2.4. **Martingale convergence theorem.** Martingale convergence theorem is a special type of theorem, since the convergence follows from structural properties of the sequence of random variables.  

**Theorem 2.8** (Martingale convergence theorem). Let \((X, \mathcal{F})\) be a submartingale, which satisfies  

\[
(2.15) \quad \sup_n \mathbb{E} X_n^+ < \infty.
\]

Then the limit \(\lim_n X_n = X_\infty\) exists almost surely [and the limit is thus finite, almost surely]. Moreover, we have that \(X_\infty \in L^1\).

**Remark 2.2.** We do not claim that the convergence in Theorem 2.8 is \(X_n \overset{L^1(\mathbb{P})}{\longrightarrow} X_\infty\), when \(n \to \infty\). We shall give a counterexample to this claim in Example 2.6.

**Proof**  
Let \(u_{n}^{X}[a,b]\) be the number of upcrossings of the interval \([a,b]\) before the time \(n\) and put \(u_{\infty}^{X}[a,b] = \lim_n u_{n}^{X}[a,b]\). Using the MCT and inequality (2.14) we will get that  

\[
\mathbb{E} u_{\infty}^{X}[a,b] = \lim_n \mathbb{E} u_{n}^{X}[a,b] 
\leq \frac{1}{b-a} \sup_n \mathbb{E}(X_n - a)^+ 
\leq \frac{1}{b-a} \left( \sup_n \mathbb{E}X_n^+ + |a| \right),
\]

where we have used the fact that always we have \((x-a)^+ \leq x^+ + |a|\). From this we will get that the random variable \(u_{\infty}^{X}[a,b]\) is finite almost surely. Put  

\[
\Lambda_{a,b} = \left\{ \limsup_n X_n \geq b, \liminf_n X_n \leq a \right\}.
\]

Since \(u_{\infty}^{X}[a,b] < \infty\) almost surely, then \(\mathbb{P}(\Lambda_{a,b}) = 0\). Let now \(a, b\) be rational numbers and let \(\Lambda = \bigcup_{a < b, a, b \in \mathbb{Q}} \Lambda_{a,b}\). From above we know that \(\mathbb{P}(\Lambda) = 0\). On the other hand we have that  

\[
\Lambda = \left\{ \limsup_n X_n > \liminf_n X_n \right\}.
\]

Hence we get that the limit \(X_\infty = \lim_n X_n\) exists almost surely [but as indicated, the limit \(X_\infty\) is not necessarily finite].  

Let us show that in fact \(X_\infty \in L^1\). \(X\) is a submartingale, and so \(\mathbb{E} X_n \geq \mathbb{E} X_0\). Let us remark that  

\[
\mathbb{E}|X_n| = \mathbb{E} X_n^+ + \mathbb{E} X_n^- 
\leq 2\mathbb{E} X_n^+ - \mathbb{E} X_0,
\]

---

2 This result is somewhat analogous to the result from analysis, which says that a bounded monotone sequence always converges.

3 In what follows we denote by \(Q\) the rational numbers.
and by Fatou’s lemma and the estimate (2.15) give by using the estimate above that
\[
\mathbb{E}|X_\infty| = \mathbb{E} \left( \lim_n |X_n| \right) \leq \liminf_n \mathbb{E}|X_n| \\
\leq 2 \sup_n \mathbb{E}X_n^+ - \mathbb{E}X_0 < \infty;
\]
and so \(X_\infty \in L^1\). \(\square\)

**Corollary 2.3.** If \((X, \mathbb{F})\) is non-negative supermartingale or if \((X, \mathbb{F})\) is a martingale bounded by above (or below), then the limit \(X_\infty = \lim_n X_n\) exists and \(X_\infty \in L^1\).

**Proof** If \((X, \mathbb{F})\) is non-negative supermartingale, then \((-X, \mathbb{F})\) is a submartingale satisfying (2.15). Hence non-negative supermartingale converges almost surely and and \(X_\infty = \lim_n X_n\) is integrable.

Let \((X, \mathbb{F})\) be a martingale bounded below: \(X_n \geq -c\) for some \(c > 0\), when \(n \geq 1\). The \(Y_n = X_n + c\) is a non-negative martingale, and hence a supermartingale, and so \((Y_n, \mathbb{F})\) converges using the previous argument. This in turn implies that \((X, \mathbb{F})\) converges almost surely, and so \(X_\infty = \lim_n X_n\) with \(X_\infty \in L^1\).

Finally, let \((X, \mathbb{F})\) be a martingale bounded from above. Then \((-X, \mathbb{F})\) is a martingale bounded below, and after this observation everything should be clear. \(\square\)

### 2.5. Uniform integrable martingales.

#### 2.5.1. Uniform integrability.

**Example 2.6.** Let \(\xi_k\) be i.i.d. random variables, \(\mathbb{P}(\xi_k = 0) = \mathbb{P}(\xi_k = 2) = \frac{1}{2}\). We have that \(\mathbb{E}\xi_k = 1\) and if \(M_n = \prod_{k \leq n} \xi_k\), then \((M, \mathbb{F}^M)\) is a martingale, \(M_n \geq 0\) and \(\sup_n \mathbb{E}M_n = 1\). Martingale convergence theorem applies, and we have that there exists almost sure limit \(M_\infty = \lim_n M_n\). It is quite easy to see that \(M_\infty = 0\) almost surely. Hence \(\mathbb{E}M_\infty = 0 \neq 1 = \mathbb{E}M_n\).

This means that \(X_n\) does not converge to \(X_\infty\) in \(L^1\).

The next definition is essential for obtaining also the convergence in \(L^1\).

**Definition 2.6.** Let \(X_j, j \in \mathcal{J}\) be a family of random variables. The family \(X_j, j \in \mathcal{J}\) is uniformly integrable, if for all \(\epsilon\) there exists \(K > 0\) such that
\[
(2.16) \quad \int_{\{|X_j| > K\}} |X_j|d\mathbb{P} < \epsilon.
\]

If \(X_j, j \in \mathcal{J}\) is uniformly integrable family, then it is bounded in \(L^1\): take \(\epsilon = 1\) and \(K > 0\) such that \(\int_{\{|X_j| > K\}} |X_j|d\mathbb{P} < 1\). We obtain that \(\mathbb{E}|X_j| < K + 1\) for all \(j \in \mathcal{J}\).

On the other hand, a family of random variables, bounded in \(L^1(\mathbb{P})\), then it is not necessarily uniformly integrable. Take \(X_n = nI_{(0, 1/n)}\) and \(\mathbb{P} = Leb|(0, 1)\),

then \(\mathbb{E}X_n = 1\) and for all \(K > 0\)
\[
\int_{\{|X_n| > K\}} |X_n|d\mathbb{P} = 1,
\]
when \( n > K \).
But if the random variables are bounded in \( L^p, p > 1 \), then we have that they are uniformly integrable as well.

**Theorem 2.9.** Let the family \( X_j, j \in \mathcal{J} \) be bounded in the space \( L^p(\mathbb{P}), p > 1 \). The family \( X_j, j \in \mathcal{J} \) is uniformly integrable.

**Proof** Let \( \epsilon > 0 \), and let \( q > 1 \) be the Hlder conjugate of \( p \) and put \( M = \sup_j \mathbb{E}|X_j|^p < \infty \). With the help of Hlder inequality

\[
\mathbb{E} \left( |X_j| I_{\{|X_j| > K\}} \right) \leq ||X_j||_p (\mathbb{P}(\{|X_j| > K\}))^{\frac{1}{q}} \\
\leq (\mathbb{E}|X_j|^p)^{\frac{1}{q}} \left( \frac{\mathbb{E}|X_j|^p}{K^p} \right)^{\frac{1}{q}} = MK^{-\frac{p}{q}} < \epsilon,
\]

when \( K > \left( \frac{\epsilon M}{p} \right)^{\frac{q}{p}} \).

To continue we need the following lemma:

**Lemma 2.3.** Let \( Y \in \mathbb{L}^1(\mathbb{P}) \). Then for all \( \epsilon > 0 \) there exists \( \delta > 0 \) such that \( \mathbb{P}(A) < \delta \Rightarrow \mathbb{E}(|Y|A) < \epsilon \).

**Proof** We give an indirect proof: there exists \( \epsilon_0 > 0 \) and \( A_n \) such that \( \mathbb{P}(A_n) < 2^{-n} \) and \( \mathbb{E}(|Y|A_n) > \epsilon_0 \), when \( n \geq 1 \).

Put \( H = \limsup_n A_n \). By construction of the sets \( A_n \) we have that \( \sum_{k=1}^{\infty} \mathbb{P}(A_n) < \infty \), and so Borel-Cantelli lemma implies that \( \mathbb{P}(H) = 0 \), and so \( \mathbb{E}(|Y|I_H) = 0 \). On the other hand, by Fatou’s lemma

\[
\mathbb{E}(|Y|I_H) = \mathbb{E}(\limsup_{n}(|Y|I_{A_n})) \geq \limsup_{n} \mathbb{E}(|Y|I_{A_n}) \geq \epsilon_0;
\]

but this is a contradiction, since we should also have \( \mathbb{E}(|Y|I_H) = 0 \). Hence we are done.

**Theorem 2.10.** Let \( X \in \mathbb{L}^1((\Omega, \mathcal{F}, \mathbb{P})) \) and put \( Y^\mathcal{G} = \mathbb{E}[X|\mathcal{G}] \), when \( \mathcal{G} \subset \mathcal{F} \). Then the family \( Y^\mathcal{G} \) is uniformly integrable.

**Proof** Using Lemma 2.3 given \( \epsilon > 0 \) we can pick \( \delta > 0 \) such that, if \( \mathbb{P}(F) < \delta \), then \( \mathbb{E}(|X|F) < \epsilon \). Moreover, from the fact that \( X \in \mathbb{L}^1(\mathbb{P}) \) it follows that we can find a constant \( K > 0 \) such that \( \mathbb{E}(|X|) < \delta \).

Put \( Y^\mathcal{G} = \mathbb{E}[X|\mathcal{G}] \), when \( \mathcal{G} \subset \mathcal{F} \). From the Jensen inequality it follows that

\[
|Y^\mathcal{G}| \leq \mathbb{E}[|X||\mathcal{G}].
\]

From the estimate (2.17) we will get that \( \mathbb{E}|Y^\mathcal{G}| \leq \mathbb{E}|X| \) and further

\[
K \mathbb{P}(|Y^\mathcal{G}| > K) \leq \mathbb{E}|Y^\mathcal{G}| \leq \mathbb{E}|X|.
\]

With the help of the inequality (2.17) we finally get that

\[
\int_{\{|Y^\mathcal{G}| > K\}} |Y^\mathcal{G}|d\mathbb{P} \leq \int_{\{|Y^\mathcal{G}| > K\}} \mathbb{E}[|X||\mathcal{G}] = \int_{\{|Y^\mathcal{G}| > K\}} |X|d\mathbb{P},
\]

where the last inequality follows from the definition of conditional expectation. Now the claim follows from the Lemma 2.3, because \( \mathbb{P}(|Y^\mathcal{G}| > K) < \delta \).

**Theorem 2.11.** Let \( X_n, X \in \mathbb{L}^1(\mathbb{P}) \). The following are equivalent:

a) \( \mathbb{E}|X_n - X| \to 0, as n \to \infty. \)
b) The family $X_n$, $n \geq 1$ is uniformly integrable and $X_n \xrightarrow{P} X$.

**Proof** First we prove the implication b) $\Rightarrow$ a): Let $K > 0$ and

$$
\psi_k(x) = xI_{\{|x| \leq K\}} + KI_{\{x > K\}} - KI_{\{x < -K\}}.
$$

The family $X_n$, $n \geq 1$ is uniformly integrable, so there exists a constant $K > 0$ such that

$$
(2.19) \quad \mathbb{E}|\psi_K(X_n) - X_n| = \left| \int_{\{|X_n| > K\}} X_n d\mathbb{P} \right| \leq \int_{\{|X_n| > K\}} |X_n| d\mathbb{P} \leq \frac{\epsilon}{3}.
$$

Clearly $|\psi_K(x) - \psi_K(y)| \leq |x - y|$ and this in turn implies that the map $x \mapsto \psi_K(X)$ is continuous. Continuous mappings preserve convergence in probability, and so $\psi_K(X_n) \xrightarrow{P} \psi_K(X)$. DCT theorem implies, that in fact this convergence in $L^1$, and hence there exists $n_0$ such that if $n \geq n_0$, then

$$
(2.20) \quad \mathbb{E}|\psi_K(X_n) - \psi_K(X)| < \frac{\epsilon}{3}.
$$

We can now collect the estimates (2.19) and (2.20) and obtain

$$
\mathbb{E}|X_n - X| \leq \mathbb{E}|X_n - \psi_K(X_n)| + \mathbb{E}|\psi_K(X_n) - \psi_K(X)| + \mathbb{E}|\psi_K(X) - X| \leq \frac{\epsilon}{3} + \frac{\epsilon}{3} + \frac{\epsilon}{3} = \epsilon.
$$

Hence we have a).

Conversely, assume a): $\mathbb{E}|X_n - X| \to 0$, as $n \to \infty$. Because $L^1$ convergence implies stochastic convergence, then $X_n \xrightarrow{P} X$.

Next we will show that the family $X_n$, $n \geq 1$ is uniformly integrable. Let $\epsilon > 0$ and $n_0$ such that for $n \geq n_0$ we have $\mathbb{E}|X_n - X| < \frac{\epsilon}{2}$. Lemma 2.3 gives the following: there exists $\delta > 0$ such that if $\mathbb{P}(A) < \delta$, then

$$
(2.21) \quad \mathbb{E}|X_k I_A| < \epsilon \text{ ja } \mathbb{E}|X I_A| < \frac{\epsilon}{2},
$$

when $k = 1, \ldots, n_0$ [formally the constant $\delta$ depends from the index $k$, but we can take $\delta = \min\{\delta_k : k \leq n_0\}$, and $\delta > 0$, since the minimum is taken over a finite collection of positive numbers].

The family $X_n$, $n \geq 1$ is bounded in $L^1(\mathbb{P})$, and hence there exists $K > 0$ such that we have

$$
(2.22) \quad \sup_n \frac{\mathbb{E}|X_n|}{K} < \delta.
$$

From this inequality we obtain that $\mathbb{P}(|X_n| > K < \delta, n \geq 1$, so if $k \leq n_0$, then $\mathbb{E}|X_k I_{\{|X_k| > K\}}| < \epsilon$ and if $k > n_0$, then

$$
\mathbb{E}|X_k I_{\{|X_k| > K\}}| \leq \mathbb{E}|X_k - X| + \mathbb{E}|X I_{\{X_k\}}| < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon.
$$

Hence the family $X_n$, $n \geq 1$ is uniformly integrable and we have proved b).

□

As a corollary we get the very useful summary:

**Corollary 2.4.** Let $X_n$, $n \geq 1$ be non-negative random variables which satisfy $X_n \xrightarrow{P} X$, $X \in L^1(\mathbb{P})$. Then the following are equivalent:

a) $\mathbb{E}|X_n - X| \to 0$, as $n \to \infty$.

b) The family $X_n$, $n \geq 1$ is uniformly integrable.
c) \( \mathbb{E}X_n \to \mathbb{E}X \), as \( n \to \infty \).

**Proof**  The claim a) \( \iff \) b) is in fact Theorem 2.11.

Assume that a) holds. Because \( |\mathbb{E}X_n - \mathbb{E}X| \leq \mathbb{E}|X_n - X| \), then we have c).

Conversely, assume c). We have the identity

\[
|X_n - X| = X_n - X + 2(X_n - X)^-.
\]

By assumption \( (X_n - X)^- \to 0 \) and since \( X_n, X \) are non-negative, we get the estimate \( (X_n - X)^- \leq X \). Now we can use the dominated convergence theorem to get

\[
\mathbb{E}(X_n - X)^- \to 0 \quad \text{as} \quad n \to \infty.
\]

So the claim a) follows from this fact, and identity (2.23). \( \square \)

2.5.2. Uniform integrable martingales.

**Definition 2.7.** \( (M_n, F_n)_{n \geq 1} \) is an uniformly integrable martingale, if \( (M_n, F_n)_{n \geq 1} \) is a martingale, and the family \( M_n, n \geq 1 \) is uniformly integrable.

Let \( (M_n, F_n)_{n \geq 1} \) be uniformly integrable martingale. Uniform integrability implies that the family \( M_n, n \geq 1 \) is bounded in \( L^1(\mathbb{P}) \), and so \( \sup \mathbb{E}|M_n| < \infty \). We can use the martingale convergence theorem to conclude that we have the limit

\[
M_\infty = \lim_{n} M_n \mathbb{P} - \ a.s.;
\]

moreover, by Theorem 2.11 we get that \( \mathbb{E}|M_\infty - M_n| \to 0 \), as \( n \to \infty \).

We claim that in fact

\[
M_n = \mathbb{E}[M_\infty | F_n].
\]

By the martingale property we get, when \( F \in F_n \) and \( r > n \):

\[
\int_F M_n d\mathbb{P} = \int_F M_r d\mathbb{P}.
\]

On the other hand

\[
\left| \int_F M_r d\mathbb{P} - \int_F M_\infty d\mathbb{P} \right| \leq \int_F |M_r - M_\infty| d\mathbb{P} \leq \mathbb{E}|M_r - M_\infty| \to 0, \ \text{kun} \ r \to \infty.
\]

This shows that we can replace in the equation (2.25) the random variable \( M_r \) by the random variable \( M_\infty \). From this we obtain the claim (2.24).

As a summary of uniform integrable martingales we obtain

**Theorem 2.12.** Assume that \( (M_n, F_n)_{n \geq 1} \) is a uniformly integrable martingale. Then there exists an integrable random variable \( M_\infty \), which satisfies

\[
M_\infty = \lim_{n} M_n \mathbb{P} - \ a.s. \text{ and in } L^1(\mathbb{P}).
\]

Moreover, we have the representation (2.24).

**Corollary 2.5.** Let \( (M_n, F_n)_{n \geq 1} \) be an uniformly integrable martingale and \( \tau \) is a stopping time. Then

\[
M_\tau = \mathbb{E}[M_\infty | F_\tau].
\]

To verify this is left as an exercise.
2.6. Change of measure.

2.6.1. Radon-Nikodym-theorem.

**Definition 2.8.** probability measure \( Q \) is absolutely continuous with respect to the probability measure \( P \), if \( P(A) = 0 \Rightarrow Q(A) = 0 \). Notation: \( Q \ll P \). The measures \( Q \) ja \( P \) are equivalent if \( Q \ll P \) ja \( P \ll Q \). Notation: \( Q \sim P \).

**Example 2.7.** Let \( X \geq 0 \) be a random variable with \( \mathbb{E}_P X = 1 \). Then one can define a new probability measure \( Q \) by putting

\[
Q(A) = \int_A XdP.
\]

We know that \( Q \) is a probability measure and that \( Q \ll P \) [after the exercise III, at least].

**Theorem 2.13** (Radon-Nikodym-theorem). Let \( (\Omega, \mathcal{F}) \) be a measurable space and \( P, Q \) probability measures on it. Let \( Q \ll P \). Then there exists a random variable \( Y \geq 0 \) such that

\[
Q(A) = \int_A YdP.
\]

We say that \( Y \) is the Radon-Nikodym derivative of \( Q \) with respect to the measure \( P \). Notation: \( Y = \frac{dQ}{dP} \).

**Proof** We prove the theorem in the case when \( \mathcal{F} \) is separable [or countably generated]: there exists sets \( A_1, A_2, \ldots \) such that \( \mathcal{F} = \sigma\{A_i : i \geq 1\} \).

Let \( \mathcal{F}_n = \sigma\{A_1, \ldots, A_n\} \); we know that if \( B \in \mathcal{F}_n \), then \( B = \bigcup_{i=1}^k B_i \), where the sets \( B_i \) have the form \( B_i = \bigcap_{i=1}^{l_i} C_i \), and \( C_i = A_j \) or \( C_i = A_j^c \) for some \( A_j \in \{A_1, \ldots, A_n\} \). We say that a set \( D \in \mathcal{F}_n \) is an atom if \( G \subset D \) implies that \( G = D \) or \( G = \emptyset \). It is easy to check that the atoms \( A_{n,1}, \ldots, A_{n,r(n)} \) of the sigma-algebra \( \mathcal{F}_n \) define a partition of the set \( \Omega : \Omega = \bigcup_{k=1}^{r(n)} A_{n,k} \) and if \( G \in \mathcal{F}_n \), then \( G = \bigcup_{j=1}^p A_{n,j} \). Denote the atoms of \( \mathcal{F}_n \) by \( A_{n,j} \).

Clearly \( \mathcal{F}_n \subset \mathcal{F}_{n+1} \). Put \( Y_n(\omega) = 0 \), when \( \omega \in A_{n,j} \) and \( P(A_{n,j}) = 0 \) [then also \( Q(A_{n,j}) = 0 \), since \( Q \ll P \)] and otherwise we put

\[
Y_n(\omega) = \frac{Q(A_{n,j})}{P(A_{n,j})}.
\]

We will show that \( (Y_n, \mathcal{F}_n) \), \( n \geq 1 \) is \( P \)-martingale. Let \( F \in \mathcal{F}_n \). Because also \( F \in \mathcal{F}_{n+1} \), then there exists atoms \( A_{n+1,j,k} \) such that

\[
F = \bigcup_{k=1}^p A_{n+1,j,k}.
\]
Now we can compute
\[ Q(F) = \sum_{k=1}^{p} Q(A_{n+1,j_k}) = \sum_{k=1}^{p} \frac{Q(A_{n+1,j_k})}{\mathbb{P}(A_{n+1,j_k})} \mathbb{P}(A_{n+1,j_k}) = \int_F Y_{n+1} \mathbb{P}. \]

And with a similar computation one can show that
\[ Q(F) = \int_F Y_n d\mathbb{P}; \]
now we have shown that
\[ \int_F Y_n d\mathbb{P} = \int_F Y_{n+1} d\mathbb{P}, \]
so the sequence \((Y_n, F_n), n \geq 1\) is a martingale.

On the other hand, \(Y_n \geq 0, \mathbb{E}_\mathbb{P} Y_n = 1\), and so the martingale \((Y_n, F_n)_{n \geq 1}\) converges almost surely and we have a limit \(Y_\infty = \lim_n Y_n\) a.s.

We will show next that \((Y_n, F_n, \mathbb{P})_{n \geq 1}\) is an uniformly integrable martingale.

The following is an exercise:

**Lemma 2.4.** Let \(Q \ll \mathbb{P}\): then for all \(\epsilon > 0\) there exists \(\delta > 0\) such that \(\mathbb{P}(F) < \delta \Rightarrow Q(F) < \epsilon\).

Now we prove uniform integrability. Let \(\epsilon > 0\); then we can find \(\delta > 0\) such that \(\mathbb{P}(F) < \delta \Rightarrow Q(F) < \epsilon\). Pick a constant \(K > 0\) such that
\[ \frac{1}{K} = \frac{Q(\Omega)}{K} < \delta. \]

Then
\[ \mathbb{P}(Y_n > K) \leq \frac{\mathbb{E}_\mathbb{P} Y_n}{K} = \frac{Q(\Omega)}{K} < \delta; \]
and by using Lemma 2.4 we finally get
\[ \int_{\{Y_n > K\}} Y_n d\mathbb{P} = Q(Y_n > K) < \epsilon. \]

Hence the family \(Y_n, n \geq 1\) is uniformly integrable with respect to the measure \(\mathbb{P}\), and \(Y_\infty = \lim_n Y_n\) also in \(L^1(\mathbb{P})\). Moreover \(\mathbb{E}_\mathbb{P} Y_\infty = 1\).

Define a probability measure \(R\) by
\[ R(A) = \int_A Y_\infty d\mathbb{P}, \text{ when } A \in \mathcal{F}. \]

We claim that \(R = Q\).

We know that \(Y_n = \mathbb{E}_\mathbb{P}[Y_\infty | \mathcal{F}_n]\), so if \(A \in \mathcal{F}_n\), the by definition of the conditional expectation
\[ Q(A) = \int_A Y_n d\mathbb{P} = \int_A Y_\infty d\mathbb{P} = R(A). \]

hence the probability measures \(R\) ja \(Q\) are the same on \(\mathcal{H} = \cup_n \mathcal{F}_n\). But the family \(\mathcal{H}\) is a monotone class, and by extension theorem we get that the measures \(Q\) and \(R\) are the same on \(\mathcal{F}\). \(\square\)

**Remark 2.3.**

- The Radon-Nikodym derivative \(Y = \frac{dQ}{d\mathbb{P}}\) is unique almost surely.
- One can show that Theorem 2.13 holds for arbitrary \(\mathcal{F}\).

27.3. 2008 [with some leftover on Tu 1.4. 2008]