We recall the optional stopping time result from the last lecture:

**Theorem 19.1** Let \( Y_n \) be a submartingale adapted to \( \mathcal{F}_n \), and suppose that \( T \) is a stopping time such that \( P(T \leq N) = 1 \), for some nonnegative integer \( N \). Then \( E(Y_0) \leq E(Y_T) \leq E(Y_N) \).

See Theorem (4.4.1) in the text [1, page 249] for a proof. Note that the proof can easily be modified to come up with the expected analogues for supermartingales and martingales, in which case the \( \leq \) inequalities in the conclusion above become \( \geq \) and \( = \), respectively.

**Example.** If \( S_n \) is a simple random walk with \( S_0 = 1 \), and \( N = \inf \{ n : S_n = 0 \} \) a stopping time, then it is clear that \( E S_0 = 1 \), while \( E S_N = 0 \) since we are only counting \( S_n \) when \( N = n \), i.e., \( S_n = 0 \). However, \( S_n \) is a martingale \( (E(S_{n+1} | \mathcal{F}_n) = E(S_{n+1}) + S_n = S_n) \).

This example shows that the result need not hold for unbounded stopping times: while \( P(N < \infty) = 1 \), \( E N = 1 \) and Wald’s equation fails. It is possible to obtain the desired analogue for unbounded stopping times if the \( X_n \) are uniformly integrable, which will be dealt with in a future lecture (see sections 4.5 and 4.7 in [1]).

This theorem may be applied to obtain some maximal inequalities.

**Theorem 19.2 (Doob’s maximal inequality)** Let \( Y_n \) be a submartingale with \( Y_N \geq 0 \) (often \( Y_m = \phi(X_n) \), where \( X_n \) is a martingale, \( \phi \) is convex), and \( b > 0 \) a constant. Let \( M_N = \max_{0 \leq n \leq N} Y_n \). Then

\[
b P(M_N \geq b) \leq E(Y_N 1_{(M_N \geq b)}) \leq E(Y_N).
\]

**Proof:** The second inequality is an obvious consequence of the assumption that \( Y_N \geq 0 \). To obtain the first, we use the stopping time \( T \) defined by

\[
T(\omega) = \begin{cases} 
\inf \{ n : Y_n(\omega) \geq b \} & \text{if } \exists n \leq N \text{ such that } Y_n(\omega) \geq b, \\
N & \text{otherwise.}
\end{cases}
\]

By the way we defined \( T \), it is clear that

\[
E Y_T = E(Y_T 1_{(M_N \geq b)}) + E(Y_T 1_{(M_N < b)}) \geq b P(M_N \geq b) + E(Y_N 1_{(M_N < b)}).
\]

However, because \( T \) is bounded by \( N \), we may use Theorem (19.1) to see that

\[
E Y_T \leq E Y_N = E(Y_N 1_{(M_N \geq b)}) + E(Y_N 1_{(M_N < b)}),
\]

and combining these two inequalities, we see that

\[
b P(M_N \geq b) \leq E Y_T - E(Y_N 1_{(M_N < b)}) \leq E(Y_N 1_{(M_N \geq b)}).
\]
Remark. In [1, page 250], Doob’s maximal inequality is proven for submartingales without the condition \( Y_N \geq 0 \), but the result is not really stronger: if \( Y_n \) is a submartingale, and \( Y_n = \max_{0 \leq n \leq N} Y_n^+ \), \( b > 0 \), then

\[
bP(Y_n \geq b) \leq E(Y_1 1_{(Y_n \geq b)}) \leq E(Y_N^+).
\]

The proof is more or less the same as the one just shown.

An easy consequence of this inequality is another proof of

**Theorem 19.3 (Kolmogorov’s inequality)*** Suppose that \( X_n \) are independent with \( EX_n = 0 \), \( \text{var}(X_n) < \infty \). If \( S_N = X_1 + \ldots + X_N \), then

\[
P \left( \max_{1 \leq n \leq N} \left| S_n \right| \geq x \right) \leq x^{-2} \text{var}(S_N).
\]

**Proof:** It is clear that \( S_n \) is a martingale (\( E(S_{n+1} | \mathcal{F}_n) = S_n + E(X_{n+1} | \mathcal{F}_n) = S_n + E(X_{n+1}) = S_n \)), so it follows, because \( x \to x^2 \) is a convex function, that \( S_n^2 \) is a submartingale. Now letting \( b = x^2 \), we apply Doob’s maximal inequality to \( S_n^2 \) to see that, when \( x \geq 0 \),

\[
x^2 P \left( \max_{1 \leq n \leq N} |S_n| \geq x \right) = x^2 P \left( \max_{1 \leq n \leq N} S_n^2 \geq x^2 \right) \leq E(S_n^2) = \text{var}(S_n).
\]

We can also use Doob’s inequality to prove the following theorem about convergence in \( L^p \) for \( p > 1 \):

**Theorem 19.4** If \( X_n \) is a martingale with \( \sup_n E|X_n|^p < \infty \) where \( p > 1 \), then \( X_n \to X \) almost surely and in \( L^p \).

**Proof:** See [1, pages 251-253].

The question of \( L^1 \) convergence for martingales does not follow so cleanly as does the case \( p > 1 \); we will show below that if \( X_n \) is a martingale with \( \sup_n E|X_n| < \infty \), then there exists a limit \( X \) such that \( X_n \to X \) almost surely, but there is not necessarily \( L^1 \) convergence. One example of this is furnished by the example of the symmetric simple random walk discussed above in connection with the optional stopping time result. We have that \( S_0 = 1 \), and \( S_n \) is a martingale, but clearly, if \( N = \inf\{m : S_m = 0\} \), then \( S_N = 0 \), and hence \( ES_N = 0 \). We can get \( L^1 \) convergence if the \( X_n \) satisfy a condition called uniformly integrability, which will be discussed in a future lecture; for the moment, we will discuss a few more examples of how convergence fails.

**Example (Double-or-nothing game).** Suppose that you begin with a dollar in your pocket, and you play a game such the probabilities of winning and losing are both equal to \( \frac{1}{2} \). You bet all the money in your pocket on a given game; if you lose and thereby have no money, you stop playing, while if you win and double your pocket money, you play again, betting all your money. So the first time you play, there is a \( \frac{1}{2} \) chance that you will end up with 2 dollars, and also a \( \frac{1}{2} \) chance that you will end up with nothing. If you win the first time, after the second time you play there is a \( \frac{1}{2} \) chance you will end up with 4 dollars, and a \( \frac{1}{2} \) chance you will end up with nothing. So more generally, after \( n \) games, the probability is \( 2^{-n} \) that you will end up with \( 2^n \) dollars, and \( 1 - 2^{-n} \) that you will have nothing. Let \( S_n \) denote your winnings after \( n \) games. A simple application of the Borel-Cantelli Lemma (applicable because \( \sum_{n=1}^\infty 2^{-n} < \infty \)) shows that \( S_n \to 0 \) almost surely. However, it is clear that \( ES_n = 2^n \cdot 2^{-n} + 0 \cdot (1 - 2^{-n}) = 1 \), so \( S_n \) does not converge in \( L^1 \).
Example (Critical Branching). Let \( \xi^n_i, i, n \geq 0 \) be i.i.d. nonnegative integer-valued random variables. Then define \( \{Z_n\} \), called a Galton-Watson process, by \( Z_0 = 1 \),

\[
Z_{n+1} = \begin{cases} 
\xi_{n+1}^1 + \cdots + \xi_{n+1}^{Z_n} & \text{if } Z_n > 0 \\
0 & \text{if } Z_n = 0
\end{cases}
\]

\( Z_n \) represents the number of people in some population in the \( n \)th generation, and each person in a given generation generates some number of offspring for the next generation; across people and generations, the number of offspring generated by a fixed person in a fixed generation are independent. If we let \( \mu = E(\xi^n_1) \), then \( Z_n/\mu^n \) is a martingale, but if \( \mu \leq 1 \), then \( Z_n/\mu^n \to 0 \) almost surely, even though \( Z_0 = 1 \). A detailed discussion of this process, including the results just mentioned, is in \cite[Section 4.3d]{1}.

There are variants of the convergence theorem for sub- and super-martingales; the most common is

**Theorem 19.5** Let \( \{S_n\} \) be a nonnegative supermartingale. Then there exists a random variable \( S_\infty \) such that \( S_n \to S_\infty \) almost surely, with \( ES_\infty \leq ES_0 \).

**Proof Sketch:** Let \( N \) be a fixed positive integer, \( b > 0 \), and let

\[
T = \inf\{0 \leq n \leq N : S_n \geq b, \text{ if } S_n < b \text{ for all } 0 \leq n \leq N\}.
\]

Because of the optional stopping result, we see that

\[
ES_0 \geq ES_T \geq bP(S_T \geq b) + 0 \cdot P(S_T < b),
\]

because \( S_n \) is a supermartingale, and because \( S_T \geq 0 \). Hence,

\[
P \left( \max_{0 \leq n \leq N} S_n \geq b \right) = P(S_T \geq b) \leq ES_0/b.
\]

Letting \( N \) go to \( \infty \), we should have that \( P(\sup_n S_n \geq b) \leq ES_0/b \).

Now suppose, for example, that \( S_0 = a \geq 0 \). Given that we are at \( a \), the probability that \( S_n \) will make an upcrossing of \([a, b]\), i.e., that \( S_n \geq b \) for some \( n \), should be \( \leq ES_0/b = a/b \). Given that we go below \( a \) at some later time, the chances that we should make another upcrossing are again \( \leq a/b \). More generally, we have the following version of Dubins’ Inequality: for \( S_n \geq 0 \) a supermartingale, the number \( U_{ab} \) of upcrossings of \([a, b]\) with \( 0 < a < b < \infty \), satisfies \( P(U_{ab} \geq n) \leq (a/b)^n \). It follows that

\[
EU_{ab} = \sum_{n=1}^{\infty} P(U_{ab} \geq n) \leq \sum_{n=1}^{\infty} (a/b)^n < \infty.
\]

Now \( U_{ab} = \infty \) means that we go below \( a \) and above \( b \) an infinite number of times, so \( P(U_{ab} = \infty) = P(\liminf S_n < a < b < \limsup S_n) = 0 \). Now letting \( a, b \) run over all rational pairs, we see that

\[
P \left( \bigcup_{a, b \in \mathbb{Q}^+} \{\liminf S_n < a < b < \limsup S_n\} \right) = 0,
\]

so it follows that \( \limsup S_n = \liminf S_n \) almost surely, i.e., \( \lim_{n \to \infty} S_n \) exists almost surely.

Call this limit \( S_\infty \). Since \( ES_0 \geq ES_n \), and since \( ES_\infty \leq \liminf_{n \to \infty} EX_n \) by Fatou’s Lemma, it follows that \( ES_\infty \leq ES_0 \). In particular, this also shows that \( S_\infty = \infty \) is impossible. \( \blacksquare \)
Lecture 19: Martingale Inequalities and Convergence Theorems

A detailed proof which utilizes different results on upcrossings is provided in [1, pages 234-236]. More details on Dubins’ inequality are discussed in Exercise 4.2.14 [1, page 238].

Example (Fair coin-tossing game). Suppose that you have a dollar in your pocket, and you play a coin-tossing game with a fair coin where you win a dollar if you get heads, and lose a dollar if you get tails; you don’t play at all if you lose all your money. If \( S_n \) denotes the amount of money in your pocket at time \( n \), it is clear that \( S_n \) is a nonnegative martingale, and that \( S_n \to 0 \) almost surely.

Example (Unfair coin-tossing game). Suppose that we have a biased coin, with probability \( p \) of heads, \( q \) of tails. Let us define i.i.d. random variables \( X_i \) which equal 1 when the \( i \)th coin toss is a head, and equal -1 when the \( i \)th coin toss is a tail; Say that \( S_0 = a > 0 \) where \( a \) is an integer, \( S_n = S_0 + X_1 + \ldots + X_n \). \( S_n \) is not a martingale, but it is possible to find a function \( h(x) \) such that \( h(S_n) \) is a martingale. Clearly, if \( h(S_n) \) is to be a martingale, we must have that

\[
h(x) = ph(x + 1) + qh(x - 1),
\]

because given \( S_n \), \( S_{n+1} \) is \( S_n + 1 \) with probability \( p \) and \( S_n - 1 \) with probability \( q \). We will try \( h(x) = r^x \), for an \( r \) to be determined shortly. From the above equation, we have that

\[
pr^{x+1} + qrr^{x-1} = rr^x + q\]

Solving this by the quadratic equation (and recalling that \( p + q = 1 \)), we obtain the solutions \( r = 1 \), \( r = q/p \); the former solution is trivial, so we use \( h(x) = (q/p)^x \). Hence, \( M_n = (q/p)^{S_n} \) is a martingale.

Let \( T = \inf \{ n : S_n = 0 \text{ or } b \} \) for \( b > a \) an integer. Now \( P(T < \infty) = 1 \), \( 0 \leq S_{T\wedge n} \) together imply that \( (q/p)^{S_{T\wedge n}} \) is bounded between \( (q/p)^0 = 1 \) and \( (q/p)^b \). Hence, we can use the dominated convergence theorem to see that

\[
E \left( \frac{q}{p} \right)^{S_T} = \lim_{n \to \infty} E \left( \frac{q}{p} \right)^{S_{T\wedge n}} = \left( \frac{q}{p} \right)^a,
\]

because \( (q/p)^{S_{T\wedge n}} \) is a martingale [1, page 234]. Finally, we use the equations

\[
P(S_T = b)(q/p)^b + P(S_T = 0)(q/p)^0 = ES_T = (q/p)^a
\]

\[
P(S_T = b) + P(S_T = 0) = 1
\]

to see that \( P(S_T = b) = \frac{(q/p)^a - 1}{(q/p)^b - 1} \), and \( P(S_T = 0) = \frac{(q/p)^b - (q/p)^a}{(q/p)^b - 1} \).

References