t$ from the information available at time t . Next recall that the σ -algebra associated with the stopping time τ is

$$\mathcal{F}_\tau = \{A \in \mathcal{F} : A \cap \{\tau \leq t\} \in \mathcal{F}_t \text{ for all } t \in T\} \quad (17.3.1)$$

So \mathcal{F}_τ is the collection of events up to the random time τ just as \mathcal{F}_t is the collection of events up to the deterministic time $t \in T$. In terms of a gambler playing a sequence of games, the time that the gambler decides to stop playing must be a stopping time, and in fact this interpretation is the origin of the name. That is, the time when the gambler decides to stop playing can only depend on the information that the gambler has up to that point in time.

Optional Stopping

The basic martingale equation $\mathbb{E}(X_t | \mathcal{F}_s) = X_s$ for $s, t \in T$ with $s \leq t$ can be generalized by replacing both s and t by bounded stopping times. The result is known as the *Doob's optional stopping theorem* and is named again for Joseph Doob. Suppose that $\mathbf{X} = \{X_t : t \in T\}$ satisfies the basic assumptions above with respect to the filtration $\mathfrak{F} = \{\mathcal{F}_t : t \in T\}$

Suppose that ρ and τ are bounded stopping times relative to \mathfrak{F} with $\rho \leq \tau$.

1. If \mathbf{X} is a martingale relative to \mathfrak{F} then $\mathbb{E}(X_\tau | \mathcal{F}_\rho) = X_\rho$.
2. If \mathbf{X} is a sub-martingale relative to \mathfrak{F} then $\mathbb{E}(X_\tau | \mathcal{F}_\rho) \geq X_\rho$.
3. If \mathbf{X} is a super-martingale relative to \mathfrak{F} then $\mathbb{E}(X_\tau | \mathcal{F}_\rho) \leq X_\rho$.

Proof in discrete time

1. Suppose that $\tau \leq k$ where $k \in \mathbb{N}_+$ and let $A \in \mathcal{F}_\tau$. For $j \in \mathbb{N}$ with $j \leq k$, $A \cap \{\tau = j\} \in \mathcal{F}_j$. Hence by the martingale property,

$$\mathbb{E}(X_k; A \cap \{\tau = j\}) = \mathbb{E}(X_j; A \cap \{\tau = j\}) = \mathbb{E}(X_\tau; A \cap \{\tau = j\}) \quad (17.3.2)$$

Since k is an upper bound on τ , the events $A \cap \{\tau = j\}$ for $j = 0, 1, \dots, k$ partition A , so summing the displayed equation over j gives $\mathbb{E}(X_k; A) = \mathbb{E}(X_\tau; A)$. By definition of conditional expectation, $\mathbb{E}(X_k | \mathcal{F}_\tau) = X_\tau$. But since k is also an upper bound for ρ we also have $\mathbb{E}(X_k | \mathcal{F}_\rho) = X_\rho$. Finally using the tower property we have

$$X_\rho = \mathbb{E}(X_k | \mathcal{F}_\rho) = \mathbb{E}[\mathbb{E}(X_k | \mathcal{F}_\tau) | \mathcal{F}_\rho] = \mathbb{E}[\mathbb{E}(X_\tau | \mathcal{F}_\tau) | \mathcal{F}_\rho] = \mathbb{E}(X_\tau | \mathcal{F}_\rho) \quad (17.3.3)$$

2. If \mathbf{X} is a sub-martingale, then by the Doob decomposition theorem, $X_n = Y_n + Z_n$ for $n \in \mathbb{N}$ where $\mathbf{Y} = \{Y_n : n \in \mathbb{N}\}$ is a martingale relative to \mathfrak{F} and $\mathbf{Z} = \{Z_n : n \in \mathbb{N}\}$ is increasing and is predictable relative to \mathfrak{F} . So

$$\mathbb{E}(X_\tau | \mathcal{F}_\rho) = \mathbb{E}(Y_\tau | \mathcal{F}_\rho) + \mathbb{E}(Z_\tau | \mathcal{F}_\rho) \quad (17.3.4)$$

But $\mathbb{E}(Y_\tau | \mathcal{F}_\rho) = Y_\rho$ by part (a) and since \mathbf{Z} is increasing, $\mathbb{E}(Z_\tau | \mathcal{F}_\rho) \geq \mathbb{E}(Z_\rho | \mathcal{F}_\rho) = Z_\rho$. Hence $\mathbb{E}(X_\tau | \mathcal{F}_\rho) \geq X_\rho$.

3. The proof when \mathbf{X} is a super-martingale is just like (b), except that the process \mathbf{Z} is decreasing.

Proof in continuous time

Suppose that \mathbf{X} is a martingale. We need to show that $\mathbb{E}(X_\tau; A) = \mathbb{E}(X_\rho; A)$ for every $A \in \mathcal{F}_\rho$. Let $\rho_n = \lceil 2^n \rho \rceil / 2^n$ and $\tau_n = \lceil 2^n \tau \rceil / 2^n$ for $n \in \mathbb{N}$. The stopping times ρ_n and τ_n take values in a countable set T_n for each $n \in \mathbb{N}$, and $\rho_n \downarrow \rho$ and

$\tau_n \downarrow \tau$ as $n \rightarrow \infty$. The process $\{X_t : t \in T_n\}$ is a discrete-time martingale for each $n \in \mathbb{N}$. By the right continuity of \mathbf{X} ,

$$X_{\rho_n} \rightarrow X_\rho, \quad X_{\tau_n} \rightarrow X_\tau \text{ as } n \rightarrow \infty \quad (17.3.5)$$

Suppose next that $\tau \leq c$ where $c \in (0, \infty)$ so that $\rho \leq c$ also. Then $\rho_n \leq c+1$ and $\tau_n \leq c+1$ for $n \in \mathbb{N}$ so the discrete stopping times are uniformly bounded. From the discrete version of the theorem, $X_{\rho_n} = \mathbb{E}(X_{c+1} | \mathcal{F}_{\rho_n})$ and $X_{\tau_n} = \mathbb{E}(X_{c+1} | \mathcal{F}_{\tau_n})$ for $n \in \mathbb{N}$. It then follows that the sequences $\{X_{\rho_n} : n \in \mathbb{N}\}$ and $\{X_{\tau_n} : n \in \mathbb{N}\}$ are uniformly integrable and hence $X_{\rho_n} \rightarrow X_\rho$ and $X_{\tau_n} \rightarrow X_\tau$ as $n \rightarrow \infty$ in mean as well as with probability 1. Now let $A \in \mathcal{F}_\rho$. Since $\rho \leq \rho_n$, $\mathcal{F}_\rho \subseteq \mathcal{F}_{\rho_n}$ and so $A \in \mathcal{F}_{\rho_n}$ for each $n \in \mathbb{N}$. By the theorem in discrete time,

$$\mathbb{E}(X_{\tau_n}; A) = \mathbb{E}(X_{\rho_n}; A), \quad n \in \mathbb{N} \quad (17.3.6)$$

Letting $n \rightarrow \infty$ gives $\mathbb{E}(X_\tau; A) = \mathbb{E}(X_\rho; A)$. The proofs in parts (b) and (c) are as in the discrete time.

The assumption that the stopping times are bounded is critical. A [counterexample](#) when this assumption does not hold is given below. Here are a couple of simple corollaries:

Suppose again that ρ and τ are bounded stopping times relative to \mathfrak{F} with $\rho \leq \tau$.

1. If \mathbf{X} is a martingale relative to \mathfrak{F} then $\mathbb{E}(X_\tau) = \mathbb{E}(X_\rho)$.
2. If \mathbf{X} is a sub-martingale relative to \mathfrak{F} then $\mathbb{E}(X_\tau) \geq \mathbb{E}(X_\rho)$.
3. If \mathbf{X} is a super-martingale relative to \mathfrak{F} then $\mathbb{E}(X_\tau) \leq \mathbb{E}(X_\rho)$.

Proof

Recall that $\mathbb{E}(X_\tau) = \mathbb{E}[\mathbb{E}(X_\tau | \mathcal{F}_\rho)]$, so the results are immediate from the optional stopping theorem.

Suppose that τ is a bounded stopping time relative to \mathfrak{F} .

1. If \mathbf{X} is a martingale relative to \mathfrak{F} then $\mathbb{E}(X_\tau) = \mathbb{E}(X_0)$.
2. If \mathbf{X} is a sub-martingale relative to \mathfrak{F} then $\mathbb{E}(X_\tau) \geq \mathbb{E}(X_0)$.
3. If \mathbf{X} is a super-martingale relative to \mathfrak{F} then $\mathbb{E}(X_\tau) \leq \mathbb{E}(X_0)$.

The Stopped Martingale

For our next discussion, we first need to recall how to stop a stochastic process at a stopping time.

Suppose that \mathbf{X} satisfies the assumptions above and that τ is a stopping time relative to the filtration \mathfrak{F} . The *stopped process* $X^\tau = \{X_t^\tau : t \in [0, \infty)\}$ is defined by

$$X_t^\tau = X_{t \wedge \tau}, \quad t \in [0, \infty) \quad (17.3.7)$$

Details

In continuous time, our standard assumptions ensure that \mathbf{X}^τ is a valid stochastic process and is adapted to \mathfrak{F} . That is, X_t^τ is measurable with respect to \mathcal{F}_t for each $t \in [0, \infty)$. Moreover, \mathbf{X}^τ is also right continuous and has left limits.

So $X_t^\tau = X_t$ if $t < \tau$ and $X_t^\tau = X_\tau$ if $t \geq \tau$. In particular, note that $X_0^\tau = X_0$. If X_t is the fortune of a gambler at time $t \in T$, then X_t^τ is the revised fortune at time t when τ is the stopping time of the gambler. Our next result, known as the *elementary stopping theorem*, is that a martingale stopped at a stopping time is still a martingale.

Suppose again that \mathbf{X} satisfies the assumptions above, and that τ is a stopping time relative to \mathfrak{F} .

1. If \mathbf{X} is a martingale relative to \mathfrak{F} then so is \mathbf{X}^τ .
2. If \mathbf{X} is a sub-martingale relative to \mathfrak{F} then so is \mathbf{X}^τ .
3. If \mathbf{X} is a super-martingale relative to \mathfrak{F} then so is \mathbf{X}^τ .

General proof

If $s, t \in T$ with $s \leq t$ then $\tau \wedge s$ and $\tau \wedge t$ are bounded stopping times with $\tau \wedge s \leq \tau \wedge t$. So the results follows immediately from the [optional stopping theorem](#) above.

Special proof in discrete time

In discrete time, there is a simple direct proof using the martingale transform. So suppose that $T = \mathbb{N}$ and define the process $\mathbf{Y} = \{Y_n : n \in \mathbb{N}_+\}$ by

$$Y_n = \mathbf{1}(\tau \geq n) = 1 - \mathbf{1}(\tau \leq n-1), \quad n \in \mathbb{N}_+ \quad (17.3.8)$$

By definition of a stopping time, $\{\tau \leq n-1\} \in \mathcal{F}_{n-1}$ for $n \in \mathbb{N}_+$, so the process \mathbf{Y} is predictable. Of course, \mathbf{Y} is a bounded, nonnegative process also. The transform of \mathbf{X} by \mathbf{Y} is

$$(\mathbf{Y} \cdot \mathbf{X})_n = X_0 + \sum_{k=1}^n Y_k (X_k - X_{k-1}) = X_0 + \sum_{k=1}^n \mathbf{1}(\tau \geq k) (X_k - X_{k-1}), \quad n \in \mathbb{N}_+ \quad (17.3.9)$$

But note that $X_k^\tau - X_{k-1}^\tau = X_k - X_{k-1}$ if $\tau \geq k$ and $X_k^\tau - X_{k-1}^\tau = X_\tau - X_\tau = 0$ if $\tau < k$. That is, $X_k^\tau - X_{k-1}^\tau = \mathbf{1}(\tau \geq k)(X_k - X_{k-1})$. Hence

$$(\mathbf{Y} \cdot \mathbf{X})_n = X_0 + \sum_{k=1}^n (X_k^\tau - X_{k-1}^\tau) = X_0 + X_n^\tau - X_0^\tau = X_n^\tau, \quad n \in \mathbb{N}_+ \quad (17.3.10)$$

But if \mathbf{X} is a martingale (sub-martingale) (super-martingale), then so is the transform $\mathbf{Y} \cdot \mathbf{X} = \mathbf{X}_\tau$.

The elementary stopping theorem is bad news for the gambler playing a sequence of games. If the games are fair or unfavorable, then no stopping time, regardless of how cleverly designed, can help the gambler. Since a stopped martingale is still a martingale, the the mean property holds.

Suppose again that \mathbf{X} satisfies the assumptions above, and that τ is a stopping time relative to \mathfrak{F} . Let $t \in T$.

1. If \mathbf{X} is a martingale relative to \mathfrak{F} then $\mathbb{E}(X_{t \wedge \tau}) = \mathbb{E}(X_0)$
2. If \mathbf{X} is a sub-martingale relative to \mathfrak{F} then $\mathbb{E}(X_{t \wedge \tau}) \geq \mathbb{E}(X_0)$
3. If \mathbf{X} is a super-martingale relative to \mathfrak{F} then $\mathbb{E}(X_{t \wedge \tau}) \leq \mathbb{E}(X_0)$

Optional Stopping in Discrete Time

A simple corollary of the [optional stopping theorem](#) is that if \mathbf{X} is a martingale and τ a bounded stopping time, then $\mathbb{E}(X_\tau) = \mathbb{E}(X_0)$ (with the appropriate inequalities if \mathbf{X} is a sub-martingale or a super-martingale). Our next discussion centers on other conditions which give these results in discrete time. Suppose that $\mathbf{X} = \{X_n : n \in \mathbb{N}\}$ satisfies the basic assumptions above with respect to the filtration $\mathfrak{F} = \{\mathcal{F}_n : n \in \mathbb{N}\}$, and that τ is a stopping time relative to \mathfrak{F} .

Suppose that $|X_n|$ is bounded uniformly in $n \in \mathbb{N}$ and that τ is finite.

4. If \mathbf{X} is a martingale then $\mathbb{E}(X_\tau) = \mathbb{E}(X_0)$.
5. If \mathbf{X} is a sub-martingale then $\mathbb{E}(X_\tau) \geq \mathbb{E}(X_0)$.
6. If \mathbf{X} is a super-martingale then $\mathbb{E}(X_\tau) \leq \mathbb{E}(X_0)$.

Proof

Assume that \mathbf{X} is a super-martingale. The proof for a sub-martingale are similar, and then the results follow immediately for a martingale. The main tool is the [mean property](#) above for the stopped super-martingale:

$$\mathbb{E}(X_{\tau \wedge n}) \leq \mathbb{E}(X_0), \quad n \in \mathbb{N} \quad (17.3.11)$$

Since $\tau < \infty$ with probability 1, $\tau \wedge n \rightarrow \tau$ as $n \rightarrow \infty$, also with probability 1. Since $|X_n|$ is bounded in $n \in T$, it follows from the bounded convergence theorem that $\mathbb{E}(X_{\tau \wedge n}) \rightarrow \mathbb{E}(X_\tau)$ as $n \rightarrow \infty$. Letting $n \rightarrow \infty$ in the displayed equation gives $\mathbb{E}(X_\tau) \leq \mathbb{E}(X_0)$.

Suppose that $|X_{n+1} - X_n|$ is bounded uniformly in $n \in \mathbb{N}$ and that $\mathbb{E}(\tau) < \infty$.

4. If \mathbf{X} is a martingale then $\mathbb{E}(X_\tau) = \mathbb{E}(X_0)$.
5. If \mathbf{X} is a sub-martingale then $\mathbb{E}(X_\tau) \geq \mathbb{E}(X_0)$.
6. If \mathbf{X} is a super-martingale then $\mathbb{E}(X_\tau) \leq \mathbb{E}(X_0)$.

Proof

Assume that \mathbf{X} is a super-martingale. The proofs for a sub-martingale are similar, and then the results follow immediately for a martingale. The main tool once again is the [mean property](#) above for the stopped super-martingale:

$$\mathbb{E}(X_{\tau \wedge n}) \leq \mathbb{E}(X_0), \quad n \in \mathbb{N} \quad (17.3.12)$$

Suppose that $|X_{n+1} - X_n| \leq c$ where $c \in (0, \infty)$. Then

$$|X_{\tau \wedge n} - X_0| = \left| \sum_{k=1}^{\tau \wedge n} (X_k - X_{k-1}) \right| \leq \sum_{k=1}^{\tau \wedge n} |X_k - X_{k-1}| \leq c(\tau \wedge n) \leq c\tau \quad (17.3.13)$$

Hence $|X_{\tau \wedge n}| \leq c\tau + |X_0|$. Since $\mathbb{E}(\tau) < \infty$ we know that $\tau < \infty$ with probability 1, so as before, $\tau \wedge n \rightarrow \tau$ as $n \rightarrow \infty$. Also $\mathbb{E}(c\tau + |X_0|) < \infty$ so by the dominated convergence theorem, $\mathbb{E}(X_{\tau \wedge n}) \rightarrow \mathbb{E}(X_\tau)$ as $n \rightarrow \infty$. So again letting $n \rightarrow \infty$ in the displayed equation gives $\mathbb{E}(X_\tau) \leq \mathbb{E}(X_0)$.

Let's return to our original interpretation of a martingale \mathbf{X} representing the fortune of a gambler playing fair games. The gambler could choose to quit at a random time τ , but τ would have to be a stopping time, based on the gambler's information encoded in the filtration \mathfrak{F} . Under the conditions of the theorem, no such scheme can help the gambler in terms of expected value.

Examples and Applications

The Simple Random Walk

Suppose that $\mathbf{V} = (V_1, V_2, \dots)$ is a sequence of independent, identically distributed random variables with $\mathbb{P}(V_i = 1) = p$ and $\mathbb{P}(V_i = -1) = 1 - p$ for $i \in \mathbb{N}_+$, where $p \in (0, 1)$. Let $\mathbf{X} = (X_0, X_1, X_2, \dots)$ be the partial sum process associated with \mathbf{V} so that

$$X_n = \sum_{i=1}^n V_i, \quad n \in \mathbb{N} \quad (17.3.14)$$

Then \mathbf{X} is the simple random walk with parameter p . In terms of gambling, our gambler plays a sequence of independent and identical games, and on each game, wins €1 with probability p and loses €1 with probability $1 - p$. So X_n is the gambler's total net winnings after n games. We showed in the Introduction that \mathbf{X} is a martingale if $p = \frac{1}{2}$ (the fair case), a sub-martingale if $p > \frac{1}{2}$ (the favorable case), and a super-martingale if $p < \frac{1}{2}$ (the unfair case). Now, for $c \in \mathbb{Z}$, let

$$\tau_c = \inf\{n \in \mathbb{N} : X_n = c\} \quad (17.3.15)$$

where as usual, $\inf(\emptyset) = \infty$. So τ_c is the first time that the gambler's fortune reaches c . What if the gambler simply continues playing until her net winnings is some specified positive number (say €1 000 000)? Is that a workable strategy?

Suppose that $p = \frac{1}{2}$ and that $c \in \mathbb{N}_+$.

1. $\mathbb{P}(\tau_c < \infty) = 1$
2. $\mathbb{E}(X_{\tau_c}) = c \neq 0 = \mathbb{E}(X_0)$
3. $\mathbb{E}(\tau_c) = \infty$

Proof

Parts (a) and (c) hold since \mathbf{X} is a null recurrent Markov chain. Part (b) follows from (a) since trivially $X_{\tau_c} = c$ if $\tau_c < \infty$.

Note that part (b) does not contradict the [optional stopping theorem](#) because of part (c). The strategy of waiting until the net winnings reaches a specified goal c is unsustainable. Suppose now that the gambler plays until the net winnings either falls to a specified negative number (a loss that she can tolerate) or reaches a specified positive number (a goal she hopes to reach).

Suppose again that $p = \frac{1}{2}$. For $a, b \in \mathbb{N}_+$, let $\tau = \tau_{-a} \wedge \tau_b$. Then

1. $\mathbb{E}(\tau) < \infty$
2. $\mathbb{E}(X_\tau) = 0$
3. $\mathbb{P}(\tau_{-a} < \tau_b) = b/(a+b)$

Proof

1. We will let X_0 have an arbitrary value in the set $\{-a, -a+1, \dots, b-1, b\}$, so that we can use Markov chain techniques. Let $m(x) = \mathbb{E}(\tau \mid X_0 = x)$ for x in this set. Conditioning on the first state and using the Markov property we have

$$m(x) = 1 + \frac{1}{2}m(x-1) + \frac{1}{2}m(x+1), \quad x \in \{-a+1, \dots, b-1\} \quad (17.3.16)$$

with boundary conditions $m(-a) = m(b) = 0$. The linear recurrence relation can be solved explicitly, but all that we care about is the fact that the solution is finite.

2. The [optional sampling theorem](#) applies, so $\mathbb{E}(X_\tau) = \mathbb{E}(X_0) = 0$.
3. Let $q = \mathbb{P}(\tau_{-a} < \tau_b)$ so that $1 - q = \mathbb{P}(\tau_b < \tau_{-a})$. By definition, $X_\tau = -a$ if $\tau_{-a} < \tau_b$ and $X_\tau = b$ if $\tau_b < \tau_{-a}$. So from (b), $q(-a) + (1 - q)b = 0$ and therefore $q = b/(a+b)$.

So gambling until the net winnings either falls to $-a$ or reaches b is a workable strategy, but alas has expected value 0. Here's another example that shows that the first version of the optional sampling theorem can fail if the stopping times are not bounded.

Suppose again that $p = \frac{1}{2}$. Let $a, b \in \mathbb{N}_+$ with $a < b$. Then $\tau_a < \tau_b < \infty$ but

$$b = \mathbb{E}(X_{\tau_b} \mid \mathcal{F}_{\tau_a}) \neq X_{\tau_a} = a \quad (17.3.17)$$

Proof

Since $X_0 = 0$, the process \mathbf{X} must reach a before reaching b . As before, $\tau_b < \infty$ but $\mathbb{E}(\tau_b) = \infty$ since \mathbf{X} is a null recurrent Markov chain.

This result does not contradict the [optional stopping theorem](#) since the stopping times are not bounded.

Wald's Equation

Wald's equation, named for Abraham Wald is a formula for the expected value of the sum of a random number of independent, identically distributed random variables. We have considered this before, in our discussion of conditional expected value and our discussion of random samples, but martingale theory leads to a particularly simple and elegant proof.

Suppose that $\mathbf{X} = (X_n : n \in \mathbb{N}_+)$ is a sequence of independent, identically distributed variables with common mean $\mu \in \mathbb{R}$. If N is a stopping time for \mathbf{X} with $\mathbb{E}(N) < \infty$ then

$$\mathbb{E}\left(\sum_{k=1}^N X_k\right) = \mathbb{E}(N)\mu \quad (17.3.18)$$

Proof

Let \mathcal{F} denote the natural filtration associated with \mathbf{X} . Let $c = \mathbb{E}(|X_n|)$, so that by assumption, $c < \infty$. Finally, let

$$Y_n = \sum_{k=1}^n (X_k - \mu) \quad n \in \mathbb{N}_+ \quad (17.3.19)$$

Then $\mathbf{Y} = (Y_n : n \in \mathbb{N}_+)$ is a martingale relative to \mathcal{F} , with mean 0. Note that

$$\mathbb{E}(|Y_{n+1} - Y_n|) = \mathbb{E}(|X_{n+1} - \mu|) \leq c + |\mu|, \quad n \in \mathbb{N}_+ \quad (17.3.20)$$

Hence a discrete version of the [optional stopping theorem](#) applies and we have $\mathbb{E}(Y_N) = 0$. Therefore

$$0 = \mathbb{E}(Y_N) = \mathbb{E}\left[\sum_{k=1}^N (X_k - \mu)\right] = \mathbb{E}\left(\sum_{k=1}^N X_k - N\mu\right) = \mathbb{E}\left(\sum_{k=1}^N X_k\right) - \mathbb{E}(N)\mu \quad (17.3.21)$$

Patterns in Multinomial Trials

Patterns in multinomial trials were studied in the chapter on Renewal Processes. As is often the case, martingales provide a more elegant solution. Suppose that $\mathbf{L} = (L_1, L_2, \dots)$ is a sequence of independent, identically distributed random variables taking values in a finite set S , so that \mathbf{L} is a sequence of multinomial trials. Let f denote the common probability density function so that for a generic trial variable L , we have $f(a) = \mathbb{P}(L = a)$ for $a \in S$. We assume that all outcomes in S are actually possible, so $f(a) > 0$ for $a \in S$.

In this discussion, we interpret S as an *alphabet*, and we write the sequence of variables in concatenation form, $\mathbf{L} = L_1 L_2 \cdots$ rather than standard sequence form. Thus the sequence is an infinite string of letters from our alphabet S . We are interested in the first occurrence of a particular finite substring of letters (that is, a “word” or “pattern”) in the infinite sequence. The following definition will simplify the notation.

If $\mathbf{a} = a_1 a_2 \cdots a_k$ is a word of length $k \in \mathbb{N}_+$ from the alphabet S , define

$$f(\mathbf{a}) = \prod_{i=1}^k f(a_i) \quad (17.3.22)$$

so $f(\mathbf{a})$ is the probability of k consecutive trials producing word \mathbf{a} .

So, fix a word $\mathbf{a} = a_1 a_2 \cdots a_k$ of length $k \in \mathbb{N}_+$ from the alphabet S , and consider the number of trials $N_{\mathbf{a}}$ until \mathbf{a} is completed. Our goal is compute $\nu(\mathbf{a}) = \mathbb{E}(N_{\mathbf{a}})$. We do this by casting the problem in terms of a sequence of gamblers playing fair games and then using the [optional stopping theorem](#) above. So suppose that if a gambler bets $c \in (0, \infty)$ on a letter $a \in S$ on a trial, then the gambler wins $c/f(a)$ if a occurs on that trial and wins 0 otherwise. The expected value of this bet is

$$f(a) \frac{c}{f(a)} - c = 0 \quad (17.3.23)$$

and so the bet is fair. Consider now a gambler with an initial fortune 1. When she starts playing, she bets 1 on a_1 . If she wins, she bet her entire fortune $1/f(a_1)$ on the next trial on a_2 . She continues in this way: as long as she wins, she bets her entire fortune on the next trial on the next letter of the word, until either she loses or completes the word \mathbf{a} . Finally, we consider a sequence of independent gamblers playing this strategy, with gambler i starting on trial i for each $i \in \mathbb{N}_+$.

For a finite word \mathbf{a} from the alphabet S , $\nu(\mathbf{a})$ is the total winnings by all of the players at time $N_{\mathbf{a}}$.

Proof

Let X_n denote the total fortunes of all of the gamblers after trial $n \in \mathbb{N}_+$. Since all of the bets are fair, $\mathbf{X} = \{X_n : n \in \mathbb{N}_+\}$ is a martingale with mean 0. We will show that the conditions in the discrete version of the [optional sampling theorem](#) hold. First, consider disjoint blocks of trials of length k , that is

$$((L_1, L_2, \dots, L_k), (L_{k+1}, L_{k+2}, \dots, L_{2k}), \dots) \quad (17.3.24)$$

Let $M_{\mathbf{a}}$ denote the index of the first such block that forms the letter \mathbf{a} . This variable has the geometric distribution on \mathbb{N}_+ with success parameter $f(\mathbf{a})$ and so in particular, $\mathbb{E}(M_{\mathbf{a}}) = 1/f(\mathbf{a})$. But clearly $N_{\mathbf{a}} \leq kM_{\mathbf{a}}$ so $\nu(\mathbf{a}) < k/f(\mathbf{a}) < \infty$. Next note that all of the gamblers have stopped playing by time N , so clearly $|X_{n+1} - X_n| \leq 1/f(a)$ for $n \in \mathbb{N}_+$. So the optional stopping theorem applies, and hence $\mathbb{E}(X_{N_{\mathbf{a}}}) = 0$. But note that $\nu(\mathbf{a})$ can also be interpreted as the expected amount of money invested by the gamblers (1 unit at each time until the game ends at time $N_{\mathbf{a}}$), and hence this must also be the total winnings at time $N_{\mathbf{a}}$ (which is deterministic).

Given \mathbf{a} , we can compute the total winnings precisely. By definition, trials $N - k + 1, \dots, N$ form the word \mathbf{a} for the first time. Hence for $i \leq N - k$, gambler i loses at some point. Also by definition, gambler $N - k + 1$ wins all of her bets, completes word \mathbf{a} and so collects $1/f(\mathbf{a})$. The complicating factor is that gamblers $N - k + 2, \dots, N$ may or may not have won all of their bets at the point when the game ends. The following exercise illustrates this.

Suppose that \mathbf{L} is a sequence of Bernoulli trials (so $S = \{0, 1\}$) with success probability $p \in (0, 1)$. For each of the following strings, find the expected number of trials needed to complete the string.

1. 001
2. 010

Solution

Let $q = 1 - p$.

1. For the word 001, gambler $N - 2$ wins $\frac{1}{q^2 p}$ on her three bets. Gambler $N - 2$ makes two bets, winning the first but losing the second. Gambler N loses her first (and only) bet. Hence $\nu(001) = \frac{1}{q^2 p}$

- For the word 010, gambler $N - 2$ wins $\frac{1}{q^2 p}$ on her three bets as before. Gambler $N - 1$ loses his first bet. Gambler N wins $1/q$ on his first (and only) bet. So $\nu(010) = \frac{1}{q^2 p} + \frac{1}{q}$

The difference between the two words is that the word in (b) has a *prefix* (a proper string at the beginning of the word) that is also a *suffix* (a proper string at the end of the word). Word \mathbf{a} has no such prefix. Thus we are led naturally to the following dichotomy:

Suppose that \mathbf{a} is a finite word from the alphabet S . If no proper prefix of \mathbf{a} is also a suffix, then \mathbf{a} is *simple*. Otherwise, \mathbf{a} is *compound*.

Here is the main result, which of course is the same as when the problem was solved using renewal theory.

Suppose that \mathbf{a} is a finite word in the alphabet S .

- If \mathbf{a} is simple then $\nu(\mathbf{a}) = 1/f(\mathbf{a})$.
- If \mathbf{a} is compound, then $\nu(\mathbf{a}) = 1/f(\mathbf{a}) + \nu(\mathbf{b})$ where \mathbf{b} is the longest word that is both a prefix and a suffix of \mathbf{a} .

Proof

The ingredients are in place from our previous discussion. Suppose that \mathbf{a} has length $k \in \mathbb{N}_+$.

- If \mathbf{a} is simple, only player $N - k + 1$ wins, and she wins $1/f(\mathbf{a})$.
- Suppose \mathbf{a} is compound and \mathbf{b} is the largest proper prefix-suffix. player $N - k + 1$ wins $1/f(\mathbf{a})$ as always. The winnings of players $N - k + 2, \dots, N$ are the same as the winnings of a new sequence of gamblers playing a new sequence of trials with the goal of reaching word \mathbf{b} .

For a compound word, we can use (b) to reduce the computation to simple words.

Consider Bernoulli trials with success probability $p \in (0, 1)$. Find the expected number of trials until each of the following strings is completed.

- 1011011
- $11 \cdots 1$ (k times)

Solutions

Again, let $q = 1 - p$.

- $\nu(1011011) = \frac{1}{p^5 q^2} + \nu(1011) = \frac{1}{p^5 q^2} + \frac{1}{p^3 q} + \nu(1) = \frac{1}{p^5 q^2} + \frac{1}{p^3 q} + \frac{1}{p}$
- Let $\mathbf{1}_j$ denote a string of j 1s for $j \in \mathbb{N}_+$. If $k \geq 2$ then $\nu(\mathbf{1}_k) = 1/p^k + \nu(\mathbf{1}_{k-1})$. Hence

$$\nu(\mathbf{1}_k) = \sum_{j=1}^k \frac{1}{p^j} \quad (17.3.25)$$

Recall that an *ace-six flat die* is a six-sided die for which faces 1 and 6 have probability $\frac{1}{4}$ each while faces 2, 3, 4, and 5 have probability $\frac{1}{8}$ each. Ace-six flat dice are sometimes used by gamblers to cheat.

Suppose that an ace-six flat die is thrown repeatedly. Find the expected number of throws until the pattern 6165616 occurs.

Solution

From our main theorem,

$$\begin{aligned} \nu(6165616) &= \frac{1}{f(6165616)} + \nu(616) = \frac{1}{f(6165616)} + \frac{1}{f(616)} + \nu(6) \\ &= \frac{1}{f(6165616)} + \frac{1}{f(616)} + \frac{1}{f(6)} = \frac{1}{(1/4)^6(1/8)} + \frac{1}{(1/4)^3} + \frac{1}{1/4} = 32\,836 \end{aligned}$$

Suppose that a monkey types randomly on a keyboard that has the 26 lower-case letter keys and the space key (so 27 keys). Find the expected number of keystrokes until the monkey produces each of the following phrases:

1. it was the best of times
2. to be or not to be

Solution

1. $27^{24} \approx 2.258 \times 10^{34}$
2. $27^5 + 27^{18} \approx 5.815 \times 10^{25}$

The Secretary Problem

The secretary problem was considered in the chapter on Finite Sampling Models. In this discussion we will solve a variation of the problem using martingales. Suppose that there are $n \in \mathbb{N}_+$ candidates for a job, or perhaps potential marriage partners. The candidates arrive sequentially in random order and are interviewed. We measure the quality of each candidate by a number in the interval $[0, 1]$. Our goal is to select the very best candidate, but once a candidate is rejected, she cannot be recalled. Mathematically, our assumptions are that the sequence of candidate variables $\mathbf{X} = (X_1, X_2, \dots, X_n)$ is independent and that each is uniformly distributed on the interval $[0, 1]$ (and so has the *standard uniform distribution*). Our goal is to select a stopping time τ with respect to \mathbf{X} that maximizes $\mathbb{E}(X_\tau)$, the expected value of the chosen candidate. The following sequence will play a critical role as a sequence of thresholds.

Define the sequence $\mathbf{a} = (a_k : k \in \mathbb{N})$ by $a_0 = 0$ and $a_{k+1} = \frac{1}{2}(1 + a_k^2)$ for $k \in \mathbb{N}$. Then

1. $a_k < 1$ for $k \in \mathbb{N}$.
2. $a_k < a_{k+1}$ for $k \in \mathbb{N}$.
3. $a_k \rightarrow 1$ as $k \rightarrow \infty$.
4. If X is uniformly distributed on $[0, 1]$ then $\mathbb{E}(X \vee a_k) = a_{k+1}$ for $k \in \mathbb{N}$.

Proof

1. Note that $a_1 = \frac{1}{2} < 1$. Suppose that $a_k < 1$ for some $k \in \mathbb{N}_+$. Then $a_{k+1} = \frac{1}{2}(1 + a_k^2) < \frac{1}{2}(1 + 1) = 1$.
2. Note that $0 = a_0 < a_1 = \frac{1}{2}$. Suppose that $a_k > a_{k-1}$ for some $k \in \mathbb{N}_+$. Then $a_{k+1} = \frac{1}{2}(1 + a_k^2) > \frac{1}{2}(1 + a_{k-1}^2) = a_k$.
3. Since the sequence is increasing and bounded above, $a_\infty = \lim_{k \rightarrow \infty} a_k$ exists. Taking limits in the recursion relation gives $a_\infty = \frac{1}{2}(1 + a_\infty^2)$ or equivalently $(a_\infty - 1)^2 = 0$.
4. For $k \in \mathbb{N}$,

$$\mathbb{E}(X \vee a_k) = \int_0^1 (x \vee a_k) dx = \int_0^{a_k} a_k dx + \int_{a_k}^1 x dx = \frac{1}{2}(1 + a_k^2) = a_{k+1} \quad (17.3.26)$$

Since $a_0 = 0$, all of the terms of the sequence are in $[0, 1]$ by (a). Approximations of the first 10 terms are

$$(0, 0.5, 0.625, 0.695, 0.742, 0.775, 0.800, 0.820, 0.836, 0.850, 0.861, \dots) \quad (17.3.27)$$

Property (d) gives some indication of why the sequence is important for the secretary problem. At any rate, the next theorem gives the solution. To simplify the notation, let $\mathbb{N}_n = \{0, 1, \dots, n\}$ and $\mathbb{N}_n^+ = \{1, 2, \dots, n\}$.

The stopping time $\tau = \inf \{k \in \mathbb{N}_n^+ : X_k > a_{n-k}\}$ is optimal for the secretary problem with n candidates. The optimal value is $\mathbb{E}(X_\tau) = a_n$.

Proof

Let $\mathcal{F} = \{\mathcal{F}_k : k \in \mathbb{N}_n^+\}$ be the natural filtration of \mathbf{X} , and suppose that ρ is a stopping time for \mathcal{F} . Define $\mathbf{Y} = \{Y_k : k \in \mathbb{N}_n\}$ by $Y_0 = 0$ and $Y_k = X_{\rho \wedge k} \vee a_{n-k}$ for $k \in \mathbb{N}_n^+$. We will show that \mathbf{Y} is a super-martingale with respect to \mathcal{F} . First, on the event $\rho \leq k-1$,

$$\mathbb{E}(Y_k | \mathcal{F}_{k-1}) = \mathbb{E}[(X_\rho \vee a_{n-k}) | \mathcal{F}_{k-1}] = X_\rho \vee a_{n-k} \leq X_\rho \vee a_{n-k+1} = Y_{k-1} \quad (17.3.28)$$

where we have used the fact that $X_\rho \mathbf{1}(\rho \leq k-1)$ is measurable with respect to \mathcal{F}_{k-1} and the fact that the sequence \mathbf{a} is increasing. On the event $\rho > k-1$,

$$\mathbb{E}(Y_k | \mathcal{F}_{k-1}) = \mathbb{E}(X_k \vee a_{n-k} | \mathcal{F}_{k-1}) = \mathbb{E}(X_k \vee a_{n-k}) = a_{n-k+1} \leq Y_{k-1} \quad (17.3.29)$$

where we have used the fact that X_k and \mathcal{F}_{k-1} are independent, and part (d) of the [previous result](#). Since \mathbf{Y} is a super-martingale and ρ is bounded, the optional stopping theorem applies and we have

$$\mathbb{E}(X_\rho) \leq \mathbb{E}(X_\rho \vee a_{n-\rho}) = \mathbb{E}(Y_\rho) \leq \mathbb{E}(Y_0) = a_n \quad (17.3.30)$$

so a_n is an upper bound on the expected value of the candidate chosen by the stopping time ρ .

Next, we will show that in the special case that $\rho = \tau$, the process \mathbf{Y} is a martingale. On the event $\tau \leq k-1$ we have $\mathbb{E}(Y_k | \mathcal{F}_{k-1}) = X_\tau \vee a_{n-k}$ as before. But by definition, $X_\tau \geq a_{n-\tau} \geq a_{n-k+1} \geq a_{n-k}$ so on this event,

$$\mathbb{E}(Y_k | \mathcal{F}_{k-1}) = X_\tau = X_\tau \vee a_{n-k+1} = Y_{k-1} \quad (17.3.31)$$

On the event $\tau > k-1$ we have $\mathbb{E}(Y_k | \mathcal{F}_{k-1}) = a_{n-k+1}$ as before. But on this event, $Y_{k-1} = a_{n-k+1}$. Now since \mathbf{Y} is a martingale and τ is bounded, the optional stopping theorem applies and we have

$$\mathbb{E}(X_\tau) = \mathbb{E}(X_\tau \vee a_{n-\tau}) = \mathbb{E}(Y_\tau) = \mathbb{E}(Y_0) = a_n \quad (17.3.32)$$

Here is a specific example:

For $n = 5$, the decision rule is as follows:

1. Select candidate 1 if $X_1 > 0.742$; otherwise,
2. select candidate 2 if $X_2 > 0.695$; otherwise,
3. select candidate 3 if $X_3 > 0.625$; otherwise,
4. select candidate 4 if $X_4 > 0.5$; otherwise,
5. select candidate 5.

The expected value of our chosen candidate is 0.775.

In our original version of the secretary problem, we could only observe the *relative ranks* of the candidates, and our goal was to maximize the probability of picking the best candidate. With $n = 5$, the optimal strategy is to let the first two candidates go by and then pick the first candidate after that is better than all previous candidates, if she exists. If she does not exist, of course, we must select candidate 5. The probability of picking the best candidate is 0.433

This page titled 17.3: Stopping Times is shared under a CC BY 2.0 license and was authored, remixed, and/or curated by Kyle Siegrist (Random Services) via source content that was edited to the style and standards of the LibreTexts platform.