Chapter 7

CONDITIONAL EXPECTATION AND MARTINGALES

7.1 Conditional Expectation.

Throughout this section we will assume that random variables X are defined on a probability space $(\Omega, \mathcal{F}, P)$ and have finite second moments so $E(X^2) < \infty$. This allows us to define conditional expectation through approximating one random variable by another, measurable with respect to a courser (or less informative) sigma-algebra. We begin with the coursest sigma algebra of all, the trivial one $\{\Omega, \varphi\}$, with respect to which only constants are measurable.

What constant is the best fit to a random variable in the sense of smallest mean squared error? In other words, what is the value of c solving

$$\min_c E[(X-c)^2]?$$

Expanding,

$$E[(X-c)^2] = var(X) + (EX - c)^2$$

and so the minimum is achieved when we choose $c = EX$.

A constant is, of course, a random variable but a very basic one, measurable with respect to the trivial sigma-field $\{\Omega, \varphi\}$. Now suppose that we wished to approximate the value of a random variable X, not with a constant, but with another random variable Z, measurable with respect to some other sigma field $\mathcal{G} \subset \sigma(X)$. How course or fine the sigma algebra $\mathcal{G}$ is depends on how much information we have pertinent to the approximation of X. How good is our approximation will be measured using the mean squared error

$$E[(X-Z)^2]$$

and we wish to minimize this over all possible $\mathcal{G}-$random variables Z. The minimizing value of Z is the conditional expected value of X.

Theorem 115 *(conditional expectation as a projection) Let $\mathcal{G} \subset \mathcal{F}$ be sigma-algebras and X a random variable on $(\Omega, \mathcal{F}, P)$. Assume $E(X^2) < \infty$. Then there exists an almost surely unique $\mathcal{G}$-measurable Y such that*

$$E[(X-Y)^2] = inf_Z E(X-Z)^2 \tag{7.1}$$

where the infimum is over all $\mathcal{G}$-measurable random variables.

Definition 116 *We denote the minimizing Y by $E(X|\mathcal{G})$.*

The next result assures us that the conditional expectation is unique, almost surely. In other words two random variables Y which solve the above minimization problem differ on a set of probability zero.

Theorem 117 *For two such minimizing Y_1, Y_2 , i.e. random variables Y which satisfy (7.1), we have $P[Y_1 = Y_2] = 1$. This implies that conditional expectation is almost surely unique.*

Proof. Suppose both Y_1 and Y_2 are $\mathcal{G}$-measurable and both minimize $E[(X-Y)^2]$. Then for any $A \in \mathcal{G}$ it follows from property (d) below that

$$\int_A Y_1 dP = \int_A Y_2 dP$$

or

$$\int_A (Y_1 - Y_2) dP = 0.$$

Choose $A = [Y_1 - Y_2 \geq 0]$ and note that

$$\int (Y_1 - Y_2) I_A dP = 0$$

and the integrand $(Y_1-Y_2)I_A$ is non-negative together imply that $(Y_1-Y_2)I_A = 0$ almost surely. Similarly on the set $A = [Y_1 - Y_2 < 0]$ we can show that $(Y_1 - Y_2)I_A = 0$ almost surely. It follows that $Y_1 = Y_2$ almost surely. ■

Example 118 *Suppose $\mathcal{G} = \{\varphi, \Omega\}$. What is $E(X|\mathcal{G})$?*

The only random variables which are measurable with respect to the trivial sigma-field are constants. So this leads to the same minimization discussed above, $\min_c E[(X-c)^2] = \min_c\{var(X)+(EX-c)^2\}$ which results in $c = E(X)$.

Example 119 *Suppose $\mathcal{G} = \{\varphi, A, A^c, \omega\}$ for some event A. What is $E(X|\mathcal{G})$? Consider the special case: $X = I_B$.*

In this case suppose the random variable Z takes the value a on A and b on the set A^c. Then

$$\begin{aligned} E[(X-Z)^2] &= E[(X-a)^2 I_A] + E[(X-b)^2 I_{A^c}] \\ &= E(X^2 I_A) - 2aE(XI_A) + a^2 P(A) \\ &+ E(X^2 I_{A^c}) - 2bE(XI_{A^c}) + b^2 P(A^c). \end{aligned}$$

Minimizing this with respect to both a and b results in

$$\begin{aligned} a &= E(XI_A)/P(A) \\ b &= E(XI_{A^c})/P(A^c). \end{aligned}$$

These values a and b are usually referred to in elementary probability as $E(X|A)$ and $E(X|A^c)$ respectively. Thus, the conditional expectated value can be written

$$E(X|\mathcal{G})(\omega) = \begin{cases} E(X|A) & \text{if } \omega \in A \\ E(X|A^c) & \text{if } \omega \in A^c \end{cases}$$

As a special case consider X to be an indicator random variable $X = I_B$. Then we usually denote $E(I_B|\mathcal{G})$ by $P(B|\mathcal{G})$ and

$$P(B|\mathcal{G})(\omega) = \begin{cases} P(B|A) & \text{if } \omega \in A \\ P(B|A^c) & \text{if } \omega \in A^c \end{cases}$$

Note: *Expected value is a constant, but the conditional expected value $E(X|\mathcal{G})$ is a random variable measurable with respect to $\mathcal{G}$. Its value on the atoms of $\mathcal{G}$ is the average of the random variable X over these atoms.*

Example 120 *Suppose $\mathcal{G}$ is generated by a finite partition $\{A_1, A_2, ..., A_n\}$ of the probability space Ω.. What is $E(X|\mathcal{G})$?*

In this case, any $\mathcal{G}$-measurable random variable is constant on the sets in the partition $A_j, j = 1, 2, ..., n$ and an argument similar to the one above shows that the conditional expectation is the simple random variable:

$$\begin{aligned} E(X|\mathcal{G})(\omega) &= \sum_{i=1}^{n} c_i I_{A_i}(\omega) \\ \text{where } c_i &= E(X|A_i) = \frac{E(XI_{A_i})}{P(A_i)} \end{aligned}$$

Example 121 *Consider the probability space $\Omega = (0,1]$ together with $P =$ Lebesgue measure and the Borel Sigma Algebra. Suppose the function $X(\omega)$ is Borel measurable. Assume that $\mathcal{G}$ is generated by the intervals $(\frac{j-1}{n}, \frac{j}{n}]$ for $j = 1, 2, ..., n$. What is $E(X|\mathcal{G})$?*

In this case

$$\begin{aligned} E(X|\mathcal{G})(\omega) &= n \int_{(j-1)/n}^{j/n} X(s)ds \quad \text{when } \omega \in (\frac{j-1}{n}, \frac{j}{n}] \\ &= \text{average of } X \text{ values over the relevant interval.} \end{aligned}$$

7.1.1 Properties of Conditional Expectation.

(a) If a random variable X is $\mathcal{G}$-measurable, $E(X|\mathcal{G}) = X$.

(b) If a random variable X independent of a sigma-algebra $\mathcal{G}$, then $E(X|\mathcal{G}) = E(X)$.

(c) For any square integrable $\mathcal{G}$-measurable Z, $E(ZX) = E[ZE(X|\mathcal{G})]$.

(d) (special case of (c)): $\int_A XdP = \int_A E(X|\mathcal{G}]dP$ for all $A \in \mathcal{G}$.

(e) $E(X) = E[E(X|\mathcal{G})]$.

(f) If a $\mathcal{G}$-measurable random variable Z satisfies $E[(X - Z)Y] = 0$ for all other $\mathcal{G}$-measurable random variables Y, then $Z = E(X|\mathcal{G})$.

(g) If Y_1, Y_2 are distinct $\mathcal{G}-$measurable random variables both minimizing $E(X - Y)^2$, then $P(Y_1 = Y_2) = 1$.

(h) *Additive* $E(X + Y|\mathcal{G}) = E(X|\mathcal{G}) + E(Y|\mathcal{G})$.
Linearity $E(cX + d|\mathcal{G}) = cE(X|\mathcal{G}) + d$.

(i) If Z is $\mathcal{G}-$measurable, $E(ZX|\mathcal{G}) = ZE(X|\mathcal{G})$ a.s.

(j) If $\mathcal{H} \subset \mathcal{G}$ are sigma-algebras, $E[E(X|\mathcal{G})|\mathcal{H}] = E(X|\mathcal{H})$.

(k) If $X \leq Y$, $E(X|\mathcal{G}) \leq E(Y|\mathcal{G})$ a.s.

(l) *Conditional Lebesgue Dominated Convergence.* If $X_n \to X$ in probability and $|X_n| \leq Y$ for some integrable random variable Y, then $E(X_n|\mathcal{G}) \to E(X|\mathcal{G})$ in probability.

Notes. In general, we define $E(X|Z) = E(X|\sigma(Z))$ and conditional variance $var(X|\mathcal{G}) = E\{(X - E(X|\mathcal{G}))^2|\mathcal{G}\}$. For results connected with property (l) above providing conditions under which the conditional expectations converge, see Convergence in distribution of conditional expectations, (1994) E.M. Goggin, *Ann. Prob* 22, 2. 1097-1114.

Proof. (Proof of the above properties)
(a) Notice that for any random variable Z that is $\mathcal{G}$-measurable, $E(X - Z)^2 \geq E(X - X)^2 = 0$ and so the minimizing Z is X (by definition this is $E(X|\mathcal{G})$).
(b) Consider a random variable Y measurable with respect $\mathcal{G}$ and therefore independent of X. Then

$$\begin{aligned}
E(X - Y)^2 &= E[(X - EX + EX - Y)^2] \\
&= E[(X - EX)^2] + 2E[(X - EX)(EX - Y)] + E[(EX - Y)^2] \\
&= E[(X - EX)^2] + E[(EX - Y)^2] \text{ by independence} \\
&\geq E[(X - EX)^2].
\end{aligned}$$

It follows that $E(X-Y)^2$ is minimized when we choose $Y = EX$ and so $E(X|\mathcal{G}) = E(X)$.
(c) for any $\mathcal{G}-$measurable square integrable random variable Z, we may define a quadratic function of λ by

$$g(\lambda) = E[(X - E(X|\mathcal{G}) - \lambda Z)^2]$$

By the definition of $E(X|\mathcal{G})$, this function is minimized over all real values of λ at the point $\lambda = 0$ and therefore $g'(0) = 0$. Setting its derivative $g'(0) = 0$ results in the equation

$$E(Z(X - E(X|\mathcal{G}))) = 0$$

or $E(ZX) = E[ZE(X|\mathcal{G})]$.
(d) If in (c) we put $Z = I_A$ where $A \in \mathcal{G}$, we obtain $\int_A XdP = \int_A E(X|\mathcal{G}]dP$.
(e) Again this is a special case of property (c) corresponding to $Z = 1$.
(f) Suppose a $\mathcal{G}$-measurable random variable Z satisfies $E[(X-Z)Y] = 0$ for all other $\mathcal{G}$-measurable random variables Y. Consider in particular $Y = E(X|\mathcal{G}) - Z$ and define

$$\begin{aligned} g(\lambda) &= E[(X - Z - \lambda Y)^2] \\ &= E((X-Z)^2 - 2\lambda E[(X-Z)Y] + \lambda^2 E(Y^2) \\ &= E(X-Z)^2 + \lambda^2 E(Y^2) \\ &\geq E(X-Z)^2 = g(0). \end{aligned}$$

In particular $g(1) = E[(X - E(X|\mathcal{G}))^2] \geq g(0) = E(X-Z)^2$ and by Theorem 117, $Z = E(X|\mathcal{G})$ almost surely.
(g) This is just deja vu (Theorem 117) all over again.
(h) Consider, for an arbitrary $\mathcal{G}-$measurable random variable Z,

$$\begin{aligned} E[Z(X+Y-E(X|\mathcal{G})-E(Y|\mathcal{G}))] &= E[Z(X-E(X|\mathcal{G}))] + E[Z(Y-E(Y|\mathcal{G}))] \\ &= 0 \text{ by property (c).} \end{aligned}$$

It therefore follows from property (f) that $E(X+Y|\mathcal{G}) = E(X|\mathcal{G}) + E(Y|\mathcal{G})$. *By a similar argument we may prove* $E(cX + d|\mathcal{G}) = cE(X|\mathcal{G}) + d$.
(i) This is Problem 2.
(j) This is Problem 4 (sometimes called the tower property of conditional expectation: If $\mathcal{H} \subset \mathcal{G}$ are sigma-algebras, $E[E(X|\mathcal{G})|\mathcal{H}] = E(X|\mathcal{H})$).
(k) We need to show that if $X \leq Y$, $E(X|\mathcal{G}) \leq E(Y|\mathcal{G})$ a.s.
(l) *Conditional Lebesgue Dominated Convergence.* If $X_n \to X$ in probability. and $|X_n| \leq Y$ for some integrable random variable Y , then it is easy to show that $E|X_n - X| \to 0$. Therefore

$$\begin{aligned} E|E(X_n|\mathcal{G}) - E(X|\mathcal{G})| &= E|E(X_n - X|\mathcal{G})| \\ &\leq E\{E(|X_n - X||\mathcal{G})\} \\ &\leq E|X_n - X| \to 0 \end{aligned}$$

implying that $E(X_n|\mathcal{G}) \to E(X|\mathcal{G})$ in probability. ■

> **Notes.** In general, we define $E(X|Z) = E(X|\sigma(Z))$ and conditional variance $var(X|\mathcal{G}) = E\{(X - E(X|\mathcal{G}))^2|\mathcal{G}\}$. For results connected with property (l) above providing conditions under which the conditional expectations converge, see Convergence in distribution of conditional expectations, (1994) E.M. Goggin, *Ann. Prob* 22, 2. 1097-1114.

7.2 Conditional Expectation for integrable random variables.

We have defined conditional expectation as a projection only for random variables with finite variance. It is fairly easy to extend this definition to random variables X on a probability space $(\Omega, \mathcal{F}, P)$ for which $E(|X|) < \infty$. We wish to define $E(X|\mathcal{G})$ where the sigma algebra $\mathcal{G} \subset \mathcal{F}$. First, for non-negative integrable X choose simple random variables $X_n \uparrow X$. Since simple random variables have only finitely many values, they have finite variance, and we can use the definition above for their conditional expectation. Then $E(X_n|\mathcal{G}) \uparrow$ and so it converges. Define $E(X|\mathcal{G})$ to be the limit. In general, for random variables taking positive and negative values, we define $E(X|\mathcal{G}) = E(X^+|\mathcal{G}) - E(X^-|\mathcal{G})$. There are a number of details that need to be ironed out. First we need to show that this new definition is consistent with the old one when the random variable happens to be square integrable. We can also show that the properties (a)-(i) above all hold under this new definition of conditional expectation. We close with the more common definition of conditional expectation found in most probability and measure theory texts, essentially property (d) above. It is, of course, equivalent to the definition as a projection in section 7.1 and the definition above as a limit of the conditional expectation of simple functions.

Theorem 122 *Consider a random variable X defined on a probability space $(\Omega, \mathcal{F}, P)$ for which $E(|X|) < \infty$. Suppose the sigma algebra $\mathcal{G} \subset \mathcal{F}$. Then there is a unique (almost surely P) $\mathcal{G}-$measurable random variable Z satisfying*

$$\int_A X dP = \int_A Z dP \quad \text{for all } A \in \mathcal{G}$$

Any such Z we call the conditional expectation and denote by $E(X|\mathcal{G})$.

7.3 Martingales in Discrete Time

In this section all random variables are defined on the same probability space $(\Omega, \mathcal{F}, P)$. Partial information about these random variables may be obtained from the observations so far, and in general, the "history" of a process up to time t is expressed through a sigma-algebra $H_t \subset \mathcal{F}$. We are interested in stochastic processes or sequences of random variables called martingales, intuitively, the total fortune of an individual participating in a "fair game". In order for the game to be "fair", the expected value of your future fortune given the history of the process up to and including the present should be equal to your present wealth. In a sense you are neither tending to increase or decrease your wealth over time- any fluctuations are purely random. Suppose your fortune at time s is denoted X_s . The values of the process of interest and any other related processes up to time s generate a sigma-algebra H_s. Then the assertion that the game is fair implies that the expected value of our future fortune given this history of the process up to the present is exactly our present wealth $E(X_t|H_s) = X_s$ for $t > s$. Suppose $\mathcal{T}$ is some set indexing "time" for a martingale. Normally $\mathcal{T}$ is either an interval on the real line or the non-negative integers.

Definition 123 $\{(X_t, H_t); t \in \mathcal{T}\}$ *is a* martingale *if*

(a) H_t is increasing (in t) family of sigma-algebras

(b) Each X_t is H_t- measurable and $E|X_t| < \infty$.

(c) For each $s < t$, $s, t \in \mathcal{T}$, we have $E(X_t|H_s) = X_s$ a.s.

Example 124 *Suppose Z_t are independent random variables with expectation 0. Define $H_t = \sigma(Z_1, Z_2, \ldots Z_t)$ and $S_t = \sum_{i=1}^t Z_i$. Then $\{(S_t, H_t)$,with $t = 1, 2, ...\}$ is a martingale. Suppose that $E(Z_t^2) = \sigma^2 < \infty$. Then $\{(S_t^2 - t\sigma^2, H_t), t = 1, 2, ...\}$ is a martingale.*

Example 125 *Suppose Z_t are independent random variables with $Z_t \geq 0$. Define $H_t = \sigma(Z_1, Z_2, \ldots Z_t)$ and $M_t = \prod_{i=1}^t Z_i$. Suppose that $E(Z_i^\lambda) = \phi(\lambda) < \infty$. Then*

$$\{(\frac{M_t^\lambda}{\phi^t(\lambda)}, H_t), t = 1, 2, ...\}$$

is a martingale.

This is an example of a parametric family of martingales indexed by λ obtained by multiplying independent random variables.

Example 126 *Let X be any integrable random variable, and H_t an increasing family of sigma-algebras. Put $X_t = E(X|H_t)$. Then (X_t, H_t) is a martingale.*

Definition 127 *Let $\{(M_n, H_n); n = 1, 2, ...\}$ be a martingale and A_n be a sequence of random variables measurable with respect to H_{n-1}. Then the sequence A_n is called* ***non-anticipating****. (an alternate term is* ***predictable****)*

In gambling, we must determine our stakes and our strategy on the $n'th$ play of a game based on the information available to use at time $n-1$. Similarly, in investment, we must determine the weights on various components in our portfolio at the end of day (or hour or minute) $n-1$ *before* the random marketplace determines our profit or loss for that period of time. In this sense gambling and investment strategies must be determined by non-anticipating sequences of random variables (although both gamblers and investors often dream otherwise).

Definition 128 *(Martingale Transform). Let* $\{\{(M_t, H_t), t = 0, 1, 2, ...\}$ *be a martingale and let* A_n *be a bounded non-anticipating sequence with respect to* H_n*. Then the sequence*

$$\tilde{M}_t = A_1(M_1 - M_0) + ... + A_t(M_t - M_{t-1}) \tag{7.2}$$

is called a Martingle transform of M_t*.*

The martingale transform is sometimes denoted $A \circ M$.

Theorem 129 $\{(\tilde{M}_t, H_t), t = 1, 2, ...\}$ *is a martingale.*

Proof.

$$\begin{aligned} E[\widetilde{M}_j - \widetilde{M}_{j-1}|H_{j-1}] &= E[A_j(M_j - M_{j-1}|H_{j-1}] \\ &= A_j E[(M_j - M_{j-1}|H_{j-1}] \text{ since } A_j \text{ is } H_{j-1} \text{ measurable} \\ &= 0 \text{ a.s.} \end{aligned}$$

Therefore

$$E[\widetilde{M}_j|H_{j-1}] = \widetilde{M}_{j-1} \text{ a.s.}$$

■

Consider a random variable τ that determines when we stop betting or investing. Its value can depend arbitrarily on the outcomes in the past, as long as the decision to stop at time $\tau = n$ depends only on the results at time $n, n-1$, ...etc. Such a random variable is called an optional stopping time.

Definition 130 *A random variable* τ *taking values in* $\{0, 1, 2, ...\} \cup \{\infty\}$ *is a (optional)* stopping time *for a martingale* (X_t, H_t) *if for each* n *,* $[\tau \leq n] \in H_n$ *.*

If we stop a martingle at some random stopping time, the result continues to be a martingale as the following theorem shows.

Theorem 131 *Suppose that* $\{(M_t, H_t), t = 1, 2, ...\}$ *is a martingale and* τ *is an optional stopping time. Define* $Y_n = M_{n\wedge\tau} = M_{\min(n,\tau)}$*. Then* $\{(Y_n, H_n), n = 1, 2, ..\}$ *is a martingale.*

Proof. Notice that

$$M_{n\wedge\tau} = M_0 + \sum_{j=1}^{n} (M_j - M_{j-1}) I(\tau \geq j).$$

Letting $A_j = I(\tau \geq j)$ this is a bounded $H_{j-1}-$measurable sequence and therefore $\sum_{j=1}^{n}(M_j - M_{j-1})I(\tau \geq j)$ is a martingale transform. By Theorem 129 it is a martingale. ■

Example 132 *(Ruin probabilities). Consider a random walk* $S_n = \sum_{i=1}^n X_i$ *where the random variables* X_i *are independent identically distributed with* $P(X_i = 1) = p,\ P(X_i = -1) = q,\ P(X_i = 0) = 1 - p - q$ *for* $0 < p + q \le 1, p \neq q$. *Then* $M_n = (q/p)^{S_n}$ *is a martingale. Suppose that* $A < S_0 < B$ *and define the optional stopping time* τ *as the first time* S_n *hits either of two barriers at* A *or* B. *If* $p \neq \frac{1}{2}$ *then since by the Law of large numbers we have*

$$\frac{S_n}{n} \to p - q \ a.s.$$

this guarantees that one of the two boundaries is eventually hit with probability 1. Then $M_{n\wedge\tau}$ *is a martingale. Since* $E(M_\tau) = \lim_{n\to\infty} E(M_{n\wedge\tau}) = (q/p)^{S_0}$ *by dominated covergence, we have*

$$(q/p)^A p_A + (q/p)^B p_B = (q/p)^{S_0} \tag{7.3}$$

where p_A *and* $p_B = 1 - p_A$ *are the probabilities of hitting absorbing barriers at* A *or* B *respectively. Solving, it follows that*

$$((q/p)^A - (q/p)^B)p_A = (q/p)^{S_0} - (q/p)^B \tag{7.4}$$

or that

$$p_A = \frac{(q/p)^{S_0} - (q/p)^B}{(q/p)^A - (q/p)^B}. \tag{7.5}$$

In the case $p = q$, *a similar argument (or alternatively taking limits as* $p \to \frac{1}{2}$*) provides*

$$p_A = \frac{B - S_0}{B - A}. \tag{7.6}$$

Definition 133 *For an optional stopping time* τ *define*

$$H_\tau = \{A \in H; A \cap [\tau \le n] \in H_n,\ \textit{for all } n\}. \tag{7.7}$$

Theorem 134 H_τ *is a sigma-algebra.*

Proof. Clearly since the empty set $\varphi \in H_n$ for all n, so is $\varphi \cap [\tau \le n]$ and so $\varphi \in H_\tau$. We also need to show that if $A \in H_\tau$ then so is the complement A^c. Notice that for each n,

$$\begin{aligned}&[\tau \le n] \cap \{A \cap [\tau \le n]\}^c \\ &\quad= [\tau \le n] \cap \{A^c \cup [\tau > n]\} \\ &\quad= A^c \cap [\tau \le n]\end{aligned}$$

and since each of the sets $[\tau \le n]$ and $A \cap [\tau \le n]$ are H_n-measurable, so must be the set $A^c \cap [\tau \le n]$. Since this holds for all n it follows that whenever $A \in H_\tau$ then so A^c. Finally, consider a sequence of sets $A_m \in H_\tau$ for all $m = 1, 2, ...$ We need to show that the countable union $\cup_{m=1}^\infty A_m \in H_\tau$. But

$$\{\cup_{m=1}^\infty A_m\} \cap [\tau \le n] = \cup_{m=1}^\infty \{A_m \cap [\tau \le n]\}$$

and by assumption the sets $\{A_m \cap [\tau \leq n]\} \in H_n$ for each n. Therefore

$$\cup_{m=1}^{\infty}\{A_m \cap [\tau \leq n]\} \in H_n$$

and since this holds for all n, $\cup_{m=1}^{\infty} A_m \in H_\tau$. ∎

Definition 135 $\{(X_t, H_t); t \in T\}$ *is a* submartingale *if*

(a) H_t is increasing (in t) family of sigma-algebras.

(b) Each X_t is H_t measurable and $E|X_t| < \infty$.

(c) For each $s < t,$, $E(X_t|H_s) \geq X_s$ a.s.

Note that every martingale is a submartingale. There is a version of Jensen's inequality for conditional expectation as well as the one proved before for ordinary expected value.

Theorem 136 *(Jensen's Inequality-conditional version) Let ϕ be a convex function. Then for any random variable X and sigma-field H,*

$$\phi(E(X|H)) \leq E(\phi(X)|H). \tag{7.8}$$

Proof. Consider the set $\mathcal{L}$ of linear function $L(x) = a + bx$ that lie entirely below the graph of the function $\phi(x)$. It is easy to see that for a convex function

$$\phi(x) = \sup\{L(x); L \in \mathcal{L}\}.$$

For any such line, since $\phi(x) \geq L(x)$,

$$E(\phi(X)|H) \geq E(L(X)|H) = L(E(X)|H)).$$

If we take the supremum over all $L \in \mathcal{L}$, we obtain

$$E(\phi(X)|H) \geq \phi(E(X)|H)).$$

∎

Example 137 *Let X be any random variable and H be a sigma-field. Then for $1 \leq p \leq k < \infty$*

$$\{E(|X|^p|H)\}^{1/p} \leq \{E(|X|^k|H)\}^{1/k}. \tag{7.9}$$

In the special case that H is the trivial sigma-field, this is the inequality

$$||X||_p \leq ||X||_k \quad where \ ||X||_p = (E|X|^p)^{1/p}. \tag{7.10}$$

Proof. Consider the function $\phi(x) = |x|^{k/p}$. This function is convex provided that $k \geq p$ and by the conditional form of Jensen's inequality,

$$E(|X|^k|H) = E(\phi(|X|^p)|H) \geq \phi(E(|X|^p|H)) = |E(|X|^p|H)|^{k/p} \text{ a.s.}$$

∎

Example 138 *(Constructing Submartingales). Let S_n be a martingale with respect to H_n. Then $(|S_n|^p, H_n)$ is a submartingale for any $p \geq 1$ provided that $E|S_n|^p < \infty$.*

Proof. Since the function $\phi(x) = |x|^p$ is convex for $p \geq 1$, it follows from the conditional form of Jensen's inequality that

$$E(|S_{n+1}|^p|H_n) = E(\phi(S_{n+1})|H_n) \geq \phi(E(S_{n+1}|H_n)) = \phi(S_n) = |S_n|^p \text{ a.s.}$$

■

Theorem 139 *Let X_n be a submartingale and suppose ϕ is a convex nondecreasing function with $E\phi(X_n) < \infty$. Then $\phi(X_n)$ is a submartingale.*

Proof. Since the function $\phi(x)$ is convex ,

$$E(\phi(S_{n+1})|H_n) \geq \phi(E(S_{n+1}|H_n)) \geq \phi(S_n) \text{ a.s.}$$

since $E(S_{n+1}|H_n) \geq S_n$ a.s. and the function ϕ is non-decreasing. ■

Corollary 140 *Let (X_n, H_n) be a submartingale. Then $((X_n - a)^+, H_n)$ is a submartingale.*

Proof. The function $\phi x) = (x-a)^+$ is convex and non-decreasing. ■

Theorem 141 *(Doob's Maximal Inequality) Suppose (M_n, H_n) is a non-negative submartingale. Then for $\lambda > 0$ and $p \geq 1$,*

$$P(\sup_{0 \leq m \leq n} M_m \geq \lambda) \leq \lambda^{-p} E(M_n^p)$$

Proof. We prove this in the case $p = 1$. The general case we leave as a problem. Define a stopping time

$$\tau = \min\{m; M_m \geq \lambda\}$$

and on the set that it never occurs that $M_m \geq \lambda$ we can define $\tau = \infty$. Then $\tau \leq n$ if and only if the maximum has reached the value λ by time n or

$$P[\sup_{0 \leq m \leq n} M_m \geq \lambda] = P[\tau \leq n].$$

Now on the set $[\tau \leq n]$, the maximum $M_\tau \geq \lambda$ so

$$\lambda I(\tau \leq n) \leq M_\tau I(\tau \leq n) = \sum_{i=1}^{n} M_i I(\tau = i). \tag{7.11}$$

By the submartingale property, for any $i \leq n$ and $A \in H_i$,

$$E(M_i I_A) \leq E(M_n I_A).$$

Therefore, taking expectations on both sides of (7.11), and noting that for all $i \leq n$,

$$E(M_i I(\tau = i)) \leq E(M_n I(\tau = i))$$

we obtain

$$\lambda P(\tau \leq n) \leq E(M_n I(\tau \leq n)) \leq E(M_n).$$

■

Theorem 142 *(Doob's L^p Inequality) Suppose (M_n, H_n) is a non-negative submartingale and put $M_n^* = \sup_{0 \leq m \leq n} M_m$. Then for $p > 1$, and all n*

$$||M_n^*||_p \leq \frac{p}{p-1} ||M_n||_p$$

One of the main theoretical properties of martingales is that they converge under fairly general conditions. Conditions are clearly necessary. For example consider a simple random walk $S_n = \sum_{i=1}^n Z_i$ where $Z_i$are independent identically distributed with $P(Z_i = 1) = P(Z_i = -1) = \frac{1}{2}$. Starting with an arbitrary value of S_0, say $S_0 = 0$ this is a martingale, but as $n \to \infty$ it does not converge almost surely or in probability.

On the other hand, consider a Markov chain with the property that

$$P(X_{n+1} = j | X_n = i) = \frac{1}{2i+1} \quad \text{for} \quad j = 0, 1, ..., 2i.$$

Notice that this is a martingale and beginning with a positive value, say $X_0 = 10$, it is a non-negative martingale. Does it converge almost surely? If so the only possible limit is $X = 0$ because the nature of the process is such that $P[|X_{n+1} - X_n| \geq 1 | X_n = i] \geq \frac{2}{3}$ unless $i = 0$. The fact that it does converge a.s. is a consequence of the martingale convergence theorem. Does it converge in L_1 i.e. in the sense that $E[|X_n - X|] \to 0$ as $n \to \infty$? If so, then clearly $E(X_n) \to E(X) = 0$ and this contradicts the martingale property of the sequence which implies $E(X_n) = E(X_0) = 10$. This is an example of a martingale that converges almost surely but not in L_1.

Lemma 143 *If $(X_t, H_t), t = 1, 2, ..., n$ is a (sub)martingale and if α, β are optional stopping times with values in $\{1, 2, ..., n\}$ such that $\alpha \leq \beta$ then*

$$E(X_\beta | H_\alpha) \geq X_\alpha$$

with equality if X_t is a martingale.

Proof. It is sufficient to show that

$$\int_A (X_\beta - X_\alpha) dP \geq 0$$

for all $A \in H_\alpha$. Note that if we define $Z_i = X_i - X_{i-1}$ to be the submartingale differences, the submartingale condition implies

$$E(Z_j | H_i) \geq 0 \text{ a.s. whenever } i < j.$$

Therefore for each $j = 1, 2, ...n$ and $A \in H_\alpha$,

$$\begin{aligned}\int_{A\cap[\alpha=j]}(X_\beta - X_\alpha)dP &= \int_{A\cap[\alpha=j]}\sum_{i=1}^{n} Z_i I(\alpha < i \le \beta)dP \\ &= \int_{A\cap[\alpha=j]}\sum_{i=j+1}^{n} Z_i I(\alpha < i \le \beta)dP \\ &= \int_{A\cap[\alpha=j]}\sum_{i=j+1}^{n} E(Z_i|H_{i-1})I(\alpha < i)I(i \le \beta)dP \\ &\ge 0 \text{ a.s.}\end{aligned}$$

since $I(\alpha < i)$, $I(i \le \beta)$ and $A \cap [\alpha = j]$ are all measurable with respect to H_{i-1} and $E(Z_i|H_{i-1}) \ge 0$ a.s. If we add over all $j = 1, 2, ..., n$ we obtain the desired result. ■

The following inequality is needed to prove a version of the submartingale convergence theorem.

Theorem 144 *(Doob's upcrossing inequality) Let M_n be a submartingale and for $a < b$, define $N_n(a, b)$ to be the number of complete upcrossings of the interval (a, b) in the sequence $M_j, j = 0, 1, 2, ..., n$. This is the largest k such that there are integers $i_1 < j_1 < i_2 < j_2 ... < j_k \le n$ for which*

$$M_{i_l} \le a \quad \text{and } M_{j_l} \ge b \quad \text{for all } l = 1, ..., k.$$

Then

$$(b - a)EN_n(a, b) \le E\{(M_n - a)^+ - (M_0 - a)^+\}$$

Proof. By Corollary 140, we may replace M_n by $X_n = (M_n - a)^+$ and this is still a submartingale. Then we wish to count the number of upcrossings of the interval $[0, b']$ where $b' = b - a$. Define stopping times for this process by $\alpha_0 = 0$,

$$\begin{aligned}\alpha_1 &= \min\{j; 0 \le j \le n, X_j = 0\} \\ \alpha_2 &= \min\{j; \alpha_1 \le j \le n, X_j \ge b'\} \\ &... \\ \alpha_{2k-1} &= \min\{j; \alpha_{2k-2} \le j \le n, X_j = 0\} \\ \alpha_{2k} &= \min\{j; \alpha_{2k-1} \le j \le n, X_j \ge b'\}.\end{aligned}$$

In any case, if α_k is undefined because we do not again cross the given boundary, we define $\alpha_k = n$. Now each of these random variables is an optional stopping time. If there is an upcrossing between X_{α_j} and $X_{\alpha_{j+1}}$ (where j is odd) then the distance travelled

$$X_{\alpha_{j+1}} - X_{\alpha_j} \ge b'.$$

If X_{α_j} is well-defined (i.e. it is equal to 0) and there is no further upcrossing, then $X_{\alpha_{j+1}} = X_n$ and

$$X_{\alpha_{j+1}} - X_{\alpha_j} = X_n - 0 \ge 0.$$

Similarly if j is even, since by Lemma 143, $(X_{\alpha_j}, H_{\alpha_j})$ is a submartingale,

$$E(X_{\alpha_{j+1}} - X_{\alpha_j}) \geq 0.$$

Adding over all values of j, and using the fact that $\alpha_0 = 0$ and $\alpha_n = n$,

$$E\sum_{j=0}^{n}(X_{\alpha_{j+1}} - X_{\alpha_j}) \geq b' EN_n(a,b)$$
$$E(X_n - X_0) \geq b' EN_n(a,b).$$

In terms of the original submartingale, this gives

$$(b-a)EN_n(a,b) \leq E(M_n - a)^+ - E(M_0 - a)^+.$$

■

Doob's martingale convergence theorem that follows is one of of the nicest results in probability and one of the reasons why martingales are so frequently used in finance, econometrics, clinical trials and lifetesting.

Theorem 145 *(Sub)martingale Convergence Theorem. Let (M_n, H_n); $n = 1, 2, \ldots$ be a submartingale such that $sup_{n\to\infty} EM_n^+ < \infty$. Then there is an integrable random variable M such that $M_n \to M$ a.s.*

Proof. The proof is an application of the upcrossing inequality. Consider any interval $a < b$ with rational endpoints. By the upcrossing inequality,

$$E(N_a(a,b)) \leq \frac{1}{b-a}E(M_n - a)^+ \leq \frac{1}{b-a}[|a| + E(M_n^+)]. \tag{7.12}$$

Let $N(a,b)$ be the total number of times that the martingale completes an upcrossing of the interval $[a,b]$ over the infinite time interval $[1,\infty)$ and note that $N_n(a,b) \uparrow N(a,b)$ as $n \to \infty$. Therefore by monotone convergence $E(N_a(a,b)) \to EN(a,b)$ and by (7.12)

$$E(N(a,b)) \leq \frac{1}{b-a}\limsup[a + E(M_n^+)] < \infty.$$

This imples

$$P[N(a,b) < \infty] = 1.$$

Therefore,

$$P(\liminf M_n \leq a < b \leq \limsup M_n) = 0$$

for every rational $a < b$ and this implies that M_n converges almost surely to a (possibly infinite) random variable. Call this limit $M.We$ need to show that this random variable is almost surely finite. Because $E(M_n)$ is non-decreasing,

$$E(M_n^+) - E(M_n^-) \geq E(M_0)$$

and so

$$E(M_n^-) \le E(M_n^+) - E(M_0).$$

But by Fatou's lemma

$$E(M^+) = E(\liminf M_n^+) \le \liminf EM_n^+ < \infty$$

Therefore $E(M^-) < \infty$, and so M is integrable and consequently finite almost surely. ■

Theorem 146 *(L^p martingale Convergence Theorem) Let (M_n, H_n); $n = 1, 2, \ldots$ be a martingale such that $sup_{n\to\infty} E|M_n|^p < \infty, p > 1$. Then there is a random variable M such that $M_n \to M$ a.s. and in L^p.*

Example 147 *(The Galton-Watson branching process). Consider a population of Z_n individuals in generation n each of which produces a random number ξ of offspring in the next generation so that the distribution of Z_{n+1} is that of $\xi_1 + ... + \xi_{Z_n}$ for independent identically distributed ξ. This process $Z_n, n = 1, 2, ...$ is called the Galton-Watson process. Let $E(\xi) = \mu$. Assume we start with a single individual in the population $Z_0 = 1$ (otherwise if there are j individuals in the population to start then the population at time n is the sum of j independent terms, the offspring of each). Then*

- The sequence Z_n/μ^n is a martingale.
- If $\mu < 1$, $Z_n \to 0$ and $Z_n = 0$ for all sufficiently large n.
- If $\mu = 1$ and $P(\xi \neq 1) > 0$, then $Z_n = 0$ for all sufficiently large n.
- If $\mu > 1$, then $P(Z_n = 0$ for some $n) = \rho$ where ρ is the unique value < 1 satisfying $E(\rho^\xi) = \rho$.

Definition 148 $\{(X_t, H_t); t \in T\}$ *is a* supermartingale *if*

(a) H_t is increasing (in t) family of sigma-algebras.

(b) Each X_t is H_t measurable and $E|X_t| < \infty$.

(c) For each $s < t$, $s, t \in T$, $E(X_t|H_s) \le X_s$ a.s.

Theorem 149 *Suppose $A_n \ge 0$ is a predictable (non-anticipating) bounded sequence and X_n is a supermartingale. Then the supermartingale transform $A \circ X$ is a supermartingale.*

Theorem 150 *Let (M_n, H_n); $n = 1, 2, \ldots$ be a supermartingale such that $M_n \ge 0$. Then there is a random variable M such that $M_n \to M$ a.s. with $E(M) \le E(M_0)$.*

Example 151 *Let S_n be a simple symmetric random walk with $S_0 = 1$ and define the optional stopping time $N = \inf\{n; S_n = 0\}$. Then*

$$X_n = S_{n \wedge N}$$

is a non-negative (super)martingale and therefore $X_n \to$ almost surely. The limit must be 0 since otherwise, $|X_{n+1} - X_n| = 1$ and so convergence is impossible. However, in this case, $E(X_n) = 1$ whereas $E(X) = 0$ so the convergence is not in L_1.

Definition 152 $\{(X_t, H_t); t \in T\}$ *is a* reverse martingale *if*

(a) H_t is decreasing (in t) family of sigma-algebras.

(b) Each X_t is H_t- measurable and $E|X_t| < \infty$.

(c) For each $s < t$, $E(X_s|H_t) = X_t$ a.s.

Example 153 *Let X be any integrable random variable, H_t be any decreasing family of sigma-algebras. Put $X_t = E(X|H_t)$. Then (X_t, H_t) is a reverse martingale.*

Theorem 154 *(Reverse martingale convergence Theorem). If (X_n, H_n); $n = 1, 2, \ldots$ is a reverse martingale,*

$$X_n \to E(X_1 | \cap_{n=1}^{\infty} H_n) \quad a.s. \tag{7.13}$$

Example 155 *(The Strong Law of Large Numbers) Let Y_i be independent identically distributed, $H_n = \sigma(\bar{Y}_n, Y_{n+1}, Y_{n+2}, ...)$, where $\bar{Y}_n = \frac{1}{n}\sum_{i=1}^{n} Y_i$. Then H_n is a decreasing family of sigma fields and $\bar{Y}_n = E(Y_1|H_n)$ is a reverse martingale. It follows from the reverse martingale convergence theorem that $\bar{Y}_n \to Y$ where Y is a random variable measurable with respect to $\cap_{n=1}^{\infty} H_n$. But $\cap_{n=1}^{\infty} H_n$ is in the tail sigma-field and so by the Hewitt-Savage 0-1 Law, Y is a constant almost surely and $Y = E(Y_i)$.*

Example 156 *(Hewitt-Savage 0-1 Law) Suppose Y_i are independent identically distributed and A is an event in the tail sigma-field. Then $P(A) = 0$ or $P(A) = 1$.*

7.4 Uniform Integrability

Definition 157 *A set of random variables $\{X_i, i = 1, 2, ...\}$ is uniformly integrable if*

$$\sup_i E(|X_i| I(|X_i| > c) \to 0 \quad as \ c \to \infty$$

7.4.1 Some Properties of uniform integrability:

1. Any finite set of integrable random variables is uniformly integrable.

2. Any infinite sequence of random variables which converges in L^1 is uniformly integrable.

3. Conversely if a sequence of random variables converges almost surely and is uniformly integrable, then it also converges in L^1.

4. If X is integrable on a probability space (Ω, H) and H_t any family of sub-sigma fields, then $\{E(X|H_t)\}$ is uniformly integrable.

5. If $\{X_n, n = 1, 2, ...\}$is uniformly integrable, then $\sup_n E(X_n) < \infty$.

Theorem 158 *Suppose a sequence of random variables satisfies $X_n \to X$ in probability. Then the following are all equivalent:*

1. $\{X_n, n = 1, 2, ...\}$ is uniformly integrable

2. $X_n \to X$ in L^1.

3. $E(|X_n|) \to E(|X|)$

Theorem 159 *Suppose X_n is a submartingale. Then the following are all equivalent:*

1. $\{X_n, n = 1, 2, ...\}$ is uniformly integrable

2. $X_n \to X$ almost surely and in L^1.

3. $X_n \to X$ in L^1.

Theorem 160 *Suppose X_n is a martingale. Then the following are all equivalent:*

1. $\{X_n, n = 1, 2, ...\}$ is uniformly integrable

2. $X_n \to X$ almost surely and in L^1.

3. $X_n \to X$ in L^1.

4. There exists some integrable X such that $X_n = E(X|H_n)$ a.s.

7.5 Martingales and Finance

Let $S(t)$ denote the price of a security at the beginning of period $t = 0, 1, 2, ...T$. We assume that the security pays no dividends. Define the *(cumulative) returns process* associated with this security by R_S where

$$\Delta R_S(t) = R_S(t) - R_S(t-1) = \frac{\Delta S(t)}{S(t-1)} = \frac{S(t) - S(t-1)}{S(t-1)}, \quad R_S(0) = 0.$$

Then $100\Delta R_S(t)\%$ is the percentage return in an investment in the stock in the $t-1'$st period. The returns process is a more natural characterisation of stock prices than the original stock price process since it is invariant under artificial scale changes such as stock splits etc. Note that we can write the stock price in terms of the returns process;

$$S(t) = S(0)\prod_{i=1}^{t}(1 + \Delta R_S(i)).$$

Now consider another security, a *riskless discount bond* which pays no coupons. Assume that the price of this bond at time t is $B(t)$, $B(0) = 1$ and $R_B(t)$ is the return process associated with this bond. Then $\Delta R_B(t) = r(t)$ is the interest rate paid over the $t-1$'st period. It is usual that the interest paid over the $t-1$st period should be declared in advance, i.e. at time $t-1$ so that if $S(t)$ is adapted to a filtration $\mathcal{F}_t$, then $r(t)$ is *predictable* , i.e. is $\mathcal{F}_{t-1}-$measurable. The *discounted stock price process* is the process given by

$$S^*(t) = S(t)/B(t).$$

Consider a *trading strategy* of the form $(\beta(t), \alpha(t))$ representing the total number of shares of bonds and stocks respectively held at the beginning of the period $(t-1, t)$. Since our investment strategy must be determined by using only the present and the past values of this and related processes, both $\beta(t)$ and $\alpha(t)$ are predictable processes. Then the value of our investment at time $t-1$ is $V_{t-1} = \beta(t)B(t-1) + \alpha(t)S(t-1)$ and at the end of this period, this changes to $\beta(t)B(t) + \alpha(t)S(t)$ with the difference $\beta(t)\Delta B(t) + \alpha(t)\Delta S(t)$ representing the *gain* over this period. An investment strategy is *self-financing* if the value after rebalancing the portfolio is the value before- i.e. if all investments are paid for by the above gains. In other words if $V_t = \beta(t)B(t) + \alpha(t)S(t)$ for all t. An *arbitrage opportunity* is a trading strategy that makes money with no initial investment; i.e. one such that $V_0 = 0$, $V_t \geq 0$ for all $t = 1, \ldots T$ and $E(V_T) > 0$. The basic theorem of no-arbitrage pricing is the following:

7.5.1 Theorem

There are no arbitrage opportunities in the above economy if and only if there is a measure Q equivalent to the underlying measure P i.e. $P << Q$ and $Q << P$ such that under Q the discounted process is a martingale; i.e. $E_Q(S^*(t)|\mathcal{F}_{t-1}] = S^*(t-1)$ a.s. for all $t \leq T$.

Proof; See Pliska (3.19)) page 94.

Note: The measure Q is called the equivalent martingale measure and is used to price derivative securities. For any attainable contingent claim X; (a for any random variable X which can be written as a linear function of the available investments), the arbitrage-free price at time t is given by the conditional expected value under Q of the discounted return X given $\mathcal{F}_t$.

7.6 Problems

1. Let $(\Omega, \mathcal{F}, P)$ be the unit interval with the Borel sigma-algebra and Lebesgue measure defined thereon. Define $\mathcal{F}_n$ to be the sigma field generated by the intervals $(\frac{j-1}{2^n}, \frac{j}{2^n}]$, $j = 1, 2, \ldots 2^n$. Let X be a bounded continuous function on the unit interval.

   (a) Find $E(X|\mathcal{F}_n)$.

   (b) Show $\mathcal{F}_n \subset \mathcal{F}_{n+1}$ for all n .

   (c) Verify that $E(X|\mathcal{F}_n)$ converges pointwise and identify the limit.

   (d) Verify directly that $E\{E(X|\mathcal{F}_n)\} = E(X)$.

   (e) What could you conclude if X had countably many points of discontinuity?

2. Prove property (i), that if Z is $\mathcal{G}-$measurable, $E(ZX|\mathcal{G}) = ZE(X|\mathcal{G})$ a.s.

3. Suppose that X is integrable so that $E(|X|) < \infty$. Prove for constants c, d that $E(cX + d|\mathcal{G}) = cE(X|\mathcal{G}) + d$ (First give the proof in case $E(X^2) < \infty$).

4. Prove property (j): if $\mathcal{H} \subset \mathcal{G}$ are sigma-algebras, $E[E(X|\mathcal{G})|\mathcal{H}] = E(X|\mathcal{H})$. Does the same hold if $\mathcal{G} \subset \mathcal{H}$?

5. Prove: if $X \leq Y$, then $E(X|\mathcal{G}) \leq E(Y|\mathcal{G})$ a.s.

6. Prove: $var(X) = E\{var(X|\mathcal{G})\} + var\{E(X|\mathcal{G})\}$.

7. Prove that if X and Y are simple random variables, $X = \sum c_i I_{A_i}$ and $Y = \sum_j d_j I_{B_j}$ then
$$E(X|Y)(\omega) = \sum_j \sum_i c_i P(A_i|B_j) I_{B_j}(\omega).$$

8. Suppose X is a normal$(0,1)$ variate and $Y = XI(X \leq c)$. Find $E(X|Y)$.

9. Suppose X and Y are independent exponentially distributed random variables each with mean 1. Let I be the indicator random variable $I = I(X > Y)$. Find the conditional expectations

   (a) $E(X|I)$

(b) $E(X+Y|I)$

10. Suppose X is a random variable having the Poisson(λ) distribution and define the indicator random variable $I = I(X$ is even). Find $E(X|I)$.

11. Consider the pair of random variables (X_n, Y_n) where $X_n = X$, and $Y_n = (1/n)X$ for all $n = 1, 2, \ldots$. Show that (X_n, Y_n) converges almost surely to some (X, Y) but it is NOT true in general that $E(X_n|Y_n) \to E(X|Y)$ almost surely or that $E(X_n|Y_n) \to E(X|Y)$ weakly.

12. Suppose Y_i are independent identically distributed. Define $\mathcal{F}_n = \sigma(Y_{(1)}, ..., Y_{(n)}, Y_{n+1}, Y_{n+2}, ...)$, where $(Y_{(1)}, ..., Y_{(n)})$ denote the order statistics. Show $\mathcal{F}_n$ is a decreasing family of sigma fields, find $s_n^2 = E(\frac{1}{2}(Y_1 - Y_2)^2|\mathcal{F}_n)$ and show it is a reverse martingale. Conclude a limit theorem.

13. Let X be an arbitrary absolutely continuous random variable with probability density function $f(x)$. Let $\alpha(s) = f(s)/P[X \geq s]$ denote the hazard function. Show

$$X_t = I(X \geq t) - \int_{-\infty}^{min(X,t)} \alpha(s)ds$$

is a martingale with respect to a suitable family of sigma-algebras.

14. Suppose $(X_t, \mathcal{F}_t)$ is a martingale and a random variable Y is independent of every $\mathcal{F}_t$. Show that we continue to have a martingale when $\mathcal{F}_t$ is replace by $\sigma(Y, \mathcal{F}_t)$.

15. Suppose τ is an optional stopping time taking values in a interval $\{1, 2, ..., n\}$. Suppose $\{(X_t, \mathcal{F}_t); t = 1, 2, ..., n\}$ is a martingale. Prove $E(X_\tau) = E(X_1)$.

16. Prove the general case of Doob's maximal inequality, that for $p > 1, \lambda > 0$ and a non-negative submartingale M_n,

$$P(\sup_{0 \leq m \leq n} M_m \geq \lambda) \leq \lambda^{-p} E(M_n^p)$$

17. Consider a stock price process $S(t)$ and a riskless bond price process $B(t)$ and their associated returns process $\Delta R_S(t)$ and $\Delta R_B(t) = r(t)$. Assume that the stock price takes the form of a binomial tree; $S(t) = S(t-1)[d + (u-d)X_t]$ where X_t are independent Bernoulli random variables adapted to some filtration $\mathcal{F}_t$ and where $d < 1 < 1+r(t) < u$ for all t. We assume that under the true probability measure P, $P(X_t = 0)$ and $P(X_t = 1)$ are positive for all t.

    Determine a measure Q such that the discounted process $S^*(t) = \frac{S(t)}{B(t)}$ is a martingale under the new measure Q and such that Q is equivalent to P i.e. $P << Q$ and $Q << P$. Is this measure unique? What if we were to replace the stock price process by one which had three branches at each step, i.e. it either stayed the same, increased by a factor u or decreased by factor d at each step (a trinomial tree)?

18. Prove that if, under a measure Q, the expected return from a stock is the risk-free interest rate; i.e. if
$$E_Q[\Delta R_S(t)|\mathcal{F}_{t-1}] = r(t) \text{ a.s.}$$
then the discounted price process $S^*(t)$ is a martingale under Q.

19. Prove that for an optional stoping time τ, $\sigma(\tau) \subset H_\tau$.

20. Let $X_1, X_2, ...$ be a sequence of independent random variables all with the same expected value μ. Suppose τ is an optional stopping time with respect to the filtration $H_t = \sigma(X_1, X_2, ..., X_t), t = 1, 2, ...$ and assume that
$$E(\sum_{i=1}^{\tau} |X_i|) < \infty.$$
Prove that
$$E(\sum_{i=1}^{\tau} X_i) = \mu E(\tau).$$

21. Find an example of a martingale $X_t, t = 1, 2, ...$ and an optional stopping time τ such that
$$P[\tau < \infty] = 1$$
but X_τ is not integrable.

22. Let X_n be a submartingale and let a be a real number. Define $Y_n = \max(X_n, a)$. Prove that Y_n is a submartingale. Repeat when $Y_n = g(X_n)$ where g is any convex function.

23. Let X_n be a simple symmetric random walk (i.e. it jumps up or down by one unit with probability 1/2 independently at each time step. Define
$$\tau = \min\{n \geq 5; X_{n+1} = X_n + 1\}.$$
    (a) Is τ a stopping time? What about $\rho = \tau - 1$?
    (b) Compute $E(X_\tau)$. Is $E(X_\tau) = E(X_1)$?

24. Let X_n be a stochastic process such that
$$E(X_{n+1}|X_0, ..., X_n) = X_n + m$$
for some constant m.
    (a) Find a martingale Y_n of the form $Y_n = X_n + cn$.
    (b) Let τ be any stopping time with finite expected value. Compute $E(X_\tau)$ in terms of $E(\tau)$.

25. Consider two independent random variables Y and X on the probability space $(\Omega, \mathcal{F}, P)$ and a sigma-algebra $\mathcal{G} \subset \mathcal{F}$. Prove or provide a counter-example to the statement that this implies $E(X|\mathcal{G})$ is independent of $E(Y|\mathcal{G})$.

26. Consider a sequence of random variables $X_1, X_2, ...$ such that $(X_1, X_2, .., X_n)$ is absolutely continuous and has joint probabiity density function $p_n(x_1, ..., x_n)$. Suppose $q_n(x_1, ..., x_n)$ is another sequence of joint probability density functions and define
$$Y_n = \frac{q_n(X_1, ..., X_n)}{p_n(X_1, ..., X_n)}$$
if the denominator is > 0 and otherwise $Y_n = 0$. Show that Y_n is a supermartingale that converges almost surely.