Math-Stat-491-Fall2014-Notes-V

Hariharan Narayanan

December 7, 2014

Martingales

1 Introduction

Martingales were originally introduced into probability theory as a model for 'fair betting games'. Essentially we bet on events of known probability according to these known values and the payoff is also according to these values. Common sense tells us that in that case, in the long run, we should neither win nor lose. The theory more or less captures this idea but there are paradoxes - for instance if you play until you are ahead you will gain but you might need unbounded resources to reach that stage. However, the value of martingale theory far exceeds the original reason for their coming into being. These notes are intended to atleast partially justify this statement.

2 Preliminaries

The theory of martingales makes repeated use of the notion of conditional expectation. We review the theory briefly.

The conditional expectation of a random variable X given that a random variable $Y = y$, is given by

$$E[X \mid \{Y = y\}] \equiv \int x \times p_{X|Y}(x \mid y)dx.$$

In the discrete case this becomes

$$E[X \mid \{Y = y\}] \equiv \Sigma_i x_i \times p_{X|Y}(x_i \mid y).$$

Therefore we have

$$E[X] = \Sigma_i x_i \times p_X(x_i) = \Sigma_j [\Sigma_i x_i \times p_{X|Y}(x_i \mid y_j)] \times p_Y(y_j).$$

Now the expression $E[X \mid \{Y = y\}]$ defines a function of y. Consider the function $f(\cdot)$ on the sample space defined by

$$f(\omega) \equiv E[X \mid \{Y = Y(\omega)\}].$$

To compute the value of $f(\cdot)$ on a point ω in the sample space, we first compute $Y(\omega)$, then compute the expectation $E[X \mid \{Y = Y(\omega)\}]$. This random variable $f(\cdot)$ is usually denoted $E[X \mid Y]$. Similarly, the random variable $(E[X \mid Y])^2$ is simply the square of the random variable $E[X \mid Y]$.

Note that, while $X \mid \{Y = y\}$ is a random variable whose sample space is the set of all ω which map to y under the random variable Y, the expression $X \mid Y$ does not have a predetermined meaning.

Next, what meaning should we assign $Var[X \mid Y]$? Formally, $Var[Z] \equiv E[Z^2] - (E[Z])^2$. So we define

$$Var[X \mid Y] \equiv E[X^2 \mid Y] - (E[X \mid Y])^2.$$

Thus $Var[X \mid Y]$ is the random variable given by the expression on the right hand side.

We will now prove the *law of total expectation:*
$E[E[X \mid Y]]$ of the random variable $E[X \mid Y]$ is equal to $E[X]$. In the discrete case, we could first partition the sample space into preimages under Y of the values $Y(\omega) = y_i$, find the probability associated with each set, multiply it by $E[X \mid \{Y = y_i\}]$ and then compute the sum of all such products. Thus we get

$$E[E[X \mid Y]] = \Sigma_i E[X \mid \{Y = y_i\}] \times p_Y(y_i).$$

Hence

$$E[E[X \mid Y]] = \Sigma_i \Sigma_j x_j p_{X|Y}(x_j \mid y_i) \times p_Y(y_i).$$

This, by using definition of conditional probability becomes

$$E[E[X \mid Y]] = \Sigma_i \Sigma_j x_j p_{XY}(x_j, y_i).$$

Interchanging summation we get

$$E[E[X \mid Y]] = \Sigma_j \Sigma_i x_j p_{XY}(x_j, y_i).$$

We can rewrite this as

$$E[E[X \mid Y]] = \Sigma_j x_j [\Sigma_i p_{XY}(x_j, y_i)],$$

i.e., as

$$E[E[X \mid Y]] = \Sigma_j x_j p_X(x_j)] = E[X].$$

In martingale theory we also come across expressions of the kind $E[E[X \mid Y, Z] \mid Z]$ and a useful result is a form of the law of total expectation:

$$E[X|Z] = E[E[X \mid Y, Z] \mid Z].$$

We will call this the *second* form of the law of total expectation.

To prove this evaluate both sides for some value of Z, say $Z = z_1$. LHS $= E[X|Z = z_1]$. Now $(X \mid Z = z_1)$ is a random variable G defined on the subset A, say, of the sample space where $Z(\omega) = z_1$. If originally a subset B of A had a measure $\mu(B)$ now it has the measure $\mu(B)/\mu(A)$, on this new sample space. The RHS $= E[E[(X \mid Z = z_1) \mid Y]] = E[E[G \mid Y]]$. By the law of total expectation, $E[G] = E[E[G \mid Y]]$, so that for $Z = z_1$, both sides are equal, proving the required result.

Another way of looking at this result is as follows. Let us call the event corresponding to $Y = y_1, Z = z_1$, in the sample space, $A(y_1, z_1)$. The event $A(\cdot, z_1)$, corresponding to $Z = z_1$ is $\bigcup_{y_i} A(y_i, z_1)$, the union being that of pairwise disjoint subsets. Let us call the event corresponding to $X = x_i, Z = z_1$ in the sample space, $B(x_i, z_1)$. Computing $E[X \mid Z = z_1] = \Sigma_{x_i} x_i p(x_i \mid z_1) = \Sigma_{x_i} x_i p(x_i, z_1)/p(z_1)$, is in words, computing the expectation over each $B(x_i, z_1)$, and then summing it over all X_i. We could compute instead $\Sigma_{x_i}\Sigma_{y_j} x_i p(x_i \mid y_j, z_1)p(y_j) = \Sigma_{x_i}\Sigma_{y_j} x_i p(x_i, y_j, z_1)/p(z_1)$. Here we are breaking $B(x_i, z_1)$ into the family of disjoint subsets $B(x_i, z_1) \cap A(y_j, z_1)$, computing the expectation over each of these smaller sets according to probability $p(x_i \mid y_j, z_1)$, multiplying by $p(y_j)$ and then summing the values. This latter computation corresponds to the RHS of the equation,

$$E[X|Z = z_1] = E[E[X \mid Y, Z = z_1] \mid Z = z_1].$$

2.1 Cautionary Remark

Martingale theory is about 'expectations'. For reasons of easy readability we will not state conditions such as 'if X is integrable '. Such a condition essentially means $E[|X|] < \infty$. All the results, in these notes on martingales, may be taken to be valid, unless otherwise stated, only when the expectation operation yields a finite value on $|f(X_1, \cdots, X_n)|$, where X_i are the random variables and $f(\cdot)$, the function under consideration.

3 Definition and elementary properties of martingales

Let $\mathcal{S} \equiv X_0, \cdots, X_n, \cdots$ be a sequence of random variables. A second sequence of random variables $\mathcal{M} \equiv M_0, \cdots, M_n, \cdots$ is said to be a *martingale* with respect to $\mathcal{S}$, if $E[M_{n+1} \mid X_0, \cdots, X_n] = M_n$. We will replace the sequence $X_0, \cdots, X_n$, by $\mathcal{F}_n$ for short so that the definition reads

$$E[M_{n+1} \mid \mathcal{F}_n] = M_n.$$

1. (**Expectation remains invariant with time**)
   Applying 'expectation' operation on both sides of the defining equation of a martingale and using the law of total expectation, we have

   $$E[M_{n+1}] = E[E[M_{n+1} \mid \mathcal{F}_n]] = E[M_n].$$

2. (**In the martingale definition M_{n+k} can replace M_{n+1}**)
   Next, we note that $\mathcal{F}_{n+k} = \mathcal{F}_n, X_{n+1}, \cdots, X_{n+k}$.Using the *second* form of the law of total expectation, we have $E[M_{n+k} \mid \mathcal{F}_n] = E[E[M_{n+k} \mid \mathcal{F}_{n+1}] \mid \mathcal{F}_n]$. Now $E[M_{m+r} \mid \mathcal{F}_m] = M_m, r = 1$. Suppose it is true that $E[M_{m+r} \mid \mathcal{F}_m] = M_m, r < k, \forall m$, then we must have $E[M_{n+k} \mid \mathcal{F}_{n+1}] = M_{n+1}$. We therefore have,

   $$E[M_{n+k} \mid \mathcal{F}_n] = E[E[M_{n+k} \mid \mathcal{F}_{n+1}] \mid \mathcal{F}_n] = E[M_{n+1} \mid \mathcal{F}_n] = M_n.$$

   Further, by the law of total expectation

   $$E[M_{n+k}] = E[E[M_{n+k} \mid \mathcal{F}_n]] = E[M_n].$$

3. (**Orthogonality of martingale increments**)
   Let $j \leq k \leq l$. We then have

   $$E[M_j(M_l - M_k)] = E[E[M_j(M_l - M_k)]|\mathcal{F}_j]] = E[M_j E[(M_l - M_k)|\mathcal{F}_j]] = E[M_j(M_j - M_j)] = 0.$$

Next let $i \leq j \leq k \leq l$. We then have

$$E[(M_j - M_i)(M_l - M_k)] = E[E[(M_j - M_i)(M_l - M_k)]|\mathcal{F}_j]] = E[(M_j - M_i)E[(M_l - M_k)|\mathcal{F}_j]]$$

$$= E[(M_j - M_i)(M_j - M_j)] = 0.$$

Note that this happens even though $(M_j - M_i), (M_l - M_k)$, are not independent random variables.

4. (**Variance consequence of orthogonality**)

$$Var[M_{m+k} - M_m] = E[(M_{m+k} - M_m)^2] - (E[M_{m+k} - M_m])^2.$$

$$E[(M_{m+k} - M_m)^2] = E[(M_{m+k})^2 + (M_m)^2] - 2E[(M_{m+k})(M_m)]$$

$$= E[(M_{m+k})^2 + (M_m)^2] - 2E[(M_m)(M_m)] = E[(M_{m+k})^2] + E[(M_m)^2] - 2E[(M_m)^2]$$

$$= E[(M_{m+k})^2] - E[(M_m)^2].$$

$$(E[M_{m+k} - M_m])^2 = (E[M_{m+k}] - E[M_m])^2 = (E[M_{m+k}])^2 + (E[M_m])^2 - 2(E[M_{m+k}])(E[M_m])$$

$$= (E[M_{m+k}])^2 + (E[M_m])^2 - 2(E[M_m])(E[M_m]) = (E[M_{m+k}])^2 - (E[M_m])^2.$$

So

$$Var[M_{m+k} - M_m] = Var[M_{m+k}] - Var[M_m].$$

Next since $Var[M_2 - M_1] = Var[M_2] - Var[M_1]$, it follows that $Var[M_2] = Var[M_1] + Var[M_2 - M_1]$. Similarly,

$$Var[M_n] = Var[M_1] + \Sigma_2^n Var[M_i - M_{i-1}].$$

4 Examples of Martingales

(a) (**Generalized random walk**)

Let $\mathcal{S} \equiv X_0, \cdots, X_n, \cdots$ be a sequence of independent random variables with common mean μ. Let $M_n \equiv \Sigma_0^n X_i - n\mu$. As before, let $\mathcal{F}_n$ denote $X_0, \cdots, X_n$. We then have,

$$E[M_{n+1} \mid \mathcal{F}_n] = E[M_n + X_{n+1} - \mu \mid \mathcal{F}_n] = M_n + E[X_{n+1} - \mu \mid \mathcal{F}_n] = M_n,$$

making M_n a martingale with respect to $\mathcal{S}$.

(b) (**Products of independent random variables**)

Let $\mathcal{S} \equiv X_0, \cdots, X_n, \cdots$ be a sequence of independent random variables with common mean μ, with $\mathcal{F}_n$ defined as before. Let $M_n \equiv (\mu)^{-n}\Pi_0^n X_i$. Then

$$E[M_{n+1} \mid \mathcal{F}_n] = E[(\mu)^{-(n+1)}\Pi_0^{n+1} X_i \mid \mathcal{F}_n]$$

$$= (E[(\mu)^{-1}X_{n+1} \mid \mathcal{F}_n])(\mu)^{-n}\Pi_0^n X_i = (\mu)^{-n}\Pi_0^n X_i = M_n.$$

(c) (**Branching processes**)

We know that in the case of these processes we have $\mu X_n = E[X_{n+1} \mid X_n] = E[X_{n+1} \mid \mathcal{F}_n]$. If we take $M_n \equiv \mu^{-n} X_n$, then we find

$$E[M_{n+1} \mid \mathcal{F}_n] = E[\mu^{-(n+1)} X_{n+1} \mid X_n] = \mu^{-n} X_n = M_n.$$

(d) (**Martingales for Markov chains**)

Suppose $\mathcal{S} \equiv X_0, \cdots, X_n, \cdots$ is a Markov chain. One way of associating a fair game with the Markov chain is to bet on the next state when the chain is at a state i. The transition matrix would be available to the bettor. The pay off would be $1/p(i,j), p(i,j) > 0$, if one bets 1 dollar on j. The expected gain would be 0, since the game is fair. One could define a martingale M_n for this situation as a function $f(i,n)$ of the present state i and the time n, i.e., $M_n \equiv f(i,n)$. In order that M_n becomes a martingale we need $E[M_{n+1} \mid \mathcal{F}_n] = M_n$. Now $E[M_{n+1} \mid \mathcal{F}_n] = \Sigma_j p(i,j) f(j, n+1)$, and $M_n = f(i,n)$. Therefore we must have $f(i,n) = \Sigma_j p(i,j) f(j,n+1)$, as a necessary condition for $M_n \equiv f(i,n)$ to be a martingale.

On the other hand if the function $f(i,n)$ satisfies $f(i,n) = \Sigma_j p(i,j) f(j,n+1)$, the above argument shows that $M_n \equiv f(i,n)$ is a martingale. Thus the condition

$$f(i,n) = \Sigma_j p(i,j) f(j,n+1), \qquad (*)$$

is necessary and sufficient for $M_n \equiv f(i,n)$ to be a martingale on the Markov chain. We give below a couple of instances of such functions.

i. Consider the 'gamber's ruin' Markov chain with $p(i,i+1) = p, p(i,i-1) = 1-p$. It can be verified that $f(i,n) \equiv ((1-p)/p)^i$ satisfies the condition (*) above so that $M_n \equiv f(i,n)$ is a martingale.

ii. When $p = 1/2$ in the gambler's ruin example another possible $f(i,n)$ is the function $i^2 - n$.

5 Optional stopping theorem

The optional stopping theorem is an important result in martingale theory. We state this result under different hypotheses in this section. However the proof may be skipped at a first reading. For this course proofs may be treated as 'optional'. The theorem needs the definition of 'stopping time' before it can be stated. Below we define stopping time for a stochastic process $\mathcal{S} \equiv X_0, \cdots, X_n, \cdots$. We distinguish between a random time and a stopping time.

The positive integer valued, possibly infinite, random variable T, is said to be a *random time* for $\mathcal{S}$, if the event $\{T = n\}$ is determined by the random variables $X_0, \cdots, X_n$, i.e., if we know the values of $X_0, \cdots, X_n$, we can say whether $\{T = n\}$ is true or not. If $Pr\{T < \infty\} = 1$, then the random time N is said to be a *stopping time*.

The stopping theorem essentially states that $E[M_T] = M_0$ under certain conditions even if T is unbounded. It is intuitively clear when T is bounded but most applications correspond to the unbounded situation. Hence the effort put in to tackle that case.

We begin with two simple lemmas whose statements can be regarded as obvious but are still being proved to indicate the steps involved in the proof. The first lemma states that under the condition that the stopping time is equal to k the expected value of a martingale for $n \geq k$ is the same as its expected value at k. Note that in this lemma $1_{T=k}$ can be replaced by any random variable 1_A, as long as the event A is fully determined by $\mathcal{F}_k$.

Lemma 5.1. *Let M_n be a martingale, and T, a stopping time, with respect to $\mathcal{F}_n$.*

Then for all $n \geq k$,

$$E[M_n 1_{T=k}] = E[M_k 1_{T=k}].$$

Proof: We have, using the second form of law of total expectation and the fact that the value of the random variable $1_{T=k}$ is fully determined by the value of $\mathcal{F}_k$,

$$E[M_n 1_{T=k}] = E[E[M_n 1_{T=k} \mid \mathcal{F}_k]] = E[1_{T=k} E[M_n \mid \mathcal{F}_k]] = E[1_{T=k} M_k].$$

The next lemma handles the case where the stopping time cannot exceed n. Note that $M_{T\wedge n}$ refers to a stopping time $T' = min(T, n)$. So even if T is unbounded, T' is bounded.

Lemma 5.2. *Let M_n be a martingale, and T, a stopping time, with respect to $\mathcal{F}_n$.*

$$For\ all\ n = 1, 2\cdots, E[M_0] = E[M_{T\wedge n}] = E[M_n].$$

Proof:

$$E[M_{T\wedge n}] = \Sigma_{k=0}^{n-1} E[M_T 1_{T=k}] + E[M_n 1_{T\geq n}]$$

$$= \Sigma_{k=0}^{n-1} E[M_k 1_{T=k}] + E[M_n 1_{T\geq n}]$$

$$= \Sigma_{k=0}^{n-1} E[M_n 1_{T=k}] + E[M_n 1_{T\geq n}]$$

$$= E[M_n].$$

The next lemma is a technical result needed in the proof of the optional stopping theorem for dominated (i.e., bounded by a random variable with finite expectation) martingales.

Lemma 5.3. *Let Z be an arbitrary random variable satisfying $E[|Z|] < \infty$, and let T be an integer valued random variable such that $Pr\{T < \infty\} = 1$. Then $\lim_{n\to\infty} E[Z 1_{T\leq n}] = E[Z]$ and $\lim_{n\to\infty} E[Z 1_{T>n}] = 0$.*

Proof: We have

$$E[|Z|] \geq E[|Z| 1_{T\leq n}] = \Sigma_{k=0}^{n} E[|Z| \mid T = k] Pr\{T = k\}.$$

If we let $n \to \infty$, since $Pr\{T < \infty\} = 1$, we get

$$\Sigma_{k=0}^{\infty} E[|Z| \mid T = k] Pr\{T = k\} = E[|Z|].$$

Hence

$$\lim_{n\to\infty} E[|Z| 1_{T\leq n}] = E[|Z|]\ and\ \lim_{n\to\infty} E[|Z| 1_{T>n}] = 0.$$

Next we have

$$0 \leq |E[Z] - E[Z 1_{T\leq n}| \leq |E[Z 1_{T>n}]| \leq E[|Z| 1_{T>n}]$$

and

$$\lim_{n\to\infty} E[|Z|1_{T>n}] = 0.$$

So $\lim_{n\to\infty} E[Z1_{T\le n}] = E[Z]$ and $\lim_{n\to\infty} E[Z1_{T>n}] = 0$.

We will now prove the Optional Stopping Theorem for three important situations. In each case the final conclusion is $E[M_T] = E[M_0]$. The first is for the important case where the martingale is dominated by a random variable whose expectation is finite. We use this to prove the second version which addresses an important special and commonly occurring situation and the third version which is usually called the Optional Stopping Theorem..

Theorem 5.1. (Optional Stopping Theorem for dominated martingales)
Let M_n be a martingale with respect to $\mathcal{F}_n$, and let T be a stopping time, i.e., $Pr\{T<\infty\} = 1$. Let Z be a random variable such that $|M_{T\wedge n}| < Z\ \forall n$ and $E[Z] < \infty$. Then $E[M_T] = E[M_0]$.

Proof: We have, since $Pr\{T<\infty\} = 1$,

$$M_T = \Sigma_{k=0}^{\infty} M_k 1_{T=k} = \Sigma_{k=0}^{\infty} M_{T\wedge k} 1_{T=k}.$$

But $|M_{T\wedge n}| < Z$ so that

$$M_T = \Sigma_{k=0}^{\infty} M_{T\wedge k} 1_{T=k} \le \Sigma_{k=0}^{\infty} Z 1_{T=k} \le Z.$$

Hence,

$$|M_T| \le \Sigma_{k=0}^{\infty} |M_{T\wedge k}| 1_{T=k} \le \Sigma_{k=0}^{\infty} Z 1_{T=k} \le Z.$$

Therefore, $E[|M_T|] \le E[Z] < \infty$ and $E[M_T] \le E[|M_T|] < \infty$. Next we have

$$|E[M_{T\wedge n}] - E[M_T]| \le E[|(M_{T\wedge n} - M_T)|] = E[|(M_{T\wedge n} - M_T)| 1_{T>n}] \le 2E[Z1_{T>n}].$$

(Here we have used the fact that $M_{T\wedge n}1_{T\le n} = M_T 1_{T\le n}$.) We know by Lemma **??**, $\lim_{n\to\infty} E[Z1_{T>n}] = 0$. Hence, $\lim_{n\to\infty} E[M_{T\wedge n}] = E[M_T]$. By Lemma **??**, it follows that $E[M_T] = E[M_0]$.

Theorem 5.2. (Optional Stopping Theorem 2)
Let $\mathcal{S} \equiv X_0, \cdots, X_n, \cdots$ and let M_n be a martingale for $\mathcal{S}$. Let T be a random time for the process $\mathcal{S}$. If

1. *T is bounded or*

2. *$E[T] < \infty$ and there is a finite M such that $E[|M_{n+1} - M_n| \mid \mathcal{F}_n] < M$,*

then $E[M_T] = E[M_0]$.

Proof:

1. T bounded case is proved in Lemma **??**.

2. Define $W_0 = |M_0|$ and $W_n = |M_n - M_{n-1}|, n = 1, 2, \cdots$ and let $Z = W_0 + \cdots + W_T$.
   Then $Z \ge |M_0 + (M_1 - M_0) + \cdots (M_T - M_{T-1})| = |M_T|$ and

$$E[Z] = \Sigma_{n=0}^{\infty} \Sigma_{k=0}^{n} E[W_k 1_{T=n}]$$

$$= \Sigma_{k=0}^{\infty} \Sigma_{n=k}^{\infty} E[W_k 1_{T=n}]$$

$$= \Sigma_{k=0}^{\infty} E[W_k 1_{T \geq k}].$$

Now the random variable $1_{T \geq k}$ is fully determined by, i.e., is a function of, $\mathcal{F}_{k-1}$. We are given that $E[W_k \mid \mathcal{F}_{k-1}] \leq M$ hold if $k - 1 < T$. Hence

$$= \Sigma_{k=0}^{\infty} E[W_k 1_{T \geq k}] = \Sigma_{k=0}^{\infty} E[E[W_k 1_{T \geq k} \mid \mathcal{F}_{k-1}]]$$

$$= \Sigma_{k=0}^{\infty} E[1_{T \geq k} E[W_k \mid \mathcal{F}_{k-1}]] \leq M \Sigma_{k=0}^{\infty} Pr\{T \geq k\} \leq M(1 + E[T]) < \infty.$$

(Here we have used the fact that $E[T] = \Sigma_{k=1}^{\infty} Pr\{T \geq k\}$, and $Pr\{T \geq 0\} = 1$.) We thus have $E[Z] < \infty$. Since $|M_{T \wedge n}| \leq |M_T| < Z \; \forall n$, we can use Theorem **??** and conclude that $E[M_T] = E[M_0]$.

Theorem 5.3. (Optional Stopping Theorem Main)
Let $\mathcal{S} \equiv X_0, \cdots, X_n, \cdots$, let M_n be a martingale and let T be a stopping time for $\mathcal{S}$. If

1. $Pr\{T < \infty\} = 1$,
2. $E[|M_T|] < \infty$.
3. $\lim_{n \to \infty} E[M_n 1_{T > n}] = 0$,

then

$$E[M_T] = E[M_0].$$

Proof: We have for all n,

$$E[M_T] = E[M_T 1_{T \leq n}] + E[M_T 1_{T > n}] = E[M_{T \wedge n}] - E[M_n 1_{T > n}] + E[M_T 1_{T > n}].$$

Now $E[M_{T \wedge n}] = E[M_0]$, by Lemma **??**. We are given that $\lim_{n \to \infty} E[M_n 1_{T > n}] = 0$. By Lemma **??**, since $M_T \leq |M_T|$, and $E[|M_T|] < \infty$, we conclude that $\lim_{n \to \infty} E[M_T 1_{T > n}] = 0$. Thus $E[M_T] = \lim_{n \to \infty} E[M_{T \wedge n}] = E[M_0]$.

6 Applications of OST

The optional stopping theorem is a subtle result which needs some discussion as to its limitations and its power. Basically it says that if you stop a stochastic process according to a criterion based only on the past then the expected value of a martingale at stopping is the same as at starting. Here is an example of a wrong use of OST.

Example: Consider the unbounded 'gambler's ruin' Markov chain on $0, \cdots, +\infty$. We know that position at time n is a martingale M_n. Suppose you start at position i, then your expected position $E[M_n]$ will remain i. But let us use a stopping rule 'stop as soon as you reach 0.' After stopping the expected position is 0, which need not be i.
What is going wrong here?
We cannot apply OST because in this case $E[T]$ is not finite.

On the other hand in the 'gambler's ruin' Markov chain on $0, \cdots, K$, if we use the stopping rule 'stop as soon as you reach 0 or K.', we can show that $E[T]$ is bounded, for instance solving the set of linear equations for hitting time of Markov chains and getting a finite value. Further $|M_n| < K+1$. So we can apply OST here. The expected value $E[T]$, at stopping equals that at starting, the position $i \leq K$. The expected value at stopping is $pK + (1-p) \times 0 = pK$, where p is the probability of hitting K starting from i. We thus have $pK = i$, and therefore $p = i/K$. This is an indirect but simple way of computing p.

6.1 Wald's Equation

Theorem 6.1. (Wald's equation)
Let $X_i, i \geq 1$, be i.i.d random variables with $E[X] = E[X_i] < \infty$. Let T be a stopping time for $X_1, \cdots$, with $E[T] < \infty$. Then, $E[\Sigma_1^T X_i] = E[T]E[X]$.

Proof: Let $\mu = E[X]$. $M_n \equiv \Sigma_1^n (X_i - \mu)$, is a martingale since $E[M_{n+1} \mid \mathcal{F}_n] = E[M_n + [X_{n+1} - \mu] \mid \mathcal{F}_n] = M_n + 0$. If OST is applicable we have, $E[M_T] = E[\Sigma_1^T (X_i - \mu)] = E[\Sigma_1^T (X_i) - T\mu] = E[\Sigma_1^T (X_i)] - E[T\mu]$. If $E[M_T] = 0$, we get $E[\Sigma_1^T (X_i)] = E[T\mu]$. This would happen if OST is applicable since then $E[M_T] = E[M_1] = E[(X_1 - \mu)] = 0$.

In order to show OST is applicable, since it is given that $E[T] < \infty$, we need merely show that $E[|M_{n+1} - M_n| \mid \mathcal{F}_n]$ is bounded. We have $E[|M_{n+1} - M_n| \mid \mathcal{F}_n] = E[(X_{n+1} - \mu) \mid \mathcal{F}_n] \leq E[|X|] + \mu < M$, say, since $E[|X|]$ is bounded.

This completes the proof of Wald's equation.

Remark: Suppose T is a random variable independent of the X_i, then Wald's theorem is obviously true since $E[\Sigma_1^T (X_i)] = E[E[\Sigma_1^T (X_i) \mid T]] = E[E[TX_i \mid T]] = E[TX_i]$, and by independence of T, X_i $E[TX_i] = E[T]E[X_i]$. The unusual feature of Wald's equation is that independence of T and the X_i is not required and something much weaker is adequate.

6.2 Expected length of sequences with a known end subsequence

Let $X_0, X_1, \cdots$ be a sequence of i.i.d. random variables. The problem is to find the expected length of a sequence of values that these random variables take till the first time a specified subsequence occurs.

One way of solving this problem is to pick a 'suitable' martingale through a betting scheme assuming the game is fair. It is of course necessary that the martingale should capture the essentials of the problem, i.e, the end sequence.

We illustrate the association of a betting scheme with a martingale as follows:
You bet \$1 on an event A and the event occurs you get a reward of \$$1/p(A)$. If you bet \$1 at time 0, on an event that is expected to happen at time n, we can take $M'_0 = M'_1 = \cdots = M'_{n-1} = 1$; M'_n has value $1/p(A)$ on event A of the sample space and zero on the complement. Then $E[M'_{n+1} \mid X_0, \cdots, X_n, M'_0] = 1 = M'_n$. The expected gain in this transaction is $1/p(A) \times p(A) + 0 \times (1 - p(A)) - 1 = 0$. This is what the expression $E[M'_{n+1} \mid\mid \mathcal{F}_n] - M'_n = 0$ means. The other and more convenient way of defining the martingale for our problem is to take the profit of the gambler or the casino as the martingale. This new martingale M_n is related to the above martingale by $M_n = M'_n - 1$.

Let us now work with a specific example. Let $X_i, i = 0, 1, 2, \cdots$ take values $0, 1, 2, \cdots$ with probabilites $p_0, p_1, \cdots$, respectively. Let us find the expected length of a sequence which ends with $0, 2, 0$.

We will build our martingale M_n as a sum of such elementary martingales, which may be considered as individual bettors profits. Let m_n^k be the martingale defined as follows:
$m_i^k = 0, i = 0, 1, \ldots, k-1; k \geq 1$.
m_k^k takes value $1/p_0 - 1$ on the event corresponding to $X_k = 0$ and -1 if the event does not occur.
m_{k+1}^k takes value $1/p_0 \times 1/p_2 - 1$, if $X_k = 0, X_{k+1} = 2$ and -1 if the event does not occur.
m_{k+2}^k takes value $1/p_0 \times 1/p_2 \times 1/p_0 - 1$, if $X_k = 0, X_{k+1} = 2, X_{k+2} = 0$ and -1 if the event does not occur.
$m_i^k = m_{k+2}^k, i > k+2$.
We now define a composite martingale by

$$M_n = m_n^0 + \cdots m_n^n.$$

In the terminology introduced in class, the money that the i^{th} bettor transfers to the Casino is m_n^i, and the money that the casino receives totally is the Martingale M_n. Now we use the stopping rule 'stop the process when you reach $0, 2, 0.'$ What would $E[M_T]$ be?

In order to apply Theorem **??** we need $E[T] < \infty$ and $E[|M_{n+1} - M_n| \mid \mathcal{F}_n] < M$. The former can be shown by Markov chain arguments and the latter follows from the way we have defined the martingale.

At stopping time all bettors have lost \$1 and some have gained, i.e, $M_T = \Sigma_{i=0}^{i=T} m_T^i = -1 \times T + terms\ corresponding\ to\ payoff$. The $(T-2)^{th}$ bettor has pay off $1/p_0 \times 1/p_2 \times 1/p_0$, the $(T)^{th}$ bettor has pay off $1/p_0$, and none of the others have any payoff. Therefore $E[M_T] = E[T] - 1/p_0 - 1/p_0 \times 1/p_2 \times 1/p_0$, i.e., $E[T] = 1/p_0 + 1/p_0 \times 1/p_2 \times 1/p_0$.